Reinvestigating who benefits and who loses from universal child care in Canada

Michael J. Kottelenberg
Queen's University

Steven F. Lehrer*
Queen's University and NBER

August 2014

Abstract

We extend earlier research evaluating the Quebec Family Policy by providing the first evidence on the distributional effects of universal child care on two specific developmental outcomes. Our analysis uncovers substantial policy relevant heterogeneity in the estimated effect of access to subsidized child care across two developmental score distributions for children from two-parent families. Whereas past research reported findings of negative effects on mothers and children from these families, igniting controversy, our estimates reveal a more nuanced image that formal child care can indeed boost developmental outcomes for children from some households: particularly disadvantaged single-parent households. In addition, we document significant heterogeneity that differs by child gender. We present suggestive evidence that the heterogeneity in policy effects that emerges across child gender and family type is consistent with differences in the home learning environments generated by parents behaviors that are previously present and are shaped by responses to the policy.

*We would like to thank Habiba Djebbari and seminar participants at the University of Toronto, NYU-Shanghai, CEA Annual meetings, Waterloo, York University, University of Sydney, University of Melbourne, Chinese University of Hong Kong, John Deutsch Institute's conference on Economic Relations Between Children and Parents, 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.

1 Introduction

In many early education and care policy debates, proponents position early education and care as being a social policy silver bullet: simultaneously boosting human capital development of children, supporting parents who would otherwise struggle to meet escalating child care costs and helping mothers re-enter the labor force. Bolstering these calls is not only a growing body of research in neurobiology summarized in Shonkoff and Phillips (2000) that clarifies the extent to which the interaction between genetics and early experience shape brain architecture, but also evidence from several influential randomized experiments including the High/Scope Perry Preschool program. Researchers, including Belfield et al. (2006) and Heckman et al. (2013), that evaluated the Perry program indeed found remarkable gains in high school graduation and employment rates along side decreases in teen pregnancy, delinquency and crime rates. In response, opponents often express concerns related to the external validity of these studies and argue that the studies do not shed any light on the effectiveness of programs and policies that universally provide early education and child care.

Baker (2011) surveys the small but growing literature on studies evaluating universal child care policies and suggests the available evidence does not provide policymakers with clear guidance. On the one hand, Berlinski, Galiani and Gertler (2009), Gormley and Gayer (2005) and Havnes and Mogstad (2011) report positive effects from expansions of early education in Argentina, Oklahoma and Norway respectively. On the other hand, Datta Gupta and Simonsen (2010) and Baker, Gruber and Milligan (2008) respectively provide evidence that the introduction of preschool in Denmark and universal child care in the Canadian province of Quebec led to statistically significant declines in child outcomes. The paper by Baker, Gruber and Milligan (2008) has been quite influential and also showed that the policy led to declines in both child and maternal health as well as family functioning. However, with the sole exception of Havnes and Mogstad (2014), studies evaluating universal policies focus primarily on estimating a mean effect of the program under study. We believe that improving our understanding of who truly benefits and who loses from these policies is increasingly important as several countries world-wide are currently considering implementing a variety of universal early childhood education programs.

Our study provides new evidence on the effectiveness of universal child care policy by exploring distributional treatment effects from the only large scale universal subsidization of child care in North America, the 1997 Quebec Family Policy. Specifically, using the Athey and Imbens change in changes estimator we present evidence that there was considerable heterogeneity in the policy impacts on developmental test scores, which is missed by looking only at average causal effects.[1] Our main findings appear to reconcile much of the conflicting evidence between results from small scale experimental studies and Baker et al. (2008). We find that the Quebec Family Policy significantly boost developmental test scores for children from single parent households particularly for those who are most disadvantaged and located at the lower quantiles of the distribution. However, children from two-parent families between the 10th and 50th quantile generally receive significant negative impacts from child care. As this group is a large fraction of the sample in the Baker et al. (2008) study, it is not surprising that the mean impacts reported are negative in sign. Against traditional notions surrounding the implementation of universal child care, those in the top half of the distribution are generally unaffected by the policy. This finding is new to the child-care literature, as there are no experiments that target the most advantaged in society. Surprisingly, our estimates uncover markedly different

[1] Havnes and Mogstad (2014) also use both a quantile difference in differences and change in changes proposed in Athey and Imbens (2006) but do not control for covariates. Ex ante we believe that controlling for covariates is important since the policy effects could operate through two channels, i) the true effect of child care on specific outcomes, and ii) these changes in the composition of those who attend child care. Adding impetus to controlling for covariates, Heckman *et al.* (2010) reevaluation of an early childhood education program clearly shows that conclusions may not be robust to balancing observed covariates between those who gained access to child care and those that did not. In our results, we present estimates with and without controls that illustrate why we need to add controls and Kolmogov-Smirnov tests clearly demonstrate that the figures are statistically different when we control for covariates.

patterns of treatment effect heterogeneity by child gender. Taken together these results reinforce 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.[2]

To shed light on the mechanism underlying the heterogeneous pattern of distributional effect estimates we next examine how child rearing practices and family functioning changed in response to the policy. This analysis is motivated by the large observed differences in these practices across the unconditional test score distribution, whereas the effects of child care attendance and maternal labour supply do not exhibit significant differences. Using an instrumental variables estimator, we find large and statistically significant declines from child care attendance on many parenting practices for two-parent families. However, despite higher take-up of child care following the policy, single parents are found to generally have a more muted response in these parenting behaviors, which may explain why the Quebec Family Policy did not lead to declines in developmental outcomes for children from these households. Since research using US data has shown gains in both parenting and family functioning after the provision of subsidized child care for economically disadvantaged subpopulations, we believe that this set of results is suggestive that in the design of any child care policy should consider efforts to ensure that parental investment into home learning environments is not reduced in response to formal child care attendance.

This paper is structured as follows. Section 2 summarizes the Quebec Family Policy and describes the data we use in this study. We discuss the empirical methods used to estimate distributional effects in section 3 and present the main results in section 4. We also present sug-

[2] 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 (e.g. Ding and Lehrer (2011), Angrist, Oreopoulos and Williams (2013), among others). More generally, recent studies including Bitler, Gelbach and Hoynes (2006) and Djebbari and Smith (2008) have shown that policy changes generate both winners and losers and as such it is important to report distributional treatment effects in empirical work.

gestive evidence that this policy may have had unintended consequences of influencing parental child rearing decisions and explore how these responses varied over the unconditional developmental score distributions. Finally, in the concluding section we summarize our findings, and argue that policymakers should either consider supplementing universal child care policies with awareness programs to minimize unintended parental responses or consider policies that would target the provision of child care to those who would gain the most.

2 Background

In 1997, the Quebec government implemented a bold set of policies with highly subsidized child-care services as the cornerstone, in hopes strengthening governmental support of parents.[3] Children that were newly-born until the age of 4, irrespective of their parents' income gained access to child care at a rate of $5 per day (becoming $7 per day in 2004).[4] This program was implemented gradually since there needed to be a corresponding increase in the number of child care spaces in Quebec; access was extended to children aged 4 in 1997, aged 3 in 1998, aged 2 in 1999 and aged 0-2 in 2000. 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.

[3]This policy also increased child care services through other channels that are not being investigated in this paper including the introduction of full-day kindergarten and, although not officially part of the policy, more child care spaces for school aged children. See Tougas (2002) for many more institutional details regarding this policy.

[4]The Quebec Family Policy's extension of highly subsidized universally available child care to children aged 0-4 is unique in North America. 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. That said, Oklahoma and Georgia provide universal access for pre-kindergarten that targets children aged 4. See Kottelenberg and Lehrer (2014) for a discussion of how the effects of Quebec's child care policy vary for children of different ages.

2.1 Data

To examine the distributional impacts of the Quebec Family policy we use the National Longitudinal Survey of Children and Youth (NLSCY). This longitudinal dataset conducts biennial assessments of a representative sample of Canadian children selected from Statistics Canada's Labour Force Survey.[5] The first NLSCY data collection cycle was carried out in 1994–1995 and the study completed eight cycles of data collection. During each new cycle, a new cohort of children aged 0-1 is added to the sample. The unit of analysis for the NLSCY is the child or youth. Thus, there is information on approximately 2,000 children at each age level in each cycle of the NLSCY data collection. In our analyses we drop data collected in the waves that began in 1998 and 2000 since the program was in the implementation stage.[6]

During each cycle, an interviewer from Statistics Canada meets with the person in the household who is most knowledgeable about the child. The respondent in each household completed a personal interview that assessed child care usage, parental labor supply, parental and family characteristics along with the child's physical, cognitive, behavioral, and social development.[7] This respondent provides information for each selected child (up to four in a household) when he or she is between 0 and 17 years of age in the household.[8]

In this paper, our primary focus is to explore the impacts of access to child care on two contemporaneous developmental scores: Motor and Social Development (MSD) and the revised

[5]The Labour Force survey is conducted monthly and is used by Statistics Canada to produce unbiased estimates of various statistics including the unemployment rate. That data uses a stratified multistage probability sample design and interviews approximately 59,000 Canadian households each month. Since households residing in institutional facilities, on Aboriginal reserves, and in the two territories are not targeted by the Labour Force Survey, children from these households are excluded from the NLSCY by design.

[6]That is, data from cycles 1 and 2 collected prior to 1997 are used for the pre-policy period and data collected after 2001 in cycles 5 to 8 are in the post-policy time period.

[7]In two parent families 90.7% of respondents are the mother. We also note that 95.0% of single parent families are headed by mothers.

[8]We use each child's final survey weight provided in the NLSCY that has been adjusted for nonresponse, and post-stratified by province, age and sex to match known population totals at the time of sample selection for the full sets of estimates.

Peabody Picture Vocabulary Test (PPVT).[9] Parents of children between the ages of 0 to 3 responded to 15 questions on their Motor and Social Development. The exact questions varied by age of the child and took the form of simply asking whether or not the child is able to perform a specific task.[10] This scale 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. In the NLSCY, children aged 4 and 5 were administered the Peabody Picture and Vocabulary Test (PPVT) which is designed to estimate a child's receptive or passive vocabulary development. The children hear a word said aloud and are asked to point to one of four pictures that they believe corresponds to the word.

Table 1 presents summary statistics on our variables of interest separately for two-parent and single parent families.[11] Each column refers to a unique group-time subgroup. Comparing the parental and family characteristics between two parent and single parent families, it is not

[9] We do not consider estimating distributional treatment effects with the change in changes estimator for the other outcomes considered in Baker et al. (2008) for two main reasons. First, as detailed in Section 4.2 of Athey and Imbens (2006) strong additional assumptions are required to point identify causal effects with the change in changes model with discrete outcomes. While, there is a temptation to treat some of the measures considered in Baker et al. (2008) such as family well-being as being continuous, evidence from a set of Monte Carlo simulations in the supplement to Athey and Imbens (2006) clearly demonstrate that this is a poor choice when dealing with data that has a discrete nature, even when there are a substantial number of values that the outcomes can take on.

[10] To overcome concerns that older children will be able to do more tasks due to maturity, standardized scores by age of child in months are calculated. Within each 1 month age group, the MSD scores are assigned such that they come from a normal distribution with mean 100 and standard deviation of 15. This was designed to compare the MSD scores of children across the 0 to 3 age group, not controlling for age.

[11] Prior research investigating the Quebec Family policy including Baker, Gruber and Milligan (2008) and Kottelenberg and Lehrer (2013) focused solely on two-parent families since concurrent program reforms (some of which are described in Milligan and Stabile (2007) complicate the inference on labor supply and child care usage outcomes for single mothers. After all, many single women in Quebec qualified for child care subsidies before the Quebec family policy was introduced, even though those subsidies were less substantial. In contrast to Lefebvre and Merrigan (2008, 2009) we analyze single parent and two parent families separately given the differential incentives generated by the policy. Further, we argue that by focusing only on two developmental scores, the effects of any concurrent reforms are minimized since they did not directly target these outcomes. Finally, as further motivation for looking at single-parent families we note that a large fraction of the sample in the Perry experiment comes from single parent families. Thus, we can better contrast our distributional estimates with those obtained in the broader early childhood education and care literature.

surprising that single parent's exhibit lower human capital accumulation, have smaller family sizes and are more likely to reside in an urban location. Between groups the trends in most of these characteristics over time are similar, with the exception of the single mother being an immigrant in the rest of Canada. Turning to the outcome variables in the last panel, there is large change in the child care usage across all regions in Canada post-policy, particularly for single parent families. In addition, there were large increases in maternal labor supply on the extensive margin for all groups over time with the exception of two-parent families in the rest of Canada. Child outcomes are generally lower in single parent families but the test score distributions are fairly similar. There are significant gains in PPVT scores over time only for single parent families in the rest of Canada. This may reflect, the large gain in the percentage of single mothers who are university educated. Interestingly, all groups experience a drop in MSD scores over time.

3 Empirical Strategy

Since there have been few empirical applications of the Athey and Imbens (2006) change in changes estimator, we begin by providing a brief discussion of how this expands on knowledge from a the linear difference in differences (DID) estimator. We define D to be a binary treatment variable and assume that the treatment happens between periods of data collection, meaning that every member of the population is untreated in the pre-treatment period. We use binary indicators t and g to respectively denote if the data was collected in the pre ($t = 0$) and post ($t = 1$) policy periods from the comparison ($g = 0$) and treatment ($g = 1$) groups. Often researchers are interested in discovering the mean effect of switching D from zero to one on some outcome variable Y. The linear DID estimator recovers this effect by subtracting the difference of the mean outcomes of treated and controls after the treatment from the outcome difference in these groups before treatment (potentially conditioning on a set of covariates X).[12]

The idea of the change in changes estimator is to compare the cumulative distribution functions of the outcomes in the four groups defined on the basis of g and t, as opposed to just the simple means as in the linear DID estimator. The empirical cumulative distribution functions will be denoted by $F_{Y_{gt}}$, for each group and time period. This estimator first constructs a counterfactual distribution, $F_{Y_{11}^{cf}}$, to predict the hypothetical cumulative distribution function of the treated group in the post-treatment period as if they were not treated. $F_{Y_{11}^{cf}}$ is non-parametrically estimated from the empirical cumulative distribution functions, $F_{Y_{gt}}$, in the pre ($t = 0$) and post ($t = 1$) policy periods from the comparison ($g = 0$) and treatment ($g = 1$) groups. Quantile-outcome pairs are identified by

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

Treatment effects are calculated by comparing the constructed counterfactual and observed outcome distributions for the $g = 1$ and $t = 1$ group at specific quantiles denoted by τ. Thus, the quantile specific treatment effect is given by $\Delta y(\tau) = F^{-1}_{Y_{11}}(\tau) - F^{-1}_{Y_{11}^{cf}}(\tau)$, which is the increase (decrease) in score at quantile τ in the unconditional distribution In our analyses, we additionally conduct Kolmogorov-Smirnov tests of the equality between the observed ($F_{Y_{11}}$) and counterfactual ($F_{Y_{11}^{cf}}$) empirical distribution functions to formally test if the policy had an influence on the distribution as a whole.

To more formally introduce the changes-in-changes estimator proposed by Athey and Imbens (2006) we begin by noting that the functional form dependence in how D, t and g influence Y is a point of departure from the linear DID estimator.[13] The estimator assumes that in each

[12] Identification of this causal parameter relies on assumptions of i) common support, ii) common trend, and iii) no anticipation effects. See Kottelenberg and Lehrer (2013a) for a discussion of the validity of these assumptions in the context of evaluating the effects of the Quebec Family policy.

[13] We did not consider using a quantile difference-in-differences (QDID) estimator to estimate the counterfactual distribution since Athey and Imbens (2006) discuss several advantages of the change in changes estimator.

time period the unobserved inputs to the production process are strictly monotonically related to the outcome. This allows the same realization of Y to have a one to one correspondence to a specific realization of unobserved factors in each time period, regardless of the group / province. Further, to identify causal effects, Athey and Imbens (2006) make an assumption of time invariance on the distribution of unobserved factors within each group.[14] However, they do not require that the distribution of unobserved factors to be identical across groups. For example, the treatment group may contains more high ability children than the control group.[15] The combination of assumptions of the time-invariance in unobservables within each group and strict monotonicity in the effects of unobserved factors on outcomes permits a relaxation of the common trend assumption underlying the linear DID estimator. That is, the time effects are allowed to differ across individuals with different observed characteristics.[16]

The steps involved in calculating the quantile treatment effects are illustrated in Figure 1. In the first panel we determine what quantile (τ) of $F^{-1}_{Y^{cf}_{11}}$ for a given test score (y) is the point being calculated by using the transformation $F_{Y_{00}}(F^{-1}_{Y_{10}}(\tau))$. Recall, in the same time period, specific realization of unobserved factors corresponds to a unique y, indicating there is a one-to-one correspondence between τ and the unobserved factor. The second panel calculates the time effect for that specific level of unobserved factors, from comparing the quantile functions of $F_{Y_{01}}$

First, the QDID requires that the distribution of unobservables is identical in all subpopulations, eliminating potential source of intrinsic heterogeneity. Second, the QDID model must place additional restrictions on the data for the transformation required to be monotone.

[14]This assumption is needed to assume that any changes in the outcomes within the control group are due to time effects and not from changes in the distribution of unobserved factors.

[15]Selection on unobservables is possible and making 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. Thus, this less restrictive model allows for the possibility that Quebec adopted their child care system based on higher expected returns than would otherwise be expected in the rest of Canada.

[16]While group and time implicitly impose constant effects with linear DID estimators, their effects in the change in change model can systematically differ across individuals. Groups are also no longer treated symmetrically; rather, treatment may affect the treated group differently than the control group since they have a different distribution of unobserved factors.

and $F_{Y_{00}}$. Given the time-invariance of the unobserved factor in a group, the $F^{-1}_{Y_{01}}$ transformation in equation (1) denotes adding the time effect to the value corresponding to the τ^{th} quantile of $F_{Y_{10}}$ as shown in the third panel.[17] This creates the value of y at the τ^{th} quantile of $F_{Y_{11}^{cf}}$. The last panel then calculates the quantile specific treatment effect $\Delta y(\tau)$ by taking the difference between $F^{-1}_{Y_{11}}(\tau)$ and $F^{-1}_{Y_{11}^{cf}}(\tau)$.

3.1 Controlling for Covariates

A potentially important challenge in applying this approach to evaluate child care policy is the presence of compositional changes, in that the introduction of a subsidized child care program in Quebec not only increased formal child care arrangement and in doing so more diverse individuals attended child care.[18] It is plausible that parents with low reservation prices for child care may have sent their children to these centers over time, in which case the decline in developmental outcomes can simply reflect the compositional change in the treatment group. The assumption of no compositional changes can be stated in terms of time invariance of unobservables within groups (see Assumption 3.3 in Athey and Imbens, 2006). In this paper, we aim to balance the observables across time and group, in the hope that this ensures the identifying assumption holds.[19]

[17] The second transformation $F^{-1}_{Y_{01}}$ in equation 1 adds this time effect to a person with the same value of the outcome variable in $F_{Y_{10}}$ but they may correspond to a different unobserved factor than in $F_{Y_{00}}$. Thus, we are saying that individuals in the treated place with a particular outcome at different points in time would be expected to experience the same time effect if the treatment was not offered. Any difference is due to the treatment. Identification of causal effects relies only on an assumption of strict monotonicity in the effects of unobserved characteristics on outcomes, a time invariance condition and that there is some overlap in the support of the unobserved factors between the treated and control.

[18] With a linear difference in differences estimator, one would simply add controls to the estimating equation to ensure balance.

[19] To investigate the sensitivity of our results, we later consider indirect tests of this assumption proposed by Bitler et al. (2005). An alternative restatement of the rationale for this exercise is that the effect of the child care policy could operate through two channels, i) the true effect of child care on specific outcomes, and ii) changes in the composition of those who attend child care. By controlling for covariates we aim to rule out the second channel.

To balance observables, we rely on Firpo's (2007) extension to quantiles of the inverse propensity scores method introduced in Hirano, Imbens, and Ridder (2003).[20] We create the following weight,

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

where N is the number of observations and $\hat{p}(X_i)$ is the estimated probability of a child being in Quebec after the implementation of the policy. This predicted probability is obtained through a series logit estimation, which incorporates the full set of 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-linear DID models.[21] 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 the observable covariates from the different sub-populations are balanced.[22] Using the estimated inverse

[20]The main empirical strategy in Havnes and Mogstad (2014) used to recover distributional effects and control for covariates applies the re-centered influence function (RIF) regression estimator of Firpo, Fortin and Lemieux (2009) to a linear DID equation. They justify this estimator on the basis of computational convenience and assume an identical production technology across groups. We do not follow this strategy since the estimator is developed for cross sectional data and not for two-way panel data settings with additive group and time fixed effects. While other empirical papers, including Meyer et al. (1995) and Poterba et al. (1995), used quantile estimates of a linear DID equation, those authors imposed that the quantile treatment effects across time are identical across all quantiles. In illustrating their strategy, Havnes and Mogstad (2014) compare individuals with the same outcome and assume that the time and group effects are constant for that outcome value but both effects can vary across the outcome distribution. However, having different time effects violates their identifying assumption (also discussed in footnote 13) that unobserved factors are equally distributed in the treatment and control groups that Athey and Imbens (2006) prove is needed to point identify impacts at each percentile for this type of model. Last, we note that Powell (2014) has developed an estimator for unconditional quantile regression with panel data but the interpretation of the estimates is conditional on the fixed effects which differs from the Firpo, Fortin and Lemieux (2009) estimator that is based on the unconditional distribution.

[21]The results presented in the next section are robust to calculating $\hat{p}(X_i)$ with a logit estimator of a specification that does not include the set of interactions between the covariates. All of our estimation code is available upon request.

[22]This method of reweighting the distribution functions is identical to the first three steps described in Appendix B of Bitler er al (2006). Note, we follow the implementation details provided in the supplement to Athey and Imbens (2006), thereby ensuring that the support conditions are satisfied in the sample. In practice, this influenced very few observations essentially only influencing estimates of the highest and lowest percentile of the quantile treatment effects.

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. As such, concerns that unaccounted for observables are driving the observed treatment effects are reduced and potentially eliminated.

To conduct statistical inference at each percentile we use Fisher's exact permutation test since it has better small sample properties. Specifically at each percentile we first conduct a test to measures the significance of the deviation of $\Delta y(\tau)$ estimate from a null hypothesis that the effect is 0.[23] In addition, we conduct a second test to measures the significance of the deviation of our estimate of $\Delta y(\tau)$ from the mean policy estimate from the change in changes estimator.[24] In effect this second test aims to determine if there is treatment effect heterogeneity. While unreported the mean policy estimate from the change in changes estimator are quite similar to the mean intent to treat effects using a linear DID estimator reported in either Baker et al. (2008) or Kottelenberg and Lehrer (2013a).

4 Empirical Results I: Distributional Effects

Table 2 presents results from Kolmogorov-Smirnov tests of the equality of the cumulative distribution function of child development scores ($F_{Y_{11}}$ vs. $F_{Y_{11}^{cf}}$) separately for single-parent and two-parent families. For each sample, with or without covariates, the results clearly indicate

[23]The exact procedure for this specific test follows. We conduct 999 experiments where we estimate the treatment effect using the change in changes estimator in which the treatment is randomly assigned to an individual. For each percentile, the randomly assigned treatment effects are then sorted from smallest to largest. In the results sections we denote the percentiles for which the estimated quantile treatment effect larger than 90% of the experiments. This inference is more conservative than simply using a bootstrap and we thank Guido Imbens for suggesting that it is more appropriate in this setting.

[24]A similar process is undertaken to conduct this second test. In effect, this test is conditional on the estimated mean effect from the change in changes estimator as we first subtract this value from the observed developmental scores of children from Quebec after the policy was introduced. We then repeat the same set of experiments outlined in the previous footnote. The estimated quantile treatment effect which are larger than 90% of the experiments correspond to percentiles at which treatment effects are significantly different from the mean effect at 10% level. Significant results are denoted separately in the graphs presented in the results section.

that the policy made a difference on the distributions of both the MSD scores for children aged 0-3 and the PPVT scores for 4 year olds. In all cases, the observed ($F_{Y_{11}}$) and counterfactual ($F_{Y_{11}^{cf}}$) empirical distribution functions differ.[25]

While the results in Table 2 suggest that the policy did alter the distribution of child development, it does not shed any light on the sign and magnitude of the estimated policy effect as well as identify at which percentiles is the $\Delta y(\tau)$ statistically significant. Figure 2 presents change in changes estimates of the effects of the policy on MSD scores. Notice first that the unweighted figures for the sample of two-parent families demonstrate that the negative impacts reported in Baker et al. (2008) are driven by the lower half of the distribution. Further, all the evidence of treatment effect heterogeneity comes from both the very bottom and the near the top of the distribution where the $\Delta y(\tau)$ are significantly different from the mean effects. While there are a fewer significant results once we use weights to balance observables, we note that these results suffer from reduced power due to the additional randomness added from incorporating estimation error in the weighting process. As a whole, these results suggest that the significant negative effects reported in Baker et al. (2008) come from the bottom 50 percent of the distribution and that there are some percentiles in the top quantile of the distribution where we find evidence that children do not see statistically significant declines in their performance on the MSD scores.

The results for the children from single-parent families show strikingly different patterns. The impact of child care access is positive at nearly every percentile and is particularly large for individuals at the very bottom of the distribution. These results are consistent with evidence from the Perry Preschool program that shows that among the most disadvantaged, who come

[25] The results for the subsample of 4 year old children of single parents relies on few observations from Quebec pre-policy. Due to the limited sample size, we have less confidence in results for this subgroup. With those caveats stated, we also note that Monte Carlo evidence in Abadie (2002) has indeed shown that the statistic for this test requires a relatively large number of data points to properly reject the null hypothesis; which makes the results in Table 2 somewhat more striking.

primarily from single parent households, there are gains from access to subsidized child care. Further comparing the panels demonstrates that when we account for observables the magnitude of the estimated effects tend to increase in size. As a whole, the panels in Figure 2 demonstrate that there is substantial heterogeneity in the effects of introducing child care on MSD scores and that some individuals do achieve gains.

In Figure 3, we present change in changes estimates of the policy effects on PPVT scores for two parent families only.[26] Baker et al. (2008) report a statistically insignificant effect from access to universal child care on this outcomes. However, our distributional estimates appear similar to the pattern for MSD scores among two parent households, where without covariates we see large negative statistically significant effects from the 5th to the 40th percentile. Many of the effects from the 5th to 20th percentile are double the size of the mean policy estimate. In contrast, estimated effects from roughly the 45th to 60th and 70th to 90th percentile are zero and are statistically different than the mean policy estimates. Finally, we see the pattern of $\Delta y(\tau)$ is robust to controlling for covariates.

Kottelenberg and Lehrer (2013b) show that the negative impacts from a linear DID estimate reported in Baker et al. (2008) differ sharply by child gender. Figure 4 shows that not only do the mean effects differ between the genders, the pattern of treatment effect heterogeneity varies markedly. Among children aged 0-3, we find that there is substantial heterogeneity among girls, where those in the bottom quartile exhibit significant declines in MSD scores but there are many percentiles in the top two quartiles where girls gain from access to universal child care. However, there is no heterogeneity among girls on the PPVT. The exact opposite pattern in $\Delta y(\tau)$ occurs for boys who exhibit little if any heterogeneity on MSD scores, whereas they exhibit substantial heterogeneity on the PPVT score. Boys at percentiles in the bottom half

[26] With the limited sample size, that is fewer than the number of percentiles, we do not present results of the policy effects on PPVT scores for single parent families but do note in Table 2 that in this case $F_{Y_{11}}$ and $F_{Y_{11}^{cf}}$ are statistically different.

of the PPVT score distribution exhibit significant declines, whereas boys between the 50th and 85th percentile are unaffected by the policy but do achieve an affect that is statistically different from the mean.

As indicated in above each graph in Figure 4, Kolmogorov-Smirnov tests of the equality of the cumulative distribution function of child development scores reject the null hypothesis in all cases whether we account for observables or not. As a whole, the analysis presented in Figures 2, 3 and 4 demonstrates that i) the variability in the policy effects within groups defined by child characteristics is substantial as compared to the variability in policy effects across groups, and ii) that many of the negative and statistically significant effects reported in Baker et al. (2008) are driven by children from two parent families in the bottom half of the distribution.

4.1 Examining the sensitivity to the assumptions underlying the analyses

The assumption of rank invariance that underlies the change in changes estimates presented in Figures 1 to 3 is potentially quite strong. This assumption requires that the rank of the potential outcome for a given individual within group and across time, would be the same under treatment as under non-treatment. If this assumption holds, we can interpret the horizontal distance between the actual and counterfactual cumulative distribution functions as a causal effect for individuals at that particular percentile of the control group outcome distribution. If this assumption is violated we are unable to make any inference on the impact on any particular person and the horizontal distance only tells us by how much more did the outcome distribution change among the treated relative to the control over time. Thus, concerns of whether the presented results can be given a causal interpretation are reasonable.

We follow Bitler et al. (2005) who propose an indirect test for whether rank invariance holds. The intuition behind the test is to check if the distribution of observable characteristics at quantile ranges of the outcome distribution vary significantly between Quebec and the rest of Canada in either pre or post policy periods. If differences are found there would be evidence against rank invariance. As such, we examine the full set of demographic controls and conduct this test for both groups g and time periods t. Our results indicate that in only 4.9% of all of the variables investigated, there is a sufficient change in the characteristics between the groups over time for two-parent families.[27] Given the strength of this assumption, this result increases our confidence in the empirical results and their interpretation.

4.2 Empirical Results II: Household Responses

The difference in the distributional pattern and sign of the estimated effect between single and two parent families in Figure 2 is particularly striking. To shed light on this finding, we draw on the framework of Becker's (1981) model of household production. We suppose that parents make simultaneous decisions about employment and household production—including fertility and child rearing—subject to their budget constraints, time constraints, skill sets, and preferences.[28] Child developmental outcomes could be influenced though several channels including time investments by parents, time and financial investments in other child care, and monetary investments in other goods and services such toys, books, or lessons. No restrictions are imposed on how parent time investments and child care enter the skill development production functions, thus allowing them to serve as potential substitutes as posited by those who champion the notion that child care may level the playing field between children of different backgrounds.

In equilibrium, parents make leisure and child time investments into their children such that the marginal rate of substitution between leisure and parental time investment is equal to the

[27]The results do not differ sharply for either the subsamples of single parent families or by child gender.

[28]Without imposing structural assumptions it is not possible to isolate through which causal pathways is the reduction in the price of child care affecting family well-being.

price of leisure divided by the price of child care. The Quebec Family policy reduces the price of child care faced by parents in Quebec relative to other investments parents can make to the production process. Thus, subsidized child care may lead parents to shift their investments away from other child rearing practices even if there were no change in labor supply.

A strength of the NLSCY data, one that has not been fully utilized in earlier research evaluating the Quebec family policy, is the collection of information on the nature and quantity of parental time spent with their children. Since children's readiness to learn at school has been linked to their experiences in the home and community, we make use of questions examining the extent to which parents are active with their child in recreational and educational contexts. Second, the NLSCY measures parenting behavior by scores on 5 different parenting scales: 1) a family dysfunction score; 2) a punitive aversive score; 3) a hostile/ineffective score; 4) a consistency score; 5) and a positive interactions score.[29] Since it is difficult to reasonably interpret several of the parenting measures as a continuous variable, we examine several of the individual questions that are contained within the scales and discretize the outcome if the parent reports having achieved a given level of certain activity.

We first present some descriptive evidence of heterogeneity in parental responses across quintiles of the unconditional test score distribution. Summary statistics on uptake, parenting scales and parent-child activities are presented across quintiles of the unconditional test score distribution for children aged 0-3 and aged 4 in Tables 3A and Table 3B.[30] These tables are also broken down by family type. Among two parent families there are few and relatively small

[29]For example, the positive interaction scale includes 5 questions such as "How often do you and s/he laugh together?" Each of the four scales were derived by factor-analyzing parenting items included in the NLSCY (Special Surveys Division, 1996) and have been shown to have high levels of internal consistency (e.g. Jenkins et al., 2003). There are five ordinal responses to the questions on these scales that range from "never" to "many times each day."

[30]We do not focus on maternal labor supply here. Lefebvre and Merrigan (2008, 2009) and Baker et al. (2008) find evidence that the policy led to increased maternal labor supply. This result contrasts Fitzpatrick (2010) who finds no evidence that either the Georgia or Oklahoma universal child care programs increased the labor supply of mothers of 4 year olds, despite providing a 100% price subsidy for child care on the extensive margin of employment.

differences in the uptake variables across the quintiles. Parenting scales are generally lower among single parent families and are concentrated in the bottom quintiles despite lower child care use and labor supply. The most heterogeneity across the test score distribution appears in the parent child activity measures. At the lower quantiles single parents are much less likely to engage in activities with their child than two-parent families. At the higher quantiles these differences between family types are much less marked. These summary statistics, together with the predictions of the simple model outlined above, lead us to hypothesize that changes in child rearing at home may explain some of the heterogeneity in the estimated policy effects.

To investigate this hypothesis, we next examine whether there are differential trends in parenting practices over time across the unconditional test score distribution between Quebec and the Rest of Canada. Since there are numerous groups, time periods and quantiles of the developmental score distribution, we summarize the changes by presenting nonparametric DID estimates over subsamples of the MSD and PPVT distribution.[31] Presented in Table 4A and 4B, these estimates show that among single parent families the gain in maternal labour supply on the extensive margin came from children who scored in the bottom quintile. There were large gains in child care use and hours spent in child care for children from these families, and these children were the only quintile to exhibit statistically significant declines in nearly every parent child activity. Yet, as documented in Table 3A and 3B these families had much lower means on these variables and Figure 2 documents large policy effects. This set of results is consistent with a hypothesis that gaining access to child care may level the playing field for the most disadvantaged chhildren since they may move from less stable home environments to formal child care settings.

Among two parent families there are very few differential effects across the quintiles for children aged 0-3. We do see declines in parent-child activities; these are concentrated among

[31]The overall intent to treat effects without covariates reported in Baker et al (2008) for the full sample of two parent families are simply weighted averages of the effects reported by quintile of the unconditional developmental score distribution.

those in the bottom two quintiles. There are also negative consequences to parenting scales, primarily for children who score in the middle and top quintiles. Among children aged 4 we observed a significant positive effect in being read to daily for those in the top quintile. This was accompanied by no significant changes in the parenting scales. All declines in parenting scales are concentrated in the middle 3 quintiles where significant negative effects emerge.

This set of results suggests that the introduction of universal child care led to substantial changes in the manner in which parents invested in their children. Thus, the heterogeneity in the policy estimates on developmental outcomes may not be strictly due to participation in child care. Instead, we postulate that the availability of subsidized child care may change the nature and quantity of time mothers spend with children and that this may also contribute to the changes in child developmental scores. Our hypothesis is also consistent with findings from several recent studies in the economics literature.[32]

Havnes and Mogstad (2014) speculate that the negative effects from the Quebec family policy reported in Baker et al. (2008) arise from the middle and higher income families which disproportionately use this policy. Our estimates in Tables 4A and 4B suggest that there is no differential uptake and that it is unlikely that middle or higher income two parent families choose lower quality care. Taken with the higher returns to disadvantaged children from single-parent families, it is likely that the negative effects to middle and higher income families result from child care being less than a perfect substitute for investments that were previously made in the home.

To identify a causal impact of how parenting practices change on average for those parents

[32]For example, Gelber and Isen (2012) present convincing evidence that a significant portion of the positive effects of Head Start on child outcomes derive from changes in how parents invest into their children. Cascio and Whitmore (2013) use data from the American Time Use Surveys and present suggestive evidence that on average less-educated mothers in Georgia and Oklahoma spend 46 fewer minutes per weekday in the presence of their 4-year-olds following the introduction of universal Pre-K in those states. Last, Bettinger et al. (2014) present evidence of long term gains on older sibling's academic performance from a program in Norway that increased parents' incentives to stay home with children under the age of 3. The authors conclude that parental care is not easily substituted by formal and informal methods of care.

who send their children to child care due to the policy, we use an instrumental variables estimator. That is, we express the relationship between a specific parental outcome of interest Y and a set of explanatory variables as a linear equation:

$$Y_{ipt} = \beta_o + \delta' Ccare_{ipt} + \beta_2' PROV_p + \beta_3' YEAR_t + \beta_4' X_{ipt} + \varepsilon_{ipt} \tag{3}$$

where Ccare is an indicator for using child care and i, p, and t index individual child, province, and year. The vector of covariates X, includes controls for child, parent, family, and geographic characteristics,[33] PROV and YEAR are respectively a series of province and time dummies.

The main empirical challenge in estimating Equation (3) is that the decision to send one's child to formal care reflects a behavioral choice that is determined in conjunction with other unobserved confounders that may also affect Y. To overcome the endogeneity of child care attendance we treat the Quebec Family Policy as an encouragement design and estimate the following first stage equation:

$$Ccare_{ipt} = \gamma_0 + \gamma_1' Policy_{pt} + \gamma_2' PROV_p + \gamma_3' YEAR_t + \gamma_4' X_{ipt} + u_{ipt}. \tag{4}$$

The *Policy* variable is simply an indicator for living in Quebec after 1998, the year the Quebec Family Policy was introduced. Provided that the policy does not influence child rearing through any other channel than child care attendance, it satisfies the exclusion restriction. Last, note the main specification used in Baker et al. (2008) is simply the reduced form of the last two equations.

Table 5 presents IV estimates of Equation (3) for groups defined by household structure and child gender. Odd columns refer to children aged 0-3 and even columns contain results for 4-

[33]We use the exact same set of controls as Baker et al. (2008) a subset of which is presented in Table 1. To reduce issues related to mis-specifying the functional form of the estimating equation, all variables included are discretized. For example, we create a host of discrete dummy variables in the following categories: number of siblings, community size, and parental age, education, and immigration status.

year olds. Among the two parent households, we see large and statistically significant effects of child care attendance on each of the parenting scales with the exceptions of parent consistency and family dysfunction for 4 year olds. In all cases, the statistically significant effects suggest a worsening of the parental activity. Conversely, there are fewer significant impacts of child care on parenting in single parent families presented in columns 3 and 4. However, in these households, the estimated effect of child care attendance has i) a different trend for positive interaction, and ii) a sharp increase with child age for aversive parenting.

Looking across the first four columns, we see large differences in the estimated effects of child care attendance on parent child activities between two parent and single parent families. In both groups, child care attendance reduces laughing with a child, an effect that grows with child age. Reading to a child daily declines for those 0-3 but increases for children aged 4 in both types of families.[34] This may help explain why there are differential effects of the policy on MSD and PPVT scores. Last, two parent families are found to be less likely to engage in special activities and focused time with a child attending child care. We conjecture these declines may arise since parents assume these activities are being conducted in formal child care.

In the last four columns of Table 5, we present IV estimates by child gender among the two parent households corresponding to Figure 4. Notice that parents of girls aged 0–3 exhibit much larger declines from attending child care in four parenting activities, particularly being read to daily. This may provide an explanation for why there is much more heterogeneity in the policy effects presented in the top panel of Figure 4 for girls. Much like we observed larger heterogeneity in policy effects on PPVT scores for boys in Figure 4, the IV estimates in Table 5 are suggestive of larger parental responses than for 4 year old girls. As a whole, the results in the last four columns are consistent with larger responses in parental child rearing practices

[34] The increase in reading is consistent with US evidence presented in Cascio and Whitmore (2013) that the introduction of state universal child care policy that targeted only 4 year olds led to increased time spent by parents reading. Last, we should note that the larger magnitude of the IV estimates in Table 5 perhaps reflecting heterogeneity among the sample that identifies the local average treatment effect.

being associated with more heterogeneity in the effects of the policy on developmental scores in Figure 4.

The bottom row of Table 5 presents results from an F test for the joint significance of the instrument from the first stage regression. They are well above current cutoffs (i.e. Staiger and Stock, 1997) for weak instruments. This is not surprising since the policy leads to substantial uptake.[35]

As a whole, the results in Tables 3 to 5 illustrate that i) the children in single parent families who score poorly receive very low levels of home investments relative to other children, ii) the Quebec Family policy led to differential reductions, on average, in how parents invested into their children, iii) the reductions were quite large and statistically significant among two-parent families who sent their children to child care in response to the availability of this policy, and iv) the trends in the reductions in certain parent-child activities and parenting scales across the unconditional developmental score distributions generally mirror the shape of the estimated policy effects for two-parent families in Figures 2, 3 and 4. In contrast, in two-parent families the effects of the policy on maternal labor supply and child care usage do not exhibit a pattern consistent with Figures 2, 3 and 4. Without imposing further structural assumptions we can only propose that a significant portion of the negative effects of subsidized child care on developmental outcomes derives from changes in how parents in two-parent families invest into their children.[36] For single-parent families, there are gains in the bottom MSD score quintiles that exhibit very low parent-child activities and large uptake in both child care and labor supply. It may be that formal child care provides additional resources for children from these

[35]Kottelenberg and Lehrer (2013) also use an IV estimator with a similar data extract and present evidence of the robustness of their results to controlling for maternal labor supply and other potential violations of the exclusion restriction assumption.

[36]That said, it is important to note that an additional limitation of the parental investment measures investigated is that they only measure investments as a flow at a certain point in time, rather than a stock that has accumulated since birth. Thus, we are also unable to rule out that differences in the stock of parental investments explain the heterogeneity in policy effects.

households.

5 Conclusion

Quebec's subsidized child care policy is unique and is often portrayed as a model for other provinces, states and countries to follow. Yet, the pattern of results reported in prior studies that have evaluated this policy are at best mixed and have led many to speculate about the underlying causal mechanism. In this paper we extend earlier research evaluating the Quebec Family Policy by providing the first evidence on the distributional effects of universal child care on two specific developmental outcomes. Our analysis uncovers substantial policy relevant heterogeneity in the estimated effect of access to subsidized child care across two developmental score distributions for children from two-parent families. Our estimates first reveal a more nuanced image that formal child care can indeed boost developmental outcomes for children from single parent households, particularly for the most disadvantaged within. We also find striking differences in the distributional effects across child gender that differ across developmental scores.

Our empirical analysis is suggestive that much of the heterogeneity in policy effects emerges from differences in home learning environments that were present prior to the policy and were altered in response to the policy. In particular, we find that the subsamples exhibiting the largest negative effects from gaining access to child care also have large declines in parenting practices such as reading to the child daily for those who took up child care in response to the policy. Gains in developmental scores, on the other hand, were experienced by samples in which there were no or small positive changes to child rearing practices following the introduction of the policy or in which the pre-existing home environment was extremely poor. As a whole, this suggests that while formal child care is not a perfect substitute for home learning environments, given the large number of hours spent in child care centers, it can provide a remedy for children from the most disadvantaged home environments.

While future research is needed to improve our understanding of how child care, parents and government policies interact to influence developmental outcomes, we believe there are three issues this paper highlights that can inform current child care policy debates. First, while the substantial heterogeneity in policy effects may appear to complicate the issue, it really points out that the evidence base surveyed in Baker (2011) is consistent with targeting child care coverage as a more effective policy option relative to universal coverage.

Second, the success of child care with disadvantaged populations witnessed in many experimental studies has been hypothesized to level the playing field, since it on average does more than substitute for the parental care that these children would otherwise receive at home. We find that among children from single parent families, subsidized child care appears to substitute for lower levels of parental care or informal care arrangements, in effect leveling the playing field. These children witness large gains in developmental scores after the policy. However, our analysis also indicates that once the policy was introduced, children from two-parent households in Quebec on average shifted from receiving otherwise strong one-on-one parental care at home to reductions in many parent-child activities coupled with potentially less effective higher adult-child ratio care giving away from home. In addition, Baker et al. (2008) and Kottelenberg and Lehrer (2013a) show that these children, while in the home, were on average subjected to working parents affected by potentially higher levels of stress. Thus, since there are many significant household responses to these policies, we suggest attention in policy design must be devoted to mitigating household responses that influence the home learning environment.

Third, the observed heterogeneity in policy impacts is not captured by comparing estimates of mean effects across subgroups of children defined by observed demographic characteristics. The treatment effect heterogeneity that we observe in the full sample also exists in subpopulations defined on the basis of child age and we see substantial different patterns for each developmental score by child gender. This reinforces that since the treatment effect heterogeneity is not fully characterized by predetermined variables, an improved understanding of why the policy works, when it does, and why it fails when it does not, is needed.

References

[1] Abadie, A. (2002), "Bootstrap tests for distributional treatment effects in instrumental variable models," *Journal of the American Statistical Association* 97(1), 284–292.

[2] Angrist, J. D. P. Oreopoulos, and T. Williams (2013), "When Opportunity Knocks, Who Answers?: New Evidence on College Achievement Awards," *Journal of Human Resources*, 49(3), 572–610.

[3] Athey, S., and G. W. Imbens , (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica* 74: 431–497.

[4] Baker, M. (2011). "Innis Lecture: Universal early childhood interventions: what is the evidence base?" *Canadian Journal of Economics* 44(4): 1069–1105.

[5] Baker, M., J. Gruber, and K. Milligan (2008). "Universal Child Care, Maternal Labor Supply, and FamilyWell-Being," *Journal of Political Economy* 116(4): 709–745.

[6] Becker, G. S. (1981), "*A Treatise on the Family*," Cambridge, Mass.: Harvard University Press

[7] Belfield, Clive R., Milagros Nores, Steven Barnett, and Lawrence Schweinhart. 2006. "The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup." *Journal of Human Resources*, 40(1): 162–90.

[8] Berlinski, S., S. Galiani, and P. Gertler (2009). "The effect of pre-elementary education on elementary school performance." *Journal of Public Economics*, 93(1–2): 219–234.

[9] Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006). "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiements," *American Economic Review* 96(4).

[10] Bitler, M. P., J. B. Gelbach, H. W. Hoynes (2005). "Distributional impacts of the self-sufficiency project.," *NBER Working paper #11626*.

[11] Cascio, E., and Schanzenbach, D. (2013). "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity* Fall, 127-178.

[12] Datta Gupta, N. and M. Simonsen (2010), "Non-cognitive child outcomes and universal high quality child care," *Journal of Public Economics* 94(1-2): 30–43.

[13] Ding, W. and S. F. Lehrer (2011) "Experimental estimates of the impacts of class size on test scores: robustness and heterogeneity," *Education Economics* 19(3), 229-252.

[14] Djebbari, H. and J. Smith (2008), "Heterogeneous impacts in PROGRESA", *Journal of Econometrics* 145(1-2), 64-80.

[15] Firpo, S., N. Fortin, and T. Lemieux (2009), "Unconditional Quantile Regressions," *Econometrica*, 77(3), 953–973.

[16] Firpo, S. (2007), " Efficient semiparametric estimation of quantile treatment effects." *Econometrica* 75(1), 259–276.

[17] Fitzpatrick, M. D., (2010), "Preschoolers Enrolled and Mothers at Work? The Effects of Universal Pre-Kindergarten." *Journal of Labor Economics* 28(1): 51-85.

[18] Gelber, A. M. and A. Isen (2013), "Children's Schooling and Parents' Investment in Children: Evidence from the Head Start Impact Study," " *Journal of Public Economics* 101(1), 25-38.

[19] Gormley, W. T., and T. Gayer (2005). "Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program" *Journal of Human Resources*, 40(3): 533–558.

[20] Havnes, T. and M. Mogstad (2014), "Is universal child care leveling the playing field?," *forthcoming in Journal of Public Economics.*

[21] Havnes, T. and M. Mogstad (2011), "No Child Left Behind: Subsidized Child Care and Children Long-Run Outcomes," *American Economic Journal: Economic Policy*, 3(2): 97–129.

[22] Heckman, J., Malofeeva, L., Pinto, R. and Savelyev, P. (2013). "Understanding the mechanisms through which an in uential early childhood program boosted adult outcomes," *forthcoming in the American Economic Review.*

[23] Heckman, J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010), "Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool 20 Program," *Quantitative Economics* 1(1), 1–46.

[24] Hirano, K., G. W. Imbens, and G. Ridder (2003). "Effcient Estimation of Average Treatment Effects Using the Estimated Propensity Score," *Econometrica*, 71(4), 1161-1189.

[25] Jenkins, J. M., J. Rasbash and T,. G. O'Connor (2003), "The role of the shared family context in differential parenting," *Developmental Psychology* 39(1), 99-113.

[26] Kottelenberg, M. J., and S. F. Lehrer (2014). "Do the Perils of Universal Childcare Depend on the Child's Age?," *CESifo Economic Studies* , 60(2): 338-365.

[27] Kottelenberg, M. J., and S. F. Lehrer (2013a). "New Evidence on the Impacts of Access to and Attending Universal Child-Care in Canada," *Canadian Public Policy*, 39(2): 263–286.

[28] Kottelenberg, M. J., and S. F. Lehrer (2013b). "The Gender Effects of Universal Child Care in Canada: Much ado about Boys?," Queen's University:*mimeo*.

[29] Lefebvre, P., P. Merrigan, and Francis Roy-Desrosiers (2011). "Quebec's Childcare Universal Low Fees Policy 10 Years After: Effects, Costs and Benefits," *CIRPEE Working Paper 11–01*.

[30] Lefebvre, P., and P. 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.

[31] Meyer, B., K. Viscusi and D. Durbin (1995), "Workers' Compensation and Injury Duration: Evidence from a Natural Experiment," *American Economic Review* 85(3), 322-340.

[32] Milligan, K., and M. Stabile (2007), "The Integration of Child Tax Credits and Welfare: Evidence from the Canadian National Child Benefit Program," *Journal of Public Economics* 91(1-2), 305-326.

[33] Powell, D. (2014), "Did the economic stimulus payments of 2008 reduce labor supply? Evidence from quantile panel data estimation," *RAND Working Paper 720*.

[34] Poterba, J., S. Venti, and D. Wise (1995), "Do 401(k) contributions crowd out other personal saving?" *Journal of Public Economics* 58(1), 1-32.

[35] Shonkoff, J., and D. Phillips D (eds) (2000). "From Neurons to Neighborhoods: The Science of Early Childhood Development." Committee on Integrating the Science of Early Childhood Development, Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education, National Research Council and Institute of Medicine. Washington, DC, National Academy Press.

[36] Tougas, J. (2002). "Reforming Quebec's early childhood care and education: The first five years," *Technical Report*, Childcare Resource and Research Unit, Toronto.

Table 1: Summary Statistics

	Two Parent Families				Single Parent Families			
	Quebec		Rest of Canada		Quebec		Rest of Canada	
	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy	Pre-Policy	Post-Policy
Child and Family Characteristics								
Child is Male	0.509	0.514	0.509	0.515	0.548	0.521	0.519	0.515
	(0.5)	(0.5)	(0.5)	(0.5)	(0.498)	(0.5)	(0.5)	(0.5)
Num of Older Siblings	0.715	0.676	0.795	0.753	0.530	0.648	0.683	0.680
	(0.739)	(0.714)	(0.762)	(0.745)	(0.724)	(0.738)	(0.774)	(0.772)
Num of Younger or Same Aged Siblings	0.268	0.222	0.255	0.246	0.129	0.179	0.160	0.138
	(0.488)	(0.44)	(0.476)	(0.468)	(0.335)	(0.454)	(0.403)	(0.366)
Lives in Rural Area	0.151	0.137	0.154	0.112	0.112	0.118	0.082	0.097
	(0.358)	(0.344)	(0.361)	(0.316)	(0.316)	(0.323)	(0.274)	(0.297)
Lives in Large City (>500K)	0.579	0.593	0.428	0.454	0.636	0.608	0.424	0.403
	(0.494)	(0.491)	(0.495)	(0.498)	(0.482)	(0.489)	(0.494)	(0.49)
Mother's Characteristics								
Age	30.926	31.501	31.737	32.677	29.269	31.274	28.829	29.589
	(4.878)	(4.939)	(5.123)	(5.437)	(6.962)	(6.565)	(6.679)	(7.342)
Immigrant Status	0.089	0.146	0.214	0.253	0.125	0.203	0.172	0.166
	(0.285)	(0.353)	(0.41)	(0.435)	(0.331)	(0.402)	(0.378)	(0.372)
High School Drop Out	0.133	0.099	0.106	0.074	0.391	0.243	0.287	0.221
	(0.34)	(0.299)	(0.307)	(0.261)	(0.489)	(0.429)	(0.452)	(0.415)
University Educated	0.203	0.353	0.206	0.352	0.090	0.120	0.044	0.107
	(0.402)	(0.478)	(0.404)	(0.478)	(0.287)	(0.326)	(0.205)	(0.31)
Father's Characteristics								
Age	33.507	34.285	34.136	35.293				
	(5.402)	(5.79)	(5.704)	(6.275)				
Immigrant Status	0.097	0.168	0.208	0.253				
	(0.296)	(0.374)	(0.406)	(0.435)				
High School Drop Out	0.168	0.123	0.138	0.096				
	(0.374)	(0.329)	(0.345)	(0.295)				
University Educated	0.194	0.303	0.214	0.313				
	(0.395)	(0.46)	(0.41)	(0.464)				
Key Outcome Variables								
Child in Care	0.415	0.649	0.405	0.452	0.263	0.694	0.410	0.576
	(0.493)	(0.477)	(0.491)	(0.498)	(0.441)	(0.461)	(0.492)	(0.494)
Mother Works	0.530	0.678	0.591	0.620	0.275	0.489	0.365	0.495
	(0.499)	(0.467)	(0.492)	(0.485)	(0.447)	(0.5)	(0.482)	(0.5)
MSD Score	99.317	96.526	100.395	98.564	98.189	97.159	100.754	97.322
	(15.031)	(15.171)	(15.343)	(14.979)	(16.034)	(15.743)	(16.044)	(16.458)
PPVT Score	99.764	100.359	100.511	101.513	94.497	93.162	95.072	100.004
	(15.139)	(15.235)	(15.277)	(15.382)	(15.276)	(15.516)	(14.231)	(14.181)
Sample Sizes								
Sample Size Age 0-3	2628	3730	10561	20533	313	378	1584	2972
Sample Size Age 4	526	765	2090	4368	61	145	359	687

— Note: Each row corresponds to a variable of interest and contains the mean and standard deviation (in parentheses) specific to the geographic region, time period and family type as denoted in the column header. The two final rows provides the sample size for these measurements based on child age cuts that will be used in the analyses. We report the summary statistics for single parent mothers and fathers together as 95% percent of single parents are mothers and we find no serious deviations between the summary statistics for these groups. The NLSCY survey weights, designed to accurately reflect the make up of the Canadian population, are applied in these and all calculations throughout the paper.

Table 2: Kolmogorov-Smirnov Tests on the Effect of Universal Childcare on the Distribution of Child Development Scores

	MSD Score		PPVT Score	
Sample	Unweighted	Weighted	Unweighted	Weighted
Two Parent Families	0.024**	0.016**	0.023**	0.000***
Single Parent Familes	0.025**	0.020**	0.000***	0.000***

— Note: The p-value are reported with ***, ** and * indicating significance difference of the equality of the distributions $F_{Y_{11}}$ and $F_{Y_{11}^{cf}}$ at the 1%, 5% and 10% level. Results are presented for each family type sample and developmental score for change in change estimates that are both unconditional and use propensity score weighting to control for covariates.

Figure 1: The Change-in-Changes Model

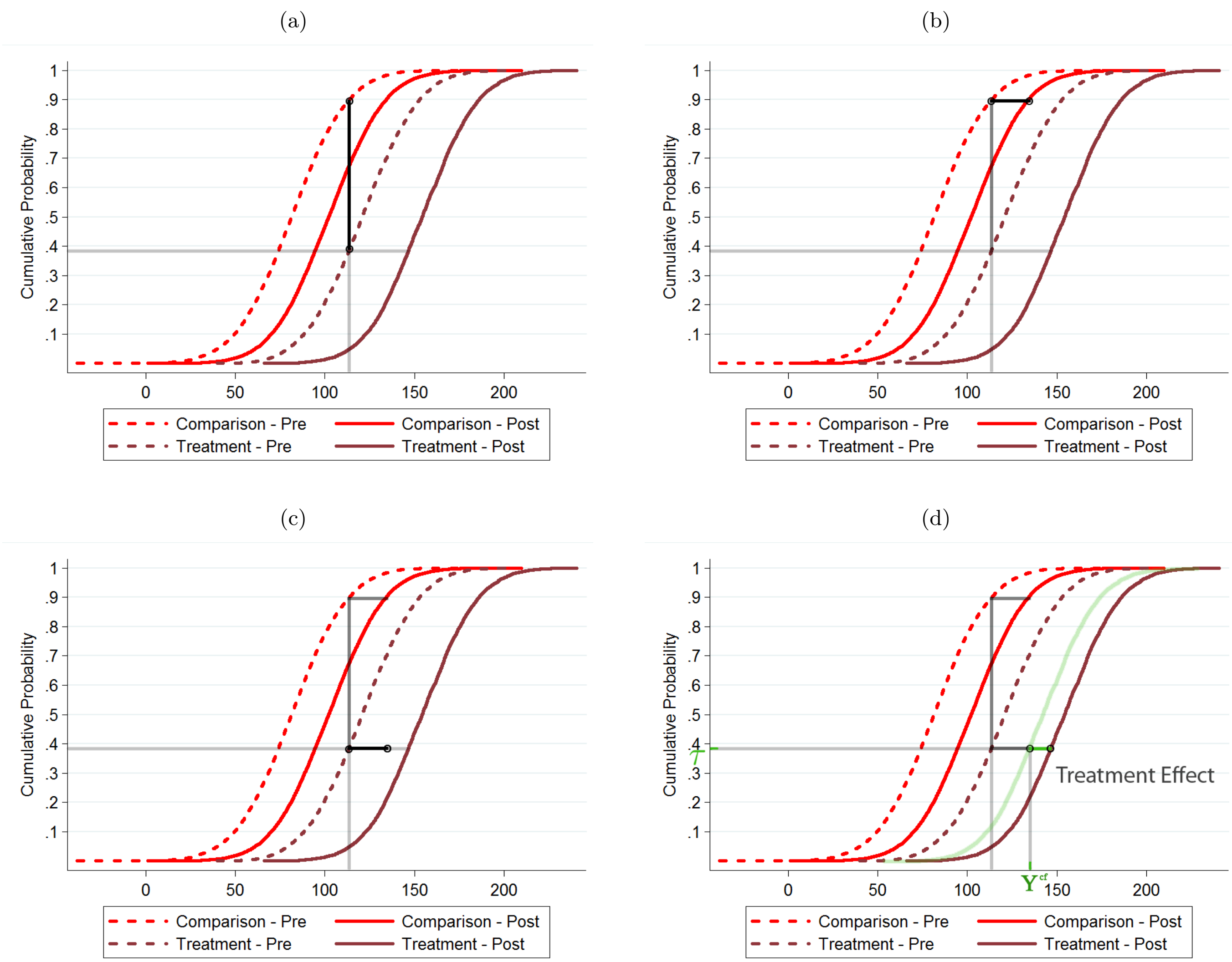
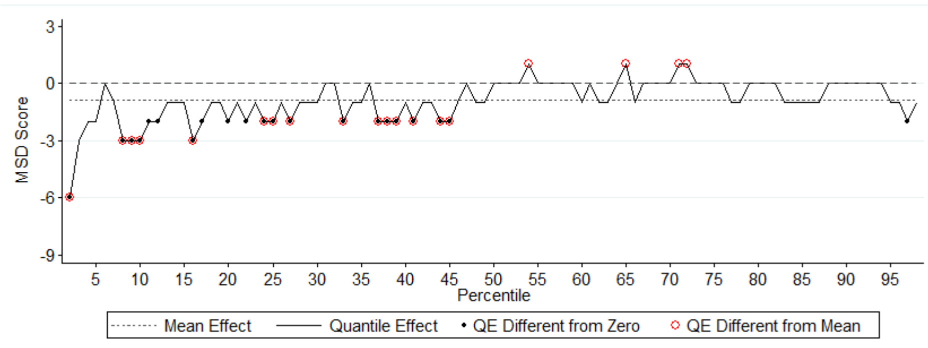
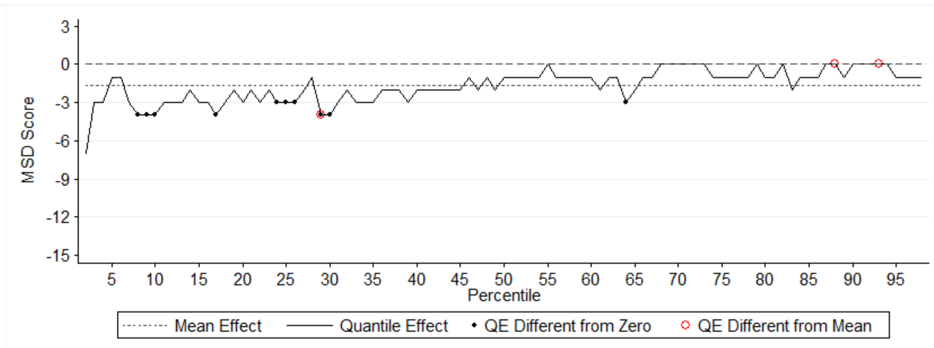


Figure 2: Change-in-Change Estimates for Motor and Social Development

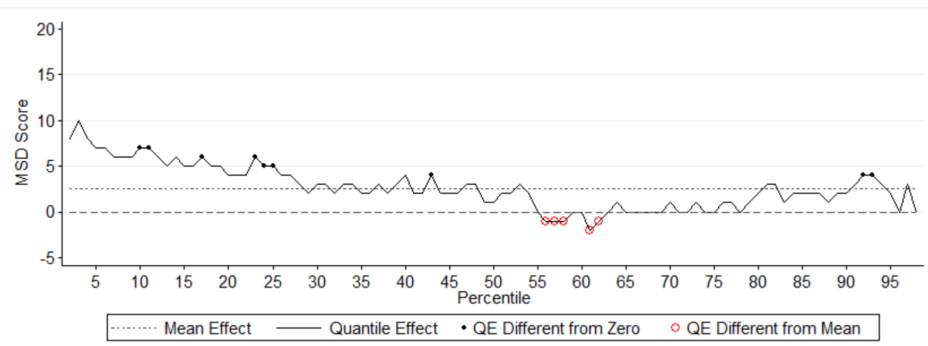
(a) Two Parent Families - Without Weights



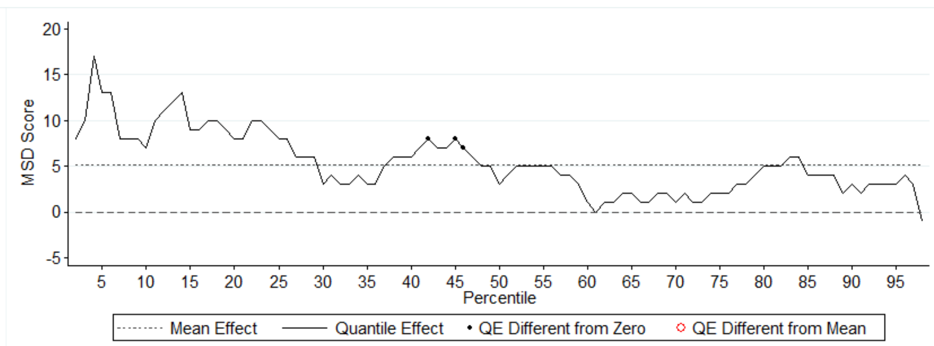
(b) Two Parent Families - Weights



(c) Single Parent Families - Without Weights



(d) Single Parent Families - Weights

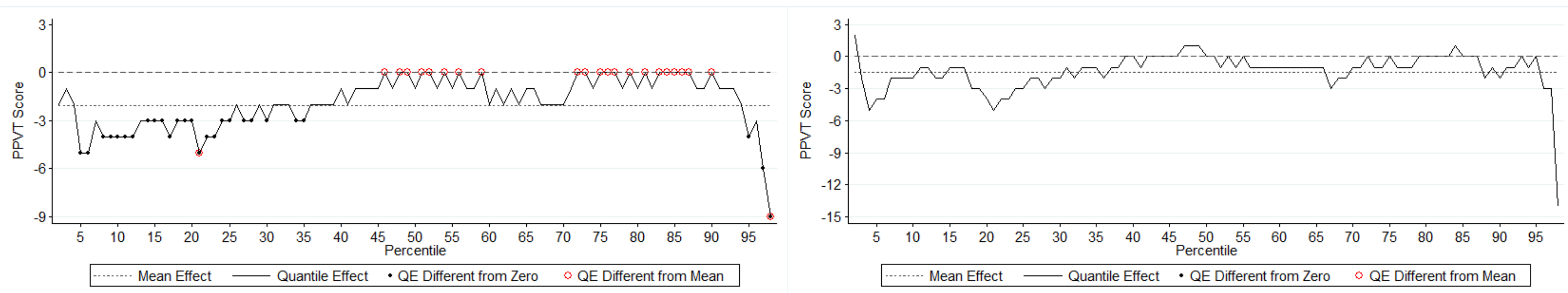


— Note: Change in change estimates are presented for each family type sample that are both unconditional and use propensity score weighting to control for covariates. The quantile effects, $F^{-1}_{Y^{cf}_{11}}(\tau) - F^{-1}_{Y_{11}}(\tau)$, is given by the solid line. The dash line gives the mean effect from the change-in-change estimate. Quantile effects that are significantly different from zero at the 10% level are marked with a solid dot. Quantile effects that are significantly different from the mean effect at the 10% level are marked by an open red circle.

Figure 3: Change-in-Change Estimates for Peabody Picture and Vocabulary Test

(a) Two Parent Families - Without Weights

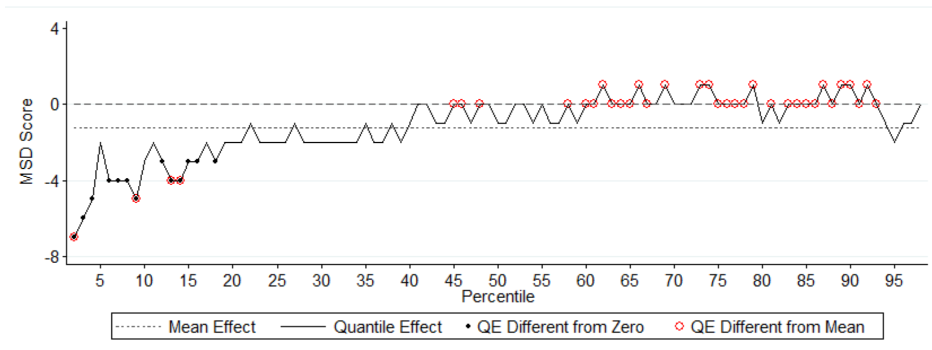
(b) Two Parent Families - Weights



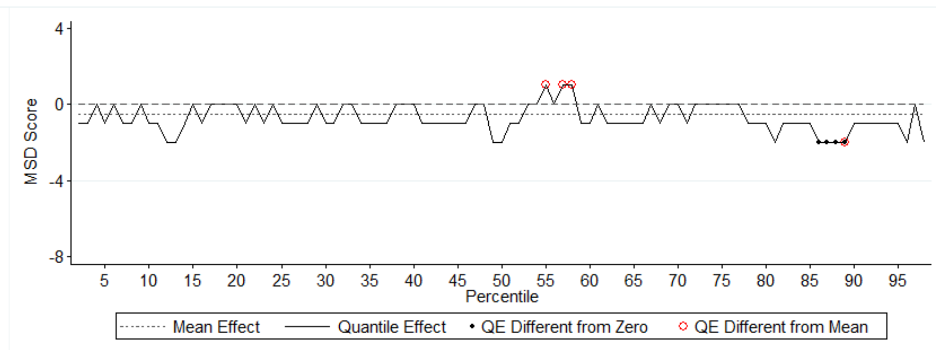
— Note: Change in change estimates are presented for each family type sample that are both unconditional and use propensity score weighting to control for covariates. The quantile effects, $F^{-1}_{Y^{cf}_{11}}(\tau) - F^{-1}_{Y_{11}}(\tau)$, is given by the solid line. The dash line gives the mean effect from the change-in-change estimate. Quantile effects that are significantly different from zero at the 10% level are marked with a solid dot. Quantile effects that are significantly different from the mean effect at the 10% level are marked by an open red circle.

Figure 4: Unweighted Change-in-Change Estimates for Two Parent Families by Gender

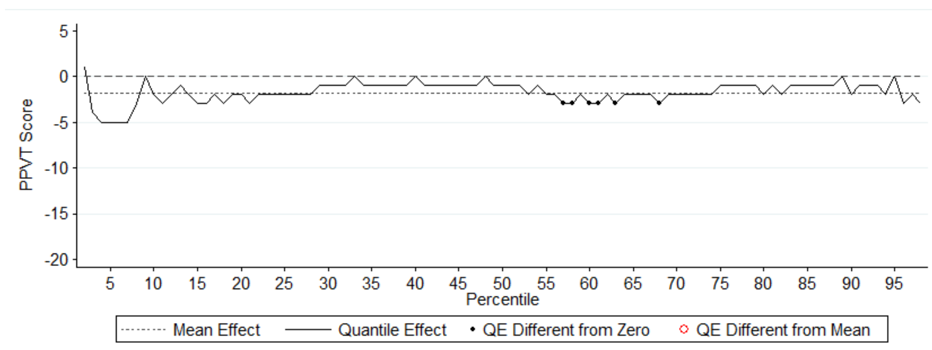
(a) Motor and Social Development - Girls
KS-Test: 0.006*** (0.004***)



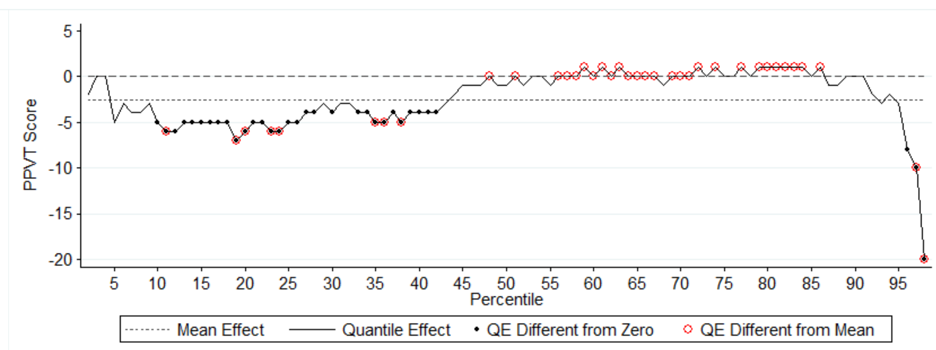
(b) Motor and Social Development - Boys
KS-Test: 0.003*** (0.004***)



(c) Peabody Picture and Vocabulary Test - Girls
KS-Test: 0.006*** (0.001***)



(d) Peabody Picture and Vocabulary Test - Boys
KS-Test: 0.007*** (0.000***)



— Note: The unconditional change in change estimates are presented for each development score by child gender of two parent parent families. The quantile effects, $F^{-1}_{Y^{cf}_{11}}(\tau) - F^{-1}_{Y_{11}}(\tau)$, is given by the solid line. The dash line gives the mean effect from the change-in-change estimate. Quantile effects that are significantly different from zero at the 10% level are marked with a solid dot. Quantile effects that are significantly different from the mean effect are marked by an open red circle. P-values from the Kolmogorov-Smirnov Test of distributions for the unconditional estimation (propensity score weighted estimation) above the diagram. ***, ** and * indicate significance difference of the equality of the distributions $F_{Y_{11}}$ and $F_{Y^{cf}_{11}}$ at the 1%, 5% and 10% level.

Table 3A: Summary Statistics by Quintile of Motor and Social Development Score for Children Age 0-3

	Two Parent Families					Single Parent Families				
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
Uptake Variables										
In Child Care	0.379	0.44	0.475	0.5	0.473	0.465	0.494	0.5	0.502	0.512
	(0.485)	(0.496)	(0.499)	(0.5)	(0.499)	(0.499)	(0.5)	(0.5)	(0.5)	(0.5)
Hours in Care	11.48	13.7	14.74	15.42	14.59	15.53	16.55	17.43	16.87	16.99
	(17.91)	(18.66)	(19.01)	(19.45)	(18.89)	(20.46)	(20.14)	(20.74)	(20.16)	(20.39)
Mother Works	0.544	0.595	0.612	0.637	0.611	0.358	0.414	0.407	0.444	0.434
	(0.498)	(0.491)	(0.487)	(0.481)	(0.488)	(0.48)	(0.493)	(0.491)	(0.497)	(0.496)
Mother Works 30 or more hours	0.362	0.407	0.411	0.427	0.406	0.255	0.274	0.278	0.295	0.312
	(0.481)	(0.491)	(0.492)	(0.495)	(0.491)	(0.436)	(0.446)	(0.448)	(0.456)	(0.464)
Parenting Scales										
Family Dysfunction Index	8.284	8.015	7.932	7.773	7.565	10.98	10.25	10.62	10.19	9.787
	(5.254)	(4.922)	(5.005)	(5.034)	(5.14)	(5.085)	(4.944)	(5.203)	(5.614)	(5.592)
Aversive Parenting	6.541	6.1	5.909	5.859	5.707	7.238	6.415	6.528	6.262	5.961
	(3.378)	(3.3)	(3.271)	(3.148)	(3.143)	(3.587)	(3.476)	(3.322)	(3.241)	(3.244)
Ineffective Parenting	9.325	8.992	8.817	8.685	8.31	9.844	8.898	8.935	8.904	8.165
	(3.939)	(3.574)	(3.512)	(3.372)	(3.433)	(3.724)	(3.7)	(3.455)	(3.667)	(4.016)
Parent Consistency	14.44	14.88	14.93	15.06	15.39	13.07	14.03	14.34	14.29	14.73
	(3.454)	(3.14)	(3.242)	(3.175)	(3.18)	(3.84)	(3.484)	(3.574)	(3.504)	(3.685)
Positive Interaction	16.11	16.34	16.62	16.58	17	15.46	16.18	16.37	16.59	17.35
	(2.562)	(2.377)	(2.328)	(2.282)	(2.23)	(3.053)	(2.35)	(2.361)	(2.303)	(2.05)
Parent-Child Activities										
Spends 5 minutes of focused time	0.666	0.712	0.721	0.725	0.775	0.646	0.685	0.707	0.766	0.82
- many times a day	(0.472)	(0.453)	(0.449)	(0.446)	(0.418)	(0.479)	(0.465)	(0.455)	(0.423)	(0.385)
Laughs with child - many times a day	0.782	0.83	0.839	0.844	0.872	0.777	0.811	0.854	0.854	0.92
	(0.413)	(0.375)	(0.368)	(0.363)	(0.334)	(0.417)	(0.392)	(0.353)	(0.354)	(0.272)
Does a special activity that the child enjoys	0.636	0.639	0.644	0.615	0.668	0.537	0.548	0.543	0.585	0.657
- Once or twice a day or more	(0.481)	(0.48)	(0.479)	(0.487)	(0.471)	(0.499)	(0.498)	(0.498)	(0.493)	(0.475)
Plays a sport, game, or hobby with child	0.733	0.758	0.753	0.735	0.794	0.65	0.677	0.673	0.746	0.808
- Once or twice a day or more	(0.443)	(0.429)	(0.432)	(0.441)	(0.404)	(0.477)	(0.468)	(0.469)	(0.435)	(0.394)
Reads to child - daily	0.596	0.667	0.719	0.716	0.746	0.564	0.635	0.667	0.688	0.704
	(0.491)	(0.471)	(0.45)	(0.451)	(0.435)	(0.496)	(0.482)	(0.472)	(0.464)	(0.457)

— Note: Each row corresponds to a variable of interest and contains the mean and standard deviation (in parentheses) specific to the quintile of the developmental score and family type as denoted in the column header.

Table 3B: Summary Statistics by Quintile of Peabody Picture Vocabulary Test for Children Age 4

	Two Parent Families					Single Parent Families				
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
Uptake Variables										
In Child Care	0.41	0.498	0.545	0.52	0.563	0.452	0.484	0.618	0.658	0.619
	(0.492)	(0.5)	(0.498)	(0.5)	(0.496)	(0.499)	(0.501)	(0.487)	(0.476)	(0.487)
Hours in Care	12.21	12.57	15.5	13.83	15.21	16.58	15.8	20	21.69	19.72
	(18.54)	(16.76)	(18.37)	(17.46)	(17.12)	(22.14)	(20.12)	(19.24)	(18.44)	(20.47)
Mother Works	0.597	0.649	0.697	0.705	0.712	0.413	0.464	0.552	0.606	0.577
	(0.491)	(0.477)	(0.46)	(0.456)	(0.453)	(0.493)	(0.5)	(0.498)	(0.49)	(0.495)
Mother Works 30 or more hours	0.401	0.408	0.435	0.439	0.483	0.29	0.316	0.418	0.49	0.415
	(0.49)	(0.492)	(0.496)	(0.496)	(0.5)	(0.455)	(0.466)	(0.494)	(0.501)	(0.494)
Parenting Scales										
Family Dysfunction Index	8.657	7.81	7.37	6.733	6.821	10.45	9.779	8.811	9.007	8.375
	(5.174)	(5.211)	(5.108)	(4.963)	(5.047)	(5.394)	(5.298)	(5.046)	(5.14)	(4.815)
Aversive Parenting	6.113	6.007	6.095	5.878	5.807	7.155	6.582	6.112	5.903	5.398
	(3.094)	(2.93)	(3.049)	(3.014)	(3.036)	(3.417)	(3.676)	(2.932)	(3.142)	(3.169)
Ineffective Parenting	8.812	8.745	8.851	8.724	8.704	10.08	9.147	8.984	8.461	8.185
	(3.758)	(3.685)	(3.402)	(3.423)	(3.379)	(4.132)	(4.108)	(3.547)	(3.43)	(3.524)
Parent Consistency	14.33	15.35	15.47	15.59	15.62	13.5	13.69	14.92	15.17	15.23
	(3.213)	(2.991)	(3.088)	(3.118)	(3.071)	(3.845)	(3.715)	(3.564)	(3.125)	(3.134)
Positive Interaction	14.99	15.11	15.33	15.42	15.48	14.45	15.43	14.85	15.48	15.18
	(2.82)	(2.391)	(2.37)	(2.39)	(2.284)	(3.12)	(2.483)	(2.837)	(2.492)	(2.544)
Parent-Child Activities										
Spends 5 minutes of focused time - many times a day	0.504	0.476	0.521	0.508	0.525	0.467	0.523	0.491	0.521	0.544
	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.5)	(0.501)	(0.501)	(0.499)
Laughs with child - many times a day	0.716	0.704	0.736	0.713	0.733	0.637	0.788	0.759	0.793	0.765
	(0.451)	(0.457)	(0.441)	(0.452)	(0.443)	(0.482)	(0.41)	(0.429)	(0.406)	(0.425)
Does a special activity that the child enjoys	0.365	0.331	0.348	0.362	0.351	0.24	0.384	0.338	0.374	0.269
- Once or twice a day or more	(0.482)	(0.471)	(0.477)	(0.481)	(0.478)	(0.428)	(0.487)	(0.474)	(0.485)	(0.444)
Reads to child - daily	0.639	0.69	0.702	0.721	0.731	0.595	0.652	0.644	0.694	0.767
	(0.481)	(0.463)	(0.458)	(0.449)	(0.444)	(0.492)	(0.477)	(0.48)	(0.462)	(0.424)

— Note: Each row corresponds to a variable of interest and contains the mean and standard deviation (in parentheses) specific to the quintile of the developmental score and family type as denoted in the column header.

Table 4A: Effect of Access to Universal Childcare by Quintile of the Motor and Social Development Score for Children Aged 0-3

	Two Parent Families					Single Parent Families				
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
Uptake Variables										
Mother Works	0.134	0.101	0.176	0.142	0.087	0.284	0.001	0.06	-0.002	0.114
	(0.001)***	(0.013)**	(0.000)***	(0.001)***	(0.011)**	(0.000)***	(0.99)	(0.62)	(0.984)	(0.324)
Mother Works 30 or more hours	0.086	0.109	0.17	0.156	0.056	0.144	-0.083	0.226	-0.082	0.123
	(0.044)**	(0.064)*	(0.000)***	(0.003)***	(0.116)	(0.019)**	(0.436)	(0.002)***	(0.366)	(0.303)
In Child Care	0.175	0.264	0.214	0.21	0.15	0.445	0.382	0.069	0.251	0.148
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.619)	(0.1)*	(0.216)
Hours in Care	8.666	10.842	9.498	9.569	6.158	17.517	14.815	4.297	10.179	6.909
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.222)	(0.085)*	(0.112)
Parenting Scales Activities										
Family Dysfunction Index	0.045	-0.054	1.179	0.523	0.056	0.258	-0.895	-1.337	-0.047	0.856
	(0.93)	(0.886)	(0.014)**	(0.497)	(0.89)	(0.745)	(0.153)	(0.136)	(0.972)	(0.464)
Aversive Parenting	0.359	0.063	0.313	0.543	0.483	0.484	0.885	-0.387	0.657	0.196
	(0.396)	(0.884)	(0.306)	(0.123)	(0.041)**	(0.352)	(0.157)	(0.603)	(0.049)**	(0.81)
Ineffective Parenting	-0.107	0.44	1.124	0.39	1.154	0.637	-0.206	0.298	3.782	0.912
	(0.867)	(0.149)	(0.000)***	(0.35)	(0.000)***	(0.405)	(0.816)	(0.664)	(0.000)***	(0.622)
Parent Consistency	-0.006	0.112	0.342	-0.192	-0.63	2.101	-0.877	0.572	-1.096	-3.721
	(0.989)	(0.825)	(0.324)	(0.637)	(0.029)**	(0.017)**	(0.127)	(0.528)	(0.422)	(0.002)***
Positive Interaction	0.061	-1.294	-0.405	-0.288	-1.01	-1.169	-2.094	-1.049	-0.941	0.2
	(0.861)	(0.002)***	(0.223)	(0.152)	(0.000)***	(0.202)	(0.003)***	(0.006)***	(0.039)**	(0.697)
Parent-Child Activities										
Spends 5 minutes of focused time	-0.036	-0.118	-0.065	0.023	-0.132	-0.185	-0.106	-0.002	0.028	-0.047
- many times a day	(0.353)	(0.004)***	(0.223)	(0.462)	(0.000)***	(0.014)**	(0.415)	(0.988)	(0.635)	(0.543)
Laughs with child - many times a day	-0.055	-0.06	-0.028	-0.018	-0.096	-0.229	-0.069	-0.042	-0.085	-0.009
	(0.066)*	(0.039)**	(0.276)	(0.38)	(0.000)***	(0.005)***	(0.528)	(0.558)	(0.195)	(0.832)
Does a special activity that the child enjoys	0.005	-0.055	-0.007	-0.05	-0.097	-0.112	-0.448	-0.14	-0.186	0.116
- Once or twice a day or more	(0.875)	(0.063)*	(0.894)	(0.244)	(0.002)***	(0.046)**	(0.001)***	(0.006)***	(0.003)***	(0.261)
Plays a sport, game, or hobby with child	0.125	0.021	-0.014	0.056	-0.007	-0.075	-0.083	-0.174	-0.052	0.243
- Once or twice a day or more	(0.002)***	(0.42)	(0.621)	(0.034)**	(0.83)	(0.315)	(0.381)	(0.143)	(0.644)	(0.044)**
Reads to child - daily	-0.178	-0.093	-0.047	-0.089	-0.032	-0.312	0.042	0.249	0.019	-0.312
	(0.000)***	(0.084)*	(0.109)	(0.007)***	(0.5)	(0.023)**	(0.614)	(0.062)*	(0.884)	(0.038)**

— Note: Each row corresponds to a variable of interest and contains the linear difference in differences estimate of mean and p-values (in parentheses) specific to the quintile of the motor-social development score and family type as denoted in the column header. P-values are adjusted for multiple testing and are corrected at the province-year level. These estimates are of the β coeffcient as specified in $Y_{ipt} = \beta_0 + \beta_1 Policy_{ipt} + \beta_2' PROV_p + \beta_3' YEAR_t + \epsilon_{ipt}$ and illustrates if Quebec experienced a disproportionate change in that activity. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

Table 4B: Effect of Access to Universal Childcare by Quintile of the Peabody Picture and Vocabulary Test score for Children Aged 4

	Two Parent Families					Single Parent Families				
	Q1	Q2	Q3	Q4	Q5	Q1	Q2	Q3	Q4	Q5
Uptake Variables										
Mother Works	-0.063	0.235	0.153	0.236	-0.035	0.005	-0.072	-0.182	-0.128	0.191
	(0.668)	(0.011)**	(0.033)**	(0.000)***	(0.619)	(0.98)	(0.773)	(0.479)	(0.411)	(0.418)
Mother Works 30 or more hours	0.016	-0.027	0.083	0.141	0.045	-0.162	-0.2	-0.224	-0.036	0.2
	(0.9)	(0.763)	(0.245)	(0.128)	(0.594)	(0.406)	(0.351)	(0.333)	(0.886)	(0.188)
In Child Care	0.162	0.385	0.117	0.183	0.091	0.347	0.285	-0.082	0.228	0.633
	(0.062)*	(0.000)***	(0.129)	(0.061)*	(0.222)	(0.011)**	(0.022)**	(0.665)	(0.447)	(0.000)***
Hours in Care	10.192	8.788	13.299	9.241	4.885	23.215	5.174	-0.586	8.009	27.222
	(0.000)***	(0.013)**	(0.000)***	(0.000)***	(0.051)*	(0.001)***	(0.367)	(0.918)	(0.427)	(0.000)***
Parenting Scales Activities										
Family Dysfunction Index	-0.843	-0.607	1.629	-0.557	-0.591	2.717	1.275	-2.296	-0.942	4.235
	(0.378)	(0.368)	(0.144)	(0.574)	(0.497)	(0.066)*	(0.486)	(0.228)	(0.563)	(0.015)**
Aversive Parenting	0.493	0.245	0.53	-0.231	0.141	1.91	0.977	0.341	0.716	2.357
	(0.327)	(0.17)	(0.097)*	(0.605)	(0.579)	(0.094)*	(0.2)	(0.575)	(0.225)	(0.007)***
Ineffective Parenting	0.257	1.279	0.736	1.437	0.327	-0.606	1.698	1.255	-1.225	3.864
	(0.755)	(0.075)*	(0.012)**	(0.142)	(0.496)	(0.649)	(0.315)	(0.222)	(0.379)	(0.005)***
Parent Consistency	-0.757	-0.039	-1.206	-1.122	0.305	0.415	0.782	-1.845	-0.101	-1.433
	(0.306)	(0.955)	(0.004)***	(0.023)**	(0.362)	(0.628)	(0.544)	(0.061)*	(0.908)	(0.183)
Positive Interaction	-0.197	-0.602	-0.583	0.001	0.056	-2.472	-0.528	-2.011	0.166	-2.535
	(0.745)	(0.297)	(0.002)***	(0.998)	(0.85)	(0.179)	(0.458)	(0.019)**	(0.864)	(0.003)***
Parent-Child Activities										
Spends 5 minutes of focused time	0.071	0.007	-0.039	0.162	0.031	-0.411	-0.033	0.286	0.041	-0.169
- many times a day	(0.382)	(0.922)	(0.627)	(0.126)	(0.541)	(0.101)	(0.8)	(0.062)*	(0.775)	(0.334)
Laughs with child - many times a day	-0.025	-0.178	-0.016	-0.073	-0.022	-0.352	-0.2	-0.166	0.108	-0.651
	(0.734)	(0.000)***	(0.799)	(0.484)	(0.748)	(0.18)	(0.04)**	(0.37)	(0.597)	(0.000)***
Does a special activity that the child enjoys	-0.222	-0.026	-0.078	-0.083	-0.06	-0.133	-0.083	-0.139	-0.05	0.076
- Once or twice a day or more	(0.002)***	(0.771)	(0.201)	(0.265)	(0.511)	(0.317)	(0.533)	(0.039)**	(0.829)	(0.596)
Reads to child - daily	-0.044	-0.019	0.073	0.19	0.276	0.411	0.43	0.455	0.157	0.016
	(0.646)	(0.715)	(0.234)	(0.02)**	(0.007)***	(0.001)***	(0.061)*	(0.012)**	(0.418)	(0.923)

— Note: Each row corresponds to a variable of interest and contains the linear difference in differences estimate of mean and p-values (in parentheses) specific to the quintile of the Peabody and Picture Vocabulary Score score and family type as denoted in the column header. P-values are adjusted for multiple testing and are corrected at the province-year level. These estimates are of the β coeffcient as specified in $Y_{ipt} = \beta_0 + \beta_1 Policy_{ipt} + \beta_2' PROV_p + \beta_3' YEAR_t + \epsilon_{ipt}$ and illustrates if Quebec experienced a disproportionate change in that activity. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

Table 5: IV Estimates of the Causal Effect of Attending Childcare

	Two Parent Families		Single Parent Families		Two Parent Families - Girls		Two Parent Families - Boys	
	Age 0-3	Age 4	Age 0-3	Age 4	Age 0-3	Age 4	Age 0-3	Age 4
Parenting Scales Activities								
Family Dysfunction Index	2.281	-1.707	0.322	1.825	2.117	-3.801	2.407	2.054
	(0.01)***	(0.508)	(0.77)	(0.844)	(0.177)	(0.195)	(0.004)***	(0.204)
Aversive Parenting	1.665	0.725	1.397	3.181	2.562	0.106	0.600	1.987
	(0.000)***	(0.041)**	(0.09)*	(0.08)*	(0.000)***	(0.842)	(0.188)	(0.024)**
Ineffective Parenting	2.814	3.572	4.689	1.555	3.304	2.125	2.146	6.598
	(0.000)***	(0.000)***	(0.005)***	(0.71)	(0.000)***	(0.028)**	(0.082)*	(0.001)***
Parent Consistency	-0.396	-0.898	-2.283	0.362	-1.216	-0.698	0.679	-1.286
	(0.643)	(0.234)	(0.224)	(0.924)	(0.122)	(0.423)	(0.503)	(0.261)
Positive Interaction	-2.858	-1.354	-2.419	-3.826	-3.558	-1.034	-2.147	-2.216
	(0.000)***	(0.061)*	(0.005)***	(0.02)**	(0.000)***	(0.138)	(0.017)**	(0.024)**
Parent-Child Activities								
Spends 5 minutes of focused time	-0.361	0.04	-0.032	-0.089	-0.644	-0.043	-0.304	-1.009
- many times a day	(0.002)***	(0.922)	(0.77)	(0.846)	(0.000)***	(0.81)	(0.018)**	(0.006)***
Laughs with child - many times a day	-0.284	-0.425	-0.311	-0.536	-0.432	-0.228	-0.178	-0.984
	(0.000)***	(0.000)***	(0.1)*	(0.035)**	(0.000)***	(0.042)**	(0.001)***	(0.015)**
Does a special activity that the child	-0.253	-0.421	-0.393	-0.237	-0.356	-0.264	-0.259	-0.969
enjoys - Once or twice a day or more	(0.000)***	(0.003)***	(0.14)	(0.298)	(0.000)***	(0.005)***	(0.000)***	(0.006)***
Plays a sport, game, or hobby with child	0.132	0.13	0.218	-0.367	0.027	0.02	0.079	0.226
- Once or twice a day or more	(0.041)**	(0.317)	(0.231)	(0.027)**	(0.783)	(0.879)	(0.054)*	(0.512)
Reads to child - daily	-0.537	0.621	-0.442	0.852	-1.203	0.072	-0.352	-0.32
	(0.003)***	(0.004)***	(0.14)	(0.08)*	(0.009)***	(0.795)	(0.003)***	(0.444)
F-Test	455.93***	277.9***	76.88***	66.96***	149.82***	212.54***	502.93***	251.48***

— Note: For the outcome variable in each row we present estimates of treatment effect of childcare, the coefficient on Ccare using an instrumental variable estimator for samples denoted in the column heading that differ by family type, child age and child gender. The specification includes all of the covariates used to reweight the change in changes estimates presented in Figures 1 and 2. Standard errors are corrected at the province-year level and are presented in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively. The final row of the table presents the F-test for the joint significance of the instrument from the first stage regression.