Academic performance, Educational Trajectories and the Persistence of Date of Birth Effects. Evidence from France*

Julien Grenet
(*CEP, London School of Economics*)

September 2010

Abstract

This paper performs a precise quantification of date of birth effects throughout individuals' schooling and working life using high-quality French data. With respect to previous studies, our paper innovates by controlling explicitly for the confounding effects that may arise from the correlation between birth seasonality and socio-economic background. Our results indicate that although the test score gap between December and January-born pupils tends to narrow as they grow older, it remains significant until the end of secondary education. In the French context, we show that the combination of extensive use of grade retention and early school tracking is particularly harmful for the youngest pupils in their academic cohort. Being born in December rather than in January doubles the probability of being held back in school and the negative signal associated with grade retention has a strong influence on the type of senior secondary school that students attend after completing junior high school. We argue that both these institutional features explain why date of birth effects persist in adulthood. Our estimates reveal that women and men born in December have a higher probability of holding a vocational qualification and a lower probability of holding an academic qualification than those born in January. Furthermore, men born at the end of the year incur a small but significant penalty in terms of labour market outcomes, in the form of lower wages and higher unemployment rates.

*Contact: Centre for Economic Performance, London School of Economics, Houghton Street, London WC2A 2AE. E-mail: j.a.grenet(at)lse.ac.uk.

1 Introduction

The relationship between pupil performance and date of birth is a well documented fact. A large number of empirical studies carried out using both national and international data have shown that the youngest pupils in their school cohorts tend to have lower academic achievements than their older peers. Yet the existing evidence has not not unambiguously established the mechanisms that underlie date of birth effects. In particular, little attention has been given to the institutional features of educational systems through which these effects can propagate into adulthood. The present study aims at contributing to fill this gap by taking advantage of high-quality French data that allows us to follow closely the impact of date of birth on individual outcomes throughout schooling and working life.

Potential effects of date of birth Following Crawford et al. (2007), five distinct mechanisms could a priori explain why educational and labour market outcomes are affected by date of birth. (1) Age of starting school effect: individuals born in different dates of the year typically start school at different ages. This is because the school year typically begins at the same date for all pupils so the youngest in their school cohort are also be younger when school starts. They might therefore be less prepared for learning than their older peers. (2) Absolute age or "maturity" effect: because examinations usually take place on the same date for all pupils in a given grade level, the lower performance of the youngest in their cohort could simply come from the fact that their intellectual maturity is less developed on the day of the test. (3) Age position effect: younger children could suffer from the fact that they are younger relative to their peers (relative age effect), which would imply that if even if all pupils were tested on their birthday, the younger would still obtain lower scores than their older peers. (4) Length of schooling effect: individuals born in different months of the academic year might not have spent the same time in school. Two institutional channels could explain this phenomenon. First, some educational systems provide multiple entry points for admission in primary school. These entry points are usually defined so that the older children in their academic cohort start school earlier than their younger peers[1]. The consequence of such admission policies is that individuals born late in the academic year will usually have spent less time in primary school than those born earlier. The second institutional feature that may induce different lengths of schooling by date of birth relates to compulsory schooling laws. In most countries, students have to stay in school until their reach the compulsory schooling age. In France, for instance, students have to wait until their 16$^{\text{th}}$ birthday before they can drop out of school. Angrist and Krueger (1991) first pointed out that this mechanism could induce the youngest individuals in their academic cohort to have slightly more schooling on average then their older peers. (5) Educational trajectories effect: date of birth may have lasting effects on individual outcomes by determining different educational trajectories. Because of their

[1] In the UK, for instance, admission policies vary across local education authorities. Some provide a single entry point, so that all children start school in the September of the academic year in which they turn 5. Others, however, provide two or three entry points that depend on the child's month of birth.

lower average academic performance, the youngest pupils in their cohorts have a higher probability of being held back in school. They might also be more often tracked into less academically-oriented studies, which later in life may affect their labour market outcomes.

Empirical challenges Although the different mechanisms underlying date of birth effects are fairly well identified, the assessment of their respective importance raises a number of empirical challenges.

A first major difficulty is that the different effects of date of birth are hard to disentangle. Indeed, for individuals who are still enrolled in school, there is an exact linear dependence between age at school start, age at test and time spent in school. In most empirical contexts, this feature hinders the identification of these factors' impact on pupil performance.

Sample selection issues are a second major concern when attempting to estimate date of birth effects, especially when considering educational outcomes. Two distinct mechanisms are involved here. First, pupils do not always start school the year they are supposed to and early or late school enrollment is likely to be correlated with a child's date of birth[2]. Second, the younger pupils in their academic cohort are more often held back in school than their older peers. As a result of these two mechanisms, pupils belonging to the same academic cohort are often distributed differently across grade-levels depending on their month of birth. In this situation, the observed impact of date of birth on test scores in a given grade level reflects not only the impact of age on academic performance, but also the effect of grade retention and of time spent in school.

A third empirical concern arises from the fact that the age at which pupils are tested in a given grade level is endogenous. The endogeneity of age in the test score equation is driven by the same two factors that we just mentioned. The age at which a pupil is tested in a particular grade level depends on whether she started school at the normal age or not, which itself is correlated with ability and other unobservable determinants of test scores. The age at test also depends on whether a pupil was held back or not, which again is correlated with ability. Hence both the age of school entry and the grade retention status induce a negative correlation between age at test and the error term in the test score equation, inducing a downward bias in the estimation of how age differences affect academic performance.

Finally, the estimation of date of birth effects faces the important question of whether or not date of birth is a randomly assigned variable. The concern here is that the timing of births could vary across different socio-economic groups, thereby inducing a spurious correlation between date of birth and individual outcomes.

Existing evidence A number of estimation strategies have been proposed in the literature to address some of these empirical challenges and provide a quantitative assessment of the channels through which date of birth may influence educational and labour market outcomes. Most existing papers use institutional variations in

[2]This is because pupils who start school one year earlier than expected are usually born soon after the school cutoff date whereas children who start school later are generally born soon before the cutoff date.

educational systems across countries or regions to isolate the different factors that are likely to shape date of birth effects.

Existing evidence on the impact of age of school entry on educational performance is mixed. Most studies estimate the impact of school entry by exploiting local variations in the age at which pupils can start school. While some authors a small positive effect of starting school at an older age on test scores (Puhani and Weber, 2005; Datar, 2006), others find no significant effect (Mayer and Knutson, 1998; Fertig and Kluve, 2005). The problem is that these estimates could potentially reflect other determinants of test scores such as length of schooling or age at test. To avoid the estimates from being contaminated by differences in the maturity of pupils on the date of the test, other studies choose to estimate the impact of age of school entry on outcomes later in life. Observed differences in the outcomes of adult workers born in different months of the same academic year are assumed to reflect differences in the age at which these individuals started school. Using this approach on US data, Angrist and Krueger (1992) find no impact of age of school entry on post-high school education. Using Swedish data, Fredriksson et Ockert (2006) find a weak positive impact of age of school entry on the number of years of schooling but no impact on wages. The problem with this approach is that individuals born in different times of the year might have experienced different educational trajectories because to their unequal maturity on the day of the tests. This phenomenon could explain persistent differences in labour market outcomes independently from the direct effect of age of school entry.

While there is no strong evidence that the age at which individuals start school influences their educational and labour market outcomes, there is a growing body of literature suggesting that pupils born soon after the school cutoff date are significantly disadvantaged from being the youngest in the classroom. In contrast to the conclusions of earlier work on this topic, which are plagued by sample selection and endogeneity bias, recent studies have shown that maturity effects have a stronger and more persistent effect on pupil performance than is commonly believed. To overcome the methodological limitations of previous work, Bedard and Dhuey (2006) use the assigned relative age of pupils in their academic cohort as an instrument for the age at which they took the examination in the test score equation. Their estimations are performed using test score data from the TIMSS[3] survey for pupils in fourth and eighth grade levels across OECD countries. The results indicate that maturity differences have a strong and persistent impact on academic performance: the youngest members of each cohort score 0.12 to 0.35 of a standard deviation lower than the oldest members in fourth grade, and 0.08 to 0.35 of a standard deviation lower in eighth grade, depending on the country. Using data from the PISA[4] survey, in which pupils from several OECD and non-OECD countries are tested at the age of 15 independently of their grade level, Strøm (2004) finds that the December vs. January gap in test scores amounts to 0.2 of a standard deviation. One limitation of these two studies is that the impact of age differences on test scores is measured essentially for pupils who are enrolled in secondary education. The estimated effect of age on academic performance must therefore be considered as a reduced form re-

[3]Trends in International Mathematics and Science Study.

[4]Program for International Student Assessment.

flecting not only the unequal maturity of students born at different times of the year, but also their potentially different educational trajectories (through grade retention and tracking). Moreover, these studies do not provide a quantitative assessment of the rate at which maturity effects fade out as pupils grow older. A precise evaluation of the impact of age on test scores throughout the school curriculum is carried out by by Crawford et al. (2007) using UK data. The main advantage of the British context is that grade retention is seldom used in the UK, which enables the authors to avoid most of the sample selection and endogeneity issues that we discussed earlier. Their estimates indicate that an eleven months age difference translates into a test score gap of 0.8 of a standard deviation at the age of 5. This gap narrows as pupils grow older but remains significant and equal to 0.12 of a standard deviation for 15-year-old students. Crawford et al. also provide some evidence that absolute age is the main driving force behind the lower performance of younger pupils and that their age position in the classroom has no additional effect on their test scores[5].

In contrast to the strong empirical evidence indicating that age differences have a significant impact on pupil performance, relatively little is known about the exact mechanisms through which date of birth effects might propagate themselves into adulthood. Taken alone, maturity effects seem unlikely to explain the persistence of date of birth effects once individuals have entered the labour market. Existing studies suggest that it is rather the structure of education systems that is responsible for the lasting effect of date of birth. One strand of literature has emphasized the role of compulsory schooling laws in explaining why individuals born at different times of the year differ in a number of outcomes. In their famous 1991 paper, Angrist and Krueger argue that in the US, season of birth is related to educational attainment because individuals born at the beginning of the year can drop out of school before those born at the end, simply because they reach the minimum school leaving age earlier. According to these authors, the lower average length of schooling of the oldest individuals in their academic cohort is the main explanation for their lower average earnings. Angrist and Krueger's findings have raised a number of criticisms regarding the weakness of the quarter-of-birth instrument (Staiger and Stock, 1997; Bound et al. 1995), the potential non-random assignment of date of birth (Bound and Jaeger, 2000; Buckles and Hungerman 2008) and the possibility that other mechanisms than compulsory schooling could explain the persistence of date of birth effects in adulthood. This latter concern is based on the observation that the association between date of birth and educational attainment exists in a number of countries where compulsory schooling laws cannot affect the length of schooling (because students who reach the minimum school leaving age have to stay in school until the end of the academic year irrespective of their month of birth), such as the Netherlands (Plug, 2001) or Sweden (Fredriksson and Ockert, 2006). Furthermore, recent studies suggest that other institutional features could explain the persistence of date of birth effects in adulthood. For instance, Leuven et al. (2006) show that in the Netherlands, pupils born soon after the school cutoff date spend

[5]The authors use the variation in the age composition of cohorts within a particular school to separately identify the impact of age position from that of absolute age. A pupil's age position is measured as the fraction of pupils who are older in the same school and academic cohort. This index is then included as a control variable in the test score equation.

more time in school than the younger children from the same academic cohort, because they are allowed to start school earlier and attend the first grade for more than a year. The authors show that this particular feature of the Dutch school systems has a small but positive impact on the Grade 2 test scores of disadvantaged pupils born soon after the school cutoff date. However, the data does not enable to authors to evaluate whether or not this length of schooling effect persists over time. Other studies argue that date of birth effects could propagate into adulthood by determining different educational trajectories for individuals born at different times of the year. Recent work suggests that early maturity differences could hurt the educational prospects of the youngest individuals in their academic cohorts by increasing their probability of being held back (Eide and Showalter, 2001; Martin et al., 2004) and, in the presence of early school tracking, of attending vocational rather than academic studies (Jürges and Schneider, 2004). Yet none of these studies have attempted to investigate whether this differentiation in educational careers by date of birth can also be found in labour market outcomes.

The paper's contribution The present study aims at improving our understanding of date of birth effects by shedding light on the features of education systems that explain why early relative maturity differences between pupils can have lasting effects on their educational and labour market outcomes. The French educational system provides a particularly valuable empirical setting to analyse date of birth effects, since it combines both the extensive use of grade retention and the practice of early secondary school tracking – two features that are likely to affect pupils differently depending on their date of birth. Moreover, date of birth effects in France can be studied using a variety of rich data sets, allowing for a precise quantification of their impact throughout individuals' schooling and working life.

This paper makes several contributions to the existing literature on date of birth effects.

One of the key identifying assumption underlying the estimation of date of birth effects is that birth patterns are unrelated to other factors that may affect educational and labour market outcomes. Yet most existing studies provide little or no evidence in support of this crucial hypothesis. In contrast, our paper provides a detailed investigation of this issue and shows that while the timing of births cannot be considered independent from observable background characteristics, the estimation biases that arise from this correlation can be appropriately neutralized in the French context.

A second important contribution of our study is to provide a precise quantification of the impact of age differences on test scores throughout the school curriculum and to evaluate whether grade retention amplifies or, on the contrary, attenuates the academic achievement gap between pupils born at different times of the year. To solve the sample selection and endogeneity issues that arise when attempting to estimate the effect of age on test scores, we use an instrumental variable strategy similar to that proposed by Bedard and Dhuey (2006). Our results indicate that maturity effects are strong in primary school and gradually narrow as pupils grow older. They nonetheless remain significant throughout secondary schooling. In Year 1, the test score gap induced by an age difference of eleven months is equal to 0.7 of a standard deviation. By the time pupils reach Year 9, the gap has fallen to 0.2 of a standard deviation. Our results also indicate that while the date of birth penalty is similar for girls and boys, pupils from disadvantaged backgrounds suffer more from being younger on the day of the test than those from more privileged backgrounds. We investigate the role of grade retention in mitigating or exacerbating the test score gap between pupils born at different times of the year by comparing our estimates with those obtained by Crawford et al. (2007) for the UK, where hardly any pupils are held back if they do not meet the academic targets. Estimates of the impact of age differences on test scores are similar in both countries and fade out at the same rate, which supports the view that grade retention neither widens nor narrows the test score gap between pupils born at different times of the year.

A third contribution of this paper is to identify the institutional channels that are most likely to explain why early maturity effects can have a lasting influence on individuals' educational and labour market outcomes. In the French context, we show that the combination of extensive use of grade retention and early school tracking is particularly harmful for the youngest pupils in their academic cohort. Our results show that being born in December rather than in January doubles the probability of being held back in primary school and that the negative signal associated with grade retention has a strong influence on the type of senior secondary school that students attend after Year 9. According to our estimates, the probability of attending vocational studies after junior high school is 3 percentage points higher for students born in December to than for those born in January and their probability of attending academic studies is reduced by a similar amount.

Finally, our study evaluates the degree to which the differentiation of educational careers affects the academic and professional prospects of individuals born at different times of the year. Our estimates indicate that although date of birth does not seem to have a significant impact on the number of years of schooling, it does influence the type of qualifications that individuals hold when they enter the labour market. In line with the empirical evidence on secondary school tracking, we find that the fraction of adults holding a vocational qualification is 2 to 3 points higher among the December-born than among the January-born. However, a closer look reveals that date of birth influences the type (vocational vs. academic) rather than the level of qualifications. This explains that despite the relatively strong differentiation in educational trajectories by date of birth, we find that the youngest individuals in their academic cohorts incur only a minor penalty in terms of labour market outcomes. Their employment status only slightly negatively affected for men and their wages are only 1 to 2% lower than those of January-born workers.

Our results have a number of policy implications. First, they suggest that the age of pupils should be taken into account when assessing their academic performance. Several options are available to avoid unjustly penalizing the youngest pupils, which include the age normalisation of test scores, the introduction of a greater flexibility into the testing system (pupils being tested at different times depending on their date of birth) and the grouping of pupils by age in grade level classes. Second, our results suggest that grade retention and early school tracking have a particularly detrimental effect on the academic prospects of children who happen to be born soon before the school cutoff date. The strong perverse effects of these two educational practices should therefore be taken into account when assessing their costs and benefits.

The remainder of this paper is organized as follows. Section 2 provides some brief background on the French educational system. The estimation strategy is presented in section 3 and the data is described in section 4. Section 5 discusses possible ways to control for the the correlation between month of birth and socio-economic background. Section 6 presents the results and section 7 concludes.

2 Some background on the French school system

The French school system is highly centralized an fairly homogenous until pupils reach the age of 14. Education is predominantly public, public schools accounting for 86% of primary school enrollment and 79% of secondary school enrollment in 2005.

Schooling is compulsory from the age of 6. Pupils enter primary school in September of the year they turn 6, which means that a school cohort is formed of all children born between January 1 and December 31 of a given year. Compulsory education extends until the age of 16 and students can leave school on their 16$^{\text{th}}$ birthday.

The organization of primary and secondary education is schematically presented in figure 1. Pupils spend 5 years in primary school (age 6 to 10), 4 years in junior high school school or *Collège* (age 11 to 14) and 3 years in senior high school or *Lycée* (age 15 to 17). The primary and junior secondary curriculum is common to all pupils. At the end of junior high school, Year 9 pupils prepare for the junior high school certificate (*Diplôme National du Brevet*) which comprises a continuous assessment component and a final examination component. Senior secondary education is divided into two tracks: academic and vocational. The vocational track leads after two years to a secondary schooling vocational qualification and, after two further years, to a vocational A-level (*Baccalauréat professionnel*). The academic track includes three years of senior high school during which students prepare the academic A-level (*Baccalauréat*) examinations that take place at the end of Year 10 in French and at the end of Year 11 in all other subjects.

Two specific features of the French educational system are worth emphasizing as they will be shown to play an important role in the shaping date of birth effects: grade retention and secondary school tacking. Grade retention is commonly practiced and about 50% of French pupils are held back once or twice during their primary and secondary education. One would expect this fraction to be higher among pupils born late in the year than among those born earlier. Secondary school tracking takes place once students have completed junior secondary schooling. In the third term of Year 9, the school's board of teachers makes a recommandation on whether a pupil should attend a vocational or an academic senior high school[6]. Parents can choose to appeal against this recommendation in front of the school's principal, who makes the final decision. In case of a persistent disagreement, parents

[6] Note that continuation of schooling in an academic senior high school is not dependent on the award of the junior high school certificate.

have the possibility of having their child repeat Year 9 for one year at the maximum[7]. Boards of teachers base their recommendation mainly on past academic performance. However, they also use a number of other criteria which include absenteeism, discipline problems and past grade retention. the empirical part of our study suggests that this latter factor plays an important role in explaining why students born late in the year have a higher probability of being tracked into vocational high schools than those born earlier.

3 Estimation strategy

In this section, we present our empirical framework for estimating the impact of date of birth on educational outcomes. Subsection 3.1 discusses the instrumental variable strategy that we implement to evaluate the effect of age differences on test scores, while subsection 3.2 explains how we measure the impact of date of birth on educational trajectories and labour market outcomes.

3.1 Measuring the impact of age differences on test scores

The fact that pupils belonging to the same school cohort perform differently because of their age differences at the time of the tests has been widely documented. Most existing studies, however, tend to underestimate the impact of absolute age on test scores and overstate the rate at which the penalty faced by the youngest pupils in their cohort narrows as they grow older. This is because they often ignore the selection and endogeneity issues that arise when attempting to evaluate the effect of age on pupil performance.

Limitations of some existing empirical strategies Three distinct empirical approaches have been commonly used in the literature to measure date of birth effects on pupil performance. For the sake of clarity, we will consider here that a school cohort comprises all pupils who are born between January and December of a given calendar year.

The first approach consists in regressing the test scores of all pupils that belong to the same grade level on their month of birth to estimate the specific penalty incurred by the youngest children in their cohort with respect to the oldest. The problem with this approach it that the observed age difference between January and December born pupils is likely to to be smaller than their theoretical age difference (11 months). In this case, and provided that biological age differences are the main driving force explaining the higher average performance of older pupils, the OLS regression of test scores on month of birth would generate downward biased estimates.

There are two main reasons why the observed age differences between pupils of the same class tend to be lower than their theoretical age differences. First, a small number of pupils start school before or after the normal age and these pupils' months of birth are not uniformly distributed over the calendar year. Pupils who

[7] Art. D 331-37 of the *Code de l'Éducation*.

start school at a younger age than required by law are usually born towards the beginning of the year whereas those who start school at an older age are more likely to be born towards the end of the year. This is because parents who choose to enroll their children in primary school before the year they turn six will usually do so as long as these children are not "too young" in comparison to their classmates, which is only possible if they are born early in the calendar year. The converse is true for parents who choose to postpone the enrolment of their child.

The second and main reason why observed age differences are likely to be lower than theoretical age differences is that the probability of repeating a year is higher for pupils born late in the year than for pupils born early, as will be shown later in this study. This phenomenon tends to reduce the relative age differences between pupils enrolled in the same grade level. The symmetric case of pupils who skip a grade level also tends to reduce within-class age differences, since grade skippers are more likely to be born at the beginning of the year. An important implication of the correlation between month of birth, grade retention and grade skipping is therefore that observed age differences between December and January born pupils in the same class narrow as one moves from earlier to later school years. Hence the above methodology does not only lead to underestimate the impact of absolute age differences on pupil performance. It also overstates the rate at which these effects disappear as pupils grow older.

To overcome these limitations, an alternative estimation strategy has been used in the literature. It consists in restricting the sample to pupils who started school at the normal age and neither repeated nor skipped a grade level, and to run an OLS regression of test scores on month of birth using this sample

The problem is that estimates based on this methodology are plagued by severe sample selection bias. The main sample restriction here is the exclusion of pupils who have repeated a grade. Because these pupils are over-represented among children born late in the year, the average ability of non-repeaters tends to increase with their month of birth. Running an OLS regression of these pupils' test scores on their month of birth therefore results in downward-biased estimates of the impact of absolute age on pupil performance. Moreover, the underestimation of age effects worsens as one moves from earlier to later school grade levels.

To avoid these shortcomings, a third empirical strategy involves regressing pupils' test scores directly on their absolute age at test instead of their month of birth. The main benefit of this methodology is that it ensures that age effects are measured accurately and without facing sample selection issues. However, this approach is flawed by endogeneity bias since the age at which a pupil takes a test cannot be considered independent from other unobservable determinants of academic achievement, such as ability. On the day of the test, the oldest pupils are usually repeaters whereas the youngest are grade skippers. Because both grade retention and grade skipping are highly correlated with pupil ability, the OLS estimation of the effect of age on test scores is subject to omitted variable bias. The estimated coefficient on age at test can even become negative if a large fraction of pupils are exposed to grade retention.

An alternative IV approach In this paper, we implement an instrumental variable estimation strategy similar to that proposed by Beddard and Dhuey (2006).

We begin with the following structural equation for the relationship between the test score s_{ig} obtained in grade level g by pupil i and the absolute age a_{ig} at which she took the test (expressed in months):

$$s_{ig} = \alpha_g + \beta_g a_{ig} + \epsilon_{ig} \tag{1}$$

where β_g, the causal impact of absolute age on academic performance in grade level g, is the parameter of interest and ϵ_{ig} is the error term.

As explained above, age at test is probably correlated with the error term in equation (1) because the older pupils are repeaters whereas the younger are grade skippers.

To solve this endogeneity problem, we use "assigned relative age" as an instrument for age in the test score equation. The assigned relative age of a pupil is computed as the difference in months between the school cohort's cutoff month (December) and the pupil's month of birth. The instrument z_i is therefore a simple linear transformation of pupil i's month of birth m_i:

$$z_i = 12 - m_i$$

The values taken by this instrument range from 0 (for pupils born in December) to 11 (for pupils born in January).

We estimate parameter β in equation (1) using Two Stage Least Squares based on the following first stage equation for observed age at test:

$$a_{ig} = \gamma_g + \delta_g(12 - m_i) + \eta_{ig} \tag{2}$$

where η_i is the error term.

The reduced form relationship between test scores and assigned relative age is:

$$s_{ig} = \lambda_g + \mu_g(12 - m_i) + \nu_{ig} \tag{3}$$

where μ_g measures the impact of relative age on test scores, net of grade repetition, early and late entry, while ν_{ig} is the error term.

For this IV strategy to provide a consistent estimate of the causal impact of absolute age on test scores, two conditions must be satisfied. First, a pupil's assigned relative age must be correlated with the absolute age at which she took the test. This condition is easily satisfied since the majority of pupils enter school on time and are never retained. By construction, the correlation should be very close to one for pupils in their first year of primary school and gradually decrease after as more and more pupils experience grade retention.

The second crucial assumption is that assigned relative age (or, equivalently, month of birth) is uncorrelated with the unobservable determinants of test scores. Two factors could violate this condition. First, children born at different times of the year could have different parental backgrounds and therefore different ability levels if birth seasonality varies from one socio-economic group to another. Second, grade retention could have an independent impact on test scores. This effect, which could be either positive or negative depending on whether grade retention helps low-achieving pupils to catch up with non-repeaters or, on the contrary, harms them by imposing a form of stigma.

In the empirical part of this paper, we explore these issues in several ways. In section 5, we show that while there is evidence that individuals born at different times of the year have slightly different socio-economic backgrounds, the correlation between month of birth and parental income can be to a large extent neutralized by restricting the sample to individuals born in January and December only.

We investigate the issue of grade retention effects by comparing our estimates to those obtained by Crawford et al. (2007) for the UK, where there is hardly no grade retention. The main purpose of this exercise, which is made possible by the very similar design of school tests in France and the UK, is to compare the rate at which the estimated absolute age effects fade out in both countries as pupils grow older. If grade retention has a negative impact on the performance of French pupils, then one would expect the estimated absolute age effects to narrow at a slower rate in France than in the UK. If, on the contrary, grade retention is beneficial to pupils, then age effects should fade out more rapidly in France. This implication of grade retention effects can be easily established within the "nearly exogenous" framework used by Angrist and Krueger (1994) in their study of World War II veterans' earnings. Suppose that because of their higher exposure to grade retention, pupils born late in the year incur a specific penalty (or a specific premium) that is reflected in their test scores. Under this scenario, a pupil's month of birth would enter the right hand side of structural equation (1):

$$s_{ig} = \alpha_g + \beta_g a_{ig} + \theta_g m_i + \epsilon_{ig} \tag{4}$$

where the coefficient θ is positive if grade retention improves pupils' test scores and is negative otherwise.

It can easily be verified that if equation (4) is estimated by TSLS using $(12 - m_i)$ as an instrument for age at test a_{ig}, then the probability limit of the absolute age effect IV estimator $\hat{\beta}_g^{IV}$ is given by:

$$\text{plim } \hat{\beta}_g^{IV} = \beta_g + \frac{\theta_g}{\hat{\delta}_g^{OLS}} \tag{5}$$

where $\hat{\delta}^{OLS}$ is the coefficient from the OLS regression of a_{ig} on $(12 - m_i)$.

The IV estimator of the impact of absolute age on test scores will overestimate the true effect if grade retention improves pupil performance ($\theta_g > 0$) and will underestimate it otherwise ($\theta_g < 0$). Note however that the magnitude of the bias can only increase as pupils grow older, since $\hat{\delta}_g^{OLS}$ will decrease as observed age differences within a class shrink (because of grade retention) whereas θ_g should increase due to the growing stock of repeaters among pupils. The bias should equal zero for pupils tested at the beginning of primary education (none of them having been retained) and reach its maximum for pupils tested in their final year of schooling. The existence of a specific premium (resp. penalty) attached to grade retention should therefore imply that absolute age effects fade out at a faster (resp. slower) rate in

France than in the UK. We use this prediction in section 6 to assess whether or not one should worry that our estimates for the impact of absolute age on pupil performance are contaminated by grade retention effects.

3.2 Estimating the impact of date of birth on educational trajectories and labour market outcomes

Evaluating the impact of date of birth effect on educational trajectories and labour market outcomes is more straightforward than estimating the effect of age differences on test scores. This is because the sample selection and endogeneity issues that we discussed earlier are less of a concern.

To assess the impact of date of birth on educational trajectories, we focus on two distinct outcomes: grade retention and secondary school tracking. We measure the effect of date of birth on grade retention by comparing the fraction of repeaters among pupils who are born in different months of the same calendar year. This exercise is performed at different points in time, to determine whether or not the probability of repeating a grade is influenced by date of birth beyond primary schooling. The impact of date of birth on secondary school tracking is measured by looking at the type of studies that pupils born in different month of the same year attend after completing junior secondary education. These estimations are not plagued by sample selection issues since our data enables us to follow the educational career of either a representative sample or the complete population of pupils born in a particular year. We provide a number of robustness checks to ensure that our estimations are not contaminated by the potential correlation between month of birth and the socio-economic background of pupils.

We evaluate the impact of date of birth on educational attainment and labour market outcomes in a similar fashion. The outcome of interest y_i is regressed on month of birth and a set of control variables, which typically include year of birth and year of survey dummies. Regressions are performed separately for males and females and a number of robustness checks are provided to ensure that we are measuring the causal impact of month of birth and not the confounding effect of unobservable characteristics that could be correlated with date of birth.

4 Data

Our evaluation of the impact of date of birth on educational and labour market outcomes in France relies on a number of rich datasets that enable us to follow these effects throughout individuals' learning and working life.

4.1 Educational data

To measure date-of-birth effects on pupil performance and educational trajectories at the different stages of education, we use four distinct educational datasets.

Pupil datasets Our analysis for primary schooling relies on the *Panel Primaire de l'Éducation nationale* (PPEN). This survey was conducted by the French ministry of Education between 1997 and 2002 on a representative sample of 9,639 pupils who started primary school in September 1997. The main advantage of this survey is that it contains the individual test scores of all pupils at the beginning of Year 1 and Year 3. The PPEN also includes information on pupils socio-demographic characteristics (date of birth, gender, parental occupation, etc.) and on the characteristics of their school (public or private status, school classified as disadvantaged[8]).

Our main statistical source for secondary schooling is the *Panel Secondaire de l'Éducation nationale* (PSEN), whose design is very similar to that of the PPEN. This survey was conducted between 1995 and 2001 on a representative sample of about 17,800 pupils who started junior high school in September 1995 in a public or private establishment. The survey reports three series of test scores in Year 6, Year 9 and Year 11. The information provided on pupil and school characteristics is similar to that included in the PPEN. Note that the PSEN survey suffers from a minor limitation when it comes to evaluating date of birth effects: due to its particular sampling design, the survey does not include pupils born in March, July and October[9].

We supplement the analysis of date of birth effects on pupil performance using the *Diplôme National du Brevet* (DNB) database which contains the grades obtained by all Year 9 pupils at the 2004 session of the junior high school certificate final examinations. The main advantage of this dataset is its large sample size: in 2004, we know the grades obtained by roughly 813,000 Year 9 pupils. In addition to pupils' birthdates, the DNB provides a limited amount of information on their socio-demographic characteristics including nationality, gender and two-digit occupation of the household head.

Test scores The main advantage of the PPEN, PSEN and DNB datasets is that they report test scores at the different stages of primary and secondary education, from Year 1 to Year 11. To ensure comparability, we normalize the scores so that they have a mean of zero and a standard deviation of one.

Pupils are tested twice in primary school, once at the beginning of Year 1 and once at the beginning of Year 3. All pupils in the PPEN dataset took the Year 1 test in October 1997. Those who were not held back in Year 1 or Year 2 took the Year 3 test in October 1999 while those who were held back took the test in October 2000. Pupils tested in Year 1 received a global score summarizing their performance in five different areas of skills[10]. In Year 3, pupils received a global score in French and a global score in Mathematics.

In junior high school, test scores are available in Year 6 and Year 9, i.e. the first and final years of junior high school. All pupils in the PSEN dataset were tested at the beginning of Year 6 (in October 1995) in French and Mathematics. Year 9

[8]Public schools are classified as disadvantaged if they are labelled *Zone d'Éducation Prioritaire* (Priority Education Zone). This programme was launched in 1982 to channel additional resources to schools located in disadvantaged areas and to encourage new teaching projects. For an assessment of this programme, see Bénabou et al. (2008).

[9]To reach a sampling rate of 1/40, the sample was constructed by selecting all pupils born on the 17$^{\text{th}}$ day of each month, excluding those born in March, July and October.

[10]general knowledge, language and literacy skills, quantitative skills, spatial and temporal reasoning, socio-cognitive ability.

test scores are available from the PSEN and the DNB datasets. The PSEN records the scores obtained by pupils at the continuous assessment component of the junior high school certificate in French, Mathematics and first foreign language. The results are available in 1999 (pupils who were not held back during junior high school), 2000 (pupils held back once) and 2001 (pupils held back twice). To avoid the complications induced by the few pupils who retook these tests because they repeated Year 9, we use only the scores obtained in their first attempt. The DNB database is comprehensive and records the results of all pupils on the junior high school certificate final examinations which took place in June 2004. Scores are available separately in French, Mathematics and History/Geography. We also know pupils' overall scores on the junior high school certificate[11]. Due to its comprehensiveness, the DNB dataset allows for a very precise estimation of absolute age effects on pupil performance.

Beyond junior high school, the impact of age differences on academic achievements is more difficult to measure. The only available test scores in senior high school are the results obtained on the first part of the *Baccalauréat* examinations (in French) by the subsample of students in the PSEN panel who were not held back between Year 6 and Year 11 and were admitted into academic senior high schools after completing junior secondary education. Hence, contrary to primary and junior high school test scores, Year 11 examination results are potentially subject to selection bias[12]. This point should be kept in mind when interpreting our results.

Educational trajectories In addition to test scores, our data enables us to investigate the impact of date of birth on educational trajectories. To study grade retention in primary school, we use the PPEN panel dataset. Grade retention and tracking in secondary education are analyzed using the 2000-2005 *Scolarité* micro database, which is almost comprehensive for pupils enrolled in French private and public secondary schools[13]. For each pupil, this administrative source contains information on date of birth, gender, nationality, place of residence, occupation of household head as well as on the pupil's school and grade level in year t and year $t-1$. Although the *Scolarité* database cannot be used to construct a panel, the information provided on the school and grade of each pupil in year $t-1$ is sufficient to measure the impact of month of birth on grade retention and secondary school tracking.

[11]The overall score is computed as a weighed average of the final examination test scores in French, Mathematics and History/Geography and the continuous assessment test scores.

[12]This is both because students born towards the end of the year have a higher probability of being held back and because they are more likely to be admitted into vocational high schools (see section 6). For this reason, using the sample of non-repeaters who were admitted into academic senior high school to evaluate the impact of age differences on test scores is likely to yield downward-biased estimates.

[13]As from the 1993-1994 wave of the *Scolarité* database, the coverage rate for pupils enrolled in public schools is almost 100%. The coverage rate for pupils enrolled in private schools was small until 1999-2000 but rose rose to 86% in 2000-2001 and reached 98% in 2002-2003.

4.2 Other data

The persistence of date of birth effects in adulthood requires statistical sources that cover individuals in working age and include the month of birth in the set of available variables. We use two datasets that fulfill these requirements: the *Enquête Emploi* and the French Census microdata samples.

The French *Enquête Emploi* (EE) is a labour force survey which is conducted every year by the INSEE (the national Statistical Institute) on a sample of approximately 60,000 households. The survey contains detailed information on the respondents' qualifications and labour market outcomes. We use the 1990-2002 waves only, since previous surveys record individual earnings in brackets rather than in continuous form. Educational attainment is measured by two variables: age left full time education and highest qualification held. Labour market outcomes include hourly wage, occupation, employment status, sector of activity, etc.

For the purpose of our study, we also utilize the public use microdata samples extracted from the 1982 and 1999 French Censuses. Although the information provided by the Census on qualifications and labour market outcomes is much scarcer that in the *Enquête Emploi*, the microdata samples have the advantage of covering one quarter of the entire French population, which permits a very high level of precision when estimating the persistence of date of birth effects in adulthood. The Census files are also particularly useful to investigate whether or not date of birth can be considered as a randomly assigned variable.

5 Is date of birth randomly assigned?

Despite its widespread use in the literature, date of birth has been increasingly criticized as a potentially defective instrument because of the correlation that might exist between the timing of births and individual observable and unobservable characteristics, which would violate the random assignment assumption and bias the estimations (Bound and Jaeger, 2000; Buckles and Hungerman 2008).

As an attempt to evaluate how exogenous an individual's date of birth is to her socio-demographic background characteristics, we use the 1982 and 1999 French Census microdata samples. Our analysis indicates that two features could a priori invalidate the use of date of birth as an instrument: the excess number of January births and the fact that birth seasonality varies with socio-economic background. However, we show that at least in the French context, both these sources of bias can to a large extent be mitigated when evaluating date of birth effects on educational and labour market outcomes.

5.1 A measurement problem: the excess number of January births

The correlation between date of birth and socio-demographic characteristics can simply arise from the mismeasurement of this variable in the subpopulation of individuals who were born in countries where registries of births are not perfectly accurate. In these countries, individuals whose exact day and month of birth is unknown are sometimes assigned a conventional day and month of birth, usually the first of January of their year of birth.

The 1999 Census indicates that this phenomenon is far from being negligible in the French case. The graphs displayed in figure 2 show the monthly deviations in the number of observed births with respect to a uniform distribution over the year, computed separately for individuals born in France and outside of France between January 1, 1900 and December 31, 1998. Graph 2(a) exhibits a particularly striking spike in January for individuals born outside of France: there are 30% more individuals born in this particular month than expected under the assumption of a uniform distribution of births over the year.

This excess number of January births represents a potentially serious threat to the identification of date of birth effects, since the socio-economic status of individuals born outside of France is on average less favorable that that of individuals born in France. This phenomenon would tend to artificially inflate the estimated impact of date of birth on educational and labour market outcomes.

This problem can nonetheless be easily overcome by restricting the sample to individuals born in France. Graph 2(b) shows that for this group of individuals, there is no evidence of an abnormal prevalence of January births. All our estimations will therefore be carried out after excluding foreign-born individuals from the samples.

5.2 The socio-economic component of birth seasonality

A more serious concern regarding the exogeneity of date of birth is the fact that an individual's date of birth could be correlated with her parents' socio-economic characteristics.

The non-uniform distribution of month of birth in the population has been abundantly documented by demographers[14]. The evolution of this distribution across decades of birth can be precisely followed using the 1982 and 1999 Census micro-data. The graphs displayed in figure 3 show how the distribution of months of births in the population deviates from a uniform distribution, separately for each decade of birth from the 1950s to the 1990s. Two factors could a priori explain the observed patterns: a non-uniform distribution of births over the calendar year and differentiated mortality rates by month of birth (especially for infants). While our data does not enable us to disentangle these two components, existing evidence seems to indicate that at least in the postwar period, the observed patterns are essentially driven by the the non-uniform distribution of births over the year[15].

Two lessons can be drawn from the the graphs in figure 3. First, the amplitude of the seasonal pattern of births has always been fairly strong, with a maximum deviation usually in the order of 6% and up to 10% in the 1970s. Second, the seasonal pattern of births has changed over the period: while in the 1950s the

[14]For the French case, see Dupaquier (1976), Saboulin (1978), Prioux (1988) and Reignier-Loilier (2004).

[15]In France, infant mortality rates have fallen below 3% since 1950 and existing studies do not show a strong association between month of birth and the probability of survival beyond the age of one (Inizan and Bouvier-Colle, 1990). The correlation between month of birth and mortality rates later in life is documented in some countries, but only for cohorts born at the beginning of the 20$^{\text{th}}$ century (Doblhammer et Vaupel, 2001).

number of birth had a tendency to decrease over the year, later decades saw the emergence of a distinctive spike of births during the second semester of the year (which corresponds to a spike of conceptions during the summer).

The analysis of the driving forces of birth seasonality are well beyond the scope of this paper. Factors as different as culture, biological cycles, climate and even psychology have been shown to exert an influence on the timing of births during the year[16]. Two important explanations for the change in the seasonal pattern of births that is observed in the 1960s and 1970s should nevertheless be mentioned. The diffusion of modern contraceptives and the increase in the number of paid holidays are considered to have contributed the emergence of a major spike of conceptions during the summer and a secondary spike in December. An important consequence of these transformations is that they could have exacerbated the socio-demographic differentiation of birth seasonality, since both the use of contraceptives and the timing of paid holidays exhibit strong variation across different socio-economic groups.

The socio-economic component of birth seasonality for the youngest cohorts can be analyzed using the 1999 Census microdata. Although the Census does not report the socio-economic status of each respondent's parents, this information can be easily inferred for children who are still living at home by using the common family identifier. To restrict the sample to individuals who are most likely to live with their parents, we focus on 0 to 18 year-olds only, i.e. children born between January 1, 1980 and December 31, 1998. Because a large fraction of these children live in female-headed households, we use the socio-economic status of the mother to examine how birth seasonality varies with parental background. The birth seasonality patterns of eight distinct socio-economic groups (farmers, teachers and professors, managers, professionals, self-employed, employees, manual workers, intermediary occupations, unemployed or economically inactive) are displayed in figure 4. These graphs show that in the French context, there is fairly strong socio-economic component to birth seasonality, with schematically four distinct patterns: farmers are more likely have their children between February and July (i.e. May to October conceptions) than during the rest of the year; mothers who are managers or have an intermediary occupation have a higher probability of conceiving their children during the summer and, in the case of managers, during the Christmas holidays; self-employed, employees, manual workers, unemployed or economically inactive mothers exhibit a weaker birth seasonality pattern with a local spike in July (i.e. conceptions in October); finally, teachers and professors stand out as the socio-economic group with both the strongest and the most distinctive birth seasonality pattern, with a 20% spike in April births. One possible explanation for this particular pattern is that by giving birth in April, female teachers and professors can benefit from a 10-week maternal leave that ends right at the beginning of the two months school summer holidays.

By inducing a spurious correlation between date of birth, pupil performance and labour market outcomes, the differentiation of birth seasonality across socio-economic groups could directly threaten the identification of date of birth effects. To get a broad idea of how large these biases are likely to be, a natural approach

[16]On these different aspects, see Bumpass et al. (1978), Marini and Hodson (1981), Lutinier (1987), Sandron (1998), Leridon (1988), Oppenheimer (1988), Retherford and Sewell (1989), Bobak and Gjonca (2001) and Goldin and Katz (2002).

consists in constructing a synthetic measure of the socio-economic component of birth seasonality. We do so by combining the 1982 and 1999 Censuses with the *Enquête Emploi* to evaluate how parental income varies depending on the month of birth of individuals born between 1960 and 1998. The 1982 Census is used to identify the parents of children born between January 1, 1960 and December 31, 1979 whereas the 1999 Census is used to identify the parents of children born between January 1, 1980 and December 31, 1998. The wages of mothers and fathers found in the Census are imputed using their 3-digits occupation code (456 distinct occupations) to assign them the average wage of their occupation in the 1990-2002 *Enquête Emploi* (computed separately for women and men). The main limitation of this methodology is that it restricts the analysis to children whose parents earn a wage, thereby excluding the self-employed, professionals, unemployed as well as the economically inactive. Another potential shortcoming of the imputation procedure is that it does not account for potential changes in the occupational wage hierarchy between the early 1980s and the 1990s. Due to the large sample size of the Census microdata, this approach nevertheless offers a fairly precise visualisation of the correlation between parental income and month of birth since 1960.

Figure 5 shows the average imputed wage of fathers and mothers of children born in the 1960s and 1970s as a function of the child's month of birth. The same information is displayed in figure 6 for children born in the 1980s and 1990s. These graphs reveal that the relationship between month of birth and parental income exhibits a distinctive sinusoidal pattern which is remarkably stable across periods and parental gender. Individuals born in April-May (i.e. conceived in July-August) have on average the highest earning parents whereas those born in August (i.e. conceived in November) have the lowest earning parents. The particular shape of this curve is in line with the previously discussed socio-economic differentiation of birth seasonality. In particular, the higher average parental wages of April and May-born individuals is consistent with the a higher birth rate of managers and the lower birth rate of manual workers and employees during that month.

5.3 Can the socio-economic component of birth seasonality be controlled for?

While date of birth cannot be considered orthogonal to socio-economic background, we believe that the biases arising from this correlation can to a large extent be controlled for.

The first point to note is that the association between month of birth and parental income is fundamentally non-linear and is roughly rotational symmetric around the calendar year's midpoint. This type of pattern seems unlikely to induce large estimation biases if date of birth effects are linear in month of birth, since the socio-economic component of month of birth would simply translate into symmetric deviations away from a linear trend that can be consistently estimated.

The second point to note from figures 5 and 6 is that the correlation between month of birth and parental income seems fairly week for the extreme months of the year (January and December) since the average wage of fathers and mothers of children born during these months does not differ greatly from the average wage.

This observation suggests that the socio-economic component of date of birth can be almost completely neutralized when the samples are restricted to January and December-born individuals. When possible, we will use this approach to test the robustness of our findings.

Finally, it must be stressed that although the socio-economic differentiation of birth seasonality is significant, it remains quite small. The graphs displayed in figures 5 and 6 indicate that the parental wage gap between different months of the year does not exceed 3.5%. A correlation of this magnitude would only hinder the estimation of small date of birth effects. Yet the analysis below shows that this is not the case, in particular when we consider the impact of date of birth on pupil performance.

6 Results

In this section, we present the results of our estimations for the impact of date of birth on three sets of outcomes: pupil performance, educational trajectories and labour market outcomes.

6.1 Absolute age differences and pupil performance

The PPEN and PSEN datasets can be used to evaluate the impact of absolute age differences on pupil performance throughout primary and secondary schooling.

Theoretical vs. observed age differences As explained in section 3, running an OLS regression of test scores on month of birth is not a suitable estimation strategy to evaluate the December birth penalty in terms of academic performance, since observed age differences between pupils belonging to the same grade level are lower than theoretical age differences.

The solid line in graph 7(a) of figure 7 shows the absolute age (in months) at which the 9,639 pupils from the PPEN panel took the Year 1 test (in October 1997) by month of birth. The dotted line shows the age at which they would have normally taken the test had they entered primary school in the year they turned six. The discrepancy between these two lines comes from the fact that a small fraction (1.5%) of pupils born in January or February actually enter primary schooling the year they turn five whereas an equally small fraction (1.5%) of pupils born in the final months of the year start primary schooling the year they turn seven. The consequence of these two phenomena is that the observed average age difference between December and January-born pupils is 9.6 months instead of an expected 11 months.

Because of the higher fraction of grade-skippers among pupils born early in the year and the higher fraction of grade-repeaters among those born late in the year, absolute age differences tend to shrink as pupils are tested later in their education. Graph 7(b) in figure 7 shows the observed vs. normal age at which the 800,000 pupils from the DNB database took the Year 9 final examinations (in June 2004), separately by month of birth. The dotted line (observed age) lies above the solid line (normal age) because of grade retention and has a smaller slope because the fraction of repeaters increases with month of birth. The observed age difference between December and January born pupils is now 6.7 months instead of 11.

The consequence of this discrepancy between observed and normal age differences on the day of the test is that an OLS regression of test scores on month of birth would underestimate the true impact of absolute age differences on pupil performance.

IV estimates To overcome this important limitation, we replace month of birth by age at the test in the test score equation and account for the endogeneity of absolute age by using assigned relative age (12-month of birth) as an instrument in equation (3).

Table 1 reports the results separately for each available test scores. The average age (in years) of pupils on the day of the test is reported in column 1. For comparative purposes, column 2 reports the OLS naive estimates of the impact of observed age at test (in months) on test scores. The fact that all of the coefficients in this column are negative except for Year 1 tests is a clear indication of the endogeneity bias that arises from the correlation between month of birth, early or postponed school entry, grade skipping and grade retention. The lower average performance of pupils born late in the year comes essentially from the fact that they have a higher probability of being repeaters. Columns 3 and 4 report the first stage and reduced form estimates of δ_g and μ_g while columns 5 to 7 report the IV estimates of β_g for alternative specifications.

The point estimates for the impact of assigned relative age on absolute age (column 3) are close to 1 in primary school (Year 1 to 3 tests) and tend to decrease as pupils are tested later in their education, since children born late in the year have an increasing probability of being behind their assigned grade level. The higher correlation between observed and assigned age for pupils tested in Year 11 than for pupils tested in Year 9 is a consequence of sample selection: Year 11 test scores are available only for pupils who were not held back in junior high school whereas earlier tests scores are available for all pupils.

Several interesting conclusions can be drawn from the IV estimates reported in column 5, which are obtained using the full sample of pupils and not including socio-demographic controls. First, the impact of absolute age on pupil performance is very large when pupils enter primary education: at the beginning of Year 1 (average age at test: 6.3 years), being a month older at the time of the test reduces test scores by 0.6 of a standard deviation, which translates into a 0.7 penalty for December vs. January born pupils[17]. Second, the effect of absolute age differences on academic achievement tends to decrease as pupils grow older but remains significant at the end of junior high school. Between Year 1 and Year 3, the December vs. January penalty is divided by two. It is again divided by two between Year 3 and Year 9 (average age: 15.4 years) but remains as high as 0.1 to 0.2 s.d. depending on the subjects tested. Our estimates also seem to indicate that age effects are still present in senior high school: although probably underestimated because of sample selection bias, the coefficient for Year 11 scores at the *Baccalauréat* written examination in French is significant at the 10% level and equal to 0.1 s.d. Third, our results suggest that age effects fade out more rapidly in Mathematics than in French or in

[17]The December vs. January penalty is computed by multiplying the point estimates by 11.

History/Geography. In Year 9, age effects are significantly smaller in Mathematics than in other subjects whereas they are not significantly different in Year 3.

Robustness checks To examine the robustness of our results to the socio-economic component of date of birth, we perform two series of tests. Column 6 of table 1 reports the IV estimates that are obtained after including controls for socio-demographic background[18] (in particular the two-digit socio-economic status of parents) and and for school characteristics[19]. The estimates are very similar to those without controls, which suggests that our results are not driven by the potentially confounding effect of socio-economic differences in birth seasonality.

Our second robustness check involves using the same specification as in column 4, but restricting the samples to pupils born in January or December. The analysis carried out in section 5 showed that the average parental income of pupils born in these months was very close to the mean income, which makes the comparison legitimate. The coefficients, which are displayed in column 6, are hardly different from those reported in columns 4 and 5 and support the conclusions drawn from our previous estimates.

Subgroup analysis Do absolute age differences affect the performance of all pupils equally? To answer this question, we perform separate estimations for different subgroups. The results are displayed in table 2.

In columns 1 and 2 are reported the IV estimates for girls and boys separately, controlling for socio-demographic and school characteristics. The coefficients indicate that the impact of age differences on academic achievement does not differ significantly by pupil gender.

The comparison of age effects across socio-economic backgrounds leads to a different conclusion. Columns 3 and 4 report the estimates obtained after splitting the initial sample of pupils into two groups of comparable size, using the household head's occupation as a measure of social status. The first subsample (labeled "‘privileged") includes pupils from middle to high socio-economic backgrounds[20] whereas the second group (labeled "‘disadvantaged") includes pupils from middle to low backgrounds[21]. Comparing the estimates in both columns reveals that in primary school, age effects are more pronounced for the less privileged pupils. The penalty incurred by younger pupils is 20 to 50% higher for the disadvantaged than for the privileged, with a significant difference for Year 3 test scores in Mathematics. The difference

[18]For samples constructed from the 1997 PPEN and 1995 PSEN datasets, the list of socio-demographic controls includes the pupil's gender and the socio-economic status of both the mother and the father. For the sample constructed from the 2004 DNB dataset, the list of controls includes the pupil's gender, the socio-economic status of the household head and a dummy variables that is equal to one for disabled students.

[19]For samples constructed from the 1997 PPEN and 1995 PSEN datasets, the list of school characteristics include urban/rural location, public/private status, classification as disadvantaged, total grade-level enrollment and year of examination. For the sample constructed from the 2004 DNB dataset, the list of controls includes the school's public/private status and the total number of Year 9 final examination candidates from that school.

[20]Self-employed, professionals, managers, intermediary occupations.

[21]farmers, employees, manual workers, unemployed and economically inactive.

seems to go the opposite way later in later stages of schooling, with the estimated coefficients on Year 9 scores being significantly higher for the disadvantaged pupils. This result should not, however, be interpreted as evidence of a more rapid fading out of age effects for pupils with lower socio-economic backgrounds. Instead, it could simply reflect the fact that because they repeat more often, these pupils are older on average than the privileged ones when they sit the Year 9 examinations. They are therefore less subject to maturity effects.

Summary Figure 8 provides a graphical summary of our estimates for the impact of age at the test on test scores. On this graph is reported the estimated impact of an 11-month age difference on test scores (computed from column 5 of table 1) at each grade level, grouped into three categories: impact on the global score, on the score in Mathematics and on the score in French. Despite the variety of sources that we use, the evolution of age effects across grade levels is remarkably consistent. Overall, age effects tend to fade out as pupils grow older but at a somewhat faster rate in Mathematics than in French.

Comparison with UK estimates Our results unambiguously show that the age at which pupils are tested significantly impacts their academic performance and that the date of birth penalty persists throughout primary and secondary education. At this stage, however, our estimates cannot distinguish the direct effect of age on pupil performance from the indirect effect of grade retention that arises from the fact that the youngest pupils in the classroom tend to repeat more often than their older classmates.

As explained in section 3, the comparison of our estimates with those reported by Crawford et al. (2007) using UK data can shed some light on the respective contributions of absolute age and grade retention effects in the French educational context. This is because almost no children are held back in the UK if they do not meet the academic targets. Comparing the rate at which age effects fade out in both countries can therefore provide an indirect assessment of the magnitude of grade retention effects.

Crawford et al. measure the impact of age differences on academic performance using the scores obtained by pupils at the various Key Stages: Foundation Stage (sat at age 5), Key Stage 1 (age 7), Key Stage 2 (age 11), Key Stage 3 (age 14), Key Stage 4 (age 16) and Key Stage 5 (age 18). Foundation stage test scores are global, Key Stage 1 scores are split in reading, writing and maths while Key Stages 2 & 3 scores are split in English, maths and science. Key Stage 4 scores are given in the form of a capped average point score that takes into account the student's eight highest GCSE grades. To achieve maximum comparability of our estimates with those reported by Crawford et al., we decided to use the coefficients associated with the following test scores: global test score in Year 1, average of the coefficients for Maths and French in Year 3 and Year 6, and global final exam score in Year 9. In the UK, the average age at the time of the test is estimated under the assumption that all pupils tested at a given grade level belong to the same school cohort (in practice, this is almost always the case). In France, the average age at the time of the test is computed directly from the PPEN, PSEN and DNB datasets.

Table 3 shows the results of this comparison and a graphical representation is displayed in figure 9. The X-axis is the average age at test (in years) and the Y-axis is the December vs. January estimated age difference penalty in terms of standardized test scores. The black triangles indicate Crawford et al.'s estimates for the UK whereas the grey squares correspond to the estimates for France. The decreasing patterns that can be observed in both countries are remarkably similar and coincide almost perfectly. This is a remarkable result, since in the presence of grade retention effects one would expect the two curves to be close for pupils under the age of 6 (all of whom are non-repeaters in both countries) but to diverge for older pupils. On the contrary, this comparison supports the view that either the impact of grade retention on test scores is small in comparison to the impact of age or that its positive and negative effects tend to neutralize each other.

6.2 Month of birth and educational trajectories

If pupils born late in the year are disadvantaged only by their lower level of maturity at the time of examinations, then one would not expect this penalty to have a persistent impact on outcomes later in life. By the time pupils finish secondary schooling, age effects seem indeed too small to have a lasting influence and their educational and labour market prospects.

The problem is that a number of institutional features of educational systems could well perpetuate date of birth effects in adulthood, by determining differentiated educational trajectories. In this section, we show that in the French context, two educational practices play a key role in carrying out the date of birth penalty: grade retention and secondary school tracking

Grade retention The correlation between month of birth and grade retention was previously discussed in relation to the estimation of the impact of age differences on pupil performance. The purpose of the following section is to quantify precisely the extent to which pupils born at different times of the year face a different probability of being held back in primary and secondary school. We do so by using both the PPEN and the *Scolarité* datasets.

We measure grade retention by age m as the probability for a pupil to have been held back one year or more by that age. The fraction of repeaters among 6 to 10 year old pupils is computed using the PPEN panel of children who entered primary school in September 1997. Repeaters are easily identified since we know their grade level every year until 2002. Computing the fraction of repeaters among 11 to 15 year old pupils is less straightforward, since the *Scolarité* database contains only those who are enrolled in secondary education[22]. For the purpose of our study, we choose to focus on the educational career of pupils born in 1989, whether they are enrolled in the public or the private sector. Among these pupils, the fraction of repeaters by age 13, 14 and 15 can be easily computed from the 2002-2003, 2003-2004 and 2004-2005 waves of the *Scolarité* dataset, since the coverage rate is close to 100%

[22]Because of data limitations, we were not able to compute the fraction of repeaters among pupils aged 16. The reason is that the coverage rate of the *Scolarité* database declines steadily for pupils enrolled in post-compulsory schooling.

for this age group[23]. In the absence of any information on the age of primary school entry, we assume that that all pupils entered Year 1 in 1993, i.e. the year they turned six. To evaluate the fraction of repeaters among 11 and 12 year old pupils, we add pupils who are missing in the 2000-2001 and 2001-2002 waves of the *Scolarité* datasets using the 2002-2003 wave which contains all pupils aged 13. Calculations are performed separately for each month of birth. All pupils added to 2000-2001 and 2001-2002 waves are considered as having been held back one year or more.

The graph displayed in figure 10 shows the fraction of repeaters among pupils aged 7, 11 and 15 as a function of their month of birth. The positive relationship between grade retention and month of birth is spectacular. By the age of 15, 51% of pupils born in December have been held back vs. 35% among those born in January. Another remarkable feature is that month of birth does not seems to influence grade retention beyond the age of 11, since the curves showing the fraction of repeaters among 11 and 15 year old pupils by month of birth are close to parallel.

To provide a more precise quantitative assessment of the dependence between grade retention and month of birth, we estimate a Probit model where the dependent variable is 1 if a pupil was held back by the age of m and where month of birth is included as a regressor. Separate regressions are performed for each distinct value of m between 7 and 15 and the results are displayed in table 4. Columns 1 reports the marginal effects evaluated at the mean and the associated December vs. January penalty is reported in column 2. The results of separate regressions by pupil's socio-economic background are reported in columns 3-4 (privileged pupils) and columns 5-6 (disadvantaged pupils). A graphical summary of the estimated December vs. January penalty is presented in figure 11.

When all pupils are included in the sample, the regressions indicate that grade retention is significantly influenced by month of birth throughout primary and secondary education but that most of the effect takes place in primary school. By the age of 11, pupils born in December have a probability of being held back which is twice as large (14 points higher) as that of pupils born in January. Furthermore, our estimates show that the impact of month of birth on grade retention rises sharply at the ages of 7, 8 and 11, which correspond to the most frequently repeated grade levels (Year 1, Year 2 and Year 5). Finally, the date of birth grade retention penalty seems to end with primary schooling, since no further impact of month of birth can be observed in junior high school.

Regressions performed separately by socio-economic background (columns 3 to 6) provide, however, a more complex picture. They reveal that until the age of 11, the marginal impact of month of birth on grade repetition is twice as large for disadvantaged pupils as it is for the more privileged. Since the fraction of repeaters among the disadvantaged is about twice the fraction of repeaters among the privileged, this result seems to be indicative of a multiplicative effect of month of birth on the factors that favor grade repetition. Furthermore, the graphic displayed in figure 11 shows that the period during which pupils are held back because of their date of birth lasts longer for the privileged than for the disadvantaged. While the December vs. January gap in retention rates is stable beyond the age of 11 for the former, it continues to increase for the latter. This phenomenon could be interpreted

[23] Pupils cannot be held back more than twice in primary school.

as a consequence of the fact that because they are more pronounced, the learning difficulties of children from lower socio-economic backgrounds are detected earlier than those of children from higher backgrounds. Since pupils are rarely held back more than once, the less privileged could therefore be better "protected" from grade retention in junior high school than the more privileged. This analysis challenges the view according to which date of birth effects cease to influence grade retention beyond elementary education. For pupils from privileged backgrounds, these effects are still being felt in secondary education.

Secondary school tracking By having a higher probability of being held back, pupils born late in the year not only face a slower progression in their studies. It is also likely that that the negative signal associated with grade retention influences the type of senior secondary school that they attend after Year 9, i.e. vocational vs. academic education.

To provide a precise assessment of the impact of date of birth on secondary school tracking, one would ideally need a panel dataset covering the entire primary and secondary schooling of French pupils. Knowing from purely cross sectional data how students from a given cohort are distributed between academic and vocational studies is indeed misleading to evaluate the impact of date of birth on educational trajectories since students born late in the year are more likely to be held back than those born early. At the age of 15, the latter would therefore be overrepresented amont students enrolled in vocational secondary schools while the former would be underrepresented simply because many of them would still be attending junior secondary education.

Although the *Scolarité* database does not allow us to follow the entire education of a panel of pupils, it can nevertheless be used to estimate the impact of date of birth on secondary school tracking. This is because each wave of the data reports both a student's grade level in both the current (t) and previous ($t-1$) academic year. Using the information contained in the 2000-2001 to 2004-2005 waves of the *Scolarité* dataset, we could find out the type of school that almost all students born in 1986 attended after finishing junior high school, whether they started Year 9 on time, one year early or one to three years late. To reconstruct the secondary school tracking of this cohort, we selected in each of the five waves of the *Scolarité* dataset all the students born in 1986, enrolled in the first year of academic or vocational secondary education during year t and enrolled in the final year junior secondary schooling during year $t-1$. Overall, we were able to infer the secondary school tracking of 82% of pupils born in 1986 (out the total number that was recorded in 2000-2001 when these students were aged 14). The remaining 18% are those who dropped out of school at the end of Year 9.

Figure 12 shows the type of senior secondary school that students born in 1986 attended after finishing junior high school, as a function of their month of birth. A striking feature of these graphs is that students' date of birth strongly influence their probability of attending an academic school rather than a vocational high school. 26.5% of those born in December attend a vocational high school after Year 9, but only 23.7% among those born in January. Conversely, only 55.2% of the December born attend academic high schools instead of 58.3% for the January born. However, there seems to be no causal link between a student's month of birth and the probability of dropping out of school after Year 9, apart from the the socio-economic component of date of birth that explains the sinusoidal pattern observed in graph 12(c). Recall that students born in the second quarter of the year come from slightly more privileged backgrounds on average than pupils born in the third quarter (see figure 6). Hence they are less likely to dropout after junior high school.

Table 5 reports the estimated impact of month of birth on secondary school tracking using a multinomial Probit regression model. Column 1 indicates the marginal effects evaluated at the mean for the entire cohort of students born in 1986 and the associated December vs. January penalty is reported in column 2. To mitigate the potentially confounding effects that arise from the correlation between date of birth and socio-economic background, we report in columns 3 and 4 the estimates where the sample is restricted to students born in January or December. The results are very similar across samples and show that the probability for students born in December to attend vocational studies after junior high school is 3 to 4 percentage points higher than for students born in January, and that their probability of attending academic studies is reduced by a similar amount. Being born in December rather than in January does not however have a significant impact on the probability of dropping out of school after Year 9.

Separate regressions for different subgroups of students are presented in table 6. The comparison of estimates for female vs. male students shows no evidence that the impact of month of birth on secondary school tracking varies by gender. This effect seem however to be twice as large for students from disadvantaged socio-economic backgrounds than for the more privileged ones. Our estimates indicate that students from low socio-economic backgrounds have a 4 percentage points lower probability of attending academic studies, a 3 points higher probability of attending vocational studies and a 1 point higher probability of dropping out of school if they are born in December rather than January.

Compulsory education and length of schooling A final channel through which date of birth might affect educational trajectories is by inducing different average lengths of schooling for students born at different times of the year.

The mechanism that is involved here relates to compulsory schooling laws and was first documented by Angrist and Kruger (1991). Since under the French compulsory schooling law, all individuals born in the same year start school at the same date but can decide to drop out of school on their 16$^{\text{th}}$ birthday, December-born students might be expected to spend on average more time in school than those born in January.

Unfortunately, our data does not enable us to produce direct evidence on the impact of compulsory schooling laws on the length of schooling, as the exact date at which individuals decided to end their studies is not provided in any of the statistical sources available. This phenomenon can nevertheless be approached indirectly using the Census microdata since individuals were asked to state whether they were still in education at the date of the interview, which took place in March of the Census year. We use this information to compute the attendance rate of all 14 to 18 year olds in the 1982 and 1999 Censuses.

The 1982 Census was conducted at the beginning of March 1982, so individuals born in January or February of year t where then aged $1982 - t$ whereas those born later in the year were aged $1982 - t + 1$. Graph 13(a) of figure 13 shows the school attendance rate in March 1982 of individuals born between 1964 and 1967 as a function of their month of birth. The impact of compulsory schooling is apparent for the 1966 cohort, since the school attendance rate drops by 4 points for those who have just turned 16. However, this graphs also reveals that school attendance rates increase with month of birth *independently* from compulsory schooling: the school attendance rate of individuals born in December is 5 to 8 percentage points higher than for those born in January of the same year. This seems to indicate that the decision to leave education is at least partly determined by biological age-related factors.

Graph 13(b) is constructed in a similar fashion using the 1999 Census. Because of the steady rise in average educational attainment since 1982, the impact of compulsory schooling can no longer be detected for individuals who are born in 1984 and turned 16 before the Census but the positive dependence between school attendance rates and month of birth remains.

These findings suggest that if individuals who are born late in the year do tend to spend a bit more time in school than individuals born early in the year, this association does not appear to be induced primarily by compulsory schooling laws. If anything, this phenomenon can be seen as partly reducing the educational penalty faced by the youngest individuals in their cohort since by delaying their entry into the labour market, they receive slightly more education.

6.3 How persistent are date of birth effects in adulthood?

The previous section has provided empirical evidence on the strong impact of date of birth on pupil performance and educational trajectories. Individuals born late in the year perform lower on the date of examinations and have a higher probability of being held back and of being tracked into vocational secondary schools. The analysis suggest however that the date of birth penalty could be partly mitigated by the fact that the youngest students in their cohorts tend to leave school later than the oldest. The question that arises then is whether the differentiation of educational careers induced by date of birth has persistent effects in adulthood, both in terms of qualifications and labour market outcomes.

By definition, the persistence of date of birth effects in adulthood can only be evaluated for earlier cohorts than those we have considered until now. Despite this time gap, we have no reasons to believe that the mechanisms that shape date of birth effects today are different from those that shaped them 30 years ago. To ensure that our sample contains only working-age adults, we choose to focus on individuals born between 1945 and 1965.

6.3.1 Month of birth and educational attainment

In relation to the previous section, we begin by investigating the impact of date of birth on educational attainment, which is measured both by terms of years of schooling and of qualifications.

Number of years of schooling Because they experience different educational trajectories, individuals born at different times of the year could also end up with different average years of schooling.

The number of years that an individual has spent in school can be proxied by the age at which she left full time education minus six. We construct our sample from the 1990 to 2002 *Enquête Emploi* by selecting all individuals who were born between 1945 and 1965 and were interviewed during the first of the three waves of the survey[24]. Table 7 reports several regression estimates for the December vs. January gap in total number of years of education, separately for women and men. The specifications in columns 1 and 4 include no control variables. Survey year and year of birth dummy variables are included in columns 2 and 3. the coefficients reported in columns 3 and 6 are estimated from the full specification when the sample is restricted to individuals born in January and December.

The regression results indicate that the impact of month of birth on the number of years of schooling is not significant for men and is weakly positive for women (in the order of one additional month of schooling for women born in December vs. January). Hence, contrary to what has been found in other countries (e.g. Fredriksson et Ockert, 2006 for Sweden), the lower average performances of French pupils born late in the year do not seem to induce them to leave school earlier than others.

Academic and vocational qualifications Although the number of years that individuals have spent in education seems mostly unrelated to their date of birth, their educational attainment could nevertheless vary in another important dimension. Individuals born late in the year have indeed a higher probability of attending vocational secondary schools and should therefore be more likely to hold the type of qualifications that these studies lead to.

To address this question, we group the highest qualification held by respondents of the *Enquête Emploi* into five categories: no qualification or primary school certificate, junior high school certificate, secondary schooling vocational qualification, academic A-level and above, vocational A-level and above[25].

Figure 14 plots the distribution of qualifications held by adults born between 1945 and 1965 as a function of their month of birth. Most strikingly, these graph show that the type of qualifications that people hold is strongly influenced by their date of birth. On the whole, people born late in the year have a higher probability of holding a vocational qualification and a lower probability of holding an academic qualifications than individuals born early in the year. The gap is largest for secondary schooling vocational qualifications, which are held by 33.5% of the December born versus 31% among the January born. Symmetrically, individuals born in December are less likely to hold a junior high school certificate or an academic A-level (17 vs. 18.5%). The probability of holding no qualification or a primary

[24] This is to ensure that each individual appears only once in the sample.

[25] In French, these five categories are: *aucun diplôme ou certificat d'études d'études primaires, BEPC seul, CAP ou BEP, Baccalauréat général ou diplôme de l'enseignement supérieur général, Baccalauréat professionnel ou diplôme de l'enseignement supérieur technique.*

school certificate[26] were does not appear however to depend on date of birth beyond its familiar socio-economic component. These observations are consistent with our previous results: since individuals who are born late in the year have a higher probability of attending vocational schools, it is not surprising to find that they are also more likely to hold the qualifications that these studies lead to.

The strength of the dependency between qualifications and month of birth can be further confirmed using the Census. Although the information on qualifications is not as detailed as in the *Enquête Emploi*, it is sufficient to evaluate the impact of month of birth on the probability of holding a junior high school certificate and on the probability of holding a secondary schooling vocational qualification, which are a priori the most sensitive to this factor. The graphs displayed in figure 15 are constructed from the 1999 Census and show the fraction of individuals holding a junior high school certificate (graph 15(a)) and the fraction of those holding a secondary schooling vocational qualification (graph 15(b)), as a function of their year and quarter of birth. These two graphs exhibit a very distinctive saw-tooth pattern, which confirms our previous findings: individuals born late in the year have a significantly higher probability of holding a vocational rather than an academic secondary school qualification.

In table 8, we use a multinomial Probit model to provide a more precise quantitative assessment of the impact of date of birth on qualifications. The specifications are identical to those in table 7 and are estimated for women and men separately using the *Enquête Emploi*. The coefficients are similar across specifications and indicate that individuals born in December have a significantly higher probability of holding a vocational (i.e. secondary school vocational qualification, vocational A-level and above) rather than an academic (i.e. junior high school certificate, academic A-level or above) qualification.

A closer look at the estimates suggests however that the date of birth penalty is more pronounced for men than for women in terms of qualifications. Indeed, the qualifications held by men born late in the year are not only more vocational but also of a lower level on average than those held by men born early in the year. Junior high school certificates and secondary school vocational qualifications have roughly the same return on the labour market, so holding the former instead of the latter is unlikely to represent a strong disadvantage. The problem for men born late in the year is that their lower probability of holding an academic A-level or above (column 3 shows a -0.022 percentage points difference between December and January born men) is not compensated by a higher probability of holding a vocational A-level or above (the December-January gap is insignificant). In contrast, the evidence seems to indicate that women born late in the year are not only more likely to hold a secondary school vocational qualification (column 6 shows a 1.8 percentage point December-January gap) but are also more likely to hold vocational A-level qualifications and above (the December-January gap being close to 1 percentage point). The situation thus seems less unfavorable for women than for men and could be explained by both the better average academic performance of female students and by the fact that they are overrepresented in many higher education studies

[26] Primary school certificates gradually had almost entirely disappeared in the early 1970s, so they account for only a very small fraction of the qualifications held by 1945-1965 cohorts.

which are classified as vocational (especially social services and nursing schools).

In light of these findings, it appears that while date of birth influences the type of qualifications that individuals hold when they enter the labour market, its impact on the level of these qualifications is only apparent for men. One would therefore expect the labour market outcomes of men born at different times to show more variation than those of women.

6.3.2 Month of birth and labour market outcomes

To study the impact of month of birth on labour market outcomes, we use the *Enquête Emploi* which provides detailed information on employment status and wages.

Employment We investigate the impact of date of birth on employment by looking at three different outcomes: unemployment, part-time work and public/private employment.

We use a Probit model to estimate the December vs. January gap for each of these variables using the sample of individuals born between 1945 and 1965. Regressions are performed separately for women and men and the results are reported in table 9 for different specifications. Columns 1 and 4 include no control variables. Survey year and year of birth controls are included in columns 2 and 5. Finally, the coefficients reported in columns 3 and 6 are estimated from the full specification when the sample is restricted to individuals born in January and December.

The estimates are very similar across specifications and suggest that month of birth has only a very limited impact on employment outcomes. December and January born individuals have essentially the same probability of working part-time and the only significant effects associated with the other variables are found for men: those born in December have a slightly higher probability of being unemployed (+0.5 percentage point) and a slightly lower probability of working in the public sector (-1 percentage point) than those born in January.

Wages Wages are the final labour market outcome that we consider in our assessment of date of birth effects. Hourly wages are computed from the *Enquête Emploi* by combining information on respondents' monthly pay and usual weekly hours of work (expressed in 2005 Euros using the consumer price index series).

We estimate the impact of month of birth on wages by running a number of OLS regressions separately for male and female workers. The results from different specifications are presented in table 10. The coefficients reported in columns 1 and 6 measure the unconditional wage gap (in logs) between workers born in December and January of the same year. They indicate that, on average, the latter earn slightly lower wages than the former and that the wage gap is higher for males (-2.3%) than for females (-0.7%). The estimates are hardly affected by the inclusion of the full set of survey year and of year of birth dummies (columns 2 and 7).

Two factors could a priori explain the observed December-January wage gap. First, individuals born at the end of the year tend to slightly delay their entry into the labour market compared to those born at the beginning of the year (see section 6.2). Second, the December-January wage gap might could partly reflect the December-January qualifications gap that we documented in section 6.3.1. In an attempt to evaluate the respective contributions of these two factors, we add the highest qualification held by the worker to the set of regressors. The inclusion of this variable reduces the wage gap by about half for male workers but leaves it unchanged for female workers. This is consistent with our previous finding that men born at the end of the year are more penalized then women in terms of their qualifications. Once the impact of date of birth on qualifications is accounted for, the December-January wage gap is in the same order of magnitude for female and male workers, i.e. about 1%. This "residual" wage gap is most likely to reflect the slightly later labour market entry of individuals who are born at the end of the year.

Given its fairly small size, the estimated December-January wage gap could well be contaminated by the socio-economic component of date of birth. To address this issue, we restrict the sample to individuals born in January and December. The regression coefficients are reported in columns 4 and 9 (without controlling for qualifications) and in columns 5 and 10 (including the set of qualification dummies). The estimated December-January wage gap is slightly higher than when using the full sample of workers and is equal to 3% for men (half of which is driven by the qualifications gap) and 1.5% for women.

The analysis carried out in this section suggests that although individuals born at different times in the year experience different educational trajectories, their labour market outcomes are only marginally influenced by their date of birth. The inconsistency here is only apparent and can be explained by the fact that date of birth influences the type rather than the level of qualification held by individuals. Furthermore, the empirical evidence reveals that male workers born in December incur a 1.5% wage penalty in comparison to the January born, and that that this wage penalty is specifically driven by their slightly lower level of qualifications.

7 Conclusion

In this study, we use a variety of microdatasets to systematically investigate the impact of date of birth on educational an labour market outcomes in the French context.

The main lesson from our empirical analysis is that a number of institutional features of educational systems play a key role in explaining why early differences in the maturity of pupils born at different times of the year can become persistent in adulthood, even after accounting for the potential non-random assignment of date of birth. In the French school system, both grade retention and early school tracking are shown to induce a strong differentiation in the educational trajectories of individuals depending on their month of birth. Being born in December rather than in January doubles the probability of being held back in school and increases by 3 points the probability of holding a vocational instead of an academic qualification. Our results also show that men born towards the end of the year incur a small but significant penalty in terms of labour market outcomes, in the form of lower wages and higher unemployment rates.

This study therefore suggests that the educational penalty incurred by the youngest pupils in their academic cohort should be viewed a serious matter and that some policy intervention is needed to ensure that individuals are not unjustly penalized by their date of birth. Several options are available, which include the age normalisation of test scores, the introduction of a greater flexibility into the testing system and the grouping of pupils by age in grade level classes. More generally, our results call into question the practice of grade retention and early school tracking, which are shown to have strong adverse effects on the educational prospects of children who happen to be born soon before the school cutoff date.

References

Angrist, J. and Krueger, A. (1991). "Does Compulsory School Attendance Affect Scooling and Earnings". *Quarterly Journal of Economics*, 106(4):979–1014.

Angrist, J. and Krueger, A. (1992). "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples". *Journal of the American Statistical Association*, 87(418):328–336.

Angrist, J. and Krueger, A. (1994). "Why Do World War II Veterans Earn More than Nonveterans?". *Journal of Labor Economics*, 12(1):74–97.

Beddard, K. and Dhuey, E. (2006). "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects". *Quarterly Journal of Economics*, 121(4):1437–1472.

Bénabou, R., Kramarz, F., and Prost, C. (2008). "The French Zones d'Éducation Prioritaire: Much Ado About Nothing". *Economics of Education Review* forthcoming.

Bobak, M. and Gjonca, A. (2001). "The Seasonality of Live Birth is Strongly Influenced by Socio-demographic Factors". *Human Reproduction*, 16(7):1512–1517.

Bound, J. and Jaeger, D. (2000). "Do Compulsory School Attendance Laws Explain the Association between Quarter of Births and Earnings?". In Bernanke, B. and Rogoff, K., editors, *Research in Labor Economics: Worker Well- Being*, pages 83–108. JAI, New York.

Buckles, K. and Hungerman, D. (2008). "Season of Birth and Later Outcomes: Old Questions, New Answers"'. NBER Working Paper 14573.

Bumpass, L., Rindfuss, R., and Janosik, R. (1978). "Age and Pace of Subsequent Fertility". *Demography*, 15:75–86.

Crawford, C., Dearden, L., and Meghir, C. (2007). "When You Are Born Matters: The Impact of Date of Birth on Child Cognitive Outcomes in England". Institute for Fiscal Studies.

Datar, A. (2006). "Does Delaying Kindergarten Entrance Give Children a Head Start?". *Economics of Education Review*, 25:43–62.

de Saboulin, M. (1978). "Un nouveau calendrier des première naissances". *Économie et Statistique*, 100.

Doblhammer, G. and Vaupel, J. (2001). "Life Span Depends on Month of Birth". *Proceedings of the National Academy of Sciences*, 98(5):2934–2939.

Dupaquier, M. (1976). "Le mouvement saisonnier des naissances en France de 1853 à 1973". Thèse de démographie.

Eide, E. and Showalter, M. (2001). “The effect of Grade Retention on Educational and Labor Market Outcomes”. *Economics of Education Review*, 20(6):563–576.

Fertig, M. and Kluve, J. (2005). “The Effect of Age at School Entry on Educational Attainment in Germany”. IZA Discussion Paper No. 1507.

Fredriksson, P. and Öckert, B. (2006). “Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance”. IFAU Working Paper 2006:12.

Goldin, C. and Katz, L. (2002). “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage”. *Journal of Political Economy*, 110(4):730–770.

Inizan, J. and Bouvier-Colle, M.-H. (1990). “Caractère saisonnier du syndrome de mort subite du nourrisson”. *Population (French edition)*, 1:9–25.

J., B., D., J., and R., B. (1995). “Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak”. *Journal of the American Statistical Association*, 40(430):443–450.

Jürges, H. and Schneider, K. (2004). “What Can Go Wrong Will Go Wrong: Birthday Effects and Early Tracking in the German School System”. CESifo working paper No. 2055.

Leridon, H. (1988). “Les conceptions du premier janvier ou les ‘étrennes de septembre’ ”. *Population*, 3.

Leuven, E., Lindhal, M., Oosterbeek, H., and Webbink, D. (2006). “Expanding schooling opportunities for 4-year-olds”’. IZA Discussion Paper No. 2434.

Lutinier, B. (1987). “La saison des mariages”. *Economie et Statistique*, 204:21–28.

Marini, M. and Hodson, P. (1981). “Effect of the Timing of Marriage and First Birth on the Spacing of Subsequent Births”. *Demography*, 18(4):529–548.

Martin, R., Foels, P., Clanton, G., and Moon, K. (2004). “Season of Birth is Related to Child Retention Rates, Achievement, and Rate of Diagnosis of Specific LD”. *Journal of Learning Disabilities*, 37(4):307–317.

Mayer, S. and Knutson, D. (1998). “Does Age at Enrollment in First Grade Affect Children’s Cognitive Test Scores?”. *Compare*, 31(3):381–391.

Oppenheimer, V. (1988). “A Theory of Marriage Timing”. *American Journal of Sociology*, 94:563–591.

Plug, E. (2001). “Season of Birth, Schooling and Earnings”. *Journal of Economic Psychology*, 22(5):641–660.

Prioux, F. (1988). “Mouvement saisonnier des naissances : influence du rang et de la légitimité dans quelques pays d’Europe occidentale”. *Population*, 3:587–609.

Puhani, P. and Weber, A. (2005). "Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany". IZA Discussion Paper No. 1827.

Retherford, R. and Sewell, W. (1989). "How Intelligence Affects Fertility". *Intelligence*, 13:169–185.

Régnier-Loilier, A. (2004). "Les naissances 'au fil des saisons' ". *Revue Européenne des Sciences Sociales*, 42(129).

Sandron, F. (1998). *"Les naissances de la pleine lune et autres curiosités démographiques"*. Éditions L'Harmattan, Paris.

Staiger, D. and Stock, J. (1997). "Instrumental Variables Regression with Weak Instruments". *Econometrica*, 65(3):557–586.

Strøm, B. (2004). "Student Achievement and Birthday Effects"'. Unpublished manuscript.

Figure 1: *The French school curriculum.*

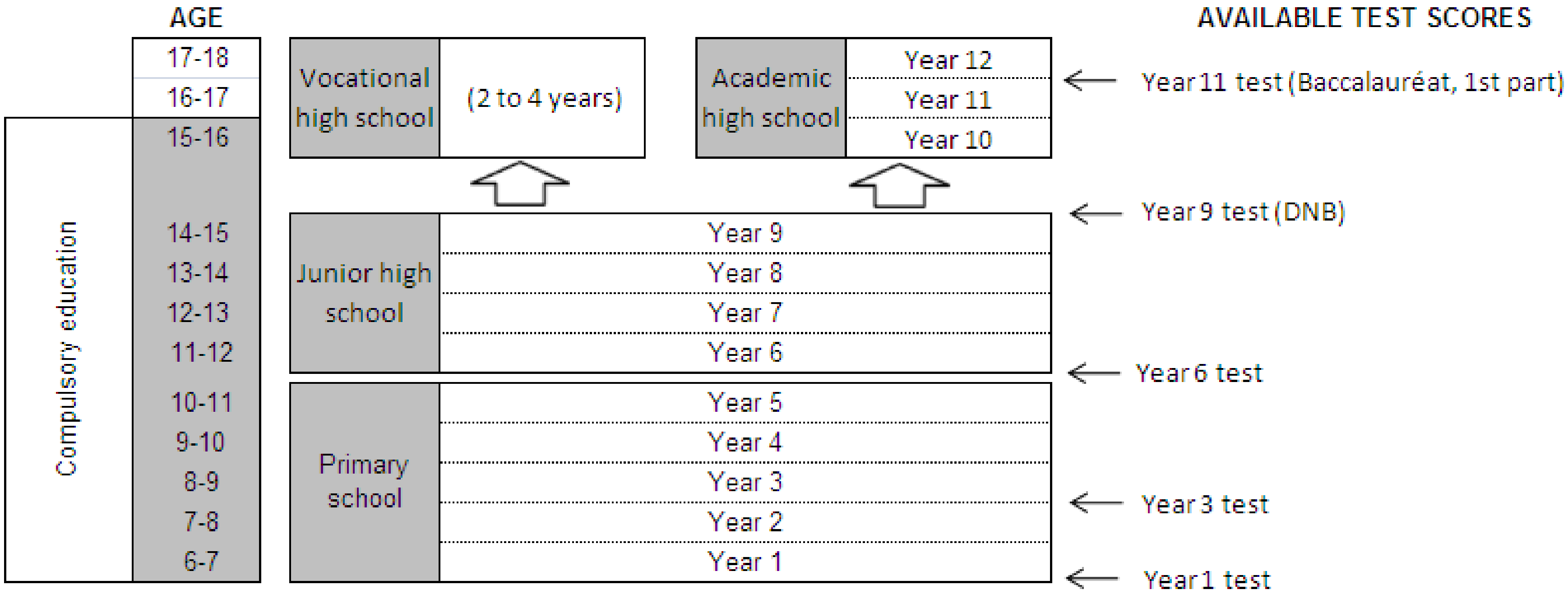
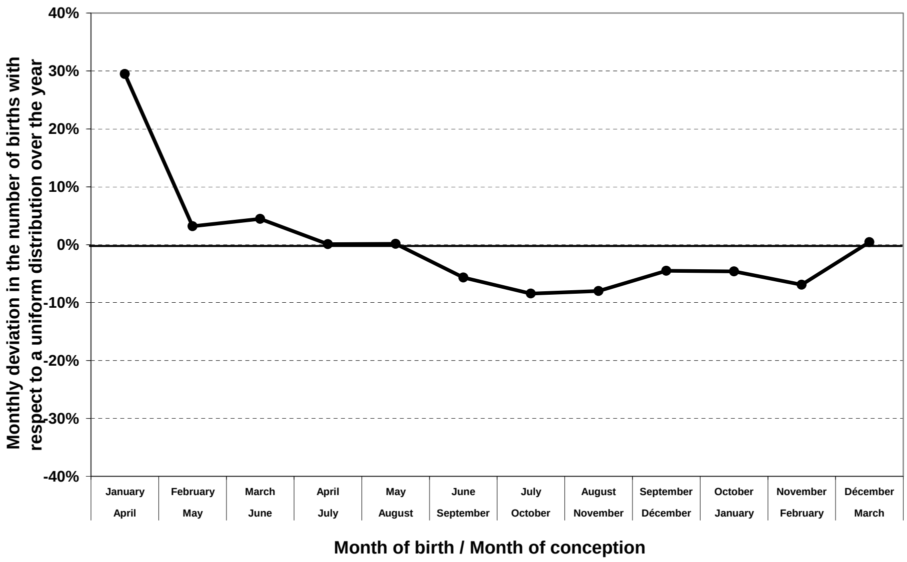
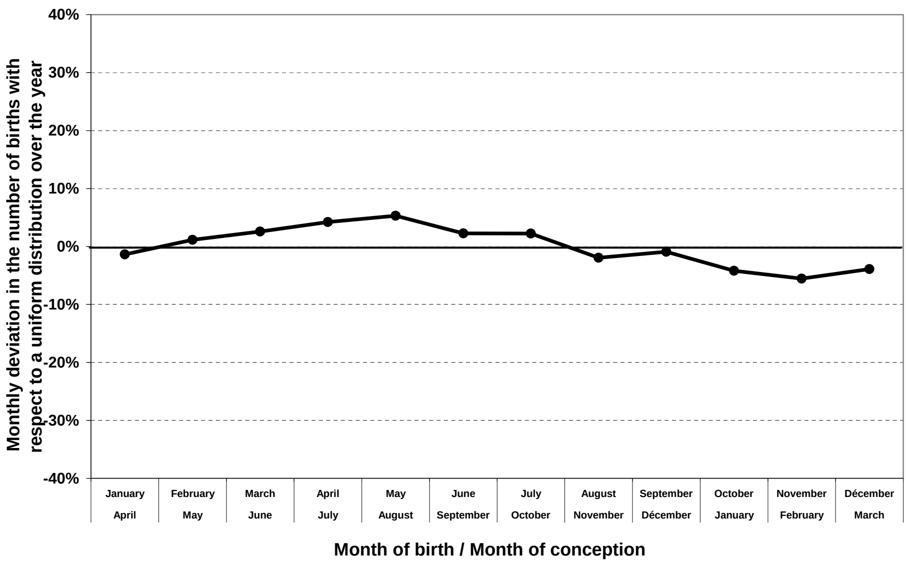


Figure 2: *Monthly deviation in the number of births with respect to a uniform distribution over the year. Individuals born in France and outside France between January 1, 1900 and December 31, 1998. Source: 1999 Census (1/4).*

(a) Individuals born outside France



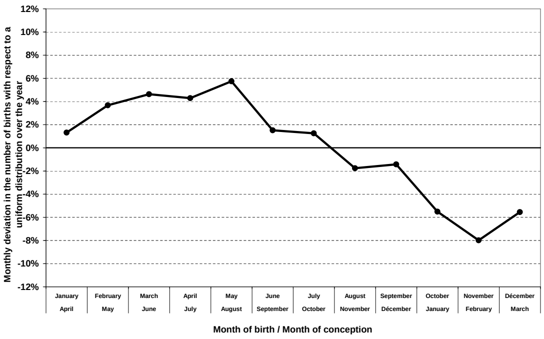
(b) Individuals born in France



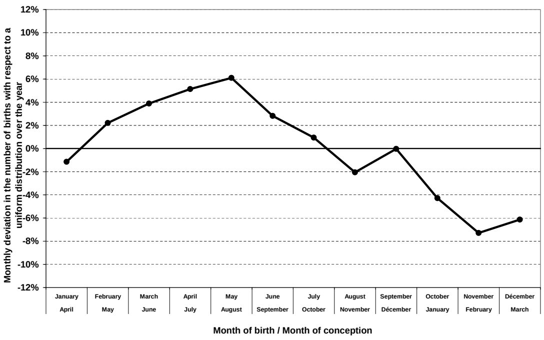
Notes: These graphs show the monthly deviation in the number of births for individuals born in France and outside France between January 1, 1900 and December 31, 1998, with respect to a uniform distribution of births over the year. Denoting λ_m the deviation associated with month m, we have: $\lambda_m = (J_m/365,2)(N_m/N)$, where J_m is the number of days in month m (28.2 for February because of leap years, 30 or 31 for other months), N_m the number of births that occurred during month m for individuals born outside France (resp. in France) and N the total number of individuals born outside France (resp. in France).

Figure 3: *Birth seasonality by decade of birth. Monthly deviation in the number of births with respect to a uniform distribution over the year. Decades 1950-1959 to 1990-1998. Source: 1982 and 1999 Censuses (1/4).*

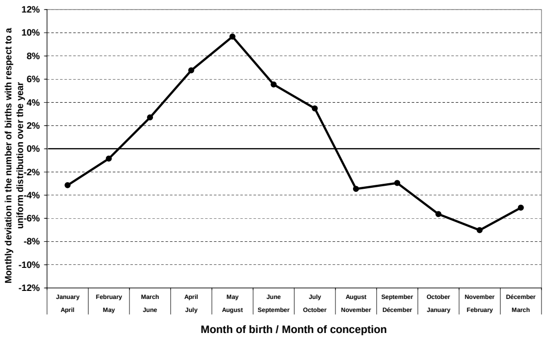
(a) Decade 1950-1959 (1982 Census)



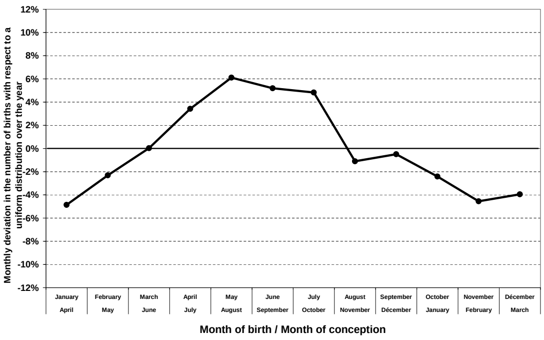
(b) Decade 1960-1969 (1982 Census)



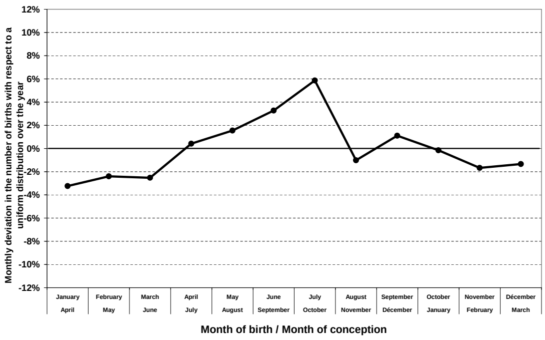
(c) Decade 1970-1979 (1982 Census)



(d) Decade 1980-1989 (1999 Census)



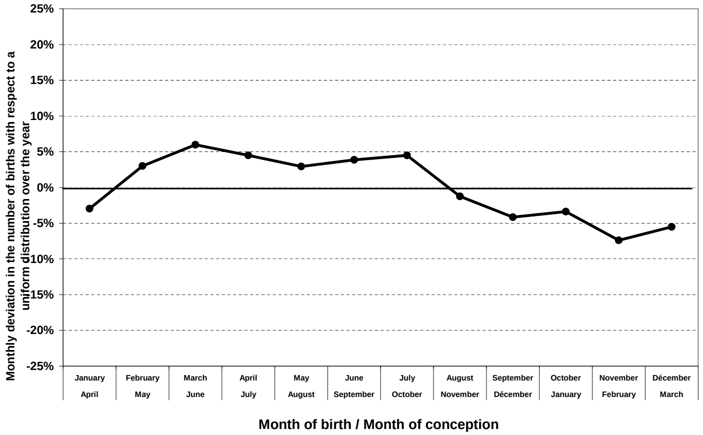
(e) Decade 1990-1998 (1999 Census)



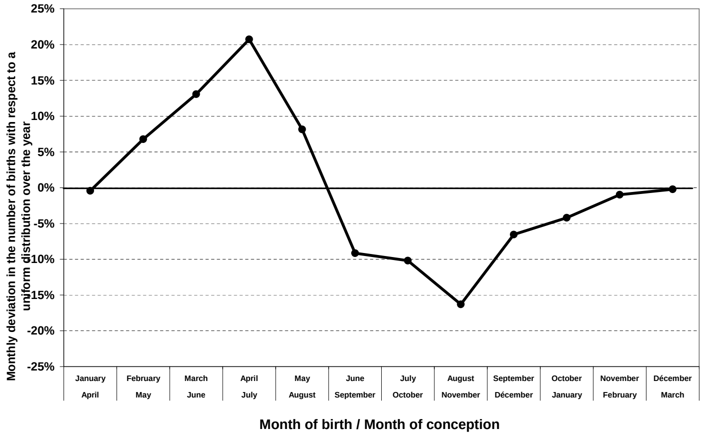
Notes: Each graph corresponds to a particular decade of birth. Each dot represents the monthly deviation in the number of birth during a decade with respect to a uniform distribution of births over the year. Denoting $\lambda_{m,d}$ the deviation associated with month m of decade d, we have: $\lambda_{m,d} = (J_m/365,2)(N_{m,d}/N_d)$, where J_m denotes the number of days in month m (28.2 for February because of leap years, 30 or 31 for other months), $N_{m,d}$ the number of births that occurred during month m of decade d and N_d the total number of births during decade d.

Figure 4: *Birth seasonality by mother's occupation. Monthly deviation in the number of births with respect to a uniform distribution over the year. Births that occurred beteen January 1, 1980 and December 31, 1998. Source: 1999 Census (1/4).*

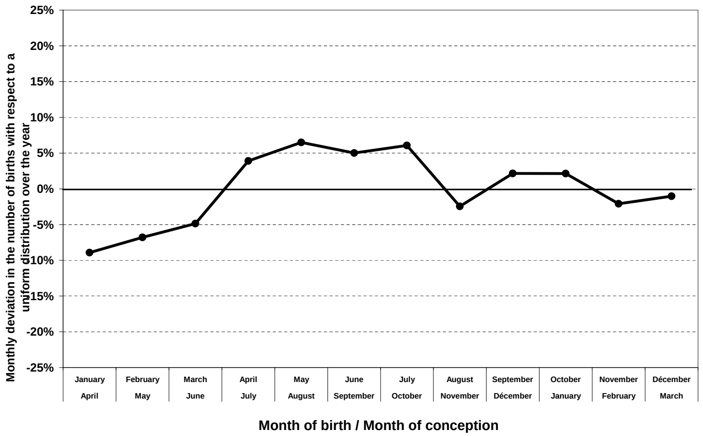
(a) Farmers



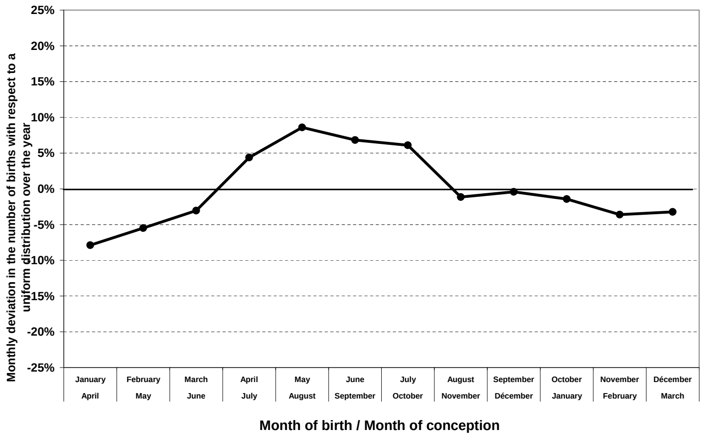
(b) Teachers and Professors



(c) Managers



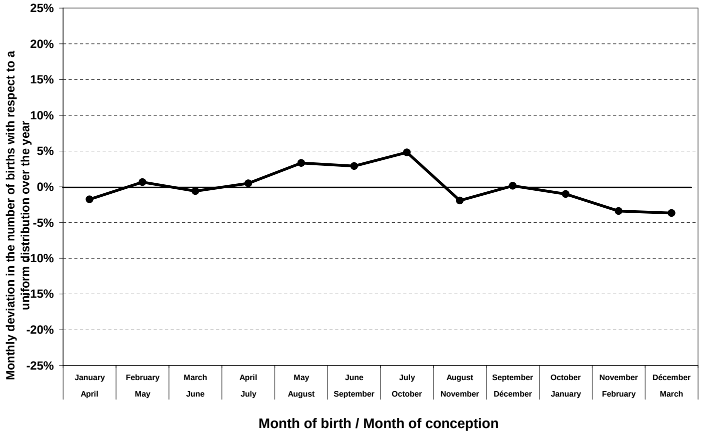
(d) Intermediary occupation



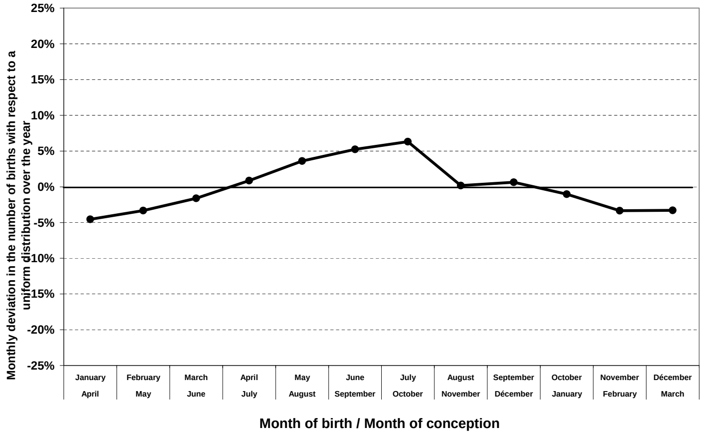
Notes: Each graph represents birth seasonality by mother's occupation for births that occurred between January 1, 1980 and December 31, 1998. Each dots shows, for a particular occupation, the monthly deviation in the number of birth during a decade with respect to a uniform distribution of births over the year. Denoting $\lambda_{m,o}$ the deviation associated with month m for women with occupation o, we have: $\lambda_{m,o} = (J_m/365,2)(N_{m,o}/N_o)$, where J_m denotes the number of days in month m (28.2 for February because of leap years, 30 or 31 for other months), $N_{m,o}$ the number of births that occurred during month m for women with occupation o and N_o the total number of births for these women between 1980 and 1998.

Figure 4: *Birth seasonality by mother's occupation. Monthly deviation in the number of births with respect to a uniform distribution over the year. Births that occurred beteen January 1, 1980 and December 31, 1998. Source: 1999 Census (1/4).*

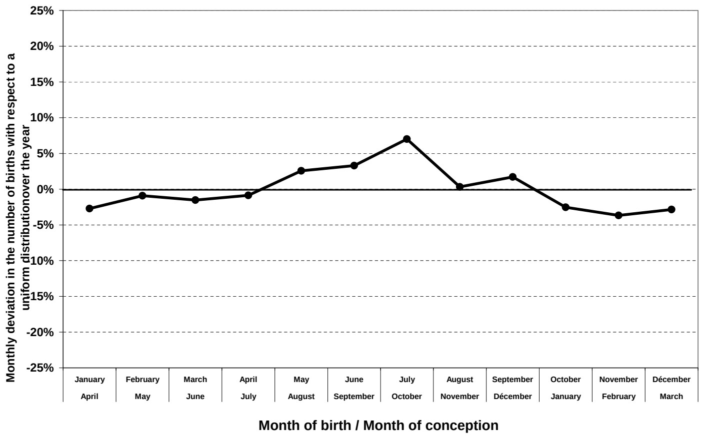
(e) Self-employed



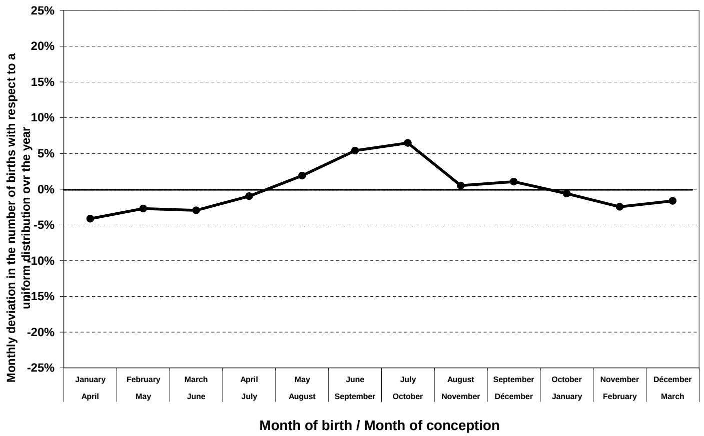
(f) Employees



(g) Manual workers



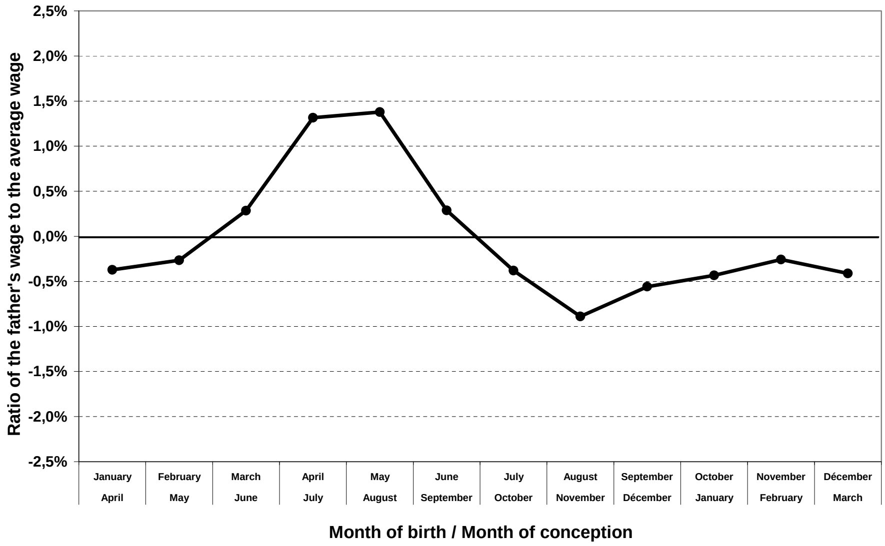
(h) Unemployed and economically inactive



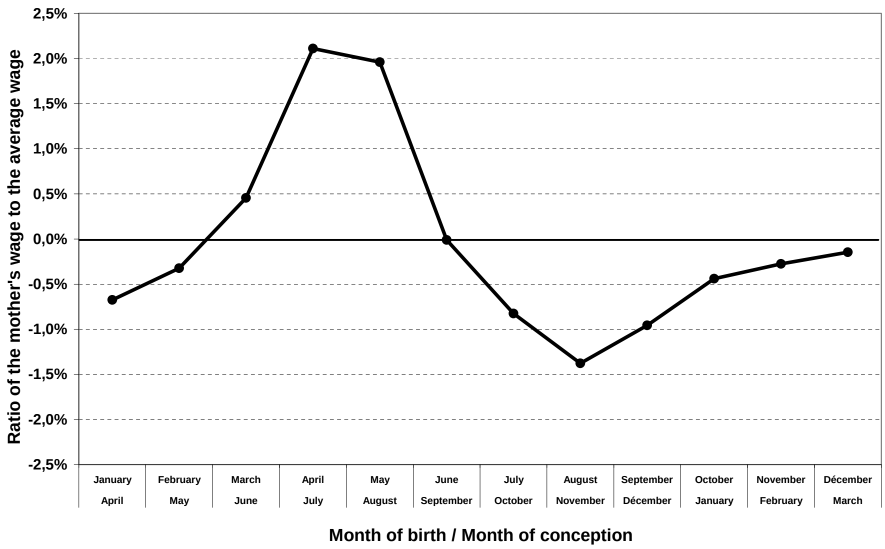
Notes: Each graph represents birth seasonality by mother's occupation for births that occurred between January 1, 1980 and December 31, 1998. Each dots shows, for a particular occupation, the monthly deviation in the number of birth during a decade with respect to a uniform distribution of births over the year. Denoting $\lambda_{m,o}$ the deviation associated with month m for women with occupation o, we have: $\lambda_{m,o} = (J_m/365, 2)(N_{m,o}/N_o)$, where J_m denotes the number of days in month m (28.2 for February because of leap years, 30 or 31 for other months), $N_{m,o}$ the number of births that occurred during month m for women with occupation o and N_o the total number of births for these women between 1980 and 1998.

Figure 5: *Average imputed wage of the father and mother of children born between January 1, 1960 and December 31, 1979 by the child's month of birth, divided by the average wage of fathers and mothers. Sources: 1982 Census (1/4) and Labour Force Survey (1990-2002).*

(a) Father's wage (imputed) / average wage



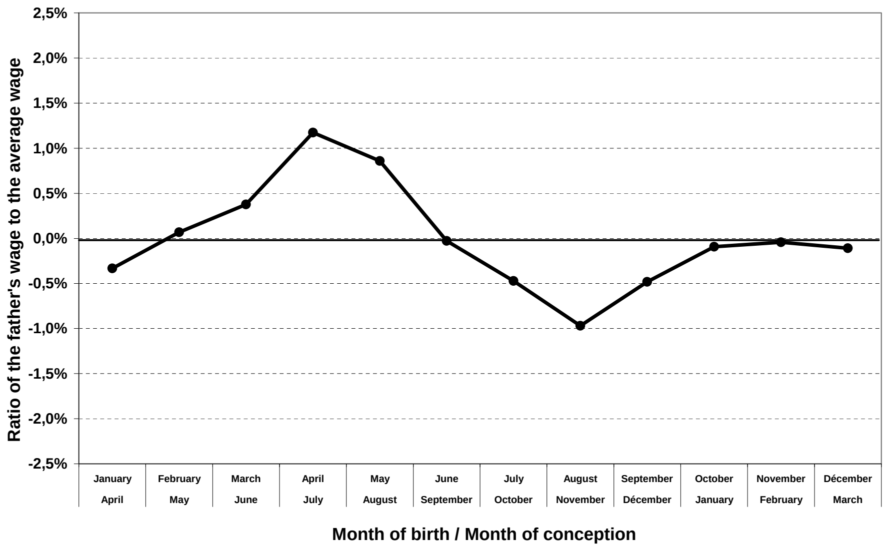
(b) Mother's wage (imputed) / average wage



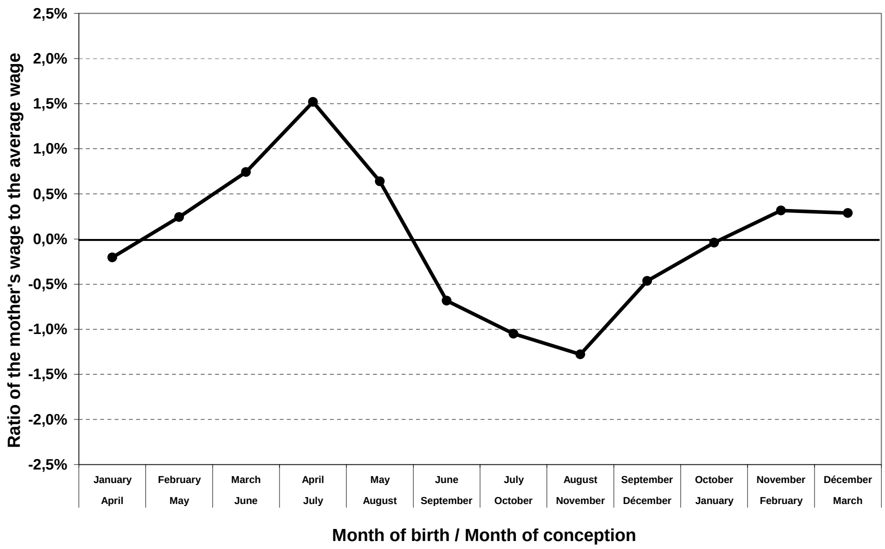
Notes: Each graph is constructed from the 1982 Census and the 1990-2002 *Enquête Emploi*. Each dots corresponds to the average wage of fathers (mothers) of children born in a particular month between January 1, 1960 and December 31, 1979. The wages of fathers and mothers are imputed from the average wage in their occupation group (456 categories) from the *Enquête Emploi*.

Figure 6: *verage imputed wage of the father and mother of children born between January 1, 1980 and December 31, 1998 by the child's month of birth, divided by the average wage of fathers and mothers. Sources: 1982 Census (1/4) and* Enquête Emploi *(1990-2002).*

(a) Father's wage (imputed) / average wage



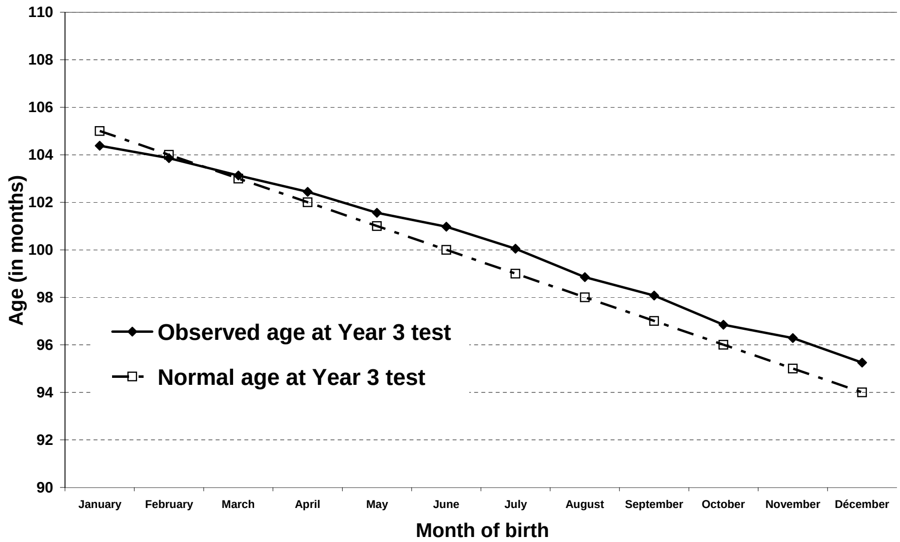
(b) Mother's wage (imputed) / average wage



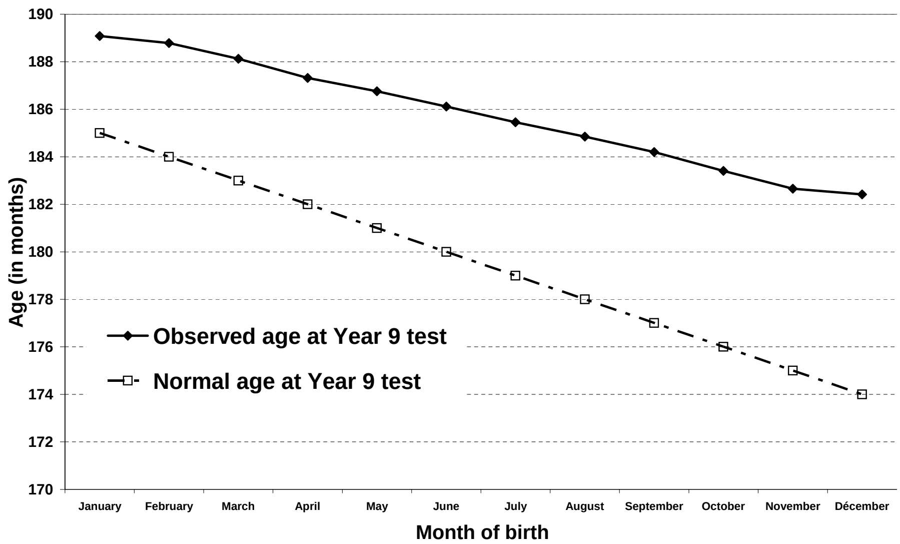
Notes: Each graph is constructed from the 1999 Census and the 1990-2002 *Enquête Emploi*. Each dots corresponds to the average wage of fathers (mothers) of children born in a particular month between January 1, 1980 and December 31, 1998. The wages of fathers and mothers are imputed from the average wage in their occupation group (456 categories) from the *Enquête Emploi*.

Figure 7: *Normal and observed age (in month) at which pupils sit Year 3 and Year 9 tests. Sources:* Panel Primaire de l'Éducation nationale *(1997) and* DNB *database (2004).*

(a) Average age at Year 3 test

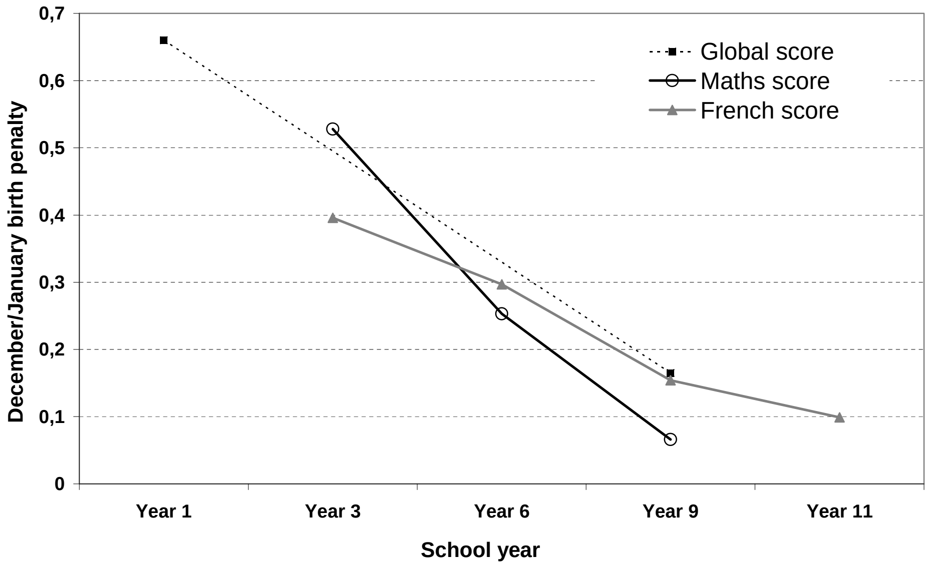


(b) Average age at Year 9 test



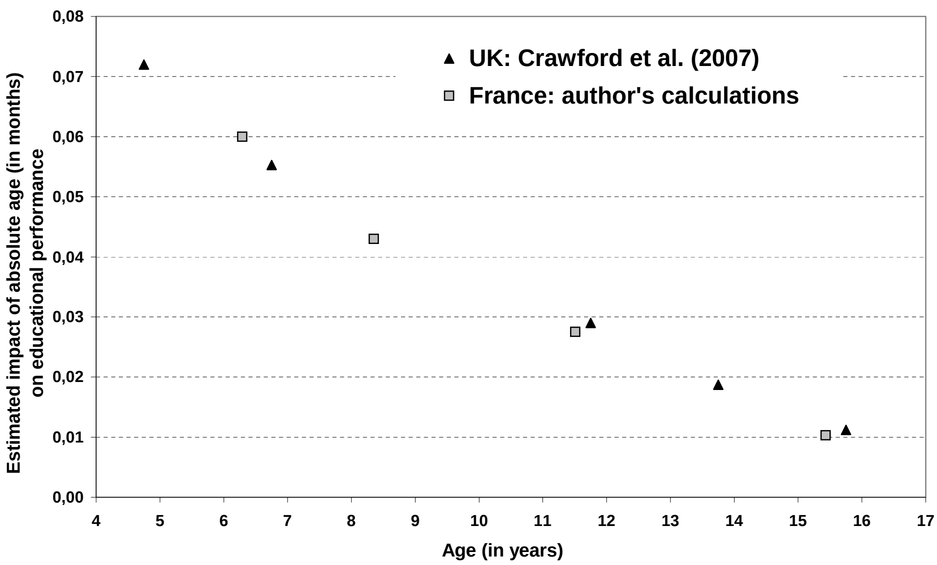
Notes: The normal age at which pupils sit the Year 3 test is the age they would have had in October 1992 had they started primary schooling at the normal age (i.e. born in 1990) without repeating a year. The observed age at which pupils sit the Year 3 tests is computed from the date at which the examination took place (October 1999 or 2000) using the pupil's month and year of birth. The normal age at which pupils sit the Year 9 test is the age they would have had in October 2004 had they started primary schooling at the normal age (i.e. born in 1989) without repeating a year. The observed age at which pupils sit the Year 9 tests is computed from the date at which the examination took place (June 2004) using the pupil's month and year of birth.

Figure 8: *Estimated impact of an 11-month age difference on standardized test scores between Year 1 and Year 11. Sources:* Panel primaire de l'Éducation nationale *(1997),* Panel secondaire de l'Éducation nationale *(1995) and* DNB *database (2004).*



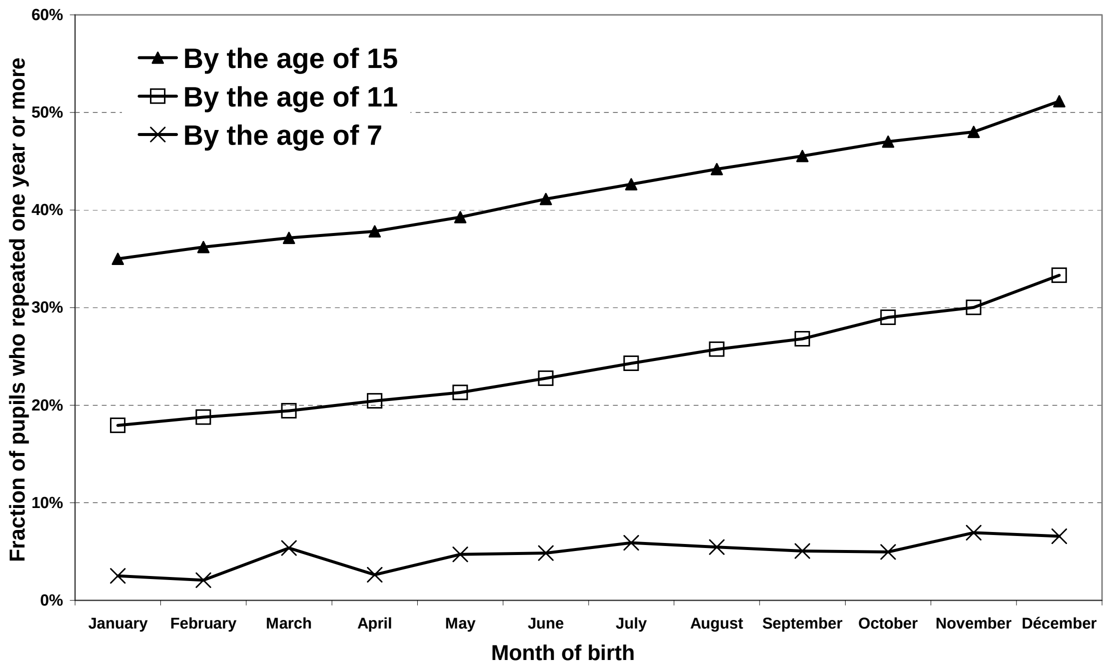
Notes: Each dot corresponds to the estimated effect of an 11-month age difference on test scores in Year 1, 3, 6, 9 and 11. The reported coefficients are computed from column 5 of table 1.

Figure 9: *Comparison of the impact of absolute age (in months) on standardized test scores in France and the United Kingdom, at different ages (in years). Sources: author's calculation and estimates from Crawford et al. (2007).*



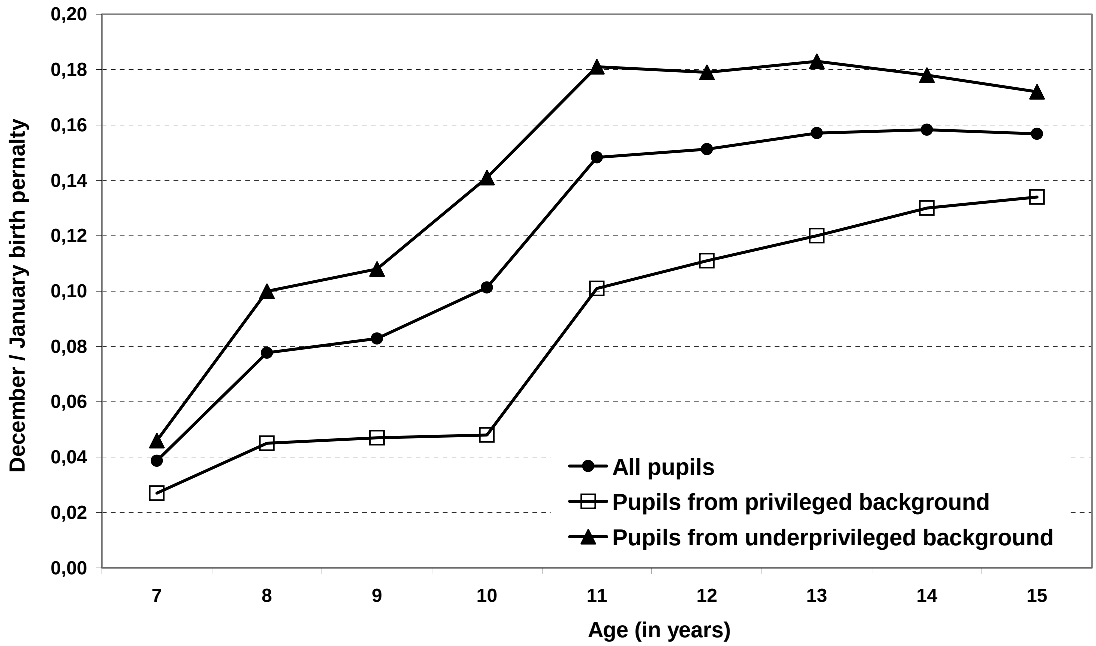
Notes: The coefficients displayed in this graph come from table 3.

Figure 10: *Fraction of pupils who have repeated one year or more by the age of 7, 11 and 15, by month of birth. Sources:* Panel primaire de l'Éducation nationale *(1997) and* Scolarité *database (2004-2005).*



Notes: The fraction of pupils who repeated a year by the age of 7 is computed using the *Panel primaire de l'Éducation nationale* (1997) whereas the corresponding calculations for age 11 and age 15 are performed using the *Scolarité* database covering school year 2004-2005.

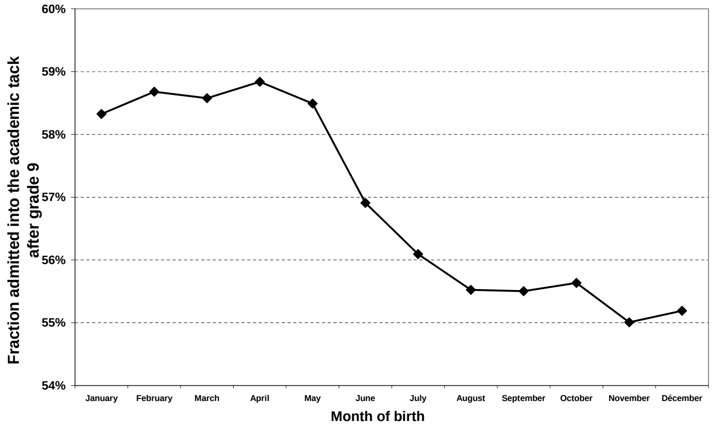
Figure 11: *Estimated penalty incurred by December-born vs. January-born pupils, in terms of the probability of having repeated one year or more at a given age, by socio-economic background. Sources:* Panel primaire de l'Éducation nationale *(1997) and* Scolarité *database (2004-2005).*



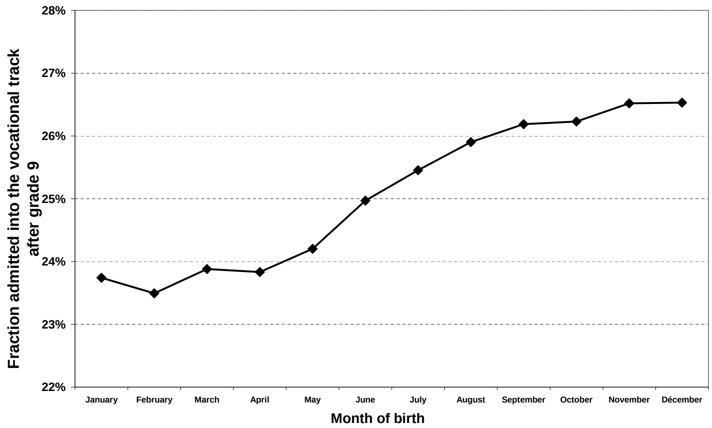
Notes: This graph shows, by age and socio-economic background, the impact of being born in December rather than January on the probability of having repeated a year or more in school. Each dot corresponds to the coefficients displayed in column 3 of table 4. This graph indicates for instance that at the age of 13, the probability of having repeated a year is 15 points higher for December-born than for January-born pupils. Age 7 to age 10 estimates were computed from the *Panel Primaire de l'Éducation nationale* (1997) and age 11 to age 15 estimates from the *Scolarité* database covering school year 2004-2005.

Figure 12: *Tracking of pupils born in 1986 at the end of Year 9. Sources:* Scolarité *database covering school years 2000-2001 to 2004-2005.*

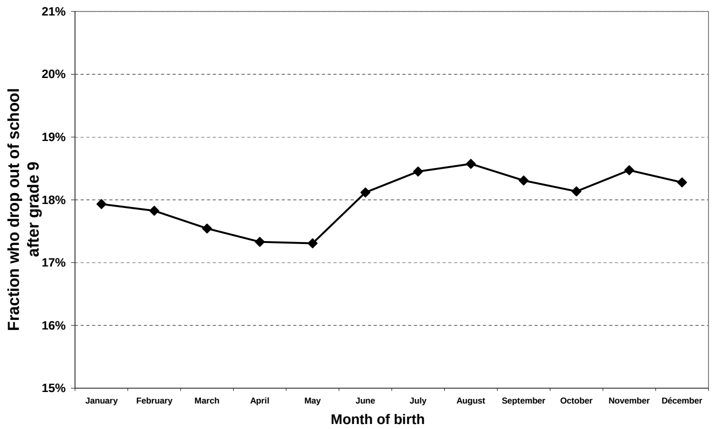
(a) Fraction of pupils admitted into the academic track



(b) Fraction of pupils admitted into the vocational track



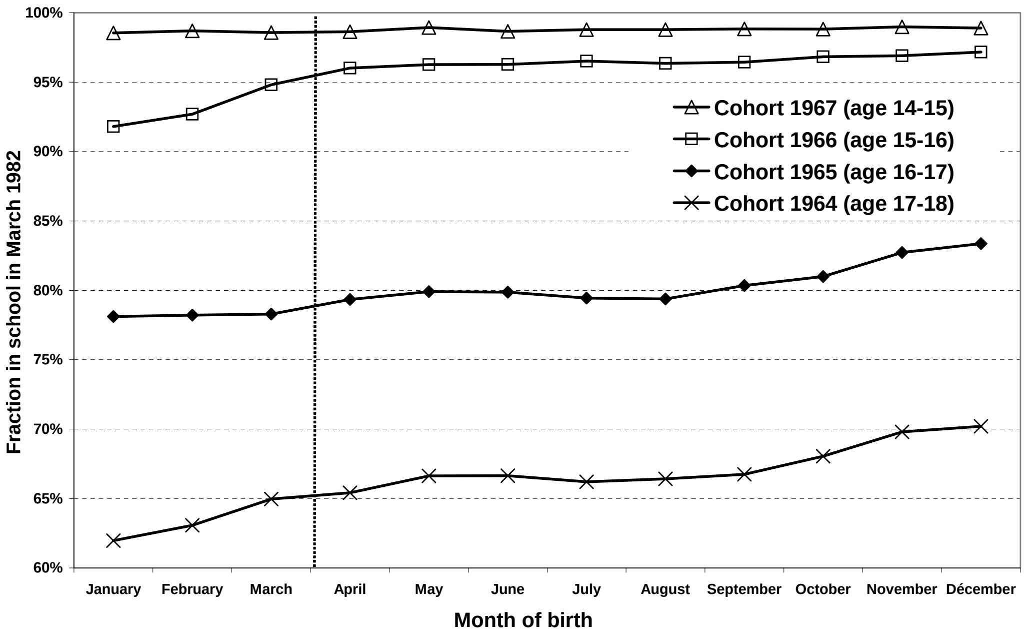
(c) Fraction of pupils who drop out of school after Year 9



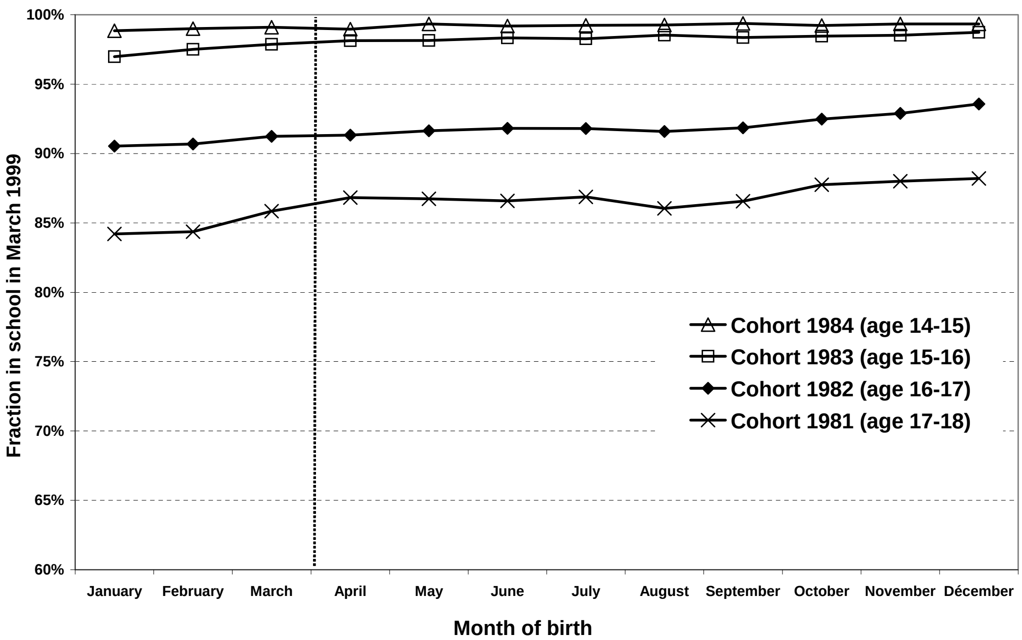
Notes: These graphs show the fraction of pupils born in 1986 who were eventually admitted into the academic track or the vocational track after year 9 and the fraction of those who dropped out of school after Year 9.

Figure 13: *Fraction of 14 to 18-year-olds who are enrolled in school in March of the Census year, by month of birth. Sources: 1982 and 1999 Censuses (1/4).*

(a) Fraction in school in March 1982



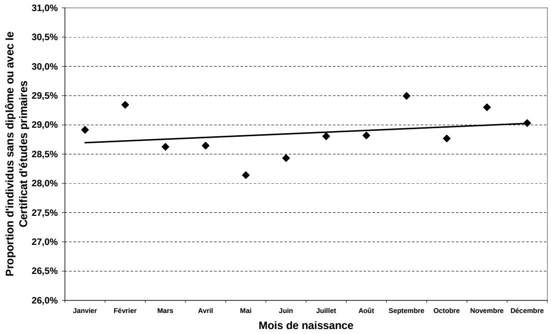
(b) Fraction in school in March 1999



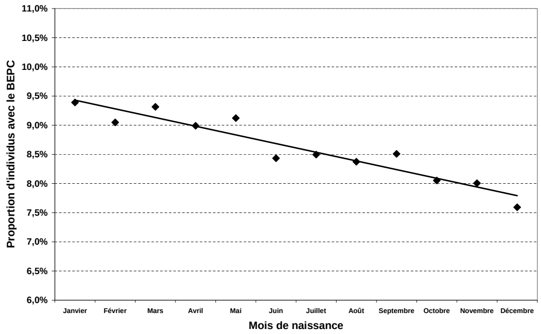
Notes: These graphs show the fraction of 14 to 18 year-olds who are enrolled in school in March of the Census year. In March 1982, individuals born in 1966 are aged 16 if they are born before March and 15 if they are born after.

Figure 14: *Highest held qualification of individuals born between 1945 and 1965, by month of birth. Source:* Enquête Emploi *(1990-2002).*

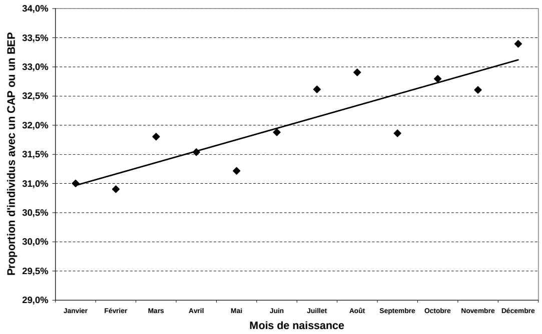
(a) No qualification or primary school certificate



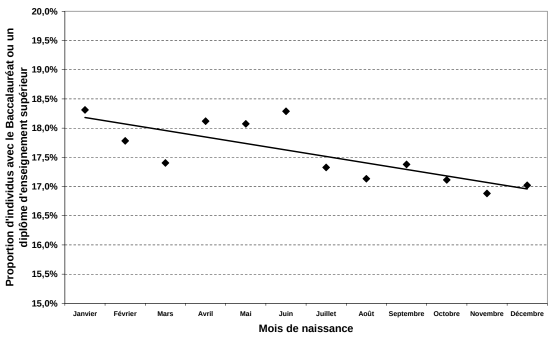
(b) Junior high school certificate



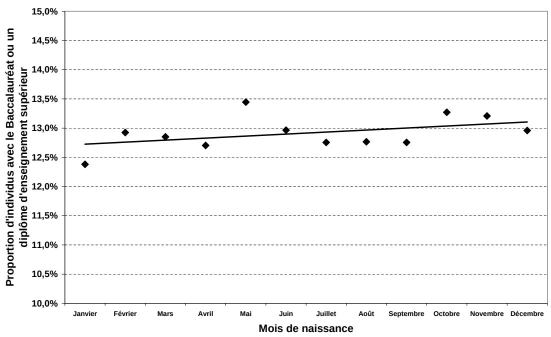
(c) Secondary school vocational qualification



(d) Academic A-level and above



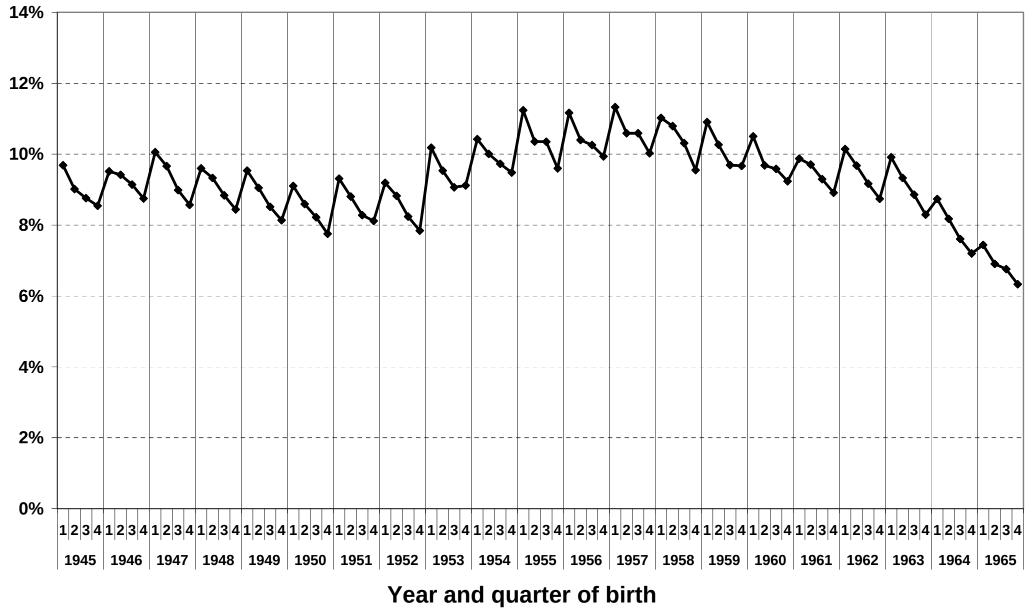
(e) Vocational A-level and above



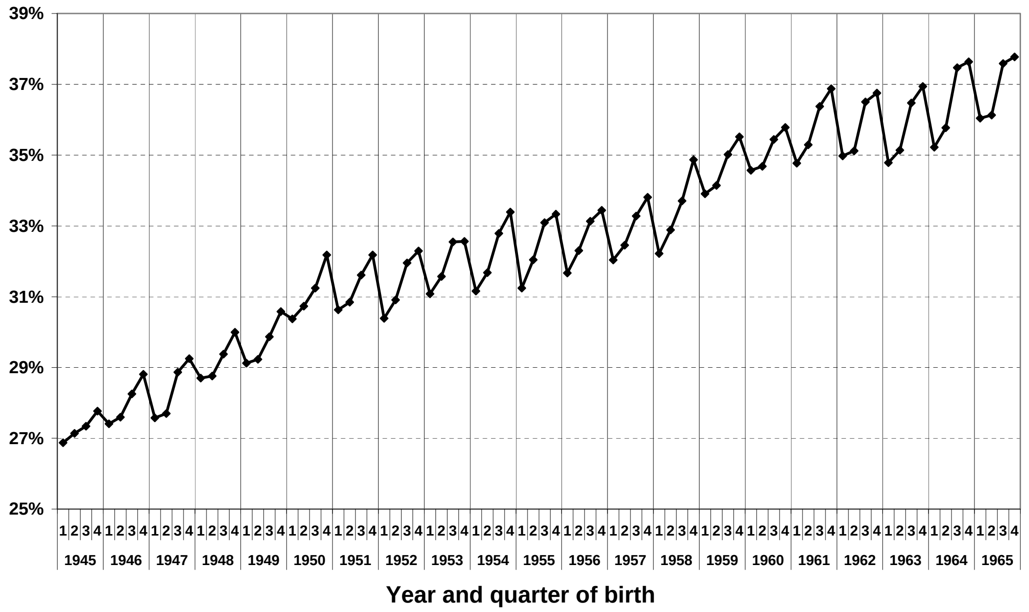
Notes: These graphs are constructed using the *Enquête Emploi* (1990-2002) dataset and show the distribution of the highest held qualification held by individuals born between 1945 and 1965, by month of birth.

Figure 15: *Fraction of individuals whose highest held qualification is either a Junior high school certificate or a secondary school vocational qualification, by year and quarter of birth. Source: 1999 Census.*

(a) Junior high school certificate



(b) Secondary school vocational qualification



Notes: These graphs are construced using the 1999 Census and show the fraction of individuals born between 1945 and 1965 whose highest held qualification is either a junior secondary schooling certificate or a vocational qualification, by year and quarter of birth.

Table 1: *Impact of absolute age (in months) on test scores in Year 1, 3, 6, 9 and 11. Instrument: assigned relative age (= 12 - month of birth). Sources:* Panel primaire de l'Éducation nationale *(1997),* Panel secondaire de l'Éducation nationale *(1995) and* DNB *database (2004).*

	Source	Average age (in years)	Dep. var.: score	Dep. var.: age in months	Dep. var.: score	Dep. var.: absolute age in months	Dep. var.: absolute age in months	Pupils born in Jan. or Dec.	Nb of obs. All pupils / Born in January and December only
				Instrument: assigned relative age					
			OLS	First stage	Reduced form	IV	IV	IV	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	
PRIMARY SCHOOL									
Year 1: Global score	PPEN 1997	6.3	0.029***	0.908***	0.056***	0.062***	0.060***	0.059***	9,342 / 1,537
			(0.003)	(0.006)	(0.003)	(0.003)	(0.003)	(0.005)	
Year 3: Maths	PPEN 1997	8.4	-0.011***	0.851***	0.042***	0.050***	0.048***	0.046***	7,653 / 1,257
			(0.003)	(0.011)	(0.003)	(0.004)	(0.004)	(0.006)	
Year 3: French	PPEN 1997	8.4	-0.013***	0.851***	0.031***	0.036***	0.036***	0.037***	7,653 / 1,257
			(0.003)	(0.011)	(0.003)	(0.004)	(0.004)	(0.006)	
MIDDLE SCHOOL									
Year 6: Maths	PSEN 1995	11.5	-0.056***	0.724***	0.019***	0.026***	0.023***	0.020***	16,790 / 3,601
			(0.001)	(0.014)	(0.002)	(0.003)	(0.003)	(0.004)	
Year 6: French	PSEN 1995	11.5	-0.056***	0.724***	0.022***	0.029***	0.027***	0.026***	16,790 / 3,601
			(0.001)	(0.014)	(0.002)	(0.003)	(0.003)	(0.004)	
Year 9: Maths (CA)	PSEN 1995	15.4	-0.051***	0.724***	0.002	0.002	0.003	-0.002	10,894 / 2,318
			(0.001)	(0.018)	(0.003)	(0.004)	(0.003)	(0.005)	
Year 9: French (CA)	PSEN 1995	15.4	-0.049***	0.724***	0.011***	0.015***	0.016***	0.012**	10,894 / 2,318
			(0.001)	(0.018)	(0.003)	(0.004)	(0.003)	(0.005)	
Year 9: Foreign language (CA)	PSEN 1995	15.4	-0.051***	0.724***	0.007***	0.010***	0.011***	0.012**	10,894 / 2,318
			(0.001)	(0.018)	(0.003)	(0.004)	(0.003)	(0.005)	
Year 9: Maths (FE)	DNB 2004	15.4	-0.023***	0.647***	0.004***	0.006***	0.006***	0.005***	781,391 / 127,822
			(0.000)	(0.004)	(0.000)	(0.001)	(0.000)	(0.001)	
Year 9: French (FE)	DNB 2004	15.4	-0.017***	0.647***	0.009***	0.014***	0.014***	0.015***	781,391 / 127,822
			(0.000)	(0.004)	(0.000)	(0.001)	(0.000)	(0.001)	
Year 9: History/Geography (FE)	DNB 2004	15.4	-0.020***	0.647***	0.008***	0.013***	0.014***	0.014***	781,391 / 127,822
			(0.000)	(0.004)	(0.000)	(0.001)	(0.001)	(0.001)	
Year 9: global score	DNB 2004	15.4	-0.026***	0.647***	0.010***	0.016***	0.015***	0.014***	781,391 / 127,822
			(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.001)	
HIGH SCHOOL									
Year 11: French (written)	PSEN 1995	16.3	-0.019***	0.775***	0.006*	0.008*	0.009*	0.011	5,460 / 1,117
			(0.003)	(0.016)	(0.004)	(0.005)	(0.005)	(0.005)	
Year 11: French (oral)	PSEN 1995	16.3	-0.018***	0.775***	-0.005	-0.006	-0.004	-0.005	5,460 / 1,117
			(0.003)	(0.016)	(0.004)	(0.005)	(0.005)	(0.008)	
Socio-demographic and school controls			No	No	No	No	Yes	Yes	

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. PPEN 1997: *Panel primaire de l'Éducation nationale* (1997). PSEN 1995: *Panel secondaire de l'Éducation nationale* (1995). DNB 2004: *Diplôme national du Brevet* 2004 database. CA: continuous assessment. FE: final examination. Each coefficient comes from a separate regression. Coefficients reported in column 1 come from the naive regression of test scores on observed age (in month) when sitting the exam. Coefficients in columns 3 to 5 correspond to the different stages (first stage, reduced form and IV) of the instrumental variables estimation of the impact of absolute age (in months) on test scores, using the assigned relative age (=12 - month of birth) as an instrument for observed age. IV estimates reported in column 5 control for pupil socio-demographic characteristics as well as school type. Coefficients reported in column 6 are estimated on the subsample of pupils born in January or December.

Table 2: *Subgroup analysis of the impact of absolute age (in months) on test scores in Year 1, 3, 6, 9 and 11. Instrument: assigned relative age (= 12 - month of birth). Sources:* Panel primaire de l'Éducation nationale *(1997),* Panel secondaire de l'Éducation nationale *(1995) and* DNB *database (2004).*

	Source	Dependent variable: test score			
		Boys	Girls	Privileged background	Underprivileged background
		IV	IV	IV	IV
		(1)	(2)	(3)	(4)
PRIMARY SCHOOL					
Year 1: Global score	PPEN 1997	0.060***	0.062***	0.054***	0.063***
		(0.004)	(0.004)	(0.004)	(0.004)
Year 3: Maths	PPEN 1997	0.052***	0.046***	0.037***	0.057***
		(0.006)	(0.005)	(0.005)	(0.005)
Year 3: French	PPEN 1997	0.039***	0.035***	0.030***	0.040***
		(0.005)	(0.005)	(0.005)	(0.005)
MIDDLE SCHOOL					
Year 6: Maths	PSEN 1995	0.020***	0.025***	0.021***	0.025***
		(0.004)	(0.004)	(0.004)	(0.004)
Year 6: French	PSEN 1995	0.026***	0.027***	0.030***	0.026***
		(0.004)	(0.004)	(0.004)	(0.004)
Year 9: Maths (CA)	PSEN 1995	0.003	0.003	0.006	-0.000
		(0.005)	(0.005)	(0.005)	(0.005)
Year 9: French (CA)	PSEN 1995	0.014***	0.017***	0.019***	0.013***
		(0.005)	(0.004)	(0.005)	(0.005)
Year 9: Foreign language (CA)	PSEN 1995	0.009*	0.011**	0.012**	0.009*
		(0.005)	(0.005)	(0.005)	(0.005)
Year 9: Maths (FE)	DNB 2004	0.005***	0.006***	0.008***	0.004***
		(0.001)	(0.001)	(0.001)	(0.001)
Year 9: French (FE)	DNB 2004	0.013***	0.015***	0.017***	0.012***
		(0.001)	(0.001)	(0.001)	(0.001)
Year 9: History/Geography (FE)	DNB 2004	0.012***	0.013***	0.014***	0.011***
		(0.001)	(0.001)	(0.001)	(0.001)
Year 9: global score	DNB 2004	0.015***	0.016***	0.017***	0.014***
		(0.001)	(0.001)	(0.001)	(0.001)
HIGH SCHOOL					
Year 11: French (written)	PSEN 1995	0.012	0.010*	0.014**	0.006
		(0.008)	(0.006)	(0.007)	(0.007)
Year 11: French (oral)	PSEN 1995	-0.001	-0.006	0.003	-0.012
		(0.008)	(0.006)	(0.007)	(0.007)
Socio-demographic and school controls		Yes	Yes	Yes	Yes

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. PPEN 1997: *Panel primaire de l'Éducation nationale* (1997). PSEN 1995: *Panel secondaire de l'Éducation nationale* (1995). DNB 2004: *Diplôme national du Brevet* 2004 database. CA: continuous assessment. FE: final examination. Each coefficient comes from a separate regression performed on a subgroup of pupils. IV estimates are obtained by using assigned relative age as an instrument for observed age when sitting the examination. Pupils are classified as having a privileged or underprivileged background depending on the household head's occupation. All regressions include controls for pupil socio-demographic characteristics and type of school attended.

Table 3: *Comparison of the estimated impact of absolute age (in months) on standardized test scores in France and the United Kingdom, by age of pupils (in years). Sources: author's calculation for France (cf. table 1) and estimates by Crawford et al. (2007) for the United Kingdom.*

Year of test		Average age at the test (in years)	Estimated impact of age (in month) on test scores	
France	UK		France	UK
(1)	(2)	(3)	(3)	(4)
	Foundation Stage	4.75		0.072
Year 1		6.29	0.062	
	Key Stage 1	6.75		0.055
Year 3		8.35	0.043	
Year 6		11.51	0.028	
	Key Stage 2	11.75		0.029
	Key Stage 3	13.75		0.019
year 9		15.43	0.010	
	Key Stage 4	15.75		0.011

Notes: This table compares the estimated impact of absolute age (in months) on test scores in France and the United Kingdom, separately by age (in years). Estimates for France come from column 4 of table 1. Estimates for the UK come from Crawford et al. (2007). Coefficients reported in column 3 correspond, for Year 1, 3, 6 and 9, to the average estimated impact of age on the score in maths and in French (final examination for Year 9 test). Coefficients reported in column 4 are computed from tables 5.1 (Foundation Stage and Key Stage 1), B.1 (Key Stage 2 and Key Stage 3) and 5.3 (Key Stage 4) in Crawford et al. (2007).

Table 4: *Probit estimation of the marginal impact of month of birth on the probability of having repeated one year or more at a given age, by socio-economic background. Sources:* Panel primaire de l'Éducation nationale *(1997) and* Scolarité *database (2004-2005).*

DEPENDENT VARIABLE: PROBABILITY OF HAVING REPEATED ONE YEAR OR MORE

		All pupils		Privileged background		Underprivileged background	
Age	Source	Marginal effects (1)	Dec. vs. Jan. penalty (2)	Marginal effects (3)	Dec. vs. Jan. penalty (4)	Marginal effects (5)	Dec. vs. Jan. penalty (6)
by age 7	PPEN 1997	0.003***	0.038	0.002***	0.027	0.004***	0.046
		(0.001)		(0.001)		(0.001)	
by age 8	PPEN 1997	0.007***	0.077	0.004***	0.045	0.009***	0.100
		(0.001)		(0.001)		(0.001)	
by age 9	PPEN 1997	0.008***	0.083	0.004***	0.047	0.010***	0.108
		(0.001)		(0.001)		(0.002)	
by age 10	PPEN 1997	0.009***	0.101	0.004***	0.048	0.013***	0.141
		(0.001)		(0.001)		(0.002)	
[Number of observations]		[8,663]		[3,749]		[4,914]	
by age 11	BASCO 2004-2005	0.013***	0.149	0.009***	0.101	0.016***	0.181
		(0.000)		(0.000)		(0.000)	
[Number of observations]		[799,309]		[327,140]		[472,169]	
by age 12	BASCO 2004-2005	0.014***	0.152	0.010***	0.111	0.016***	0.179
		(0.000)		(0.000)		(0.000)	
[Number of observations]		[799,309]		[327,140]		[472,169]	
by age 13	BASCO 2004-2005	0.014***	0.158	0.011***	0.120	0.017***	0.183
		(0.000)		(0.000)		(0.000)	
[Number of observations]		[799,309]		[327,140]		[472,169]	
by age 14	BASCO 2004-2005	0.014***	0.159	0.012***	0.130	0.016***	0.178
		(0.000)		(0.000)		(0.000)	
[Number of observations]		[792,577]		[326,434]		[466,143]	
by age 15	BASCO 2004-2005	0.014***	0.158	0.012***	0.134	0.016***	0.172
		(0.000)		(0.000)		(0.000)	
[Number of observations]		[752,877]		[317,149]		[435,728]	

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. PPEN 1997: *Panel Primaire de l'Éducation nationale* (1997). BASCO 2004-2005: *Scolarité* database covering school year 2004-2005. Each coefficient comes from a separate regression. The impact of month of birth on the probability of having repeated a year or more by a given age is estimated using a probit model. The penalty incurred by December-born pupils with respect to their January-born classmates is computed by multiplying by a factor 11 the estimated marginal effect of month of birth on the probability of having repeated a year or more. Pupils are considered as having a privileged or underprivileged background depending on the household head's occupation.

Table 5: *Impact of month of birth on the tracking of pupils after Year 9. Multinomial probit model. Sources:* Scolarité *database covering years 2000-2001 to 2004-2005.*

	All pupils		Pupils born in January or in December	
	Marginal effects (1)	Dec/Jan penalty (2)	Marginal effects (7)	Dec/Jan penalty (8)
Dependent variable:				
Admitted into the academic track	-0.004***	-0.044	-0.003***	-0.031
	(0.000)		(0.000)	
Admitted into the vocational track	0.003***	0.035	0.003***	0.028
	(0.000)		(0.000)	
Drop out after Year 9	0.001***	0.009	0.000	0.003
	(0.000)		(0.000)	
Number of observations	772,561		127,553	

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the tracking of pupils after Year 9, using a probit model. The sample is constructed from the *Scolarité* database covering school years 2000-2001 to 2004-2005 for pupils born in 1986.

Table 6: *Impact of month of birth on the tracking of pupils after Year 9. Subgroup analysis using a multinomial probit model. Sources:* Scolarité *database covering years 2000-2001 to 2004-2005.*

	Pupils born in January or December							
	Boys		Girls		Privileged background		Underprivileged background	
	Marginal effects	Dec/Jan penalty	Marginal effects	Dec/Jan penalty	Marginal effects	Dec/Jan penalty	Marginal effects	Dec/Jan penalty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:								
Admitted into the academic track	-0.003***	-0.033	-0.003***	-0.033	-0.002***	-0.019	-0.004***	-0.041
	(0.000)		(0.000)		(0.000)		(0.000)	
Admitted into the vocational track	0.003***	0.028	0.003***	0.028	0.002***	0.022	0.003***	0.032
	(0.000)		(0.000)		(0.000)		(0.000)	
Drop out after Year 9	0.000	0.005	0.000	0.002	-0.000	-0.003	0.001***	0.008
	(0.000)		(0.000)		(0.000)		(0.000)	
Number of observations	64,475		63,078		50,426		77,127	

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the tracking of pupils after Year 9, using a probit model. The sample is constructed from the *Scolarité* database covering school years 2000-2001 to 2004-2005 for pupils born in 1986.

Table 7: *Impact of month of birth on the number of years of schooling. Individuals born between 1945 and 1965. Source:* Enquête Emploi *(1990-2002).*

	Men			Women		
	December/January penalty			December/January penalty		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:						
Age left full-time education	-0.059	-0.057	-0.050	0.067**	0.072**	0.101**
	(0.055)	(0.036)	(0.055)	(0.023)	(0.032)	(0.048)
CONTROL VARIABLES:						
Survey year dummies	No	Yes	Yes	No	Yes	Yes
Year of birth dummies	No	Yes	Yes	No	Yes	Yes
SAMPLE:						
Month of birth	All	All	Jan & Dec	All	All	Jan & Dec
Number of observations	97,746	97,746	16,195	102,532	102,532	17,078

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the age left full-time education. The sample is constructed from the *Enquête Emploi* (1990-2002) dataset and includes all individuals born between 1945 and 1965. Standard-errors are clustered at the year of birth level.

Table 8: *Impact of month of birth on highest held qualification. Multinomial probit model. Individuals born between 1945 and 1965. Source:* Enquête Emploi *(1990-2002).*

	Men			Women		
	Dec/Jan penalty			Dec/January penalty		
	(1)	(2)	(3)	(4)	(5)	(6)
HIGHEST HELD QUALIFICATION:						
No qualification or primary school certificate	0.004	0.003	0.004	-0.001	-0.001	-0.004
	(0.005)	(0.005)	(0.007)	(0.004)	(0.005)	(0.007)
Junior high school certificate	-0.012***	-0.012***	-0.016***	-0.021***	-0.021***	-0.025***
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Vocational secondary school qualification	0.024***	0.024***	0.034***	0.021***	0.021***	0.018***
	(0.006)	(0.005)	(0.007)	(0.005)	(0.004)	(0.007)
Academic A-Level and above	-0.016***	-0.016***	-0.022***	-0.006**	-0.006	0.001
	(0.005)	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)
Vocational A-Level and above	0.001	0.001	-0.003	0.007**	0.007**	0.011**
	(0.003)	(0.003)	(0.005)	(0.004)	(0.003)	(0.005)
CONTROLS:						
Survey year dummies	No	Yes	Yes	No	Yes	Yes
Year of birth dummies	No	Yes	Yes	No	Yes	Yes
SAMPLE:						
Month of birth	All	All	Jan & Dec	All	All	Jan & Dec
Number of observations	97,868	97,868	16,224	102,598	102,598	17,082

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the highest held qualification. The sample is constructed from the *Enquête Emploi* (1990-2002) dataset and includes all individuals born between 1945 and 1965. Standard-errors are clustered at the year of birth level.

Table 9: *Impact of month of birth on employment status. Probit model. Individuals born between 1945 and 1965. Source:* Enquête Emploi *(1990-2002).*

	Men			Women		
	Dec/Jan penalty			Dec/Jan penalty		
	(1)	(2)	(3)	(4)	(5)	(6)
IMPACT OF MONTH OF BIRTH ON:						
The probability of being unemployed	0.004**	0.004*	0.004*	0.001	0.001	-0.000
	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)
Number of observations	93,890	93,890	15,550	81,343	81,343	13,604
Probability of working part-time	0.001	0.001	0.001	-0.002	-0.002	-0.006
	(0.001)	(0.002)	(0.002)	(0.003)	(0.005)	(0.009)
Number of observations	88 547	88 547	14 684	73 459	73 459	12 302
Probability of being a civil servant	-0.011***	-0.011***	-0.018***	-0.007	-0.007	-0.013*
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.008)
Number of observations	89,097	89,097	14,768	74,129	74,129	12,425
CONTROL VARIABLES:						
Survey year dummies	No	Yes	Yes	No	Yes	Yes
Year of birth dummies	No	Yes	Yes	No	Yes	Yes
SAMPLE:						
Month of birth	All	All	Jan & Dec	All	All	Jan & Dec

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the employment status. The sample is constructed from the *Enquête Emploi* (1990-2002) dataset and includes all individuals born between 1945 and 1965, who earn a positive wage and are surveyed during the survey's first wave. Standard-errors are clustered at the year of birth level.

Table 10: *Impact of month of birth on hourly wage (in 2005 Euros). Individuals born between 1945 and 1965. Source:* Enquête Emploi *(1990-2002).*

	Men					Women				
	Dec/Jan penalty					Dec/Jan penalty				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable:										
Log of hourly wage	-0.023***	-0.023***	-0.012***	-0.031***	-0.014**	-0.007**	-0.008*	-0.009*	-0.014*	-0.014**
	(0.006)	(0.005)	(0.004)	(0.007)	(0.006)	(0.003)	(0.005)	(0.004)	(0.008)	(0.006)
CONTROL VARIABLES:										
Survey year dummies	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Year of birth dummies	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Highest held qual. (5 cat.)	No	No	Yes	No	Yes	No	No	Oui	No	Yes
SAMPLE:										
Month of birth	All	All	All	Jan & Dec	Jan & Dec	All	All	All	Jan & Dec	Jan & Dec
Number of observations	63,551	63,551	63,551	10,603	10,603	60,231	60,231	60,231	10,091	10,091

Notes: *: significant at the 10% level; **: significant at the 5% level; ***: significant at the 1% level. This table shows the estimated impact of month of birth on the log of hourly wage. The sample is constructed from the *Enquête Emploi* (1990-2002) dataset and includes all individuals born between 1945 and 1965, who earn a positive wage and are surveyed in the survey's first wave. Qualifications are grouped into five categories: primary schooling certificate or below; junior secondary schooling certificate; vocational qualification; Academic A-level or above; technical A-Level or above. Standard-errors are clustered at the year of birth level.