View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

DISCUSSION PAPER SERIES

IZA DP No. 6117

Family Background, Self-Confidence and Economic Outcomes

Antonio Filippin
Marco Paccagnella

November 2011

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Family Background, Self-Confidence and Economic Outcomes

Antonio Filippin
University of Milan
and IZA

Marco Paccagnella
Bank of Italy

Discussion Paper No. 6117
November 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 6117
November 2011

ABSTRACT

Family Background, Self-Confidence and Economic Outcomes*

In this paper we analyze the role played by self-confidence, modeled as beliefs about one's ability, in shaping task choices. We propose a model in which fully rational agents exploit all the available information to update their beliefs using Bayes' rule, eventually learning their true type. We show that when the learning process does not convergence quickly to the true ability level, even small differences in initial confidence can result in diverging patterns of human capital accumulation between otherwise identical individuals. As long as inital differences in the level of self-confidence are correlated with the socioeconomic background (as a large body of empirical evidence suggests), self-confidence turns out to be a channel through which education and earnings inequalities are transmitted across generations. Our theory suggests that cognitive tests should take place as early as possible, in order to avoid that systematic differences in self-confidence among equally talented people lead to the emergence of gaps in the accumulation of human capital.

JEL Classification: D83, J24, J62

Keywords: self-confidence, family background

Corresponding author:

Antonio Filippin
University of Milan
Department of Economics and Business
via Consevatorio, 7
20122 Milan
Italy
E-mail: antonio.filippin@unimi.it

* We would like to thank Marcel Jansen, Marco Leonardi, Michele Pellizzari, Jan Van Ours and seminar participants at Bocconi University, Universitat Autonoma de Madrid, Tilburg University, ASSET, EALE, ESSLE 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

Gaps in economic outcomes such as educational attainments and earnings tend to persist across generations and it is well-known that the socio-economic status of the parents is usually a very good predictor of the outcomes of their offspring.

Bowles, Gintis, and Osborne (2001) stress that "the advantages of the children of successful parents go considerably beyond the benefits of superior education, the inheritance of wealth, or the genetic inheritance of cognitive ability." They propose additional variables comparable to what now goes under the label of "non-cognitive skills" as factors that can supplement the ohterwise low explanatory power of the traditional variables used to fit the variance of earnings[1]. Moreover, they claim that the contribution of parental socio-economic status to earnings is in part determined by such non-cognitive skills, genetically transmitted or learned from parents that act as role models.

Since then many other authors have emphasized the role played by non-cognitive skills in explaining economic success and gaps in attainments. The current literature on the economic relevance of non-cognitive skills tends to treat these measures as inputs that enter the "black-box" of the skill production function. Cunha and Heckman (2007) propose a particular formulation of the technology of skill formation featuring self-productivity and dynamic complementarities among a multidimensional vector of cognitive and non-cognitive skills. They argue that insufficient investment in some of these skills early in life has long-lasting consequences that are very difficult or costly to revert. Heckman, Stixrud, and Urzua (2006), Cunha and Heckman (2007, 2008) and Cunha, Heckman, and Schennach (2010) have shown that gaps between children from different backgrounds open up very early in life, as soon as in pre-school age, and then tend to persist and stay roughly constant over the lifetime. Note how this finding clearly locates the rising of the problem in a period in which the role of the parents is the most important. Other recent papers that have focused on assessing the economic returns of different types of inputs (cognitive vs. non-cognitive skills) include Heineck and Anger (2010) and Lindqvist and Vestman (2011).

In this paper we want to analyze the role possibly played by a single non-cognitive skill, namely self-confidence, defined as the beliefs over one's unknown level of cognitive ability. Hence, our model entails the simplest possible multidimensional vector of skills, containing only two elements: a cognitive skill (innate ability) and a non-cognitive one (self-confidence). The use of such a framework is neither meant to deny the importance of other skills, nor the well-established fact that cognitive and non-cognitive abilities are multidimensional in nature, nor to downplay the significance of the interaction among them. It simply reflects our goal to

[1]In general, non-cognitive skills are defined as "personality traits that are weakly correlated with measures of intelligence" (Brunello and Schlotter, 2011). In this broad concept economists have usually investigated the so-called "Big Five" factors, following Nyhus and Pons (2005): agreeableness, conscentiousness, emotional stability, extraversion and autonomy. Other commonly used measures include the locus of control (Caraloc) and the Lawseq self-esteem score, but also attitudes toward risk and educational aspirations and expectations.

isolate and highlight a very specific mechanism, i.e. the role that a wrong self-confidence plays through the distortion of task choices. In other words, our purpose is to go into the "black-box" of the skill production function, identifying a precise and specific channel through which inherited differences in self-confidence can endogenously (i.e., through individual choices) explain the emergence and persistence of gaps in the accumulation of human capital.

The working idea of our model is that by acting as role-models, parents transmit to their children beliefs about their (unknown) ability. Such beliefs affect educational and task choices and, through this channel, contribute to widen the gap in the accumulation of human capital while the learning process (of acutal ability) takes place. The consequences of initially "wrong" beliefs can thus have long lasting effects, even if agents eventually learn their true level of ability.

An advantage of our approach is that the single non-cognitive skill we study has a clear and simple economic interpretation, and that we make transparent the channel through which it affects the accumulation of human capital (and thus, indirectly, earnings).

For self-confidence to have important effects we do not need to assume that agents enjoy holding a good image of themselves (i.e. that self-confidence enters directly the utility function), something that would imply that some degree of overconfidence is optimal[2]. Our theoretical framework assumes full rationality, given that agents extract all the available information from the signals received in order to update their beliefs, and this implies that they eventually learn their true type. Similarly, we exclude any other form of self-deception. The Bayesian learning mechanism is based on observing success or failure in the endeavour undertaken, given that the probability of success depends on the true level of ability as well as on the difficulty of the task, which is chosen endogenously in accordance with (updated) beliefs about one's ability.

We finally 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. As long as it correlates with the family background, self-confidence constitutes therefore a channel through which gaps in educational attainments and earnings perpetuate across generations.

This finding also helps to explain why the early gaps based on the socio-economic background do not narrow when the role of the family becomes less important during life, and it suggests that policies aimed at providing early and accurate feedbacks on the cognitive skills of disadvantaged children can be important in promoting inter-generational income mobility.

The outline of the paper is as follows. In section 2 we survey the relevant literature comparing our theoretical approach with others in the literature. We also provide evidence sopporting both the important role played by self-confidence and the correlation between self-

[2] Such an assumption is quite common in the behavioral economics literature (eg. in Köszegi, 2006 and Weinberg, 2009). We discuss this issue in more details in section 2.3.

confidence and family background. In Section 3 we present a simple and parsimonious theoretical model that highlights how self-confidence can affect the accumulation of human capital via task choice. In section 4 we present the results of a simple simulation in order to better assess the implications of our model in terms of the emergenge of gaps in educational attainments between people from different backgrounds. Section 5 comments upon our results and draws some conclusions.

2 Motivation

In this section we survey the related theoretical and empirical literature to motivate the relevance of our work. In subsection 2.1 we document how incorrect beliefs about ability are very common, and how this fact can have important practical consequences. In subsection 2.2 we discuss the different definitions and interpretations of confidence that have been used in the literature and why we have chosen to focus on a definition based on the *levels* of beliefs, rather than on their *precision*. In subsection 2.3 we discuss an assumption commonly made in the literature, i.e. that people actually care about their beliefs, and explain why decided not to make this assumption. Finally, in subsection 2.4 we justify one of the main assumption of our model by providing evidence that suggests the existence of a relevant link between self-confidence and the socio-economic background.

2.1 Imperfect knowledge of one's ability

We define self-confidence as the beliefs an agent holds about his own ability, following Bénabou and Tirole (2002), Hvide (2002), Köszegi (2006), Sjögren and Sällström (2004), and Weinberg (2009), among the others. This derives from the assumption that ability is unknown to the agent, instead of being his private information as in standard signaling models.

There is an extensive literature showing that agents hold a rough estimate of their cognitive skills. For instance, Dunning, Heath, and Suls (2004) survey the psychological literature documenting the presence of a weak correlation between actual and perceived performance in several domains, while Falk, Huffman, and Sunde (2006) provide experimental evidence that people are substantially uncertain about their relative ability and that this have indeed important consequences on search decisions.

Indirect evidence supporting the idea that individuals have imperfect knowledge about one's own ability can be inferred by observing that people with similar observable characteristics make different choices. In figure 1 we show that there is a considerable degree of overlapping in the distribution of PISA 2006 test scores across people enrolled in different high-school tracks (which are very likely to yield very different returns on the labor market)[3].

[3] A similar figure appears in Checchi and Flabbi (2007).

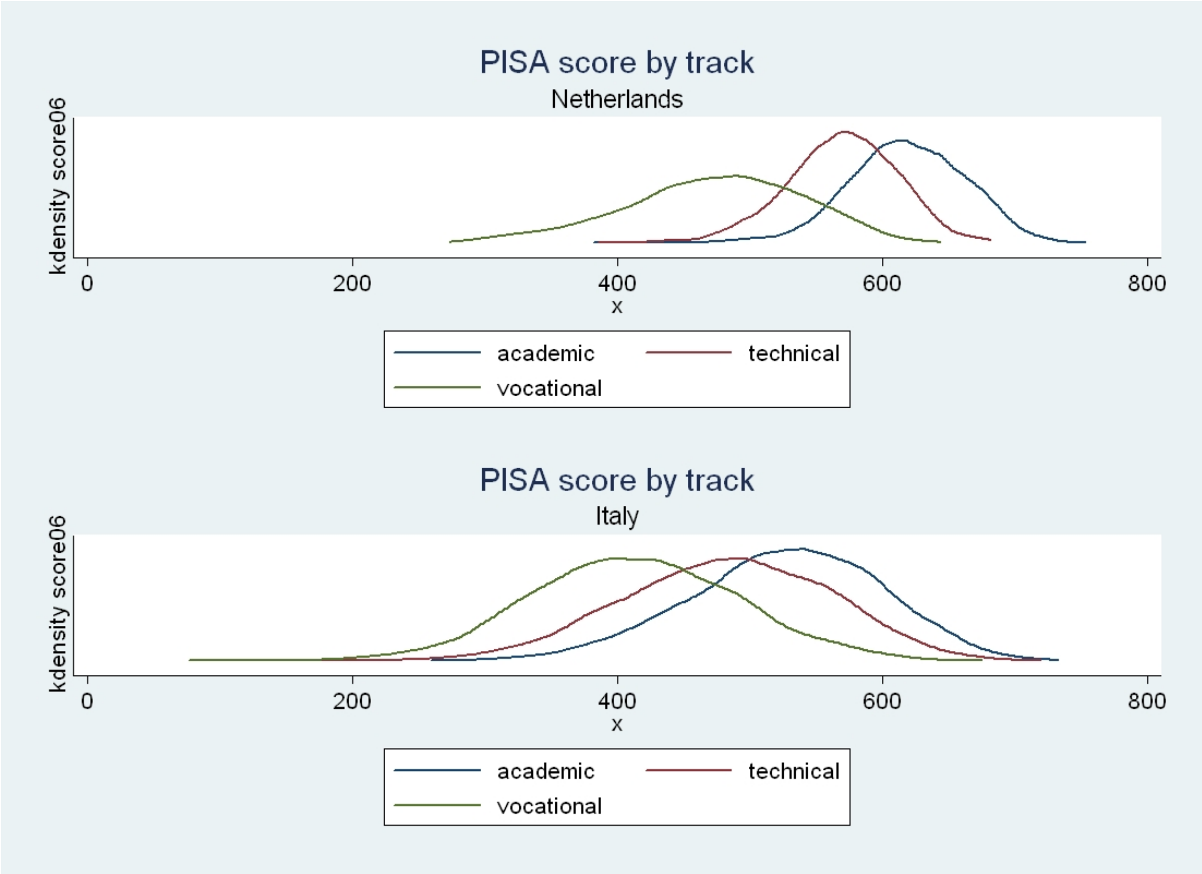

Figure 1: Netherlands (top), Italy (bottom)

Furthermore, comparing the top and the bottom panel (which refer respectively to the Netherlands and to Italy) we see quite a different degree of overlapping in the two countries. The main difference between the two educational systems is that while in Italy students and parents are perfectly free to choose the high-school track[4], in the Netherlands there is an aptitude test[5] administered at age 12 that, although not mandatory, has gained a considerable influence in recommending the secondary track most suitable for the pupil. While we recognize that there might be many other reasons to choose different tracks, the level of ability should also play an important role. Figure 1 shows that the degree of overlapping across tracks, which should in part be driven by ability mismatch, is much more pronounced in Italy where the uncertainty about one's ability is higher. There is also ample evidence (eg in Giuliano, 2008) that the socio-economic background has a strong influence on the choice of high-school track. These two stylized facts, examined toghether, suggests that systematic differences not related to cognitive ability do influence educational choices.

Recent related literature has also investigated the role played by subjective returns to education in determining educational choices in developing countries. Jensen (2010) finds that,

[4]Although it is common to receive suggestions from school teachers.
[5]The so-called Cito test.

in the Dominican Republic, perceived returns to secondary school are extremely low, despite high measured returns, and that providing information about measured returns significantly increased enrollment rates; Nguyen (2008) finds instead that providing information on the returns to schooling improves school performance in Madagascar. Attanasio and Kaufmann (2009) and Kaufmann (2010) show that, in Mexico, subjective earning expectations and risk perceptions are important determinants of college attendance choices. However, they argue that,in that particular context, an important determinant of differential enrollment rates between poor and rich students is the presence of credit constraints. Imperfect information about ability (affected by different levels of confidence) is conceptually very similar to imperfect estimation of the market returns to education (due to incomplete information or to differences in the information set), and our model could easily be reinterpreted in this term. However, although being observationally equivalent, the two competing explanations can have very different consequences in terms of policy prescriptions.

2.2 Level vs. precision of beliefs

Another important issue is the definition of confidence in terms of the mean vs. the spread of the distribution of beliefs[6]. The former implies that an overconfident holds too high an estimate of his ability. The latter refers to an evaluation that is too precise, and better fits for instance an investment decision about which an agent can underestimate the variance of the future return. Confidence in terms of precision of beliefs could be adapted to a framework in which one's ability is the random variable, although it would be meaningless to talk about overconfidence as long as the true level of ability is a point estimate.

The two concepts, however, are not correlated, and their interaction can also determine counterintuitive results. For instance, it may happen that an agent that is quite confident along both dimensions could think to have a lower probability of success than another who is totally agnostic about his ability. To avoid such a possibility we assume that the probability of success is linear in ability, and we prefer to adopt a notion of confidence that refers to the level instead of the precision of beliefs. In our framework beliefs are defined as a probability distribution over the whole support of ability, so the second moment enters the picture but it can only affect the speed of convergence.

The focus on the level rather than on the precision of beliefs is one of the main difference between our model and the one proposed by Sjögren and Sällström (2004) (who also describe the endogenous evolution of self-confidence for rational agents that choose tasks and update their beliefs in a Bayesian fashion after observing the outcomes of their choices). A second major difference is that, in our framework, agents eventually discover their true type, while Sjögren and Sällström (2004) show how people can remain "trapped" with wrong beliefs due

[6]Both definitions are used in the literature. The first for instance by Hvide (2002), Bénabou and Tirole (2002), and Weinberg (2009), the second by Sjögren and Sällström (2004), while Köszegi (2006) and Belzil (2007) use both.

to insufficient experimentation and learning[7].

The probability of success that characterizes our model implies that our agents make decisions under uncertainty; our work can thus be linked to the literature that sees education as a risky investment and that investigates the role played by risk aversion. Belzil (2007) estimates both the degree of over and under estimation of labor market skills and the dispersion of the distribution of subjective beliefs. He finds evidence of frequent (but moderate) over-estimation and cases of severe under-estimation (particularly among the most able individuals). He also finds that only 25% of unobserved ability heterogeneity is perceived by the individuals as ex-ante risk and that 36% of the population act on the basis of a degenerate subjective ability distribution. 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. Since the effect of underconfidence can be confounded with that of risk aversion, in order to isolate the role of confidence we assume that agents are risk neutral.

2.3 Self-confidence in the utility function

There is a wealth of contributions in the psychological literature showing that confidence affects the task choice (see Weinberg (2009) and literature therein). In this paper we also focus on the role that confidence plays through this channel, and more specifically on how task choice shapes human capital acquisition. In our model confidence affects utility only indirectly through the choice of task, which determines how much human capital the agent gets, and there is no direct influence like there would be in case the agent enjoys thinking that his ability is high. Examples of models in which beliefs about one's ability enter directly the utility function are Weinberg (2009) and Köszegi (2006). While such models rationalize many interesting features of human behaviour (along with the result that moderate levels of overconfidence turn out to be optimal), we decide to stick to a simpler theoretical framework in which this does not happen. The main reason is that once agents are supposed to enjoy holding a good self-image, they should also be capable of tailoring the information acquisition during their learning process in such a way to preserve it, for instance by means of beliefs that are "pragmatic" (Hvide, 2002) or more generally self-serving, as well as with selective memory (Bénabou and Tirole, 2002)[8].

Manipulating the information acquisition can only be effective *in the short run*, unless agents end up stuck in a self-confirming equilibrium in which their learning process reaches a fixed point although their beliefs are wrong. In other words, beliefs are wrong but never disconfirmed by the evidence either because further experimenation is not available or because

[7]To achieve this result, they have to assume that there are non-informative task, in which the probability of succes is equal to one.

[8]Bénabou and Tirole (2002) also assume that discount rates are lower at shorter horizons than at more distant ones (time-inconsistency). Belzil (2007), however, find a predominance of the future component of intertemporal utility over the present component in schooling decision, and interpret it as evidence supportive of the standard time-consistent model.

agents continue to indefinitely self-deceive themselves[9]. Although such an outcome cannot be exclueded, we find more interesting to analyze the effect of holding a wrong self-image when the true type is eventually learned. Including beliefs in the utility function would only incentivate some form of self-deception that, even allowing the manipulation of information acquisition, would only have the transitory effect of slowing down the learning process, and therefore we prefer to avoid such a complication. Our model thus adheres to a perfectly rational framework, with agents characterized by standard preferences and that unbiasedly exploit the whole information available.

2.4 The inter-generational transmission of confidence

Recent empirical findings provide support for one of the key assumptions of our model, namely that self-confidence is correlated with the family background. Cesarini, Johannesson, Lichtenstein, and Wallace (2009), using Swedish data on a sample of twins and defining overconfidence as the difference between the perceived and actual rank in cognitive ability, argue that genetic differences explain 16-34% of the variation in overconfidence, and that common environmental differences explain an additional 5-11%. A series of studies on different longitudinal UK datasets (collected in Goodman and Gregg, 2010) find a strong intergenerational correlation not only in cognitive skills, but also in a variety of attitudes that can be considered proxies of confidence. In particular, Gregg and Washbrook (2011) find that, even after controlling for long-run family background factors and prior attainment, children are more likely to perform well in tests at age 11 if they have strong beliefs in their own ability and have a more internal locus of control[10], and they also find that children from poorer families are less likely to have these attributes. Chowdry, Crawford, and Goodman (2011) find that richer parents have higher expectations of their children's educational attainments and that young people from poorer families have lower ability beliefs, a more external locus of control and lower educational aspirations and expectations. After controlling for attainment at age 11, 15% of the socio-economic gap in attainment at age 16 is accounted for by child attitudes, and an additional 12% is accounted for by parental attitudes. Chevalier, Gibbons, Thorpe, Snell, and Hoskins (2009) find that working class undergraduates underestimate their performance relative to others, but also that working class secondary school pupils have greater confidence and a more positive self-evaluation of their math ability. This finding may be due to differences in peer-groups and to the "big fish, small pond effect". Here we provide additional evidence about the link between socio-economic background and self-confidence using data from the OECD-PISA study. This dataset contains what we believe is a good proxy for self-confidence, namely "Science Self-

[9]Models in Köszegi (2006) and Weinberg (2009) are characterized by a small number of periods. Hvide (2002) justifies pragmatic beliefs in the long run with a thought experiment in which "the agent takes into account what pays rather than what is true."

[10]People with an external locus of control tend to think that luck or fate, rather than their own actions, are what matters in life. It is likely that this is related to low levels of confidence in own ability.

Efficacy", an index built from student's answers to questions about the ease with which they believe they could perform eight science-related tasks. This variable is a good proxy for beliefs about academic ability because 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 student's abilities and performance" (OECD, 2009)[11].

We thus regress our measure of confidence on family background, adding controls at the individual, school and family level; results are presented in Table 1.

The relationship between Self-efficacy and family background is significant and positive as expected, displaying a convex correlation. 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. The inclusion of PISA score captures some variance of self-efficacy, but the positive relationship with family background remains strong. Notice that controlling for the PISA score is likely to bias downward the role played by self-confidence, because if our model is correct the PISA score already encompasses the gap in the human capital accumulated up to that point also because of a different self-confidence. In other words, two students with the same innate ability but characterized by a different self-confidence should also display a different PISA score.

Adding further controls at the student level (column 3) and at the parent and school level (column 4) does not change significantly the results, which we interpret as suggestive evidence that family background has a direct impact on self-confidence, over and above the one operating through the transmission of cognitive skills.

The PISA dataset has the advantage of being a large-scale, international, representative sample, but includes extremely heterogeneous students at an early stage of their education career. Therefore, we replicate a similar analysis using another dataset with opposite characteristics, coming from a survey of a much more homogeneous population at later stage of their academic career. This dataset has been collected by circulating a questionnaire to all second 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[12]. We use expected wage as a proxy for self-confidence, while family-background is proxied by parents' educational levels and the students' tuition category (a function of family income). Wage expectations 10 year after graduation are probably the best measure of self-confidence, since after

[11] 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 does not ask students to compare their ability to that of others, we believe it is a better proxy for the notion of confidence that we use in the model of Section 3.

[12] The same data are used in Filippin and Ichino (2005), to which we refer for further details on the characteristics of the dataset.

Table 1: Results: Science Self-Efficacy

	(1) Baseline	(2) Pisa score	(3) Effort	(4) Parents
Index of socio-ec. status	0.318***	0.145***	0.111***	0.119***
	[0.014]	[0.014]	[0.015]	[0.026]
Index of socio-ec. status^2	0.033***	0.022*	0.030**	0.026
	[0.009]	[0.009]	[0.009]	[0.014]
Female	-0.157***	-0.141***	-0.157***	-0.031
	[0.012]	[0.011]	[0.011]	[0.017]
PISA score in Science		0.004***	0.004***	0.003***
		[0.000]	[0.000]	[0.000]
Out of school - Science			0.112***	
			[0.009]	
Self study - Science			0.114***	
			[0.006]	
Interest in learning science				0.226***
				[0.012]
Personal value of science				0.222***
				[0.013]
Parents' value of science				0.001
				[0.013]
Science career motivation				-0.029**
				[0.009]
Science activities at age 10				0.062***
				[0.008]
School-level characteristics	NO	NO	NO	YES
R^2	0.119	0.230	0.255	0.355
Observations	225,098	225,098	216,304	29,970

BRR standard errors in brackets. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.
All regressions include country dummies and control for immigrant status, tracking and the interaction between tracking and the socio-economic status. In column 4 we also control for school-level variables like school size, student-teacher ratio, ability sorting and a dummy for public schools.

such a spell of time wages should be expected to reflect productivity more precisely[13]. Notice that also in this case the proxies for ability are likely to bias downward the role played by self-confidence, since they also control for the gap in the human capital accumulated up to that point. Table 2 reports results from regressing the log of expected wage ten years after graduation on family background variables and individual controls. While parental education does not seem to have a significant impact on expected wages, the effect of family income (proxied by tuition category) is significant and J-shaped, with a minimum in the third category[14]. Results are almost unchanged when both measures of family background are included.

Table 2: Expected Wage 10 Years After Graduation

	(1) Parental Ed.	(2) Income	(3) Income squared	(4) Full
Parent graduate	0.056			0.030
	[0.030]			[0.031]
Parent primary ed.	0.001			-0.006
	[0.057]			[0.057]
Income bracket		0.030***	-0.138**	-0.142***
		[0.009]	[0.042]	[0.043]
Income bracket^2			0.022***	0.023***
			[0.005]	[0.005]
Female	-0.085**	-0.094**	-0.094***	-0.091**
	[0.029]	[0.029]	[0.028]	[0.028]
Family firm	0.212***	0.174**	0.164**	0.165**
	[0.057]	[0.058]	[0.057]	[0.057]
Average grade	0.022**	0.021**	0.019*	0.019*
	[0.008]	[0.008]	[0.008]	[0.008]
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. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.
All regressions include dummies for degree program, type of high school, region of residence and expected sector of employment.

[13] For the sake of brevity we only report results using expected wages 10 years after graduation. Results using short-term expectations are not significantly different, and are available upon request.

[14] 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). Our results imply that students in the lowest income category are more confident than those from the middle class (third income bracket). A possible explanation is that Bocconi is a very expensive university where rich families are over-represented. Students from poor families are instead under-represented because they could not afford the tuition fees without financial help, which is awarded only if strict requirements in terms of academic performance are fulfilled. Therefore, the subsample of students in lower income brackets is likely to suffer a stronger self-selection problem because only particularly good and strongly motivated students are able to enroll.

Bocconi is recognized as 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 any other difference in students' former endowments. In contrast, we find that the different socio-economic background still shapes wage expectations. Hence, the same observed (and observable) signals have a different impact on different people. Our interpretations is that inherited beliefs about one's own ability survive a string of commonly-believed-to-be very good signals. Unfortunately, we cannot attribute a causal interpretation to this result, because such a correlation could be a spurious spillover of different networking abilities or different preferences correlated to the family background. However, the same correlation appears in the wage realizations of a similar but richer survey of Bocconi graduates in which a larger set of controls is available[15]. Moreover, our results are similar to what has been recently found by Delaney, Harmon, and Redmond (2011), who use a dataset collected from seven Irish universities (and thus certainly more representative of the population of Irish undergraduate students), and that also include many different measures of non-cognitive skills such as risk attitudes, time preferences and personality traits.

3 The Model

In this section we present in more details a multi-period model based on the assumptions already discussed in sections 2.1-2.4, in which agents choose a task on the basis of their beliefs, which are updated in a Bayesian manner after observing the outcome of every choice. Our purpose is to highlight the role played by confidence in explaining educational attainments via task choice.

As already explained, we assume that children do not 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 underestimates her ability: $\hat{\mu}\,(a) < a$. Similarly, the overconfident is characterized by $\hat{\mu}(a) > a$. Students make educational choices by choosing "tracks" (ψ). 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 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 drops out or must repeat a grade) 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 done had she been less ambitious.

We assume that the probability of success is given by

[15] Results are not displayed to save space but they are available upon request.

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

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

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 $\mu(a)$, posterior beliefs after receiving the signal implicit in the outcome $o = \{s; f\}$ are equal to:

$$\mu(a|o) = \frac{p(o)\,\mu(a)}{\int p(o)\,\mu(a)\,da.} \tag{2}$$

Successful outcome (s) in the track chosen allows agents to add human capital $k(\psi|s)$ to working life productivity, and agents maximize their instantaneous utility by choosing the track that optimally balances their expected acquisition of human capital with a convex cost of acquiring it $U\left[p(s)\,k(\psi) - \psi^2\right]$, given their confidence about unobserved ability.

If the track chosen is totally uninformative (e.g. $p(s) = 1$) the student does not gather evidence that contradicts his/her wrong beliefs. For instance, 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 accumulation of human capital when agents do 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, as anticipated in section 2.2, we 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 for instance with respect to the model in Sjögren and Sällström (2004), who assume that the probability of succesfully acquiring skills of type c_1 is $p(s) = a^{c_1}$, where $a \in [0, 1]$ 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 or less attractive depending on whether the probability of success is convex or concave in ability. 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 an agent characterized by quite a precise belief of being above the average. In contrast, we choose to remove such discontinuities by assuming linearity in ability in equation (1) and to focus on the effect of the level of confidence[16]. Hence, we assume the following functional form of the probability of success:

[16]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.

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

This specification implies that the importance of ability is proportional to the difficulty of the track. Notice that for the probability of success to be properly defined we need ability to have a finite support, and for the sake of simplicity we assume both $a \in [0, 1]$ and $\psi \in (0, 1]$. 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 2).

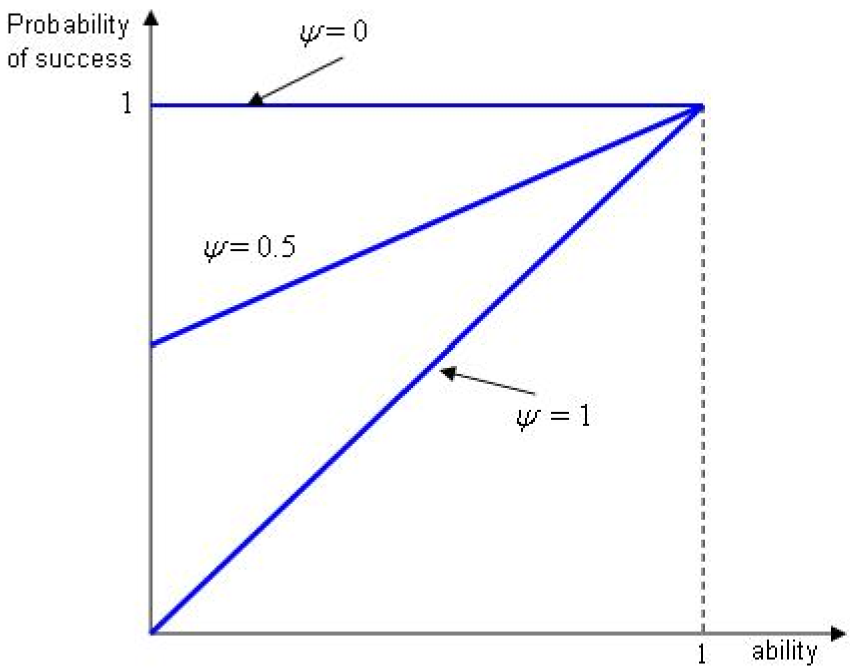

Figure 2: Different tracks in terms of importance of ability

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 ability mismatch. $\hat{a}$ represents the optimal level of ability for that track, i.e. the level of ability that characterizes the student that maximizes her utility by choosing exactly that track. 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 for the same reasons outlined in section 2.3. Of course, the effect of the mistake in evaluating 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 $m > 0$). To complete the picture, we assume that a failure leaves the stock of human capital unchanged, i.e. $k(\psi, \mu(a), a|f) = 0$[17].

Students are free to self-select into different tracks given the best estimate of their ability, trading off a lower human capital in case of success with a higher probability of acquiring it. If ability was known, the first-order conditions would imply[18]:

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

Given that $f(m)$ implies 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 information. Equation 5 therefore implies that both under- and over-confidence determine a suboptimal track choice and a loss of utility due to the mismatch $\mu(a) \neq \hat{a}$.

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:

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

[17]This assumption is made without loss of generality as compared to the case in which the human capital accumulated in case of failure is positive but strictly lower: $k(\psi, \mu(a), a|s) > k(\psi, \mu(a), a|f)$.

[18]To analyze the role played by self-confidence in shaping the gap in educational attainments when agents are eventually learning their true level of ability 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 24 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.

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 that a small degree 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[19]. As overconfidence increases, the sign of $\delta E(k)/\delta m$ depends on the magnitude of 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 that this event happens. In contrast, if the effect of overevaluating one's ability increases substantially 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.

Agents update their beliefs given the signal received (success or failure) at the end of each period[20]. 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, necessary to ensure 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, something necessary because with a Bayesian learning process agents can never assign a positive probability to events excluded by the prior.

The density function of the $Beta\,[\alpha, \beta]$ 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:

[19]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.

[20]Note that was the agent receiving a perfectly informative signal like the exact amount of human capital acquired when successful he could invert $k\,(\psi, \mu(a), a|s)$ deriving with certainty her true ability level. However, 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.

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

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 $\beta > \alpha > 1$[21]. 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 $Beta\,[2.5, 2.5]$, 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 $\mu\,(a) \sim Beta\,[2.5, 2.5]$ 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:

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

By contrast, if a failure was observed:

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

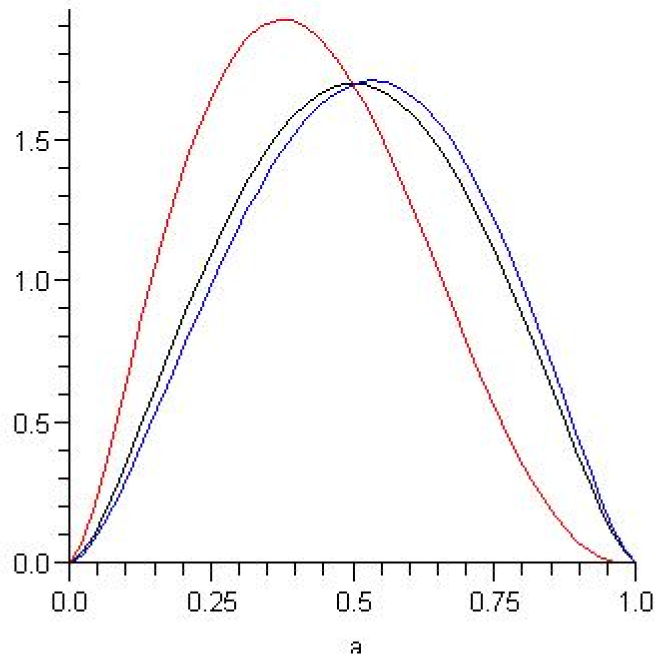

Figure 3: Beliefs updating of the median student after the first signal

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

The mass of probability is reallocated according to the realization of the signal, towards the upper bound if successful (see Figure 3, right curve) and toward the lower bound if not (see Figure 3, left curve), keeping contstant the support of the density. Notice that the bad event has a stronger effect when updating beliefs[22].

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 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 as long as the agent continues receiving informative signals.

4 Simulation

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 effect in her educational attainments of a wrong confidence in both directions. In fact, the higher human capital accumulated when the student is not too overconfident, i.e. when the mismatch effect does not prevail, 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 eventually drives her confidence towards the true level of ability.

We represent underconfidence with a distribution of prior beliefs

$$\mu(a) \sim Beta[1.5, 3] \tag{12}$$

skewed to the right. This 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

$$\mu(a) \sim Beta[3, 1.5] \tag{13}$$

skewed to the left, which implies a level of confidence $\hat{\mu}(a) = 2/3$, corresponding to the

[22]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 expected posterior confidence is always unchanged at 0.5, the upward and downward revisions would be symmetric only when $\psi = 1$, i.e. when the two events are equally likely.

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[23]. Prior beliefs of the three different types of student are summarized in Figure 4. 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.

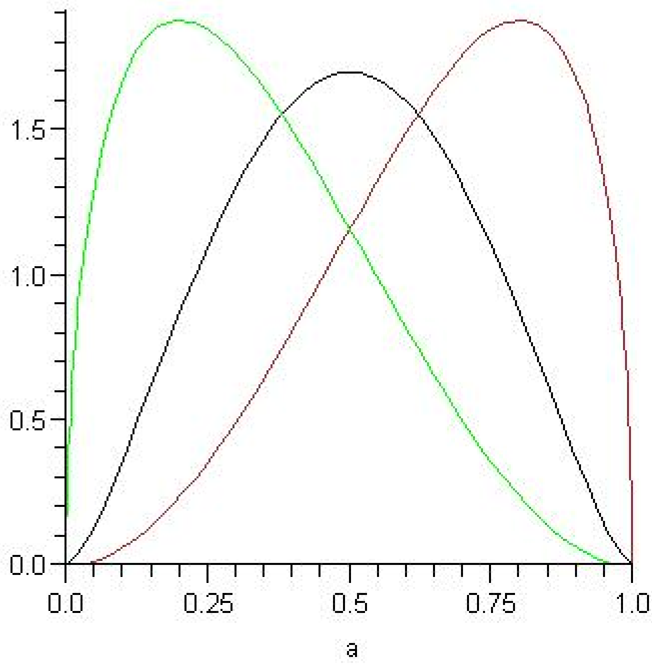

Figure 4: Prior beliefs given the different levels of confidence

We analize what happens to the human capital accumulated by the three types while the learning process takes place, iterating the updating of beliefs 45 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. In fact, at the end of the 45^{th} iteration confidence is about $.425$ for the underconfident and $.558$ for the overconfident, in both cases significantly different than $.5$ ($|p| < 0.001$)[24].

[23] 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.

[24] 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 true level of 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 optimiziation. In fact, there

Figure 5 displays the average gap, period by period, across repetitions, in the accumulation of human capital of the types who start with wrong priors as compared to the student starting with 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 also the cost of mismatching), the human capital accumulated is not much higher, while it increases as compared to the student with correct beliefs, as long as confidence converges towards the true type and the cost of mismatch decreases. Given the chosen specification of the model, the gap between the overconfident and the underconfident turns out to be about 6%.

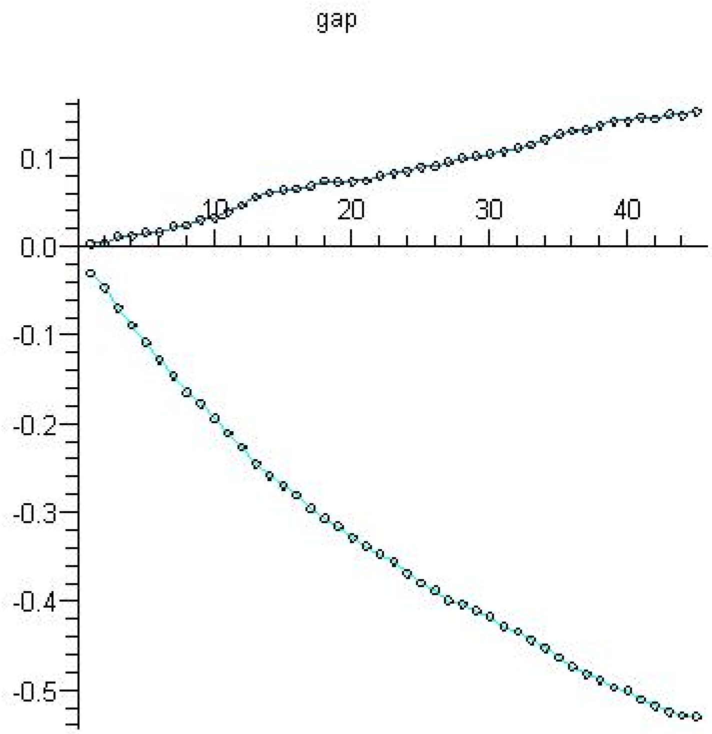

Figure 5: Gap in the accumulation of human capital

To summarize, 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

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 the best estimate of their ability $\mu(a)$. Notice that the perceived probability of success is increasing in $\mu(a)$. Hence, internalizing the different informativeness of the signal would imply a lower revision of the optimal choice at low levels of ability. In any case, maximizing utility period by period implies choices that marginally differ in terms of magnitude, and therefore a negligible mistake, particularly at low levels of ability.

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 if they lead to significantly different choices during the learning process. As long as the family background shapes children's beliefs about their ability, confidence can be a transmission mechanism that increases the intergenerational persistence of outcomes.

Notice that in the model the probability of success increases with innate ability only, while the human capital accumulated plays no role. As already noticed in Section 2.4, this simplifying assumption downplays the role of nurture, since achievements are also determined by the whole history of intermediate outcomes, in turn also driven by self-confidence, as well as by the environment in which the children grow. Therefore, what found by the model is once more a lower bound of the role of self-confidence, since the cumulative effect of the gap in the human capital accumulated during the learning process of one's ability is not taken into account. The role of nurture therefore implies that tests meant to measure students' ability are instead 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 probably 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 endow parents with measures of the innate level of ability of the children that are not confounded with the role that the family background can play through self-confidence among the several ways.

5 Conclusions

In line with some recent contributions, we claim that the socio-economic background affects not only the actual stock of cognitive skills possessed by a child (innate ability) but also the beliefs about such (unobserved) cognitive skills. There is indeed a vast literature supporting the hypothesis that people have imperfect knowledge of their ability and that many personality traits related to the concept of self-confidence are influenced by the family background in which a child grows up.

We provide further suggestive evidence about the link between confidence and family background using two very different sources: the PISA datasets, which is a representative cross-national survey of 15-year old pupils, and a very homogeneous dataset of students from Bocconi University surveyed at a later stage of their career. We show that in both samples the link between confidence and background is strong, and survives the inclusion of good controls of unobserved and observed ability. Our proxies of ability are likely to bias downward the estimated link between confidence and background, since they capture not only innate ability but also the gap in human capital that has been accumulated up to that point.

We then propose a model in which fully rational agents, who maximize the expected acquisition of human capital, choose tasks according to their perceived ability. True ability and the difficulty of the chosen track affect the probability of success. After observing whether they succeed or not, students update their beliefs, fully exploiting the available information, following Bayes' rule. We simulate the model with a bootstrapping procedure and we show that choices distorted by over- and under-confidence lead to a significant gap in the accumulation of human capital during the process in which agents eventually learn their true level of ability.

In our model agents do not derive additional utility by holding a good self-image; the consequence of this assumption is that if a perfectly informed and benevolent planner could force individuals to choose the "right" task, the effect of wrong confidence would disappear. Nevertheless, even in a setting in which agents are fully rational and have standard preferences, a moderate degree of over-confidence can be beneficial in terms of the accumulation of human capital over the life course, although at a price of a lower utility (since overconfident and underconfident agents do not make, by construction, utility-maximizing choices). Underconfidence, on the other hand, is suboptimal in terms of both utility maximization and human capital accumulation.

The intergenerational transmission of beliefs can thus constitute a further channel through which socio-economic differences perpetuate from one generation to the other because, even if two individuals had the same innate cognitive ability, differences in beliefs would lead them to make different choices in terms of investment in education. The results of our analysis suggest that policy interventions aimed at providing early and precise feedbacks about the cognitve skills of children from disadvantaged backgrounds can be beneficial in helping to narrow the gaps in educational attainments, by avoiding that equally talented people make different choices only because they have inherited different beliefs about their potential.

References

ATTANASIO, O., AND K. KAUFMANN (2009): "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Papers 15087, National Bureau of Economic Research, Inc.

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.

BÉNABOU, R., AND J. TIROLE (2002): "Self-Confidence And Personal Motivation," *The Quarterly Journal of Economics*, 117(3), 871–915.

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

BRUNELLO, G., AND M. SCHLOTTER (2011): "Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in Education & Training Systems," IZA Discussion Papers 5743, Institute for the Study of Labor (IZA).

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 Tracks," IZA Discussion Papers 2876, Institute for the Study of Labor (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.

CHOWDRY, H., C. CRAWFORD, AND A. GOODMAN (2011): "The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16," *Longitudinal and Life Course Studies*, 2(1), 5–76.

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. M. SCHENNACH (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931.

DELANEY, L., C. HARMON, AND C. REDMOND (2011): "Parental Education, Grade Attainment and Earnings Expectations among University Students," Discussion Paper Series 5646, IZA.

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

FALK, A., D. HUFFMAN, AND U. SUNDE (2006): "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.

GIULIANO, P. (2008): "Culture and the Family: An Application to Educational Choices in Italy," *Rivista di Politica Economica*, 98(4), 3–38.

GOODMAN, A., AND P. GREGG (eds.) (2010): *Poorer children's educational attainment: how important are attitudes and behaviour?*, Joseph Rowntree Foundation Report. London. Joseph Rowntree Foundation.

GREGG, P., AND E. WASHBROOK (2011): "The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 11," *Longitudinal and Life Course Studies*, 2(1), 41–58.

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.

HEINECK, G., AND S. ANGER (2010): "The returns to cognitive abilities and personality traits in Germany," *Labour Economics*, 17(3), 535–546.

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

JENSEN, R. (2010): "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 125(2), 515–548.

KAUFMANN, K. M. (2010): "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns," Working Paper Series 362, IGIER, Bocconi University.

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

LINDQVIST, E., AND R. VESTMAN (2011): "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics*, 3(1), 101–28.

NGUYEN, T. (2008): "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," Working paper, MIT.

NYHUS, E., AND E. PONS (2005): "The effect of personality on earnings," *Journal of Economic Psychology*, 26, 363–384.

OECD (2009): *Top of the Class. High Performers in Science in PISA 2006 Programme*. OECD Publishing.

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

WEINBERG, B. A. (2009): "A Model Of Overconfidence," *Pacific Economic Review*, 14(4), 502–515.