Extended maternity leave and children's long-term development*

Luc BEHAGHEL[1] and Maria Florencia PINTO[2]

[1]Paris School of Economics, INRAE. Email: luc.behaghel@psemail.eu
[2]CEDLAS, IIE-FCE, UNLP. Email: florencia.pinto@econo.unlp.edu.ar

February 24, 2023

Abstract

Countries around the world are increasingly expanding legal maternity leaves, with the dual objective of protecting mothers' jobs during childbirth recovery and enhancing child development. Using exhaustive census data, we find that a three-year paid leave in France had zero average effects on children's long-term schooling achievement, and no detectable impact heterogeneity. The lack of positive effects on children adds to the case against a policy that has strong adverse effects on mothers' careers.

*We thank two anonymous referees for useful comments and suggestions. We are also grateful to Inés Berniell, Jérémie Gignoux, Libertad Gonzalez, Julien Grenet, Xavier d'Haultfœuille, Sylvie Lambert, Karen Macours, Marco Mannacorda, Eric Maurin, Thomas Piketty, Maxime Tô, and seminar participants at PSE applied seminar, and at the joint J-PAL Europe - IPP seminar for valuable comments. Clotaire Boyer provided excellent research assistance. We thank INSEE and Centre Maurice Halbwachs for access to data. This work has received funding from CEPREMAP, the EUR project ANR-17-EURE-0001, and the *Chaire Politiques Éducatives et Mobilité Sociale*. All remaining errors are ours. Declarations of interest: none.

1 Introduction

How families adjust to the arrival of a newborn, reorganizing their schedule and budget, may have durable consequences on all members: parents, in particular by attenuating or amplifying the "child penalty" (Kleven *et al.*, 2019); but also children, whose early years have long-run consequences on health and education (Almond and Currie, 2011). Which policies can facilitate appropriate adjustments is therefore a pressing issue. Over the last few decades, most OECD countries have introduced paid maternity leaves of around four months into their legislation, reflecting the consensus that short maternity leaves help make work and family life more compatible, avoiding a complete withdrawal of mothers from the labor market and providing income support around childbirth (Figure 1a). Whether extended parental leaves enhance welfare is more debatable, and countries have followed different paths over the last 30 years: many have introduced such extended parental leaves, but two countries (Hungary and Sweden) have shortened their existing ones slightly; overall, there remains considerable heterogeneity, with a group of countries offering three-year leaves with job protection, while another, larger group, only enforces leaves of less than a year (Figure 1b).

This paper investigates the long-run impact of extended (three-year) paid parental leave provisions on children's schooling achievement. It studies the expansion in 1994 of a flagship French policy, the *Allocation parentale d'éducation* (Parental education allowance, APE) which was first introduced in 1985. We focus on children's outcomes, as the impact of extended parental leaves on mothers' careers has already been extensively studied. The impact on children is theoretically indeterminate: while the literature on early childhood and its long-term impacts suggests that substituting maternal care for other forms of daycare until age three may have large, long-lasting consequences, the effect is hard to sign a priori. On the one hand, early socialization with peers and interactions with more adults including professional educators may be more beneficial than staying at home, as argued by Canaan (2022). On the other hand, maternal care may benefit the child by providing one-to-one interactions with a trusted adult, critical to early development according to a long-standing strand of research in psychology (see the review in Fort *et al.*, 2019). These trade-offs likely vary depending on the parents' socioeconomic background, the quality of alternative formal and informal childcare available, and the length of the leave period.

The 1994 expansion of the APE allowed parents with two children to remain out of the labor market for up to three years after the birth of their second child, providing them with job protection and a 460-euro allowance (or a reduced-rate allowance in case of part-time work). The APE was taken up by around 20% of eligible mothers (and virtually no fathers), mostly full-time and for a duration of three years –a large increase compared to 16 weeks of paid maternal leave otherwise– inducing a substantial substitution between maternal care and other types of care until the age of three. The reform had a sharp eligibility cutoff, as it was only accessible to parents with a second child born after July 1, 1994.[1] We study the effects of this reform on children's long-term schooling achievement. [Using a difference-in-differences (DD) approach, we detect no] statistically or economically significant

[1]Families with three children or more had been eligible since the APE was introduced in 1985, while single-child families remained ineligible, and only had access to 16 weeks of paid maternity leave.

effect of the reform on various measures of schooling achievement (grade repetition or school dropout by the end of middle school, graduation from high school –with the associated national diploma, the *baccalauréat*– be it on time or up to two years late). With 95% confidence, we can rule out an increase higher than 0.017 in the probability of graduating on time from middle high school (a 3% increase from a baseline of 0.7). For high school graduation, the upper bound of the confidence interval is 0.037 (less than a 10% increase from the baseline of 0.4).[2] We explore effect heterogeneity by proxies of maternal and formal care quality, in particular the mother's education and the availability of public daycare managed by certified professionals (*crèches*), but do not find sizable differences. In addition, we show that the absence of detectable effects and of heterogeneity in impact is not driven by sample selectivity, measurement error, or the choice of the comparison group.

Our paper fills an important gap in a long-standing debate on the overall welfare effect of the APE after its adverse short- and long-term effects on mothers' careers were well documented (Lequien, 2012, Moschion, 2010, Piketty, 1998). Most closely related to our work, Canaan (2022) finds negative effects of the 1994 APE expansion on rough measures of the child's language skills at the end of kindergarten. We reject large positive effects on several long-run measures of educational attainment between ages 15 and 20. To the extent that educational attainment measures are representative of other margins of child development, this completes the case that this extension of parental leave did not reach its goal of improving the joint welfare

[2]See Table 3 and Section 4 for details. A 10% increase would remain modest compared to aggregate changes observed in that period: in five years, the probability of graduating on time from high school increased by about 0.15 or 40%.

of mothers and their children. Importantly from a methodological perspective, using exhaustive rolling census data allows us to get tight estimates. We show how supplementing the census data with ancillary survey data with complete family structures allows us to correct for the mismeasurement of birth rank. [Lastly, while using RD as an alternative empirical strategy confirms the DD results, we show how RD can be imprecise and even misleading despite large samples in the presence of large month-of-birth effects.]

More broadly, paper contributes to the international literature on the impact of maternal care on children's middle- and long-run educational outcomes. The few existing studies find all possible effects: negative (Baker and Milligan, 2010, Baker *et al.*, 2015, Canaan, 2022, Dustmann and Schönberg, 2012), zero (Danzer and Lavy, 2018), or positive (Carneiro *et al.*, 2015, Fort *et al.*, 2019, Rossin, 2011). A possible interpretation of these seemingly conflicting results is that maternal care has a comparative advantage at very early ages (up to 12 months), in line with the findings of Berger *et al.* (2005), Carneiro *et al.* (2015), and Rossin (2011), and when the mother is from an advantaged socioeconomic background (Danzer and Lavy, 2018).[3] Our results are partly consistent with this view: the zero relative effect of three years of maternal care may result from a positive effect until age one, offset by a negative effect from one to three. However, the absence of detectable heterogeneity by mother's education in our setting is less consistent with the idea that the quality of the mother's time is critical.

[3]In particular, Rossin-Slater (2018) concludes that extensions of existing paid leave policies have no impact on measures of child well-being, while the introduction of short paid leaves can improve children's outcomes.

The remainder of the paper is organized as follows. In Section 2, we present the APE and its reform in 1994. Section 3 introduces the data used, while Section 4 presents our main results, as well as the heterogeneity analysis and robustness checks. Section 5 provides a further discussion of our results, especially with regard to the broader literature on parental leaves and child development, and concludes.

2 The *Allocation Parentale d'Education*

The *Allocation Parentale d'Éducation* (APE) was created in 1985 to allow parents (either the mother or the father) to interrupt or decrease their labor market activity after giving birth to a third- or higher-order child. The allowance consisted of a lump-sum transfer received monthly until the child's third birthday, and it was aimed at compensating the time that parents take off from their jobs to educate their young children.[4] It was a job-protected parental leave, meaning that the worker could go back to her previous job and receive a salary as high as her salary before the interruption plus the average wage increase in the company over the period (Lequien, 2012). In order to be eligible for the benefit, individuals had to have been working (or receiving unemployment benefits) for at least two years out of the ten years prior to the birth of the third child.

In July 1994, the French government extended the APE to families having a second child. Specifically, parents whose second child was born on or after July 1, 1994, were eligible for the allowance, again until the child turned three years old.

[4]The initial duration of the allowance is one year, and it can be extended twice until the child turns three years old (Gosset-Connan, 2004).

The implementing law for this reform (*Loi Famille*) was passed on July 25, 1994, and was not announced in advance, thus precluding potential anticipation effects among parents who could have timed the birth of their second child to become eligible (Lequien, 2012). Moreover, even though the law also affected other family policies, the July 1 cutoff was specific to the reform of the APE (Canaan, 2022). The requirements to benefit from the APE were somewhat more stringent for parents of a second child, who had to have been working for at least two out of the last five years. The amount of the monthly transfer varied depending on whether the parent took it at the full rate (i.e., stopped working) or at a reduced rate (i.e., continued working part-time). The benefit was available for either of the two parents (and both parents could even choose to take the APE at a reduced rate at the same time), but, in practice, 98 percent of APE beneficiaries were mothers (Moschion, 2010, Piketty, 2005). The full-rate benefit for parents who decided to stop working came to about 460 euros per month (about half the minimum wage in 1997), while the reduced-rate benefit was 300 euros if they worked at most 50 percent of the time, and 230 euros if they worked between 50 and 80 percent of the time.

The total number of APE beneficiaries increased sharply after this extension. While there were roughly 175,000 APE beneficiaries in 1994, the number had tripled by 1997, reaching about 500,000 (see Figure 2), and 60 percent of them were receiving the allowance for their second child (Piketty, 2003). About three-quarters of all beneficiaries withdrew completely from the labor market and received the full-rate allowance (Jacquot, 2000, Piketty, 2005), while one quarter stayed in the labor force and received the reduced-rate allowance.[5]

By making it easier for mothers to withdraw from the labor market, the introduction and expansion of the APE presumably increased the share of children below age three who were taken care of by their parents during week days. Daniel and Ruault (2004) report that this was by far the commonest type of childcare in 2002 in France (64%). Alternative modes of day care include childminders, known as *assistantes maternelles* (18%), formal daycare facilities known as *crèches* (8%), grandparents (4%), and other forms of arrangements (6%). The main goal of this paper is to study the impact on children's education outcomes of reinforcing this predominance of maternal care until age three.

Some authors have studied the effects of the APE on various aspects of the labor market. Using information from the French Labor Force Survey (*Enquête Emploi*), Piketty (2005) studies the direct effects of this measure on mothers' labor supply and employment. He finds that labor force participation of mothers of two children (with one aged under three, thus APE eligible) decreased by 15 to 23 percentage points as a consequence of the APE extension, a 22-33 percent drop compared to the initial level of 69 percent.[6] Moreover, the incentives to withdraw from the labor market did not affect all mothers in the same way: while more educated women (high school graduates or higher) reduced their employment rate by 16 percent, the drop

[5]The type of women that chose each regime were different, though. Women from higher socioeconomic groups (determined by the income of their spouse) were more likely to choose the APE at a reduced rate than those from more disadvantaged backgrounds (Afsa, 1998).

[6]Canaan (2022) complements the analysis of the labor market effects of the APE by studying labor supply and working hours of fathers. While she finds no evidence of a change in their labor force participation, she does find an increase in fathers' hours of work as a response to the APE, which thereby induces a higher gender specialization within the household.

among lower qualified women reached 47 percent. Importantly, this effect does not appear to have been long lasting: once the youngest child turned three years old (and the parents stopped receiving the APE), mothers' labor force participation returned to its pre-APE level. Nevertheless, Lequien (2012) shows evidence of a long-lasting impact on earnings: using administrative data on private sector workers, he found a negative impact (although not always significant) on mothers' wages after their return to work that seemed to persist even ten years after the birth of the second child. It is useful to place these effects in their broader context. Piketty (1998) documents the heterogeneity of the labor force participation behavior of women in France during the 1990s. In 1994, about 35% of mothers with three children were in employment, compared to 65% (resp. 75%) for mothers of two (resp. one) children. Labor force participation was increasing slowly, by about five percentage points in all three groups over the previous decade. In that context, the drop in the labor force participation rate by about 20 percentage points for mothers with two children was a major break in the trend. It illustrates the sensitivity of mothers to financial incentives: using the APE reform along with several other quasi-experiments, Piketty estimates the (extensive margin) labor supply elasticity of women to be in the 0.6 - 1 range, compared to 0 - 0.1 for men.

By decreasing the costs associated with childbearing and reducing the perceived benefits of returning to work, the APE may also have affected fertility decisions. However, neither Piketty (2005) nor Canaan (2022) find evidence of a fertility response (measured either by birth spacing or by the total number of children).

3 Data

Our analysis relies on data from the French Population Census, carried out by the National Institute of Statistics and Economic Studies (Institut national de la statistique et des études économiques - INSEE). It gathers demographic information on the whole population, and education and labor market-related information for individuals aged 15 and above. Importantly for our analysis, it reports the exact date of birth of every individual in the household, allowing the identification of those born before and after the APE reform.[7] It also reports the birth order of all children in the household, allowing us to distinguish children who are second in their sibship from other children, and construct our APE-eligibility indicator. Note, however, that if some children have left the household, the sample of remaining children may be selected, and their birth rank may be underestimated. We postpone this important discussion to Section 4.3, where we show evidence against sample selection bias and show that measurement in children's birth order results in an attenuation bias that can be corrected using ancillary data sources.

In 2004, in order to produce local-level information at a higher frequency and spread the cost of implementation over time, the INSEE switched from the traditional exhaustive census (usually every eight or nine years) to a "rolling" census, based on annual census surveys (known as *Enquêtes Annuelles de Recensement* - EARs) that enumerate about one fifth of the population every year.[8] Of particular interest to our

[7]Enumeration takes place over a period of four to five weeks starting in the third week of January (Desplanques and Rogers, 2008).

[8]The sampling is as follows. Small communes (of fewer than 10,000 inhabitants) are divided into five rotating groups. Each year, households are exhaustively enumerated in all municipalities corresponding to a given rotation group. After a five-year period, all five groups have been covered

study is the fact that the rolling structure of the census allows us to observe several cohorts of individuals at a given age, as opposed to a traditional census at one point in time where only one cohort can be observed at a given age. For instance, we can observe the proportion of youth that graduate from high school on time for adjacent cohorts observed in different EARs – something that would be impossible with the traditional census structure.

Since the census gathers information on education for individuals aged 15 years old and above, we construct several measures of educational attainment that are relevant to that age range, based on the question *"What is the highest diploma obtained?"*. The first outcome we measure is having obtained the middle high school certificate (*Diplôme National du Brevet* or BEPC). In the French educational system, students take an examination to obtain this certificate at the end of middle high school (9th grade). Though taking the exam is not a strong indicator of performance in itself, it is a proxy for grade-on-time: all students that were not held back should have taken this examination in the year they turn 15.[9] We look at the probability of having obtained this certificate either on time or with a two-year delay. In our sample, 69 percent of children obtained the middle high school certificate on time according to their age, and this proportion increases to 84 percent when allowing for

---

and the small communes have been exhaustively enumerated. Large municipalities (with 10,000 inhabitants or more) are visited every year but enumeration is not exhaustive. Households in large municipalities are divided into five rotation groups, and each year, 8% of the households from each community (corresponding to a given rotation group) are enumerated. After five years, 40% of dwellings in large municipalities are enumerated.

[9]French children start primary school the year in which they turn six years old (an academic cohort corresponds to all children born between January and December of a given year). Students would be in a higher or lower grade when they turn 15 as a consequence of grade retention or grade skipping.

a delay.[10] In addition, we also measure school dropout by the end of middle school by looking at whether children attend any educational institution at age 16.

After middle high school, students move up to senior high school where they can follow either a general academic track (three years long) or a vocational track (four years long). At the end of senior high school, they take the *baccalauréat* examination, which involves several oral and written examinations and takes place over several days. This diploma is one of the most important schooling outcomes in France: it gives recipients the right to attend a university (it is technically the first university degree) and has significant consequences in terms of success in the labor market (Maurin and McNally, 2008). For the 1994 cohort, this will happen in June 2012 (if tracked into general education) and it will be observed in the 2013, 2014, or 2015 EAR depending on whether students pass it on time, with a one-year or a with two-year delay, respectively (those in the vocational track will take this examination for the first time in 2014 instead of 2013). We study the probability of obtaining the *baccalauréat* (referred to as "(senior) high school graduation"), either on time or with a two-year delay.[11] In our sample, 38 percent of youth report having obtained this diploma on time, and 70 percent do so when we allow for a two-year delay. The share of youth graduating from high school with the *baccalauréat* has been steadily increasing in France over the years, from 62% in 1998 to more than 80% in 2018, due

[10] Students born in 1994 will prepare for this certificate during the 2008-2009 academic year, and will take it for the first time in June 2009 (see Table A1). Given that the census enumeration period takes place at the beginning of the calendar year, a student who earned the certificate on time will only be observed in the 2010 EAR, and someone who earned it one (two) year(s) later will be observed in 2011 (2012).

[11] An interesting and perhaps more informative measure of educational success is having obtained the *Baccalauréat general*. Unfortunately, the census only reports this category separately until 2014, which makes it useless for our analysis.

to both demand and supply factors. On the supply side, a 2008 educational reform affecting the cohorts born after 1990 increased access for students in the vocational track to the *baccalauréat professionnel*. The reform has contributed to the upward trend shown in Figure 4. It remains orthogonal to the APE reform, however. The outcomes observed and the rounds of EARs used for the other cohorts are described further in Table A1.[12]

In addition to these key outcome variables, the census data provide limited background information on the household at the time of the census. We create a few control variables for robustness checks. To avoid controls that could be endogenous to the APE reform ("bad controls"), we restrict ourselves to the mother's education –a proxy for socioeconomic status that is in most cases predetermined at the time of the child's birth– and to geographical controls (dummies for the department of residence and an indicator for urban/rural place of residence). A potential concern is that the household's place of residence could have been affected by the APE reform; however, French national exams (*brevet* and *baccalauréat*) are graded separately by regions (groups of departments), making comparisons across regions potentially fragile. Checking the robustness of estimates to the inclusion of geographical controls is therefore important.

[12] At the time of data collection for this paper, the latest available round of the EAR was 2017.

4 Results

4.1 Main results

The 1994 reform of the APE provides quasi-experimental variation to identify the impact of access to extended paid parental leave by comparing families whose second child was born after July 1, 1994, to families whose second child was born prior to this date. However, this requires us to account for potential confounding effects associated with the birth date of the child (henceforth, birth date effects).[13] [In order to neutralize birth date effects, we follow a difference-in-differences (DD) approach under the assumption that birth date effects would have evolved in the same way in the absence of reform for children in families impacted than for a comparable group of unaffected families.[14] [15] First-born children emerge as a natural control group.[16] Under certain assumptions, that we discuss below, the impact of the APE reform can be identified from the β_2 coefficient in the following equation:]

$$y_{iq} = \beta_0 + \beta_1 Rank2_i + \beta_2 \cdot Rank2_i \cdot After + \lambda_q + X_i\theta + \varepsilon_{iq} \tag{1}$$

[13] [Birth date effects combine the impact of a variety of factors: changing family background driven by fertility trends, changes in the schooling environment and the educational policies affecting successive cohorts of students, and differences in maturity within a cohort associated with differences in ages at school entry, as documented by Grenet (2010) in the French context.]

[14] [This strategy has been used by previous studies of the 1994 APE reform (Lequien, 2012, Piketty, 2003, 2005) using families with one child (or three children) to net out date of birth effects.]

[15] [In Section 4.3, we conduct robustness checks and provide qualitatively similar results using a regression discontinuity (RD) design approach.]

[16] [An alternative would be to use third-born children as a comparison group, but this is not appealing for statistical precision as it yields a small sample. First-born children yield a much larger sample.]

where y_{iq} measures the educational outcome of individual i born in quarter-year q,[17] $After$ is a dummy equal to one if the individual was born on or after July 1, 1994, i.e., after the APE reform, and $Rank2_i$ is a binary variable equal to one if the individual is the second-born child in the sibship and zero if she is the first in a single-child family. In the main specification, the control group thus contains only single children and the included cohorts are 1992-1996.[18] [We also control for quarter-year-of birth effects λ_q and a vector of controls X_i that includes an indicator variable that equals one if the mother of the child has graduated from high school, dummies for the department of residence and an indicator for urban/rural place of residence.] Note that the DD estimates net out week-of-birth effects twice: using children in single-child families as a control, and averaging the outcomes over two calendar years before and after the reform. The validity of this strategy relies on the assumption of parallel trends between groups in the absence of the reform, which we discuss in detail in Section 4.3.

[The main results are shown in Table 1, that displays the estimated effects for the probability of obtaining the middle school certificate, either on time or with a delay of up to 2 years, which can be considered as proxies for being in grade on time, or with some delay. The second panel refers to the probability of graduating from senior high school, also on time or up to two years later. The DD coefficients are not statistically

[17] We measure five different outcomes: having received the middle school certificate on time, or up to two years late (as a proxy for grade repetition), having dropped out of school by the end of middle high school, and high school graduation, also on time and up to two years later. Note that all these outcomes are measured on different and independent samples, depending on the year in which each cohort reaches the relevant age, as explained in Section 3.

[18] In Section 4.3 we discuss the choice of comparison group and test the robustness of the results to the different choices.

different from zero.] Children of rank two born after the reform are neither more nor less likely to obtain the middle high school certificate than children of rank one, whether in the year in which they should obtain it according to their age or with a delay of up to two years. Controlling for geographical characteristics, as well as maternal educational level, does not change the results. We also find no evidence that the APE changed the likelihood of being out of school by age 16. Importantly, the 95% confidence intervals are tightly centered around zero. We further discuss the size of the effect that can be excluded after controlling for attenuation bias in Section 4.3. Turning to the effects on senior high school graduation, we find that children of rank 2 born after July 1, 1994 were not significantly more likely to graduate on time. Similarly, we find no significant effect on high school graduation up to two years late. The results are robust to the inclusion of geographic characteristics and maternal education as controls.

Overall, the results from the DD analysis show that the APE extension to second-born children did not significantly affect the probability of grade retention, school dropout, or graduation from senior high school. [These are tight zero results as the large sample sizes from census data ensure a high precision, and they are robust to different specifications.]

4.2 Heterogeneity analysis

After showing no detectable effect on long-term educational achievement for children whose mothers had access to leaves three years long on average, we examine heterogeneity in responses. In particular, the impacts may differ depending on the relative quality of maternal time compared to the alternative mode of childcare. One may expect children from relatively higher socioeconomic groups to benefit from spending more time with their mother since their home learning environments are of higher quality relative to children from less affluent families (due to high human capital and high income). With varying settings and identification strategies, the literature has mainly documented negative impacts of substituting maternal time by institutional time for children from more affluent families (Baker *et al.*, 2008, Fort *et al.*, 2019), and positive effects for children from more disadvantaged backgrounds (Drange and Havnes, 2018, Felfe and Lalive, 2018, Kottelenberg and Lehrer, 2017). Meanwhile, Canaan (2022) does not find strong evidence of effect heterogeneity by socioeconomic status.

We run our DD regression on children from high- versus low-educated mothers. Although the mother's education is measured at the time of the survey, it is a good indicator of socioeconomic status at the time of birth as most women had completed their education before their first pregnancy. We do not find heterogeneous impacts of APE eligibility between children whose mothers have finished high school and children whose mothers have not (see Table 2 and Figure 3). For each of the outcomes, we cannot reject the null hypothesis that the estimated APE impact is the same across groups. As an exception, however, note the positive effect on HS graduation on time for children with higher-educated mothers. This suggests a positive impact of maternal time when mothers are more educated (note that if the mother's time had the same effect irrespective of the mother's education, we would expect a smaller coefficient on children of HS educated mothers, as those mothers were less likely to take up parental leave (Piketty, 2005). However, one should remain cautious, as it is unclear why this positive effect would only materialize in one outcome measure out of five. The statistically significant estimate may also be due to a sampling error. We also test whether the impact differs according to whether children live in urban or rural areas, which may capture other dimensions of socioeconomic status, but the estimations yield the same results.

We next consider heterogeneous impacts of the APE reform by gender. The evidence in the literature is not conclusive. Some studies have shown that girls tend to be more harmed by early daycare attendance than boys and benefit more from one-to-one interactions with a trusted adult than boys (Fort *et al.*, 2019), while others have found that most of the negative impacts of attending daycare in the first years are concentrated on boys (Baker *et al.*, 2015), and the positive ones on girls (Drange and Havnes, 2018, Felfe and Lalive, 2018). Our results suggest that the corresponding coefficients for both boys and girls are small, and we cannot reject the hypothesis that they are equal to zero.

Finally, we take into account the potential alternative mode of care. Danzer *et al.* (2017) find that, in Austria, the extension of parental leaves had a positive effect on a set of children's health and human capital indicators only when the reform induced a replacement of informal care (mostly by grandparents or other relatives) by maternal care, but not for children who switched from formal care to maternal care. As we do not have information on childcare use for these cohorts in their early years, we construct a proxy measure. We calculate the availability of daycare at the department level as the ratio of the number of slots in formal daycare facilities managed by certified providers *crèches* to the population aged 0 to 4.[19] Each department is considered as having a "high" availability of *crèches* if it is located in the top 25 percent of the distribution, and a "low" availability otherwise, and children are classified according to their department of birth. No clear heterogeneity pattern emerges: the impacts of APE eligibility on children's long-term educational attainment are the same regardless of whether there was a large offer of formal care in their department of birth or not.[20]

4.3 Robustness checks

Despite our use of large census data sets, our estimates show no detectable effect of access to extended paid parental leave on all measures of children's schooling achievement, and no heterogeneity in impact with respect to the mother's education or the availability of formal daycare at the department-of-birth level. In this section, we perform several robustness checks. We first discuss the parallel trend assumption in our DD analysis. We then check that the zero effects are not driven by sample selectivity or measurement error in children's birth order. [We also provide additional evidence of the effect of APE using an alternative identification strategy.] Last, we check that our results are not driven by the choice of comparison group.

[19]France is divided into about 100 departments. The *crèches* are centers that provide formal daycare services for children aged 0 to two years old. The information on the number of slots available for each department in 2004 comes from Bailleau (2010). The population numbers come from the INSEE's population estimates by department, gender and five-year age categories. We use the total population in the 0 to 4 age range as denominator as there is no available information on finer age breakdowns with such a low geographical disaggregation.

[20]The likely counterfactual is not the only potential source of heterogeneity in this case, as there could be treatment effect heterogeneity due to other differences associated with the location. At all events, we cannot reject the null hypothesis that the estimated APE impact is the same across the two groups.

Parallel trends in the DD analysis

The validity of the DD estimation relies on the assumption of parallel trends between groups in the absence of the reform. The observation of several birth cohorts before 1994 allows a suggestive test of this assumption. Figure 4 plots different average educational outcomes for first- and second-born children. In each graph, the vertical red line in July 1994 represents the time of the APE extension. Each dot represents the average outcome for children born in 12-month windows starting in July of each year, i.e., born between July 1992 and June 1993, July 1993 and June 1994, and so on.[21] The two groups present different average outcomes before the policy change, but, for most of the outcomes, follow similar trends.[22] However, Figure 4 suggests differential trends for some of the outcomes prior to the APE extension, especially when cohorts far apart from 1994 are considered.[23]

This is a priori an important caveat that requires further attention. Note that rejection is not systematic across outcomes and cohorts. In particular, the relative outcomes of second-born children improve for "Middle high school certificate on time" for cohorts born between July 90 and June 91 and July 91 and June 92, and for "Senior high school graduation up to two years late" for cohorts born between July 91 and

[21]Since the APE extension took place in July 1, and each dot covers the twelve months of a calendar year, by doing this we average out week-of-birth effects due to age at school entry.

[22]Black *et al.* (2005, 2007, 2017) have documented a relative advantage of first-born children in terms of IQ, non-cognitive abilities, educational attainment, and earnings.

[23]To show statistical significance, Figure A1 in the online Appendix displays the gap between the two groups and the corresponding 95% confidence intervals, normalized to zero for those born in the last period before the reform (June 1993 - July 1994). While we do not reject the possibility that trends evolve in parallel for the two cohorts (July 1992 to June 1994) preceding the APE reform, we do reject it for some of the outcomes when a longer period beforehand is considered (the results of a statistical test of this assumption are shown in Table A2).

June 92 and July 92 and June 93 (Figure 4). Given the five-year rotating sampling scheme of the census, these relative changes are in fact measured on the same 1/5 subsample of the population, five years apart.[24] The break in parallel trends is thus specific to one subsample, and can therefore be interpreted as a "sampling error" (for lack of a better understanding of its causes, which may be due to composition effects or idiosyncratic shocks that our data do not identify). Such a sampling error could be accounted for in the inference, as proposed by Rambachan and Roth (2020).[25] We anticipate that it would confirm the absence of a statistically significant effect, but acknowledge that it would also most likely lead to wider confidence intervals (depending on the exact assumptions made). As a simple fix, we dropped cohorts born before 1992 for the DD analysis, [but we also show in this section] that the zero-effect result holds with a regression discontinuity approach that does not rely on a parallel trend assumption. The early break in parallel trends thus does not invalidate the conclusions of the analysis.

Moreover, for the parallel trends assumption to hold, we also need no other treatment to take place simultaneously with the reform impacting first- and second-born children differently. There is a well-documented discontinuity in schooling achievement associated with the date of birth, in France as in several other countries: children entering school one year older repeat less frequently and are less likely to follow a vocational track, resulting in better schooling achievement overall. However, this

[24] For instance, for the 1992 birth cohort, "Middle high school certificate on time" and "Senior high school graduation up to two years late" are observed in the 2008 and 2013 *Enquêtes annuelles de recensement*, respectively, which cover the same individuals in the population, up to movers.

[25] We became aware of this robust inference approach after our temporary access to the data was interrupted.

discontinuity occurs at the January 1 cutoff (due to differences in age at school entry), and not on July 1, as these children enter school at about the same age (Grenet, 2010).[26] No other reform was introduced with the July 1, 1994 birth date as an eligibility cutoff. And more importantly, there is no obvious reason why the mechanisms that presumably drive date-of-birth effects –in particular, the effect of age at school entry– would differ for children of different birth ranks.

Attrition and measurement of birth order

Census data only list members of the sibship residing with their parents. As children start leaving the household, this may create two difficulties for our empirical analysis. First, it may generate sample selection bias. Second, it induces measurement error on birth order, as birth order is not reported directly and needs to be inferred from the age of children present in the household at the time of the census. In this section, we use additional data sources to show that sample selection bias is unlikely, and to show how measurement error on birth rank creates an attenuation bias that can be corrected to recover the effect of APE eligibility.

Consider sample selectivity first. It arises because the census only reports information on youth still co-residing in their parents' household. This may induce sample selection bias in our DD estimation if attrition is selective (correlated with education outcomes) and evolves differently across birth cohorts between first-born and second-born children. To explore the latter, we measure attrition in the census

[26] This is due to the fact that children in France enter pre-school (*école maternelle*, which is not mandatory but followed by the vast majority) during the calendar year in which they turn three, which implies that children born on December 31 enter school at about 2 2/3 years old, while those born on January 1 are 3 2/3 years old.

data by comparing the number of youth observed in their parents' households at the ages of 15 to 20, for different birth cohorts, to the total size of these cohorts. We measure cohort sizes using the 1999 Population Census (RP99): children born between 1992 and 1996 were two to seven years old in 1999 and thus too young to have left their parents' household.[27] The analysis, presented in detail in Appendix C.1, shows that our working sample represents between 83 and 86 percent of each cohort. However, this proportion does not change in any systematic way for individuals born before and after July 1994. In other words, even though we do have a selected sample, such sample selection is not at odds with a causal interpretation of the DD estimates since it does not evolve over the period of the estimation.

Similarly, birth rank misclassification arises because the rank of a child is inferred from the household composition at enumeration: it does not match the true rank if an older sibling has left the household. We use an additional data source (the *Enquête Famille*) that covers complete families and reports whether children co-reside with their parents. As detailed in online Appendix C.2, this allows us to quantify the degree of misclassification in the census. Importantly, we do not find evidence that the probability of birth rank misclassification changed significantly before and after 1994. In online Appendix C.3 we show that, under such stable misclassification, the naïve DD estimators suffer from attenuation bias, but that re-inflating them according to the probabilities of misclassification allows us to recover unbiased estimators.

[27]To avoid double counting in cases where a family splits and the child lives some time with each of her parents (*garde alternée*), the child is enumerated in the household where she spends the larger part of the year. If she spends equal time with her mother and her father, then she is enumerated in her place of residence during the first day of census enumeration.

The corrected estimates are displayed in Table 3. After re-inflation, we reject the hypothesis that the APE increased the probability of holding the middle school certificate on time by more than 1.7 percentage points, or by more than 1.8 percentage points if a two-year delay is allowed. Also, we rule out a decrease in school dropout at age 16 by more than 0.2 percentage points. For the case of senior high school graduation, with a 95% level of confidence we can reject the hypothesis that it increased more than 3.7 percentage points. These upper bounds can be compared to the 20 percentage-point gap in high school graduation rate between children of high-educated vs. low-educated mothers, or the 15 percentage-point increase in high school graduation rate over this five-year period. It should be noted that we estimate the reduced-form effect of being exposed to the reform, not the effect of staying home with one's parents. Due to lack of data, we do not measure exactly the "first-stage" impact of the reform on the probability of staying home with one's parents. Furthermore, the exclusion restriction needed for an instrumental variable estimation may not be valid, as the reform affects aspects other than daycare, including family income. However, as an indication, if that restriction were to hold, given a 25% take-up of the parental leave, one would have to inflate estimates by a factor of the order of four to recover local average treatment effects of maternal care.

Regression discontinuity approach

Figure 5 presents a standard RD graph. The horizontal axis measures the number of days normalized to 0 for July 1, 1994, and each dot represents the average outcome for individuals grouped in two-week bins. No jump can be noticed in the probability of holding the middle school certificate on time or up to two years late, in the likelihood of having dropped out of school by age 16, or in high school graduation. The one possible exception concerns high school graduation up to two years late: there, we can see quite a large and somewhat erratic week-of-birth effect. To a lesser extent, these fluctuations are apparent in other outcomes.

It is apparent from Figure 5 that RD estimates will only yield convincing evidence if a narrow bandwidth is used. Accordingly, we run local linear regressions over one- and two-month bandwidths on both sides of the July 1, 1994 threshold:

$$y_i = \alpha_0 + \alpha_1 dob_i + \alpha_2 After_i + \alpha_3 dob_i \cdot After_i + \theta X_i + \varepsilon_i \tag{2}$$

where y_i measures the educational outcome for individual i, dob_i is the date of birth defined as days relative to July 1, $After_i$ is a dummy variable indicating if the child was born on or after July 1, and X_i is a vector of controls including an indicator variable that equals one if the mother of the child has graduated from high school, dummies for the department of residence and an indicator for urban/rural place of residence. The effect of the APE reform is identified from the discontinuity in outcomes of second-born children captured by α_2.

Results are shown in Table 4. None of the estimates is statistically significant, except for the effect on high school graduation up to two years late when we do not control for the mother's education. Note, however, that these estimates are imprecise, such that 95%-confidence intervals include large positive and negative effects – for instance from a 2.8 percentage reduction to a 9 percentage point increase in the probability of graduating on time from middle high-school, according to column

(4). In Appendix B, we show why controlling for the mother's education and using small bandwidths is needed: there is a strong seasonality in the week-of-birth effects, implying a nonlinear relationship between the forcing variable and the outcome variables. This seasonality seems to be induced by differential fertility behaviors across socioeconomic groups, as shown by variations in the average education of mothers giving birth in different months. While the RD approach remains valid in that context, the continuous but nonlinear relationship imposes the use of small bandwidths, which in turn reduces the precision of the estimates (indeed, standard errors are multiplied by a factor of about 5 when compared to those from the DD). Overall, RD and DD lead to the same conclusion: there is no statistically detectable effect of the APE expansion on long-term schooling outcomes.]

Other comparison groups

The exact choice of control group in the DD analysis deserves attention. Even though the reform of the APE affected second-born children, first-born children could indirectly be affected by the APE if they have younger siblings, as they would also be exposed to more time with their mother if she takes the extended parental leave. To avoid an attenuation bias due to such within-family spillovers, one may prefer to exclude first-born children living with other younger siblings from the control group, and only keep children in one-child families. A counter-argument, however, is that families with one child may differ from larger families in unobserved ways that lead to differential trends, thus violating the parallel trend assumption.

How to compare the two risks –attenuation bias vs. non-parallel trends– is not clear, and a pragmatic solution is to check the robustness of the results to alternative control groups, which we do in two ways. First, we re-define our comparison group and include all children of rank 1 in our DD analysis (column 1 in Table 5), regardless of whether they have younger siblings or not. The results are robust to this change (the coefficient for on-time high school graduation is significant at the 10% level, but it is not statistically different from the main regression estimate). Alternatively, we keep in the control group only first-born children who are at least three years older than their younger sibling. Most children have started pre-school (*école maternelle*) by age three in France, and spillover effects due to maternal presence are likely to be smaller for kids who are at pre-school during the day. The results reported in column 2 are not statistically different from the main DD estimations.[28]

Last, we address an important issue, which is that the socioeconomic composition between second and first children born in the first semester of 1994 may be different. Children born in the first semester of 1994 were conceived in the wake of one of France's deepest recessions of the second half of the twentieth century, which could have induced changes in the socioeconomic background of second-born children in the first semester of 1994. As the timing of birth may be especially responsive to economic conditions for a second born, those changes may not be mirrored by first-born children. In that case, the first semester of 1994 in a DD analysis may introduce a bias. Column 3 in Table 5 excludes the whole 1994 cohort from the DD analysis.[29]

[28] Results also remain qualitatively unchanged when third-born children (whose parents have been eligible to APE throughout the period) are included in the control group along with first-born children. Due to the smaller sample size, however, the results are too imprecise when they are used alone as controls.

[29] This is similar in spirit to a "donut" RD design, which has been used to account for imbalance in covariates due to sorting. The "donut" omits from the analysis those observations that are closest

The results remain qualitatively unchanged.

5 Discussion

In our analysis of the expansion of a flagship policy in France that allowed mothers to stay at home with their child until the age of three, inducing a substantial increase in the time mothers spent at home with their child, we found no detectable effects over a range of schooling achievement outcomes measured at the end of middle or high school, and no heterogeneity by socioeconomic status, public daycare availability, or gender.

Canaan (2022) provides results most closely related to ours when studying the effects of the same APE reform not only on several parental outcomes, but also on measures of the child's language skills at the end of kindergarten. Using an RD approach similar to that in Section 4, she finds large negative drops at the cutoff on indicator variables such as "spontaneous speech is average" and "overall speech is average."[30] Combined with our findings, these results may imply that the long maternal leave associated with APE has a negative impact on language skills at age six that fades out over time or at least has no consequences on middle-school and high-school completion.[31] Data at intermediate ages would be needed to take

to the threshold where such sorting is present (Barreca *et al.*, 2011).

[30]These measures are part of a general medical check-up performed at school by doctors and nurses.

[31]However, we notice that the RD graphs (see Figures 4 (d) and (e) in Canaan (2022) display large and nonlinear month-of-birth effects, with a change in slope around the cutoff, similar to our data at later ages. These complex patterns associated with the child's birth date, which arise due to differences in fertility behavior across socioeconomic groups, stress the importance of truly local estimations to avoid the risk of false positives. Our findings that RD design in this context is highly sensitive to the chosen bandwidth might also explain the divergence between the negative impacts

a stand on whether the results at age six show the existence of a dynamic effect, first negative and then fading out. Unfortunately, we are not aware of such a data source that would include birth rank and schooling outcomes in France for cohorts born around 1994. Despite this "missing middle", our findings and Canaan's results allow us to confidently reject any large positive effects of the APE on children schooling achievement that would justify the well-documented costs imposed on mothers' careers.

The broader international literature paints a consistent picture, in which maternal care during the first months of life enhances the development of the child, and thus policies promoting short maternity leaves are highly beneficial (Rossin, 2011, Carneiro *et al.*, 2015, Berger *et al.*, 2005). In contrast, the evidence for extended parental leaves tends to show no gains in terms of child development. Dahl *et al.* (2016) and Danzer and Lavy (2018) study expansions of paid leave from 18 to 35 weeks and from 12 to 24 months, respectively, and find no impact on children's schooling outcomes. Looking at an expansion from 25 to 50 weeks, Baker and Milligan (2015) find no evidence of positive impacts over a range of measures of children's cognitive and behavioral development, and some evidence of negative effects for boys and children of more educated mothers. Dustmann and Schönberg (2012) study an expansion of job-protected leave from 18 to 36 months and document a small overall negative effect on children's educational achievement at age 14. One potential explanation for these contrasting findings depending on the length of the leave is that, by increasing the time with the mother beyond the first months of life, longer leaves

found by Canaan (2022) and the absence of statistically significant effects in our analysis.

imply a decrease in interactions with other children as well as professional caregivers that older children could benefit from. In addition, extended parental leaves usually go hand in hand with a decrease in household income due to foregone wages, which could also explain these findings.

The APE extension we analyzed in this paper increased the incentives to exit the labor market after giving birth from four months (length of the legal maternity leave) to three years. Our zero results are consistent with the view that maternal care has an initial advantage (until age one), offset by a negative effect above one, as suggested by the results of Dustmann and Schönberg (2012). However, unlike other parental leave policies analyzed in the literature that affect all families equally, our findings apply to families who have a second child, and may not necessarily be the same for one-child families.

Some studies have detected heterogeneous effects across subgroups depending on the quality of maternal time relative to the alternative mode of care. According to a longstanding strand of research in psychology, one-to-one high-quality interactions with a trusted adult are critical for early development (see the review in Fort *et al.*, 2019). The zero overall effects on schooling outcomes found by Danzer and Lavy (2018) are actually the result of a positive impact for children of highly educated mothers and a negative one among children of less educated mothers. In the same spirit, Fort *et al.* (2019) demonstrated that substituting maternal time by institutional time during the first three years of life has negative effects on IQ and personality traits at the ages of 8-14 for children from affluent families. Danzer *et al.* (2017) explicitly take into account the alternative mode of care replaced by maternal time and show that the benefits of extended parental leave are concentrated among children for whom maternal care replaced informal care (mostly by grandparents), while no impact is found for children who switched from formal care to maternal care. Taken together, these studies suggest that extended parental leaves might have beneficial effects for the offspring of highly educated women in an institutional setting with no formal childcare system for very young children. Our results, however, are less consistent with the idea that the quality of the time spent with the mother (relative to the alternative mode of care) is critical, as they are zero across subgroups of children from different socioeconomic backgrounds and with varying availability of formal daycare.

While these important interpretation issues call for further research, the absence of positive impacts on children's long-term educational outcomes implies that this extended maternal leave policy does not enhance child development. The leave has adverse effects on mothers without having detectable positive effects on their children: the estimated aggregate welfare effect is negative, which is an important result for the longstanding debate on the social value of extended maternity leave.

References

Afsa, Cédric. 1998. *L'allocation parentale d'éducation: entre politique familiale et politique pour l'emploi.* INSEE Première 569.

Almond, Douglas, and Currie, Janet. 2011. Human capital development before age five. *Pages 1315–1486 of:* Card, David, and Ashenfelter, Orley (eds), *Handbook of Labor Economics*, vol. 4. Elsevier.

Bailleau, Guillaume. 2010. *L'offre d'accueil collectif des enfants de moins de 6 ans en 2008. Enquête annuelle auprès des services de PMI.* Serie Statistiques. Document de Travail 146. Direction de la recherche, des études, de l'évaluation et des statistiques DREES.

Baker, Michael, and Milligan, Kevin. 2010. Evidence from Maternity Leave Expansions of the Impact of Maternal Care on Early Child Development. *Journal of Human Resources*, **45**(1), 1–32.

Baker, Michael, and Milligan, Kevin. 2015. Maternity leave and children's cognitive and behavioral development. *Journal of Population Economics*, **28**(2), 373–391.

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

Baker, Michael, Gruber, Jonathan, and Milligan, Kevin. 2015 (Sept.). *Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program.* Working Paper 21571. National Bureau of Economic Research. Series: Working Paper Series.

Barreca, Alan I., Guldi, Melanie, Lindo, Jason M., and Waddell, Glen R. 2011. Saving Babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics*, **126**(4), 2117–2123. Publisher: Oxford Academic.

Berger, Lawrence M., Hill, Jennifer, and Waldfogel, Jane. 2005. Maternity Leave, Early Maternal Employment and Child Health and Development in the US. *The Economic Journal*, **115**(501), F29–F47.

Black, Sandra E., Devereux, Paul J., and Salvanes, Kjell G. 2005. The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *The Quarterly Journal of Economics*, **120**(2), 669–700.

Black, Sandra E, Devereux, Paul J, and Salvanes, Kjell G. 2007 (July). *Older and Wiser? Birth Order and IQ of Young Men*. Working Paper 13237. National Bureau of Economic Research.

Black, Sandra E, Grönqvist, Erik, and Öckert, Björn. 2017 (May). *Born to Lead? The Effect of Birth Order on Non-Cognitive Abilities*. Working Paper 23393. National Bureau of Economic Research.

Borra, Cristina, Gonzalez, Libertad, and Sevilla, Almudena. 2014. *The Impact of Eliminating a Child Benefit on Birth Timing and Infant Health*. Discussion Paper 7967. Institute for the Study of Labor (IZA).

Canaan, Serena. 2022. Parental leave, household specialization and children's well-being. *Labour Economics*, **75**, 102–127.

Carneiro, Pedro, Løken, Katrine V., and Salvanes, Kjell G. 2015. A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children. *Journal of Political Economy*, **123**(2), 365–412.

Dahl, Gordon B., Løken, Katrine V., Mogstad, Magne, and Salvanes, Kari Vea. 2016. What Is the Case for Paid Maternity Leave? *The Review of Economics and Statistics*, **98**(4), 655–670.

Daniel, Audrey, and Ruault, Marie. 2004. Les modes d'accueil des enfants de moins de 6 ans : premiers résultats de l'enquête réalisée en 2002. *Études et Résultats, DREES*, **235**, 1–12.

Danzer, Natalia, and Lavy, Victor. 2018. Paid Parental Leave and Children's Schooling Outcomes. *The Economic Journal*, **128**(608), 81–117.

Danzer, Natalia, Halla, Martin, Schneeweis, Nicole E., and Zweimüller, Martina. 2017. *Parental Leave, (In)Formal Childcare and Long-Term Child Outcomes*. CEPR Discussion Papers 12064. Centre for Economic Policy Research (CEPR).

Dehejia, Rajeev, and Lleras-Muney, Adriana. 2004. Booms, Busts, and Babies' Health. *The Quarterly Journal of Economics*, **119**(3), 1091–1130.

Desplanques, Guy, and Rogers, Godfrey. 2008. Strengths and Uncertainties of the French Annual Census Surveys. *Population (English Edition, 2002-)*, **63**(3), 415–439.

Drange, Nina, and Havnes, Tarjei. 2018. Early Childcare and Cognitive Development: Evidence from an Assignment Lottery. *Journal of Labor Economics*, **37**(2), 581–620. Publisher: The University of Chicago Press.

Dustmann, Christian, and Schönberg, Uta. 2012. Expansions in Maternity Leave Coverage and Children's Long-Term Outcomes. *American Economic Journal: Applied Economics*, **4**(3), 190–224.

Felfe, Christina, and Lalive, Rafael. 2018. Does early child care affect children's development? *Journal of Public Economics*, **159**(Mar.), 33–53.

Fort, Margherita, Ichino, Andrea, and Zanella, Giulio. 2019. Cognitive and Noncognitive Costs of Day Care at Age 0–2 for Children in Advantaged Families. *Journal of Political Economy*, **128**(1), 158–205. Publisher: The University of Chicago Press.

Gans, Joshua S., and Leigh, Andrew. 2009. Born on the first of July: An (un)natural experiment in birth timing. *Journal of Public Economics*, **93**(1), 246–263.

Gosset-Connan, Stéphanie. 2004. Les usages des bénéficiaires de l'APE attribuée pour le deuxième enfant. *Revue des politiques sociales et familiales*, **75**(1), 39–48.

Grenet, Julien. 2010. Academic Performance, Educational Trajectories and the Persistence of Date of Birth Effects. Evidence from France. *Unpublished manuscript, Center for Economic Performance, London School of Economics, London, UK.*

Jacquot, Alain. 2000. Introduction. *Recherches et Prévisions*, **59**, 1–8.

Kleven, Henrik, Landais, Camille, Posch, Johanna, Steinhauer, Andreas, and Zweimüller, Josef. 2019 (Feb.). *Child Penalties Across Countries: Evidence and Explanations*. Working Paper 25524. National Bureau of Economic Research.

Kottelenberg, Michael J., and Lehrer, Steven F. 2017. Targeted or Universal Coverage? Assessing Heterogeneity in the Effects of Universal Child Care. *Journal of Labor Economics*, Mar. Publisher: University of Chicago PressChicago, IL.

Lee, David S., and Lemieux, Thomas. 2010. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, **48**(2), 281–355.

Lequien, Laurent. 2012. The Impact of Parental Leave Duration on Later Wages. *Annals of Economics and Statistics*, 267–285.

Maurin, Eric, and McNally, Sandra. 2008. Vive la Révolution! Long-Term Educational Returns of 1968 to the Angry Students. *Journal of Labor Economics*, **26**(1), 1–33.

McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**(2), 698–714.

Moschion, Julie. 2010. Reconciling Work and Family Life: The Effect of the French Paid Parental Leave. *Annals of Economics and Statistics*, 217–246.

Piketty, Thomas. 1998. L'impact des incitations financières au travail sur les comportements individuels : une estimation pour le cas français. *Economie et Prévisions*, **132-133**, 1–35.

Piketty, Thomas. 2003. *L'impact de l'allocation parentale d'éducation sur l'activité féminine et la fécondité, 1982-2002.* Working Paper 2003-09. CEPREMAP.

Piketty, Thomas. 2005. *L'impact de l'allocation parentale d'éducation sur l'activité féminine et la fécondité en France, 1982-2002.* In: Lefèvre, C. (Ed.), Histoires de familles, histoires familiales: Les Cahiers de l'INED 156. Institut National des Etudes Démographiques (INED).

Rambachan, Ashesh, and Roth, Jonathan. 2020. *An Honest Approach to Parallel Trends.* Working Paper. Harvard University.

Rossin, Maya. 2011. The effects of maternity leave on children's birth and infant health outcomes in the United States. *Journal of Health Economics*, **30**(2), 221–239.

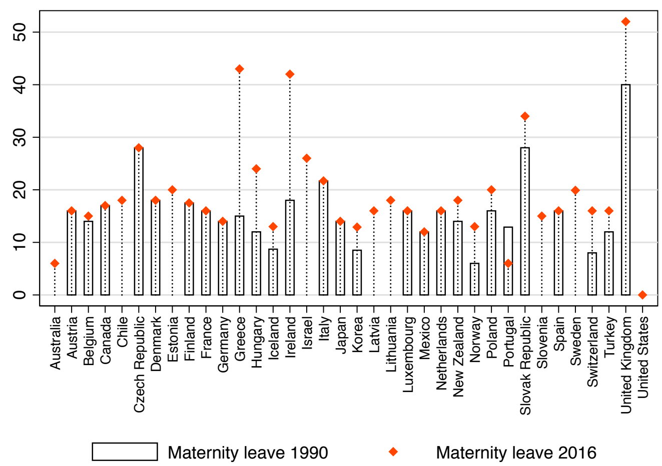
Rossin-Slater, Maya. 2018. Maternity and Family Leave Policy. *In:* Averett, Susan L., Argys, Laura M., and Hoffman, Saul D. (eds), *The Oxford Handbook of Women and the Economy.* New York: Oxford University Press.

Tamm, Marcus. 2013. The Impact of a Large Parental Leave Benefit Reform on the Timing of Birth around the Day of Implementation. *Oxford Bulletin of Economics and Statistics*, **75**(4), 585–601.

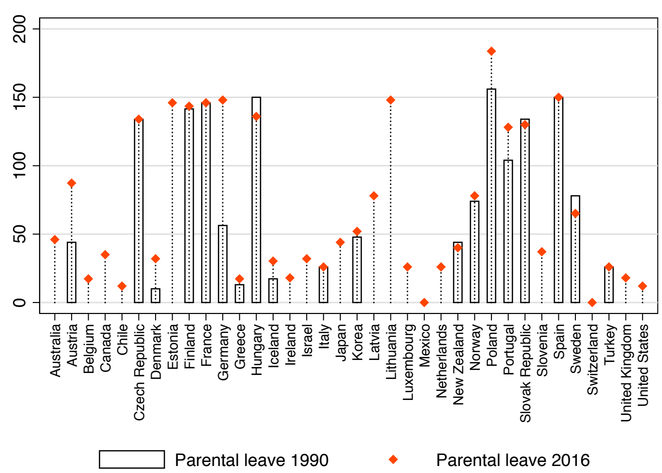
Figures and Tables

Figure 1: Weeks of maternity leave and job-protected parental leave available to women

(a) Paid maternity leave

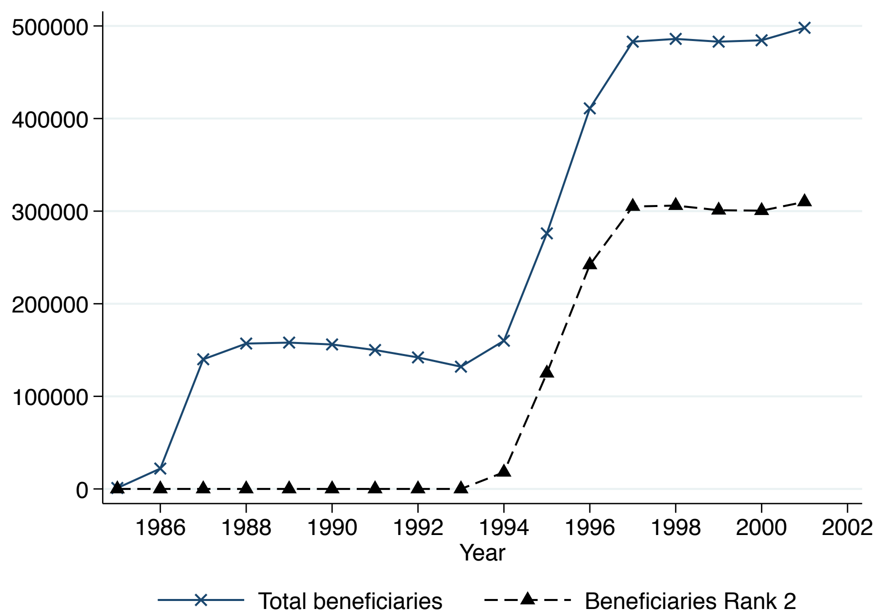


(b) Job-protected parental leave



Note: Paid maternity leave refers to paid job-protected leave available for mothers just before and after childbirth. Parental leave with job protection refers to the number of weeks after maternity leave which a woman can take with her job protected, disregarding payment conditions. Data reflect entitlements at the national or federal level only. Source: OECD Family database, accessed February 2020.

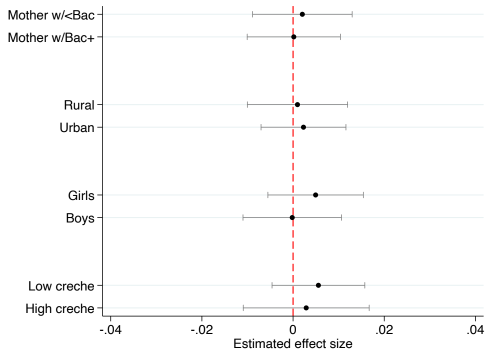
Figure 2: Number of APE beneficiaries



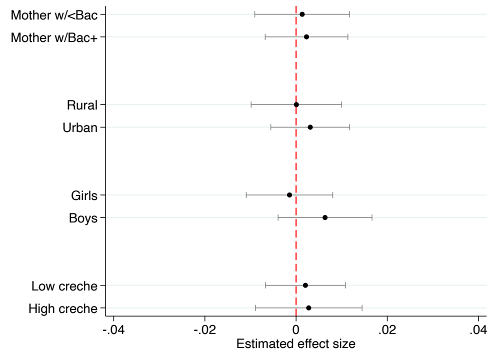
Note: The total number of beneficiaries corresponds to APE beneficiaries from all social security regimes (the *Caisse des allocations familiales* -CAF- from Metropolitan France, the special regimes, and the *Mutualité Sociale Agricole* - MSA), while the number of APE beneficiaries of Rank 2 refers to those APEs payed by the main regime (CAF) only. The APEs payed by the other regimes would represent an additional 10 percent (Piketty, 1998). Source: Piketty (2003) based on *Caisse Nationale des Allocations Familiales* (CNAF).

Figure 3: Heterogeneous effects

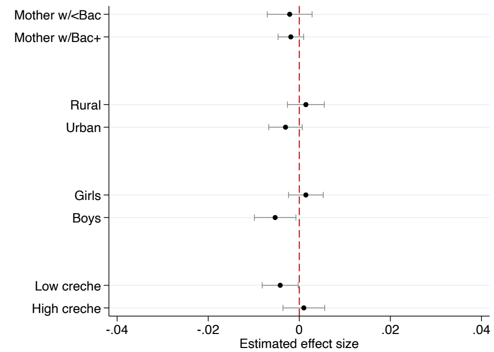
(a) Middle HS certif. on time



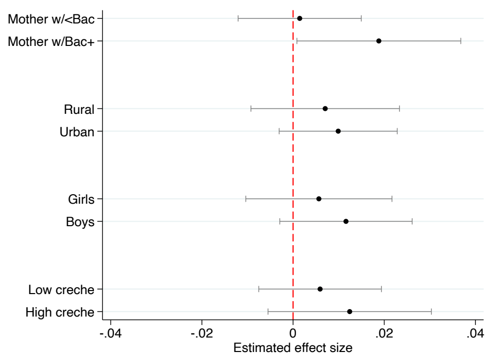
(b) Middle HS certif. up to 2 years late



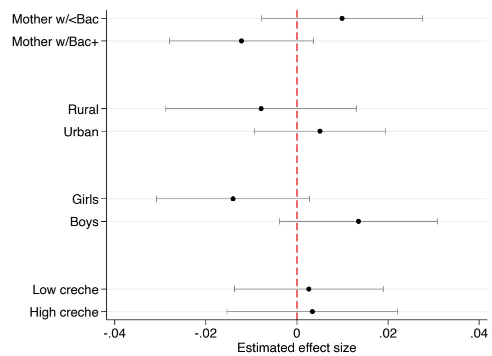
(c) Dropout at age 16



(d) Senior HS grad. on time



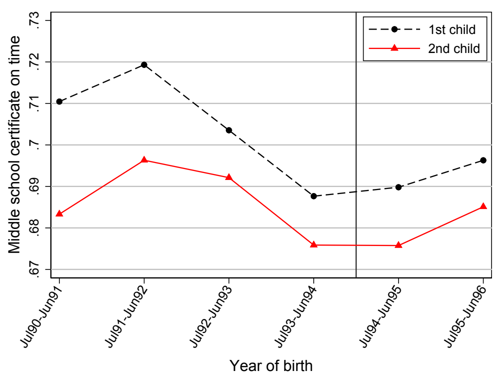
(e) Senior HS grad. up to 2 years late



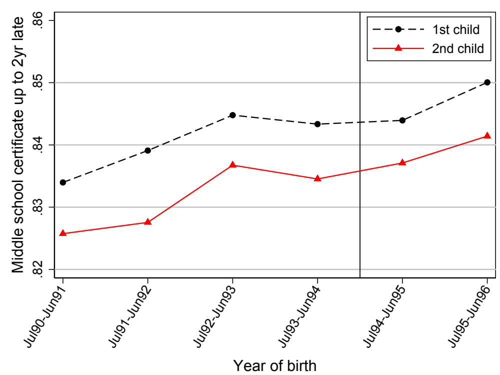
Note: The dots and bars represent the estimated coefficients and 95% confidence intervals (robust standard errors) for the interaction between the APE eligibility indicator and a dummy equal to one for cohorts born after the APE reform, run on the indicated subsample. Also included (but not reported) are the eligibility indicator, year dummies and a set of controls including maternal education, a dummy for urban / rural place of residence and dummies for the department of residence. Source: French Population Census, EARs 2008-2016.

Figure 4: Evolution of children's educational outcomes

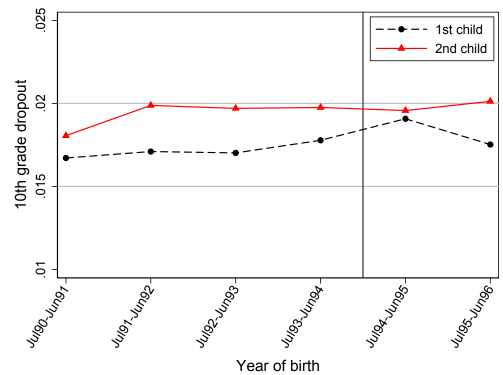
(a) Middle HS certif. on time



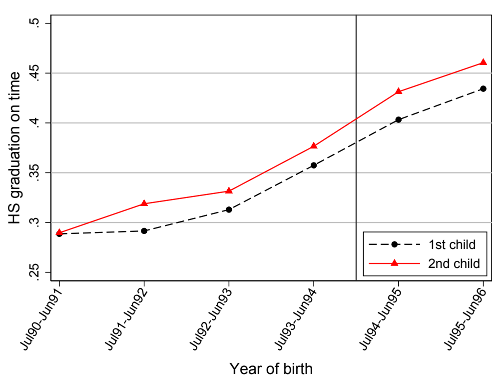
(b) Middle HS certif. up to 2 years late



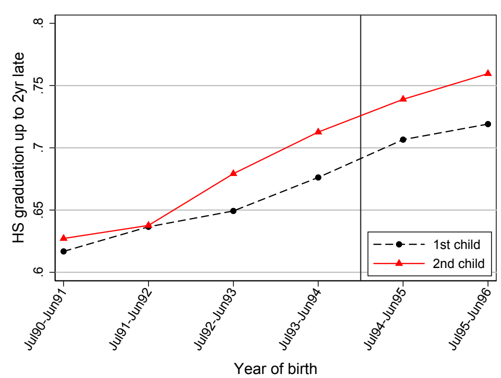
(c) Dropout at age 16



(d) Senior HS grad. on time

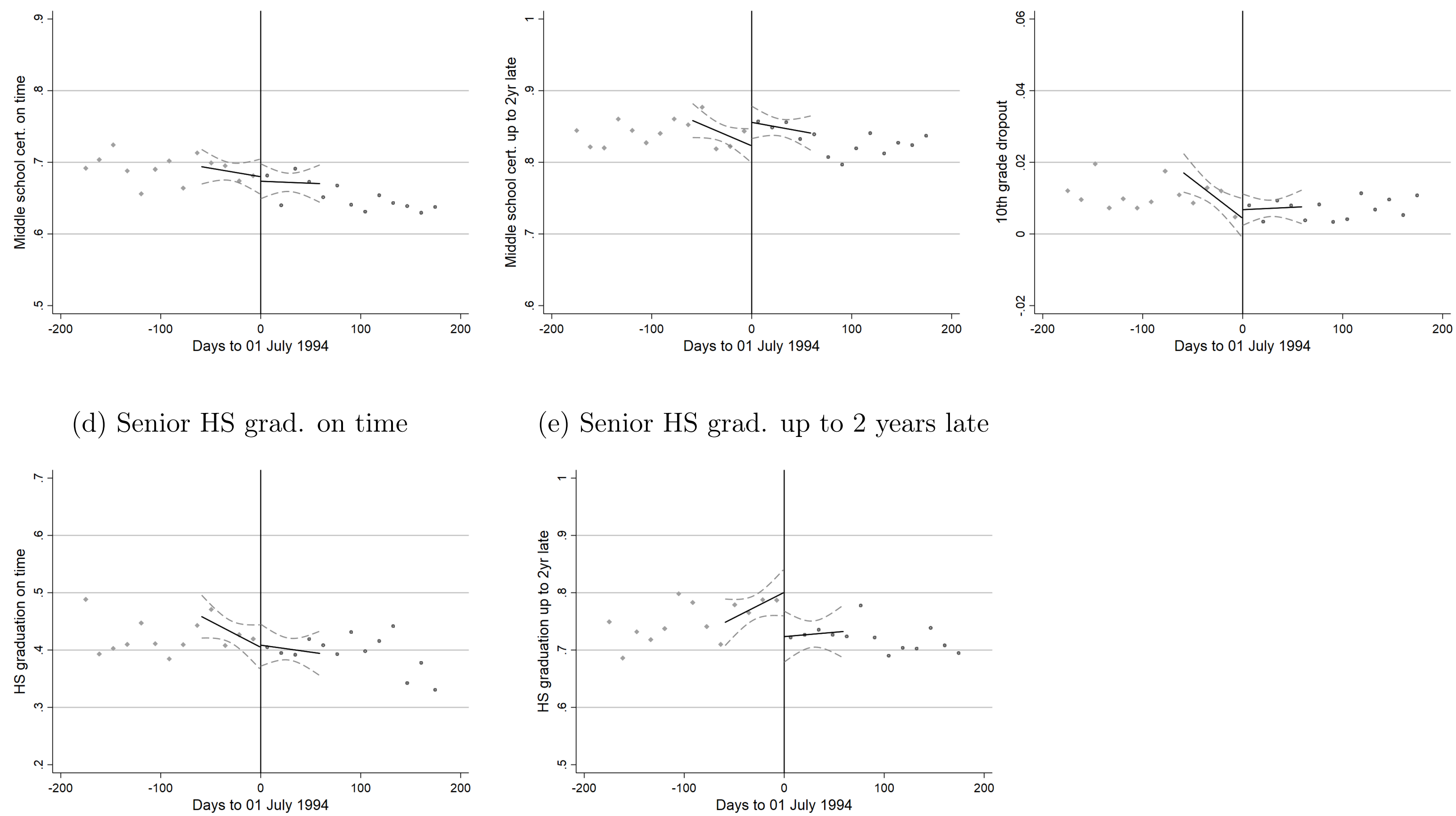


(e) Senior HS grad. up to 2 years late



Note: The dashed black line represents the average value of each outcome for first-born children in single-child households, while the solid red line refers to second-born children. Source: French Population Census, EARs 2008-2016.

Figure 5: Educational outcomes of second-born children on each side of the threshold



(a) Middle HS certif. on time

(b) Middle HS certif. up to 2 years late

(c) Dropout at age 16

(d) Senior HS grad. on time

(e) Senior HS grad. up to 2 years late

Note: The solid lines correspond to linear regressions fitted separately before and after the reform of the APE using a 2-month bandwidth, and the dashed lines represent the 95% confidence intervals. Each dot represents the bi-weekly average outcome. The sample corresponds to children of rank 2 born in 1994. Source: French Population Census, EARs 2010-2015.

Table 1: Differences-in-differences estimates

	(1)	(2)	(3)
Middle HS certificate on time	0.0001	0.001	0.002
	(0.004)	(0.004)	(0.004)
N	364,660	359,330	342,997
Middle HS certificate up to 2 years late	0.0004	0.0003	0.002
	(0.004)	(0.004)	(0.004)
N	307,114	302,055	287,147
School dropout by age 16	-0.0001	-0.002	-0.002
	(0.001)	(0.002)	(0.002)
N	365,708	360,367	343,936
HS graduation on time	0.008	0.008	0.008
	(0.006)	(0.006)	(0.006)
N	201,975	199,211	187,120
HS graduation up to 2 years late	0.002	0.002	0.002
	(0.006)	(0.006)	(0.006)
N	132,897	132,148	122,485
Qob dummies	Yes	Yes	Yes
Geographic controls	No	Yes	Yes
Mother's education	No	No	Yes

Note: The estimates correspond to the interaction between $Rank_2$ and $After$ in equation 1. The comparison group includes children in single-child households. Mother's education refers to a dummy equal to one if she is (at least) a high school graduate, and geographic controls include dummies for the department of residence and an indicator for urban/rural place of residence. All controls refer to the time of the survey. The sample includes individuals born between July 1992 and June 1996. Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

Table 2: Heterogeneous effects

	By maternal education		By gender		By area		By *crèche* availability	
Outcome:	HS graduate	HS dropout	Boys	Girls	Urban	Rural	High	Low
Middle HS certificate on time	0.000	0.002	-0.000	0.005	0.002	0.001	0.003	0.006
	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.007)	(0.005)
Middle HS certificate up to 2 years late	0.002	0.001	0.006	-0.001	0.003	0.000	0.003	0.002
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.004)
Dropout by age 16	-0.002	-0.002	-0.005**	0.001	-0.003	0.001	0.001	-0.004**
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
HS graduation on time	0.019**	0.001	0.012	0.006	0.010	0.007	0.012	0.006
	(0.009)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.009)	(0.007)
HS graduation up to 2 years late	-0.012	0.010	0.014	-0.014	0.005	-0.008	0.003	0.003
	(0.008)	(0.009)	(0.009)	(0.009)	(0.007)	(0.011)	(0.010)	(0.008)

Note: The estimates correspond to the interaction between $Rank_2$ and $After$ in equation 1, run on the indicated subsample. Mother with Bac refers to whether the mother graduated from senior high school. High *crèche* is a dummy equal to one if the department of residence is on the top 25% of the *crèche* availability distribution. Controls as in column 7 of Table 1 column 3 are included. Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

Table 3: 95% confidence intervals of naïve and re-inflated ITT (from DD regressions)

	Naïve ITT	True ITT
	(1)	(2)
Middle HS certificate on time	[-0.005; 0.010]	[-0.009; 0.017]
Middle HS certificate up to 2 years late	[-0.005; 0.009]	[-0.009; 0.018]
School dropout by age 16	[-0.005; 0.001]	[-0.009; 0.002]
Senior HS graduation on time	[-0.003; 0.019]	[-0.006; 0.037]
Senior HS graduation up to 2 years late	[-0.011; 0.014]	[-0.026; 0.035]

Note: The table reports the 95% confidence intervals for the coefficients estimated in the DD regressions. The first column reports the *naïve* C.I.s, and the second one adjusts them for birth rank measurement error, following the procedure described in C.3.

Table 4: Regression discontinuity estimates

	Bandwidth = 2 months		Bandwidth = 1 month	
	(1)	(2)	(3)	(4)
Middle HS certificate on time				
Born after 01/07/1994	0.000	0.004	0.019	0.031
	(0.021)	(0.021)	(0.030)	(0.030)
N	10,778	10,224	5,511	5,251
Middle HS certificate up to 2 years late				
Born after 01/07/1994	0.029	0.022	0.018	0.010
	(0.020)	(0.020)	(0.027)	(0.028)
N	7,253	6,882	3,795	3,599
School dropout by age 16				
Born after 01/07/1994	0.002	0.002	0.005	0.005
	(0.004)	(0.004)	(0.005)	(0.005)
N	10,804	10,249	5,523	5,263
HS graduation on time				
Born after 01/07/1994	-0.006	-0.022	0.005	0.007
	(0.032)	(0.031)	(0.045)	(0.043)
N	5,162	4,934	2,637	2,516
HS graduation up to 2 years late				
Born after 01/07/1994	-0.079**	-0.042	-0.056	-0.040
	(0.035)	(0.033)	(0.048)	(0.047)
N	3,031	2,981	1,567	1,537
Geographic controls	No	Yes	No	Yes
Mother's education	No	Yes	No	Yes

Note: Results from local linear regressions using a 2-month bandwidth on each side of the cutoff (columns 1 and 2) and 1 month (columns 3 and 4). The sample includes individuals who are second in their sibship. Mother's education refers to a dummy equal to one if she is (at least) a high school graduate, and geographic controls include dummies for the department of residence and an indicator for urban/rural place of residence. All controls are refer to the time of the survey. Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

Table 5: Robustness checks

Outcome:	Control group: All rank-1 children	Control group: Rank-1 children 3+yrs than rank-2 sibling	Exclude 1994 cohort
	(1)	(2)	(3)
Middle school certificate on time	0.002	0.002	0.004
	(0.004)	(0.004)	(0.004)
Middle school cert. up to 2 years late	0.001	0.002	0.002
	(0.004)	(0.004)	(0.004)
School dropout by age 16	-0.0002	-0.002	0.001
	(0.001)	(0.002)	(0.001)
Senior HS graduation on time	0.010*	0.008	0.009
	(0.005)	(0.005)	(0.006)
Senior HS graduation up to 2 years late	0.005	0.006	0.009
	(0.006)	(0.006)	(0.007)
Qob dummies	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
Mother's education	Yes	Yes	Yes

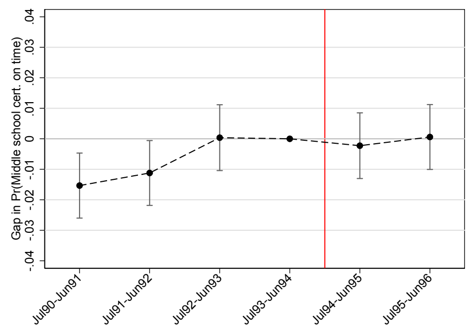
Note: The estimates correspond to the interaction between $Rank_2$ and $After$ in equation 1. Column (1) includes only first-born children who are at least 3 years older than their younger sibling in the comparison group. Column (2) includes all first-born children in the comparison group, including those with younger siblings. Column (3) excludes the 1994 cohort from the analysis. Mother's education refers to a dummy equal to one if she is (at least) a high school graduate, and geographic controls include dummies for the department of residence and an indicator for urban/rural place of residence. All controls refer to the time of the survey. Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

Appendices

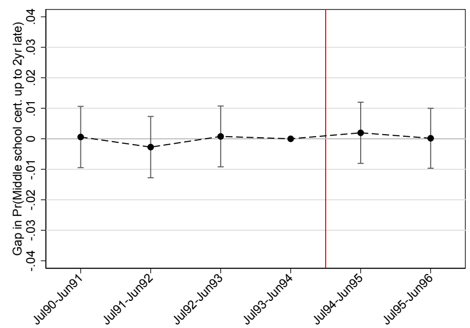
A Appendix Figures and Tables

Figure A1: Difference in educational outcomes between second and first-born children

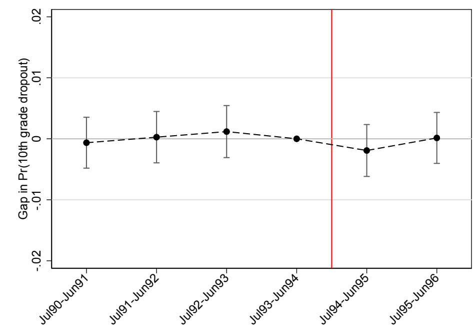
(a) Middle HS certif. on time



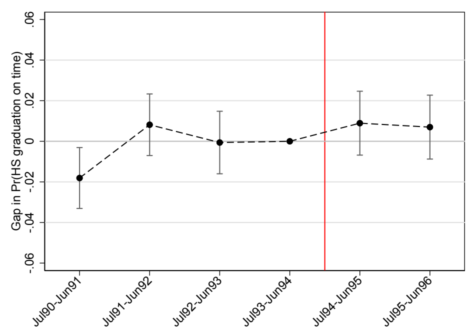
(b) Middle HS certif. up to 2 yrs late



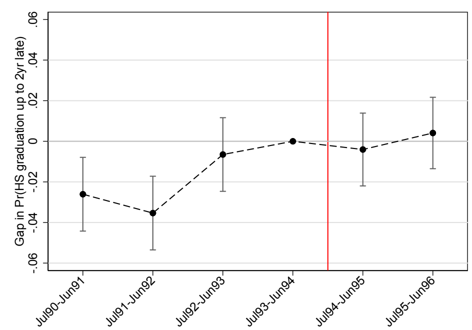
(c) Dropout at age 16



(d) Senior HS grad. on time



(e) Senior HS grad. up to 2 yrs late



Note: The dots and bars represent the point estimates and the 95% confidence intervals for year dummies interacted with an indicator for being rank-2 (in a regression including year dummies and the rank-2 indicator as well). The coefficient is normalized to zero for the last period before the reform (July 1993 - June 1994). Source: French Population Census, EARs 2008-2016.

Table A1: Outcomes analyzed and EARs used

	Born in				
	1992	1993	1994	1995	1996
Start 9th grade in	Sep-06	Sep-07	Sep-08	Sep-09	Sep-10
Take BEPC first time in	Jun-07	Jun-08	Jun-09	Jun-10	Jun-11
Observed in EARs (Jan.):					
on time	2008	2009	2010	2011	2012
up to 1 year late	2009	2010	2011	2012	2013
up to 2 years late	2010	2011	2012	2013	2014
Start 12th grade in	Sept-09	Sep-10	Sep-11	Sep-12	Sep-13
Take Bac first time in	Jun-10	Jun-11	Jun-12	Jun-13	Jun-14
Observed in EARs (Jan.):					
on time	2011	2012	2013	2014	2015
up to 1 year late	2012	2013	2014	2015	2016
up to 2 years late	2013	2014	2015	2016	2017

Note: BEPC refers to *Diplôme National du Brevet*, a certificate obtained after an examination at the end of middle school. Bac (for *baccalauréat*) is the exam taken at the end of senior high school.

Table A2: Test of parallel trends before APE extension

	Middle school certificate				Senior HS graduation on time				School dropout by age 16	
	on time		up to 2 years late		on time		up to 2 years late			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Linear trend	0.004***	0.004***	0.001	0.001	0.002	0.000	0.005***	0.005***	0.000*	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)	(0.000)
Year dummies										
Jul90-Jun91	-0.015***	-0.017***	0.001	0.000	-0.018**	-0.014*	-0.029***	-0.029***	-0.001	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.009)	(0.001)	(0.001)
Jul91-Jun92	-0.011**	-0.012**	-0.003	-0.004	0.008	0.009	-0.038***	-0.034***	0.001	0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.009)	(0.001)	(0.001)
Jul92-Jun93	-0.000	0.000	0.001	0.001	-0.001	0.000	-0.009	-0.008	0.000	0.000
	(0.006)	(0.006)	(0.005)	(0.005)	(0.008)	(0.008)	(0.009)	(0.009)	(0.001)	(0.001)
With controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Each row reports the coefficients for a linear time trend (Panel A) or year dummies (Panel B) interacted with the indicator of being the second-born child. All columns include a dummy for being the second-born child and cohort fixed effects. The omitted category is the last cohort prior to the APE reform (July 1993 - June 1994). Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

B Implementation of the RD approach

In this appendix, we provide further details on the validity and implementation of the RD approach presented in Section 4.

The first requirement is that no other treatment creates a discontinuity at the same cutoff. As discussed in Section 4, this condition is fulfilled in the case of APE, as as no other reform was introduced with the July 1, 1994 birth date as an eligibility cutoff. The well-documented discontinuity in schooling achievement due to a discontinuity in the age at school entry occurs in France at the January 1 cutoff, not at July 1.

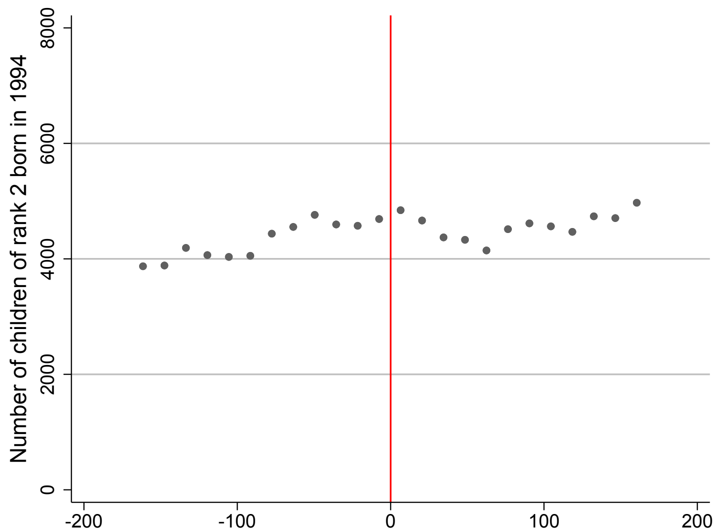
Another concern would be a discontinuous change in the characteristics of children born just before and just after the cutoff. Following Lee and Lemieux (2010), this will only occur in the case of *precise* manipulation of the date of birth, i.e. if some parents could strategically time the date of birth of their second child to become eligible for APE benefits.[32] This seems unlikely in this case: as mentioned in Section 2, the law that modified the APE was passed on July 25, with retroactive effects on parents of children born from July 1, 1994 onward, and was not announced in advance. We check empirically for manipulation of the date of birth. Figure B1 plots the bi-weekly number of rank-2 children born in 1994, and does not suggest any strong heap on the right-hand side of the cutoff. The McCrary (2008) test yields an estimated discontinuity of 0.024 and a standard error of 0.022, showing no evidence of a significant discontinuity in the density of births at the cutoff.

Taken together, these two elements – the general effect of age at school entry and the absence of information to sort precisely at the reform cutoff – suggest that, if APE has no effect on schooling achievement, schooling outcomes should be continuous and monotonically decreasing with the date of birth, between January 1 and December 31, 1994. The results in Figure 5 do not suggest any jump at the July 1 threshold. What we can detect, however, are the potential non-linear effects of the date of birth, possibly induced by differences in fertility by socioeconomic groups. As documented

[32] There is evidence showing that economic incentives can indeed induce parents to postpone or to anticipate the date of birth of their newborns. See, for instance, Borra *et al.* (2014), Gans and Leigh (2009), Tamm (2013).

by Grenet (2010), the fertility behaviors of different socioeconomic groups in France display different seasonality patterns, which may introduce birth date effects on schooling achievement within a yearly cohort. To give one striking example, teachers are more likely than other occupations to have children in the spring – then the end of the legal maternal leave (outside APE) coincides with the beginning of the summer school vacation. By contrast with the effect driven by age at school entry, these effects have no reason to be monotonic. Moreover, they may vary across cohorts, in particular in times of recession.[33]

Figure B1: Number of second-born children, by date of birth (bi-weekly)



Note: The horizontal axis measures the number of days normalized to 0 for the first two weeks after the APE reform. Each dot represents the number of children of rank 2 born per two-week period, and the vertical line indicates the time of the APE reform. Data corresponds to individuals still living in their parents' household by the ages of 15 to 20 years old. The formal McCrary (2008) test yields an estimated discontinuity of 0.024 and a standard error of 0.022, consistent with the absence of bunching in the density of births at the threshold. Source: French Population Census, EARs 2008-2016.

Seasonal composition effects do not undermine identification in the RD approach

[33]The literature has shown that fertility responses to busts vary significantly by socioeconomic and ethnic background (Dehejia and Lleras-Muney, 2004).

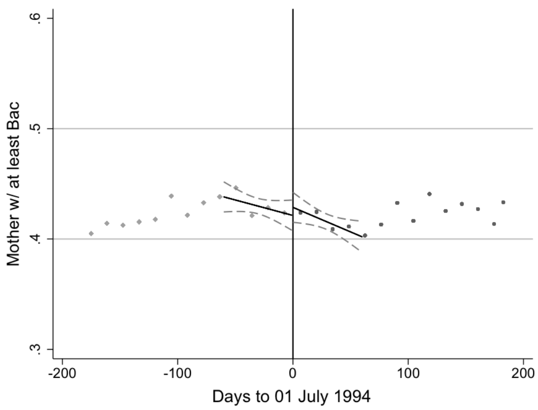
as there is no reason to believe that they are discontinuous at the July 1 birth date. Yet, they may matter for estimation: the specification used should account for possibly non linear effects of the forcing variable. We illustrate the point by running one of the usual tests for precise sorting in RD, asking whether one can detect a discontinuity at the cutoff for some predetermined variables. The 2008-2017 waves of census that we use report few measures that can be considered as predetermined as of 1994, but they contain a critical one: mother's education. Is it the case that differences in fertility behavior in 1994 between more or less educated mothers could induce patterns susceptible of confounding the effect of APE? From inspecting Figure B2 we do not detect any jump in the share of educated mothers at the cutoff, although date-of-birth effects appear non-linear. We illustrate how this may lead to statistically significant estimates when using large bandwidths in Table B1. The results with a 6-month bandwidth indicate that children born after the 1st of July cutoff are 1.4 percentage points less likely to have an educated mother, while if we take a smaller bandwidth of either one or two months, the coefficient changes sign and becomes insignificant. Clearly, the estimates using the six-month bandwidth are spurious as they are driven by the inability of (not so local) linear regressions to fit non-linear date-of-birth effects.

In order to account for these potential birth effects and avoid capturing discontinuities where there are none, our choice is to use local linear regressions with narrow bandwidths. We therefore provide estimations using bandwidths of one and two months on each side of the cutoff date in Table **??**. These specifications imply a loss of statistical precision, but avoid false positives. Overall, the regression results do not suggest any significant change in neither of the outcomes, with point estimates close to zero. Moreover, the point estimates remain around the same magnitude and statistically insignificant after controlling for geographic indicators and maternal education. For the case of graduation with delay, the specification with no controls and a two-month bandwidth estimates a negative and significant effect, but it is almost halved and becomes non significant when controls are added, as well as when a narrower bandwidth of one month is used.

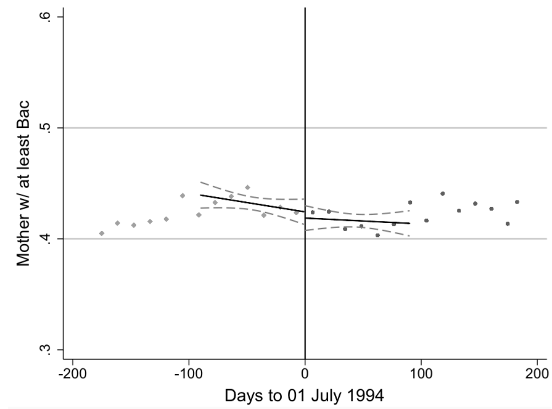
Table B2 performs placebo tests, further validating the RD design: we find no

Figure B2: Share of educated mothers (HS graduates) of children of rank 2 born in 1994

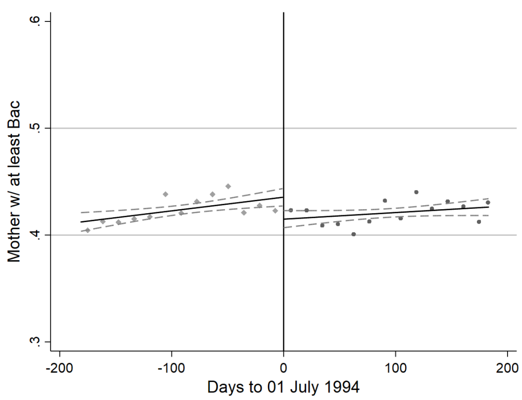
(a) Local linear regression (bw: 2 months)



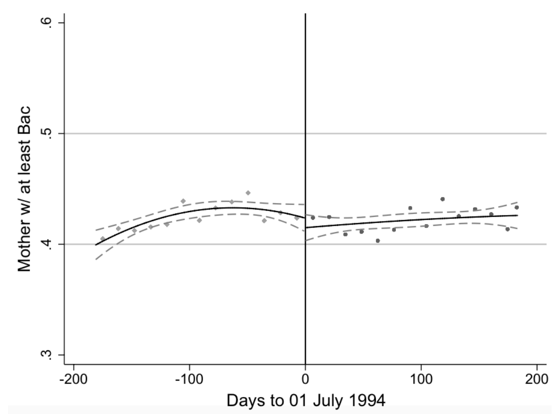
(b) Local linear regression (bw: 3 months)



(c) Local linear regression (bw: 6 months)



(d) Second order polynomial (bw: 6 months)



Note: Each figure plots the share of educated mothers for children of rank 2 born in 1994. Educated mothers are defined as having completed at least senior high school. The solid lines represent regression lines from local linear regressions with bandwidths of 2, 3, and 6 months (subfigures a, b, and c, respectively) and from a quadratic regression using a 6-month bandwidth (subfigure d). Source: French Population Census, EARs 2010-2015.

Table B1: RD estimates of the share of educated mothers (HS or more) for children of rank 2 born in 1994

	Bandwidth used:							
	6 months		2 months		1 month		6 months	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Born after 01/07/1994	-0.020***	-0.014**	0.007	0.012	0.001	0.004	-0.009	-0.004
	(0.007)	(0.007)	(0.012)	(0.012)	(0.017)	(0.016)	(0.010)	(0.010)
N	110,840	109,272	38,305	37,766	19,641	19,372	110,840	109,272
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes
Controls for 2nd-order polynomial of dob	No	No	No	No	No	No	Yes	Yes

Note: Columns 1 to 6 report the results from linear regressions using 6, 2, and 1 month bandwidth on each side of the July 1, 1994 cutoff date. Columns 7 and 8 use a 6-month bandwidth and add a second-order polynomial of the date of birth. The sample includes mothers of rank-2 children born in the specified period. Educated mothers are defined as having completed at least senior high school. Geographic controls include dummies for the department of residence and an indicator for urban/rural place of residence. Source: French Population Census, EARs 2010-2015.

discontinuity in educational outcomes at the July 1 birth date, for first-born children, or second-born children born in 1993 or 1995.

To sum up, even though our data does not reject any of the underlying assumptions of the RD approach, we suggest caution when using this empirical strategy in the analysis of APE effects. The complex and non-linear effects of the month of birth on the outcomes analyzed rule out the use of linear regressions with large bandwidths, as they could capture discontinuities where there are none. Conducted with caution, the RD results, albeit imprecise, are consistent with a zero effect.

Table B2: RD estimates for groups of placebo children

	Rank 1 children (in single-child households), 1994 cohort		Rank 2 children, 1993 cohort		Rank 2 children, 1995 cohort	
	bw = 2m	bw = 1m	bw = 2m	bw = 1m	bw = 2m	bw = 1m
	(1)	(2)	(3)	(4)	(5)	(6)
Junior HS certificate on time:						
Born after 01/07	-0.026*	-0.030	-0.012	-0.016	0.013	0.032
	(0.016)	(0.022)	(0.022)	(0.031)	(0.020)	(0.029)
Junior HS certificate up to 2 years late:						
Born after 01/07	0.002	-0.002	-0.004	-0.000	0.029	0.026
	(0.013)	(0.018)	(0.022)	(0.031)	(0.020)	(0.029)
HS graduation on time:						
Born after 01/07	0.005	0.018	-0.026	-0.020	-0.002	0.027
	(0.021)	(0.030)	(0.030)	(0.043)	(0.031)	(0.044)
HS graduation up to 2 years late:						
Born after 01/07	-0.021	0.014	-0.022	-0.007	-0.001	0.038
	(0.024)	(0.034)	(0.037)	(0.052)	(0.032)	(0.046)
Mother's education	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Coefficients from local linear regressions using the specified bandwidth (2 or 1 month around the cutoff). Columns 1 and 2 refer to children of rank 1 born in 1994, while columns 3 and 4 (5 and 6) refer to children of rank 2 born in 1993 and 1995, respectively. Mother's education refers to a dummy equal to one if she is (at least) a high school graduate and geographic controls include dummies for the department of residence and an indicator for urban/rural place of residence. Robust standard errors in parentheses. *** Significant at 1%; ** significant at 5%; * significant at 10%. Source: French Population Census, EARs 2008-2016.

C Sample selectivity and birth rank mismeasurement

This appendix provides a detailed discussion of the data sources and the methodology used to assess the extent of sample selectivity and birth rank measurement error in our main working sample.

C.1 Sample selectivity

Table C1 provides a direct measure of attrition in our main sample, by comparing the size of cohorts included in our analysis at age 15-20 to their original size as measured in the 1999 population census. Since children born between 1992 and 1996 were 2 to 7 years old in 1999, it is reasonable to think that this is measured before any potential selective attrition. Each column refers to a different cohort-semester. The top panel refers to the number of children observed in the 1999 population census,[34] and the bottom panel to number of individuals from the same cohorts observed in later census rounds, when they are 15-20 years old.[35]

[34] Only children living in regular households are considered, excluding children in collective dwellings.

[35] Given that the 1999 population census was conducted under a traditional format (i.e. exhaustive enumeration), all children are observed at the same point in time, so different cohorts are observed at different ages. For years after 2004, the rotating-sample census based on yearly EARs enumerates about a fifth of a cohort per year. To recover an entire cohort (and compare it with the cohort size observed in 1999) we pool the data from five consecutive EARs.

Table C1: Size of 1992-1996 cohorts as observed in census data at different ages

		Birth cohort									
		1992		1993		1994		1995		1996	
	Age	s1	s2	s1	s2	s1	s2	s1	s2	s1	s2
RP1999	2-3									361,228	369,313
	3-4							352,801	370,762		
	4-5					349,696	360,026				
	5-6			350,118	356,474						
	6-7	370,459	369,637								
EARs 2008-2016	15-16	72,894	72,683	68,916	70,848	69,295	70,624	70,105	72,764	71,189	72,635
	16-17	73,418	73,398	68,700	70,022	68,443	70,091	69,120	73,392	71,476	72,607
	17-18	68,294	69,165	63,883	66,664	64,864	66,688	65,330	70,059	65,738	69,533
	18-19	52,484	54,418	49,687	52,693	49,047	52,842	49,253	54,259	48,668	51,008
	19-20	46,472	48,424	43,081	46,937	42,811	46,584	41,829	46,511	41,792	44,727
	15-20	313,562	318,087	294,267	307,165	294,460	306,829	295,637	316,985	298,862	310,510
% of cohort in 1999		84.6%	86.1%	84.0%	86.2%	84.2%	85.2%	83.8%	85.5%	82.7%	84.1%

Note: The table reports the number of individuals observed in their parents' household (excludes children in collective dwellings and institutions), by year and semester of birth, at different ages. In the second panel, the number of individuals from a given cohort observed at a given age come from one EAR at the time. Source: 1999 Population Census and 2008-2016 EARs.

As we can note from the first column, in 1999 there were about 370 thousand children born both in the first and in the second semester of 1992 (between six and seven years old at the time of census).[36] By the time they were 15 to 20 years old (second panel), there were about 314,000 and 318,000 living in their parents' household, which represent 85 and 86 percent of the cohort size in 1999. Additionally, as expected, the number of youth still in their parents' household decreases with age: while there are almost 73 thousand youth born in 1992 who still co-reside with their parents at the ages of 15-16, the number decreases to 46 to 48 thousands by the ages of 19-20, which simply reflects the fact that children tend to leave their parents' household as they grow up.[37]

The same reading can be done for other cohorts. Overall, our working sample (i.e. youth residing in their parents' household) represents between 83 and 86 percent of the relevant cohorts. Importantly, we find no evidence of differential sample selectivity, as this proportion displays no systematic evolution for individuals born before and after July 1994.

C.2 Birth rank measurement error

A second concern in our data is that we do not directly observe birth rank (which defines the APE eligibility), and derive it from the rank among children in the parents' household, which may not yield the true rank for two reasons. First, the census data does not provide information on family relationships that allow to link each individual with her mother in the household, which is necessary to establish birth rank among siblings. Second, if an older sibling left the household, the rank among

[36] As a benchmark, we compared the estimated cohort sizes from the RP99 with the "full" cohort sizes, computed as the total number of births in France obtained from the vital statistics records reported by INSEE (before any mortality or out-migration). When only children living in regular households are considered, the cohort sizes estimated from the 1999 population census represent 93% of the total number of births. This percentage increases to 96% when children living in non-regular households are included. Importantly, this proportion is relatively constant and has not changed before and after 1994.

[37] Each cohort is observed at a given age in only one annual census survey. For instance, the 1992 cohort is observed in the EAR 2008 at the ages of 15 to 16, in the EAR 2009 at the ages of 16 to 17, and so on.

children in the household will not exactly match the true rank. We analyze each of these issues in turn in the subsections below, and show that: (i) measurement error due to "missing family links" is not important, and (ii) measurement error due to "missing members", while existent, does not change before and after the APE reform, enabling us to adjust our estimates accordingly.

C.2.1 Missing family links

The census microdata is made available by the French Institute of Statistics in two types of files: the main file (*Exploitation Principale*, henceforth EP), that comprises information about a limited set of questions for the entire population, and a supplementary file (*Exploitation Complémentaire*, henceforth EC), that provides detailed information on a number of topics but covers only a subsample of about 30 percent of households. The detailed information about family structure, that enables us to link each child in the household to her mother (and thus establish her birth rank), is only available in the EC, and therefore for a limited subsample. In order to work with the full sample, we predict sibling relationships based on the following rule:

1. We drop all households where there are no individuals aged 15 to 30 years old (households with younger kids only are not relevant to our analysis since information on educational attainment is only available starting at the age of 15). Even though our interest is on youth aged 15 to 20 years old, we keep households with older children in order to establish birth order.

2. We also drop households with individuals aged 15 to 30 where these individuals are the household heads. These are youth that already left their parents' household, and thus we cannot observe their rank in the sibship.

3. Among individuals aged 15 to 30 whose age difference with the household head is between 15 and 49 years old, we assign birth order based on their date of birth.

Given that the EP does not provide the household head's spouse identifier either, we identify the potential mother (and/or father) of these individuals based on the following criteria. Either (i) she is the household head and she is above 30 years of age, or (ii) she is not the household head, but she is above 30 years of age, has an age difference with the household head that is no more than 15 years, and she is the opposite gender of the household head. There were a few cases (between one and two percent of the households) where this rule identified more than one mother or father in the household. We corrected the cases of multiples mothers or fathers as follows: (i) if one of them is the household head, keep that one; (ii) if none of them is the household head, but one is closer in age to the household head, keep that one; (iii) if none of them is the head of the household and all of them have the same age difference, then keep the one that appears first in order in the household members enumeration.

Although the use of this constructed measure of family links might introduce some error,[38] we can assess the accuracy of our constructed family link, using the EC subsample that contains family relationships. Table C2 compares children's birth order determined from the true family link with the one arising from our constructed measure, for younger children (aged 15 to 17 years old) and older children (aged 18 to 20). As we can see, the probability that someone who is predicted to be rank 1 based on our constructed variable is truly rank 1 is 99 percent. For second and third or higher order children it is 98 and 97 percent, respectively. Moreover, these probabilities do not differ between younger (ages 15-17) and slightly older (ages 18-20) children, implying that birth rank misclassification due to missing family links is not a major concern.

C.2.2 Missing family members

Measuring birth rank from children still in the household will not reflect the true rank when older siblings have left the household. For instance, in the case of two siblings where the first-born is no longer co-residing in the household, what we "observe" as a first-born child using census data (i.e. the eldest child in the household) is in fact

[38]For instance, in cases of recomposed family, youth leaving in the same household may actually be half-sisters or half-brothers and not share the same mother. Or, in households with more than one adult woman, the youth that we classify as siblings could in fact be cousins.

Table C2: Birth rank measurement error due to missing links

True birth order	Birth order without links		
	1	2	3+
Ages 15-17			
1	99.2	1.5	0.2
2	0.8	97.7	3.0
3+	0.0	0.8	96.8
Ages 18-20			
1	99.1	1.9	0.2
2	0.9	96.7	3.2
3+	0.1	1.4	96.6

Note: The table compares birth order as constructed from our main data source (*Exploitation principale*) without direct measure of family links, to the birth order observed in supplementary data (*Exploitation complémentaire*). Source: French Population Census, EARs 2008-2016.

the second one. In this section, we explain in detail how we measure the extent of birth rank misclassification using auxiliary data. Table C3 presents a matrix with the probabilities of birth rank misclassification, based on data from the Family Survey (*Enquête Famille*, henceforth EF).[39] We construct a measure of "observed" birth rank among children co-residing with their parents (following the same steps we follow with the census data) and compare it to the "true" rank for individuals in the relevant age ranges.

Each column reports the probability of being the true first, second, or higher order in the sibship conditional on being *observed* as first, second or higher order in the household, for different cohorts (each cohort observed at a different age). For individuals born in 1995 who are observed as first in their sibship by the age of 15, the probability that their true birth rank is one is 74 percent, while in 17 percent

[39] This survey gathers information about fertility and families' structure on a nationally representative sample of about 400,000 individuals aged 18 and over. The EF is conducted about every ten years, with the last two rounds taking place in 1999 and 2011. Its main advantage for our analysis is that it records women's complete birth history, and reports the year of birth of all children of the respondent (or the respondent's spouse), regardless of whether they reside in their parents' household or not.

of the cases they are actually the second in the sibship, and 9 percent of the times they are rank three or higher. For those observed as second in their household, their rank is correctly measured in 75 percent of the cases, while for those observed as third or higher, it is correct 93 percent the times. As expected, the probability of misclassification increases with age: while there is a 26 percent chance that someone aged 15 years old (born in 1995) whom we observe as first-born is actually the second or higher order child, the probability increases to 39 percent for someone aged 20 (born in 1990). This is not surprising since the older a person is, the more likely it is that she has older siblings who left the household.

Table C3: Measurement error (missing members and missing family links). True rank conditional on observed rank

Age	Observed rank	Cohort	True birth order			Cohort	True birth order		
			1st	2nd	3rd+		1st	2nd	3rd+
20	1st	1990	61%	24%	15%	1978	60%	28%	12%
	2nd		3%	64%	33%		3%	68%	28%
	3rd+		2%	5%	93%		0%	12%	88%
19	1st	1991	64%	21%	15%	1979	61%	26%	13%
	2nd		4%	66%	30%		3%	68%	30%
	3rd+		1%	5%	94%		1%	7%	92%
18	1st	1992	68%	21%	11%	1980	63%	23%	14%
	2nd		4%	71%	25%		3%	66%	32%
	3rd+		1%	4%	95%		1%	5%	93%
17	1st	1993	67%	21%	12%	1981	64%	23%	13%
	2nd		4%	73%	23%		2%	67%	31%
	3rd		1%	8%	91%		1%	6%	94%
16	1st	1994	70%	20%	10%	1982	69%	21%	11%
	2nd		4%	74%	22%		3%	71%	26%
	3rd+		2%	10%	89%		1%	4%	95%
15	1st	1995	74%	17%	9%	1983	75%	16%	8%
	2nd		4%	75%	21%		3%	76%	21%
	3rd+		1%	6%	93%		1%	4%	95%

Note: This table compares children's true birth order with their "observed" birth order (based on the rank among children still in the household who are considered as siblings based on the constructed family link). Cohorts 1990 to 1995 are observed in 2011, while cohorts 1978 to 1983 are observed in 1999. Source: *Enquête Famille* 1999 and 2011.

However, what really matters for our identification strategy, which compares outcomes of children of different ranks born before and after the reform of the APE, is whether the probability of birth rank misclassification changed before and after.[40] Given that the EF2011 data is collected at a given point in time, we cannot compare different cohorts at the same age. For this reason, we complement the analysis using data from the EF1999. Although it means that we are comparing cohorts 12 years apart, it is still useful to detect potential changes in birth rank misclassification after the APE reform. The last three columns in Table C3 report the probabilities of misclassification for individuals aged 15 to 20 in 1999 (born between 1978 and 1983). As we can see, they do not differ much: a 15 year-old individual observed as rank 2 in the household has a 76 percent chance of being actually rank two if born in 1983, compared to 75 percent if born in 1995. The proportions are similar in 2011 and 1999 for other birth orders as well. Since the 1995 cohort is the only one observed in this data source that was born after the APE reform, we cannot do the same before-after comparison for older individuals (all of them were born before).[41] Nevertheless, it is reassuring that we do not detect any strong pattern that would suggest differential pre-trends in the probability of misclassification for youth in this age range.[42]

[40] See Subsection C.3 for more details.

[41] We exclude children born in 1994 as we do not have information on the month of birth to know if they were born after July 1 or not.

[42] We perform the same analysis of birth order misclassification separately for urban and rural areas. A priori, we could expect cohabitation with parents to be more prevalent in urban than in rural areas, for instance due to a larger offer of higher education programs or job opportunities. In that case, birth order misclassification should be lower in urban areas. The results in Table C4 confirm this intuition, with a misclassification probability that is smaller in urban than rural areas (for instance, the probability that a child born in 1990 observed as rank one is actually rank one is 63 percent in urban areas, compared to 56 percent in rural areas).

Table C4: Measurement error - Missing members + Missing family links - True rank conditional on observed rank, EF2011

Cohort	Observed birth order	True birth order (EF2011) Urban 1st	2nd	3rd	Rural 1st	2nd	3rd
1990	1st	63%	23%	14%	56%	27%	17%
	2nd	3%	64%	33%	2%	62%	36%
	3rd	1%	5%	93%	5%	4%	91%
1991	1st	65%	20%	15%	61%	24%	15%
	2nd	4%	66%	30%	4%	68%	29%
	3rd	0%	4%	96%	7%	11%	82%
1992	1st	68%	21%	11%	67%	22%	11%
	2nd	4%	72%	24%	4%	67%	29%
	3rd	1%	5%	94%	0%	3%	97%
1993	1st	69%	19%	12%	63%	24%	13%
	2nd	3%	74%	23%	7%	67%	26%
	3rd	1%	7%	92%	0%	12%	88%
1994	1st	71%	19%	10%	67%	23%	10%
	2nd	4%	74%	22%	3%	74%	23%
	3rd	1%	10%	89%	2%	10%	88%
1995	1st	75%	17%	8%	71%	19%	10%
	2nd	4%	75%	22%	4%	76%	20%
	3rd	1%	6%	93%	2%	5%	93%

Note: This table compares children's true birth order with their "observed" birth order (based on the rank among children still in the household who are considered as siblings based on the constructed family link), which contains measurement error due to missing members and missing links). Source: *Enquête Famille* 2011.

C.3 Adjusting estimates by birth rank measurement error

In this section, we show how the naïve DD estimator based on misclassified birth order can be corrected to recover an unbiased estimator of the effect of APE eligibility.

The parameter of interest is

$$
\begin{aligned}
ITT \equiv & [E(y|R=2, A=1) - E(y|R=1, A=1)] \\
& - [E(y|R=2, A=0) - E(y|R=1, A=0)],
\end{aligned}
$$

where R is the true birth order, A is an indicator variable equal to 1 for cohorts born after APE extension, and y is an education outcome. Under parallel trends, ITT is interpreted as the causal impact of APE eligibility on the child's education, i.e. the intention-to-treat effect.

We do not observe R in the census data, but only a proxy R'. With this proxy, we estimate the following parameter

$$
\begin{aligned}
DD \equiv & [E(y|R'=2, A=1) - E(y|R'=1, A=1)] \\
& - [E(y|R'=2, A=0) - E(y|R'=1, A=0)],
\end{aligned}
$$

which is different from ITT, given that $R' \neq R$ (e.g., someone with $R' = 1$ could in fact have R equal to 1, 2, or 3 and above, which we denote 3+).

Denote by $D(A) \equiv E(y|R' = 2, A) - E(y|R' = 1, A)$ the first difference (across groups) for $A \in \{0, 1\}$. Though it is computed based on mismeasured birth order R', $D(A)$ is related to the true birth order as

$$
\begin{aligned}
D(A) = & \sum_{b=1}^{3+} E(y|R'=2, R=b, A) \cdot Pr(R=b|R'=2, A) \\
& - \sum_{b=1}^{3+} E(y|R'=1, R=b, A) \cdot Pr(R=b|R'=1, A).
\end{aligned}
$$

We make two assumptions:

(i) Misclassification is stable over the period:

$$
\Pr(R = b|R' = b', A) = \Pr(R = b|R' = b') \equiv \pi_{bb'}.
$$

(ii) Misclassification is independent from education, conditional on true birth order (in each period) implying:

$$E(y|R', R, A) = E(y|R, A).$$

Assumption (i) is testable and, as shown in Table C3, it is not rejected by the data. Assumption (ii) may seem questionable: for instance, if older siblings leave the parents' household earlier when they stop school earlier, and if educational attainment is correlated across siblings, then birth rank will be more frequently underestimated for less educated youth in our sample. We use assumption (ii) first as it simplifies exposition greatly, but show below that it can be replaced by the milder assumption of the absence of changes in the correlation between education and misclassification.

Under assumptions (i) and (ii), $D(A)$ can be rewritten as:

$$\begin{aligned} D(A) = &\{E(y|R = 2, A) - E(y|R = 1, A)\} \cdot (\pi_{11} - \pi_{12}) \\ &+ \{E(y|R = 2, A) - E(y|R = 3+, A)\} \cdot (\pi_{31} - \pi_{32}), \end{aligned}$$

where we have also used the fact that $\pi_{12} + \pi_{22} + \pi_{32} = 1$ and $\pi_{11} + \pi_{21} + \pi_{31} = 1$.

Taking the second difference (over time) between $D(1)$ and $D(0)$ yields

$$\begin{aligned} DD =& D(1) - D(0) \\ =& ITT_1 \cdot (\pi_{11} - \pi_{12}) + ITT_3 \cdot (\pi_{31} - \pi_{32}) \\ =& (\pi_{11} - \pi_{12} + \pi_{31} - \pi_{32}) ITT \end{aligned} \tag{C1}$$

where $ITT_b \equiv \{E(y|R = 2, A = 1) - E(y|R = 2, A = 0)\} - \{E(y|R = b, A = 1) - E(y|R = b, A = 0)\}$; under parallel trends, $ITT_1 = ITT_3 \equiv ITT$ yielding the third equality. Equation C1 shows that the intention-to-treat parameter can be recovered be re-inflating the naïve DD estimator according to the misclassification probabilities:

$$ITT = \frac{DD}{(\pi_{11} - \pi_{12} + \pi_{31} - \pi_{32})}.$$

Each of the probabilities $\pi_{bb'}$ is estimated using *Enquête Famille* (reported in Table C3). As an example, for someone aged 15 years old (age for on-time middle school certificate), $\pi_{11} - \pi_{12} + \pi_{31} - \pi_{32} = .74 - .04 + .09 - .21 = .58$, meaning we need to multiply the diff-in-diff estimates by $1/.58$ in order to recover the true effect of APE eligibility on the probability of getting the middle HS certificate on time.

Note that the measurement error also affects precision, so the standard errors need to be adjusted as:

$$Var(ITT) = \frac{Var(DD)}{(\pi_{11} - \pi_{12} + \pi_{31} - \pi_{32})^2}$$

If assumption (ii) does not hold, $D(A)$ can be rewritten as

$$\begin{aligned} D(A) = &\{E(y|R = 2, R' = 2, A) - E(y|R = 1, R' = 1, A)\} \cdot (\pi_{11} - \pi_{12}) \\ &+ \{E(y|R = 2, R' = 2, A) - E(y|R = 3+, R' = 2, A)\} \cdot (\pi_{31} - \pi_{32}) \\ &- \pi_{21}\{E(y|R = 2, R' = 2, A) - E(y|R = 2, R' = 1, A)\} \\ &+ \pi_{12}\{E(y|R = 1, R' = 2, A) - E(y|R = 1, R' = 1, A)\} \\ &+ \pi_{31}\{E(y|R = 3, R' = 1, A) - E(y|R = 3, R' = 2, A)\}, \end{aligned}$$

where the last three lines capture the fact that educational attainment may be correlated with birth rank misclassification. As a consequence, equation C1 becomes

$$\begin{aligned} DD = &ITT \cdot (\pi_{11} - \pi_{12} + \pi_{31} - \pi_{32}) \\ &- \pi_{21}d_{21} + \pi_{12}d_{12} + \pi_{31}d_{31}, \end{aligned}$$

where the terms $d_{bb'} \equiv \{E(y|R = b, R' = b', A = 1) - E(y|R = b, R' = b', A = 1)\} - \{E(y|R = b, R' = b', A = 0) - E(y|R = b, R' = b', A = 0)\}$ measure changes in the correlation between youth's education and misclassification. Given the stability of misclassification documented in Table C3, these terms are likely to be small at best; they become negligible when weighted by misclassification probabilities $\pi_{bb'}$ that are below 20% in practice. Equation C1 therefore remains a valid approximation.