ONLINE APPENDIX

"The Long-Run Impacts of a Universal Child Care Program"

Michael Baker, University of Toronto and NBER, baker@chass.utoronto.ca
Jonathan Gruber, MIT and NBER, gruberj@mit.edu
Kevin Milligan, University of British Columbia and NBER, kevin.milligan@ubc.ca

Appendix A provides background information on the quality of care provided by the Quebec program and on child eligibility.

Appendix B has detail on, and discussion of, the behavioral scores we analyze in the paper.

Appendix C details the availability and specification of the control variables in each data source.

Appendix D provides supplemental analysis of the non-cognitive scores, including heterogeneity in the contemporaneous (preschool) outcomes, and splits of the age 5-9 outcomes by gender.

Appendix E provides supplementary analysis of our crime data, including additional figures, an analysis of leads and lags of the policy variable and separate estimates by gender.

Appendix A: Quality of Care and Program Eligibility

Childcare Quality in the Quebec Program and Comparisons to Other Jurisdictions

Japel et al. (2005) offer a view of the care provided by the Quebec program just after it was extended to all age groups (2000-2003). They report that just over 60 percent of the venues evaluated provided care at the level of "minimal quality", meaning the health and safety of the children were provided for, but the educational component was weak. Just over one-quarter provided services that were judged good, very good or excellent. CPE-based care was found superior to that provided in other settings. It is important to note that this evaluation included unregulated home-based care which is not part of the Quebec program (which provides regulated care).[1] Also, that the authors report that their conclusions are consistent with Drouin et al. (2004), a 2003 study of child care quality undertaken by the government of Quebec through the Institut de la statistique du Quebec. [2]

The Quebec government undertook a second review of child care quality in 2014 (Gingras et al 2015 and Lavoie et al. 2015). This review focused on CPE provided care, and non-subsidized care (which was not reviewed in the 2003 government study). The quality of CPE based care was found to differ little from the levels recorded in 2003, with children less than 18 months receiving "good" care and those 18 months through 5 years receiving "acceptable" care.

[1] The unregulated providers were found to offer the lowest quality care.
[2] This study, which only reviews regulated places, reports an average quality of "fair", using a different scale than Japel et al (2005). We focus on the Japel study for the early period because, as explained below, they use a rating instrument that can be compared to studies of child care in other countries.

The Japel et al. (2005) study adopts the widely-used Early Childhood Environment Rating Scale (ECERS-R) to evaluate center provided care, and the Family Day Care Rating Scale (FDCRS) to evaluate home based care.[3] These instruments use a 7-point scale, with scores <3 denoted inadequate, 3 to <5 denoted minimal and 5 to 7 denoted acceptable. They report overall scores, using these instruments, of 4.58 in CPE center based care, 4.41 in CPE home based care, 3.69 in for profit care and 3.60 in unregulated care.

Helmerhorst et al. (2014) provide a view of the quality of child care in the Netherlands in 2008 for a nationally representative sample using both the ECERS-R and the Infant/Toddler Environment Rating Scale–Revised (a 7 point scale for care of younger children).[4] They report an overall mean score of 3.0. Sylva et al (2006) report a mean score on the ECESR-R for a nationally representative sample of preschools (ages 3-5) in the UK of 4.34. In the US, a study of licensed full day child care facilities in Nebraska, Iowa, Missouri and Kansas (Torquati et al 2011) report a mean ECESR-R of 4.48 and a mean FDCRS of 4.25. An earlier cross country study of child care quality in the early 1990s (Cryer et al. 1999), using the ECESR, a predecessor to the ECESR-R, reports mean scores of 4.47 for a German sample, 4.31 for the sample from Portugal, 4.04 for the Spanish sample and 4.25 for the sample for the US.

Finally, child care in Scandinavian countries is often viewed as some of the highest quality. We are not aware of any national level evaluations using the ECESR-R for these

[3] The authors report (p. 11) "Although the items [in the FDCRS] reflect the reality of this type of child care, the structure of the scale is almost identical to that of the ECERS-R (32 items grouped into 6 subscales: Space and Furnishings for Care and Learning, Basic Care, Language and Reasoning, Learning Activities, Social Development, Adult Needs)".

[4] The studies for comparison are drawn from those reported in Vermeer et al. (2016) as having a high representativeness of the underlying population.

countries. Sheridan et al. (2009) report an overall score of 4.52 for a sample of venues in Gothenburg using a Swedish adaptation of the ECESR.

The message of this comparison is the quality of care in the Quebec program, as measured by the ECERS-R, is not largely different from the quality of care offered in many other developed countries. We note, however, that we are unable to make meaningful comparisons of the Quebec care to the care offered in many Scandinavian countries such as Norway. Also the quality of care in for-profit and unregulated venues is markedly lower. While the proportion of regulated places in for-profit venues has been growing recently, up until 2009, which would cover the treated children in our samples, it was 20 percent or less.[5]

Eligibility

Eligibility for the program was extended sequentially by age between 1997 (4 year olds) and 2000 (0 and 1 year olds). This means that cohorts born before 1993 were not eligible, those born between 1993 and 1999 were partially eligible, and those born from 2000 onward were eligible for all ages between 0 and 4. This eligibility pattern is displayed in Appendix Figure 1.

References

Canadian Test Centre, Educational Assessment Services (1992) *Canadian Achievement Tests, 3rd ed.* (Ontario: Canadian Test Centre, Educational Assessment Services)

Cryer, Debby, Tietze, Wolfgang, Burchinal, Margaret, Leal, Teresa and Jesus Palacios (1999) "Predicting Process Quality from Structural Quality in Preschool Programs: A Cross-Country Comparison", *Early Childhood Research Quarterly*, 14(3) 339–361

Drouin, C., N. Bigras, C. Fournier, H. Desrosiers, and S. Bernard (2004) Grandir en qualité 2003: Enquête québécoise sur la qualité des services de garde éducatifs. Quebec: Institut de la statistique du Québec.

[5] The numbers of child care places in different venues over time can be found at https://www.mfa.gouv.qc.ca/fr/services-de-garde/portrait/places/Pages/index.aspx (accessed July 6, 2018).

Gingras, Lucie, Amélie Lavoie, and Nathalie Audet (2015) "La qualité éducative dans les installations de centres de la petite enfance. Faits saillants, Enquête québécoise sur la qualité des services de garde éducatifs 2014", Québec, Institut de la statistique du Québec, 16 p.

Helmerhorst , Katrien O. W. Riksen-Walraven, J. Marianne A. Gevers Deynoot-Schaub, Mirjam J. J. M. Tavecchio, Louis W. C. and Ruben G. Fukkink (2015) "Child Care Quality in The Netherlands Over the Years: A Closer Look", *Early Education and Development*, 26(1): 89-105

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

Lavoie, Amélie, Lucie Gingras, and Nathalie Audet (2015) "La qualité éducative dans les garderies non subventionnées. Faits saillants, Enquête québécoise sur la qualité des services de garde éducatifs 2014", Québec, Institut de la statistique du Québec, 16 p.

Sheridan, Sonja, Giota, Joanna, Han, You-Me and Jeong-Yoon Kwon, (2009) "A cross-cultural study of preschool quality in South Korea and Sweden: ECERS evaluations" *Early Childhood Research Quarterly* 24: 142–156

Sylva, Kathy, Siraj-Blatchfordb, Iram, Taggartb, Brenda, Sammonsc, Pam, Melhuishd, Edward Elliote, Karen and Vasiliki Totsika (2006) "Capturing quality in early childhood through environmental rating scales" *Early Childhood Research Quarterly* 21: 76–92

Torquatia, Julia C., Raikesa, Helen H., Huddleston-Casasa, Catherine A., Bovairda, James A. and Beatrice A. Harris (2011) "Family income, parent education, and perceived constraints as predictors of observed program quality and parent rated program quality" *Early Childhood Research Quarterly* 26: 453-464.

Vermeer, Harriet J., van IJzendoorn, Marinus H., Carcamo, Rodrigo A. and Linda J. Harrison (2016) "Quality of Child Care Using the Environment Rating Scales: A Meta-Analysis of International Studies" *International Journal of Early Childhood*, 48(1): 33-60

Appendix Figure 1: Cohort map for program eligibility

Year Of Observation	Age 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1997	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1998	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1999	0	0	1	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2000	1	1	1	2	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	1	2	2	2	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
2002	1	2	3	3	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0	0
2003	1	2	3	4	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0	0
2004	1	2	3	4	5	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0	0
2005	1	2	3	4	5	5	4	3	3	2	2	1	1	0	0	0	0	0	0	0	0
2006	1	2	3	4	5	5	5	4	3	3	2	2	1	1	0	0	0	0	0	0	0
2007	1	2	3	4	5	5	5	5	4	3	3	2	2	1	1	0	0	0	0	0	0
2008	1	2	3	4	5	5	5	5	5	4	3	3	2	2	1	1	0	0	0	0	0
2009	1	2	3	4	5	5	5	5	5	5	4	3	3	2	2	1	1	0	0	0	0
2010	1	2	3	4	5	5	5	5	5	5	5	4	3	3	2	2	1	1	0	0	0
2011	1	2	3	4	5	5	5	5	5	5	5	5	4	3	3	2	2	1	1	0	0
2012	1	2	3	4	5	5	5	5	5	5	5	5	5	4	3	3	2	2	1	1	0
2013	1	2	3	4	5	5	5	5	5	5	5	5	5	5	4	3	3	2	2	1	1
2014	1	2	3	4	5	5	5	5	5	5	5	5	5	5	5	4	3	3	2	2	1
2015	1	2	3	4	5	5	5	5	5	5	5	5	5	5	5	5	4	3	3	2	2

Notes: The figure shows eligibility for the Quebec childcare program for a child reaching the given age in the given year. The eligibility reported is the number of years of lifetime exposure, so a child age 9 who was eligible from ages 0 to 4 has 5 years of lifetime eligibility. The eligibility ranges from zero for those born before the program was introduced to 5 for those who were eligible from ages 0 to 4.

Appendix B: The NLSCY and SYC Behavioral Scores

The behavioral scores in the NLSCY and SYC are built up from individual questions with three categorical responses, typically “never/rarely”, “sometimes” “often/always”. Each response is assigned 0, 1, or 2 points corresponding to the three categories, and then added up across the questions to form the behavioral score. There are some differences in questions across the age 2-3 and age 5-9 samples. The underlying questions for the SYC were identical to the NLSCY. More detail on these scores is available in the online appendix to Baker et al. (2008). The list below indicates what questions are used to form each behavioral score we use in this paper.

Appendix Table 1: Non-cognitive Behavioral Score Components

Behavioral Scores	Ages 2-3	Ages 5-9
Hyperactivity		
Can’t sit still, is restless or hyperactive?	X	X
Is distractible, has trouble sticking to any activity?	X	X
Can't concentrate, can't pay attention for long?	X	X
Has difficulty awaiting turn in games or groups?	X	X
Is inattentive?	X	X
Cannot settle to anything for more than a few moments?	X	X
Is impulsive, acts without thinking?		X
Anxiety		
Seems to be unhappy, sad or depressed?	X	X
Is not as happy as other children?	X	X
Is too fearful or anxious?	X	X
Is worried?	X	X
Cries a lot?		X
Is nervous, highstrung or tense?	X	X
Has trouble enjoying him/herself?	X	X
Separation Anxiety		
Cries a lot?	X	
Clings to adults or is too dependent	X	
Constantly seeks help	X	
Gets upset when separated from parents?	X	

Doesn’t want to sleep alone	X	

Behavioral Scores	**Ages 2-3**	**Ages 5-9**
Aggression		
Gets into many fights?	X	X
When another child accidentally hurts him/her (such as by bumping into him/her), assumes that the other child meant to do it, and then reacts with anger and fighting?	X	X
Physically attacks people?		X
Threatens people?		X
Is cruel, bullies or is mean to others?		X
Kicks, bites, hits other children?	X	X
Is defiant?	X	
Doesn't seem to feel guilty after misbehaving?	X	
Punishment doesn't change his behaviour?	X	
Has temper tantrums or hot temper?	X	
Has angry moods?	X	
Indirect Aggression		
When mad at someone, tries to get others to dislike that person?		X
When mad at someone, becomes friends with another as revenge?		X
When mad at someone, says bad things behind the other's back?		X
When mad at someone, says to others: let's not be with him/her?		X
When mad at someone, tells the other one's secrets to a third person?		X
Prosocial		
Shows sympathy to someone who has made a mistake?		X
Will try to help someone who has been hurt?		X
Volunteers to help clear up a mess someone else has made?		X
If there is a quarrel or dispute, will try to stop it?		X
Offers to help other children (friend, brother or sister) who are having difficulty with a task?		X
Comforts a child (friend, brother, or sister) who is crying or upset?		X
Spontaneously helps to pick up objects which another child has dropped (e.g. pencils, books, etc.)?		X
Will invite bystanders to join in a game?		X
Helps other children (friends, brother or sister) who are feeling sick?		X
Takes the opportunity to praise the work of less able children?		X

Appendix C: Specification of Control Variables by Data Source

This appendix shows the control variables we use in different data sets for our regressions reported in the paper.

Appendix Table 2: Control variables available in the analysis samples

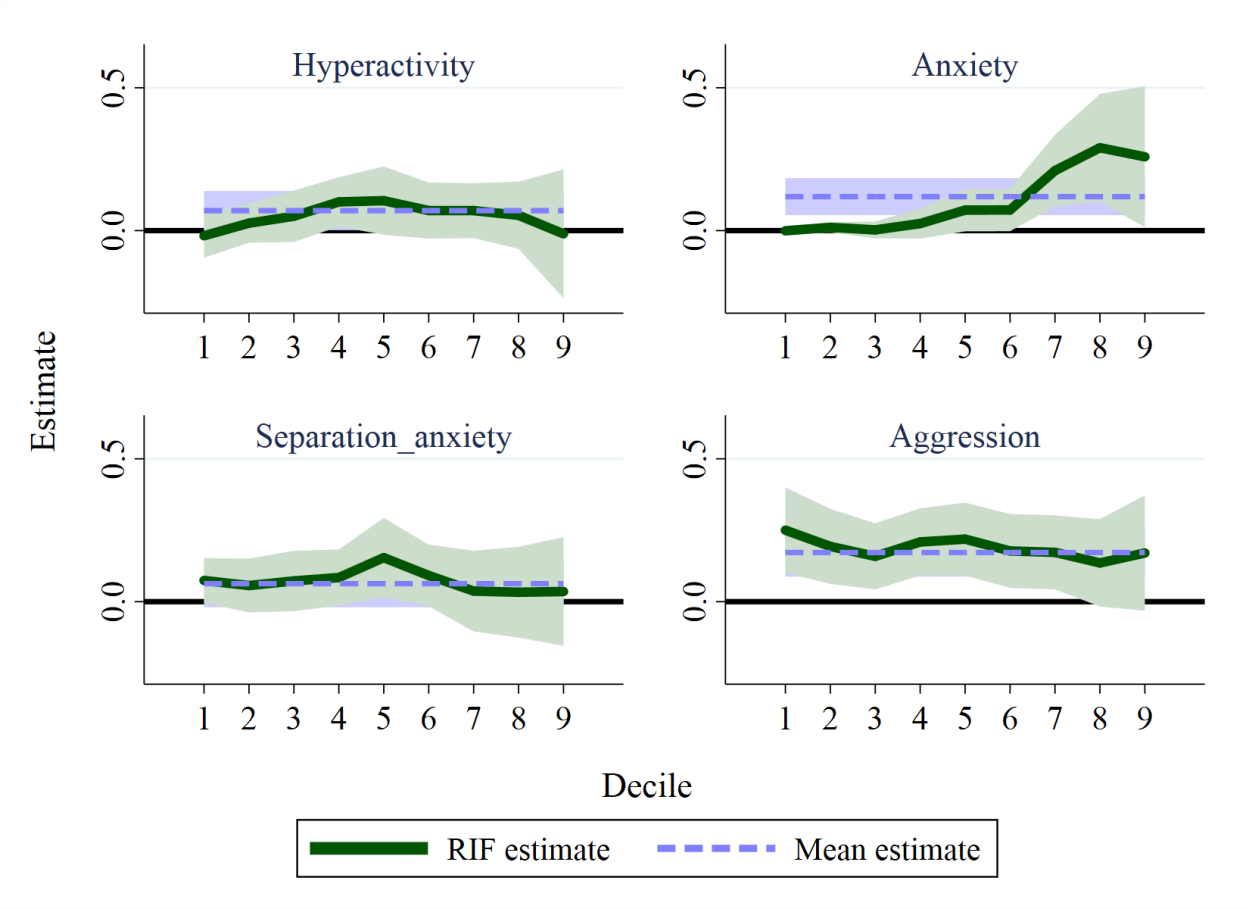
	NLSCY/SYC	CCHS	CHMS	SAIP/PCAP	PISA	UCRS
Male	Dummy	Dummy	Dummy	Dummy	Dummy	Dummy
Province	Dummies	Dummies	Dummies	Dummies	Dummies	Dummies
Year	Dummies	Dummies	Dummies	Dummies	Dummies	Dummies
Own Age	Dummies	Dummies	Dummies			Dummies
Month of Birth					Dummies	
Mother's Education	Dummies				Dummies	
Mother's Age	Dummies					
Father's Age	Dummies					
Highest Education in Family		Dummies	Dummies			
Two Parent Family	Dummy	Dummy	Dummy			
Number of Younger Siblings	Dummies					
Number of Older Siblings	Dummies					
Number of Children in Household <12		Dummies	Dummies			
Mother is Immigrant	Dummy				Dummy	
Father is Immigrant					Dummy	
Child born in Canada		Dummy	Dummy			
Family is not "white"		Dummy	Dummy			

Appendix D: Supplemental Analysis of Non-cognitive Outcomes

We provide two supplementary analyses in this appendix. First we present RIF estimates across deciles of the unconditional outcome distribution for the contemporaneous outcomes presented in Table 1 (ages 2-3). Second we provide estimates by gender for the school age results presented in Table 2 (ages 5-9).

Appendix Figure 2 shows the mean estimated impacts from Table 1 along with RIF estimates (based on the estimator in Firpo, Fortin, and Lemeiux 2009) at decile cutoffs. The figures demonstrate that the impact is close to the mean impact across deciles for three of the four non cognitive outcomes. Notably for anxiety, however, we observe a larger point estimates at higher deciles, consistent with the evidence for older ages in figure 2.

Appendix Figure 2: Quantile Regression Analysis of Non-cognitive Scores, ages 2-3



Notes: Displayed are the mean and RIF estimates for contemporaneous outcomes at ages 2-3 in the NLSCY/SCF. The dependent variable is scaled to mean zero and a unit standard deviation, so the estimates can be interpreted as fractions of a standard deviation. The mean estimates are from Table 1. The RIF estimates vary by decile cutoff and are described in the text. We show shaded 95 percent confidence intervals around each estimate.

We next provide supplemental estimates by gender for the outcome variables in Table 2. The results are reported in Appendix Table 3. We observe much stronger impacts for most measures for boys. The dependent variables are standardized separately by sex, but we also note that the means of these outcomes differ by sex. That said, the negative impacts on Hyperactivity and Aggression are primarily for boys, at one-fifth and almost one third of a standard deviation, respectively, as the estimates for females for these outcomes are small and statistically insignificant. For girls the strongest effect is on prosocial behavior, which worsens by almost 20 percent of a standard deviation.

Appendix Table 3: Gender Differences in the Impacts of the Quebec Family Plan on Non cognitive skills, ages 5-9

	Hyperactivity	Anxiety	Aggression	Indirect Aggression	Prosocial	Get Along with Teacher
Girls	0.060	0.177	0.060	0.161	0.199	-0.068
	[0.061]	[0.019]***	[0.040]	[0.041]***	[0.059]***	[0.019]***
	(0.070)	(0.027)***	(0.047)	(0.051)***	(0.050)**	(0.005)***
Boys	0.200	0.387	0.295	0.241	-0.217	-0.106
	[0.063]**	[0.074]***	[0.054]***	[0.047]***	[0.058]***	[0.031]***
	(0.061)*	(0.020)***	(0.052)**	(0.037)**	(0.047)*	(0.022)**

Notes: from NLSCY (cycles 1, 2, 7) and the SYC (cycle 9). Sample includes all families. Reported is the coefficient on a dummy indicating exposure. Significance at the 10, 5, and 1 percent levels is indicated with 1, 2, and 3 asterisks respectively, using p-values adjusted for multiple testing.

Appendix E: Supplementary Crime Analysis

In this appendix, we supplement the crime analysis in the paper with additional figures, including an analysis of convictions and disaggregate analysis by type of crime. Next, we present regression results with leads and lags of the main policy variable included in the specification. Finally we present estimates of the impact of the Program on accusation rates disaggregated by gender.

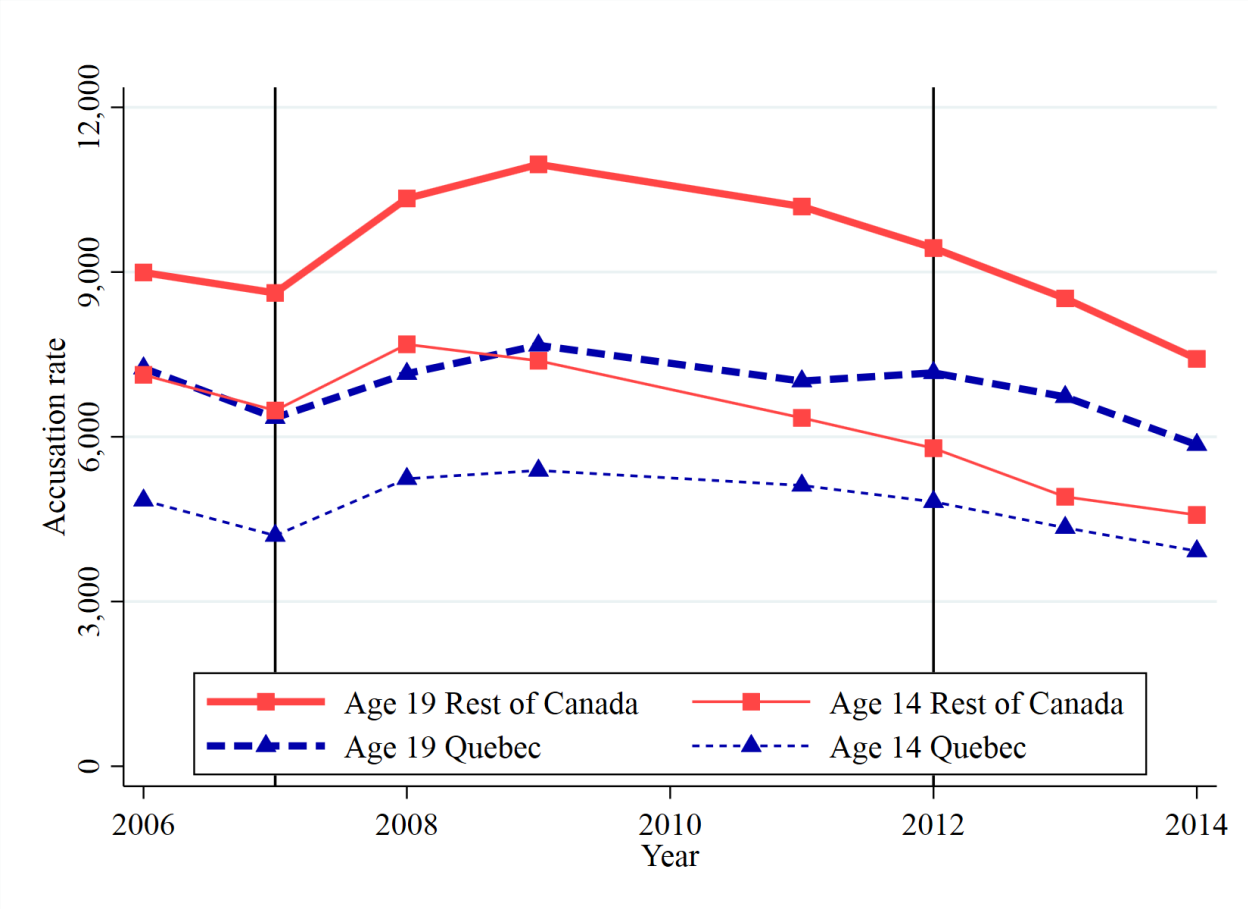
Appendix Figure 3a and 3b present time trends for accusation and conviction rates in Quebec and the rest of Canada. This provides insight to the credibility of the identifying assumptions of our analysis. In Figure 3a we display the evolution of annual total accusation rates per 100,000 of population in Quebec and the rest of Canada at ages 14 and 19.[6] Figure 3b shows the same for convictions. Cohorts born after 1993 had some program eligibility, reaching five years of eligibility for the 2000 birth cohort and beyond. We have drawn vertical lines at 2007 for 14-year olds and 2012 for 19-year olds indicting the first year a cohort with any exposure enters the relevant age category.

For the 14-year olds, the gap between the lines for Quebec and the rest of Canada diminishes as the number of years of exposure increases between the 1993 and 2000 birth cohorts (in the years after 2007). This is consistent with a dose-response relationship between years of exposure and the accusation rate. In contrast, the gap for 19-year olds is fairly constant before 2012, which is the first year 19-year olds have some eligibility. However, it starts to close

[6] We select these two specific ages so we can precisely indicate the years in which children of these ages were eligible for the program. The results look similar if other ages are chosen.

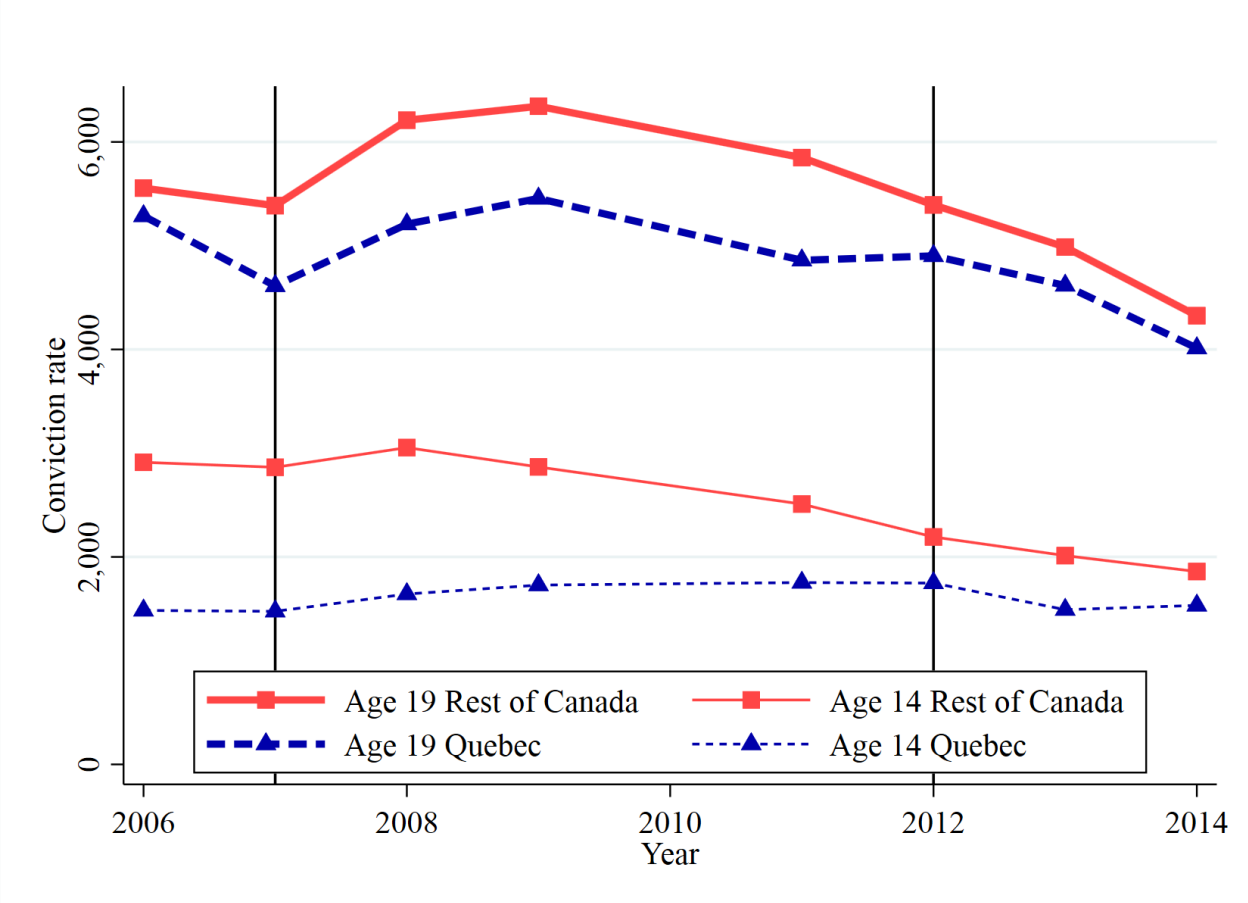
at this point. This suggests our results are not driven by common year effects across provinces, as children of different ages in a given year show trends consistent with a policy impact.

Appendix Figure 3a: Time trends for Accusation rates in Quebec and Rest of Canada, ages 14 and 19



Notes: Authors' calculations from UCR data. The graph shows the annual accusation rate for Quebec and the rest of Canada at ages 14 and 19 as indicated. The vertical lines indicate the year in which cohorts are first exposed to eligibility: 2007 for 14 year olds and 2012 for 19 year olds.

Appendix Figure 3b: Time trends for Conviction rates in Quebec and Rest of Canada, ages 14 and 19

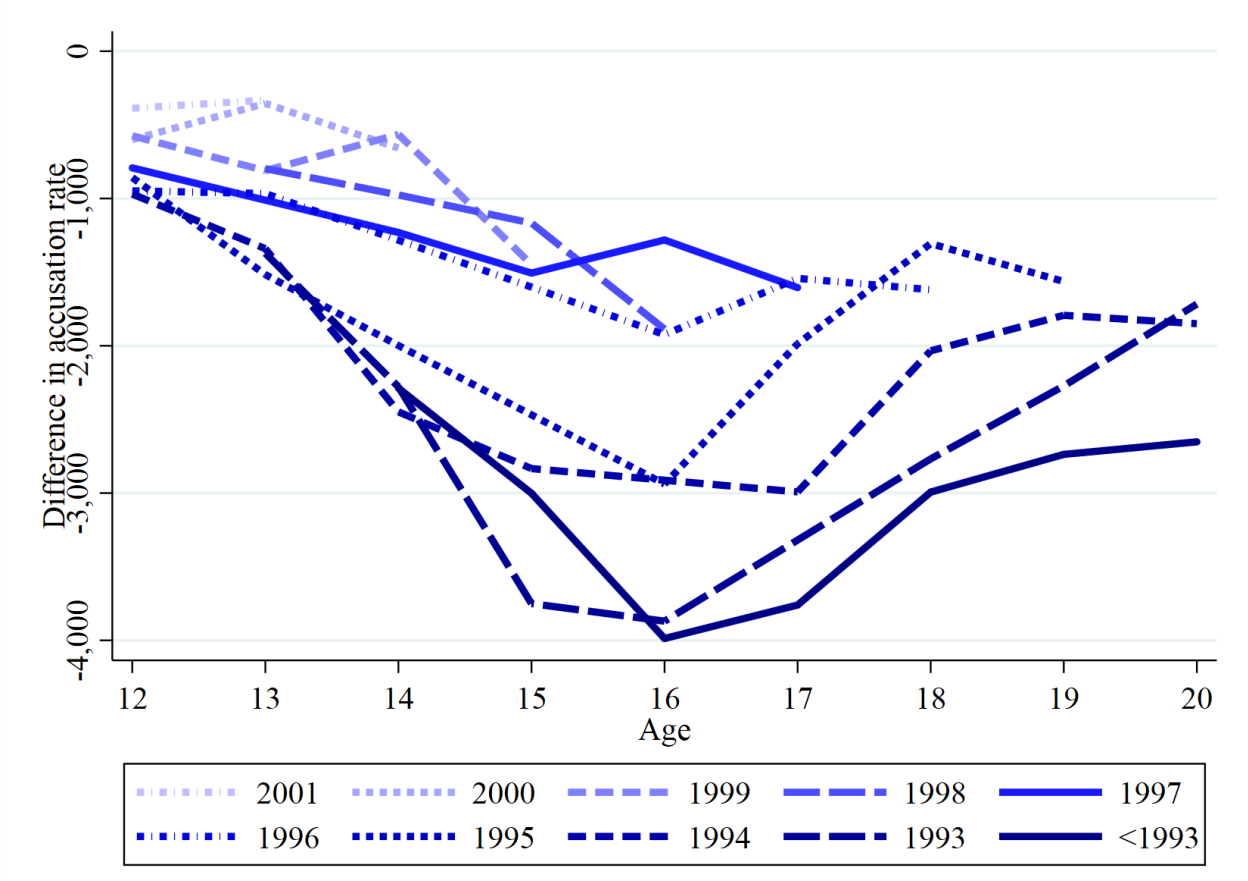


Notes: Authors' calculations from UCR data. The graph shows the annual conviction rate for Quebec and the rest of Canada at ages 14 and 19 as indicated. The vertical lines indicate the year in which cohorts are first exposed to eligibility: 2007 for 14 year olds and 2012 for 19 year olds.

We next provide supplementary figures to support the analysis in Figure 4 of the paper. In Figure 4, we show the difference between accusation rates in Quebec and the rest of Canada across the number of years a child of a particular birth cohort was exposed to eligibility for subsidized childcare in Quebec. Here we show the same graph across year of birth cohorts for both accusations and convictions. The number of years of eligibility by year of birth cohort can be referenced in Appendix Figure 1.

In Appendix Figure 4 there is a separate line for each birth cohort. The line traces the difference between the accusation rate in Quebec and the Rest of Canada, graphed against age. The differential declines cross-cohort analogous to the decline by exposure in Figure 4.

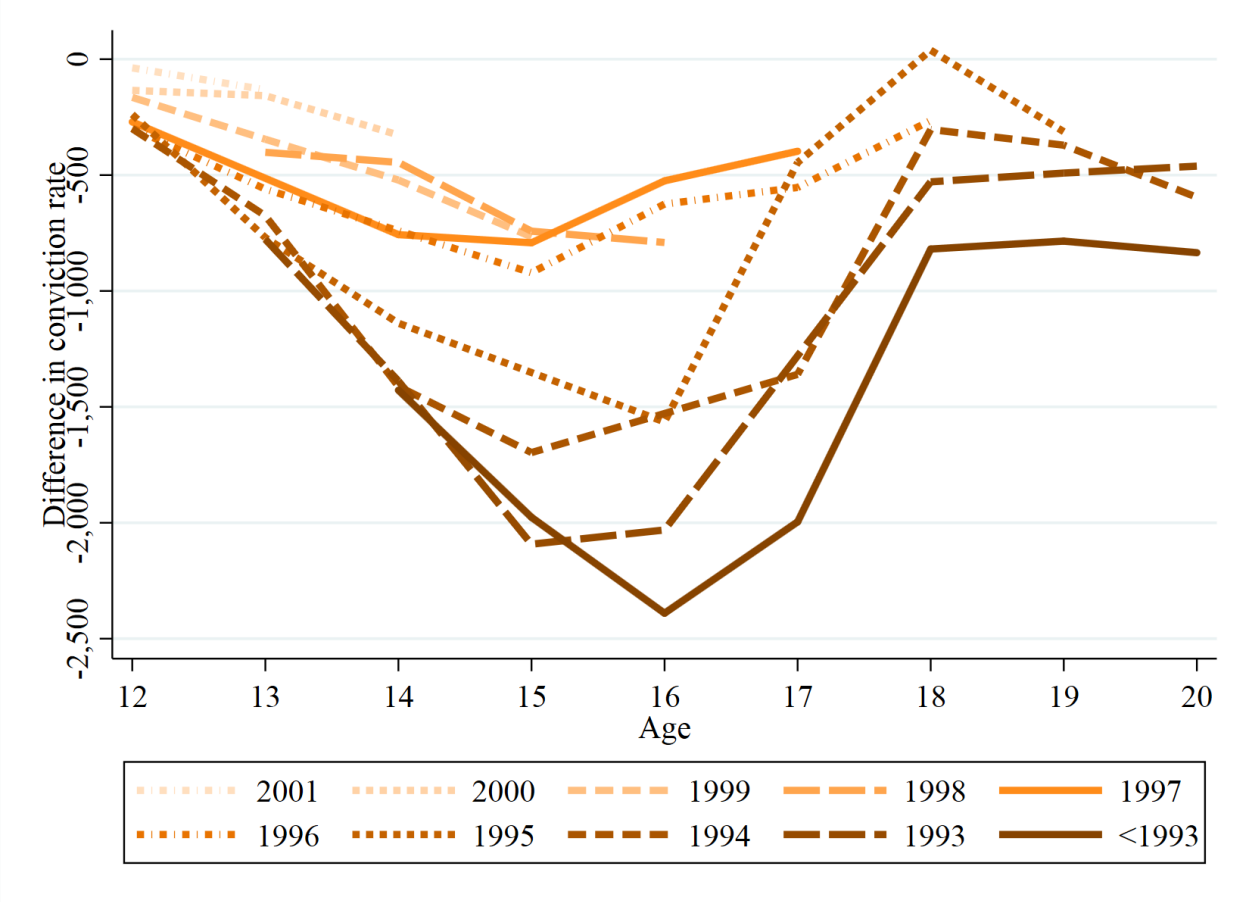
Appendix Figure 4: Quebec-Rest of Canada Differences in Accusation Rates by Year of Birth and Age



Notes: Authors' calculations from UCR data. Displayed is the difference in the annual accusation rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different birth cohort. Data are aggregated across four crime categories (crime against persons, crime against property, other criminal code violations, and drug crimes).

Appendix Figure 5 shows the same analysis as Appendix Figure 4, but now for conviction rates. Again, the differential declines cross-cohort analogous to the decline by exposure in Figure 4.

Appendix Figure 5: Quebec-Rest of Canada Differences in Conviction Rates by Year of Birth and Age



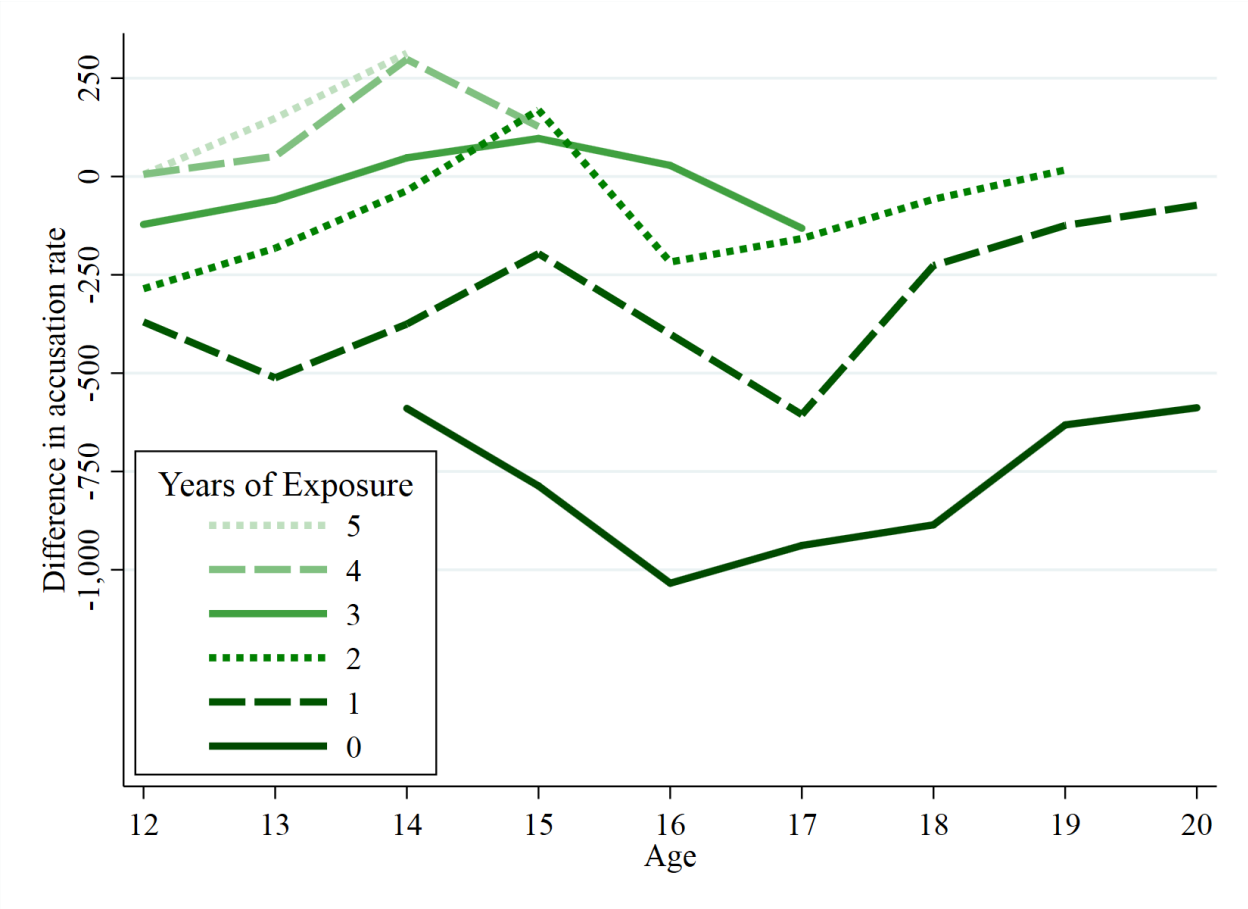
Notes: Authors' calculations from UCR data. Displayed is the difference in the annual conviction rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different birth cohort. Data are aggregated across four crime categories (crime against persons, crime against property, other criminal code violations, and drug crimes).

We next repeat the analysis presented in Figure 4 in the paper but disaggregate by type of crime. The results are reported in Appendix Figure 6 through Appendix Figure 9. In each

figure each line captures birth cohorts grouped by years of exposure according to Appendix

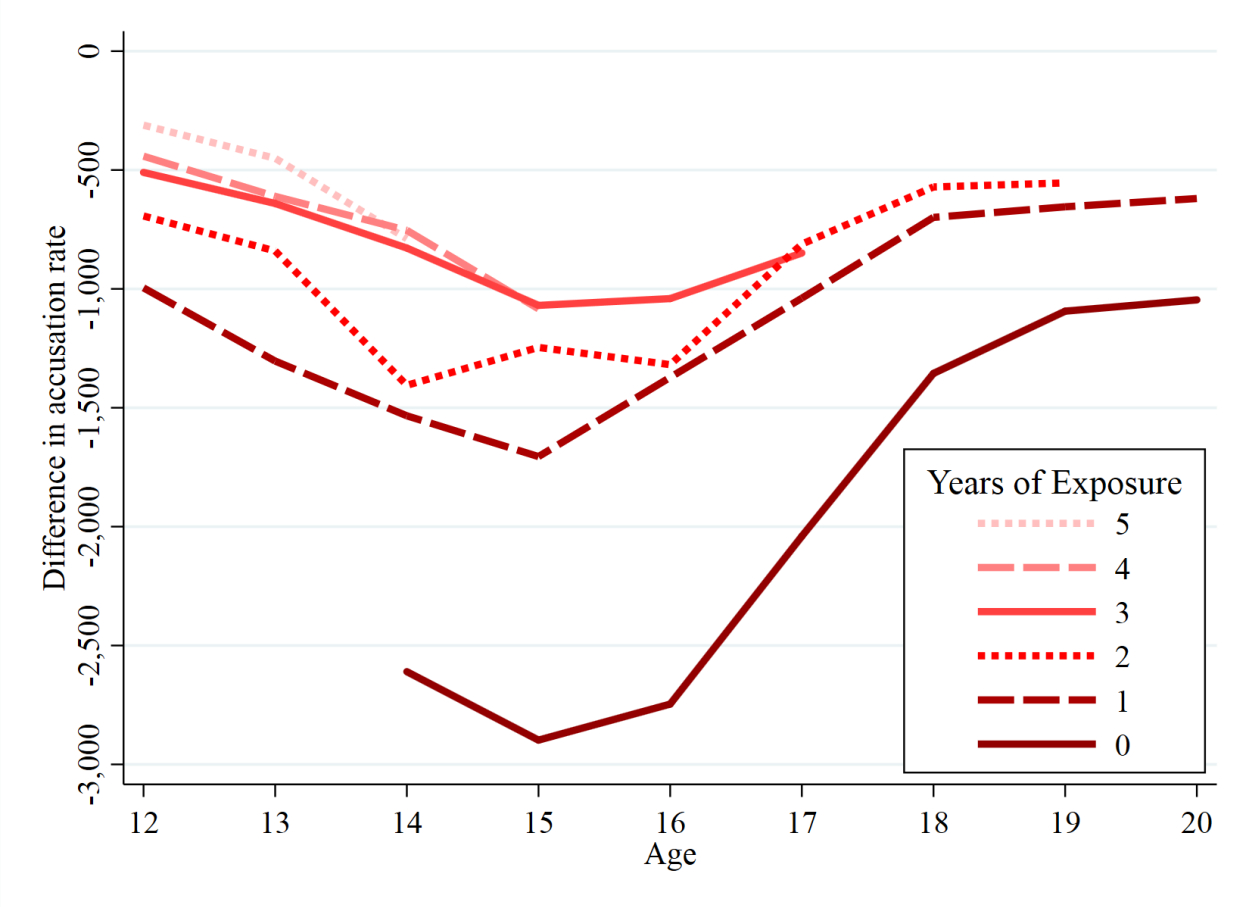
Figure 1. In each case the differential declines by exposure echoing the message of Figure 4.

Appendix Figure 6: Quebec-Rest of Canada differences in accusation rates by cohort—Crimes against persons



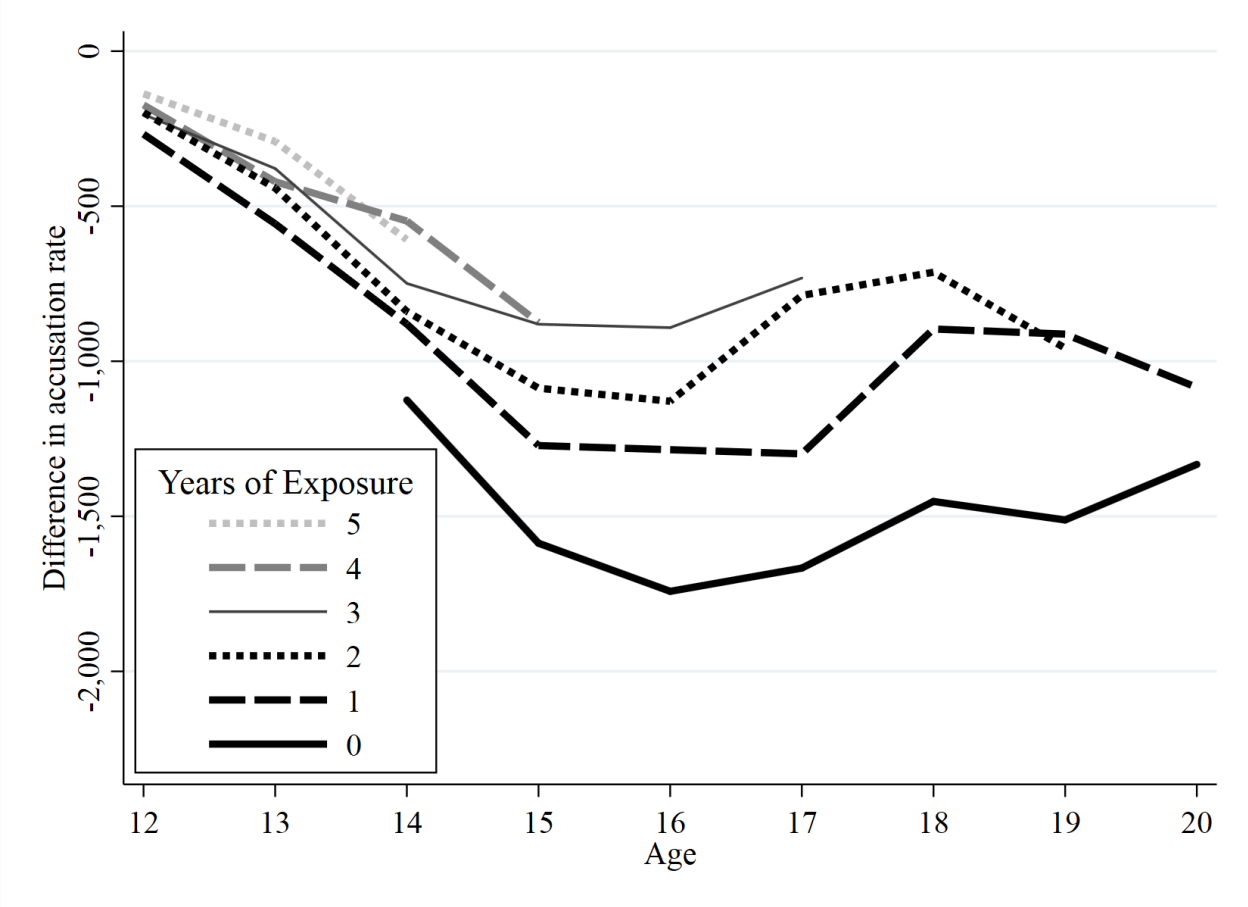
Notes: Authors' calculations from UCR data. Displayed is the difference in the annual accusation rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different set of birth cohorts, arranged by years of exposure to eligibility. This is for crimes against persons.

Appendix Figure 7: Quebec-Rest of Canada Differences in Accusation Rates by Birth Cohort—Crimes against property



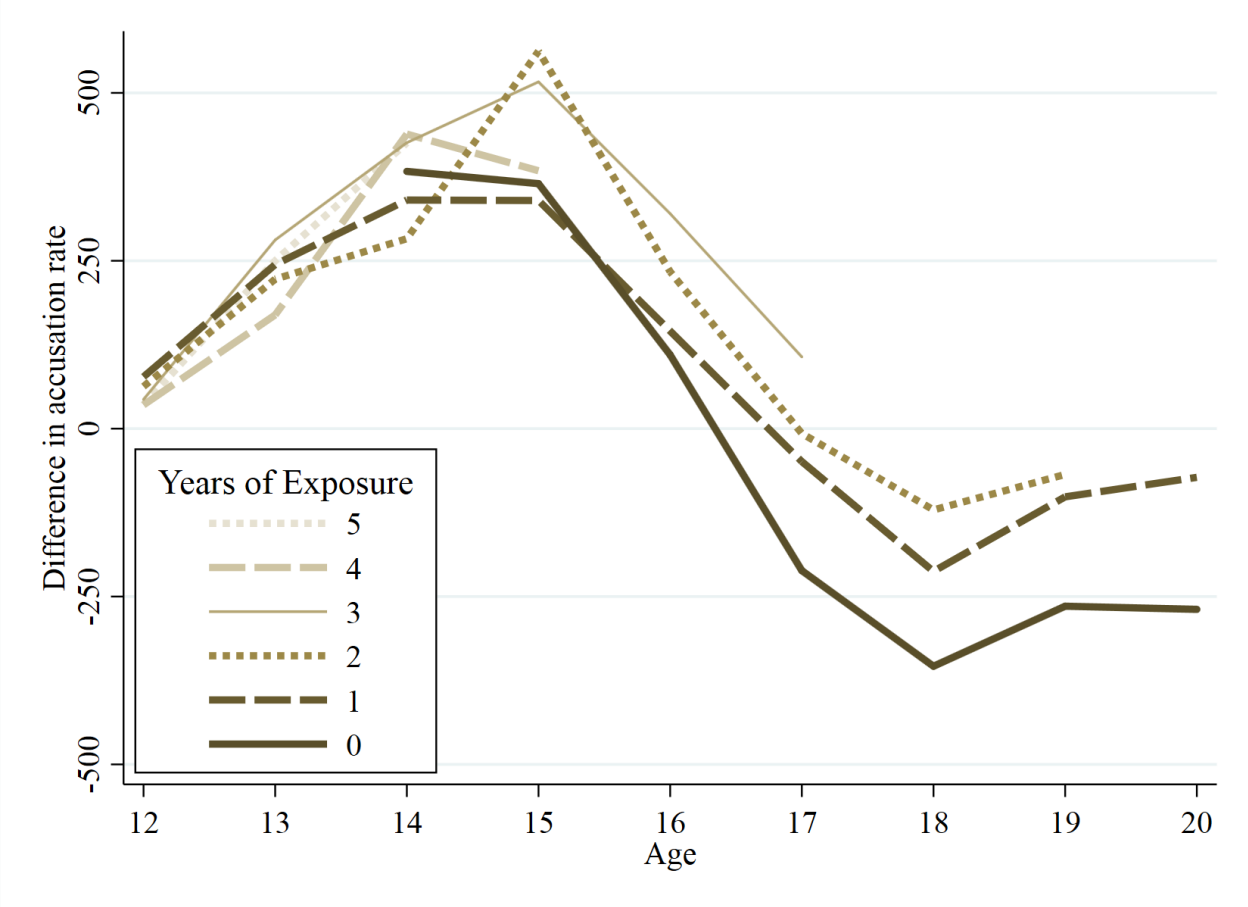
Notes: Authors' calculations from UCR data. Displayed is the difference in the annual accusation rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different set of birth cohorts, arranged by years of exposure to eligibility. This is for crimes against property.

Appendix Figure 8: Quebec-Rest of Canada Differences in Accusation Rates by Birth Cohort — Other criminal code violations



Notes: Authors' calculations from UCR data. Displayed is the difference in the annual accusation rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different set of birth cohorts, arranged by years of exposure to eligibility. This is for other criminal code violations.

Appendix Figure 9: Quebec-Rest of Canada Differences in Accusation Rates by Birth Cohort — Drug crimes



Notes: Authors' calculations from UCR data. Displayed is the difference in the annual accusation rate per 100,000 of population between Quebec and the Rest of Canada by age. Each line shows a different set of birth cohorts, arranged by years of exposure to eligibility. This is for drug crimes.

We next estimate regressions based on equation (2) in the paper, but augment the control variables with up to 2 leads and lags of the binary policy variable EXPOSURE. This allows us to see if our specification is picking up general trends in Quebec compared to the other provinces. The leads check to see if the policy had an impact on crime rates one or two years before the policy actually came into effect. We expect these coefficients to be close to zero, as these are like a placebo test—looking for an effect where there should be none. For the policy

lags, these variables check to see if the policy change had an additional impact after the policy was actually put in place; for example, if there was a slow transition. Here we would not be surprised to see some impact, as it may have taken some years of adjustment for the new child care policy eligibility to be reflected in actual child care use. Also we expect a "dosage effect" across the first birth cohorts exposed to the policy, which would also lead to lagged effects.

We code these leads and lags as $EXP_{ap(t+j)}$ for a given age *a*, province *p*, year *t*, and lag/lead *j*. This is displayed here as equation (A1).

(A1) $$CR_{aspt} = \tau + \sum_{j=-2}^{2} \rho_j EXP_{ap(t+j)} + \varphi PROV_p + \kappa YEAR_t + \omega SEX_s + \theta AGE_a + \epsilon_{aspt}$$

We use specification 3 (as noted in Table 5), which includes 2nd order interactions between province, age, sex, and year (except for YEAR*PROV) along with province-specific linear trends. We use only the 'all crime' dependent variable here.

The results for accusations are shown in Appendix Table 4. In the first column we report the estimate for the exposure variable from the third column (first row) of Table 5. In the second column we include one policy lead along with the main policy variable. The estimated coefficient is -15, which is small and not statistically significant. In the third column we include one policy lag along with the main policy variable. Here we find a statistically significant impact of 172 accusations per 100,000 in population on the lag, but the main policy effect remains, at 308, in the same range as the previous estimates. In the fourth column we include both the lead and lag of policy, along with the main policy effect. The results are similar. Finally, in the fifth column we repeat the analysis using 2-years lag and lead of policy, finding similar results.

We extend the analysis in Appendix Table 5 to cover leads and lags of convictions. The results are similar, with small and mostly insignificant impacts of policy leads and significant impacts of policy lags. This aligns with expectations, and reinforces the case that our policy variable is picking up the impact of eligibility for the Quebec childcare program on criminal accusations and convictions in teenage years.

Appendix Table 4: Impact of Exposure to the Quebec Family Plan on Crime Rates, ages 12-20, Lags and Leads, Accusations

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Accused**	**Accused**	**Accused**	**Accused**	**Accused**
Exposure	353	358	308	293	381
	[69]***	[62]***	[32]***	[28]***	[46]***
	(35)***	(43)***	(32)***	(41)***	(37)***
Lead 1		-15		35	
		[39]		[33]	
		(23)		(22)	
Lead 2					-20
					[34]
					(26)
Lag 1			172	181	
			[32]***	[37]***	
			(11)***	(6)***	
Lag 2					158
					[38]***
					(14)***
Specification					
3 controls	yes	yes	yes	yes	yes
Observations	5,632	5,632	5,632	5,632	5,632
R-squared	0.917	0.917	0.918	0.918	0.918

Notes: Authors' calculations from UCR data. The dependent variable is the number of accusations per 100,000 in population across four crime categories. Reported are regression coefficients and standard errors from specification 3 used in Table 5, augmented with leads and lags of the 'exposure' policy variable. Robust standard errors clustered on province and year are in square brackets. Standard errors using the Bester et al. (2011) method are reported in round brackets.

Appendix Table 5: Impact of Exposure to the Quebec Family Plan on Crime Rates, ages 12-20, Lags and Leads, Convictions

VARIABLES	(1) **Convicted**	(2) **Convicted**	(3) **Convicted**	(4) **Convicted**	(5) **Convicted**
Exposure	212	210	184	169	228
	[44]***	[41]***	[21]***	[16]***	[33]***
	(5)***	(4)***	(4)***	(5)***	(4)***
Lead 1		4		37	
		[25]		[21]*	
		(18)		(19)	
Lead 2					-12
					[24]
					(18)
Lag 1			108	118	
			[18]***	[22]***	
			(2)***	(5)***	
Lag 2					90
					[28]***
					(10)***
Specification 3 controls	yes	yes	yes	yes	yes
Observations	5,632	5,632	5,632	5,632	5,632
R-squared	0.902	0.902	0.903	0.903	0.903

Notes: Authors' calculations from UCR data. The dependent variable is the number of accusations per 100,000 in population across four crime categories. Reported are regression coefficients and standard errors from specification 3 used in Table 5, augmented with leads and lags of the 'exposure' policy variable. Robust standard errors clustered on province and year are in square brackets. Standard errors using the Bester et al. (2011) method are reported in round brackets.

Finally, we extend the crime analysis in Appendix Table 6 presenting estimates by gender. The estimates indicate larger absolute impacts on the crime rates for boys, particularly for other criminal code violations and drugs. In fact in our richest specification some of the estimates for girls are substantively smaller and lose some statistical significance. Therefore, by this metric the gender differences in the impacts of the Quebec program on crime rates line up with the gender differences in the impact of the program on non-cognitive development. That said, the mean rates for males are higher than for females. However, we argue that what is pertinent here is not the share of crimes committed by boys but whether there is more criminal activity when there is a reduction in population non-cognitive skills.

Appendix Table 6: Impact of Exposure to the Quebec Family Plan on Crime Rates by Gender, Ages 12-20

	All		Person		Property		Other CC		Drugs	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Accused										
(1)	378	638	376	673	581	542	378	880	176	458
	[52]***	[106]***	[50]***	[122]***	[87]***	[136]***	[65]***	[179]***	[41]***	[109]***
	(102)***	(256)*	(108)*	(233)	(120)**	(259)	(196)	(590)	(24)**	(85)*
(2)	326	844	427	862	639	1213	194	917	46	382
	[51]***	[75]***	[59]***	[89]***	[121]***	[164]***	[45]***	[102]***	[14]***	[51]***
	(55)***	(88)***	(97)**	(93)**	(136)**	(269)**	(43)**	(123)**	(36)	(105)**
(3)	158	539	200	396	282	906	126	622	26	233
	[74]**	[76]***	[71]**	[116]***	[194)	[190]***	[68]*	[112]***	[12]*	[55]***
	(40)**	(33)***	(70)	(80)**	(70)	(55)***	(56)	(183)**	(19)	(68)**
Convictions										
(1)	162	247	176	399	206	15	196	371	70	204
	[25]***	[52]***	[27]***	[78]***	[36]***	[10])	[34]***	[82]***	[18]***	[39]***
	(28)**	(64)**	(32)**	(91)*	(25)**	(41)	(60)**	(170)	(12)**	(59)*
(2)	148	490	174	465	278	625	93	519	47	353
	[26]***	[56]***	[32]***	[61]***	[49]***	[104]***	[36]***	[79]***	[10]***	[47]***
	(5)***	(47)***	(17)***	(15)***	(36)***	(50)***	(45)	(166)*	(18)	(73)**
(3)	81	337	97	235	133	541	78	393	15	181
	[35]**	[62]***	[44]*	[89]***	[66]*	[140]***	[52]	[90]***	[10]	[53]***
	(6)***	(9)***	(25)**	(32)***	(14)***	(56)***	(14)**	(29)***	(8)	(46)**

Notes: Authors' calculation from UCR data. In rows titled (1) are estimates from the difference in differences specification. In rows titled (2) are estimates that add all second order province, age, gender, year interactions, expect year*prov. In rows titled (3) are estimates that add province, year linear trend interactions. Reported is the coefficient on a dummy indicating exposure to eligibility. Significance at the 10, 5, and 1 percent levels is indicated with 1, 2, and 3 asterisks respectively, using p-values adjusted for multiple testing.