

Targeted or universal coverage? Assessing heterogeneity in the effects of universal childcare

Michael J. Kottelenberg
Huron University College

Steven F. Lehrer*
Queen's University and NBER

November 2015

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 two anonymous reviewers, 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 and previously circulated under the title "Reinvestigating who benefits and who loses from universal child care in Canada". 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.¹

Baker (2011) surveys the small but growing literature of 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

¹For example, Heckman et al. (2010) point out that to be eligible to participate in the Perry program, individuals were required to have an IQ between 70 and 85, as measured by the Stanford-Binet IQ test (1960s norming). This sample differs sharply in multiple dimensions including IQ from the national sample evaluated in this paper.

(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 (2015), 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 worldwide are currently considering implementing a variety of universal early childhood education programs.²

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 (2006) 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.³ 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 boosts developmental test scores for children from single parent households particularly for those who are most disadvantaged and located

²Three examples of debates ongoing in September 2015 for either introducing or expanding universal child-care would include Canada, where nationwide universal subsidized childcare at \$15 a day has emerged as a major difference in policy platforms prior to the National election. In Great Britain, there are debates as to whether the Government should offer free childcare to all parents of two-year-olds and not just those from disadvantaged households, as well as increase the number of hours of the early years entitlement for three to four-year-old children. Last, in several states such as Minnesota and Maryland, debates of whether to introduce universal preschool to 4 years olds versus only targeting preschool to the children who are most at risk of underachievement are underway in state legislatures.

³Havnes and Mogstad (2015) 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 the policy on specific outcomes, and ii) the 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 Kolmogorov-Smirnov tests clearly demonstrate that the figures are statistically different when we control for covariates.

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 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, to 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.⁴

To shed light on a 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

⁴Despite 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. Angrist, Oreopoulos and Williams (2013), Ding and Lehrer (2011), 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.

subsidized child care for economically disadvantaged subpopulations, we believe that this set of results is suggestive that efforts to ensure that parents do not reduce investments made into home learning environments in response to formal child care attendance should be considered when designing child care policy.

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 suggestive 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 of strengthening governmental support of parents.⁵ 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).⁶ This program

⁵This 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.

⁶The 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

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.

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.⁷ 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.⁸

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 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.

⁷The 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.

⁸That 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. Dropping cycles of the NLSCY during the period where the Quebec Family Policy is being introduced is commonly done in the literature evaluating the Quebec Family policy and was undertaken in Baker et al (2008), Kottelenberg and Lehrer (2013), among others. Ding et al (2015) provide evidence on policy effects during the period in which the Quebec Family Policy was implemented.

completes 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.⁹ 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.¹⁰

Our developmental outcomes first include motor and social development (MSD) scores for children aged 0 to 3. This score is based on parental responses to a set of 15 child age-varying questions that measure dimensions of the motor, social and cognitive development of young children. Standard scores are subsequently calculated by the month of the child’s age, in which the average score for the population is set at 100 with a standard deviation of 15.¹¹ For children aged 4 to 5, a Peabody Picture Vocabulary Test (PPVT) was conducted during the home interview. The child was asked to look at pictures on an easel and identify the picture which matched the word the interviewer read out. Standard scores that take into account the child’s age were subsequently computed within two month age groups so that the mean of the standard scores was 100 and the standard deviation was 15 for all age groupings.¹²

Table 1 presents summary statistics on our explanatory variables separately for two parent and single parent families. 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 surprising that the latter 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

⁹In two parent families 90.7% of respondents are the mother. We also note that 95.0% of single parent families are headed by mothers.

¹⁰We 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.

¹¹The standardization process relies on data from the first cycle of the NLSCY. Thus, differences in these scores over time are reflected within this measure.

¹²Canadian norms were developed for these children in the 4 to 5 age group in a manner that ensures when global comparisons are made between children who completed the test in English vs. French, performance should be equivalent.

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. This in part motivates the importance of controlling for compositional changes in our analyses. 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.

Earlier research did not investigate policy effects on single parent households, in part 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.¹³ 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.¹⁴

¹³Prior research investigating the Quebec Family policy including Baker, Gruber and Milligan (2008) and Kottelenberg and Lehrer (2013) focused solely on estimating mean impacts with two parent families and did not consider distributional treatment effects. These studies used a linear difference in differences estimator allowing them to additionally consider outcomes that are discrete or interval measured and began by presenting evidence that assumptions underlying the identification strategy such as common trend and common support are either met or in the few situations where they are violated operate in the opposite direction towards the policy effect. We did replicate this analysis corresponding to single parent families and the general pattern of these results hold. For example, using linear regression with the pre-program data we do not find any evidence of statistically significant differential trends for either childcare use ($\beta = 0.028$, $p\text{-value} = 0.53$) or maternal labor supply ($\beta = 0.071$, $p\text{-value} = 0.114$) from two-sided tests with clustered standard errors at the province-cycle level. The results of the full set of tests that were carried out in Kottelenberg and Lehrer (2013) for the samples of both two parent and single parent families used in this paper are available from the authors upon request.

¹⁴As 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.

3 Empirical Strategy

Since there have been few empirical applications of the Athey and Imbens (2006) change in changes estimator (CIC), we begin by providing a brief discussion of how this expands on knowledge from 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 comparing the realization of the mean outcome in the presence of treatment with a counterfactual mean.¹⁵ This counterfactual mean is constructed by adding the observed change over time in the mean outcome for the comparison group to the treatment group's mean outcomes prior to treatment, thereby providing an estimate of the expected change in the absence of treatment (potentially conditioning on a set of covariates X).

The CIC estimator allows for the estimation of the entire potential outcomes distribution, thereby proving a richer perspective by identifying the quantiles of the distribution where the treatment is more effective. As ingredients to the estimation procedure the cumulative distribution functions of the outcome variable in the four groups are defined on the basis of g and t are utilized.¹⁶ We define $F_{Y_{gt}}$ to be the empirical cumulative distribution functions for each group and time period. The CIC estimator mimics the general strategy of the DID estimator but at each quantile. It uses the observed outcomes from the treated group prior to treatment

¹⁵Identification of this causal parameter relies on assumptions of i) common support, ii) common trend, and iii) no anticipation effects. Without covariates, this causal parameter can be obtained by subtracting the difference of the mean outcomes of treated and controls after the treatment from the outcome difference in these groups that existed prior to the treatment.

¹⁶Recall the DID estimator only requires the means for these same groupings.

and the observed changes occurring over time in the non-treated group to construct points on a counterfactual cumulative distribution function, $F_{Y_{11}^{cf}}$. This counterfactual distribution represents the outcomes of all treated individuals as if they did not have access to the treatment.

Formally, $F_{Y_{11}^{cf}}$, is non-parametrically estimated using the empirical cumulative distribution functions, $F_{Y_{gt}}$. Quantile-outcome pairs are identified by

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

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 . The quantile specific treatment effect is given by $\Delta y(\tau) = F_{Y_{11}}^{-1}(\tau) - F_{Y_{11}^{cf}}^{-1}(\tau)$, which is the increase (decrease) in score at quantile τ in the unconditional distribution and where $F_{Y_{11}^{cf}}$ is calculated using equation (1). The identifying assumptions, discussed further below, establish that observed changes in the outcome at specific quantiles in the comparison group's cumulative distribution function can be directly related to the specific quantiles in the treated group. Thus, the CIC estimator constructs a counterfactual distribution by imputing at each quantile the expected movement in the outcome in the absence of treatment at that point of the cumulative distribution function.

To understand the identifying assumptions, we start by noting that unlike the DID estimator, the change in changes estimator removes the functional form dependence in how D, t and g influence Y .¹⁷ The outcome within each time period is only assumed to be generated by a strictly monotonic function of the unobserved inputs. As such, a realization of Y has a one-to-one correspondence to a given realization of unobserved factors within each time pe-

¹⁷We 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. First, the QDID requires that the distribution of unobservables is identical in all subpopulations, eliminating potential sources of intrinsic heterogeneity. Second, the QDID model must place additional restrictions on the data for the transformation required to be monotone.

riod, regardless of the group/province. The CIC estimator also requires the assumption that the distribution of the unobserved factors is time invariant within each group.¹⁸ However, the distribution of unobserved factors is not required to be identical across groups. For example, it can accommodate situations where the treatment group may contain more high ability children than the control group.¹⁹

The above assumptions are less restrictive than the common trend assumption underlying the linear DID estimator, where group and time implicitly impose constant effects. With the CIC estimator, group and time effects are allowed to differ across individuals with different observed (and unobserved) characteristics. Further, 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.

The steps involved in calculating the quantile treatment using the CIC estimator are illustrated in Figure 1. These steps highlight the role of each of the maintained assumptions in imputing the outcome of a treated unit in the absence of treatment. The first panel corresponds to the computation of $F_{Y_{00}}(F_{Y_{10}}^{-1}(\tau))$ in equation (1), where $F_{Y_{10}}^{-1}(\tau)$ identifies the outcome in the treatment group prior to treatment is at the quantile of interest τ . Thus, $F_{Y_{00}}(F_{Y_{10}}^{-1}(\tau))$ identifies the quantile in the pre-treatment comparison group for which the outcome level is the same. Finding comparable quantiles in the pre-treatment comparison and treatment groups can be made since a specific realization of unobserved factors is assumed to have a one-to-one correspondence to a unique y , regardless of group in the same time period.

The second panel shows the calculation of the time effect for the specific level of unobserved

¹⁸This 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.

¹⁹Selection 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.

factors at the quantile of interest τ identified in the first panel. The time effect is computed by comparing the quantile functions of $F_{Y_{01}}$ and $F_{Y_{00}}$. This comparison across time can be made since rank invariance follows from the assumptions of monotonicity and within group time-invariance of the unobserved factor, ensuring one’s rank within a group is associated with an unobserved factor that does not change over time.

In the third panel, points on the counterfactual distribution $F_{Y_{11}^{cf}}$ are calculated by returning to the initial value of y in the pre-treatment distribution of the treated and adding the time effect calculated in the second panel. This comparison across time can be made since the monotonic mapping between unobserved factors and outcomes is assumed to be the same in both groups. This step is expressed via the $F_{Y_{01}}^{-1}(\cdot)$ transformation in equation (1) and computes y_{τ}^{cf} , the post-treatment outcome associated with the initial quantile of interest τ in the absence of treatment.²⁰ The final panel illustrates the calculation of the quantile specific treatment effect, $\Delta y(\tau)$, by taking the difference between $F_{Y_{11}}^{-1}(\tau)$ and $F_{Y_{11}^{cf}}^{-1}(\tau)$.

To further compare the quantile treatment effects between single parent and two-parent families additionally requires us to make corrections for the fact that the distribution of developmental scores between children from these different households are unequal even in the absence of subsidized childcare. After all, evidence in Table 1 indicates that on average child outcomes are generally lower in single parent families. Bitler, Hoynes and Domina (2014) suggest placing the quantile treatment effects on the same scale. In our case, we use the observed quantiles of the respective unconditional distribution of developmental scores for children from Quebec in the pre-policy cycles as the reference distribution to establish anchoring points. By

²⁰The second transformation $F_{Y_{01}}^{-1}$ 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.

placing the developmental scores on identical scales, the interpretation of the quantile treatment effects is facilitated. For example, after making this correction a child at the 17th percentile of the PPVT score from a single-parent household would have the same relative position as a child at the 17th percentile in a two-parent household.²¹

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. Under the null, these distributions come from the same underlying density and the D statistics reported from this test provide the absolute value of the largest difference between the two empirical distributions. Hence a statistically significant D-value indicates we can reject the null hypothesis that the availability of subsidized childcare did not alter the distribution of developmental outcomes.

3.1 Controlling for Covariates

Lechner (2011) postulates that there have been few applications of the CIC estimator since there lacks a tractable estimator in the presence of covariates. While Athey and Imbens (2006) do propose two strategies by which researchers could include covariates to account for compositional changes, however the arguably more empiricist friendly strategy requires highly restrictive assumptions of additive separability and statistical independence between the observed covariates and unobservables, conditional on access to the policy.²² Specifically, these assumptions rule out the cases where the treatment effect at any quantile is correlated with an observed covariate. As such, to evaluate child care policy while accounting for the presence of compositional changes, we consider reweighting since it does not require this stringent assumption of additive separability.

²¹We also report uncorrected causal estimates from the CIC to shed light on how within the single parent and two-parent families does subsidized childcare influence developmental outcomes.

²²The other strategy is fully nonparametric and would be difficult to implement if the dimensionality of observed covariates is high, as it is in our application.

An 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.²³ 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).²⁴ 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.²⁵

²³To 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.

²⁴The main empirical strategy in Havnes and Mogstad (2015) 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 (2015) 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.

²⁵In other words, we do not require an assumption of additive separability. All of the results presented in

Such a technique removes differences in the unconditional distribution of scores that may arise from differences in observables across cohorts and is accomplished by weighting the data such that the observable covariates from the different sub-populations are balanced.²⁶ Using the estimated inverse propensity weights, we are able to produce distributions of the test score for the treatment and comparison in the pre- and post-policy periods given a set of equivalent observables between all four groups. As such, concerns that unaccounted for observables are driving the observed treatment effects are reduced and potentially eliminated.²⁷

Yet even when accounting for covariates the interpretation of the estimated intention to treat effects might be complicated by changes in child care usage patterns in Quebec. The policy not only altered formal child care arrangements, shifting care away from home-based care towards more center-based care (Tougas, 2002), but these changes also lead to a more diverse set of individuals attending child care. This diversity emerged from two potential sources, which may in part account for some of the treatment effect heterogeneity witnessed. First, 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. Second, a challenge in the implementation of this policy was that since there were more eligible children than subsidized spaces, waiting lists at each daycare centre were created and where individual children were positioned on these lists is largely left at the discretion of the provider.²⁸ Therefore, it is possible that the providers

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.

²⁶This method of reweighting the distribution functions is identical to the first three steps described in Appendix B of Bitler et 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.

²⁷With a linear difference in differences estimator, one would simply add controls to the estimating equation to ensure balance.

²⁸This suggests that there was two-sided selection to attend subsidized child care adding to the importance of balancing covariates. Many subsidized private daycare centres engaged in extra billing for supplementary

held time-varying beliefs about the sort of children they would like to choose from the queue, given the observables available. Together these results suggest that child care usage following the policy was different prior to the policy.

Finally, to conduct statistical inference at each percentile we use Fisher’s exact permutation test since it has good small sample properties. Specifically at each percentile we first conduct a test to measure the significance of the deviation of $\Delta y(\tau)$ estimate from a null hypothesis that the effect is 0.²⁹ In addition, we conduct a second test to measure the significance of the deviation of our estimate of $\Delta y(\tau)$ from the mean policy estimate from the change in changes estimator.³⁰ In effect this second test aims to determine if there is treatment effect heterogeneity.³¹ While unreported, the mean policy estimates from the change in changes estimator are quite similar to the mean intent to treat effects using a linear DID estimator reported in both Baker et al. (2008) and Kottelenberg and Lehrer (2013).

activities prior to 2005.

²⁹The 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 either the bootstrap or the asymptotic formulas developed in Athey and Imbens (2006) and we thank Guido Imbens for suggesting that it is more appropriate in this setting.

³⁰A 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.

³¹When conducting this test we are assuming that the average effect is known so that the test can yield an exact p-value. We considered following a testing procedure developed in Ding et al. (2015) who first point out that when the average effect is also calculated the Null hypothesis is no longer sharp. However, in our application the average effect from CIC is calculated from the set of the distributional CIC effects and its construction assumes that it is not a constant effect; whereas the test proposed in Ding et al (2015) test requires the average effect to be constant and estimated independently from the distributional effects. Thus, future research is needed to develop inference procedures for this setting and while we experimented with the computationally intensive Ding et al. (2015) test, it unsurprisingly led to fewer significant differences since it reports a maximal p-value when comparing the distributional effect to possible values for the true constant effect.

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 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 significantly differ.³²

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 or identify at which percentiles the $\Delta y(\tau)$ is statistically significant. Figure 2 presents change in changes estimates of the effects of the policy on MSD scores that measures the distributional policy parameters on the same scale for single-parent and two-parent families. 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 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

³²The 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 this caveat 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.

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 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.³³ 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. We see the pattern of $\Delta y(\tau)$ is robust to controlling for covariates.

Kottelenberg and Lehrer (2014) show that the negative impacts from a linear DID estimate reported in Baker et al. (2008) differ 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

³³With 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 $F_{Y_{11}^{ef}}$ are statistically different.

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 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 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.

Not surprisingly and as indicated in the notes listed under Figure 4, Kolmogorov-Smirnov tests of the equality of the cumulative distribution function of child development scores reject the null hypothesis for both developmental scores for girls but only for PPVT scores and when we control for covariates with boys. This indicates that there is no evidence that the policy changed the distribution of MSD scores for boys, which is why we observed very few dots in that panel of Figure 4.

In Appendix Figures A1 and A2, we present the corresponding distributional policy effect estimates that do not place the quantile treatment effects on the same scale for each family structure. These figures demonstrate that the general patterns presented in Figures 2 and 3 hold within family structure.

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 relaxes the common trend assumption underlying the linear DID estimator and 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, the median outcome in the treated has its counterfactual as the median in the untreated distribution. Thus, 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 it is possible that quantile differences of potential outcomes are zero at any quantiles, and this is due to individuals moving up and down in the distribution not that the true treatment effects are not zero. 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.³⁴ If the characteristics vary significantly, this provides 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 . We follow Abadie (2002) in using a bootstrap procedure to compute the p-values of the test statistic since there is substantial heaping in test scores at integer values which would be inconsistent with the continuity assump-

³⁴Recall, the change in changes estimator assumes that any changes in the outcomes within the control provinces are due to time effects and not from changes in the distribution of unobserved factors. Unlike the linear difference in differences estimator, the time effects may vary across individuals with different observed characteristics. Thus, this test evaluates whether we are comparing similar individuals based on variables available to the econometrician at different points in the outcome distributions.

tion underlying asymptotic tests. This test relies on a relatively weak regularity conditions to imply consistency. 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.³⁵ 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 an avenue that may explain 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.³⁶ Child developmental outcomes could be influenced through 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 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

³⁵The results do not differ sharply for either the subsamples of single parent families or by child gender.

³⁶Without 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.

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 five different parenting scales: 1) a family dysfunction score; 2) a punitive aversive score; 3) a hostile/ineffective score; 4) an inconsistency score; 5) and a positive interactions score.³⁷ 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. Evidence of differences in 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.³⁸ In these tables, the first column for each family structure presents summary statistics of various measures of parent-child activities from households in the first quintile of the respective developmental score distribution. The remaining columns provide information on how, on average these variables

³⁷For 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.”

³⁸We 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.

differ between the remaining quintiles of the distribution relative to the lowest quintile. Results from two-sided tests of these differences in the means presented in parentheses in Tables 3A and Table 3B indicate there are many significant differences across the quintiles presented for both single and two parent families. A notable exception is that among two parent families, differences in uptake variables across quintiles are relatively small. Not surprisingly, parenting scales are generally lower among single parent families and increase across the developmental score distribution.

There is substantial heterogeneity across the test score distribution in the parent child activity measures. The differences in levels of these activities do not solely emerge within family structures, but when comparing the lowest quintiles between family structures we observe that single parents are much less likely to engage in activities with their child relative to two-parent families. These differences between family types are much less marked at higher quintiles. These summary statistics and tests of differences across quintiles, 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.³⁹

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 quintiles of the developmental score distribution, we summarize the changes by presenting nonparametric DID estimates for the first quintile of the MSD and PPVT distribution in the first column of Table 4A and 4B. The remaining columns present results of tests of differences between the DID estimates in the remaining quintiles relative to the lowest quintile.⁴⁰ These estimates indicate

³⁹Even if child care uptake affects the investment of parent equally across the distribution, we would expect that children receiving fewer inputs to begin with would suffer disproportionately from these changes if there are diminishing returns to investment in children. As is explored it may also be the case that parents across the distribution disproportionately change their investment in their children as well.

⁴⁰The 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

that among single parent families, the gain in maternal labour supply on the extensive margin came from children who scored in the bottom quintile. These families increasingly utilized child care and the children experienced large gains in the hours spent in child care.⁴¹ Yet, these children are also the only quintile to exhibit statistically significant declines in nearly every parent child activity. Further, 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 children 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 observe declines in parent-child activities, though the pattern of these changes is somewhat inconsistent. For children aged 4, there are negative consequences to parenting scales, primarily for children who score in the middle and top quintiles. However, it is difficult to distinguish these effects from changes in the other quintiles. Exploring the individual parenting activities, there are both negative and positive impacts to parenting behavior occurring throughout the distribution. For example, we observe parents of children in the top most quintile reduce their frequency of doing special activities with their children between age 0 and 3 relative to the other quintiles. We also observe modest increases in time spent playing sports for children in the lowest portion of the distribution. For children aged 4, there are significant differences across the quantiles in parent-child activities.

Though it is difficult to observe a clear pattern across the quantiles, we do note that the largest declines in reading to a child daily, a variable we would closely associated with parental investment in cognitive ability, occur for children aged 3 in the bottom quintile. Further, in examining children aged 4, we again find that compared to children in the bottom portion of the

⁴¹The differences along the intensive margins are not as stark. There are fewer differences in the effect on the hours in care between the quintiles.

distribution, those in the top quintile are significantly more likely to be read to daily. In both of these cases, the observations appear consistent with the observed changes in the distribution of the cognitive tests scores for two parent families.

Altogether, these sets 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.⁴²

Havnes and Mogstad (2015) 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 little 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 possible that in part 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.

Thus far, we have provided only suggestive evidence of changes in parenting practices being a mechanism that can explain the heterogeneity in developmental outcomes due to access to subsidized childcare. While we are unable to provide clear causal evidence of this mechanism

⁴²For 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.

across the distribution, it is possible to identify a local causal average impact of how parenting practices change on average for those parents who send their children to child care due to the policy, using 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:

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

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,⁴³ $PROV$ 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:

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

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

⁴³We 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.

and child gender. Odd columns refer to children aged 0-3 and even columns contain results for 4-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 specific parental activity. Further, not only does the impact of child care on ineffective parenting increase with child age, but the estimated effect is roughly 25% higher for 4-years olds. 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.⁴⁴ 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

⁴⁴The 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.

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 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.⁴⁵

As a whole, we interpret the results in Tables 3 to 5 as documenting 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 often 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.⁴⁶ For single-parent families, there are gains in

⁴⁵Kottelenberg 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.

⁴⁶That 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

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 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 only two specific developmental outcomes.⁴⁷ 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. We also find striking differences in the distributional effects across child gender that differ across developmental scores.

Our empirical analysis is suggestive that the heterogeneity in policy effects also 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 investments explain the heterogeneity in policy effects.

⁴⁷Future work should follow Bitler et al (2014) and also consider evaluating the distributional effects of universal childcare on measures of non-cognitive skills or socioemotional outcomes, which can potentially be constructed using behavioral questions. In addition, we did not consider the potential of interactions between the policy being evaluated with concurrent policy reforms that may have varied across provinces and time; which is commonly done in research that evaluates a natural experiment. Such an approach could be useful in understanding whether the differences in policy effects across household structures is due to other policies.

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 were alternatively 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 may 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.⁴⁸ In addition, Baker et al. (2008) and Kottelen-

⁴⁸It is worth noting that a subset of these findings might appear inconsistent with both Kline and Walters (2015) and Feller et al. (2014) who independently evaluated the Head Start Impact study, where they additionally account for there being different alternatives available to parents including other preschool programs

berg and Lehrer (2013) 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 substantially different patterns for each developmental score by child gender. Observing that the treatment effect heterogeneity is not fully characterized by predetermined variables reinforces the need for an improved understanding of why the policy works, when it does, and why it fails when it does not.

and home care. These studies extend the prior literature evaluating Head Start by pointing out that parents face multiple competing choices in the randomized trial and then respectively develop empirical strategies that can identify causal parameters in this setting. In particular, these studies find large positive short run effects on the test scores of children who would have otherwise been cared for at home, and small effects for children who would otherwise attend other preschools. However, Kline and Walters (2015) interpret the selection in their study as leading to enrollment by the families most desperate to get help with child care, which is also consistent with the large negative IV estimates reported in Kottelenberg and Lehrer (2013) that suggest those children who were induced to child care due to the policy benefitted the least from childcare attendance. We hypothesize that any differences may result from the distinction in activities performed in a child care as opposed to preschool setting and that future work should more closely examine how effects differ due to parental information concerning either match quality in the quality of the practices employed in individual childcare centers and preschools operate.

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(2), 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 family well-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, C. R., M. Nores, S. Barnett, and L. Schweinhart (2006), “The High/Scope Perry Preschool Program: Cost-Benefit Analysis Using Data from the Age-40 Followup,” *Journal of Human Resources*, 40(1), 162–190.
- [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., Domina, T., and Hoynes, H. (2014), “Experimental Evidence on Distributional Effects of Head Start,” *NBER Working Paper no. 20434*.
- [10] Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006), “What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments,” *American Economic Review*, 96(4), 988-1012.
- [11] Bitler, M. P., J. B. Gelbach, H. W. Hoynes (2005), “Distributional impacts of the self-sufficiency project,” *NBER Working paper #11626*.
- [12] Cascio, E., and Schanzenbach, D. (2013), “The impacts of expanding access to high-quality preschool education,” *Brookings Papers on Economic Activity*, Fall, 127-178.

- [13] 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.
- [14] Ding, W., M. Kottelenberg and S. F. Lehrer (2015), “New evidence on the influence of universal child care on child care decisions and child development in Canada,” *mimeo*, Queen’s University.
- [15] Ding, P., A. Feller and L. W. Miratrix (2015), “Randomization inference for treatment effect variation,” forthcoming in the *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*.
- [16] 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.
- [17] Djebbari, H. and J. Smith (2008), “Heterogeneous impacts in PROGRESA”, *Journal of Econometrics*, 145(1-2), 64-80.
- [18] Feller, A., T. Grindal, L. Miratrix and L. Page (2014), “Compared to what? Variation in the impact of early childhood education by alternative care-type settings,” *mimeo*, University of California.
- [19] Firpo, S., N. Fortin, and T. Lemieux (2009), “Unconditional Quantile Regressions,” *Econometrica*, 77(3), 953–973.
- [20] Firpo, S. (2007), “Efficient semiparametric estimation of quantile treatment effects,” *Econometrica*, 75(1), 259–276.
- [21] Fitzpatrick, M. D., (2010), “Preschoolers enrolled and mothers at work? The effects of universal Pre-Kindergarten,” *Journal of Labor Economics*, 28(1), 51-85.
- [22] 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.
- [23] 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.
- [24] Havnes, T. and M. Mogstad (2015), “Is universal child care leveling the playing field?,” *Journal of Public Economics*, 127, 100–114.

- [25] 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.
- [26] Heckman, J., Malofeeva, L., Pinto, R. and Savelyev, P. (2013), “Understanding the mechanisms through which an influential early childhood program boosted adult outcomes,” *American Economic Review*, 103(6), 2052-2086.
- [27] 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.
- [28] Hirano, K., G. W. Imbens, and G. Ridder (2003), “Efficient estimation of average treatment effects using the estimated propensity score,” *Econometrica*, 71(4), 1161-1189.
- [29] 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.
- [30] Kline, P. and C. Walters (2015), “Evaluating Public Programs with Close Substitutes: The Case of Head Start,” *NBER Working Paper No. 21658*.
- [31] Kottelenberg, M. J., and S. F. Lehrer (2015), “The gender effects of universal child care in Canada: Much ado about boys?,” *mimeo*, Queen’s University.
- [32] 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.
- [33] Kottelenberg, M. J., and S. F. Lehrer (2013), “New evidence on the impacts of access to and attending universal child-care in Canada,” *Canadian Public Policy*, 39(2), 263–286.
- [34] Lechner, M. (2011), “The Estimation of Causal Effects by Difference-in-Difference Methods,” *Foundations and Trends(R) in Econometrics*, 4(3), 165-224.
- [35] 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*.
- [36] 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.
- [37] 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.

- [38] 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.
- [39] Newson, R., and ALSPAC Study Team (2003), “Multiple-test procedures and smile plots.” *Stata Journal*, 3, 109-132.
- [40] Powell, D. (2014), “Did the economic stimulus payments of 2008 reduce labor supply? Evidence from quantile panel data estimation,” *RAND Working Paper 720*.
- [41] 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.
- [42] 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.
- [43] Simes, R.J. (1986) “An improved Bonferroni procedure for multiple tests of significance.” *Biometrika*, 73(3), 751-754.
- [44] 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 of Parenting Behaviours

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

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

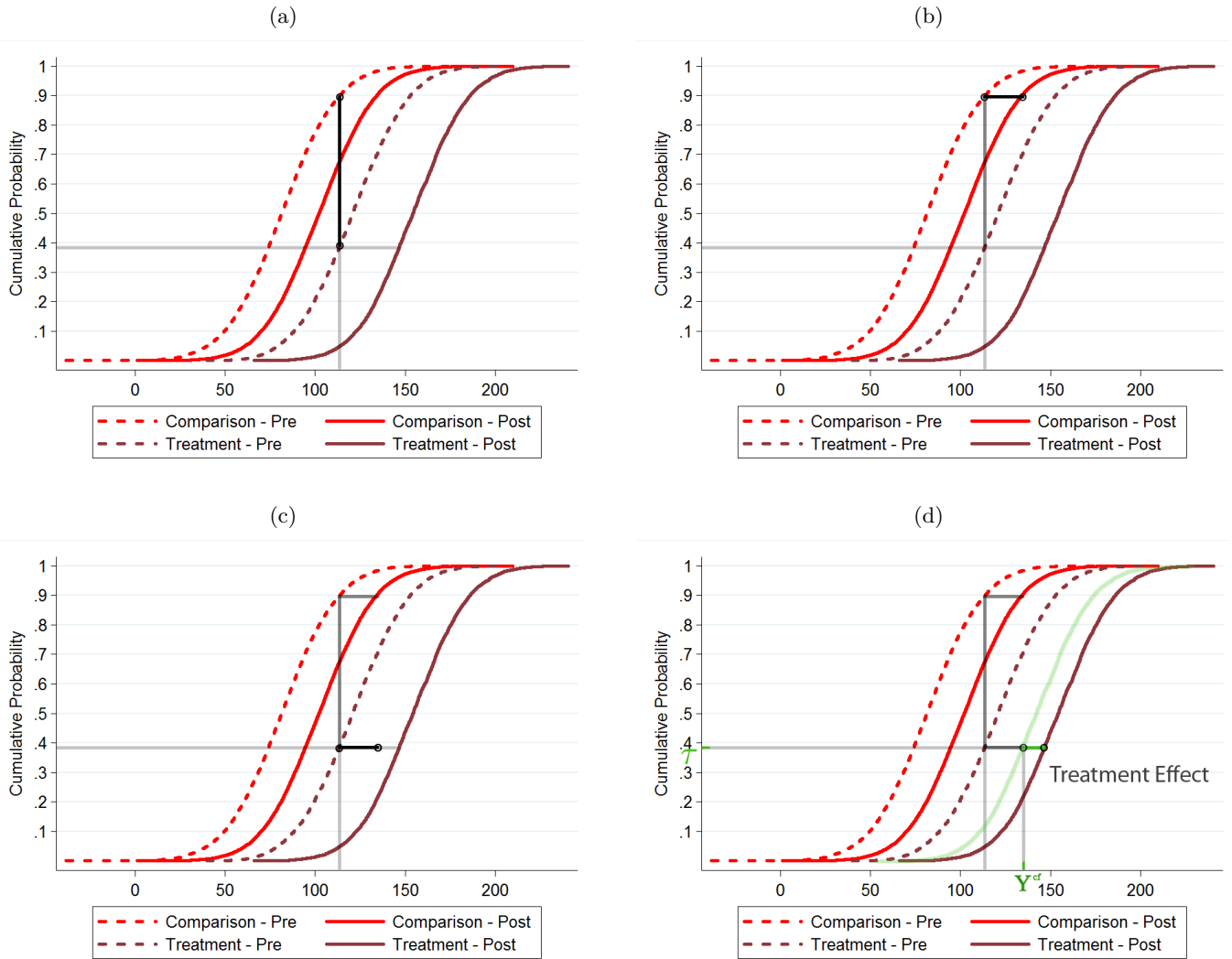


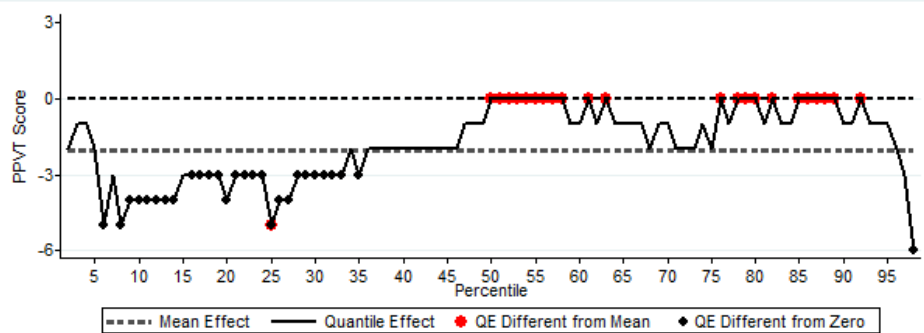
Figure 2: Change-in-Change Estimates for Motor and Social Development



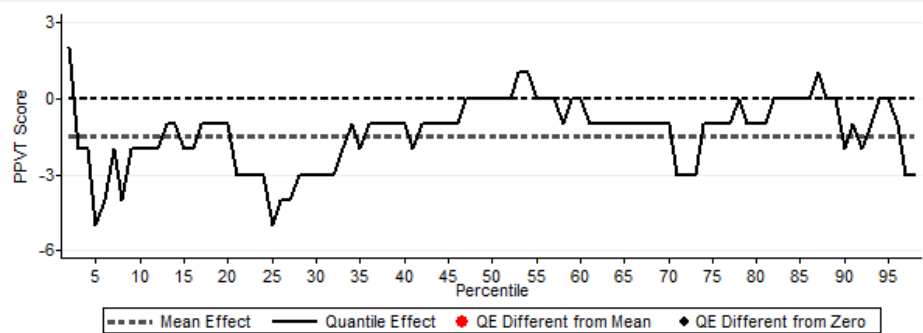
— Note: Change in change estimates are presented for each family type sample that are both unconditional and use propensity score weighting to control for covariates. The quantile effects, $F_{Y_{11}^{cf}}^{-1}(\tau) - F_{Y_{11}}^{-1}(\tau)$, is given by the solid line. The dashed line gives the mean 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

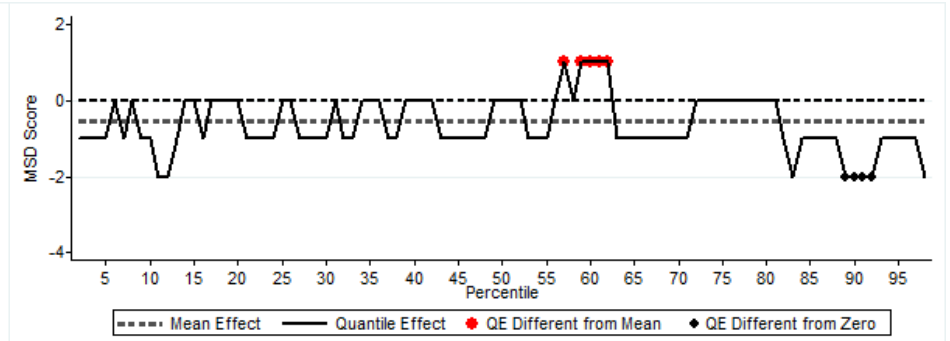
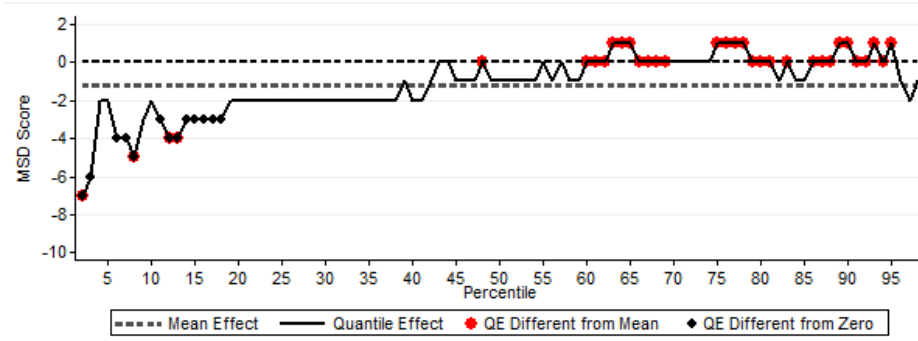


— Note: Change in change estimates are presented for each family type sample that are both unconditional and use propensity score weighting to control for covariates. The quantile effects, $F_{Y_{cf}}^{-1}(\tau) - F_{Y_{11}}^{-1}(\tau)$, is given by the solid line. The dashed line gives the mean effect 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

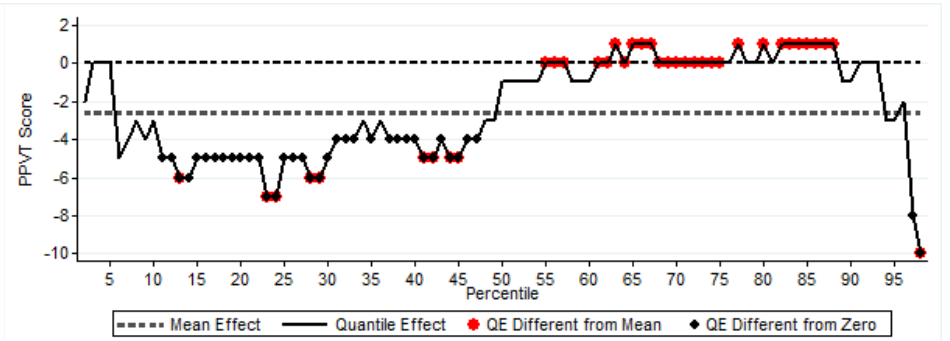
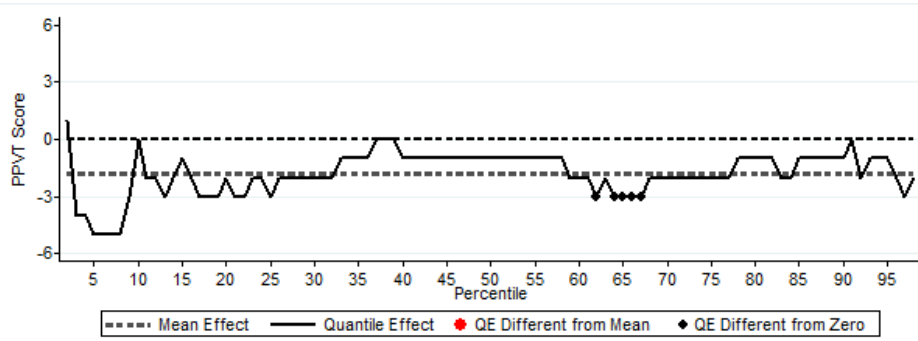
(a) Motor and Social Development - Girls
KS-Test: 0.006*** (0.004***)

(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, $F_{Y_{11}^{cf}}^{-1}(\tau) - F_{Y_{11}}^{-1}(\tau)$, is given by the solid line. The dashed 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 $F_{Y_{11}^{cf}}$ at the 1%, 5% and 10% level.

Table 3A: Summary Statistics and Test of Differences 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	0.465	0.028 (0.001)***	0.034 (0.000)***	0.037 (0.000)***	0.046 (0.000)***	0.379	0.061 (0.006)***	0.096 (0.000)***	0.121 (0.000)***	0.094 (0.000)***
Hours in Care	15.53	1.024 (0.003)***	1.902 (0.000)***	1.336 (0.000)***	1.46 (0.000)***	11.476	2.227 (0.007)***	3.267 (0.000)***	3.947 (0.000)***	3.109 (0.000)***
Mother Works	0.358	0.056 (0.000)***	0.049 (0.000)***	0.086 (0.000)***	0.076 (0.000)***	0.544	0.051 (0.02)**	0.068 (0.002)***	0.093 (0.000)***	0.067 (0.002)***
Mother Works 30 or more hours	0.255	0.018 (0.103)	0.022 (0.045)**	0.04 (0.000)***	0.057 (0.000)***	0.362	0.045 (0.158)	0.049 (0.115)	0.065 (0.037)**	0.044 (0.155)
Parenting Scales										
Family Dysfunction Index	10.98	-0.726 (0.000)***	-0.361 (0.000)***	-0.786 (0.000)***	-1.193 (0.000)***	8.284	-0.269 (0.252)	-0.351 (0.132)	-0.511 (0.029)**	-0.719 (0.002)***
Aversive Parenting	7.238	-0.823 (0.000)***	-0.71 (0.000)***	-0.976 (0.000)***	-1.277 (0.000)***	6.541	-0.441 (0.036)**	-0.632 (0.002)***	-0.681 (0.001)***	-0.833 (0.000)***
Ineffective Parenting	9.844	-0.946 (0.000)***	-0.909 (0.000)***	-0.941 (0.000)***	-1.679 (0.000)***	9.325	-0.333 (0.164)	-0.508 (0.026)**	-0.64 (0.005)***	-1.015 (0.000)***
Parent Consistency	13.069	0.956 (0.000)***	1.266 (0.000)***	1.223 (0.000)***	1.661 (0.000)***	14.443	0.442 (0.036)**	0.492 (0.017)**	0.619 (0.003)***	0.944 (0.000)***
Positive Interaction	15.46	0.723 (0.000)***	0.913 (0.000)***	1.13 (0.000)***	1.893 (0.000)***	16.11	0.231 (0.135)	0.512 (0.001)***	0.473 (0.001)***	0.892 (0.000)***
Parent-Child Activities										
Spends 5 minutes of focused time - many times a day	0.646	0.039 (0.000)***	0.061 (0.000)***	0.121 (0.000)***	0.174 (0.000)***	0.666	0.046 (0.029)**	0.055 (0.007)***	0.059 (0.004)***	0.109 (0.000)***
Laughs with child - many times a day	0.777	0.034 (0.000)***	0.077 (0.000)***	0.077 (0.000)***	0.143 (0.000)***	0.782	0.048 (0.007)***	0.056 (0.001)***	0.061 (0.000)***	0.09 (0.000)***
Does a special activity that the child enjoys - Once or twice a day or more	0.537	0.011 (0.196)	0.006 (0.465)	0.048 (0.000)***	0.12 (0.000)***	0.636	0.003 (0.895)	0.008 (0.707)	-0.021 (0.327)	0.032 (0.131)
Plays a sport, game, or hobby with child - Once or twice a day or more	0.65	0.027 (0.001)***	0.023 (0.003)***	0.096 (0.000)***	0.158 (0.000)***	0.733	0.025 (0.207)	0.02 (0.309)	0.003 (0.898)	0.061 (0.001)***
Reads to child - daily	0.564	0.071 (0.000)***	0.104 (0.000)***	0.124 (0.000)***	0.14 (0.000)***	0.596	0.07 (0.014)**	0.123 (0.000)***	0.12 (0.000)***	0.15 (0.000)***

— Note: For each variable of interest, we present the mean in the subgroup of the population defined by scores in the first quintile of the developmental score of the family type denoted in the header. For the remaining quintiles we present the difference in the mean from the first quintile with the p-value of a test for significant difference in the parenthesis. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

Table 3B: Summary Statistics and Test of Differences 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	0.452	0.032	0.166	0.206	0.167	0.41	0.088	0.135	0.11	0.153
		(0.44)	(0.000)***	(0.000)***	(0.000)***		(0.000)***	(0.000)***	(0.000)***	(0.000)***
Hours in Care	16.58	-0.78	3.417	5.113	3.137	12.214	0.357	3.29	1.617	2.992
		(0.659)	(0.052)*	(0.004)***	(0.113)		(0.652)	(0.000)***	(0.042)**	(0.000)***
Mother Works	0.413	0.051	0.139	0.193	0.164	0.597	0.052	0.099	0.108	0.115
		(0.218)	(0.001)***	(0.000)***	(0.000)***		(0.016)**	(0.000)***	(0.000)***	(0.000)***
Mother Works 30 or more hours	0.29	0.025	0.128	0.2	0.125	0.401	0.007	0.034	0.038	0.082
		(0.506)	(0.002)***	(0.000)***	(0.005)***		(0.829)	(0.273)	(0.224)	(0.01)***
Parenting Scales										
Family Dysfunction Index	10.454	-0.675	-1.643	-1.446	-2.079	8.657	-0.847	-1.288	-1.924	-1.837
		(0.134)	(0.000)***	(0.002)***	(0.000)***		(0.000)***	(0.000)***	(0.000)***	(0.000)***
Aversive Parenting	7.155	-0.574	-1.044	-1.253	-1.757	6.113	-0.106	-0.018	-0.236	-0.307
		(0.055)*	(0.000)***	(0.000)***	(0.000)***		(0.577)	(0.921)	(0.207)	(0.116)
Ineffective Parenting	10.075	-0.928	-1.091	-1.615	-1.891	8.812	-0.067	0.038	-0.089	-0.108
		(0.007)***	(0.001)***	(0.000)***	(0.000)***		(0.776)	(0.861)	(0.689)	(0.636)
Parent Consistency	13.498	0.187	1.423	1.677	1.735	14.326	1.025	1.145	1.265	1.296
		(0.558)	(0.000)***	(0.000)***	(0.000)***		(0.000)***	(0.000)***	(0.000)***	(0.000)***
Positive Interaction	14.447	0.981	0.404	1.029	0.734	14.991	0.12	0.34	0.426	0.487
		(0.000)***	(0.111)	(0.000)***	(0.005)***		(0.463)	(0.031)**	(0.008)***	(0.003)***
Parent-Child Activities										
Spends 5 minutes of focused time	0.467	0.056	0.024	0.054	0.077	0.504	-0.028	0.017	0.004	0.021
- many times a day		(0.18)	(0.572)	(0.231)	(0.101)		(0.211)	(0.453)	(0.864)	(0.355)
Laughs with child - many times a day	0.637	0.15	0.121	0.155	0.128	0.716	-0.011	0.02	-0.002	0.018
		(0.000)***	(0.002)***	(0.000)***	(0.003)***		(0.586)	(0.316)	(0.913)	(0.379)
Does a special activity that the child enjoys	0.24	0.144	0.098	0.134	0.029	0.365	-0.034	-0.016	-0.003	-0.013
- Once or twice a day or more		(0.000)***	(0.011)**	(0.001)***	(0.483)		(0.119)	(0.449)	(0.894)	(0.533)
Reads to child - daily	0.595	0.057	0.049	0.099	0.172	0.639	0.051	0.063	0.082	0.092
		(0.166)	(0.24)	(0.023)**	(0.000)***		(0.017)**	(0.003)***	(0.000)***	(0.000)***

— Note: For each variable of interest, we present the mean in the subgroup of the population defined by scores in the first quintile of the developmental score of the family type denoted in the header. For the remaining quintiles we present the difference in the mean from the first quintile with the p-value of a test for significant difference in the parenthesis. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

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 (0.001)***	-0.033 (0.566)	0.042 (0.518)	0.008 (0.893)	-0.047 (0.377)	0.284 (0.000)***	-0.283 (0.013)**	-0.224 (0.125)	-0.286 (0.026)**	-0.17 (0.23)
Mother Works 30 or more hours	0.086 (0.044)**	0.023 (0.752)	0.084 (0.195)	0.07 (0.301)	-0.03 (0.59)	0.144 (0.019)**	-0.227 (0.065)*	0.082 (0.39)	-0.226 (0.039)**	-0.021 (0.876)
In Child Care	0.175 (0.000)***	0.089 (0.328)	0.039 (0.623)	0.035 (0.656)	-0.025 (0.706)	0.445 (0.000)***	-0.063 (0.708)	-0.376 (0.046)**	-0.194 (0.33)	-0.297 (0.09)*
Hours in Care	8.666 (0.000)***	2.176 (0.585)	0.832 (0.822)	0.903 (0.808)	-2.508 (0.412)	17.517 (0.000)***	-2.702 (0.682)	-13.22 (0.031)**	-7.338 (0.344)	-10.608 (0.111)
Parenting Scales Activities										
Family Dysfunction Index	0.045 (0.93)	-0.099 (0.876)	1.134 (0.106)	0.478 (0.605)	0.011 (0.987)	0.258 (0.745)	-1.153 (0.254)	-1.595 (0.183)	-0.305 (0.845)	0.598 (0.672)
Aversive Parenting	0.359 (0.396)	-0.296 (0.624)	-0.046 (0.93)	0.184 (0.738)	0.124 (0.798)	0.484 (0.352)	0.401 (0.622)	-0.871 (0.337)	0.173 (0.779)	-0.288 (0.766)
Ineffective Parenting	-0.107 (0.867)	0.547 (0.44)	1.231 (0.086)*	0.497 (0.515)	1.261 (0.08)*	0.637 (0.405)	-0.843 (0.471)	-0.339 (0.741)	3.145 (0.018)**	0.275 (0.891)
Parent Consistency	-0.006 (0.989)	0.118 (0.86)	0.348 (0.532)	-0.186 (0.755)	-0.624 (0.232)	2.101 (0.017)**	-2.978 (0.005)***	-1.529 (0.226)	-3.197 (0.049)**	-5.822 (0.000)***
Positive Interaction	0.061 (0.861)	-1.355 (0.013)**	-0.466 (0.333)	-0.349 (0.386)	-1.071 (0.018)**	-1.169 (0.202)	-0.925 (0.424)	0.12 (0.904)	0.228 (0.824)	1.369 (0.192)
Parent-Child Activities										
Spends 5 minutes of focused time	-0.036 (0.353)	-0.082 (0.146)	-0.029 (0.66)	0.059 (0.236)	-0.096 (0.077)*	-0.185 (0.014)**	0.079 (0.599)	0.183 (0.231)	0.213 (0.026)**	0.138 (0.201)
- many times a day	-0.055 (0.066)*	-0.005 (0.905)	0.027 (0.494)	0.037 (0.308)	-0.041 (0.314)	-0.229 (0.005)***	0.16 (0.241)	0.187 (0.085)*	0.144 (0.169)	0.22 (0.017)**
Laughs with child - many times a day	0.005 (0.875)	-0.06 (0.167)	-0.012 (0.845)	-0.055 (0.303)	-0.102 (0.022)**	-0.112 (0.046)**	-0.336 (0.023)**	-0.028 (0.712)	-0.074 (0.379)	0.228 (0.052)*
Does a special activity that the child enjoys	0.125 (0.002)***	-0.104 (0.031)**	-0.139 (0.005)***	-0.069 (0.153)	-0.132 (0.011)**	-0.075 (0.315)	-0.008 (0.947)	-0.099 (0.48)	0.023 (0.865)	0.318 (0.025)**
- Once or twice a day or more	-0.178 (0.000)***	0.085 (0.252)	0.131 (0.026)**	0.089 (0.144)	0.146 (0.036)**	-0.312 (0.023)**	0.354 (0.027)**	0.561 (0.003)***	0.331 (0.08)*	0 (0.968)
Plays a sport, game, or hobby with child										
- Once or twice a day or more										
Reads to child - daily										

— Note: The first column (Q1) for the corresponding family type, as denoted in the header, presents, for each variable of interest, the linear difference-in-differences estimate of mean effect of the policy and p-value (in parentheses) specific to the first quintile of the motor-social development score. The following columns (Q2 - Q5) report the differences in estimates of the mean effect between the first quintile and the quintile denoted in the header. The p-value testing the significance of this differences is presented in parenthesis below. All reported p-values are corrected at the province-year level. To account for the change in the false discovery rate when examining multiple measures, we use Simes' (1986) improved Bonferroni correction procedure to further adjust our p-values. The method for this adjustment is outlined in Newson et al. (2003). The estimates used in this table are derived from estimates of the β coefficient 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 a given 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 (0.668)	0.298 (0.086)*	0.216 (0.186)	0.299 (0.065)*	0.028 (0.864)	0.005 (0.98)	-0.072 (0.773)	-0.182 (0.479)	-0.128 (0.411)	0.191 (0.418)
Mother Works 30 or more hours	0.016 (0.9)	-0.043 (0.782)	0.067 (0.646)	0.125 (0.427)	0.029 (0.849)	-0.162 (0.406)	-0.2 (0.351)	-0.224 (0.333)	-0.036 (0.886)	0.2 (0.188)
In Child Care	0.162 (0.062)*	0.223 (0.113)	-0.045 (0.698)	0.021 (0.872)	-0.071 (0.535)	0.347 (0.011)**	0.285 (0.022)**	-0.082 (0.665)	0.228 (0.447)	0.633 (0.000)***
Hours in Care	10.192 (0.000)***	-1.404 (0.76)	3.107 (0.519)	-0.951 (0.81)	-5.307 (0.168)	23.215 (0.001)***	5.174 (0.367)	-0.586 (0.918)	8.009 (0.427)	27.222 (0.000)***
Parenting Scales Activities										
Family Dysfunction Index	-0.843 (0.378)	0.236 (0.84)	2.472 (0.092)*	0.286 (0.835)	0.252 (0.845)	2.717 (0.066)*	1.275 (0.486)	-2.296 (0.228)	-0.942 (0.563)	4.235 (0.015)**
Aversive Parenting	0.493 (0.327)	-0.248 (0.642)	0.037 (0.95)	-0.724 (0.282)	-0.352 (0.532)	1.91 (0.094)*	0.977 (0.2)	0.341 (0.575)	0.716 (0.225)	2.357 (0.007)***
Ineffective Parenting	0.257 (0.755)	1.022 (0.35)	0.479 (0.584)	1.18 (0.356)	0.07 (0.941)	-0.606 (0.649)	1.698 (0.315)	1.255 (0.222)	-1.225 (0.379)	3.864 (0.005)***
Parent Consistency	-0.757 (0.306)	0.718 (0.478)	-0.449 (0.597)	-0.365 (0.681)	1.062 (0.191)	0.415 (0.628)	0.782 (0.544)	-1.845 (0.061)*	-0.101 (0.908)	-1.433 (0.183)
Positive Interaction	-0.197 (0.745)	-0.405 (0.628)	-0.386 (0.543)	0.198 (0.785)	0.253 (0.707)	-2.472 (0.179)	-0.528 (0.458)	-2.011 (0.019)**	0.166 (0.864)	-2.535 (0.003)***
Parent-Child Activities										
Spends 5 minutes of focused time - many times a day	0.071 (0.382)	-0.064 (0.554)	-0.11 (0.335)	0.091 (0.495)	-0.04 (0.676)	-0.411 (0.101)	-0.033 (0.8)	0.286 (0.062)*	0.041 (0.775)	-0.169 (0.334)
Laughs with child - many times a day	-0.025 (0.734)	-0.153 (0.088)*	0.009 (0.926)	-0.048 (0.707)	0.003 (0.976)	-0.352 (0.18)	-0.2 (0.04)**	-0.166 (0.37)	0.108 (0.597)	-0.651 (0.000)***
Does a special activity that the child enjoys - Once or twice a day or more	-0.222 (0.002)***	0.196 (0.087)*	0.144 (0.127)	0.139 (0.179)	0.162 (0.163)	-0.133 (0.317)	-0.083 (0.533)	-0.139 (0.039)**	-0.05 (0.829)	0.076 (0.596)
Reads to child - daily	-0.044 (0.646)	0.025 (0.819)	0.117 (0.304)	0.234 (0.063)*	0.32 (0.022)**	0.411 (0.001)***	0.43 (0.061)*	0.455 (0.012)**	0.157 (0.418)	0.016 (0.923)

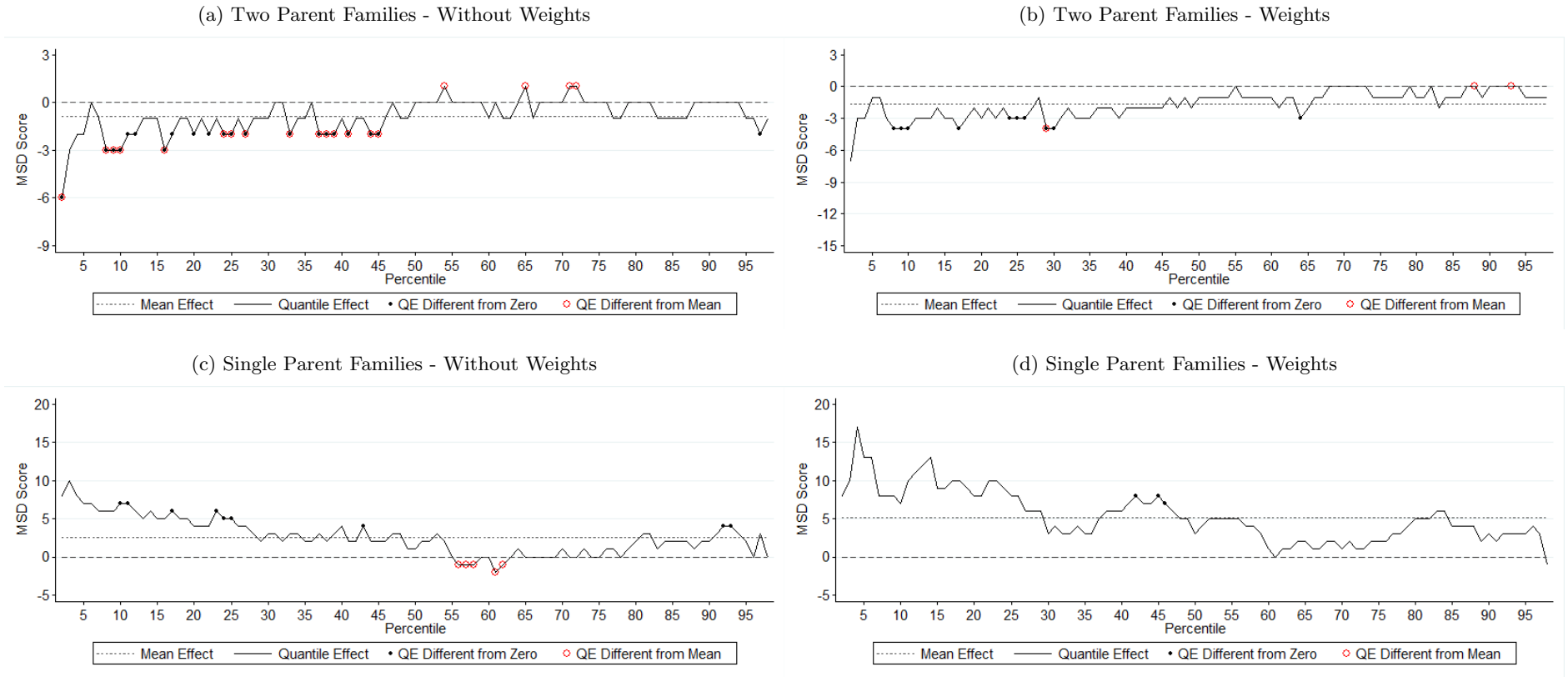
— Note: The first column (Q1) for the corresponding family type, as denoted in the header, presents, for each variable of interest, the linear difference-in-differences estimate of mean effect of the policy and p-value (in parentheses) specific to the first quintile of the motor-social development score. The following columns (Q2 - Q5) report the differences in estimates of the mean effect between the first quintile and the quintile denoted in the header. The p-value testing the significance of this differences is presented in parenthesis below. All reported p-values are corrected at the province-year level. To account for the change in the false discovery rate when examining multiple measures, we use Simes' (1986) improved Bonferroni correction procedure to further adjust our p-values. The method for this adjustment is outlined in Newson et al. (2003). The estimates used in this table are derived from estimates of the β coefficient 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 a given 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 (0.01)***	-1.707 (0.508)	0.322 (0.77)	1.825 (0.844)	2.117 (0.177)	-3.801 (0.195)	2.407 (0.004)***	2.054 (0.204)
Aversive Parenting	1.665 (0.000)***	0.725 (0.041)**	1.397 (0.09)*	3.181 (0.08)*	2.562 (0.000)***	0.106 (0.842)	0.600 (0.188)	1.987 (0.024)**
Ineffective Parenting	2.814 (0.000)***	3.572 (0.000)***	4.689 (0.005)***	1.555 (0.71)	3.304 (0.000)***	2.125 (0.028)**	2.146 (0.082)*	6.598 (0.001)***
Parent Consistency	-0.396 (0.643)	-0.898 (0.234)	-2.283 (0.224)	0.362 (0.924)	-1.216 (0.122)	-0.698 (0.423)	0.679 (0.503)	-1.286 (0.261)
Positive Interaction	-2.858 (0.000)***	-1.354 (0.061)*	-2.419 (0.005)***	-3.826 (0.02)**	-3.558 (0.000)***	-1.034 (0.138)	-2.147 (0.017)**	-2.216 (0.024)**
Parent-Child Activities								
Spends 5 minutes of focused time - many times a day	-0.361 (0.002)***	0.04 (0.922)	-0.032 (0.77)	-0.089 (0.846)	-0.644 (0.000)***	-0.043 (0.81)	-0.304 (0.018)**	-1.009 (0.006)***
Laughs with child - many times a day	-0.284 (0.000)***	-0.425 (0.000)***	-0.311 (0.1)*	-0.536 (0.035)**	-0.432 (0.000)***	-0.228 (0.042)**	-0.178 (0.001)***	-0.984 (0.015)**
Does a special activity that the child enjoys - Once or twice a day or more	-0.253 (0.000)***	-0.421 (0.003)***	-0.393 (0.14)	-0.237 (0.298)	-0.356 (0.000)***	-0.264 (0.005)***	-0.259 (0.000)***	-0.969 (0.006)***
Plays a sport, game, or hobby with child - Once or twice a day or more	0.132 (0.041)**	0.13 (0.317)	0.218 (0.231)	-0.367 (0.027)**	0.027 (0.783)	0.02 (0.879)	0.079 (0.054)*	0.226 (0.512)
Reads to child - daily	-0.537 (0.003)***	0.621 (0.004)***	-0.442 (0.14)	0.852 (0.08)*	-1.203 (0.009)***	0.072 (0.795)	-0.352 (0.003)***	-0.32 (0.444)
F-Test	455.93***	277.9***	76.88***	66.96***	149.82***	212.54***	502.93***	251.48***

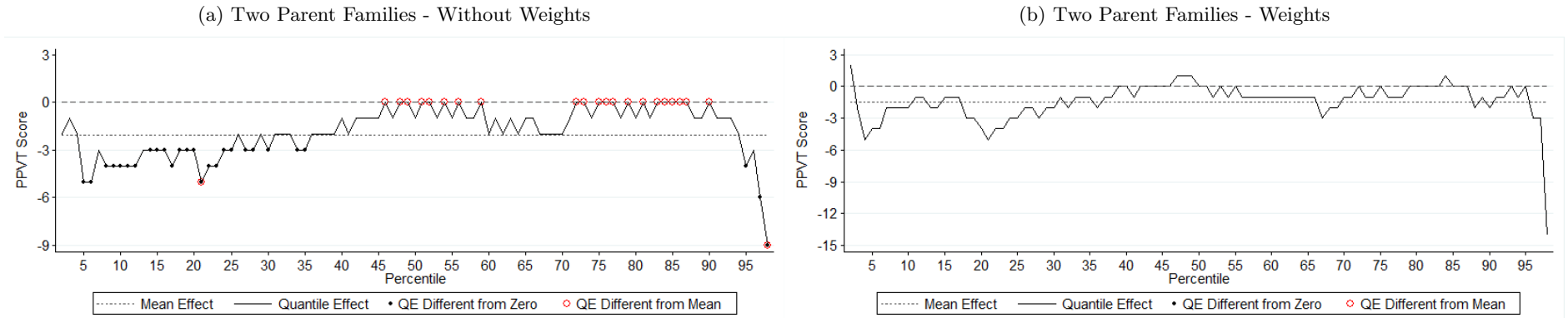
— Note: For the outcome variable in each row we present estimates of treatment effect of childcare, the coefficient on Ccare using an instrumental variable estimator for samples denoted in the column heading that differ by family type, child age and child gender. The specification includes all of the 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.

Figure A1: Change-in-Change Estimates for Motor and Social Development



— 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, $F_{Y_{11}^{cf}}^{-1}(\tau) - F_{Y_{11}}^{-1}(\tau)$, is given by the solid line. The dashed 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 A2: Change-in-Change Estimates for Peabody Picture and Vocabulary Test



— 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, $F_{Y_{cf}}^{-1}(\tau) - F_{Y_{11}}^{-1}(\tau)$, is given by the solid line. The dashed 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.