

Aspirations and the Perpetuation of Social Inequalities: Evidence from French Teenagers *

Nina Guyon[†] and Elise Huillery[‡]

October 2, 2016

Abstract

This paper provides empirical evidence that aspirations are socially determined and affect individual outcomes in the context of education. Using unique and exceptionally rich data on aspirations from French ninth graders, we show that low-SES students have lower aspirations than high-SES students, even among equally-achieving classmates, and that lower aspirations lead to lower later school outcomes. We also find that the main reason why equally-achieving students from different social origin have different aspirations is that they do not feel capable of pursuing the same academic tracks. Our data shows that it cannot be taken for granted that low-SES students assess their future academic potential at its true value, as they misperceive their present academic capacity and exhibit excessive social fatalism. These findings show that aspirations perpetuate social inequalities and cast doubts about the optimality of subsequent outcomes.

JEL Codes: I24, I21, J15, O15

Keywords: School Aspirations, Track choices, Inequality, Poverty Trap

1 Introduction

Low social mobility is a major concern with respect to both social fairness and economic efficiency: inequalities in opportunities prevent the disadvantaged from realizing their best outcomes possible and the society from benefitting from everyone's full potential¹. This paper explores the role of student aspirations as a cause

*We wish to acknowledge the support of the LIEPP (Interdisciplinary Research Center for the Evaluation of Public Policies), the Sciences Po Scientific Advisory Board, the Direction de l'évaluation, de la prospective et de la performance in the French Ministry of Education and the French National Research Agency (ANR-11-LABX-0091, ANR-11-IDEX-0005-02, et ANR-10-BLANC-1819-01-EVALPOLPUB). We are grateful to Yann Algan, Pierre-Henri Bono, Séverine Chauvel, Denis Cogneau, Marie Duru-Bellat, Marc Gurgand, Emeric Henry, John Ham, Karla Hoff, Rachel Kranton, Julien Labonne, Eric Maurin, Jessica Pan, and Eve Shira for helpful comments. We are deeply indebted to Sébastien Bauvet for excellent project management, as well as Cécile Ballini, Mbayi Mukendi, Rokhya Konate, Nicolas Hovart, and Rabeh Mbaigoto for excellent data collection. We also thank Alexandre Naud for his assistance with data entry and Cécile Ballini for her help with data management. Finally, we gratefully acknowledge the school staff and students for the time and information they gave. All errors remain our own.

[†]National University of Singapore, Department of Economics (nina.guyon@nus.edu.sg) and Sciences Po, LIEPP

[‡]Sciences Po, Department of Economics (elise.huillery@sciencespo.fr) and LIEPP

¹Intergenerational correlation in years of schooling between parents and children is between 0.30 and 0.50 in seven OECD countries including the United States (Hertz et al. 2007). Björklund and Salvanes (2010) find that, in all countries for which they had data, more than 50% of the variation in years of schooling can be attributed to factors shared by siblings. In France, the link between social background and school achievement is particularly strong: the social gap in math score in 2012 is the most extreme of all OECD countries, and it has increased by 33% since 2003 (PISA, 2012). Precisely, in 2012, being from a more advantaged background in France induces a 57-point increase in the math score. The average for OECD countries is 39 points, and it was 43 points in 2003 in France. In 2010-2012, 65% of individuals aged 25-29 from high or medium socio-economic status families completed some higher education, while only 30% from low socio-economic status families did so (Le Rhun, 2015).

of self-perpetuating social inequalities in school achievement. Numerous studies have shown that family background influences student academic achievement (Sacerdote 2007, Oreopoulos 2006) through different channels: inherited ability (Black et al. 2005), parental involvement (Avvisati et al. 2014), school and neighbourhood quality (Goux and Maurin 2007, Chetty et al. 2011, Chetty et al. 2016). Whether family background also determines academic achievement through students’ aspirations, and whether aspirations are formed on rational or irrational bases remain open empirical questions.

Aspirations are the goals that individuals form for themselves in the future. Following the seminal paper of Appadurai (2004), aspirations should be considered as a *capacity* in which people may differ: some people are more capable than others to set appropriate goals for their future, meaning goals that are in line with their potential and lead to the best possible outcomes. The emergent theoretical literature on aspirations argues that the capacity to aspire is inherently unequal between rich and poor (Appadurai 2004, Ray 2006, Ray and Génicot 2015, Dalton et al. 2016, see Section 2 for more details). An aspiration-based poverty trap occurs when inadequate aspirations induce suboptimal investment and effort to better one’s life, resulting in poor outcomes compared to what would have been achieved with adequate aspirations, that lead to lower aspirations, etc. This paper contributes to this literature by providing first empirical evidence of the role of aspiration failures in widening social inequalities in the context of education in France. After a preliminary investigation of the effect of aspirations on later school outcomes, this paper explores whether aspirations are influenced by parental socio-economic status (hereafter SES), and why. While this paper does not preclude that low-SES students have good reasons to aspire lower than high-SES students, we explore the possibility of additional irrational reasons: ignorance of some educational options, and misperceptions of one’s academic potential.

We use both administrative and unique survey data from 3,415 students in their last year of junior high school (i.e. 9th grade) in 59 schools in the Paris metropolitan area. Our survey includes two one-week apart academic tests to assess individual academic performance at the beginning of the year, and measures educational and professional aspirations, knowledge of existing academic tracks, perception of attainable tracks, scholastic self-esteem, beliefs about the influence of social origin on future academic success, and additional measures of cognitive and non-cognitive skills. To the best of our knowledge, this paper is the first to de-construct preferred options as embedded in the zone of self-perceived attainable and salient options (“aspiration window” in Ray’s terminology). Students were asked first to list all existing tracks they know (hereafter *salient* tracks), then, among the tracks they know, which ones they feel capable of pursuing (hereafter *attainable* tracks), and finally, among tracks they feel capable of pursuing, which one they prefer (hereafter *preferred* tracks). We matched this survey data with administrative data providing parental SES,

average yearly grades in grade 9, test scores at the national exam taken at the end of grade 9, and track assignments in grade 10. In France, the curriculum is uniform for all students in junior high school, but it gives way to a stratified system of high schools, which involves academic and vocational tracks. For the first time in their lives, a choice is thus to be made between different educational tracks that will, due to some irreversibility, determine future academic and professional paths.

Social inequalities in educational aspirations would not matter would aspirations not influence later outcomes. We first examine the relationship between aspirations at the beginning of grade 9 and school outcomes at the end of grade 9, and discuss the extent to which this relationship is driven by confounding factors. Although we don't have an experimental source of variation in aspirations to perfectly identify the causal impact of aspirations on later school outcomes, our original survey and administrative data allows to control for an exceptionally large number of confounding factors. We find that the relationship between educational aspirations and later school outcomes is large and robust to the inclusion of important control variables, which indicates that our estimates are close to the causal ones. Among equally-achieving classmates at the beginning of grade 9 with similar parental SES, those who aspire to academic high school get a 0.22 standard deviation higher average grade over grade 9, a 0.17 standard deviation higher test score at the end of grade 9, and a 0.18 percentage point higher probability to be assigned to academic high school in grade 10, than those who aspire to vocational high school or have no aspiration. Our data are consistent with the hypothesis that aspirations have a large impact on later school outcomes. Our estimates contribute to the nascent empirical literature (see Section 2 for a review) showing that aspirations are not just a consequence of realized outcomes, but also a cause of later outcomes.

We then provide a measure of social inequalities in aspirations. Our identification relies on the fact that family background is determined by the accident of birth. We are not interested in the pure effect of SES on aspirations *everything else being equal*, but in the *ecological* effect (Bronfenbrenner, 1979) of parent SES on aspirations, which incorporates a lot of dimensions: inherent cognitive and non-cognitive skills, social networks, parental involvement in education, school quality, etc. The reasons why aspirations correlate with parental SES are treated as consequences of parental SES. Because a change in parental SES implies a change in all these SES-related factors, differences between low-SES and high-SES students is interpreted as the effect of all SES-related factors. We thus compare educational and professional aspirations of low and high-SES students, and examine the degree to which the influence of parent SES can be attributed to differential academic capacity and neighborhood-school quality, which are presumably *rational* determinants of aspirations (they affect one's potential). But the most important innovation of the paper is to also compare students' salient and attainable tracks to examine the role of knowledge of academic tracks and *self-*

perceived academic potential in explaining social differences in aspirations. Once academic capacity, school environment, salient tracks and attainable tracks have been controlled for, the remaining social differences in aspirations may be attributed to reasons like different liquidity constraints facing the cost of education, different expected returns to education, or identity concerns, which we will not investigate in this paper. This paper rather focuses on irrational factors that *pre-determine* aspirations: knowledge of academic tracks and misperception of one’s academic potential. To examine the rationality of self-perceived academic potential, we present additional descriptive evidence on the extent to which students have fatalistic views on the link between social origin and academic success, as well as on the influence of parent SES on scholastic self-esteem.

We find clear evidence of social differences in aspirations, and not only because of differences in academic capacity and school environment. Low-SES students have very different salient, attainable and preferred tracks than high-SES, even once academic capacity and class fixed effects are controlled for. In terms of aspirations for high school, low-SES they have a 7% higher probability to mention vocational high school in their salient tracks, a 45% higher probability to mention vocational high school in their attainable tracks, and a 120% higher probability to prefer vocational high school, relative to equally-achieving high-SES classmates. Low-SES are symmetrically biased against academic high school relative to high-SES. And in terms of aspirations for higher education, low-SES students are 15% less likely to mention masters as a salient track, 28% less likely to mention masters as an attainable track, and 26% less likely to prefer masters, than equally-achieving high-SES classmates. Low-SES are symmetrically more likely to not mention any salient and attainable tracks in higher education, and 69% more likely to prefer finding a job right after high school. Interestingly, social differences in aspirations concern both high, medium and low-achieving students. While low-achieving low-SES students may appear more realistic than their high-SES counterparts, medium and high-achieving low-SES students seem to aspire below their academic potential.

Importantly, we find that differences in salient tracks explain a significant part of the differences in attainable tracks - but not all -, and that differences in attainable tracks explain the main part of the differences in preferred tracks. Hence, low-SES and high-SES students who are in the same class and have similar academic capacity differ in their aspirations predominantly because of differential confidence in their future academic potential, and also partly because of ignorance of some academic tracks. While the lack of confidence of the low-SES students is partly justified by factual constraints imposed by their social background, we further show that it cannot be taken as granted that low-SES students assess their academic potential at its right value. First, students exhibit excessively fatalistic views on the extent to which future academic success is determined by social background. Second, scholastic self-esteem proves socially-dependent: low-SES (in particular the high-achievers) have a 0.13 standard deviation lower scholastic self-esteem than *equally-achieving*

high-SES classmates. Our measure of self-esteem explains 25% of the deficit in high-achieving low-SES students mentioning masters in their attainable tracks. All these findings point to the role of social stereotypes and fatalism in shaping students' perceptions and cognition, with important consequences on educational aspirations and later outcomes.

We then show that the social differences in academic aspirations that we find cannot all be explained by differences in professional aspirations. This is because academic aspirations are not consistent with professional aspirations, which suggests that teenagers do not see education as a pure investment for a future job, probably due to ignorance of educational pathways to jobs. It is true for high-SES students, but it is stronger for low-SES students. For instance, low-SES students are as likely to prefer a job that requires a masters than their equally-achieving high-SES classmates, and only 11% more likely to prefer a job that does not require higher education, which is not consistent with the differences in academic aspirations described above. At 15 years-old, social groups differ much more in the way they plan to invest in education than in the way they think about their future occupation.

Finally, it is to notice that once socio-economic status is taken into account, students from immigrant families have generally more ambitious aspirations (both educational and professional) and similar academic paths than students from non-immigrant families, which points to the fact that the aspiration-based educational trap demonstrated in this paper concerns socio-economic inequalities but not ethnic inequalities.

Our paper lies within the scope of several different strands of the literature. First, it brings new evidence on the social determination of aspirations, a topic on which empirical evidence remains thin (Sewell et al. 1969, Sewell et al. 1970, Jencks et al. 1983, Hoxby and Avery, 2013). Thanks to our rich survey data, we make a strong contribution to this literature by exploring detailed measures of aspirations in the short, medium, and long term, and by enlightening several mechanisms thanks to the distinction between salient, attainable and preferred tracks. Our paper also adds to the empirical literature on the consequences of aspirations on economic outcomes (Oyserman et al. 2006, Beaman et al. 2012, Bernard et al. 2013, and Goux et al. 2016) by exploring the consequences of aspirations on academic progress and track assignment. As our estimates are robust to the inclusion of important confounding factors, we argue that an important part of this relationship can reasonably be considered causal. This paper also contributes to the behavioral economics of education that points to various biases in the way students make decisions on education (see Oreopoulos 2014 for a review). Our findings speak more specifically to the literature on the role of social identity in economic behavior (Akerlof and Kranton 2000, 2002, Hoff and Pandey 2006, 2012; Hoff and Fehr 2011, Hoff and Stiglitz 2010 and 2016). Because we are able to distinguish between salient, attainable and preferred tracks, and to measure scholastic self-esteem and social fatalism, we show that identity-based preferences are formed partly

on the ground of differential knowledge of existing tracks and differential self-perceived academic potential, and cast doubts on the rationality of these preferences.

The policy implications of this paper are important. Since low-SES students have lower aspirations than high-SES students, they are more likely to suffer from the negative consequences of low aspirations on top of the negative consequences of their social origin, which creates rapid divergence of school outcomes between high- and low-SES students. Therefore, educational interventions aiming at reducing social inequalities in academic performances by improving school quality in disadvantaged areas for instance, like extra-tutoring, increased social diversity, or reduced class size programs, are necessary but not sufficient to close the gap and increase upward mobility. They must be combined with actions that strengthen the capacity to aspire, otherwise aspiration failures will continue to dampen the academic outcomes of the disadvantaged. However, strengthening the capacity to aspire is *not* equivalent to raising aspirations for all low-SES students, which may produce worse outcomes among the weakest students (Goux et al. 2016). The reduction of social inequalities in education requires that aspirations be adjusted in ways that bring them in line with students' real potential, independent of their social backgrounds, which may imply downward adjustments for low-achieving students and upward adjustments for medium and high-achieving students, especially the low-SES.

The remainder of the paper is organized as follows. Section 2 presents the conceptual framework of aspirations used in the rest of the paper. Section 3 presents the data we use. Section 4 presents our empirical strategy and estimates of the impact of aspirations on later school outcomes. Section 5 presents our empirical strategy and estimates of the social differences in aspirations. Section 6 concludes.

2 Related Literature

This section presents the literature related to our paper. We first present the emerging theory of aspiration-based poverty traps. We then present the empirical literature related to our two sets of findings: the consequences of aspirations on economic outcomes on the one hand, and the social determination of aspirations on the other hand. We finally present the literature on the optimality of socially-dependent preferences.

2.1 Theories of aspiration-based poverty traps

The theoretical literature on aspirations emerged a decade ago at the intersection of anthropology and economics (Appadurai, 2004; Ray, 2006). According to these authors, the capacity to aspire is a “specific future-oriented instance of culture” that is socially determined because the experiences from which one learns “the map to explore the future” are formed in the “thick of social life”. The individuals who populate poor people’s possible selves are different from those who populate rich people’s possible selves because people

use comparisons and similarities with peers and relatives when they form their aspirations. The capacity to aspire is thus inherently unequal between rich and poor. An aspirational trap then occurs when low aspirations induce low investment and effort to better one's life, resulting in poor outcomes. Embedding this theory in a macroeconomic growth model, Genicot and Ray (2015) show that the social determination of aspirations can be the source of divergent income inequalities: aspirations are influenced by society-wide distribution of income in the current generation. In relatively equal societies, aspirations are more equally distributed and easier to satisfy which creates convergence. In contrast, unequal societies make aspirations of the poor more often frustrated, inducing lower aspirations, investment and growth for the poor, and widening society-wide inequalities. In this first class of models, social background has a direct impact on aspirations.

Dalton et al. (2016) develop a different model in which aspirations are not inherently socially determined but still participate to a poverty trap: at a given initial aspiration level, a poor person will choose a lower level of effort than a rich person because poverty imposes external constraints that make effort less productive. This lower effort induces lower realized outcomes, which results in lower aspirations in the next period. Hence, the aspiration level of the poor person diverges from the aspiration level of the rich person, and so do the realized outcomes, starting a vicious circle that locks individuals in a poverty trap. These theories of aspiration-based poverty traps both draw on a common dynamic going from aspirations to effort, realized outcomes, and back to aspirations, although they differ essentially in the action of social origin: either directly on aspirations (Appadurai 2004, Ray 2006, Génicot and Ray 2015), or on returns to effort (Dalton et al. 2016). Our paper provides the first empirical evidence on these theories.

2.2 Empirical evidence on impact of aspirations on economic outcomes

To the best of our knowledge, four papers provide empirical experimental evidence on the impact of aspirations on subsequent behavior and outcomes² (Oyserman et al. 2006, Beaman et al. 2012, Bernard et al. 2013, and Goux et al. 2016). To draw a causal link between the change in aspirations and the change in outcomes, the effect of the randomly assigned treatment on realized outcomes has to go entirely through its effect on aspirations. In Goux et al. (2016), the treatment³ credibly affects outcomes only through parental aspirations, but students' aspirations are not observed as the experiment was not designed to measure the impact of students' aspirations on later effort and outcomes. The impact of this intervention may go entirely

²Bernard et al. (2011) show that Ethiopian farmers who express fatalistic views also demand less long-term loans and loans for productive purposes, although this correlation is not interpreted as causal since third factors could drive both fatalistic views and investment behavior.

³Parents of low-achieving grade 9 students were invited to an individual meeting with the school provost in which they were informed about vocational and academic tracks existing after middle school, and sensitized to the importance of adjusting expectations to students performance.

through parents' preferences and decisions with no claim on the relationship between students' aspirations and effort at school. And the fact that the intervention has no impact on students' test score at the national examination at the end of the year actually suggests that students' aspirations were not affected.⁴

In the other experiments, although aspirations do credibly play a central role in changing behaviors and outcomes, other effects of the intervention may contribute to the improved outcomes in ways that may not be related to aspirations. In Oyserman et al. (2006), the intervention⁵ affects both aspirations and the capacity to perform better at school. In Bernard et al. (2013), Ethiopian farmers were invited to watch video documentaries about people who had succeeded in agriculture or small businesses which include both a role model effect and an informational effect on how to succeed. Finally, Beaman et al. (2012) show that the reservation of leadership positions for women in Indian village councils increased both girls' aspirations and educational attainment. However, Chattopadhyay and Duflo (2004) show that this policy also affected public good provision -increasing for instance the number of drinking water facilities- which could explain the increase in school participation of girls (given that girls are generally in charge of water duties).

Overall, for now, no intervention affecting only students aspirations with no effect on other elements that can contribute directly to better school outcomes. Our paper does not provide such an ideal experiment but uses a descriptive analysis of the correlation between aspirations and later outcomes *net* of the effect of strong determinants of school outcomes. There are assumptions for causality to hold, but we argue that this contribution improves the existing literature as the parameters we estimate should be close to the causal parameters.

2.3 Empirical evidence on the social determination of aspirations

The empirical literature showing that aspirations are influenced by individuals' social background is quite limited. The first evidence was provided in the 1960s using US data in what is known as the "Wisconsin Model": at equal IQ test score and rank in the class, 11th grade students whose father has a low education level are less likely to aspire to and reach college than those whose father has a higher education level (Sewell and al. 1969). In this seminal paper, the external validity was very limited⁶ and the measurement of academic achievement raised concerns since rank in the class depends heavily on the composition of the class. Additional papers therefore extended this first result using broader population and better measurements of academic achievement - both a test score and teachers' grades (Sewell et al. 1970, Jencks et al. 1983). More

⁴Also, this experiment focuses on low-achieving students only, whereas our paper provides results on the entire distribution of students.

⁵12 individual meetings providing low-SES students with new "Academic Possible Selves" together with strategies to attain these selves, meaning strategies to perform better at school.

⁶The sample has been constructed based on a survey done in 1957 including all high school seniors in Wisconsin. Out of this database, only 929 high school male seniors whose fathers were farmers in 1957 and who accepted to respond to a survey in 1964 are included in the sample.

recently, Hoxby and Avery (2013) show that among the highest-achieving US students (top 4% at college assessment test scores⁷), low-income students are less likely to apply to selective universities than high-income students, although the cost of attending a highly selective university would not have been larger. French sociologists developed a connected literature on social inequalities in track assignment: low-SES French students are less likely to enter selective tracks than high-SES students who have similar academic performances (Girard and Bastide 1963, Duru-Bellat 1988, Davaillon and Nauze-Fichet 2004, Felouzis 2003, Broccolichi and Sinthon 2011). Yet there is little evidence on the mechanisms behind, either teacher discrimination, parental preferences or pupil inhibition, as students' aspirations are not observed. Our paper contributes to this literature by using unique data on students' aspirations that allow to decompose aspirations in salient, attainable and preferred tracks to understand what drives social differences in aspirations.

2.4 Optimality of socially-dependent preferences

The identity literature provides various explanations for socially-dependent preferences leading to different views on their optimality in terms of welfare. A first class of theoretical models reconcile low economic outcomes and maximized utility. Individuals may increase their utility by investing in identity-reinforcing attitudes because it limits disruption and maintains a sense of unity (Akerlof and Kranton 2000 and 2002). Identity-reinforcing attitudes is also a way to invest in one's social network in order to secure interactions and cooperation in the future (Fryer 2007, Fang and Loury 2005), or to signal and affirm values and beliefs to avoid cognitive dissonance (Benabou and Tirole 2011). Finally, resisting education or alike puzzling attitude may be a way to fight the threat of losing one's culture (Carvalho and Koyama 2014). In the view of these papers, identity-based behaviors may be detrimental in terms of economic outcomes but still optimal for individuals, since they get the highest utility from economically-lower, identity-conformed, outcomes.

Another class of papers, from the behavioral economics literature, see the role of identity as driven by non-rational (or unconscious) factors, leading potentially to sub-optimal behaviors and outcomes. Avery and Kane (2004) and Oeropoulos and Dunn (2013) find that low investment in higher education by low-income students can be partly attributed to misinformation about its returns and costs. Hoff and Pandey (2006, 2011) and Hoff and Fehr (2011) show that identity shapes preferences due to stereotype susceptibility and point to the risk that endogenous preferences perpetuate social inequalities. Hoff and Stiglitz (2010) build a theoretical model of identity-based poverty trap where beliefs related to social inferiority affect the *perceived* probability of success (or self-confidence) and so change behavior in ways that make the beliefs come true. Hoff and Stiglitz (2016) describe how social identity creates mental models affecting how an individual

⁷This paper thus focuses on a very specific group of high-school students, whereas our paper provides results on the entire distribution of middle-school students.

experiences what he experiences, making decision-makers *enculturated* actors whose preferences, perception and cognition are subject to deep social influences. Our paper contributes to this literature by isolating the role of information and perception in shaping social differences in educational preferences.

3 Context and Data

3.1 Background on French Education System

In France, the curriculum is the same across schools from kindergarten to the end of junior high school. Junior high school runs from grades 6 to 9. After finishing junior high school, 60% of pupils enroll in academic high school while 40% of pupils enroll in vocational high school (Afsa, 2009). Academic high schools are more selective than vocational high schools: the distributions of test scores at the end of grade 9 show that students who enroll in academic high schools have much better academic performances than students who enroll in vocational high schools (Figure 1). Academic and vocational high schools also differ in their link to higher education. Academic high schools do not provide a professional degree, so students are expected to get some higher education: in fact, 92% of students who graduate from academic high school enroll in higher education⁸ (Afsa, 2009). In contrast, vocational high schools provide students with a professional degree allowing them to find a job with no further education: only 25% of students who graduate from the vocational 3-year track get some higher education, while no students in the vocational 2-year track enroll in higher education⁹ (Afsa, 2009). The early specialization in vocational high school makes later track changes difficult, and many higher education pathways are not accessible to students in vocational high schools¹⁰.

Hence, the choice that ninth graders make between academic and vocational high schools is a crucial milestone with important consequences for final educational and occupational attainment. The procedure of track assignment starts in the middle of grade 9 and ends in June. At the end of the Winter term, in March, families indicate their preference to the teacher conference (academic track, 3-year vocational track, 2-year vocational track, or grade repetition), and the teacher conference expresses an opinion on this

⁸Within academic high schools (*Lycée Général et Technologique*), 67% of students graduate from the *Général* track, of which almost 100% get some higher education, while 33% of students graduate from the *Technologique* track, of which 75% get some higher education (Afsa, 2009).

⁹Access to higher education requires the completion of a *Baccalauréat*; thus 2-year vocational-track (*Centre de Formation par l'Apprentissage*) students do not access higher education (their diploma is a *Certificat d'Aptitude Professionnelle*). 3-year vocational-track (*Lycée Professionnel*) students have formal access to universities with their professional *baccalauréat*, but they are not prepared for university, so in practice less than 5% do enroll in a university. The other 20% enroll in 2-year technical programs.

¹⁰The curriculum in academic high school is common across students in the first year (grade 10); then students choose among 10 different tracks in grades 11 and 12: literature, social sciences, and sciences constitute the *Général* track, while management, industrial technology, health, laboratory science, art, life sciences, and hospitality constitute the *Technologique* track. All academic tracks end with the graduation exam *Baccalauréat* and give access to any higher education pathways, although access to the more selective pathways is conditional on performance (teacher grades) and curriculum (for instance, engineering schools cannot be accessed by students who graduate in literature). In contrast, vocational high schools offer a large number of tracks as soon as grade 10, varying in terms of number of years of education (2-year or 3-year tracks), topics (construction, sanitation, mechanics, electrical technician, commerce, secretaries, agriculture, and other services), and pedagogy (with or without apprenticeship).

preference¹¹. At the end of the school year, in June, families choose a track and the teacher conference validates or invalidates this choice based on students' performances. If teachers invalidate the choice of the family (which occurs in about one case out of four according to Caille 2005), the family meets with the provost. If disagreement persists, the family can request an appeal committee whose decision is definitive. An important feature of this procedure is the leading role of families, who move first. The legal framework insists on the idea that teachers' role is corrective and must respect families' preferences and responsibility unless student performances are not compatible (Caille, 2005, p.78). Note also that the procedure of track assignment does not take into account students' performance on the national exam that is given at the end of June and anonymously graded in July. Students' performances are thus assessed on the basis of teachers' grades over the course of the academic year.

3.2 Data

3.2.1 Junior High School Sampling Strategy

59 junior high schools from the three educational districts of the Paris metropolitan area participate in the study: 6 in the Paris district, 15 in the Créteil district (east of Paris), and 38 in the Versailles district (west of Paris). The sampling strategy was not random, so our sample is not representative of the French nor of the Parisian junior high schools. The main reason is that the school provosts had to agree to participate in the study, which means that the junior high schools in our sample are headed by provosts who may be more concerned by the topic of the study than the provost of the average junior high school. We also followed two research-based selection criteria independent of the purpose of this paper and linked to the “quality” of the school in terms of success on the national exam at the end of junior high school¹². In the present paper, we focus on the effect of social background *within* classrooms, so these features of our sampling strategy are neutral for the analysis except that they affect the representativity of the sampled population. Our sample is indeed different from the national junior high school population in terms of school social composition: our sample over-represents junior high schools in which 40-60% of students are low-SES students, and those in which more than 90% are low-SES students, at the expense of the most advantaged (less than 40% low-SES) and intermediary (60-80% low-SES) junior high schools (Figure 2). This implies that our results may not generalize to all areas of France.

¹¹The legal framework is available at <http://eduscol.education.fr/pid23597-cid53993/textes-reference.html>.

¹²First, we favored junior high schools with low and high success rates rather than those in between, while excluding outliers with extremely low and high rates. Second, we selected our sample so that the geographical location of the schools ensured that their students had equivalent access, on average, to any educational track, both at the high school and higher education levels, to rule out the effect of the supply of education as a determinant of aspirations. The point is to be able to study some specific “school context” effects in another paper.

3.2.2 Data Sources

Data come from two sources: (i) a research survey administered to ninth graders in the sampled junior high schools, and (ii) administrative data collected by the statistical unit of the Ministry of Education¹³. The administrative data contains information about parental SES, teachers' grades averaged over grade 9, score at a national standardized, anonymously graded test that students took in June 2013 at the end of grade 9, and track assignment in September 2013 at the beginning of grade 10. The research survey was administered in November 2012 in two parts. First, students took a math test consisting of seven exercises covering grade 8's math curriculum. The test was administered in class by one of their teachers. However, students were informed that the math test would be graded by independent researchers, and that their scores would be kept strictly confidential. Second, one week later, students completed a 50-minute questionnaire to assess their educational aspiration windows, their educational aspirations, and their occupational aspirations¹⁴. The questions were designed to be open-ended so as to capture as truly as possible students' attainable and preferred academic tracks and occupations: the measure of the aspiration window is not distorted by a provided set of existing tracks. For instance, the use of an MCQ format could have increased the salience of some tracks that low-SES students do not naturally consider.

3.2.3 Student Sample

Of the 6,903 students registered in the 59 junior high schools of our sample, 5,672 completed both the math test and the questionnaire: the response rate was 87% in the first visit and 88% in the second visit, resulting in a combined response rate of 82%. Attrition is due to student absenteeism, which may result from sickness, voluntary class skipping, or, in very few cases, refusal to take the test and/or the questionnaire. We matched math tests and questionnaires in class after the completion of the questionnaire and then anonymized. We were not authorized to collect any students' identifiers in our independent test or survey¹⁵, so we matched the test and survey data with administrative data using school, class, year of birth, month of birth, and parents' socio-economic status¹⁶. Since this information does not constitute a unique identifier, duplicates were dropped, as well as observations with incomplete information for these characteristics. 67% were matched, resulting in a sample of 3,789 students. Finally, 10% of these students were missing values for their track

¹³MEN-MESR, Direction de l'Evaluation, de la Prospective et de la Performance, "Bases Scolarité" 2012 and 2013, and "Base DNB" 2013.

¹⁴The questionnaire was administered early in grade 9 in order to capture students' aspirations at a point of time when discussions about track assignment at school had not yet started. In particular, they might have discussed track assignment with their parents, but the family would not have made a formal choice yet. Moreover, no information about teachers' opinions is provided during the first term, so when students took the survey they were unlikely to know what teachers thought about their track assignment.

¹⁵To avoid a breach of confidentiality, we did not collect names, administrative identifiers, or complete dates of birth.

¹⁶Students were asked in the questionnaire to report their parents' occupations. We used the administrative classification of occupations to code parental SES in order to get the same variable as in the administrative data.

assignment in September 2013. The reasons for this attrition are threefold: first, some students moved to an educational district outside of Paris, Créteil, and Versailles for which we have no data; second, some students dropped out of school and do not appear in any administrative dataset; third, some students enrolled in independent private schools, which do not report information to the Ministry of Education. These students were dropped from the study since we are interested in the full trajectory going from initial aspirations and academic performance to later academic performance, grades, and track assignment.

Our final sample thus consists of 3,415 students, nearly half of the students registered in the sampled junior high schools in grade 9. Attriters are students who are more likely to be absent; who have, by chance, a classmate sharing the same month of birth, sex, and parental SES; and who are more likely to move outside the Paris region, stop education, or enroll in the private sector. We do not claim in any way that the resulting student sample is representative of the original junior high school population, and acknowledge that our findings on the role of aspirations may not apply to attriters. However, non-attriters look quite similar to the initial population in terms of family background (68% low-SES in both groups), test scores in June 2013 (144 versus 141 points), yearly grade average (85 versus 82 points), gender (52% versus 51% girls), probability of having repeated a grade (22% versus 23.5%), and probability of having skipped a grade (4% in both groups), none of these differences being important and significant.

3.2.4 Variables of Interest

Educational Aspirations: Salient, Attainable and Preferred Tracks Following Ray (2006), we define aspirations as the preferred track within the zone of attainable tracks. Students were asked first what tracks they know (salient tracks), then, among these tracks, which ones they feel capable of pursuing (attainable tracks), and finally, among these tracks, which one they prefer (preferred track, or aspiration). By construction, the preferred track is included in the attainable tracks¹⁷, itself included in the salient tracks¹⁸. These questions were asked first for high school tracks, and then for higher education tracks. The questions were totally open, which represents a fundamental value of this paper.

We coded and aggregated students' answers to create dummies indicating whether a given track appears in the self-declared set of salient (resp. attainable, preferred) tracks. At the high school level, the dummies indicate whether “No high school track is in the answer”, “Vocational high school is in the answer”, and “Academic high school is in the answer”. At the higher education level, the dummies indicate whether “No higher education track is in the answer”, “1-4 years college is in the answer”, and “5 years college or more is

¹⁷There are few exceptions since it may be natural for students who are asked about the tracks they feel capable of pursuing to just answer the “highest” (most selective) track they feel capable of pursuing, instead of all of them.

¹⁸A fifth of students were randomly chosen within each class to take a different questionnaire which includes a list of existing tracks both at the high school and higher education levels. These students were not asked what tracks they know, but directly what tracks they feel capable of pursuing. The number of observations for salient tracks is thus smaller than for attainable and preferred tracks.

in the answer”. We report in the Data Appendix detailed information on data construction.

On average, students report 3.9 salient tracks, 1.4 attainable tracks, and 0.9 preferred tracks at the high school level (including 10% of students reporting 0 tracks). 99% of students report between 0 and 5 attainable tracks at each level. At the higher education level, on average students reported 1.9 salient tracks, 0.9 attainable tracks, and 0.8 preferred tracks (including 44% of students reporting 0 tracks). Students who provided several preferences (11% at the high school level and 16% at the higher education level) may occupy several categories. Finally, the questionnaire included a question on whether the student preferred to find a job after high school or pursue higher education. We use a dummy indicating whether the student prefers to find a job after high school as an additional measure of higher education aspirations.

Professional Aspirations Students were asked which job(s) they would like to have. On average, students provided 1.7 jobs (including 20% who provided no job). We coded jobs according to the number of years of education required to practice them and created dummies indicating whether the student entered the categories “No Response”, “No higher education”, “1-2 years college”, “3-4 years college”, or “5 years college or more” (see the Data Appendix for more information on these variables).

Academic Capacity in November 2012 The starting point of this paper is that aspiration is the capacity to set goals for the future that are in line with one’s potential. A key variable is thus students’ academic capacity, that we will proxy by their current capacity at performing at academic tasks, together with other measures of cognitive and non-cognitive skills.

Academic Test Score in November 2012 This paper uses academic test scores in November 2012 to measure academic performance at the same time as aspirations, in order to test how much academic aspirations depend on current academic performances. Two tests (one 45-minute and the other 20-minute) were administered in class by the research team in November 2012 one week apart. The advantage of using two independent tests administered on two different days one week apart is to limit measurement errors due to random performance variations related to topics and personal temporary dispositions. These tests were anonymously and externally graded so scores are unaffected by teachers’ beliefs and less prone to stereotype threats and parental inputs than teachers’ grades. In fact, teacher grades incorporate social factors unrelated to individual academic capacity: (i) grades include at-home assignments for which parents provide direct help, and parents also provide preparation for in-class tests; (ii) teachers’ beliefs about students’ capacity according to their SES may bias their assessments (Hanna and Linden, 2012; Merle, 1998); (iii) stereotype susceptibility may affect student performances in class (Hoff and Pandey, 2006; Steele and Aronson, 1995, Croizet et al.

2001, 2004; Dee, 2014). The tests administered in November 2012 were meant to limit these social factors: they were clearly disconnected from any academic stake, were not going to be graded by teachers, and were explicitly anonymously graded. Moreover, the tests - which focus on math - were generally not administered during a math class. These precautions were to reduce the stereotype threat and the loss in self-confidence that could be associated with it for low-SES students. None of these tests were announced to avoid pupil preparation - as it is likely that high-SES parents encourage and support more at-home preparation than low-SES parents. We average individual scores at the two tests, and then use the relationship between the total score and the score in math at the national exam administered in June (see next paragraph) to estimate the total score that a student would have had in November 2012 given her average math score at our two tests and her invariant characteristics (see the Data Appendix for more information on this variable). We are confident about the quality of the measure of academic performance in November 2012 since the two tests were administered in strict and rigorous conditions, and designed to avoid ceiling effects, as shown on Figure 1. The correlation between the test scores in November 2012 and in June 2013 (see below for the definition of test scores in June 2013) is 0.78.

Additional Proxies of Cognitive and Non-Cognitive skills in November 2012 We use the Raven Progressive Matrices (RPM) test to measure students' cognitive ability independent of any knowledge from the past. It is a non-verbal multiple choice test. Participants have to complete a series of drawings by identifying relevant features based on the spatial organization of an array of objects, and choosing one object that matches one or more of the identified features. This task assesses the ability to consider one or more relationships between mental representations or relational reasoning (Raven 2003). The RPM is one of the most commonly used measures of fluid intelligence, or fluid reasoning, which includes such abilities as pattern recognition, abstract reasoning, and problem-solving. Psychologists see fluid intelligence (Gf) as one of the factors of general intelligence (see Jensen 1998 for a review). People with a high capacity of Gf tend to acquire more knowledge and at faster rates (Cattell 1971).

We also use a dummy indicating that the student has repeated a grade (born before 1998), and a dummy indicating that the student skipped a grade (born after 1998). Indeed, the past decision to repeat or skip a grade must be based on additional observations on students' cognitive and non-cognitive skills that the researcher cannot observe in the current test scores¹⁹. In the same vein, we use a dummy indicating that the student is a female since the literature shows that teenage girls are more self-disciplined and conscientious than boys (Duckworth and Seligman 2006). Information on year of birth and gender comes from administrative

¹⁹The French procedure to skip a grade includes IQ tests and psychological interviews to assess precisely cognitive and non-cognitive skills of the student.

data. In our sample, 51.6% of students are female, 5% have skipped a grade, and 22% have repeated a grade.

Finally, we use two measures for students' behavioral conduct: first, the number of questions that students attempted to solve in the two academic tests (whether they succeeded or not), which reflects the effort that the students invested in taking the tests²⁰. This measure may capture diligence and dutifulness, especially in this context where the test entails no challenge at all. The second measure is the score of self-perceived behavioral conduct based on the "Self-Perception Profile for Adolescents" (SPPA) scale developed by Susan Harter (Harter 1988) in its French version (Bariaud 2006). The "Behavioral Conduct" subscale taps the degree to which one likes the way s/he behaves, does the right thing, acts the way s/he is supposed to act, and avoids getting into trouble.

Later School Outcomes

Academic Test Score in June 2013 We use test score in June 2013 to measure academic performance at the end of the school year and the influence of early aspirations on later performances. Individual test scores at the national exam administered in June 2013 are reported in the administrative data. This test is also anonymously and externally graded test score, and includes math, French, and history. This test has a higher academic resonance than the test administered in November: it is a national exam organized by the Ministry of Education and taken by all pupils at the end of Junior High School. However, the test has no impact on later academic paths: decisions on assignment to high schools are made before the test, and it is not necessary to pass the test to enter high school.

Average yearly teachers' grade We also use average teachers' grades in grade 9 to test the influence of early aspirations on later performances. The average yearly grade is the average of all grades a student received from all teachers over grade 9 -from September to June. Individual average yearly grades are reported in the administrative data. In our empirical analysis, we use class fixed-effects to account for between-classroom and between-school variation in the grading system. The advantage of this measure is that it provides more precise information about academic performance than a single test since it smooths random volatility in individual single performances. While test scores reflect are not observed by any agent, the average yearly grade adds an interesting ingredient to the analysis since it reflects academic performance *as observed* by students, parents, and teachers themselves at the time of assignment to high school. The average yearly grade also incorporates more amply academic *behavior* compared to test scores: teachers' grades take into account manifestations of effort like delays in assignment submission, in-class attitudes and

²⁰For each question of the math tests, a student is considered as having "attempted to solve the question" if s/he wrote down some calculation or if provided an answer, no matter if the answer is right or not.

participation, etc (Avvisati et al. 2014). However, as mentioned above, average yearly grade also incorporates social factors like parental inputs or teachers’ priors related to social stereotypes. We see here that the ideal measure of academic performance does not exist, and both average yearly grade and test scores in June 2013 have advantages and caveats. This paper uses both to make sure the findings are robust to these different approaches of academic performance.

Track Assignment Track assignment is observed in the administrative data. We use dummies indicating whether the student “Entered vocational high school”, “Entered academic high school”, or “Repeated grade 9”.

Family Characteristics

Family Socio-Economic Status The administrative data contains socio-economic status of each parent (more precisely guardian). The socio-economic status is coded on a 32-code scale, each code being a two-digit number. In this paper we construct two classifications of the family socio-economic status: a rough classification containing two categories, and a detailed classification containing six categories: “No parent has ever worked”, “Maximum family SES is manual laborer”, “Maximum family SES is low-skilled white-collar”, “Maximum family SES is craftsman or storekeeper”, and “Maximum family SES is intermediate occupation” (more information in the Data Appendix). Overall, 31% of the families are in the high-SES category; the low-SES category contains the other 69% of families in which both parents have intermediate or low-skilled occupations. As additional information about family context, we also use dummies indicating each parent is unemployed, or whether each parent is retired (note that while information on parental SES comes from administrative data, information on parental employment status comes from our research survey).

Immigrant Family We use a dummy indicating that *both* parents were born abroad, and use it as a control variable. Data on parents’ country of birth comes from our research survey. In our sample, 38% of families are immigrants (of which 60% come from Africa²¹), and 88% of immigrant families are low-SES families, so immigrant families are largely a sub-group of low-SES families. Following Caille (2007), who shows that immigrant families have higher aspirations for their children than non-immigrant families, our main specification uses the immigrant dummy as a control variable to capture the systematic difference between immigrant and non-immigrant families when it comes with academic and professional aspirations.

²¹30% come from Northern Africa, 30% from Sub-Saharan Africa, 12% from Asia, 7% from the Middle East, 7% from the Caribbean, 5% from Portugal, 4% from Eastern Europe, 3% from Latin America, 2% from other European countries, and 0.5% from North America.

Students' Perceptions

Scholastic Self-Esteem Students' self-perception of their scholastic competence, or "scholastic self-esteem", is measured using the "Self-Perception Profile for Adolescents" (SPPA) conceived by Susan Harter (Harter 1988). The "Scholastic Competence" subscale includes five items: "being just as smart as others", "doing school work quickly", "doing well at class work", "feeling pretty intelligent", and "almost always figuring out the answers in class". The answer to each item is coded as an integer between 1 and 4, 1 meaning the student feels she is not at all like that, and 4 she feels she is very much like that. Our measure of scholastic self-esteem uses the standardized score over all five items.

Fatalism Students were asked to assess the probability of success in high school of a hypothetical *high-achieving* student on a scale from 0 to 10. The hypothetical high-achieving student was presented in different situations: "s/he lives in an advantaged neighbourhood", "s/he lives in a disadvantaged neighbourhood", "his/her parents are immigrants", and "one of his/her family member succeeded academically and professionally". We use the gap in assessed probabilities of success of the hypothetical high-achieving student between the first situation and the next ones as measures of fatalism: the larger the gap, the more the social background influences future academic path independently of current academic performance, so the more fatalistic the student.

4 Do Aspirations Matter?

In this section, we examine how aspirations influence later school outcomes. While estimating the causal link from aspirations to later school outcomes would require a randomized controlled manipulation of aspirations that we don't have, our survey data provides exceptionnally rich control variables. We also use residual variations in salient tracks as instruments for aspirations. We thus argue that our estimates are close to the causal ones and that this paper is a strong addition to the emerging empirical literature presented in Section 2.3 showing that aspirations do influence later economic outcomes.

4.1 Empirical Strategy

4.1.1 OLS Estimates

Our dependent variables are three indicators of later school outcomes: average teachers' grades in grade 9, academic test scores in June 2013, or effective track assignment in September 2013. The explanatory variables of interest are aspirations at high school and higher education levels in November 2012. Our analysis focuses on the effect of the most ambitious aspirations: academic high school at the high school level

(versus vocational or no relevant response), and masters at the higher education level (versus no relevant response, no higher education, or max. 4 years in higher education). Obviously, the naïve correlation between aspirations and later school outcomes does not reflect the causal impact of aspirations as aspirations and later school outcomes are both the product of other common factors.

What sort of confounding factors do we worry about? Any factor that affects both aspirations and later school outcomes *separately from aspirations*. There are three categories of such factors: school-neighborhood characteristics²², family characteristics, and student academic skills (at least those who have an influence on school achievement). School and neighborhood characteristics can be neutralized by the use of class fixed effects. For the two other categories of factors, neutralization is more challenging.

Our first control variable for academic skills is contemporaneous (November 2012) school performances as measured by our independent academic test scores. Academic performance is the product of both intellectual ability (cognitive skills) and efforts put in learning (displeasing or enjoyable), which result from non-cognitive skills like motivation, persistence, self-esteem, diligence, etc. In Dalton et al. (2016) model, present academic performance represents the “realized outcome” at the basis of aspirations. Beside, present academic performance is a strong predictor of later academic performance, the only two disturbing elements being natural noise (measurement errors) and differential academic progression. We worry about differential progression as long as it is *not* the consequence of differential aspirations, but the consequence of differential skills that do not express in November 2012 academic test scores.

The factors that may influence academic progression over grade 9 separately from aspirations are intelligence and Conscientiousness (in the terminology of the Factor Five Personality model). In fact, the literature shows that more intelligent and more conscientious pupils learn at faster rates than the others (Poropat 2009)²³. Neuroticism does also correlate with school achievement as shown by Heckman et al. (2006)²⁴ but we postulate that Neuroticism is *not* a confounding factor because this domain is inherent to aspirations: traits related to this domain, e.g. self-evaluation, self-esteem, self-efficacy, and optimism, affect later school outcomes only through their interaction with aspirations. For instance, low self-esteem causes inhibition, which we take as inseparable from low aspirations. Similarly, fatalism reduces the perceived return to effort and limits ambition and motivation, which also cannot be distinguished from low aspirations. Optimism works the opposite way on ambition and motivation, but also inseparably from high aspirations. In this paper, we propose that traits related to Neuroticism affect later school achievement in ways that

²²Schools in this study are all public schools whose population is defined by zoning. As a consequence, pupils in the same school are all living in the same neighborhood.

²³In contrast, Poropat (2009) does not find that Openness, Extraversion, and Agreeableness correlate with school achievement once intelligence is controlled for.

²⁴This paper focuses more precisely on self-esteem and internal/external locus of control, which both enter into the Neuroticism domain.

cannot be distinguished from aspirations. The two remaining factors that may bias our estimates are thus general intelligence and conscientiousness not expressed in November 2012 academic test scores.

In order to address this concern, we add additional control variables and examine the sensitivity of our estimates to their inclusion. We first add the additional proxies of cognitive and non-cognitive skills: scores at the RPM test (a proxy for general intelligence), a dummy for whether the student ever repeated, and a dummy for whether s/he ever skipped a grade, a dummy for female students, dummies indicating the number of questions of the two academic tests that the students tried to solve successfully or unsuccessfully (a measure of the effort invested in taking the test), and dummies for levels of self-perceived behavioral conduct.

Regarding family characteristics, our control variables include the socio-economic category of both parents, their employment status (employed, unemployed, or retired) and the region of birth of both parents. These variables proxy parental education and income levels, as well as knowledge of French curriculum. However, we do not claim that these variables fully capture the impact of family on both aspirations and later school outcomes since there should remain variations in parental involvement in their child education once these variables are taken into account (like homework assistance, monitoring of homework schedule, or management of sleep time). An additional variable that captures part of this remaining variation is whether the student repeated or skipped a grade, since higher parental involvement is associated with smaller probability of grade repetition and larger probability of grade skip²⁵. To test the sensitivity of our results to family characteristics, we use two sets of family characteristics. The first set includes only basic variables: a dummy for low-SES indicating that none of the parents are high-skilled workers, and a dummy of immigrant family indicating that both parents are born abroad. The second set includes more detailed variables: for each parent, five dummies for each sub-category of low-SES occupation, two dummies indicating that the parent is unemployed / retired, and two dummies indicating that the parent is born in a non-OECD country / colored (based on the country of birth). The detailed set of controls should better capture family involvement in education than the basic one. Therefore, the sensitivity of the coefficients on aspirations to replacing the basic set by the detailed set indicates whether unobserved family characteristics seem to bias our estimates.

As a result, we run regressions of the form:

$$Y_{ijt+1} = \alpha + \beta Asp_{it} + \sum_{d=2}^{10} \gamma_d TestScore_{dit} + \lambda OtherSkills_{it} + \delta FE_j + \mu Fam_i + \epsilon_{ijt} \quad (1)$$

, where Y_{ijt+1} is average yearly grade, test scores in June 2013, or effective track assignment in September 2013 of student i in class j . Asp_{it} is a vector of two dummies indicating that academic high school and

²⁵In France, teachers propose to the family that the child repeats a grade but parents can refuse, which may happen more often when parents are more involved in education. Parents can also ask that their child skips a grade - which may or may not be accepted by the teachers, but definitively increases the probability of grade skip. Conditional on present academic performance, variations in grade repetition and grade skip should thus correlate with parental involvement in child education.

masters are included in student i 's preferences in November 2012, the reference group (33%) being students who do not aspire neither to academic high school nor to masters (note that 93% of those who aspire to masters also aspire to academic high school). $TestScore_{dit}$ is a dummy indicating that academic test score in November 2012 is in decile d . $OtherSkills_{it}$ is a vector including RPM test scores, dummies for grade repetition and skip, gender, and dummies indicating effort invested in taking the test (number of questions that the student tried to solve). FE_j are class fixed effects. Finally, Fam_i is a (either basic or detailed) set of family characteristics described above.

4.1.2 IV Estimates

To address the remaining issues of measurement errors and omitted variable bias, we use residual salient tracks as an instrument for preferred tracks. Since information about tracks is likely to correlate with school characteristics, family characteristics, and student skills, our identification strategy relies on the lighter assumption that *conditional* on class fixed effects, observable family characteristics and skills (same control variables as in the above equation), residual salient tracks are accidental. We postulate the existence of situations where students are exposed to information in ways that do not directly influence their future academic performance. Such situations occur as soon as a student is in contact with someone that may share her experience without having a direct influence on academic performance (for instance friends of the parents, after school activity leader, adult neighbors, vacation acquaintances, etc.). The result is that residual salient tracks may vary for reasons that are orthogonal to academic performances, which creates a valid instrument for preferred tracks. We understand that the instrument is not perfect as variations in residual salient tracks are not experimentally random and requires some assumptions.

4.1.3 Shorter-Term OLS Estimates

We finally provide shorter-term estimates of the impact of aspirations on June 2013 academic test scores and September 2013 track assignement by using deciles in average yearly grade as additional control variables in equation 1. The interpretation of this estimate is different in two ways: first, it leaves a very short amount of time for aspirations to influence school outcomes since we compare students who have not only similar characteristics in November 2012, but also similar average teachers' grade *all year long*. Second, it excludes the impact of aspirations on June 2013 academic test scores and September 2013 track assignement going through average yearly grade. The estimate gives thus a lower bound of the impact of aspirations. The advantage is an excellent control for student skills and family characteristics, as average teachers' grade over the year is based on multiple academic tests on all topics including both in-class and at-home assignments, as well as student behavior and conscientiousness. Both measurement errors and omitted variable bias are

therefore well addressed.

4.2 Results

4.2.1 Aspirations and Average Yearly Grade

Table 1 presents the influence of aspirations on average yearly grade. Column 6 shows our preferred estimates of the impact of aspirations on average yearly grade where we compare students who are in the same class, show similar academic, cognitive and non-cognitive measures in November 2012, and are from similar social background. We find that students who prefer academic high school at the beginning of the year have a 0.21 standard deviations higher average yearly grade than those who prefer vocational high school or have no preference. Students who also prefer masters have an additional 0.11 standard deviations higher average yearly grade than those who do not. Both differences are significant at the 1% level.

The previous columns show how the addition of our control variables affects the estimation of the impact of aspirations on average yearly grade. November 2012 academic test scores explain a big part of the variation in average yearly grade (the R-squared goes up from 0.15 to 0.45), and reduce by 58% the coefficient on aspirations (column 2), which confirms that academic performances are a major determinant of both aspirations and later academic performances. The additional proxies for cognitive and non-cognitive skills (column 3) also add to the explanatory power of the model (from 0.45 to 0.51) and reduce the coefficient on aspiration for academic high school in a substantial way (from 0.29 down to 0.22). Beyond that, the addition of class fixed effects and family characteristics does improve the explanatory power of the model but does not affect substantially the coefficients on aspirations which is still 0.21 (columns 4-5-6). It should be noticed that this is not related to poor quality measures. In fact, Appendix Table A1 shows that the contribution of family characteristics is important when we include them first as control variables (columns 2-3). The lack of impact of class fixed effects and family characteristics on the coefficients of aspirations in Table 1 is thus reassuring since it means that our measures of academic, cognitive, and non-cognitive skills are capturing most of the differences in academic potential. We are thus confident that our estimates of the impact of aspirations on average yearly grade can be interpreted as causal.

Interestingly, all control variables are highly significant and influence the average yearly grade as expected. The only notable exception is immigrant family, which entertains no particular relationship with average yearly grade once initial academic skills, class fixed effects and family characteristics have been taken into account. This finding - confirmed in Table 2 and 3 - suggests that being born in an immigrant family does not affect academic progress and track assignment to high school separately from initial academic performances, school quality and family SES. Besides, subgroup analysis shows that the results hold for all students whatever their SES (Appendix Table A4). If anything, the impacts of aspirations are larger for

low-SES students than for high-SES ones. The results also hold for all three terciles of initial academic test score (results available upon request). Finally, Column 7 gives the IV estimates using residual salient tracks as instruments of aspirations. The coefficient on aspirations is larger, 0.62, suggesting that measurement errors may bias downward the OLS estimates.

4.2.2 Aspirations and Academic Scores in June 2013

Table 2 shows the same patterns for the effect of aspirations on academic scores at the national exam in June 2013. Controlling for initial academic capacity, class fixed effects, and family characteristics, students who aspire to academic high school have a 0.17 standard deviation higher score in June 2013 than students who prefer vocational high school or have no preference (column 6). Students who also prefer masters have an additional 0.09 standard deviations higher score in June 2013 than those who do not. Our proxies for academic capacity capture almost all the omitted variable bias, as the addition of school and family characteristics increases the explanatory power from 0.65 to 0.71 when included last (and from 0.17 to 0.42 when included first as shown in Appendix Table A1) but does not affect substantially the coefficients on aspirations (columns 4-5). Again, the results are qualitatively similar within SES (Appendix Table A4) and within terciles of initial academic performance (results available upon request). The IV estimate is larger at 0.35, suggesting a downward bias of the OLS estimates due to measurement errors (column 6).

Table 4 columns 1-3 show the shorter term effect of aspirations on academic scores in June 2013 controlling for average yearly teachers' grade. Our preferred estimate in column 3 shows that students who are in the same class and show similar social background and academic skills *both* in November 2012 and all year long, have different scores in June 2013 depending on their initial aspirations: those who initially aspired to academic high school obtain a 0.065 standard deviation higher score than those who preferred vocational high school or had no preference. Initial aspirations thus impact performances at the national exam even when school performances all year long are taken into account.

4.2.3 Aspirations and Track Assignment in September 2013

Table 3 shows that aspirations are a strong determinant of track assignment in September 2013. Controlling for initial academic capacity, class fixed effects, and family characteristics, students who initially aspire to academic high school have a 18 percentage point (37%) higher probability to enter in academic high school the next year than students who prefer vocational high school or have no preference (column 6). As shown in Appendix Tables A2 and A3, students who initially aspire to academic high school have a symmetrically lower probability to enter in vocational high school the next year than those who prefer vocational high school or have no preference: -0.174 off 0.482 (we find no difference in the probability to repeat, which concerns only

4.6% of students). Aspiring to masters adds no chance to enter academic high school, consistent with the fact that aspiration to masters is not directly relevant to assignment to high school (as noted above, almost all students who aspire to masters also aspire to academic high school). As previously, the coefficients on aspirations remain largely stable to the inclusion of class fixed effects and family characteristics, suggesting that their relationship with track assignment and aspirations is already absorbed in our measures of initial academic capacity (columns 1-5). The IV estimate is also bigger - at 0.378 - than the OLS estimate (column 6). The results hold within SES with a larger effect for the low-SES (Appendix Table A4), and within terciles of initial performance (with a larger effect for the bottom tercile) (results available upon request).

Table 4 columns 4-6 show that about a third of the impact of early aspirations on track assignment is explained by average yearly grade. Once average yearly grade is taken into account, students who initially aspire to academic high school still have a 11 percentage point (23%) higher probability to enter in academic high school the next year than students who prefer vocational high school or have no preference (column 6). This finding shows that track assignment is not based only on available information on academic performances (teachers' grades), and that students' preferences play an important role independently of their performances.

4.2.4 Interpretation

Overall, our estimates suggest that aspirations affect academic paths in two ways: first, aspirations affect academic progress from the beginning to the end of 9th grade. Second, aspirations affect decisions and choices independently of academic performances. These results provide empirical evidence that aspirations determine effort as modelled in Ray (2006), Genicot and Ray (2015) and Dalton et al. (2016): students who have lower aspirations seem to invest less effort in class and at home, and thus achieve less and less compared to initially equally-able classmates from similar social background who had higher aspirations. Low aspirations become a source of disadvantage in their own right. Ray (2006) proposes that aspirations are not optimal when they are either too modest (easily satisfied) or too ambitious (unreachable) because such aspirations induce less effort than intermediate aspirations. Our results indicate that the aspiration gap²⁶ is on average too small rather than too large for both low- and high-SES: those who aspire higher get higher outcomes. Would the average aspiration gap be too large relative to the optimal level of aspirations, academic progress would be found smaller for those who aspire higher. In France, "fatalistic" aspiration failures thus dominate "frustration" aspiration failures. The next question is whether aspiration failures interact with social inequalities, so whether students who suffer from fatalistic aspiration failures are of any social background, or more often of low social background.

²⁶The gap between one's current situation and the aspired situation.

5 Social Inequalities in Aspirations

In this section, we present our results on social inequalities in aspirations and on the mechanisms driving these inequalities. Our analysis focuses on low-SES versus high-SES students, and all tables include a dummy indicating whether students come from immigrant families to distinguish the effect of social background from the effect of immigration status. The effect of the socio-economic status is thus considered independently of whether students are immigrants.²⁷ Actually, Caille (2007) shows that first-generation immigrants have higher aspirations than non-immigrants and our data confirm this result.

5.1 Empirical Strategy

We first want to measure the size of the social gap in aspirations. Family background is determined by the accident of birth. The reasons why aspirations are correlated with parental SES are all consequences of family characteristics correlated with SES: parents' level of education, parents' involvement in child's education, parents' choices of housing and school location, the characteristics of parents' friends and networks, genetics, etc. Any difference in aspirations between low-SES and high-SES students is thus the consequence of these family characteristics correlated with parental SES, like inherent intellectual skills, social networks, parental skills, school quality, and so forth. Here, we are not interested in the pure effect of SES on aspirations *everything else being equal*, but in the *ecological* effect of SES which incorporates all SES-related dimensions (except immigration).

Aspirations are the product of a three-step process: 1. individual knowledge of existing tracks (salient tracks), 2. perceived capacity of pursuing these tracks (attainable tracks), and 3. personal preference (preferred track). We first present raw differences in salient, attainable and preferred tracks between students from low and high-SES families:

$$[Salient/Attain/Pref]Track_{ijt} = \alpha + \beta LowSES_i + \gamma Immigrant_i + \varepsilon_{ijt} \quad (2)$$

where $[Salient/Attain/Pref]Track_{ij}$ is a dummy indicating if Track (vocational high school, academic high school, no track in higher education, finding a job, and masters) is salient to / attainable for / preferred by student's i in class j . $LowSES_i$ indicates that none of the parents are high-skilled workers, and $Immigrant_i$ indicates that both parents are born abroad. We interpret β as the ecological effect of parental SES for both immigrant and non-immigrants families, and γ as the independent ecological effect of being born in an immigrant family for both low and high-SES families.

²⁷As explained in the Data section, 88% of immigrant families are low-SES families, so immigrant families are largely a sub-group of low-SES families.

Then, we want to measure the remaining gap once we control for differences in academic capacity. . Current academic performances and school environment (neighbourhood, teacher and peer quality) are obvious determinants of academic potential, and we don't expect students of different potential to aspire equally. We consider that students who belong to the same class and show similar measures of academic performances, cognitive skills, and non-cognitive skills, have similar present academic capacity²⁸. To control for this, we add to the previous model class fixed-effects FE_j and a vector of present academic capacity, $AcadCap$, which includes deciles of November 2012 academic scores, dummies for RPM scores, dummies for the number of questions that students tried to solve at the two tests (effort put into the test), and deciles of self-perception of behavioral conduct²⁹.

$$[Salient/Attain/Pref]Track_{ijt} = \alpha + \beta LowSES_i + \gamma Immigrant_i + \lambda AcadCap_{it} + \delta FE_j + \varepsilon_{ijt} \quad (3)$$

This provides evidence on the degree to which the raw differences in salient, attainable and preferred tracks between low and high-SES (resp. immigrants and non-immigrants) are explained by academic capacity and school environment. Note that we do *not* assume that similar present academic capacity means similar future academic potential. We will discuss students' anticipations about future academic potential later in the paper.

In this model, the remaining coefficient on low-SES (resp. immigrant) family represents the difference in the probability to prefer (resp. feel capable of pursuing, know) track X for low-SES students compared to same-class equally-performant high-SES (resp. non-immigrant) students. The remaining differences thus come from differences in salience of existing tracks, in self-perceived academic potential, and in preferences, all not related to present academic capacity nor school environment. To disentangle the relative contributions of these three mechanisms, we finally add dummies indicating student i 's salient tracks in the regression of attainable tracks, and dummies for student i 's attainable tracks in the regression of preferred tracks:

$$Attain.Track_{ijt} = \alpha + \beta LowSES_i + \gamma Immigrant_i + \lambda AcadCap_{it} + \delta FE_j + \mu Salient.Track_{it} + \varepsilon_{ijt} \quad (4)$$

$$Pref.Track_{ijt} = \alpha + \beta LowSES_i + \gamma Immigrant_i + \lambda AcadCap_{it} + \delta FE_j + \mu Attain.Track_{it} + \varepsilon_{ijt} \quad (5)$$

²⁸French public schools recruit in a pre-defined zone so belonging to the same school is equivalent to living in the same neighbourhood.

²⁹In this model, we do not include gender since gender is orthogonal to family SES and immigration status so its contribution to academic skills and aspirations do not bias our estimates of interest. Besides, we do not include dummies indicating whether the student skipped or repeated a grade since, as mentioned in Section 4, the probability that a student repeats or skips a grade depends not only on academic skills but also on her family characteristics independently of academic skills. With these dummies included, the results are qualitatively similar and quantitatively very close (available upon request). As we are interested in social differences in aspirations, we prefer to exclude these variables which encapsulate social differences.

We interpret the coefficient on low-SES in (4) as the difference in attainable tracks between low and high-SES students *due to* perception of one’s academic potential (we compare students who in the same class, have similar academic capacity, similar immigration background, and importantly similar salient tracks). In (5), we interpret the coefficient on low-SES as the difference in preferred tracks between low and high-SES students *not* due to knowledge of academic tracks and self-perception of academic potential (these students are in the same class, have similar academic capacity, similar immigration background, and importantly similar attainable tracks). Reasons for this remaining gap may be different liquidity constraints facing the cost of education, or different expected returns to education, or identity concerns. Comparison of the coefficient on low-SES between equation (5) and equation (3) gives the degree to which differences in preferred tracks are the consequence of self-perceived academic potential.

Finally, to explore further the role of family background in affecting perceptions and cognition, we estimate equation (3) using measures of fatalism (gaps in assessed probabilities of success of a hypothetical high-achieving student living in different social situations), and scholastic self-esteem, as dependent variables.

5.2 Social differences in aspirations

5.2.1 Aspirations for High School

Table 5 presents the differences in aspirations at the high school level: vocational or academic high school. For simplicity, we do not report and comment on results for the remaining outcome “no high school track”. Column 1 shows that low-SES students are as likely as high-SES to bring up vocational high school in their salient tracks (84% among low-SES versus 81% among high-SES), but much more likely to refer to vocational high school as attainable and preferred tracks: 29% of low-SES refer to vocational high school as attainable track, and 16% as preferred track, compared to respectively 15% and 4.5% of high-SES. On the contrary, column 5 shows that low-SES students are a bit less likely to bring up academic high school in their salient tracks (90% versus 93% among high-SES), and much less likely to refer to academic high school as attainable and preferred tracks: 72% of low-SES refer to academic high school as attainable track, and 61% as preferred tracks, compared to respectively 89% and 80% of high-SES.

A big part of these differences can be attributed to differences in academic capacity and school environment (columns 2-3 and 6-7). Among students who show similar capacity and belong to the same class, the gaps between low and high-SES are generally smaller. Regarding salient tracks, low-SES are slightly more likely to mention both vocational and academic high school in their salient tracks than equally-performant high-SES classmates (+ 7.4% for vocational high school and + 2.7% for academic high school). This indicates that high-SES have a deficit in information at the beginning of grade 9, especially regarding vocational high school. For attainable tracks, low-SES have a 45% higher probability to cite vocational high school, and a

4% lower probability to cite academic high school, than equally-performant high-SES classmates. Finally regarding preferences, low-SES have a 120% higher probability to prefer vocational high school, and a 5% lower probability to prefer academic high school, than equally-performant high-SES classmates. These results demonstrate that low-SES students have more modest aspirations than high-SES despite similar academic capacity and school environment. In spite of slightly broader knowledge of existing tracks, low-SES students' attainable and preferred tracks are biased in favor of vocational high school and against academic high school relative to high-SES. In contrast, it should be noticed that students from immigrant family have similar salient options, and a lower probability to refer to vocational high school as attainable, and similar preferred tracks, than non-immigrant students.

Once salient tracks are controlled for, low-SES students are still 41% more likely to refer to vocational high school as attainable relative to equally-performant high-SES classmates (column 4, panel 2). The fact that vocational high school is more salient to low-SES than high-SES students does not explain the larger proportion of low-SES who refer to vocational high school as attainable. The difference in self-perceived academic potential is thus the predominant explanation. High-SES students may feel not capable of doing vocational high school, but the most sensible explanation is that some students just answer their highest attainable track (i.e. academic high school) instead of all of them, meaning that students who feel capable of doing both vocational and academic high school but just cite academic high school as an attainable track. Low-SES students are thus more likely to feel capable of doing *only* vocational high school than equally-achieving high-SES classmates.

Once attainable tracks are controlled for, low-SES students are 56% more likely to prefer vocational high school than equally-performant high-SES classmates (column 4, panel 1), compared to 120% when attainable tracks are not controlled for. This result indicates that about half of the difference in aspirations for vocational high school among equally-achieving classmates is due to differences in self-perceived academic potential. The other half is to attribute to differences in preferences, which may relate to differential liquidity constraints vis-à-vis the cost of education, or differential returns to education, or identity concerns. Besides, immigrant students have the same probability to prefer vocational track than equally-performant non-immigrant classmates once attainable tracks are controlled for. Overall, our results show that aspirations for high school are different among low and high-SES students, due not only to differences in academic capacity, school environment, and preferences (which all play an important role), but also largely to differences in self-perception of academic potential. Subgroup analysis by terciles of academic test score shows that low-SES students in the bottom and medium terciles are the most affected.

5.2.2 Aspirations for Higher Education

Table 6 presents the differences in aspirations at the higher education level. For simplicity, we focus on the two extreme aspirations (“finding a job right after high school” and “masters”) and do not report results for the remaining intermediate outcome “1-4 years college” (results available upon request). We investigate whether the preference for finding a job is sourced in an absence of salient and attainable tracks in higher education, and whether the preference for masters is sourced in masters being salient and attainable.

Column 1 shows that low-SES students are much more likely to cite no salient track (41% versus 28% among high-SES), as well as no attainable track in higher education (65% versus 51% among high-SES). Also, low-SES are much more likely to prefer finding a job right after high school (22% versus 8% among high-SES). On the contrary, column 5 shows that low-SES students are less likely to bring up masters in their salient tracks (20% versus 38% among high-SES), in their attainable tracks (11% versus 26% among high-SES), and in their preferred track (9% versus 24% among high-SES). It is also to notice that immigrant students are remarkably similar to non-immigrant students with respect to salient, attainable and preferred tracks in higher education.

A major part of the differences between low and high-SES can be attributed to differences in academic capacity and school environment, but the gaps remain substantial even among equally-achieving classmates (columns 2-3 and 6-7). Regarding salient tracks in higher education, low-SES students have a 11% higher probability to cite no track (although this difference is not statistically significant), and a 15% smaller probability to bring up masters (significant at the 10% level) than equally-performant high-SES classmates (columns 3 and 7, panel 3). Regarding attainable tracks, low-SES have about the same probability to cite no attainable tracks but still a 28% lower probability to cite masters as an attainable track than equally-performant high-SES classmates (columns 3 and 7, panel 2). Regarding preferences, low-SES are 69% more likely to prefer finding a job right after high school, and a 26% less likely to prefer masters, than equally-performant high-SES classmates (columns 3 and 7, panel 1). Equally-achieving classmates thus have different salient, attainable, and preferred tracks with respect to higher education depending on parent SES, low-SES students’ being biased in favor of finding a job (despite small differences in reporting no salient or attainable option in higher education) and against masters relative to high-SES. In contrast, immigrant students preferences are biased against finding a job relative to equally-achieving non-immigrant classmates (columns 2-4 panel 1).

The larger preference for finding a job right after high school among the low-SES is not much driven by differences in salient and attainable tracks: once attainable tracks are controlled for, low-SES students are still 59% more likely to prefer finding a job than equally-performant high-SES classmates (column 4, panel 1),

instead of 69% when attainable tracks are not controlled for (column 3). In contrast, the smaller preference for masters among the low-SES is entirely driven by masters being less salient and less attainable. Differences in masters as salient account for half of the differences in masters as attainable (columns 7-8, panel 2), and differences in masters as attainable account for the totality of differences in preference for masters (columns 7-8, panel 1). Social differences in aspirations for masters are thus not only due to social differences in academic capacity and school environment, but also to differences in information and self-perceived academic potential. Subgroup analysis by terciles of academic test score shows that low-SES students in the top tercile are the most concerned by these deficits in information and self-perceived academic potential regarding masters.

5.2.3 Professional Aspirations

The first thing to notice about professional aspirations is that, for all students, academic aspirations are not even close to be consistent with professional aspirations. For instance, among high-SES students, 23.9% aspire to masters as an educational track, while 44.5% aspire to a job that requires masters. And it is also true for low-SES students. This suggests that teenagers do not see education as a pure investment for a future job, probably due to ignorance of educational pathways to jobs. At this age, educational aspirations are not entirely driven by professional aspirations. This lack of realism is important as it may lead students to make irreversible decisions in the short-term like entering vocational high school while you aspire to a job that requires at least a masters (for instance, vocational high school does not give any chance at the faculty of medicine; only 5% who complete vocational high school go to university). Since educational choices have consequences on later occupations, professional aspirations will have to adjust to educational aspirations *ex post*. The inconsistencies between educational and professional aspirations suggest that these consequences are not fully anticipated and internalized by teenagers, which will create frustration when students will realize that their educational choices do not match their professional aspirations. And because low-SES students achieve less well at school on average, they will have a higher probability to be constrained by this kind of irreversible choice.

On top of that, Table 7 then shows that social differences in professional aspirations are neither aligned with the social differences in academic aspirations that we find above. Columns 1, 4, 7 and 10 first show raw differences in professional aspirations. Low-SES and high-SES are as likely to have no professional aspiration (22%), and as likely to aspire to a job that requires 1-4 years of higher education (33%). However, low-SES students are much more likely to aspire to a job that does not require any higher education (45% versus 35% among high-SES), and symmetrically less likely to aspire to a job that requires a masters (35% versus 45% among high-SES). However, as soon as academic performance and class fixed-effects are controlled for (columns 3, 6, 9, and 12), the social gap related to the preference for a masters-level job disappears, and the

gap related to the preference for a job that does not require higher education falls by half. Social differences in professional aspirations are thus mainly due to differences in academic performances. And this remaining gap in professional aspirations is not aligned with the gap in academic aspirations that we find above. For instance, when low-SES students are as likely to prefer a job that requires a masters than their equally-achieving high-SES classmates, they are also 26% less likely to prefer masters as an educational track. And when they are only 11% more likely to prefer a job that does not require any higher education, they are 69% more likely to say that they prefer finding a job right after high school.

This is not true for immigrants and non-immigrants: our results show that immigrants are more likely to aspire to occupations that require a masters than equally-achieving non-immigrant classmates, consistent with Caille (2007) finding that immigrants exhibit an excess of ambition relative to non-immigrants, which is plausibly inherent to the decision to migrate. This result reveals that teenagers' lack of realism is even more pronounced among immigrants: 16% aspire to masters while 47.6% aspire to a job that requires masters, when among non-immigrants, 15.6% aspire to masters while 38.2% aspire to a job that requires masters. As long as students do not make irreversible decisions in the short-term like entering vocational high school, aspirations for higher education may adjust to professional aspirations later when students will receive information on higher education in high school. In this respect, our results suggest that we don't have to worry too much since immigrant students have 1) the same probability to aspire to academic high school (Table 5 panel 1 column 7), and 2) the same probability to enter in academic high school given their aspiration (Table 3 column 8) than equally-achieving non-immigrant classmates.

These findings are important to interpret the differences in educational aspirations among the low and high-SES: social groups differ much more in the way they plan to invest in education than in the way they think about their future occupation. It maybe that low-SES students plan to invest less in education than equally-achieving high-SES classmates because they aspire to lower-qualified jobs, for instance because they anticipate discrimination on the job market, or think they lack appropriate social networks to have a high-qualified occupation. But given the type of jobs they are aspiring to, their educational aspirations are too low. This lack of educational aspirations also exists for high-SES students, but it is stronger for low-SES students. This social gap may evolve later on when students will become more aware and sensitive to questions related to the job market. But at 15 years-old, social differences in educational aspirations are not entirely driven by social differences in professional aspirations.

5.2.4 Socially-Dependent Perceptions

Our previous results show that low-SES and high-SES students differ in their educational aspirations not only because they have different present academic capacity, school environment and preferences towards

education, but also because of different knowledge of academic tracks and different self-perceived academic potential. Table 8 provides additional evidence on the ways family SES influences self-perceived academic potential.

Columns 1, 2 and 3 show that students have fatalistic views on the extent to which social origin affects later success in high school. On average in our sample, students believe that the probability of success in education for a hypothetical high-achieving student living in an advantaged neighbourhood is 85.5%. Low- and high-SES students who perform similarly and belong to the same class all perceive that this probability would fall down by 36.6 points if this hypothetical high-achieving student was living in a disadvantaged neighbourhood (column 1), by 27.3 if s/he had immigrant parents (column 2), and by 8.4 points if s/he had a family member who succeeded academically and professionally (column 3). Students thus perceive social background as having a huge influence on later success in high school for presently high-achieving students, an influence that is excessive in view of the existing evidence: Broccolichi and Sinthon (2011) show that the probability to complete academic high school is 96% for the top-quintile (at entry in high school) high-SES students, while 91% for the top-quintile low-SES students (for the 4th quintile, these probabilities are respectively 90% and 82%). While parent SES does have an influence on the probability to succeed in high school for presently equally-achieving students, this influence is evidently considerably smaller than what is anticipated by the students in our data. Interestingly, these fatalistic views are perfectly equally distributed among all students whatever their own social origin. The only small difference is between students from immigrant and non-immigrant families: students from immigrant families are less pessimistic about the negative impact of immigrant families on later success in high school than non-immigrants (column 4). Overall, our data show that social fatalism regarding education is excessive and widespread among French students, which may explain the observed differences in self-perceived academic potential between low and high-SES who perform equally and are in the same class. The issue is that fatalistic anticipations are self-fulfilling: too fatalistic anticipations of future academic success may cause low aspirations, which lead to lower school outcomes than what she would have had with higher aspirations as shown in Section 4. The lack of accuracy of self-perceived academic potential can thus be the source of an aspiration failure.

Moreover, column 4 shows that low-SES students have a 0.131 standard deviation lower scholastic self-esteem than equally-achieving high-SES classmates. The scholastic self-esteem is based on perceptions of one's present academic capacity, so there is nothing here related to anticipations for the future. It means that students who have similar *objective* present academic capacities have different *subjective* present academic capacities. Interestingly, this result is true for students of all school proficiency levels, although more pronounced for students in the top tercile for which the gap in scholastic self-esteem is 0.16 standard de-

viations relative to equally-achieving high-SES classmates. For students in the top tercile, we find that the deficit in aspirations due to family SES concerns masters, partly due to ignorance, partly due to self-perception of academic potential. By introducing our measure of scholastic self-esteem in equation (4), we find that scholastic self-esteem explains 25% of the deficit in low-SES students in the top tercile mentioning masters in their attainable tracks³⁰.

These findings provides evidence that social identity creates mental models affecting how an individual experiences what he experiences, as proposed in Karla Hoff and co-authors' models: beliefs related to social inferiority affect self-confidence, self-perceived probability of success, and behavior (aspirations) in ways that make the beliefs come true (Hoff and Pandey 2006, 2012; Hoff and Fehr 2011; Hoff and Stiglitz 2016). Our empirical results support this view: scholastic self-esteem is prone to stereotype susceptibility related to parent SES, which results in different sets of attainable tracks, which result in different preferred tracks, and eventually affects school outcomes.

6 Are Aspirations Optimal?

Aspirations are optimal if they lead to maximum welfare. Our findings provide three reasons why aspirations may not lead to maximum welfare.

First, some differences in aspirations between equally-achieving low and the high-SES classmates are due to differences in knowledge and information, in particular regarding masters among students in the top-tercile. Lack of information is a simple reason why preferences may not be rational and lead to maximum welfare.

Second, social differences in aspirations for vocational high school, academic high school, and masters, among equally-achieving classmates are mostly due to differences in self-perceived attainable tracks. Low-SES are partially right to anticipate lower academic capacity in the future relative to high-SES students that perform equally in the present. Their social background is a clear factual disadvantage, as shown by the negative coefficient of low-SES family on short-term academic progress in Tables 1, 2 and 4. However, they are likely to underestimate their future academic potential since they do underestimate their present academic capacity and overestimate the way social background influences future probability of success (Table 8). It should not be taken for granted that students assess their academic potential at its true value. Social stereotypes and fatalism bias students' perceptions and cognition, which casts doubt on the accuracy of their self-perceived academic potential.

Finally, low aspirations are a source of disadvantage in their own right. In section 4, we have shown

³⁰More specifically, we add deciles of scholastic self-esteem and deciles of scholastic self-esteem interacted with SES in model (4) on the subgroup of students in the top tercile with masters as attainable as the dependent variable. In model (4), the coefficient on low-SES is -0.071, significant at the 5% level. In the augmented version, the coefficient on low-SES is -0.054, significant at the 5% level, which represents a 25% reduction in the coefficient on low-SES.

that low aspirations lead to lower school performances in the short-term. There is an negative impact of low-SES family on short-term outcomes, but there is also an additional negative impact of low aspirations. Two students in similar situations in terms of present academic capacity, school environment and family environment (same *objective* probability of success) get different later outcomes depending on their aspiration: the one who aspires high gets higher school outcomes than the one who aspires low. And this result holds for both high and low-SES students (Appendix Table A4). If school attainment has a positive return on the labor market, this mechanism can thus lead to suboptimal job market outcomes. Actually, the vast literature on the returns to education shows that returns are substantial - about 10% higher wages per additional year of higher education (see Oreopoulos and Petronijevic 2013 for a review), with no evidence that returns are different for low and high-SES students. Moreover, the cost of education in France is particularly low: 79% of students enrolled in academic high schools attend public (tuition-free) high schools³¹, and most higher education institutions are entirely free for low-SES students (including the most selective institutions). Reduced school attainment is thus likely to result in suboptimal job market outcomes.

There may be some situations where low aspirations are not detrimental in terms of long-term welfare. First, if higher school performances in the short-term do not lead to better school attainment. This may be true for very weak students who would never succeed in academic high school. Goux et al. (2016) shows that the weakest students should better enter vocational high school instead of trying academic high school in order to avoid dropping-out. However, while the result may well apply to a subsample of students in our bottom tercile³², medium and high-achievers should be able to complete academic high school and some higher education, so reduced school outcomes in grade 9 due to low aspirations should lead eventually to lower school attainment and job market outcomes for most students. Second, lower school outcomes and job market outcomes may be compatible with higher welfare because of identity concerns and social preferences. Individuals do not value only economic outcomes but also social outcomes. Since social interactions are frequently a priority for many teenagers (Coleman 1961), students may forgo worthwhile educational and professional opportunities because of social benefits. The identity literature provides several explanations like keeping close and conform to friends, affirming one's social identity to maintain a sense of unity, or fighting the threat of losing one's culture (Akerlof and Kranton 2002, Fryer 2006, Fang and Loury 2005, Benabou and Tirole 2011, Carvalho and Koyama 2013). However, because low-SES students aspire to jobs that require a number of years of education that is much higher than what they plan to invest (Tables 5, 6 and 7), and because social groups differ much more in the way they plan to invest in education than in the way they think about their future occupation, we can doubt that lower school attainment and job market outcomes

³¹[http://www.education.gouv.fr/cid57111/l-education-nationale-en-chiffres.html#Le second degré](http://www.education.gouv.fr/cid57111/l-education-nationale-en-chiffres.html#Le%20second%20degr%C3%A9)

³²Students in Goux et al. (2016) have academic performances that are slightly worse than those of students of the bottom tercile in this paper.

lead to maximum welfare.

7 Conclusion

This paper provides evidence that students' educational aspirations are influenced by their family socioeconomic status on the one hand, and that these aspirations contribute to the short-term evolution of school outcomes on the other hand. As school outcomes are themselves a determinant of aspirations, our results reveal the existence of an aspiration-based inequality trap which concerns low-SES students of all academic proficiency levels. Low-SES students start with clear factual disadvantages, but this aspiration trap drags them down even more. By contrast, being in an immigrant family rather boosts educational and professional aspirations.

A natural question is whether it reflects a market failure that would rationalize some form of policy intervention. Do students have suboptimal aspirations leading to suboptimal educational, job market and welfare outcomes? This paper provides evidence of three reasons why aspirations may not be optimal, especially those of the low-SES students: first, it cannot be taken as granted that students assess their academic potential at its true value, as suggested by the fact that the low-SES students underestimate their present academic capacity relative to their equally-achieving high-SES classmates, as well as by the fact that their views appear excessively fatalistic. Social influence on perceptions and cognition may thus affect the perception of attainable options in ways that perpetuate social inequalities. Second, we show that ignorance of some academic tracks, especially masters, plays a role in shaping socially differential aspirations. Finally, the very fact that one aspires low creates a disadvantage in its own right. Low-SES (and high-SES) students would benefit from higher aspirations in terms of short-term school outcomes, which would likely result in higher school attainment and job outcomes in the longer term except may be for the very weak students. This result suggests that most low-SES students would benefit (in terms of educational and professional outcomes) from either not realizing the influence their social background imposes on their future performances, or on the contrary realizing the influence both their social background *and* their aspirations impose on their future performances. In that sense, partial awareness (being aware of the impact of social background *but not* of the impact of aspirations) may be worse than full awareness or no awareness at all.

The question is more difficult when it comes to welfare. Whether higher aspirations, school attainment and job outcomes would make low-SES students happier remains an open question. They may feel socially isolated or at odds with their cultural values, as suggested by the identity literature. However, our results show that low-SES students have professional aspirations that are much higher in terms of required diploma than what they plan to invest in education, which is not quite consistent with the view that increased

educational and professional levels would hurt their welfare.

This paper thus questions whether preferences can be wrong, an issue that should be discussed. Most of the economic literature is based on the latin maxim that *de gustibus non est disputandum*³³, so that everyone's personal preferences are merely subjective opinions that cannot be right or wrong. If preferences are formed on a clear-sighted and informed basis, this may be true. But if preferences are formed on the ground of misperceptions and lack of information, preferences may be the root of a market failure. This paper should thus encourage further research on the long-term consequences of educational aspirations, in particular their consequences on welfare, and motivate the design of appropriate interventions to help disadvantaged people aspire at their true potential to increase upward mobility.

³³Meaning "In matters of taste, there can be no disputes".

References

- [1] Aeberhardt, R., R. Rathelot and M. Safi. 2015. “Les difficultés scolaires et professionnelles des jeunes issus de l’immigration : effet de l’origine ou effets géographiques ?”, *Population* (french edition), Institut National d’Études Démographiques (INED), vol. 70(3): 599-635.
- [2] Afsa, Cédric. 2009. “La moitié d’une génération accède à l’enseignement supérieur”, in *La France, portrait social* - édition 2009.
- [3] Akerlof, George and Rachel and Kranton. 2000. “Economics and Identity”, the *Quarterly Journal of Economics*, 65(3): 715-753.
- [4] Akerlof, George and Rachel Kranton. 2002. “Identity and Schooling: Some Lessons for the Economics of Education”, *Journal of Economic Literature*, 40(4): 1167-1201.
- [5] Appadurai, Arjun. 2004. “The Capacity to Aspire: Culture and the Terms of Recognition”, in *Culture and Public Action*, Rao Vijayendra and Michael Walton, Stanford University Press.
- [6] Austen-Smith and Roland Fryer. 2005. “An Economic Analysis of “Acting White””, the *Quarterly Journal of Economics*, 120(2): 551-583.
- [7] Avvisati, F., M.Gurgand, N. Guyon and E. Maurin. 2014. “Getting Parents Involved: a Field Experiment in Deprived Schools”, *The Review of Economic Studies*, ISSN 0034-6527, 01/2014, Vol. 81 (1), p. 57.
- [8] Bariaud, Françoise. 2006. “Le Self-perception profile for adolescents (SPPA) de S. Harter”, *L’orientation scolaire et professionnelle*, 35/2: 282-295.
- [9] Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. “Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India”, *Science*, 335: 582-586.
- [10] Bénabou, Roland and Jean Tirole. 2011. “Identity, Morals and Taboos: Beliefs as Assets”, *The Quarterly Journal of Economics*, 126 (2): 805-855.
- [11] Bernard, Tanguy, Stefan Dercon, and Alemayehu Seyoum Taffesse. 2011. “Beyond Fatalism: An Empirical Exploration of Self-Efficacy and Aspirations Failure in Ethiopia”, IFPRI Discussion Paper 01101

- [12] Bernard, Tanguy, Stefan Dercon, Kate Orbin, and Laemayehu Seyoum Taffesse. 2013. "Learning with Others: A Field Experiment on the Formation of Aspirations in Rural Ethiopia", mimeo
- [13] Björklund, Anders; Salvanes, Kjell G. 2010. "Education and family background: Mechanisms and policies", Discussion paper series // Forschungsinstitut zur Zukunft der Arbeit, No. 5002
- [14] Black, S. E., P J. Devereux and K. G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital", American Economic Review, Vol. 95(1).
- [15] Broccolichi Sylvain and Rémi Sinthon. 2011. "Comment s'articulent les inégalités d'acquisition scolaire et d'orientation ? Relations ignorées et rectifications tardives", Revue française de pédagogie, 175: 15-38.
- [16] Bronfenbrenner, Urie. 1979. "The Ecology of Human Development: Experiments by Nature and Design", Cambridge, MA: Harvard University Press. ISBN 0-674-22457-4.
- [17] Caille, Jean-Paul. 2005. "Le vécu des phases d'orientation en fin de troisième et de seconde", Éducation & formations, 72: 77-99.
- [18] Caille, Jean-Paul and Fabienne Rosenwald. 2006. "Les inégalités de réussite à l'école élémentaire : construction et évolution", in "France : portrait social", INSEE: 115-137.
- [19] Caille, Jean-Paul. 2007. "Perception du système éducatif et projets d'avenir des enfants d'immigrés", Éducation & formations, 74.
- [20] Carvalho, Jean-Paul and Mark Koyama. 2013. "Resisting Education," MPRA Paper 48048, University Library of Munich, Germany.
- [21] Cattell, R. B. (1971). Abilities: Their structure, growth, and action. New York: Houghton Mifflin.
- [22] Chetty, Raj, John Friedman, Nathaniel Hilger, Diane Whitmore Schanzenbach and Danny Yagan. 2011. "How does your kindergarten affect your earnings? Evidence from Project STAR", The Quarterly Journal of Economics, 126(4): 1593-1660.
- [23] Chetty, Raj, Nathaniel Hendren and Lawrence Katz. 2016. "The effects of exposure to better neighborhoods on children: Evidence from the Moving to Opportunity Experiment", American Economic Review, 106(4): 855-902.

- [24] Coleman, James. 1961. *The Adolescent Society*. New York: The Free Press.
- [25] Croizet, Jean-Claude, Michel Désert, Marion Dutrévis and Jacques-Philippe Leyens. 2001. "Stereotype threat, social class, gender, and academic under-achievement: when our reputation catches up to us and takes over", *Social Psychology of Education*, 4: 295-310.
- [26] Croizet, Jean-Claude, Gérard Després, Marie-Eve Gauzins, Pascal Huguet, Jacques-Philippe Leyens, Alain Méot. 2004. "Stereotype Threat Undermines Intellectual Performance by Triggering a Disruptive Mental Load", *Personality and Social Psychology Bulletin*, 30: 721-732.
- [27] Dalton, Patricio, Sayantan Ghosal and Anandi Mani. 2016. "Poverty and Aspirations Failures", *The Economic Journal*, *The Economic Journal*, 126(590): .
- [28] Dee, Thomas. 2014. "Stereotype threat and the student-athlete", *Economic Inquiry*, 52(1): .
- [29] Duckworth, Angela Lee and Martin E. P. Seligman. 2006. "Self-Discipline Gives Girls the Edge: Gender in Self-Discipline, Grades, and Achievement Test Scores", *Journal of Educational Psychology*, 98(1): 198 –208.
- [30] Duru-Bellat, Marie. 1988. *Le fonctionnement de l'orientation. Genèse des inégalités sociales à l'école*. Lausanne, Editions Delachaux et Niestlé.
- [31] Fang, Hanming, and Glenn Loury. 2005. "'Dysfunctional Identities' Can Be Rational." *American Economic Review*, 95(2): 104-111.
- [32] Fehr, Ernst and Karla Hoff. 2011. "Tastes, castes, and culture: The influence of society on preferences," *ECON - Working Papers 026*, Department of Economics - University of Zurich.
- [33] Fletcher, J., B. L. Wolfe. 2016. "The Importance of Family Income in the Formation and Evolution of Non-Cognitive Skills in Childhood", *NBER Working Paper No. 22168*.
- [34] Fryer, Roland. 2006. "A Model of Social Interactions and Endogenous Poverty Traps," *NBER Working Papers 12364*, National Bureau of Economic Research.
- [35] Jenck, Crouse and Mueser. 1983. "The Wisconsin Model of Status Attainment: A National Replication with Improved Measures of Ability and Aspirations," *Sociology of Education* (January 1983): pp. 3-19.
- [36] Jensen, A.R. (1998). *The g factor: The science of mental ability*. Praeger.
- [37] Génicot, Garance and Debraj Ray. 2015. "Aspirations and Inequality", mimeo.

- [38] Goux, Dominique and Eric Maurin. 2007. "Close Neighborhoods Matter: Neighborhoods Effects on Early Performances at School", *Labour Economics*,
- [39] Goux, Dominique, Marc Gurgand, and Eric Maurin. 2016. "Adjusting your dreams? The Effect of School and Peers on Dropout Behaviour", forthcoming *Economic Journal*
- [40] Guyon, Nina and Elise Huillery. 2015. "Explaining Aspiration Failures: Evidence from French Middle School Students", mimeo, Sciences Po.
- [41] Hanna, Rema and Leigh Linden. 2012. "Discrimination in Grading", *The American Economic Journal: Economic Policy*, 4(4): 146–168.
- [42] Harter, Susan. 1988. "The Self-Perception Profile for Adolescents", Unpublished manual. University of Denver, Denver, CO.
- [43] Heckman, J., Stixrud, N. and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior", *Journal of Labor Economics* 24 (3), 411-48.
- [44] Hertz Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith and Alina Veraschagina. 2007. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends", *The B.E. Journal of Economic Analysis and Policy (Advances)*, 7(2), Article 10.
- [45] Hoff, Karla and Pandey Priyanka. 2006. "Discrimination, Social Identity, and Durable Inequalities", *American Economic Review*, 96(2): 206-211.
- [46] Hoff, Karla and Pandey Priyanka. 2012. "Making up people - the effect of identity on preferences and performance in a modernizing society," *Policy Research Working Paper Series 6223*, The World Bank.
- [47] Hoff, Karla and Joseph Stiglitz. 2010. "Equilibrium Fictions: A Cognitive Approach to Societal Rigidity," *American Economic Review: Papers and Proceedings*, 100(2): 141-146.
- [48] Hoff, Karla and Joseph Stiglitz. 2016. "Striving for Balance in Economics: Towards a Theory of the Social Determination of Behavior", *Journal of Economic Behavior and Organization*, forthcoming.
- [49] Hoxby, Caroline and Christopher Avery. 2013. "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students", *Brookings Papers On Economic Activity*, Spring 2013.

- [50] Le Rhun, Béatrice. “Le niveau d’études selon le milieu social”, in : État de l’Enseignement supérieur et de la Recherche en France - 49 indicateurs, Kabla-Langlois Isabelle (dir.). Paris : Ministère de l’Éducation nationale, de l’Enseignement supérieur et de la Recherche, 2015 (8e éd.), fiche 20, p. 50-51.
- [51] Merle, Pierre. 1998. “Les recherches sur la notation des élèves”, *Education & Formations*, 53: 7-19.
- [52] Messick, Samuel (1989). "Meaning and Values in Test Validation: The Science and Ethics of Assessment". *Educational Researcher*. 18 (2): 5–11.
- [53] Oyserman, Bybee, Terry. 2006. “Possible Selves and Academic Outcomes: How and When Possible Selves Impel Action”, *Journal of Personality and Social Psychology*, 91(1): 188-204
- [54] Oreopoulos, P., M. Page and A. Stevens. 2006. “Does Human Capital Transfer from Parent to Child? The Intergenerational Effects of Compulsory Schooling”, *Journal of Labor Economics*, 24(4): .
- [55] Pirus, C. 2013. “Le déroulement de la procédure d’orientation en fin de troisième reste marqué par de fortes disparités scolaires et sociales ”, Note d’information de la DEPP, 13.24
- [56] Poropat, Arthur. 2009. “A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance”, *Psychological Bulletin*, 135(2): 322–338.
- [57] Program for International Student Assessment. 2012. “Results: Excellence through Equity: Giving Every Student the Chance to Succeed”, Volume II, OECD.
- [58] Ray, Debraj. 2006. “Aspirations, Poverty and Economic Change”, in A. Banerjee, R. Bénabou and D. Mookherjee (eds), *What Have We Learnt About Poverty*, Oxford University Press.
- [59] Raven, J.; Raven, J. C.; Court, J. H. (2003) [1998]. "Section 1: General Overview". Manual for Raven’s Progressive Matrices and Vocabulary Scales. San Antonio, TX: Harcourt Assessment.
- [60] Sacerdote, Bruce. 2007. “How large are the effects from changes in family environment? A study of Korean American adoptees”, *The Quarterly Journal of Economics*, 122 (1): 119-157.
- [61] Sewell, William, Archibald Haller, and Alejandro Portes. 1969. “The Educational and Early Occupational Attainment Process”, *American Sociological Review*, 34(1): 82-92.

- [62] Sewell, William, Archibald Haller, and George Ohlendorf. 1970. "The educational and early occupational status attainment process: replication and revision", *American Sociological Review*, 35: 1014-1027.
- [63] Steele, Claude and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African Americans", *Journal of Personality and Social Psychology*, 69(5): 797-811.

Data Appendix

Construction of the Educational Aspirations

Given the open nature of these questions, the actual denomination of tracks varies a lot. At the high school level, a majority of answers are very precise (e.g., “scientific *baccalauréat*” or “2-year vocational track in hairstyling”) or quite precise (e.g., “academic high school”, “vocational *baccalauréat*”), while a minority (5%) of answers are vague (e.g., “high school”, “music”). We coded the answers in order to classify them into four categories: academic high school, vocational high school, no high school, and no response. When the answer is vague and there is uncertainty about the corresponding category, we consider two extreme scenarios: for instance, “music” may be associated with, at least, no high school education or, at most, academic high school (the literature track offers a music section). Since the vague responses represent only 5% of responses, there is a very high correlation (0.93) between the overall results when these 5% of responses are classified according to the “pessimistic” scenario (in which the inferred track is the less selective) and the overall results when the responses are classified according to the “optimistic” scenario (in which the inferred track is the most selective). We present the results using the pessimistic scenario, but they are identical in the optimistic one. Then, we aggregate answers at the student level to create dummies indicating whether the student’s response is among the following categories: “No response”, “Vocational high school is among attainable tracks”, and “Academic high school is among attainable tracks”. Students who reported several answers can be in both the academic and the vocational categories. Students who wrote “I don’t know”, “None”, *only* answers that are not relevant like “Traveling”, or who did not write anything, constitute the “No response” category. Finally, 20 students reported only one vague response that is associated with no high school education in the pessimistic scenario (e.g., students whose *unique* response is “music”). These few students were grouped with the “No response” category as we consider that their answer does not inform us about which track they feel capable of pursuing.

Data construction is similar at the higher education level. Students’ answers are coded according to the implied number of years of education: “school of architecture” is coded as 5 years, “I.U.T.” is coded as 2 years, etc. A handful of answers are vague and allow for different implied levels of education, like “university”. In that case, as before, we use the lowest number of years of education compatible with the answer. In the example of “university”, the shortest degree requires 3 years, so the answer is coded as 3 years. We create dummies indicating whether the student falls into the following categories: “No response”, “No higher education is among attainable paths”, “1-2 years college is among attainable paths”, “3-4 years college is among attainable paths”, and “5 years college or more is among attainable paths”. We do not use the dummy indicating that “No higher education is among attainable paths” because the answers that fall into this category are often

imprecise and make this category too heterogeneous: a third of the answers in this category are vague like “artist”, “singer”, “pianist”, etc. These answers allow for different implied levels of education including “No higher education” but also quite high levels of education (e.g., conservatoire, “Ecole des Beaux-Arts”, etc.). For this reason, we think that this category is too heterogeneous to draw clear conclusions.

Professional Aspirations

We used the website of ONISEP, the principal French institute providing information on academic paths and jobs to associate the number of years of education required for each job. When jobs are not precise and can be associated with different levels of education (e.g., “IT engineer” which can be associated with at least a 2-year college education and at most a 5-year college education), as is the case with roughly 29% of answers, we build two extreme scenarios and use the lowest number of years of education compatible with the answer given. We check that our results are robust to the use of the optimistic scenario and indicate whenever a result is not robust.

Academic Capacity

The total test score in November 2012 is constructed using the two November math test scores on the one hand, and the relationship between the total score in June 2013, the score in math in June 2013, and invariant students’ characteristics (gender, SES, year of birth, and classroom fixed effect) on the other hand. 83% of the variation in students’ total scores on the national exam is explained by the variation in math scores. Adding gender, SES, year of birth, and classroom fixed effects raises explanatory power to 87%. We estimate the coefficients of a regression of the total score on the math score and students characteristics in June, and use these coefficients to estimate the total score that a student would have had in November given her average math score and her characteristics.

Family Socio-Economic Status

The two-category classification separates “high-SES” from “low-SES” based on whether at least one guardian of the student (parent) has an occupation that corresponds to five years or more of education. The list of these occupations is: legal, medical, teaching, and artistic freelance occupations; high-level civil servants; professors; researchers; journalists; artists; senior executives; engineers. They account for 25% of guardian 1 and 20% of guardian 2.

The six-category classification divides the low-SES families into 5 groups to get more homogeneous social groups: “No parent has ever worked”, “Maximum family SES is manual laborer”, “Maximum family SES is low-skilled white-collar”, “Maximum family SES is craftsman or storekeeper”, and “Maximum family SES is

intermediate occupation”. The social hierarchy used to define these categories relies on the increasing average level of education throughout the job categories “manual laborer”, “low-skilled white-collar”, “craftsman and storekeeper”, and “intermediate occupation”. The average level of education by job category is computed using our research survey, which contains information about both parent occupations and levels of education.

Figure 1: Test scores in June 2013 by track assignment in September 2013

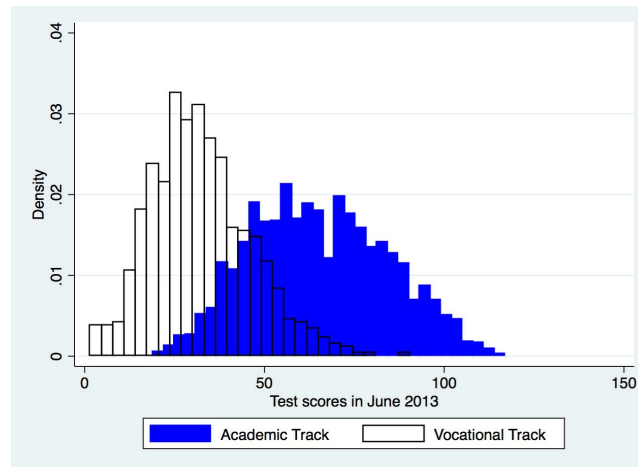


Figure 2: Proportion of low-SES families at the school level: our sample versus junior high school population

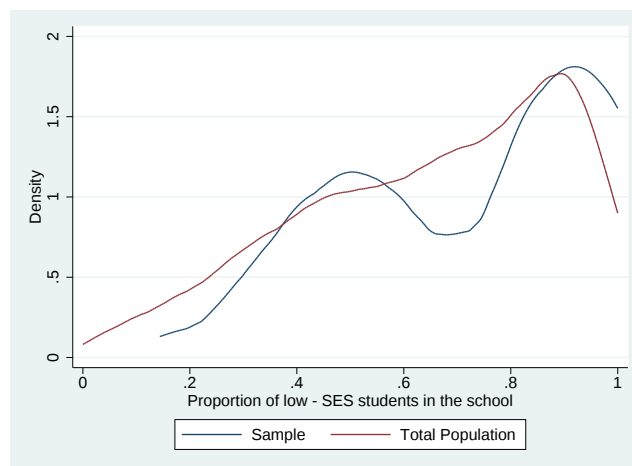


Figure 1: Test scores in November 2012

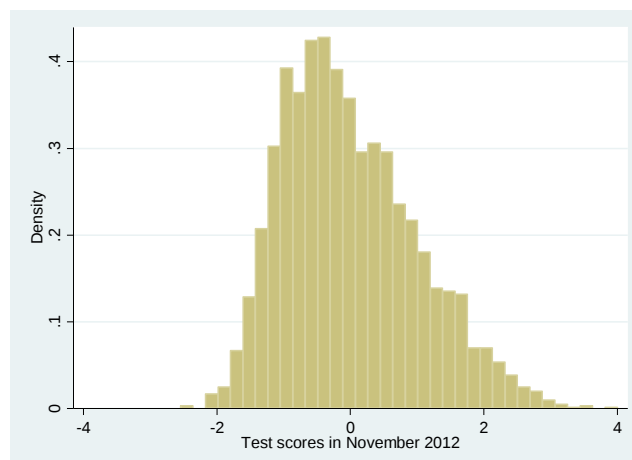


Table 1: Correlation between Aspirations and Yearly grades

Variable	Yearly grades given by teachers					
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	0.688*** (0.036)	0.290*** (0.035)	0.222*** (0.031)	0.217*** (0.029)	0.215*** (0.029)	0.210*** (0.028)
Preferred tracks include Masters	0.458*** (0.040)	0.192*** (0.035)	0.160*** (0.036)	0.126*** (0.043)	0.110*** (0.041)	0.112*** (0.041)
Repeated a grade			-0.284*** (0.038)	-0.292*** (0.041)	-0.277*** (0.040)	-0.277*** (0.041)
Skipped a grade			0.271*** (0.055)	0.219*** (0.059)	0.205*** (0.059)	0.209*** (0.057)
Girl			0.241*** (0.030)	0.239*** (0.030)	0.245*** (0.029)	0.244*** (0.029)
Low-SES family					-0.234*** (0.031)	
Immigrant family					0.012 (0.033)	
Detailed SES						Y
Detailed immigration status						Y
Deciles in test scores in Nov. 2012		Y	Y	Y	Y	Y
Score at Raven matrices			Y	Y	Y	Y
Effort put into the test			Y	Y	Y	Y
Self-Perception of Behavioral Conduct			Y	Y	Y	Y
Class fixed effects				Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.152	0.447	0.514	0.569	0.577	0.581
Mean among students whose pref. tracks do not include Academic HS:					-0.511	
Mean among students whose pref. tracks do not include Masters:					-0.088	

Each column reports the coefficients of a different OLS regression. The dependent variable is the standardized yearly average of teachers' grades. "Pref. tracks include Academic HS" is a dummy indicating that preferred tracks after middle school include academic high school. 'Low-SES family' is a dummy variable indicating that none of the parents are high-skilled workers. Controls for "Detailed SES" correspond to controls for each family's socioeconomic status (SES) stratified into six categories based on the parents' occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for "Detailed immigration status" correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students' score at Raven matrices is controlled for using dummies indicating the number of wrong answers. 'Proxy for Conscientiousness' corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 2: Correlation between Aspirations and Test score at the national exam

Variable	Test scores in June 2013					
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	0.717*** (0.035)	0.239*** (0.028)	0.189*** (0.025)	0.174*** (0.021)	0.171*** (0.021)	0.168*** (0.021)
Preferred tracks include Masters	0.499*** (0.054)	0.161*** (0.037)	0.142*** (0.037)	0.104*** (0.036)	0.087*** (0.033)	0.087*** (0.034)
Repeated a grade			-0.322*** (0.029)	-0.318*** (0.031)	-0.303*** (0.030)	-0.295*** (0.031)
Skipped a grade			0.275*** (0.051)	0.264*** (0.058)	0.249*** (0.057)	0.251*** (0.055)
Girl			0.078*** (0.023)	0.067*** (0.023)	0.069*** (0.023)	0.068*** (0.023)
Low-SES family					-0.235*** (0.030)	
Immigrant family					0.024 (0.025)	
Detailed SES						Y
Detailed immigration status						Y
Deciles in test scores in Nov. 2012		Y	Y	Y	Y	Y
Score at Raven matrices			Y	Y	Y	Y
Effort put into the test			Y	Y	Y	Y
Self-Perception of Behavioral Conduct			Y	Y	Y	Y
Class fixed effects				Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.169	0.616	0.651	0.700	0.708	0.710
Mean among students whose pref. tracks do not include Academic HS:					-0.519	
Mean among students whose pref. tracks do not include Masters:					-0.080	

Each column reports the coefficients of a different OLS regression. The dependent variable is the standardized test score at the national exam in June. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student were born outside of France. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. ‘Proxy for Conscientiousness’ corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 3: Correlation between Aspirations and Assignment to the academic track

Variable	Entered Academic HS					
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	0.364*** (0.018)	0.222*** (0.022)	0.197*** (0.021)	0.182*** (0.020)	0.180*** (0.020)	0.176*** (0.020)
Preferred tracks include Masters	0.113*** (0.017)	0.042*** (0.016)	0.038** (0.016)	0.027 (0.018)	0.019 (0.018)	0.018 (0.018)
Repeated a grade			-0.183*** (0.020)	-0.184*** (0.021)	-0.178*** (0.021)	-0.175*** (0.021)
Skipped a grade			0.035* (0.018)	0.032 (0.021)	0.025 (0.020)	0.022 (0.020)
Girl			0.051*** (0.015)	0.049*** (0.015)	0.050*** (0.015)	0.048*** (0.015)
Low-SES family					-0.103*** (0.014)	
Immigrant family					0.027 (0.017)	
Detailed SES						Y
Detailed immigration status						Y
Deciles in test scores in Nov. 2012		Y	Y	Y	Y	Y
Score at Raven matrices			Y	Y	Y	Y
Effort put into the test			Y	Y	Y	Y
Self-Perception of Behavioral Conduct			Y	Y	Y	Y
Class fixed effects				Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.171	0.354	0.392	0.418	0.425	0.428
Mean among students whose pref. tracks do not include Academic HS:				0.472		
Mean among students whose pref. tracks do not include Masters:				0.698		

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy indicating whether the student has been assigned to the academic track. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student were born outside of France. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. ‘Proxy for Conscientiousness’ corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 4: Shorter term correlation between Aspirations and Academic outcomes, controlling for teachers' grades

Variable	Test scores in June 2013			Entered Academic HS		
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	0.087*** (0.022)	0.065*** (0.020)	0.065*** (0.020)	0.128*** (0.014)	0.115*** (0.014)	0.111*** (0.014)
Preferred tracks include Masters	0.055* (0.031)	0.026 (0.026)	0.019 (0.025)	0.012 (0.012)	0.010 (0.012)	0.004 (0.012)
Repeated a grade	-0.179*** (0.025)	-0.161*** (0.024)	-0.148*** (0.024)	-0.080*** (0.014)	-0.076*** (0.016)	-0.073*** (0.016)
Skipped a grade	0.150*** (0.037)	0.149*** (0.041)	0.145*** (0.040)	0.018 (0.015)	0.025 (0.016)	0.018 (0.016)
Girl	-0.042** (0.019)	-0.067*** (0.017)	-0.065*** (0.018)	-0.010 (0.012)	-0.013 (0.013)	-0.014 (0.013)
Detailed SES			Y			Y
Detailed immigration status			Y			Y
Deciles in test scores in Nov. 2012	Y	Y	Y	Y	Y	Y
Score at Raven matrices	Y	Y	Y	Y	Y	Y
Effort put into the test	Y	Y	Y	Y	Y	Y
Self-Perception of Behavioral Conduct	Y	Y	Y	Y	Y	Y
Class fixed effects		Y	Y		Y	Y
Deciles in average yearly grade	Y	Y	Y	Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.771	0.829	0.831	0.633	0.651	0.654
Mean among students whose pref. tracks do not include...						
... Academic HS:	-0.519	-0.519	-0.519	0.472	0.472	0.472
... Masters:	-0.080	-0.080	-0.080	0.698	0.698	0.698

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy indicating whether the student has been assigned to the academic track. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. ‘Proxy for Conscientiousness’ corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 5: Academic Aspirations after Junior High School

Variable	Vocational High School				Academic High School			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel 1: Preferred tracks after JHS								
Low-SES family	0.119*** (0.013)	0.054*** (0.012)	0.054*** (0.014)	0.025*** (0.009)	-0.184*** (0.017)	-0.061*** (0.015)	-0.038** (0.019)	-0.009 (0.016)
Immigrant family	-0.026* (0.015)	-0.046*** (0.014)	-0.037*** (0.014)	-0.012 (0.011)	0.001 (0.021)	0.043** (0.019)	0.035 (0.023)	0.014 (0.018)
Attainable tracks includes Vocational HS				0.320*** (0.020)				-0.081*** (0.022)
Attainable tracks includes Academic HS				-0.187*** (0.020)				0.610*** (0.020)
Nb Obs	3317	3317	3317	3317	3317	3317	3317	3317
Adjusted R-squared	0.026	0.093	0.128	0.385	0.033	0.134	0.165	0.425
Mean among high-SES families:		0.045				0.795		
Mean among non-immigrant families:		0.115				0.693		
Panel 2: Attainable tracks after JHS								
Low-SES family	0.137*** (0.018)	0.074*** (0.017)	0.068*** (0.020)	0.062*** (0.019)	-0.168*** (0.015)	-0.054*** (0.012)	-0.038** (0.015)	-0.037** (0.018)
Immigrant family	-0.055*** (0.020)	-0.076*** (0.019)	-0.066*** (0.021)	-0.072*** (0.022)	-0.005 (0.022)	0.036** (0.018)	0.026 (0.021)	0.023 (0.021)
Salient tracks includes Vocational HS				0.264*** (0.025)				-0.023 (0.022)
Salient tracks includes Academic HS				-0.023 (0.029)				0.607*** (0.030)
Nb Obs	3317	3317	3317	2661	3317	3317	3317	2661
Adjusted R-squared	0.020	0.069	0.105	0.145	0.036	0.172	0.205	0.361
Mean among high-SES families::		0.152				0.889		
Mean among non-immigrant families::		0.236				0.797		
Panel 3: Salient tracks after JHS								
Low-SES family	0.030 (0.019)	0.066*** (0.020)	0.060** (0.024)		-0.036*** (0.013)	0.015 (0.013)	0.025* (0.014)	
Immigrant family	-0.043*** (0.017)	-0.031* (0.017)	-0.022 (0.018)		-0.010 (0.014)	0.004 (0.013)	0.014 (0.017)	
Nb Obs	2661	2661	2661		2661	2661	2661	
Adjusted R-squared	0.002	0.017	0.051		0.003	0.064	0.086	
Mean among high-SES families:		0.807				0.932		
Mean among non-immigrant families:		0.830				0.914		
Deciles in test scores in Nov. 2012		Y	Y	Y		Y	Y	Y
Dummies for score at Raven matrices in Nov. 2012		Y	Y	Y		Y	Y	Y
Effort put into the test		Y	Y	Y		Y	Y	Y
Self-Perception of Behavioral Conduct		Y	Y	Y		Y	Y	Y
Class fixed effects			Y	Y			Y	Y

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy variable indicating whether the preferred (panel 1) / attainable (panel 2) / salient (panel 3) tracks include vocational HS (columns 1-4) or Academic HS (columns 5-8). 'Low-SES family' is a dummy variable indicating that none of the parents are high-skilled workers. 'Immigrant Family' is a dummy variable indicating that both parents of a student are born outside of France. Students' test scores in Nov. 2012 and average yearly grades are controlled for by deciles. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 6: Academic Aspirations after High School

Variable	Finding a job				Masters			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel 1: Preferred tracks after HS								
Low-SES family	0.143*** (0.013)	0.065*** (0.013)	0.056*** (0.014)	0.048*** (0.013)	-0.148*** (0.024)	-0.088*** (0.023)	-0.061*** (0.023)	-0.012 (0.015)
Immigrant family	-0.021 (0.014)	-0.046*** (0.014)	-0.054*** (0.017)	-0.049*** (0.016)	0.010 (0.014)	0.027** (0.013)	0.028* (0.016)	0.009 (0.010)
Attainable tracks includes 1-4 yrs college				-0.130*** (0.014)				-0.012 (0.011)
Attainable tracks includes Masters				-0.082*** (0.015)				0.627*** (0.027)
Mean among high-SES families:	0.081	0.081	0.081	0.081	0.239	0.239	0.239	0.239
Mean among non-immigrant families:	0.163	0.163	0.163	0.163	0.155	0.155	0.155	0.155
Nb Obs	3196	3196	3196	3196	3102	3102	3102	3102
Adjusted R-squared	0.028	0.110	0.134	0.166	0.037	0.071	0.097	0.500
Panel 2: Attainable tracks after HS								
		No tracks in HE				Masters		
Low-SES family	0.136*** (0.027)	0.037 (0.025)	0.023 (0.028)	0.003 (0.021)	-0.155*** (0.019)	-0.085*** (0.018)	-0.072*** (0.021)	-0.038*** (0.014)
Immigrant family	-0.026 (0.020)	-0.059*** (0.018)	-0.036* (0.020)	-0.016 (0.016)	0.010 (0.014)	0.028** (0.012)	0.031* (0.016)	0.011 (0.013)
Salient tracks includes 1-4 yrs college				-0.477*** (0.021)				-0.005 (0.009)
Salient tracks includes Masters				-0.242*** (0.024)				0.421*** (0.026)
Mean among high-SES families:	0.509	0.509	0.509	0.509	0.261	0.261	0.261	0.261
Mean among non-immigrant families:	0.589	0.589	0.589	0.589	0.174	0.174	0.174	0.174
Nb Obs	3308	3308	3308	2654	3308	3308	3308	2654
Adjusted R-squared	0.015	0.060	0.070	0.409	0.037	0.083	0.091	0.366
Panel 3: Salient tracks after HS								
Low-SES family	0.130*** (0.028)	0.045* (0.027)	0.030 (0.034)		-0.183*** (0.029)	-0.085*** (0.026)	-0.055* (0.032)	
Immigrant family	0.024 (0.023)	-0.004 (0.021)	0.004 (0.026)		-0.013 (0.019)	0.012 (0.016)	0.025 (0.021)	
Mean among high-SES families:	0.278	0.278	0.278		0.377	0.377	0.377	
Mean among non-immigrant families:	0.347	0.347	0.347		0.278	0.278	0.278	
Nb Obs	2654	2654	2654		2654	2654	2654	
Adjusted R-squared	0.018	0.061	0.054		0.040	0.094	0.118	
Deciles in test scores in Nov. 2012		Y	Y	Y		Y	Y	Y
Dummies for score at Raven matrices in Nov. 2012		Y	Y	Y		Y	Y	Y
Effort put into the test		Y	Y	Y		Y	Y	Y
Self-Perception of Behavioral Conduct		Y	Y	Y		Y	Y	Y
Class fixed effects			Y	Y			Y	Y

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy variable indicating whether the preferred (panel 1) / attainable (panel 2) / salient (panel 3) tracks include finding a job (columns 1-4) or Masters (columns 5-8). ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student are born outside of France. Students’ test scores in Nov. 2012 and average yearly grades are controlled for by deciles. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 7: Professional Aspirations

Variable	Level corresponding to job preference after HS												
	No response				No higher ed.			1-4 yrs college			5 or more yrs college		
	(1)	(2)	(3)		(4)	(5)		(6)	(7)		(8)	(9)	(10)
Low-SES family	-0.009 (0.018)	-0.012 (0.018)	-0.023 (0.021)	0.102*** (0.024)	0.025 (0.021)	0.040* (0.023)	-0.014 (0.021)	-0.012 (0.023)	0.008 (0.023)	-0.091*** (0.022)	0.004 (0.023)	-0.003 (0.024)	
Immigrant family	-0.021 (0.016)	-0.019 (0.017)	-0.024 (0.018)	-0.059*** (0.017)	-0.086*** (0.018)	-0.074*** (0.021)	0.029 (0.022)	0.030 (0.023)	0.049** (0.024)	0.082*** (0.021)	0.111*** (0.020)	0.095*** (0.022)	
Deciles in test scores in Nov. 2012		Y	Y		Y	Y		Y	Y		Y	Y	
Dummies for score at Raven matrices in Nov. 2012		Y	Y		Y	Y		Y	Y		Y	Y	
Effort put into the test		Y	Y		Y	Y		Y	Y		Y	Y	
Self-Perception of Behavioral Conduct		Y	Y		Y	Y		Y	Y		Y	Y	
Class fixed effects			Y			Y			Y			Y	
Mean among high-SES families	0.217	0.217	0.217	0.352	0.352	0.352	0.333	0.333	0.333	0.445	0.445	0.445	
Mean among non-immigrant families	0.215	0.215	0.215	0.417	0.417	0.417	0.321	0.321	0.321	0.382	0.382	0.382	
Nb Obs	3325	3325	3325	3325	3325	3325	3325	3325	3325	3325	3325	3325	
Adjusted R-squared	0.000	0.005	0.037	0.009	0.040	0.063	0.000	0.012	0.017	0.009	0.057	0.102	

Each column reports the coefficients of a different OLS regression. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student are born outside of France. Students’ test scores in Nov. 2012 and average yearly grades are controlled for by deciles. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table 8: Explanatory factors

For a hypothetical high achieving student:				
(a) Average perceived prob. of success if s/he lives in an advantaged neighborhood: 0.855				
Variable	(b) Gap with her/his perceived prob. of success if s/he...			Scholastic self-esteem
	lives in a disadvant. neighborhood	has one parent who is an immigrant	has a family member who succeeded	
	(1)	(2)	(3)	
Low-SES family	-0.015 (0.011)	-0.004 (0.014)	-0.012 (0.010)	-0.131*** (0.041)
Immigrant family	-0.001 (0.012)	-0.035*** (0.012)	-0.008 (0.012)	-0.039 (0.038)
Deciles in test scores in Nov. 2012	Y	Y	Y	Y
Dummies for score at Raven matrices in Nov. 2012	Y	Y	Y	Y
Effort put into the test	Y	Y	Y	Y
Self-Perception of Behavioral Conduct	Y	Y	Y	Y
Class fixed effects	Y	Y	Y	Y
Nb Obs	3240	3230	3239	3266
Adjusted R-squared	0.040	0.058	0.034	0.348
Mean among high-SES families	0.366	0.273	0.084	0.297
Mean among non-immigrant families	0.346	0.276	0.071	0.082

Each column reports the coefficients of a different OLS regression. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student are born outside of France. Students’ test scores in Nov. 2012 and average yearly grades are controlled for by deciles. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Reading note: The “probability of success” is the probability that the hypothetical high achieving student follows his preferred track; on average, students estimate this probability to be 85.5%. If s/he lives in a disadvantaged neighborhood, students from high-SES families estimate this probability to be 36.6 percentage points lower (row “Mean among high-SES families”). This gap is not statistically significantly different for students from low-SES families.

Table A.1: Correlation between Aspirations and Academic outcomes: quality of the measure of parents' social background

Variable	Yearly grades given by teachers					Test scores in June 2013				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Preferred tracks include Academic HS	0.688*** (0.036)	0.602*** (0.038)	0.574*** (0.038)	0.571*** (0.039)	0.210*** (0.028)	0.717*** (0.035)	0.590*** (0.032)	0.560*** (0.031)	0.523*** (0.032)	0.168*** (0.021)
Preferred tracks include Masters	0.458*** (0.040)	0.331*** (0.042)	0.329*** (0.041)	0.276*** (0.047)	0.112*** (0.041)	0.499*** (0.054)	0.315*** (0.047)	0.315*** (0.047)	0.256*** (0.044)	0.087*** (0.034)
Low-SES family		-0.542*** (0.040)					-0.764*** (0.047)			
Immigrant family		-0.114*** (0.040)					-0.213*** (0.035)			
Repeated a grade					-0.277*** (0.041)					-0.295*** (0.031)
Skipped a grade					0.209*** (0.057)					0.251*** (0.055)
Girl					0.244*** (0.029)					0.068*** (0.023)
Detailed SES			Y	Y	Y			Y	Y	Y
Detailed immigration status			Y	Y	Y			Y	Y	Y
Deciles in test scores in Nov. 2012					Y					Y
Score at Raven matrices					Y					Y
Effort put into the test					Y					Y
Self-Perception of Behavioral Conduct					Y					Y
Class fixed effects				Y	Y				Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.152	0.225	0.243	0.292	0.581	0.169	0.329	0.360	0.424	0.710
Mean among students whose pref. tracks do not include...										
... Academic HS:	-0.511	-0.511	-0.511	-0.511	-0.511	-0.519	-0.519	-0.519	-0.519	-0.519
... Masters:	-0.088	-0.088	-0.088	-0.088	-0.088	-0.080	-0.080	-0.080	-0.080	-0.080

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy indicating whether the student has been assigned to the academic track. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student were born outside of France. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table A.2: Correlation between Aspirations and Assignment to the vocational track

Variable	Entered Vocational HS					
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	-0.348*** (0.020)	-0.224*** (0.022)	-0.194*** (0.021)	-0.177*** (0.021)	-0.175*** (0.020)	-0.172*** (0.020)
Preferred tracks include Masters	-0.097*** (0.015)	-0.035** (0.015)	-0.028* (0.015)	-0.021 (0.017)	-0.013 (0.017)	-0.012 (0.017)
Repeated a grade			0.228*** (0.019)	0.232*** (0.021)	0.227*** (0.020)	0.226*** (0.021)
Skipped a grade			-0.022 (0.016)	-0.026 (0.018)	-0.019 (0.017)	-0.018 (0.017)
Girl			-0.044*** (0.015)	-0.042*** (0.015)	-0.042*** (0.015)	-0.041*** (0.015)
Low-SES family					0.096*** (0.015)	
Immigrant family					-0.036** (0.015)	
Detailed SES						Y
Detailed immigration status						Y
Deciles in test scores in Nov. 2012		Y	Y	Y	Y	Y
Score at Raven matrices			Y	Y	Y	Y
Effort put into the test			Y	Y	Y	Y
Self-Perception of Behavioral Conduct			Y	Y	Y	Y
Class fixed effects				Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.166	0.320	0.377	0.396	0.403	0.406
Mean among students whose pref. tracks do not include Academic HS:					0.483	
Mean among students whose pref. tracks do not include Masters:					0.268	

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy indicating whether the student has been assigned to the vocational track. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student were born outside of France. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. ‘Proxy for Conscientiousness’ corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table A.3: Correlation between Aspirations and Repeating a grade

Variable	Stayed in Middle School					
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred tracks include Academic HS	-0.016** (0.008)	0.001 (0.008)	-0.003 (0.008)	-0.005 (0.009)	-0.005 (0.009)	-0.005 (0.009)
Preferred tracks include Masters	-0.016** (0.007)	-0.007 (0.007)	-0.010 (0.007)	-0.006 (0.008)	-0.006 (0.008)	-0.006 (0.008)
Repeated a grade			-0.045*** (0.008)	-0.048*** (0.009)	-0.049*** (0.009)	-0.050*** (0.009)
Skipped a grade			-0.013 (0.009)	-0.006 (0.011)	-0.006 (0.011)	-0.004 (0.010)
Girl			-0.007 (0.007)	-0.007 (0.008)	-0.007 (0.007)	-0.007 (0.007)
Low-SES family					0.008 (0.008)	
Immigrant family					0.009 (0.008)	
Detailed SES						Y
Detailed immigration status						Y
Deciles in test scores in Nov. 2012		Y	Y	Y	Y	Y
Score at Raven matrices			Y	Y	Y	Y
Effort put into the test			Y	Y	Y	Y
Self-Perception of Behavioral Conduct			Y	Y	Y	Y
Class fixed effects				Y	Y	Y
Nb Obs	3097	3097	3097	3097	3097	3097
Adjusted R-squared	0.003	0.021	0.034	0.043	0.042	0.048
Mean among students whose pref. tracks do not include Academic HS:				0.046		
Mean among students whose pref. tracks do not include Masters:				0.034		

Each column reports the coefficients of a different OLS regression. The dependent variable is a dummy indicating whether the student has repeated the last grade of middle school. “Pref. tracks include Academic HS” is a dummy indicating that preferred tracks after middle school include academic high school. ‘Low-SES family’ is a dummy variable indicating that none of the parents are high-skilled workers. ‘Immigrant Family’ is a dummy variable indicating that both parents of a student were born outside of France. Controls for “Detailed SES” correspond to controls for each family’s socioeconomic status (SES) stratified into six categories based on the parents’ occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for “Detailed immigration status” correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students’ score at Raven matrices is controlled for using dummies indicating the number of wrong answers. ‘Proxy for Conscientiousness’ corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.

Table A.4: Correlation between Aspirations and Educational Outcomes, by SES

Variable	Yearly teachers' grades (1)	Test scores in June (2)	Entered Academic HS (3)
Preferred tracks include Academic HS	0.123** (0.050)	0.129*** (0.044)	0.103*** (0.035)
Preferred tracks include Masters	0.147*** (0.045)	0.081 (0.053)	0.022 (0.016)
Pref. tracks include Academic HS * Low-SES family	0.121* (0.064)	0.050 (0.050)	0.088** (0.041)
Pref. tracks include Academic HS * Immigrant family	-0.003 (0.075)	0.020 (0.052)	0.033 (0.048)
Pref. tracks include Masters * Low-SES family	0.023 (0.065)	0.051 (0.074)	0.050 (0.035)
Pref. tracks include Masters * Immigrant family	-0.140* (0.079)	-0.085 (0.070)	-0.089** (0.044)
Low-SES family	Y	Y	Y
Immigrant family	Y	Y	Y
<i>Other controls: per se and interacted with low-SES family and immigrant family:</i>			
Repeated a grade	Y	Y	Y
Skipped a grade	Y	Y	Y
Girl	Y	Y	Y
Deciles in test scores in Nov. 2012	Y	Y	Y
Score at Raven matrices	Y	Y	Y
Effort put into the test	Y	Y	Y
Self-Perception of Behavioral Conduct	Y	Y	Y
Class fixed effects (<i>without interactions</i>)	Y	Y	Y
Nb Obs	3097	3097	3097
Adjusted R-squared	0.578	0.707	0.429
Mean among students whose pref. tracks do not include Academic HS:	-0.511	-0.519	0.472
Mean among students whose pref. tracks do not include Masters:	-0.088	-0.080	0.698

Each column reports the coefficients of a different OLS regression. The dependent variable is either the standardized yearly average of teachers' grades, either the standardized test score at the national exam in June, either a dummy indicating whether the student has been assigned to the academic track. "Pref. tracks include Academic HS" is a dummy indicating that preferred tracks after middle school include academic high school. 'Low-SES family' is a dummy variable indicating that none of the parents are high-skilled workers. Controls for "Detailed SES" correspond to controls for each family's socioeconomic status (SES) stratified into six categories based on the parents' occupation, together with controls for whether the student has one parent who is unemployed and for whether she has one parent who is retired, separately for each parent when there are two. Controls for "Detailed immigration status" correspond to controls for whether the student has one parent or both who were born in a non-OECD country, and at least one parent colored (defined base on the country of birth). Students' score at Raven matrices is controlled for using dummies indicating the number of wrong answers. 'Proxy for Conscientiousness' corresponds to dummies indicating the number of questions that the student tried to solve for each test in November. When controlling for the immigration status, for the employment status, and for the score at Raven matrices, we also add controls for missing data for each characteristic. When we use salient tracks as an instrument for preferred tracks, the instruments are two dummies, respectively for academic high school or masters being salient to the student. The standard errors are clustered at the school level and robust; they are reported in parenthesis. * indicates significance at the 10% level, ** indicates significance at the 5% level, *** indicates significance at the 1% level.