Family Background, Self-Confidence and Economic Outcomes *

Antonio Filippin
University of Milan and IZA

Marco Paccagnella
Bank of Italy and Bocconi University

Current version: August, 2011

WORK IN PROGRESS.
COMMENTS ARE WELCOME.

Abstract

In this paper we argue that self-confidence, transmitted by the socio-economic background, is a non-cognitive skill that provides a strong channel through which education and earning inequality is perpetuated from one generation to the next. Self-confidence allows to explain why the early gap observed between children from different socio-economic backgrounds does not narrow when the role of the family becomes less important during life.

We propose a model that describes the evolution of self-confidence, defined as beliefs about one's own ability, in a Bayesian framework. When the learning process does not convergence quickly to the true ability level even small differences in confidence can explain a divergent pattern of human capital accumulation by otherwise identical individuals. Cognitive tests should hence take place as early as possible in order to minimize the effect of family background on measured ability.

Empirical evidence confirms that there is a correlation between family background and self-confidence and that the learning process is slow: inherited beliefs about one's ability can survive a long string of signals.

JEL Classifications: D83, J24, J62
Key Words: Self Confidence, Family background.

*We would like to thank Marcel Jansen, Marco Leonardi, Michele Pellizzari, and seminar participants at Bocconi, Autonoma de Madrid, ASSET, EALE, and ESPE for useful comments and suggestions. All remaining errors are ours. The views experessed in this paper are those of the authors and do not involve the responsibility of the Bank of Italy.

1 Introduction

Economic outcomes such as educational attainments and earnings tend to persist across generations. In many Western countries social mobility is quite low, and parents' socio-economic status is usually a very good predictor of the socio-economic outcomes of their offspring.

Recent literature has emphasized that non-cognitive skills are also correlated with family background and that they also play a relevant role in determining future outcomes like educational attainments and earnings. The importance of family background has been further stressed by James Heckman and co-authors, who show that gaps between children from different backgrounds open up very early in life, as soon as in pre-school age, and tend then to be persistent over the lifetime. The explanation provided for such a persistence is the existence of dynamic complementarities among a multidimensional vector of cognitive and non-cognitive skills, such that an insufficient investment in some of these skills very early in life cause long-lasting consequences very difficult to revert.

In this paper we focus on one non-cognitive skill, namely self-confidence, and we propose an alternative mechanism capable of explaining the persistence of the gaps based on family background. We posit that the socio-economic background affects not only the actual stock of cognitive skills possessed by a child but also self-confidence, that we model as beliefs one holds about his unknown ability. We propose a simple model in which agents extract all the information available from the signals received (success vs. failure in the endeavour undertaken) in order to update their beliefs. The probability of success depends on the true level of ability as well as on the difficulty of the task, which is chosen endogenously given (updated) beliefs about ability. We simulate the model with a bootstrapping procedure, showing that choices distorted by under-confidence (while all the other sources of heterogeneity are neutralized) lead to a significant gap in the accumulation of human capital during the learning process of the true level of ability. In this way self-confidence can contribute to explain why the early gap based on the socio-economic background does not tend to narrow when the role of the family becomes less important during life.

The theoretical framework features different levels of confidence as a primitive without modeling a specific transmission mechanism across generations. Hence, we use the 2006 wave of the

OECD-PISA survey to verify whether there is indeed a correlation between family background and self-confidence, finding supportive evidence. Results are robust to the inclusion in the estimated specification of a very good proxy of cognitive ability, namely the PISA score itself, which is unobserved by the student at the time of the survey.

Following the simulation of the model we also want to establish whether the learning pattern is actually a slow process during which a gap in human capital can accumulate. Towards this goal we use a survey administered to 2nd-year undergraduate students from Bocconi University in Italy. This dataset is interesting because, although not representative, it allows to investigate whether differences in self-confidence survive until late in the academic career even within a homogeneous and highly selected body of students about to acquire a very good signal to be spent in the labour market. The answer is positive, since we find that students with different socio-economic backgrounds hold different expectations about post-graduation earnings. Also in this case the results are robust to the inclusion of a very good proxy for ability. To exclude that different wage expectations across family backgrounds simply reflect different "reference wages" instead of a different self-confidence we replicate the analysis using a similar survey of Bocconi graduates finding the same results.

Our interpretation is that the effect of family background is difficult to tackle because inherited beliefs about one's ability survive a long string of signals. Beliefs updating is a slow process that does not quickly lead to convergence to the true ability level. As a consequence, even relatively small differences in the degree of confidence can help in explaining different achievements of otherwise identical individuals.

The outline of the paper is as follows. Section 2 provides a brief review of the various strands of the literature our work is linked to. In Section 3 we present a theoretical model that motivates the empirical investigation carried out in Section 4. Section 5 concludes.

2 Literature Review

From an empirical point of view, a countless number of papers have studied the intergenerational persistence of inequality in earnings and educational attainments. In this section, however, we prefer to summarize some contributions that relates to the specific mechanism we propose to explain the intergenerational persistence, rather than on the persistence itself. For a recent survey we refer the interested reader to Black and Devereux (2010).

Our work is related to the vast literature on the technology of skill formation, and in particular to those papers that highlight the role played by non-cognitive skills. Cunha and Heckman (2007), Heckman, Stixrud, and Urzua (2006), Cunha and Heckman (2008) and Cunha, Heckman, and Schennach (2010) provide evidence that non-cognitive skills are indeed correlated with family background and that they do matter directly in determining earnings and educational attainments. They find that gaps correlated with the family background open up very early in life, typically before school age, and then remain pretty much constant. At first glance such a persistence is quite puzzling given that childrend spend less and less time within the family as they grow up. As long as there is not perfect segregation based on socio-economic status, one could expect those gaps to narrow when the role of the family becomes less important and the children start accumulating human capital at school, interact with their peers, etc. The data, however, contradict this intuition. The aforementioned authors attribute the persistence of the gaps to the existence of sensitive and critical periods, in which returns to investment in given skills are particularly high, and to dynamic complementarities among different skills. As a result, failing to invest at the right time either because of liquidity constraints or because of parents' poor endowment of some skills implies that it becomes very difficult (and/or very expensive) to build up that particular skill later.

In this paper we propose self-confidence as an alternative mechanism capable of explaining the persistence of the gaps based on family background. We use a much simpler framework since we model the multidimensional vector of skills as containing only two elements, a cognitive skill (ability) and a non-cognitive one (self-confidence). Ability, instead of being private information of the individual as in standard signaling models, is assumed to be unknown. We then model self-confidence as the belief about one's own cognitive skills. Imperfect knowledge about one's own ability is consistent with the evidence that there is a significant overlapping in the distribution of measured ability across educational tracks. In other words, people with the same measured ability happen to make different educational choices. This is what emerges in the figures below. Figure 1, taken from Sjögren and

Sällström (2004), shows the distribution of verbal ability of swedish graduates and non-graduates. Figure 2 displays the distribution of PISA 2006 scores for italian students enrolled in different high-school tracks.[1] In both cases a substantial overlapping in the distribution of ability clearly emerges in spite of a different average. Bowles, Gintis, and Osborne (2001) survey different beahvioural-economics-type explanations for the observed dispersion in earnings among people with similar observable characteristics, including incentive-enhancing preferences and the role of personality traits.

Self-confidence has been widely studied by both psychologists and economists. Dunning, Heath, and Suls (2004) survey the psychological literature documenting the presence of a weak correlation between actual and perceived performance in several domains. In the (theoretical) economics literature, many papers have studied mechanisms capable of generating the emergence (and persistence) of biased beliefs. Bénabou and Tirole (2002) explain a wide range of behaviours documented by psychologists by means of a rational choice model of endogenous production of self-confidence in an intrapersonal game. We follow them in modelling confidence as beliefs over own ability, but we do not resort to hyperbolic discounting. Köszegi (2006) explains the emergence of self-confidence by assuming that agents derive "ego-utility" from holding positive views abouth themselves, while Hvide (2002) proposes the notion of "pragmatic beliefs," according to which an agent forms beliefs that are the most useful to him. In our model agents have standard preferences, and we simply assume that biased beliefs are inherited or transmitted through the family background. Cesarini, Johannesson, Lichtenstein, and Wallace (2009) provide support for our hypothesis by estimating that genetic differences explain between 16% and 34% of the variation in overconfidence, and that common environmental differences (parental style, for instance) explain another 5 to 11%.

The paper most closely related to ours is Sjögren and Sällström (2004). They describe the endogenous evolution of self-confidence when rational agents update their beliefs in a Bayesian fashion and they show how people can remain "trapped" with wrong beliefs due to insufficient experimentation and learning. However, their notion of confidence refers to the *precision* of beliefs, while we focus instead on their *level*.

Other theoretical contributions worth citing are the papers studying feedback provisions in tour-

[1] A similar figure appears in Checchi and Flabbi (2007), who study the impact of family background on educational choices in tracked educational systems.

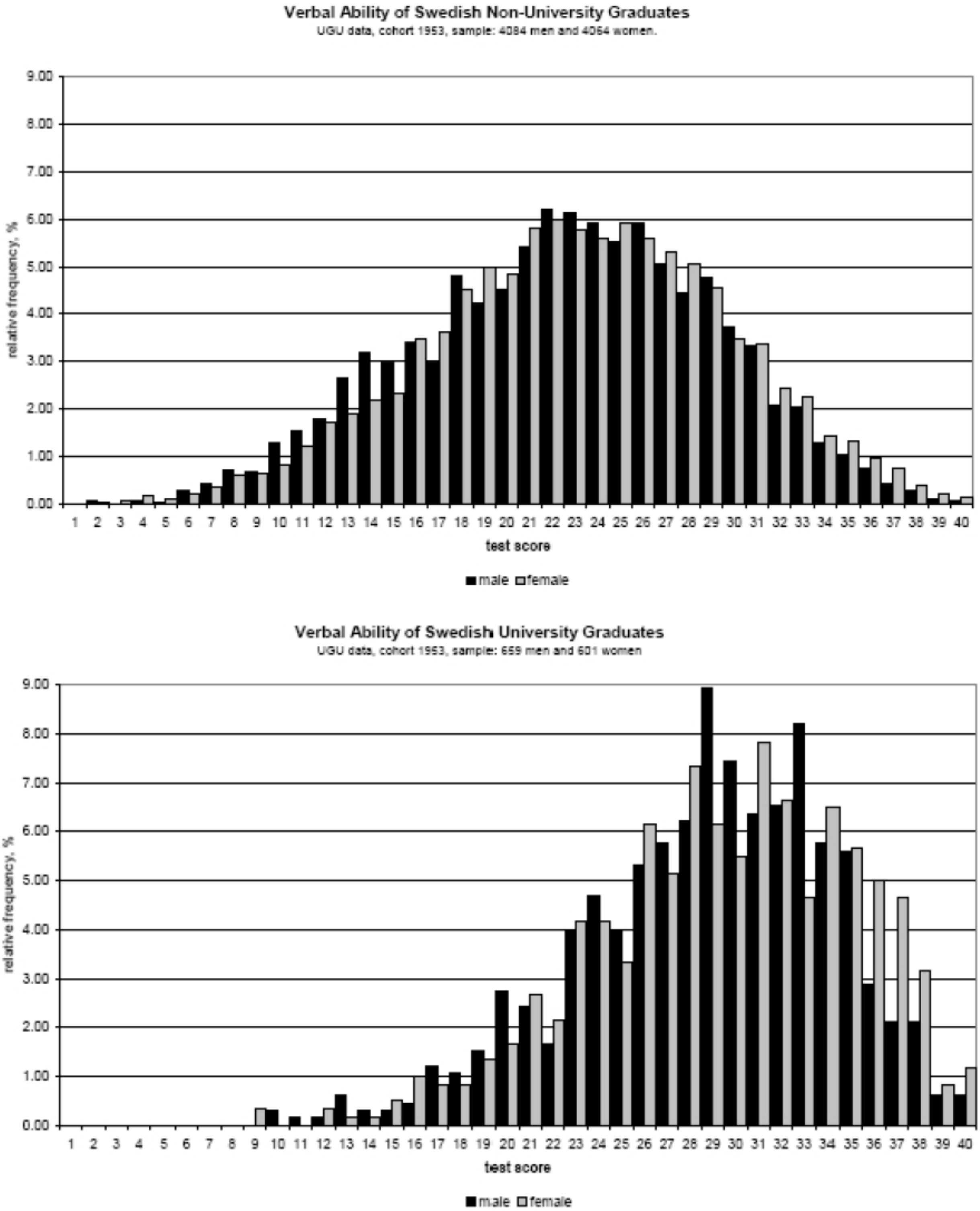

Figure 1: Verbal ability of Swedish graduates and non-graduates

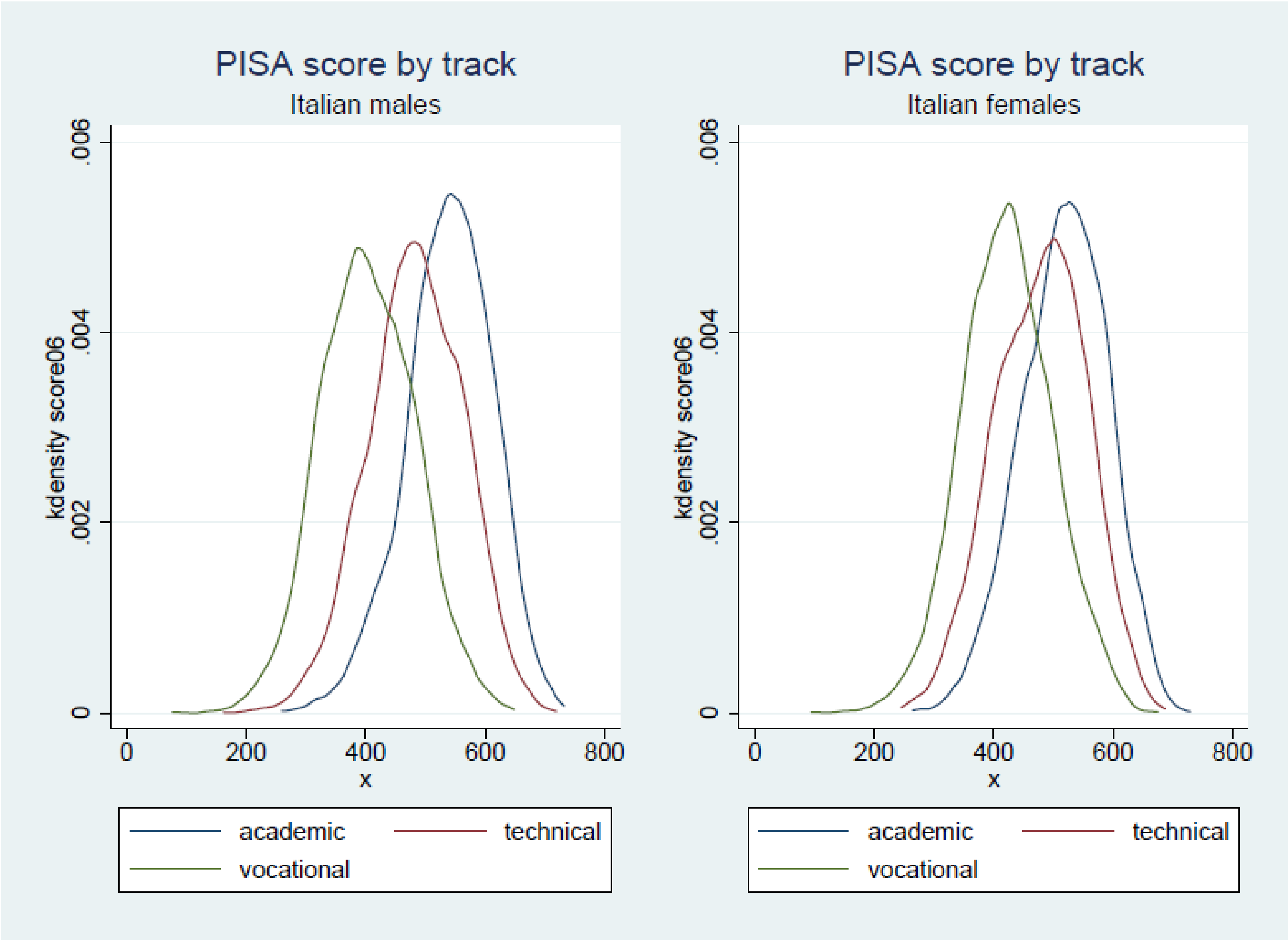

Figure 2: PISA score by track

naments in which agents are uncertain about their ability. In such a setting, it is relatively straightforward to show how people can remain trapped with wrong beliefs, lacking the incentives to experiment and learn (Ederer, 2010; Krahmer, 2007; Squintani, 2006). In our framework there is no strategic interaction among agents. Indeed, people eventually learn their true level of ability, but the slowness of this process can have persistent adverse consequences. Falk, Huffman, and Sunde (2006a) build a standard search-and-matching model of the labor market in which agents are uncertain about their relative ability to find a job, and show that workers can become discouraged and give up search. In a companion paper (Falk, Huffman, and Sunde, 2006b) they test their model in a laboratory experiment, finding that people are substantially uncertain about their relative ability and that this have indeed important consequences on search decisions.

Many empirical papers have studied the consequences of personality traits like self-confidence or self-esteem on educational and labour market outcomes, often providing mixed evidence. Chevalier, Gibbons, Thorpe, Snell, and Hoskins (2009) find that high school pupils with a more positive view of their academic abilities are more likely to expect to continue to higher education. Low social class seems to negatively affect academic self-perception of undergraduates, but not that of high-school pupils. This is partly in contrast with our findings and to those of Sullivan (2006). Marsh (2005) show that students who are better at assessing themselves have better academic outcomes, but Baumeister, Campbell, Krueger, and Vohs (2003) find no causal effect of self-esteem on academic outcomes; on the oter hand, Drago (2008) provide evidence for the existence of a positive self-esteem premium on earnings, after appropriately controlling for the large measurement error in the reported self-esteem measure.

Our model features decisions under uncertainty, although we assume risk neutrality, and therefore it is also linked to the literature that sees education as a risky investment and that investigates the role played by risk aversion. Belzil and Leonardi (2007) find that risk aversion can be a deterrent to investing in education, but that differences in risk attitude account for a modest portion of the probability of entering higher education. Belzil (2007) suggests that under-estimation of ability (which is the notion of confidence that we adopt in our paper) is particularly relevant among high-ability individuals, while in the medium-ability range over-estimation is more common. Overall, over-estimation appears to be more common than under-estimation.

3 The Model

We propose (lack of) self-confidence as one of the channels through which levels of educational attainments and earnings perpetuate across generations. By acting as role-models, parents transmit to their children beliefs about their (unknown) ability. Such beliefs shape educational choices, which in turn can either make wrong beliefs self-confirming (i.e., not disconfirmed by evidence) in a limit case, or having long lasting effects because they contribute to widen the gap in educational attainments while the learning process take place.[2]

The model presented in this paper is inspired by Sjögren and Sällström (2004) and it is characterized by a multi-period career choice that highlights the role played by confidence in explaining educational attainments.

We assume that children don't know their own ability a and hold a belief represented by the density function $\mu(a)$. We define confidence the perceived ability $\hat{\mu}(a) = \int a\mu(a)da$ and underconfident a student who underestimate her ability: $\hat{\mu}(a) < a$. Similarly, the overconfident is characterized by $\hat{\mu}(a) > a$. Students make educational choices by choosing "tracks". We think of tracks as a rather general concept, encompassing either "real" school tracks (eg. academic vs. vocational high schools) or any goal that the student sets herself. In the latter sense a track could well be interpreted as the amount of knowledge encompassed in a concept. More difficult tracks (in both interpretations) are obviously more costly in terms of effort, but they also yield higher payoffs in case of success. A failure could be interpreted either as a "true" failure in a "real" track (eg. the student decides to drop out) or as the chance that, in trying to deeply understand some difficult material, the student "wastes" energy and time, ending up learning less than she would have been, had she been less ambitious.

Let us assume that the probability of success is given by

$$p\left(s\right) = f\left(a, \psi\right) \tag{1}$$

where ψ represents how difficult is the track chosen. The probability of success is assumed to be increasing in ability $f'\left(a\right) > 0$ and decreasing in the difficulty of the track $f'\left(\psi\right) < 0$.

Students have then the possibility of updating their beliefs using Bayes' rule, when additional information can be derived from the outcome of their choice. Given a generic density of prior beliefs $\mu\left(a\right)$, posterior beliefs after receiving the signal s are equal to:

$$\mu\left(a|s\right) = \frac{p\left(s\right)\mu\left(a\right)}{\int p\left(s\right)\mu\left(a\right)da} \tag{2}$$

Success in the track chosen also allows the agent to add human capital $k\left(\psi|s\right)$ to working life pro-

[2]We provide some evidence that such a correlation exists in section 4 but we do not analyze specific transmission mechanisms across generations, leaving such a task for future research.

ductivity, and agent maximizes the instantaneous utility by choosing the track that optimally balances the expected acquisition of human capital with a convex cost of acquiring it $U\left[p\left(s\right)k\left(\psi\right)-\psi^{2}\right]$, given his/her confidence about unobserved ability.

If the track chosen is totally uninformative (e.g. $p\left(s\right)=1$) the student does not gather evidence that contradicts his/her wrong beliefs. For istance, this may happen when there is a discrete set of tracks and the less able students self-select into the easiest track characterized by no probability of failure. This is admittedly a limit situation, and therefore we prefer to concentrate on what happens to the gap in the educational attainments when agents learn from observed outcomes and proceed with Bayesian updating of their beliefs until their confidence eventually converges towards the true value of ability.

To achieve this goal we make some simplifying assumptions. First, we assume that the probability of success is linear in ability. The reason is that we want to concentrate on the role played by the *level* of one's perceived ability, and not by the *precision* of such belief. This is a major difference with respect to the model in Sjögren and Sällström (2004), who assume that the probability of succesfully acquiring skills of type c_1 is $p\left(s\right)=a^{c_1}$, where $a\in\left[0,1\right]$ is the agent's unknown ability, while $c_1>1$ measures the ability elasticity of success. In such a framework the precision of the signal is crucial, because uncertainty about ability makes riskier options more/less attractive depending on whether the probability of success is convex/concave in ability. In fact, Sjögren and Sällström only focus on the variance of agents' beliefs. They compare the outcomes of individuals who know they are of average ability with the outcomes of individuals who have a uniform prior over the whole support. For instance, what could happen with a convex probability of success is that a totally uncertain agent could think to have more chances of succeeding than agent characterized by quite a precise belief of being above the average. By contrast, we choose to remove such an effect by assuming linearity in ability in equation (1) because we want to focus on overconfidence and underconfidence and their effects in terms of unequal outcomes.[3] Hence, we assume the following functional form of the probability of success:

[3]Note that in order to neutralize the effect of the precision of beliefs it is not enough to assume the same variance of prior beliefs, because at different level of confidence the impact of the variance would be different as long as the probability of success is not linear in ability.

$$p(s) = \psi a + (1 - \psi). \tag{3}$$

Notice that this specification implies that the importance of ability is proportional to the difficulty of the track. For the probability of success to be properly defined we need ability to have a finite support $a \in [0, 1]$. For the sake of simplicity let us assume that $\psi \in (0, 1]$. The extreme value $\psi = 0$ would correspond to the uninformative case mentioned above in which ability does not matter and the signal is totally uninformative (see Figure 3).

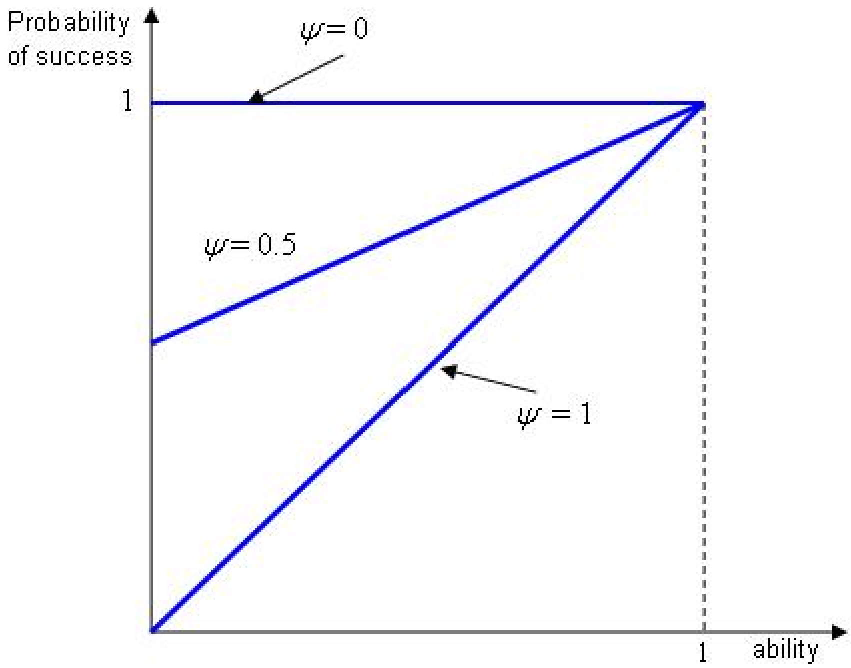

Figure 3: Different tracks in terms of importance of ability

We also assume that more difficult tracks allow students to acquire more human capital if successful, and in particular that the level of capital is equal to

$$k(\psi, \mu(a), a|s) = \frac{\psi}{1 + f(m)}, \tag{4}$$

where $m = (a - \hat{a})$ represents the mismatch of ability as compared to the optimal level $\hat{a}$ for each track, while a failure leaves the stock of human capital unchanged $k(\psi, \mu(a), a|f) = 0$.[4] We assume that $f(0) = 0$ i.e. that human capital concides with the difficulty of the track when ability perfectly fits, otherwise ψ is corrected, with the shape of $f(m)$ when $m \neq 0$ crucially affecting the results. In particular, we assume that $f'(|m|) \geq 0$ meaning that neither under- nor overconfidence can increase human capital beyond ψ. This assumption might appear counterintuitive at first glance, but it has the great advantage of preventing self-deception. Consider the case in which in the same track the human capital is lower only for the overconfident successful students, because their ability is lower than what optimal for such a track, while the opposite happens for the underconfident successful students. In this case the possibility of supplmenting the human capital provided by the chosen track with an ability higher than $\hat{a}$ implies that there is room for self-deception, i.e. that systematically underestimating one's ability might become an optimal solution, with a consequent bias in the choice of the track that we want to avoid. Of course, the effect of the mistake in evaluate ability does not need to be symmetric. In the simulation below we will assume that underconfidence has no effect ($f(m) = 0$ when $m < 0$), while overconfidence has a negative impact ($f'(m) > 0$ when $m > 0$).

Students are free to self-select into different tracks given their best estimate of ability, trading off a lower human capital in case of success with a higher probability of acquiring it. If ability was known FOC would imply:[5]

$$\psi^* = \frac{1}{2}\frac{1}{2-\hat{a}} \tag{5}$$

Given the incentive to truthfully self-report one's unknown ability, i.e. to set the mistake $\mu(a) - a = 0$, the optimal choice of track becomes an increasing function of confidence. However, even removing any bias in the self-evaluation of ability, $\mu(a)$ and $\hat{a}$ may still differ due to insufficient

[4]These assumptions are made for the sake of simplicity, without loss of generality as compared to the case in which the human capital accumulated in case of failure is positive but strictly lower: $k(\psi, \mu(a), a|s) > k(\psi, \mu(a), a|f)$.

[5]To analyze the role played by self-confidence in shaping the gap in educational attainments when agents are eventually learning their true level of ablility we need to iterate this choice for several periods. In principle, we should compute the optimal track choice by maximizing a lifetime utility function. Since additional information about one's ability is valuable per se as long as it helps making better choices in the future, agents could be willing to pay a price to receive a more informative signal, by choosing a track slightly different than what would be optimal in a static framework. However, such an effect is of a second order magnitude and it does not determine appreciable changes in the results (see footnote 11 below), thereby not justifying the corresponding increase in the complication of the model. Hence, we assume that agents are myopic and that they maximize their expected utility period by period.

information. Equation 5 therefore implies that both under- and over-confidence would determine a suboptimal choice of track and a loss of utility due to a mismatch between the perceived ability of the student and the chosen track $\mu(a) \neq \hat{a}$.

On the other hand, the effect of under- and over-confidence can differ as far as the accumulation of human capital is concerned. Rewriting confidence as the composition of optimal ability and the evaluation mistake $\mu(a) = \hat{a} + m$ we can derive that the expected human capital is given by:

$$E(k) = -\frac{1}{4}\frac{\hat{a} + 2m - 3}{(\hat{a} + m - 2)^2(1 + f(m))}. \tag{6}$$

The relationship between confidence and human capital can be summarized by means of the derivative of $E(k)$ with respect to the mistake m:

$$\frac{\delta E(k)}{\delta m} = \frac{1}{2}\frac{m - 1}{(\hat{a} + m - 2)^3(1 + f(m))} + \frac{1}{4}\frac{(\hat{a} + 2m - 3)f'(m)}{(\hat{a} + m - 2)^2(1 + f(m))^2}. \tag{7}$$

As long as a small ability mismatch has a negligible impact, i.e. as long as $f'(0)$ is sufficiently small, the derivative is positive around $m = 0$ for every value of $a \in [0, 1]$. This means a little bit of overconfidence $(m > 0)$ increases the amount of expected human capital, although at a price of lower utility because the increase of human capital would be acquired overestimating the expected return on the additional effort.[6] As overconfidence increases, the sign of $\delta E(k)/\delta m$ depends on the effect of the mismatch. In the limit case in which there is no effect, e.g. when $f(m) = 0$ in Equation 4, or in any case when such an effect is negligible, the human capital acquired would monotonically increase with overconfidence since the positive effect of the higher human capital acquired when successful dominates the negative effect of a lower chance of this event to happen. In contrast, if the effect of overevaluating one's ability increases with the size of the mistake (e.g. if $f(m) = m^2$) the relation between expected human capital and overconfidence becomes bow-shaped. As far as underconfidence is concerned, the condition that ensures that there is no incentive to self-deception is also sufficient to grant that human capital decreases monotonically as underconfidence increases.

[6]The reason is that the probability of success depends on the true level of ability, and overconfidence would grant a higher level of human capital when successful, but a positive outcome is less likely to happen than what an overconfident agent expects.

Agents update their beliefs given the signal received (success Vs. failure) at the end of each period.[7] In order to characterize the learning process and to investigate the effect of self-confidence on educational attainments we need to specify how beliefs about one's ability are shaped. The Beta distribution perfectly fits our assumption of a finite support of the ability distribution, ensuring that the probability of success is linear in ability. At the same time the Beta distribution is sufficiently general to allow prior beliefs to represent different levels of confidence while keeping the whole domain of ability in their support.

The density function of the $Beta\left[\alpha,\beta\right]$ distribution is:

$$\mu(a)=\frac{a^{\alpha-1}(1-a)^{\beta-1}}{\int_0^1 a^{\alpha-1}(1-a)^{\beta-1}da}, \tag{8}$$

while the mean is given by:

$$\hat{\mu}(a)=\int_0^1 a\mu(a)da=\frac{\alpha}{\alpha+\beta}. \tag{9}$$

When $\alpha=\beta>1$ the distribution is symmetric and bell-shaped. The distribution is skewed to the left when $\alpha>\beta>1$, and to the right when $\beta>\alpha>1$.[8] The higher α and β, the lower the variance and therefore the more precise the beliefs. We assume that ability is distributed in the population following a $Beta\left[2.5,2.5\right]$, and that the same distribution also characterizes the beliefs of the median student. This is equivalent to assume that the median student ($a=0.5$) holds correct beliefs about his/her ability, because when $\mu\left(a\right)\sim Beta\left[2.5,2.5\right]$ confidence is $\hat{\mu}(a)=0.5$.

Before analyzing the effect of over- and underconfidence let us focus on the median student in order to describe in some details the learning process. After observing the outcome, the agent updates her beliefs using Bayes rule. In particular, her posterior beliefs after observing a success are:

[7] Note that was the agent receiving a perfectly informative signal like the amount of human capital actually acquired when this is one to one related with ability, he could invert $k\left(\psi,\mu(a),a|s\right)$ deriving with certainty her true ability level. Data suggest that uncertainty about ability survives many signals, which therefore are not perfectly informative (or even if they are perfectly informative agents cannot fully exploit them). In what follows we assume that agents only observe the event success vs. failure. In other words, agents know only the potential amount of human capital ψ but not the actual amount once corrected for the mismatch of ability $1+f(m)$. An intermediate situation in which additional information can be extracted from a noisy signal of the level of human capital actually acquired (in other words when different degrees of success are observable) could be formalized at the price of a significantly increased complication of the model without appreciable additional insights. Hence, we prefer to stick to the simplest version of the information structure.

[8] The Uniform is a special case of the Beta distribution when both parameters are equal to 1.

$$\mu(a|s) = \frac{(\psi a + 1 - \psi)\mu(a)}{\int_0^1 (\psi a + 1 - \psi)\mu(a)da} \tag{10}$$

By contrast, if a failure was observed:

$$\hat{\mu}(a|f) = \frac{(\psi - \psi a)\mu(a)}{\int_0^1 (\psi - \psi a)\mu(a)da} \tag{11}$$

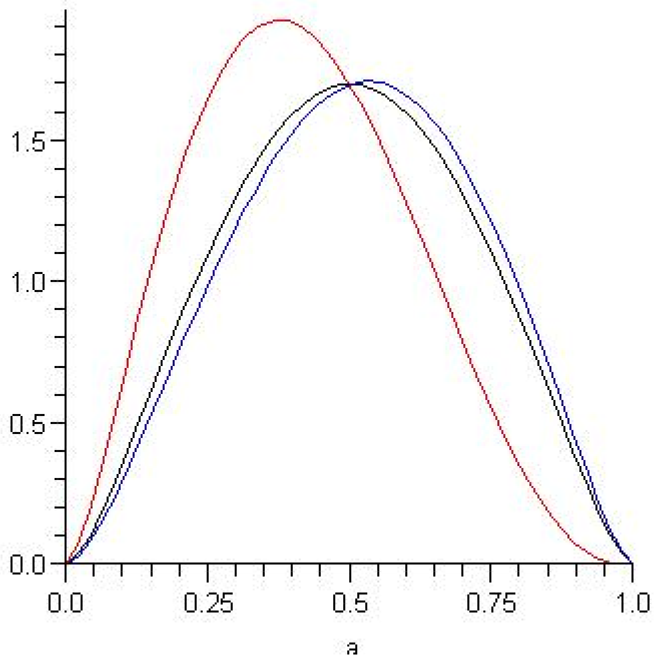

Figure 4: Beliefs updating of the median student after the first signal

The mass of probability is reallocated according to the realization of the signal, towards the upper bound if successful (see Figure 4 blue curve) and toward the lower bound if not (see Figure 4 red curve), keeping contstant the support of the density. Notice that the bad event has a stronger effect when updating beliefs.[9]

The agent will then choose again the optimal track given posterior beliefs, that will be further revised after observing the outcome in the second period, and so on and so forth. The bottom line is

[9]The reason is that a failure is far less likely given the specification of the model. In fact, the student with correct prior beliefs will revise her confidence upward a fraction $1 - 0.5\psi$ of the times, while she will revise her confidence downward in the other 0.5ψ times. While her confidence in expectation stays always unchanged at the correct level of 0.5, the upward and downward revisions would be symmetric only when $\psi = 1$ and the two events are therefore equally likely.

that, within the support of initial beliefs, the distribution of beliefs changes according to the history of signals observed. Subsequent updates bring beliefs closer and closer to the true ability level when prior beliefs are wrong as long as the agent continues receiving informative signals.

To analyze the effect of self-confidence we analyze the choices made and the human capital accumulated by an agent whose ability is always $a = 0.5$ when she holds correct prior beliefs on average $\mu(a) \sim Beta\,[2.5, 2.5]$, and comparing them with the counterfactuals in which she is underconfident and overconfident, respectively. In other words, we simulate the model picking up the median student and looking at the twofold effect in her educational attainments of a wrong confidence. In fact, the higher human capital accumulated when the student is not too overconfident and successful can be compensated by a probability of achieving it that is lower for two reasons. First, because the track is more difficult and therefore the same person is more likely to fail. Second, because the true ability is lower than confidence. In the utility maximization only the former is correctly internalized, and the student will therefore be successful less often than she expects. This is the engine that drives her confidence towards the true level of ability.

We represent underconfidence with a distribution of prior beliefs $\mu(a) \sim Beta\,[1.5, 3]$ skewed to the right, which implies a level of confidence ($\hat{\mu}(a) = 1/3$) corresponding to the 24th percentile in the true distribution. Similarly, overconfidence is summarized by a distribution of prior beliefs $\mu(a) \sim Beta\,[3, 1.5]$ skewed to the left, which implies a level of confidence ($\hat{\mu}(a) = 2/3$) corresponding to the 77th percentile in the true distribution. These parameters also imply that the three distributions have roughly the same variance, and therefore that over- and underconfidence are perfectly symmetric.[10] Prior beliefs of the three different types of student are summarized in Figure 5. As far as the ability mismatch described in Equation 4 is concerned, we choose no correction in case of underconfidence ($f(m) = 0$ if $m < 0$) and a quadratic term $f(m) = 3m^2$ if $m > 0$ that implies a discount of about 7.5% in the human capital acquired in the first period by the overconfident student if successful.

[10] Although the probability of success does not depend on the variance of beliefs, the latter could still affect the updating process, since the more precise the beliefs, the lower the change of confidence induced by the same signal received. We do not want the learning pattern to be affected by a different precision of beliefs, and therefore we assume the same variance in the prior distributions.

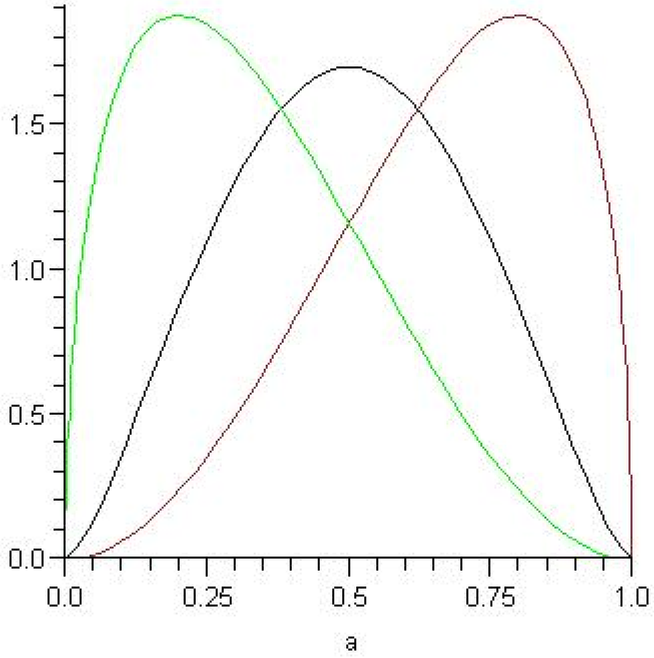

Figure 5: Prior beliefs given the different levels of confidence

We analize what happens to the human capital accumulated by the three types while the learning process takes places, iterating the updating of beliefs 30 times. Since the single realization of human capital relies upon a random component, we replicate the procedure 200 times.

The value of confidence slowly converges towards the true ability level for those starting with a wrong prior, but the learning process is far from being completed at the 45^{th} iteration. In fact, confidence is about $.425$ for the underconfident and $.558$ for the overconfident, in both cases significantly different than $.5$ ($|p| < 0.001$).[11]

Figure 6 displays the average across repetitions of the gap, period by period, in the accumulation of human capital of the types who start with wrong priors as compared to the student starting with

[11] The speed of convergence of the two types differs a little bit. In fact, the mistake in confidence becomes significantly smaller for the over confident ($|p| = 0.038$). The reason is that the higher the track chosen, the more balanced the probability of success *given the same ability* $a = 0.5$, the more informative the signal. At first glance this seems to imply that the choice of track and the educational outcomes could have been different had we internalized the different informativeness of the signals by means of dynamic optmiziation. In fact, there seems to be an additional incentive to choose a higher track thereby reducing the effect of underconfidence while increasing that of overconfidence. This is not the case, however, because such an argument holds only when the probability of success is computed holding constant the true value of ability. When choosing ψ, in contrast, agents use their best estimate of ability $\mu(a)$ and the perceived probability of success is increasing in $\mu(a)$. Hence, internalizing the different informativeness of the signal would imply a downward revision of the choice, thereby exacerbating the effect of underconfidence while reducing that of overconfidence. In any case, maximizing utility period by period implies choices that marginally differ in terms of magnitude, and therefore a negligible mistake.

correct beliefs. The human capital accumulated by the underconfident is significantly lower than the human capital acquired by the student holding correct beliefs ($|p| < 0.001$), while the opposite happens for the overconfident type ($|p| < 0.001$), though the magnitude is different in absolute terms because of the cost of the mismatch $f(m)$. Notice that at the beginning, when the overconfidence is larger and therefore so do the cost of mismatching, the human capital accumulated is not higher, while it increases as compared to the student with correct beliefs, as confidence converges towards the true type and the cost of mismatch decreases. Given the specification of the model chosen, the gap between the overconfident and the underconfident turns out to be about 6%.

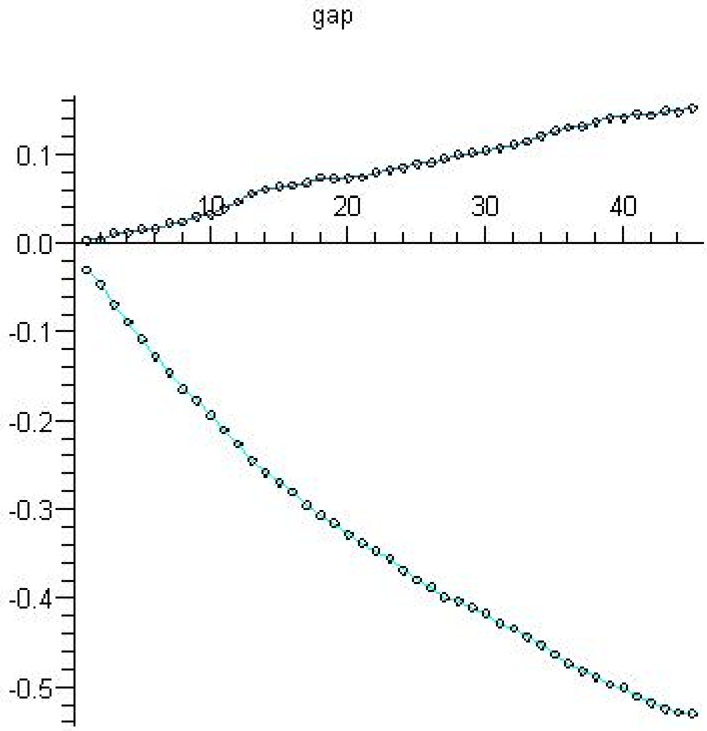

Figure 6: Gap in the accumulation of human capital

This means that self-confidence can determine significant differences in the outcomes observed. When the learning process reaches the fixed point implied by discovering the true level of ability, the three types in the simulation will start making the same choices and from that moment onwards they will be observationally equivalent. However, the level of human capital acquired is and will remain significantly different. Wrong beliefs about one's ability do not need to be self-confirming to explain unequal outcomes as long as they lead to significantly different choices during the learning process. As long as the family background shapes beliefs about children ability, confidence can therefore be a transmission mechanism that increases the intergenerational persistence of outcomes.

Notice that in the model the probability of success increases with innate ability while the human capital accumulated plays no role. This simplifying assumption downplays the role of nurture, since achievements are also determined by the whole history of intermediate outcomes, as well as by the environment in which the children grow, over and above the role of confidence. However, what found by the model is again a lower bound of the role of self-confidence since the difference of human capital caused by different beliefs about one's ability is not taken into account. The role of nurture further implies that tests meant to measure students' ability are instad capturing also the gap in human capital accumulated up to that point because of a different family background. For instance, a centralized test administered at age 15 in order to select students into different tracks would classify as different two students characterized by the same innate ability but with a different background, thereby helping to perpetuate intergenerational inequalities. A policy implication arising from the model is therefore that cognitive tests should take place as early as possible in order to deliver a measure of the innate level of ability of the children that is less affected by the family background.

To summarize, in this section we have shown through a bootstrapping procedure of a very simple model that self-confidence can have significant effects on economic outcomes even when students eventually learn their true type. However, the model starts from different levels of confidence, while the link between family background and self-confidence has yet to be shown. Moreover, for such a simulation exercise to be relevant, we have to provide some evidence that the learning process is actually slow, so that it makes sense to posit that a gap is accumulating over time. These two goals are pursued in the following section.

4 Empirical evidence

What we do in this section is twofold. First, we document the empirical relationship between socio-economic background and self-confidence, using two different sources:

1. The 2006 wave of the OECD-PISA survey, which tested the skills of 15-years old pupils in mathematics, reading and science. The goal is to verify whether there is indeed a correlation between family background and self-confidence in a large, representative dataset.

2. A survey administered to 2nd-year undergraduate students from Bocconi University. In this case the goal is to investigate whether differences in self-confidence survive later on in the academic career, even within an homogeneous body of student that are about to acquire a strong signal to be spent in the labour market.

Second, we use a survey of past graduates from Bocconi University in order to test whether the correlation between the socio-economic background and our proxy for self-confidence finds a counterpart in the realizations in the labour market.

4.1 OECD-PISA Dataset

The survey contains different variables that could capture students' self-confidence. We follow **?** using as dependent variable Science Self-Efficacy, an index built from student's answers when asked to rate the ease with which they believe they could perform eight scientific tasks. This variable is a very good proxy for self-confidence since it is meant to go "beyond how good students think they are in subjects such as science. It is more concerned with the kind of confidence that is needed for them to successfully master specific learning tasks, and is therefore not simply a reflection of a students abilities and performance" (see OECD (2009))[12].

Among the several variables that could capture the family background of pupils we prefer a comprehensive index of socio-economic status. Regressing it on the other candidates - parental education, an index of home possessions, an index of cultural possessions at home, and the highest educational level reached by parents - we can see that it is strongly correlated with all of them, as Table 1 shows.

INSERT TABLE 1 ABOUT HERE

[12]See Ferla, Valcke, and Cai (2009) for a discussion on the differences between Self-Efficacy and Self-Concept. Since Self-Efficacy solicits goal-referenced evaluation and do not ask students to compare their ability to that of others, we believe it is a better proxy to the "absolute level" of confidence that we use in the Model of Section 3

Our estimation strategy is to regress our measure of confidence on family background, adding controls at the individual, school and family level. Results are presented in Table 2, which does not report the coefficient of immigrant and country dummies included in all the specifications.

INSERT TABLE 2 ABOUT HERE

In the first column we regress self-efficacy on the index of socio-economic status, its square, a dummy equal to one when students are from a country with a tracked educational system, and its interaction with the family background variable. The relationship between Self-efficacy and family background is significant and positive as expected, displaying a convex correlation. In contrast, tracking seems to have a negative effect on Self-Efficacy. There are two non-mutually exclusive explanations for this finding. First, being "tracked down" depresses students' level of confidence as compared to their "clones" in a comprehensive educational system. Second, self-efficacy also contains an implicit *relative evaluation* component, so that being "tracked up" has a negative impact on the self-confidence of the relatively less able students because of the higher average quality of their peers. The second explanation also gathers support from the negative coefficient of the interaction term, which means that self-confidence increases more with the socio-economic status when students are enrolled in a comprehensive educational system rather than when they self-select into different tracks.

Although self-efficacy is meant to measure confidence rather than ability, the two variables are certainly correlated, therefore in the second column we also control for the score obtained by the student in the Science section of the test. This is a proxy for "true" ability, comparable across students in different countries and unobserved by the student at the time of filling in the questionnaire. Notice that if our model is correct, the PISA score already encompasses a gap in the human capital accumulated up to that point. In other words, two students with the same innate ability but characterized by a different self-confidence should display a different PISA score. Hence, using the score as a proxy for ability is likely to underestimate the role played by selfconfidence. The inclusion of PISA scores captures some variance of self-efficacy, but the positive relationship between family background and self-efficacy remains strong although the magnitude of the coefficient decreases.

We then add further student-level variables to control for student effort, like the time devoted to out-of-school-time lessons and the time devoted to studying or doing science homework (column 3) and for personal motivation, like the reported interest in learning science and the personal value assigned to science (column 4). All these regressors are positively related to self-efficacy, as expected, but they barely affect the coefficients of the socio-economic background.

Finally, in the last column we control for additional variables taken from the parents' questionnaire that could shed some light on the channels through which family background has an impact on pupils' self-efficacy.[13] Many of those variables turn out to be not significantly related to self-efficacy. A positive link emerges between self-efficacy and whether the parents reported having done science-related activities with their children when they were 10 years old. Notice also that adding such school- and parent-level variables the number of observations plummets. What matters for our purposes is that we still observe a positive and significant correlation between self-efficacy and our measure of family background.

We interpret our results as evidence that students from different socio-economic background hold different beliefs about their own ability, even after controlling for true, unobserved ability (ie., PISA test scores) and a wide range of variables at the individual, school and family level.

The PISA dataset has the advantage of being a large-scale, international, representative sample that however includes extremely heterogeneous students at an early stage of their education career. Therefore, we replicate a similar analysis using another dataset with opposite characteristics, i.e. very homogeneous although not representative sample of students surveyed later in their academic career.

4.2 Dataset of Bocconi Students

This dataset has been collected by circulating a questionnaire attached to the evaluation forms to all 2nd year Bocconi students in 2001, subsequently merged with administrative data. It contains information about students' expectations on occupation and wages 1 and 10 years after graduation, about their family background, as well as detailed information on their academic career. We refer to

[13]The regression also includes some variables taken from the school questionnaire: school size, student-teacher ratio and dummies for public schools as well as for whether students are sorted by ability within the school. The coefficients associated with these variables are generally not statistically significant and therefore they are not reported to save space.

Filippin and Ichino (2005) for further details on the characteristics of the dataset. We use expected wage as a proxy for self-confidence, while our family-background variables are derived from parents' educational levels and the students' tuition category, which is a function of family income.

INSERT TABLE 3 ABOUT HERE

Table 3 reports results from regressing the logarithm of expected wage one year after graduation on family background variables and individual controls. Other controls include a female dummy, the average university grade, the grade obtained in the high-school exit examination and whether the student expect to end up working in a family business. Notice that also in this case the proxies for ability are likely to bias downward the role played by self-confidence, because if our model is correct they should already encompasses a gap in the human capital accumulated up to that point. In other words, two students with the same innate ability but characterized by a different self-confidence should display a different avereage grade at university for instance. All regressions also control for the degree program the student is enrolled in, the type of high school attended, the region of residence and expected sector of employment although these coefficients are not reported. In column 1 we proxy for the family background by using dummies for whether at least one of the parent has a university degree and for whether at least one of the parents has only primary education. Estimated coefficients have the expected sign, but are not statistically significant. In the second column we use instead family income, proxied by the tuition category. The effect of family income is significant and J-shaped, with a minimum in the third category. At that time, there were 6 brackets, and more than 60% of the students in our sample were in the top three categories (with 35% of students in the top bracket). Results are almost unchanged when both measures of family background are included (column 3).

INSERT TABLE 4 ABOUT HERE

Wage expectations 10 years after graduation are probably a better proxy for self-confidence, since after such a spell of time wages should be expected to reflect productivity more precisely. The results, presented in Table 4, are very similar to those of Table 3 in terms of sign and magnitude, but the coefficients of family income are estimated more precisely. We observe the same J-shaped behaviour of tuition category that looks counterintuitive at first glance because it implies that students from poor families are more confident than those from the middle class (third income bracket). A possible explanation is that Bocconi is a very expensive university in Italy where rich families are over-represented. Students from poor families are instead under-represented because they could not afford the tuition fees without the financial help of social welfare institutions, which is awarded only if strict requirements in terms of academic performance are fulfilled. Therefore, the subsample of poor families suffers a stronger self-selection problem because only particularly good and strongly motivated students are enrolled.

Summarizing, Bocconi is an elite university in Italy widely known to attract very good students and well recognized in the labour market. Hence, one should expect that the signal provided by graduating at Bocconi is strong enough to more than counterbalance the effect of other differences in students' former endowments. In contrast, we find that the different socio-economic background still shapes wage expectations and the result is robust to the inclusion of a number of controls, the most obvious of which is ability as proxied by grades obtained at university. Hence, the same observed (and observable) signals have a different impact on different people because inherited beliefs about one's own ability survive a string of commonly-believed-to-be very good signals.

Of course, one could argue that students have an imperfect knowledge of the labour market, and that different wage expectations simply reflect different "reference wages" rather than self-confidence. If this is the case, however, wage realizations should then be correlated with the observable signals (e.g. graduation marks) rather than with the family background at least in the particular case of Bocconi graduates. Therefore, in the next section we check whether different expectations across family backgrounds find a counterpart in the actual wages of former Bocconi students.

4.3 Dataset of Bocconi Graduates

The data used in this section come from a survey of Bocconi graduates containing information very similar to that of Bocconi students decribed in the previous subsection. Respondents were asked a number of questions on their current and past working situation, among which current wages (from 4 to 16 years after graduation, depending on the cohort) and wages in the first job.[14]

INSERT TABLE 5 ABOUT HERE

INSERT TABLE 6 ABOUT HERE

The empirical specification is as similar as possible to that used in the previous subsection. Table 5 presents the results when we use the logarithm of wage in first job as dependent variable, while in table 6 current (log) wages are used. Parental education seems to have a significant effect in the baseline specifications (column 1 of tables 5 and 6). Statistical significance disappears when we add interactions with survey cohorts dummies, but coefficients do not change much in magnitude and keep having the expected sign. In columns 3 and of tables 5 and 6 we use, as a proxy for family background, log family income as declared by the students in their first year in Bocconi. By doing so, we loose the 1985 and 1989 cohorts because that piece of information was not available at the time. Furthermore, we loose most of the data for students at the top of the income distribution, since students paying the highest tuition fee were not requested to certify family income. Using family income, however, results are stronger. They do not depend on whether we use current wages or first-job wages as dependent variable, and they survive the inclusion of both the the Self-Employed and the Network dummies. The latter is particularly important because we expect network effects to be among the main channels through which parental background can have an impact on labour market outcomes over and above self-confidence. The minimum of the quadratic function roughly corresponds to the 2nd percentile of the distribution of family income, thereby confirming the punchline of the results about wage expectations and therefore their goodness as a proxy for self-confidence.

[14]We refer again to Filippin and Ichino (2005), especially for the details about how we obtain income measures from the income classes in the original database.

5 Conclusions

In line with some recent contributions in the literature, we claim that the socio-economic background affects not only the actual stock of cognitive skills possessed by a child, like ability, but also other non-cognitive skills. In particular, we focus on one non-cognitive skill, self-confidence, that we define as beliefs about one's ability.

Empirical evidence based on the PISA survey support such a claim, even when controlling for the PISA score itself, a proxy for cognitive ability unobserved by the student at the time of the survey. This proxy is even likely to bias downward the importance of self-confidence since it captures not only innate ability but also the gap in human capital that has been accumulated up to that point.

We then propose a model that describes the evolution of self-confidence based on prior beliefs updated at every period according to the success or failure in the endeavour undertaken. The probability of success depends on the true level of ability as well as on the difficulty of the task, which is chosen endogenously given one's self-confidence. We simulate the model with a bootstrapping procedure showing that choices distorted only by under-confidence lead to a significant gap in the accumulation of human capital during the learning process of the true level of ability.

Empirical evidence coming from a survey of students and graduates from Bocconi University confirms that indeed the learning pattern is a slow process because inherited beliefs about one's ability survives a long string of signals. Students from a poorer socio-economic background have significantly lower wage expectations even when (over)controlling for ability, and despite the strenght of the signal to be spent in the labour market they are going to acquire.

This evidence supports indirectly the implication of the model, because as long as the learning process does not convergence quickly to the true ability level even relatively small differences in confidence can determine different achievements of otherwise identical individuals. The level of human capital in the model diverges while the under-confident student eventually learns his true ability level, and this allows to explain why the early gap based on the socio-economic background does not narrow when the role of the family becomes less important during life. The model suggests as a policy implication that cognitive tests should take place as early as possible in order to minimize the correlation of the measure of ability with the family background.

References

BAUMEISTER, R., J. CAMPBELL, J. KRUEGER, AND K. VOHS (2003): "Does High Self-Esteem Cause Better Performance, Interpersonal Success, Happiness or Healthier Lifestyles?," *Psyc*, 4, 1–44.

BELZIL, C. (2007): "Subjective Beliefs and Schooling Decisions," Discussion Paper Series 2820, IZA.

BELZIL, C., AND M. LEONARDI (2007): "Can Risk Aversion Explain Schooling Attainments? Evidence from Italy," *Labour Economics*, 14(6), 957–970.

BLACK, S. E., AND P. J. DEVEREUX (2010): "Recent Developments in Intergenerational Mobility," Discussion Paper Series 4490, IZA.

BÉNABOU, R., AND J. TIROLE (2002): "Self-Confidence and Personal Motivation," *The Quarterly Journal of Economics*, 117(3), 871–891.

BOWLES, S., H. GINTIS, AND M. OSBORNE (2001): "The Determinants of Earnings: a Behavioral Approach," *Journal of Economic Literature*, XXXIX(4), 1137–1176.

CESARINI, D., M. JOHANNESSON, P. LICHTENSTEIN, AND B. WALLACE (2009): "Heritability of Overconfidence," *Journal of the European Economic Association*, 7(2-3), 617–627.

CHECCHI, D., AND L. FLABBI (2007): "Intergenerational Mobility and Schooling Decisions in Germany and Italy: The Impact of Secondary School Track," Discussion Paper Series 2876, IZA.

CHEVALIER, A., S. GIBBONS, A. THORPE, M. SNELL, AND S. HOSKINS (2009): "Students' Academic Self-Perception," *Economics of Education Review*, 28(6), 716–727.

CUNHA, F., AND J. J. HECKMAN (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.

——— (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4), 738–782.

CUNHA, F., J. J. HECKMAN, AND S. SCHENNACH (2010): "Estimating the Technology of Skill Formation," *Econometrica*, 78(3), 883–931.

DRAGO, F. (2008): "Self-Esteem and Earnings," IZA Discussion Papers 3577, Institute for the Study of Labor (IZA).

DUNNING, D., C. HEATH, AND J. SULS (2004): "Flawed Self-Assessment," *Psychological Science in the Public Interest*, 5, 69–106.

EDERER, F. (2010): "Feedback and Motivation in Dynamic Tournaments," *Journal of Economics & Management Strategy*, 19(3), 733–769.

FALK, A., D. HUFFMAN, AND U. SUNDE (2006a): "Do I Have What It Takes? Equilibrium Search with Type Uncertainty and Non-Participation," Discussion Paper Series 2531, IZA.

——— (2006b): "Self-Confidence and Search," Discussion Paper Series 2525, IZA.

FERLA, J., M. VALCKE, AND Y. CAI (2009): "Academic Self-Efficacy and Academic Self-Concept: Reconsidering Structural Relationships," *Learning and Individual Differences*, 19(4), 499–505.

FILIPPIN, A., AND A. ICHINO (2005): "Gender Wage Gap in Expectations and Realizations," *Labour Economics*, 12, 125–145.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3), 411–482.

HVIDE, H. K. (2002): "Pragmatic Beliefs and Overconfidence," *Journal of Economic Behavior & Organization*, 48, 15–28.

KÖSZEGI, B. (2006): "Ego Utiliy, Overconfidence and Task Choice," *Journal of the European Economic Association*, 4(4), 673–707.

KRAHMER, D. (2007): "Equilibrium Learning in Simple Contests," *Games and Economic Behavior*, 59, 105–131.

MARSH, H. W. (2005): "Big Fish Little Pond Effect on Academic Self-Concept," *German Journal of Educational Psychology*, 19, 119–128.

OECD (2009): "Top of the Class. High Performers in Science in PISA 2006 Programme," Oecd report.

SJÖGREN, A., AND S. SÄLLSTRÖM (2004): "Trapped, Delayed and Handicapped," Working Paper Series 613, Research Institute of Industrial Economics.

SQUINTANI, F. (2006): "Mistaken Self-Perception and Equilibrium," *Economic Theory*, 27, 615–641.

SULLIVAN, A. (2006): "Students as Rational Decision Makers: The Question of Beliefs and Attitudes," *London Review of Education*, 4, 271–290.

Tables

Table 1: Index of socio-economic status and other family background variables

	(1)	(2)	(3)	(4)
Highest parental occupational status	0.0441***			
	[0.000270]			
Highest educational level of parents		0.502***		
		[0.00350]		
Index of home possessions			0.705***	
			[0.00482]	
Cultural possessions at home				0.419***
				[0.00704]
Observations	223480	226277	229458	227775
R^2	0.686	0.607	0.544	0.248

BRR Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All regressions include a constant and country dummies

Table 2: Results: Science Self-Efficacy

	(1) Baseline	(2) Pisa score	(3) Effort	(4) Motivation	(5) Parents
Index of socio-ec. status	0.318***	0.145***	0.111***	0.118***	0.119***
	[0.0144]	[0.0141]	[0.0148]	[0.0106]	[0.0262]
Index of socio-ec. status^2	0.0325***	0.0223*	0.0298**	0.0141	0.0263
	[0.00886]	[0.00864]	[0.00849]	[0.00834]	[0.0138]
Female	-0.157***	-0.141***	-0.157***	-0.103***	-0.0312
	[0.0123]	[0.0109]	[0.0114]	[0.0102]	[0.0173]
PISA score in Science		0.00401***	0.00397***	0.00291***	0.00267***
		[0.000124]	[0.000111]	[0.0000937]	[0.000139]
Out of school - Science Q31b			0.112***		
			[0.00803]		
Self study - Science Q31c			0.114***		
			[0.00605]		
Interest in learning science				0.240***	0.226***
				[0.00799]	[0.0116]
Personal value of science				0.242***	0.222***
				[0.00659]	[0.0129]
Parents' value of science					0.000227
					[0.0128]
Par.'s sc. career motivation					-0.0285**
					[0.00901]
Science activities at age 10					0.0617***
					[0.00829]
Tracked educational system	-0.194***	-0.0844***	0.00533	-0.0450**	.
	[0.0206]	[0.0206]	[0.0207]	[0.0164]	.
Tracking*Socio-ec. status	-0.0737***	-0.0565**	-0.0395*	-0.0654***	-0.0569*
	[0.0151]	[0.0171]	[0.0175]	[0.0121]	[0.0277]
School-level characteristics	NO	NO	NO	NO	YES
Observations	225098	225098	216304	223912	29970
R^2	0.119	0.230	0.255	0.375	0.355

BRR Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Expected Wage 1 Year After Graduation

	(1) Parental Ed.	(2) Income	(3) Income squared	(4) Full
Parent graduate	0.00478			-0.000730
	[0.0307]			[0.0320]
Parent primary ed.	-0.0734			-0.0801
	[0.0579]			[0.0580]
Income bracket		0.00799	-0.0821	-0.0880*
		[0.00877]	[0.0431]	[0.0435]
Income bracket^2			0.0119*	0.0125*
			[0.00559]	[0.00560]
Female	-0.0557	-0.0576*	-0.0572*	-0.0569
	[0.0291]	[0.0290]	[0.0290]	[0.0291]
Family firm	0.213***	0.205***	0.199***	0.198***
	[0.0575]	[0.0586]	[0.0586]	[0.0586]
Average grade	0.0197*	0.0195*	0.0183*	0.0183*
	[0.00837]	[0.00837]	[0.00837]	[0.00837]
High School grade	-0.256	-0.246	-0.232	-0.229
	[0.201]	[0.201]	[0.201]	[0.201]
R^2	0.075	0.073	0.079	0.081
Observations	764	764	764	764

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All regressions include dummies for degree program, type of high school, region of residence and expected sector of employment.

Table 4: Expected Wage 10 Years After Graduation

	(1) Parental Ed.	(2) Income	(3) Income squared	(4) Full
Parent graduate	0.0556			0.0298
	[0.0304]			[0.0314]
Parent primary ed.	0.000837			-0.00627
	[0.0575]			[0.0568]
Income bracket		0.0302^{***}	-0.138^{**}	-0.142^{***}
		[0.00865]	[0.0422]	[0.0426]
Income bracket^2			0.0222^{***}	0.0225^{***}
			[0.00547]	[0.00549]
Female	-0.0846^{**}	-0.0945^{**}	-0.0937^{***}	-0.0910^{**}
	[0.0289]	[0.0286]	[0.0283]	[0.0285]
Family firm	0.212^{***}	0.174^{**}	0.164^{**}	0.165^{**}
	[0.0570]	[0.0578]	[0.0573]	[0.0574]
Average grade	0.0220^{**}	0.0215^{**}	0.0192^{*}	0.0192^{*}
	[0.00830]	[0.00825]	[0.00818]	[0.00819]
High School grade	-0.395^{*}	-0.342	-0.316	-0.319
	[0.199]	[0.198]	[0.197]	[0.197]
R^2	0.117	0.127	0.146	0.147
Observations	764	764	764	764

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All regressions include dummies for degree program, type of high school, region of residence and expected sector of employment.

Table 5: Realized Wage at First Job

	(1) Baseline Par. Edu.	(2) Full Par. Edu.	(3) Baseline Fam. Income	(4) Full Fam. Income
Parent graduate	0.0556**	0.0104	-	-
	[0.0169]	[0.0510]	-	-
Parent primary ed.	-0.0531*	-0.0850	-	-
	[0.0242]	[0.0704]	-	-
Log Family Income	-	-	-0.0989**	-0.0998**
	-	-	[0.0340]	[0.0338]
Log Family Income ^{2}	-	-	0.00655**	0.00665**
	-	-	[0.00214]	[0.00213]
Female	-0.0600***	-0.0607***	-0.0649**	-0.0619**
	[0.0161]	[0.0162]	[0.0227]	[0.0225]
High School grade	-0.171*	-0.172*	-0.133	-0.140
	[0.0828]	[0.0828]	[0.133]	[0.132]
Avg. Univ. grade	0.00576	0.00553	0.000119	-0.00342
	[0.0104]	[0.0104]	[0.0158]	[0.0158]
Top 25% graduate	0.0925***	0.0928***	0.113**	0.118**
	[0.0270]	[0.0270]	[0.0374]	[0.0372]
Network		-0.0105		0.00915
		[0.0181]		[0.0287]
Self-Employed		-0.0572**		-0.100***
		[0.0184]		[0.0260]
Survey cohort dummies	YES	YES	YES	YES
Interactions cohort*parental education	NO	YES	-	-
R^2	0.147	0.153	0.082	0.097
Observations	2096	2096	969	969

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All regressions include dummies for degree program, type of high school and region of residence

Table 6: Current Wage - Parental Education

	(1) Baseline Par. Edu.	(2) Full Par. Edu.	(3) Baseline Fam Income	(4) Full Fam Income
Parent graduate	0.0962***	0.0119	-	-
	[0.0214]	[0.0647]	-	-
Parent primary ed.	-0.0315	-0.0485	-	-
	[0.0306]	[0.0891]	-	-
Log Family Income	-	-	-0.177***	-0.177***
	-	-	[0.0389]	[0.0389]
Log Family Income ^{2}			0.0110***	0.0110***
			[0.00245]	[0.00245]
Female	-0.280***	-0.280***	-0.149***	-0.149***
	[0.0204]	[0.0204]	[0.0259]	[0.0260]
High School grade	-0.297**	-0.277**	-0.388*	-0.380*
	[0.105]	[0.105]	[0.152]	[0.152]
Avg. Univ. grade	0.0325*	0.0332*	0.0152	0.0147
	[0.0131]	[0.0131]	[0.0180]	[0.0181]
Top 25% graduate	0.0677*	0.0671	0.136**	0.137**
	[0.0342]	[0.0342]	[0.0428]	[0.0428]
Network		-0.00101		-0.0229
		[0.0262]		[0.0374]
Self-Employed		0.0850***		0.0347
		[0.0242]		[0.0390]
Survey cohort dummies	YES	YES	YES	YES
Interactions cohort*parental education	NO	YES	-	-
R^2	0.255	0.260	0.178	0.178
Observations	2096	2096	969	969

Standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All regressions include dummies for degree program, type of high school and region of residence.