CESifo WORKING PAPERS

6863
2018
January 2018

Do Preferences and Biases Predict Life Outcomes? Evidence from Education and Labor Market Entry Decisions

Uschi Backes-Gellner, Holger Herz, Michael Kosfeld, Yvonne Oswald

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6863*
Category 5: Economics of Education

Do Preferences and Biases Predict Life Outcomes? Evidence from Education and Labor Market Entry Decisions

Abstract

Evidence suggests that acquiring human capital is related to better life outcomes, yet young peoples' decisions to invest in or stop acquiring human capital are still poorly understood. We investigate the role of time and reference-dependent preferences in such decisions. Using a data set that is unique in its combination of real-world observations on student outcomes and experimental data on economic preferences, we find that a low degree of long-run patience is a key determinant of dropping out of upper-secondary education. Further, for students who finish education we show that one month before termination of their program, present-biased students are less likely to have concrete continuation plans while loss averse students are more likely to have a definite job offer already. Our findings provide fresh evidence on students' decision-making about human capital acquisition and labor market transition with important implications for education and labor market policy.

JEL-Codes: D010, D030, D910, I210, J640.

Keywords: economic preferences, education, dropout, human capital, job search.

Uschi Backes-Gellner
Department of Business Administration
University of Zurich
Plattenstr. 14
Switzerland – 8032 Zurich
ubg@business.uzh.ch

Holger Herz
Department of Economics
University of Fribourg
Bd. de Perolles 90
Switzerland – 1700 Fribourg
holger.herz@unifr.ch

Michael Kosfeld
Faculty of Economics and Business
Administration, Goethe University
Theodor-W.-Adorno-Platz 4
Germany – 60323 Frankfurt am Main
kosfeld@econ.uni-frankfurt.de

Yvonne Oswald
Department of Business Administration
University of Zurich
Plattenstr. 14
Switzerland – 8032 Zurich
yvonne.oswald@oec.uzh.ch

January 16, 2018
The paper benefited from helpful comments by seminar participants at Aarhus, Amsterdam, Cologne, Frankfurt, Hamburg, Heidelberg, Maastricht, München, Oxford, Regensburg, San Diego, Southampton, Stanford, Zurich, the IMEBESS conference in Toulouse 2015, the Economics in Education workshop in Mainz 2015 and the ESA Europe conference in Heidelberg 2015 and the CESifo behavioral economics conference 2017. We are grateful to Donata Bassey for her help in conducting the experiments.

1 Introduction

Acquiring human capital is considered among the prime factors for subsequent higher income and other important positive life outcomes. For example, Lindahl and Krueger (2001) find that an additional year of schooling raises earnings by about 10 percent. Yet, the determinants of young peoples' decisions to invest in or stop acquiring human capital are still relatively poorly understood.

Human capital theory (Mincer, 1958; Schultz, 1961; Becker, 1962) provides a straightforward economic framework for analyzing educational investment decisions. Individuals invest in their own education if the expected present value of the benefits is higher than the expected present value of the costs. Given the documented benefits of schooling, the question arises why some students stop acquiring human capital at relatively early stages. In this paper, we contribute to answering this question by analyzing empirically the role of (heterogeneity in) economic time preferences and behavioral biases in the decision to finish or drop out of upper-secondary education programs. In addition, our data enables us to investigate the predictive power of preferences and biases for job search decisions upon completion of the program.

Eckstein and Wolpin (1999) provide a number of reasons for why students may terminate education, one of them being that dropouts have lower expectations about the rewards from graduation. Such lower expectations could, for example, be stemming from an underestimation of lifetime benefits from staying in school. Similarly, Oreopoulos (2007) suggests that ignorance or the heavy discounting of substantial lifetime gains generated by additional schooling might explain dropout behavior. Consistent with this view, Golsteyn et al. (2013) document a significant association between hypothetically elicited time preferences at age 13 and lifetime outcomes such as earnings, health and education in Swedish data. In particular, they find that higher patience is positively related to good life outcomes, and argue that educational attainment modulates this positive effect.

In the case of upper-secondary education, however, schooling is no longer compulsory and hence a decision to drop out is always subsequent to a previous enrollment decision. The key question is therefore why those who drop out decided to start non-compulsory education in the first place. Two arguments have been emphasized in the recent literature: incomplete information and time inconsistency. These arguments are not only very different in nature — rational learning vs. bounded rationality —, they also yield very different, in fact conflicting, policy implications (cf. below). A careful investigation of these potential explanations seems therefore warranted.

Incomplete information, on the one hand, assumes that students at the time of enrollment are only incompletely informed about the costs and benefits of pursuing education. Acquiring education in this case involves an element of experimentation and dropouts rationally occur as a consequence of new information updates (Manski, 1989; Altonji, 1993). Such information updates could, for example, stem from learning about individual ability, the job market perspectives upon completion, or the effort and opportunity costs of completing the program. Stinebrickner and Stinebrickner (2012, 2013) and Zafar (2011) show that such information updating indeed occurs and can account for dropout decisions. Thus, if students hold only partial information at the time of enrollment, they continuously make cost-benefit trade-offs while pursuing upper-secondary education and may be tipped towards termination in light of information shocks. Using data from the National Longitudinal Survey of Youth (NLSY), Arcidiacono et al. (2016) estimate that eliminating informational frictions would indeed increase the college graduation rate by 9 percentage points.

Time inconsistency, on the other hand, assumes that students have present-biased preferences (Laibson, 1997; O'Donoghue and Rabin, 1999). In this case, the cost-benefit tradeoff of continued education can change between enrollment and the time at which education is actively pursued, even in the absence of new information. The reason is that once the costs of education become immediate and the benefits remain in the future, present-biased students prefer to discontinue education programs that they previously enrolled in, thereby acting in a time-inconsistent manner. Cadena and Keys (2015) assess this hypothesis with NLSY data documenting that a proxy for student impatience correlates with dropout from college, which is taken as evidence for time-inconsistent preferences and its impact on dropout decisions.[1]

Notice that these two explanations yield conflicting policy implications. If students have present-biased preferences and drop out of education because they overvalue immediate costs, commitment devices limiting the possibility to quit education (by making dropouts more costly) seem favorable. If, in contrast, students have incomplete information about the costs and benefits of education and learn over the course of the program whether the chosen educational path fits their preference or ability, such policy is exactly what one would *not* like to do. Instead, eliminating (or at least reducing) informational frictions would be preferable. Our paper contributes to this discussion by providing novel and detailed empirical evidence on the relationship between dropouts, economic time preferences, and behavioral biases.

We analyze a unique data set that combines behaviorally elicited information on students' economic preferences and potential behavioral biases with administrative data on education outcomes in the context of vocational training programs in Switzerland. In a vocational training program, students study part-time at vocational schools and work part-time at host companies. It constitutes the most popular form of post-secondary education in Switzerland, accounting for 70% of all post-secondary education degrees in the country. For students who are about to successfully complete the program, we also obtain detailed survey measures on labor market transition or continued higher education plans. Our behavioral measures were taken directly in the classroom, at the very beginning of the program. They include incentivized measures of time preferences (long-run patience, present bias), as well as risk and loss aversion. In addition, we obtain a number of important controls such as proxies for intelligence and other socio-demographic characteristics that are known to predict life outcomes and at the same time correlate with patience and risk aversion (Dohmen et al., 2010). In comparison to previous studies, our results thus rely on precise individual estimates of economic preferences allowing us to explicitly differentiate between long-run patience and present bias as well as a large set of covariates to control for potential confounding factors.

[1]The proxy Cadena and Keys (2015) use is whether or not a student was classified as "restless and impatient" during the interview by the interviewer. The same measure was used before by DellaVigna and Paserman (2005) to assess the relationship between impatience and job search.

In total, we were able to obtain a sample of 265 students, out of which 30 (11.5%) terminate their vocational training contracts prior to completion. The observed dropout rate is similar to the average dropout rate on the cantonal level (9 percent in 2008; Maghsoodi and Kriesi, 2013). Our results show that the association between present bias and dropout is relatively weak. While an increase in a student's present bias increases the probability of dropout, the association is very small and not statistically significant. We do, however, find that long-run patience is significantly negatively associated with dropout behavior. Controlling for a wide array of socio-demographic characteristics, a one-standard deviation increase in the measured 3-month discount rate decreases the likelihood of dropping out of the vocational training program by approximately 2.6 percentage points. Similar results are obtained if we consider information about whether a student finishes the program in time as an alternative outcome measure: long-run patience significantly correlates with this measure, whereas present bias has no predictive power. In sum, our results do not provide evidence that time inconsistency is a key driver of dropout behavior from upper-secondary education. Rather, they suggest that long-run patience, together with information updating, plays an important role.

Preferences and biases may not only matter for completion of educational programs. They might be similarly important for job search decisions, and in turn for labor market entry and early career labor market success of the students. DellaVigna and Paserman (2005) show theoretically that present bias reduces incentives to invest into job search, implying a negative effect on the transition from unemployment to employment.[2] Empirically, they find an association between a measure of impatience and the length of unemployment spells in the NLSY.[3] In addition, DellaVigna et al. (2017) propose a model of job search with reference-dependent preferences and loss aversion relative to recent income. They derive the model prediction that anticipated benefit cuts increase search efforts of the unemployed, and find transition patterns in Hungarian data that are consistent with this theory.

Arguably, at the end of their educational program, apprentices are in a comparable situation to the unemployed in terms of incentives to search for a job, because their initial work contracts are limited to the duration of the program.[4]Applying the theories by DellaVigna and Paserman (2005) and DellaVigna et al. (2017) to our setting, we should therefore expect that present-biased students are less likely and loss averse students more likely to have secured a job offer, shortly before their vocational training program ends. Further, both effects should be driven by incentives to invest into job search, which are expected to increase in loss aversion and to decrease in present-bias.

To assess these hypotheses, we administered a labor market transition survey to students about one month before the end of the vocational training program. In the survey, we asked whether students already have a definite job offer, whether they plan to continue higher education, or

[2]The effect of impatience is ambiguous, because it jointly affects the reservation wage and investment incentives.

[3]Ben Halima and Ben Halima (2009) also find evidence in French job search data that is consistent with hyperbolic discounting.

[4]While it does happen that firms continue to employ their apprentices, a large fraction is forced to enter the labor market and actively search for a job.

neither. We were able to collect survey responses from 181 students who were expected to finish their program (81% of the relevant initial sample). Of these, 92 (51%) had a definite job offer, 47 (26%) planned continued education, and 42 (23%) had neither. We also asked them whether they actively engaged in job search activities. By combining this survey data with our experimental preference measures, elicited several years before, we obtained a unique data set that, to our knowledge, allows for the first time to directly assess these job search theories using incentivized preference measurements and labor market outcomes at the individual level.

We find evidence that is consistent with both predictions. First, consistent with DellaVigna and Paserman (2005), students who are more present biased are indeed significantly less likely to have a definite job offer or concrete plans for continued higher education. A one standard deviation decrease in the estimated β increases the probability of having no job relative to having a job by around 13-18 percentage points. At the same time, long-run patience is not significantly associated with these outcomes. Second, consistent with DellaVigna et al. (2017) students who are more loss averse are more likely to have a definite job offer. A one standard deviation increase in loss aversion increases the probability that an apprentice has a definite job offer rather than no continuation plan by 11-15 percentage points. Our results therefore suggest that loss aversion and present bias play an important role for students' labor market entry and early career success.

These results have several implications for policy. First, we show that long-run patience — and *not* present bias — is significantly associated with dropping out of upper-secondary education. This suggests that policies targeted at reducing dropouts should focus on factors that influence students' long-run patience positively, in particular during early childhood (Cunha and Heckman, 2007; Falk and Kosse, 2016), together with eliminating information frictions. Commitment devices, on the contrary, that would limit the possibility to terminate non-compulsory education are likely to be ineffective and may even be harmful in light of the fact that acquiring education also involves an element of experimentation.[5] Second, such commitment devices may instead be useful when it comes to student behavior towards the end of the education program. As our results show, present bias — and *not* long-run patience — significantly correlates with student outcomes in terms of concrete options and plans to enter regular employment or higher education. Here, early deadlines and related policy instruments that increase a student's effort and commitment to ensure a successful transition out of the vocational training program seem beneficial.

Besides highlighting important mechanisms in human capital acquisition, our paper contributes to a broader literature on how predictive behaviorally elicited preference measures are for a variety of lifetime outcomes. Castillo et al. (2011) find that impatience correlates with disciplinary referrals in school. Other studies have looked at the differential effect of hyperbolic vs. exponential discounting on credit card borrowing and credit worthiness (Meier and Sprenger, 2010, 2012). Sutter

[5] Cadena and Keys (2015) argue that late dropouts, for example after the third year of college, are unlikely to be due to learning. Indeed, their impatience measure correlates particularly strongly with these late dropouts. Because of data limitations (too few late dropouts), we cannot directly assess this specific hypothesis. Assuming that these late dropouts are indeed due to present-bias and not learning, our data nonetheless strongly suggests that commitment devices for completion should, if at all, only be applied in the late phases of educational programs.

et al. (2013) analyze the effects of hyperbolic and exponential discounting on saving, smoking and alcohol consumption of school children. Chabris et al. (2008) document similar correlations between discount rates and smoking, body mass index, and exercise behavior. Our paper differs from these contributions by being the first to use incentivized behavioral experiments in combination with administrative and survey data in analyzing the role of economic time and reference-dependent preferences in human capital acquisition and labor market transition.

The remainder of the paper is structured as follows. In the next section, we illustrate how time and reference-dependent preferences matter for dropout decisions and labor market transition by means of a simple model as well as results from the relevant literature. Section 3 explains our preference measures and the administrative and survey data. Section 4 presents the empirical results. Section 5 concludes.

2 How Time and Reference-Dependent Preferences Matter

2.1 For Dropout Decisions

To illustrate how time preferences affect a student's decision to invest in or stop education, consider the following simple model. A student at time $t = 0$ decides whether to start education in period $t = 1$. Education generates both costs $c < 0$ that occur in period $t = 1$ and future benefits $b > 0$ that occur in periods $t = 2, \ldots$.[6] The student discounts future payoffs according to a discount function that is equal to one for the current period and equal to $\beta\delta^\tau$ for later periods $\tau \geq 1$ with $\beta, \delta \leq 1$, where δ denotes the long-run rate of time preference and β an individual's potential present bias. Formally, the present value of future income streams in period t equals

$$U_t = x_t + \beta \sum_{\tau=1}^{\infty} \delta^\tau x_{t+\tau}, \tag{1}$$

where x_t is equal to the cost or benefit in period t.

At time $t = 0$, the student plans to start education if and only if the discounted net future payoff is larger than zero. Formally,

$$\begin{aligned} -\beta\delta c + \beta \sum_{t=2}^{\infty} \delta^t b \quad &> \quad 0 & (2) \\ \frac{\delta}{1-\delta} \quad &> \quad \frac{c}{b}. & (3) \end{aligned}$$

Only students with sufficient long-run patience $\delta > \frac{c}{c+b}$ get enrolled in education. As both costs and benefits occur in the future, present bias β does not matter for decision-making at $t = 0$. This, however, changes in period $t = 1$.

If the student invests in education in $t = 1$, education costs arise immediately, while all benefits

[6]For expositional simplicity, we assume here that the student is infinitely lived. Arguments do not depend on this.

occur in periods $t = 2, \ldots$. The student actually invests if and only if

$$
\begin{aligned}
-c + \beta \sum_{t=1}^{\infty} \delta^t b &> 0 \qquad (4) \\
\frac{\beta\delta}{1-\delta} &> \frac{c}{b}. \qquad (5)
\end{aligned}
$$

For $\beta = 1$, this condition is identical to the *selection condition* (3) in period $t = 0$. In the absence of information shocks, time-consistent students do not change their education plan. This is different, if $\beta < 1$. In particular, if condition (3) is fulfilled but $\beta < \frac{c(1-\delta)}{b\delta}$, the student in $t = 0$ plans to start education but changes his plan in $t = 1$ and drops out.

More generally, if information shocks can occur, i.e., the student updates information about c and b in $t = 1$, the following condition becomes relevant:

$$
\frac{\beta\delta}{1-\delta} > \frac{\tilde{c}}{\tilde{b}}, \qquad (6)
$$

with $\tilde{c}$ and $\tilde{b}$ denoting updates of current costs and future benefits, respectively. Note that in the *revision condition* (6) both β and δ together determine whether a student continues or drops out. More specifically, information shocks will be more likely to lead to dropouts the smaller is the left hand side of (6). It can be shown that changes in δ have a larger effect on the LHS of (6) unless β is very small, which is a consequence of compounding. To see this, note that the marginal effect of β in the left-hand-side of (6) is equal to $\frac{\delta}{1-\delta}$. The marginal effect of δ is equal to $\frac{\beta}{(1-\delta)^2}$. The latter is larger than the former if and only if $\beta > \delta(1-\delta)$, which is always the case if $\beta > 0.25$. This implies that if information shocks occur — suppose, e.g., that the right-hand-side of (6) increases —, it is more likely that the revision condition is violated because of a student's δ rather than a student's β.

Let us summarize our hypotheses with respect to dropout. Long-run patience δ determines both the selection and the revision decision. Present bias β only plays a role in the revision decision. Thus, if time-inconsistency is the main driving force behind dropout, we should observe a negative and significant effect of β on a student's decision to terminate the program prior to completion (negative, because a higher β makes it less likely that the student drops out). If information updates are relatively more important, however, the effect of δ should be negative and significant.

2.2 For Labor Market Transition

With respect to labor market transition our hypotheses are based on the job search model of DellaVigna and Paserman (2005). There, the authors show that present bias has a negative effect on an individual's probability to secure a job, while the effect of long-run patience is ambiguous or even positive. The reason is that present bias affects only the short-run incentive to search whereas long-run patience affects both search incentives and an individual's reservation wage. In particular, it can be that individuals, who are impatient in the long-run, are more likely to accept any job offer, given they have one.

Besides time preferences, we can also make a prediction with regard to reference-dependent preferences based on DellaVigna et al. (2017). They show that loss aversion increases job search effort and thereby the probability to enter employment. The intuition for our setting is straightforward: Because a loss-averse student experiences an extra loss in utility when not having secured a job after education, this increases the incentives to search and generate a job offer.

3 Data

To analyze the role of economic preferences in explaining dropout behavior and labor market transition, we collected a well-suited data set that comprises four key features: (1) Individual preference measures elicited through incentivized experiments at the beginning of the first year of the education program; (2) important student characteristics including socio-economic background, IQ proxies as well as BIG 5 and GRIT personality measures, (3) register data on student dropouts and successful completion of the educational program; (4) and survey measures on students' plans for labor market transition about one month prior to the end of the vocational training program. In the following sections, we explain all data in detail.[7]

3.1 Student Sample

Our sample consists of students in upper-secondary education who are enrolled in a vocational training program in Switzerland. The average age of students at the time of enrollment in these programs is 16. Students study part-time at vocational schools and work also part-time at host companies. The students are employed at the host company for the duration of the education program and earn a moderate wage.[8] In Switzerland, about 70% of the graduates of lower-secondary education enroll in such vocational education (OPET, 2011). Hence, our student sample represents the largest part of young adults pursuing upper-secondary education in Switzerland.

We conducted in-class experiments within the first weeks of school in the first year of the education program, in late August and early September 2009. Experiments took place during school hours, lasting approximately one hour. In total, 265 students from 14 complete classes in three public, tuition-free vocational schools in Switzerland participated. All schools are located in the greater region of Zurich, the largest city of Switzerland.

60 percent of the students in our sample participate in training programs in the commercial sector, planning to become commercial employees; 40 percent participate in the technical sector, planning to become either electricians or polytechnicians. These three training programs are among the top ten regarding the number of students of all 230 training programs offered in Switzerland (OPET, 2011). The training program for students in the commercial sector lasts three years and

[7]Parts of this data are also used in Oswald and Backes-Gellner (2014), who study the role of financial incentives on student's school performance, and their interaction with preferences.

[8]Financial constraints are hence an unlikely cause of dropouts, in contrast to college dropouts in the US.

includes training in a broad range of skills for carrying out administrative work in various industries. In contrast, the training programs for students in the technical sector last four years and include training in different technical skills. While electricians learn specific skills for setting up, installing, and maintaining complex electrical wiring systems, polytechnicians learn how to fabricate special tools and work pieces required in the production sector, program and operate machines, and monitor different types of production.

3.2 Experimentally Elicited Preference Measures and Controls

3.2.1 Time Preferences

We elicited time preference using incentivized choice experiments. More precisely, each student made decisions on two multiple-price lists. On each list, students were asked to chose between a smaller payment of CHF X at an earlier date and a larger payment of CHF 100 at a later date three months later. On the first price list, the earlier date was the present and the delayed date was in three months. On the second price list, the earlier date was in three months and the delayed date was in six months. Each price list consisted of 20 decisions between X at the earlier date and 100 at the later date, where X varied systematically in increments of 5 Swiss Francs between CHF 5 and CHF 100.[9]

Students' decisions from these multiple-price lists provide an estimate of students' time preferences as well as potential present bias (Laibson, 1997; O'Donoghue and Rabin, 1999). Consider equation (1) from Section 2, with U_t denoting the present value of future income streams at time t, δ the long-run rate of time preference and β an individual's potential present bias. In the following, we adopt the 3-months time distance between payments as one unit of time. Hence, $t = 0$ is the present, $t = 1$ is in three months, and $t = 2$ in six months from now.

Let us start with the second price list, which only contains payments in the future. The decision maker will prefer the sooner payment x_1, if and only if

$$U_t(x_1) = \beta\delta x_1 \geq U_t(x_2) = \beta\delta^2 100, \tag{7}$$

or, equivalently, $x_1 \geq \delta 100$. For each student, we observe the lowest x_1 in three months that is

[9] Similar procedures have been used, e.g., by Burks et al. (2012), Dohmen et al. (2010), Meier and Sprenger (2010, 2012, 2015), Balakrishnan et al. (2015), and Dohmen et al. (2017). In recent years (after our experiments were conducted), the procedure to use time-dated monetary rewards to measure time preferences has been called into question based on arguments of non-credibility of future payments, curvature of the utility function, possibility of arbitrage, or credit constraints. See, e.g., Andersen et al. (2008), Andreoni and Sprenger (2012a,b), and Augenblick et al. (2015) proposing various alternatives to circumvent these caveats. Unfortunately, there is still no consensus on what procedure is best (Andreoni et al., 2015), each has its pros and cons. In particular, multiple-price lists are comparably easy to implement in the field and with non-standard subject pools, thereby reducing noise based on lack of understanding. Dohmen et al. (2017) also find no evidence that choice patterns can be explained by the potential confounds in a representative sample of adults in Germany, and Balakrishnan et al. (2015) show that measures using multiple-price lists and "convex time budgets" (Andreoni and Sprenger, 2012a) are strongly and highly correlated. In our set-up, we explicitly guaranteed credibility of future payments by an official statement from the University of Zurich. Further, we control for risk and loss aversion by means of additional behavioral measures, and we include an explicit question on credit constraints (see below).

revealed to be preferred to CHF 100 in six months. Let us denote this value by X_1. We can then define an upper bound on that student's long-run rate of time preference by $\delta := X_1/100$.

Now, consider the first price list. Here, the decision maker will prefer the sooner payment, if and only if

$$x_0 \geq \beta\delta 100. \tag{8}$$

Again, we observe X_0, the lowest x_0 that is revealed to be preferred to CHF 100 in three months. Substituting δ into this equation, we can identify a student's present bias as $\beta := X_0/X_1$. Intuitively, if a student reveals consistent time preferences, the two switch points X_1 and X_0 are the same, i.e., β is equal to one. In case of present bias, however, the student switches earlier in the first price list (now vs. three months) than in the second price list (three months vs. six months). In other words, $X_0 < X_1$ implying $\beta < 1$.[10]

Note that our estimation strategy is not feasible if a student has multiple switch points, that is, if some value x_t is preferred over CHF 100 three months later, but then CHF 100 is again revealed preferred over some higher value $x'_t > x_t$. In this case, the preference relation is intransitive. In our analysis, we exclude all students for whom we cannot identify a unique switch point, which is the case for 20 out of the 265 students (7.5%).

3.2.2 Risk and Loss Aversion

To measure risk and loss aversion, we ran two lottery tasks. In the first lottery task to assess a student's risk aversion, each student is presented with the opportunity to participate in ten different lotteries, each of the following form:

> Win CHF 10 with probability $\frac{1}{2}$ or CHF 0 with probability $\frac{1}{2}$, or reject the lottery and get a fixed payment of CHF Y.

The ten lotteries varied in the amount Y offered as a certain payment, where Y took on the values $Y \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. At the end of the experiment, one of the ten lotteries was randomly selected and paid. The higher a student's risk aversion, the lower should be the value of Y at which the student starts to reject the lottery and take the certain payment instead. Thus, the amount Y at which a student starts rejecting the lottery can be taken as a proxy for that student's degree of risk aversion. For example, a student who rejects all lotteries for a certain payment of $Y > 3$ is classified as exhibiting higher risk aversion than a student who only rejects all lotteries for a certain payment of $Y > 7$. We use the largest amount Y at which a student still prefers the lottery and define an index of risk aversion by $riskaversion = (10 - Y)$.[11]

In the second lottery task to assess students' loss aversion, each student is presented with the opportunity to participate in six different lotteries, each of the following form:

> Win CHF 6 with probability $\frac{1}{2}$ or lose CHF X with probability $\frac{1}{2}$. If the subject rejects the lottery s/he receives CHF 0.

[10] Similarly, a student reveals future bias, if $X_0 > X_1$ or, equivalently, $\beta > 1$.

[11] Reversing the index is convenient so that larger values of $riskaversion$ indeed indicate stronger risk aversion.

The six lotteries varied in the amount X that could be lost, where X took on the values $X \in \{2, 3, 4, 5, 6, 7\}$. Again at the end of the experiment, one of the six lotteries was randomly selected and paid. The higher a student's loss aversion, the lower should be the value of X at which the student starts to reject the lottery. Thus, the amount X at which a student starts rejecting the lottery can be taken as a proxy for a student's loss aversion. For example, a student who rejects all lotteries with a potential loss of $X > 3$ is classified as exhibiting higher loss aversion than a student who only rejects all lotteries with a potential loss of $X > 5$. We use the largest possible loss X at which a subject still prefers the lottery and define an index of loss aversion by $lossaversion = (7 - X)$.[12]

As before, we cannot precisely define the risk aversion and loss aversion index in case there are multiple switch points. We therefore exclude all students with multiple switch points from the analysis, which is the case for 15 out of our 265 students (5.7%).

In total, 231 students gave consistent answers in all four preference elicitation tasks (88%). Importantly, as shown below, these students do not differ significantly from students with inconsistent answers in any of our outcome variables.

3.2.3 Socio-economic and Personality Characteristics

Prior to the choice experiments, we collected several socio-economic characteristics as well as personality measures. In particular, we elicited students' final grades in English, German and Math in high school, information about their parents' educational background, their age, gender, country of birth and their native language. To assess whether credit constraints might affect student's decision making in the inter-temporal choice tasks, we also included a question on how difficult it is for a student to spontaneously raise CHF 100, which was answered on a 5-point Likert scale, with larger numbers indicating less difficulty.

We also gathered personality measures through surveys. First, we implemented the GRIT questionnaire, measuring a student's perseverance and passion for long term goals (Duckworth et al., 2007), consisting of 17 items. Second, we implemented the German 15-item version of the BIG 5 questionnaire (Gerlitz and Schupp, 2005). Once these surveys were finished, students participated in the cognitive reflection test (CRT) (Frederick, 2005) in order to get another proxy for their IQ in addition to the high school grades.[13]

In our regression analysis, we group our control variables as follows: The first set of controls includes students' training program (commercial, electrician, or polytechnician) and gender. The

[12]Again, reversing the index is convenient so that larger values of *lossaversion* indeed indicate stronger loss aversion. In principle, the rejection of actuarially fair gambles in this lottery choice task may also reflect a subject's risk aversion. However, since we simultaneously control for preferences over risk from a task that does not involve losses, we attribute individual differences that stem from this task to differences in individual loss aversion. Losses that actually occurred in the experiment were covered by earnings from the remaining choice experiments and the participation fee.

[13]Subjects also participated in a symbol-digit correspondence test, a sub-module in the non-verbal section of the Wechsler Adult Intelligence Scale (WAIS). However, due to missing observations on this test, we decided to drop this measure in our analysis in order to not loose observations. Including this IQ score and dropping the missing observations leaves our results unaltered, however. Results are available upon request.

second group of controls includes all socioeconomic variables, i.e., age, an indicator whether they are native German speakers, high school grades and CRT score, parents' educational background and the indication on potential credit constraints. The third set of controls includes personality measures, i.e., GRIT and BIG 5 scores. Table A.1 in Appendix A provides summary statistics of all control variables.

3.2.4 Procedures

Students were asked to fill out all surveys and answer all questions independently, and to remain quiet while the experiments were conducted. In all classes, students first filled out the surveys and participated in the IQ tests. Then, the choice experiments were conducted. Once the choice experiments were finished, all students were paid in private in an adjacent room.

Each student received CHF 10 for participation in the study. In addition, students earned additional money from the choice experiments. For each of the two lottery tasks, one gamble was randomly selected and paid. If the student decided to take the respective gamble, the student himself flipped a coin which determined the outcome of the gamble. For the inter-temporal decision tasks, not all students were paid. Each subject received an individual ID number, and once all subjects had finished the choice experiments, in each class two ID numbers were randomly and publicly drawn for each of the two inter-temporal decision tasks for payment. For these four subjects, again one of the 20 inter-temporal decisions was selected at random, and they were paid the respective amount at the respective time according to their choice. In case payment was in the future (i.e., either three or six months later), the respective amount was sent by mail to the home address of the student. All future payments were explicitly guaranteed by an official letter from the University of Zurich that was shown to all students.[14]

3.3 Dropout and Labor Market Transition

After the standard time to finish the vocational training program had elapsed, we collected administrative data on successful completion, dropout or delay in finishing the program. In addition, we administered a survey about one month before the end of the program to collect information on students' options and plans with respect to labor market transition or continued higher education.

3.3.1 Dropouts

The official register data on dropouts was collected from the cantonal office, the *Mittelschul- und Berufsbildungsamt* in Zurich. In particular, we received information on whether or not a contract was terminated prior to completion of the program. In addition, we observe whether a student finished the program within the expected time (three years in case of the commercial program, and four yours in case of the two technical programs). This measure differs from the dropout measure

[14] Class size varied between 13 and 23 students. In total, 21.1 percent of students were paid out for one of the two inter-temporal decision tasks.

in two regards. First, some students dropped out of the program at the very beginning and started a new program right away, so they did not suffer any economic consequences from their dropout. These are coded as “dropout”, but also as “having finished in time” (4 out of 26 students in the final sample; cf. Table 1). Second, some students did not drop out but had to prolong the program, for example because they failed important exams. These are coded as “no dropout”, but also as “not having finished in time” (23 out of 201 students in the final sample; cf. Table 1).

3.3.2 Labor Market Transition

About one month before the end of the program (2012 for students in the commercial program, 2013 for students in the technical programs), we administered a survey to assess students' concrete options and short-run plans with regard to the job market or further education. We contacted students via their school class and in addition tried to reach those who were not present via mail.[15] In the survey, we asked students whether they already have a definite job offer for the time after the program. If this were not the case, we wanted to know whether they are planning any full- or part-time education program instead. Importantly, enrollment deadlines for Swiss Universities and Universities of Applied Sciences are at the end of April, prior to our transition survey. Hence, continued education plans at such institutions had to be very concrete. Those who neither had a job offer nor planned to continue education could indicate further possibilities such as “making a break” or “planning a longer stay abroad”. However, all these other options involve neither working nor further acquiring human capital in the short run. We pool these answers in the analysis.

4 Results

Before we turn to the regression analysis of dropout behavior and labor market transition decisions, we present descriptive statistics of our data.

4.1 Descriptive Statistics

Out of our initial sample of 265 students, 4 students had to be dropped because we had no access to their register data. These students had moved out of the canton Zurich during the four years of their program and their data was transferred to another cantonal state office, which we have no access to. From the remaining 261 students, 30 terminated the program prior to completion and 54 students did not finish in time. As mentioned before, 34 out of the 261 students gave inconsistent answers in the preferences measures. We exclude these observations in our analysis, leaving us with a final sample of 227 students. Table 1 shows the joint distribution of our two outcome variables on program completion for the final sample. 26 out of the 227 students terminated their contract prior to completion of the program, which amounts to 11.5% of the sample. Moreover, 45 students did not finish the program in time (19.8%). Importantly, students with consistent answers do not differ

[15] As an incentive for participation, two iPad3 were raffled among students who filled out the survey.

significantly from those with inconsistent answers in neither of the outcome measures (dropout: 11.5% vs. 11.8%, $p = 0.96$; finished in time: 80.2% vs. 79.4%; $p = 0.92$).

Table 1: Joint Distribution of Dropout and Finished in Time

	Finished in Time		
Dropout	No	Yes	Total
No	23	178	201
Yes	22	4	26
Total	45	182	227

Table 2 shows the mean and standard deviation of our preference measures conditional on the two outcome measures. Distributions of all preference measures are included in Appendix A. The table shows that students are, on average, risk neutral. The mean measure of risk aversion is roughly equal to 5, which implies that students on average switch from accepting the coin toss with a 50% chance of winning CHF 10 to accepting a certain payment precisely when the certain payment is CHF 5. However, the standard error is relatively large, implying considerable heterogeneity in risk aversion.

Table 2: Average Preference Measures, by Dropout and Finished in Time

	All	Dropout			Finished in Time		
		No	Yes	p	Yes	No	p
Long-run Patience (δ)	0.78	0.79	0.74	0.06	0.79	0.75	0.16
	(0.16)	(0.17)	(0.15)		(0.16)	(0.17)	
Present Bias (β)	0.95	0.95	0.96	0.28	0.95	0.95	0.56
	(0.22)	(0.22)	(0.22)		(0.23)	(0.20)	
Risk Aversion	5.02	5.02	5.00	0.92	5.03	4.96	0.61
	(1.69)	(1.62)	(2.17)		(1.67)	(1.76)	
Loss Aversion	4.93	4.95	4.77	0.39	4.95	4.84	0.68
	(1.05)	(1.05)	(1.07)		(1.03)	(1.17)	
Number of Obs.	227	201	26		182	45	

Note: Standard Errors in parentheses. Column p shows the p-value of a Wilcoxon rank-sum test comparing dropouts and non-dropouts and finished and not finished in time, respectively.

Second, mean loss aversion is in the range of 4.7 to 4.9, which implies that students on average switch to rejecting the coin toss in the loss gamble, in which they could win CHF 6 with 50% probability or lose CHF X with 50% probability, when the potential loss is between 4 and 5 CHF. Hence, while our subject pool on average appears to be risk neutral, we do find evidence for relatively mild loss aversion. Heterogeneity in loss aversion in our sample is also considerable.

Third, we do find considerable discounting in our sample. Recall that δ, our measure for long-run patience, is directly inferred from the switch point in the multiple-price list in which both payments are in the future. A mean δ of 0.74 to 0.79 implies that students are willing to accept an amount X_1 in three months that is, on average, equal to 74 to 79 CHF, rather than waiting for CHF 100 in six months. Here, the difference between dropouts and non-dropouts as well as finished and not finished in time is larger, averaging 4 to 5 percentage points.

Finally, a potential present bias (β) is inferred from the ratio of switch points in the two multiple-price lists involving immediate payments and involving only delayed payments. The average β in our data is 0.95. Distributions in Appendix A show that 40 percent of the sample reveal time consistency ($\beta = 1$) and 44 percent present bias ($\beta < 1$).

The right column of Table 2 shows that, with the exception of long-run patience, none of the differences in economic preferences for the two outcome variables is statistically significant based on a non-parametric Wilcoxon rank-sum test. The difference in δ is significant on the 10 percent level for dropouts ($p = 0.06$) but fails to reach significance for finished in time ($p = 0.16$). In order to get a better grip on the role of students' time preferences in completing their vocational training program, it is necessary to control for the different preference measures simultaneously, as well as for other socio-economic and personality characteristics. We do so in the subsequent regression analyses. We first focus on program completion (Section 4.2) and then analyze labor market transition (Section 4.3).

4.2 Regression Analysis of Dropout and Finishing in Time

Result 1 (Dropout) *The stronger a student discounts the long-run future, the more likely he or she drops out of the vocational training program prior to its completion.*

Evidence for Result 1 is given in Table 3. All columns show marginal effects of logit regressions to explain whether a student terminated his or her program prior to successful completion. Standard errors are clustered at the class level, to account for potential correlation within classes. In columns (1)-(5), only the four preference measures are successively included as explanatory variables. It can be seen that a higher δ, which implies less discounting of the long-run future, is associated with a significantly lower dropout probability. This result holds even if δ is the sole regressor in column (1). Adding additional preference measures in columns (3)-(5) does not alter the point estimate and increases the significance of δ. A one standard-deviation increase in our measure of δ, which is equal to 0.15, is associated with a 2.55 percent lower probability of dropping out of the vocational training program.

A student's present bias may represent an additional source of discounting to future payments. Column (2) assesses the association between β and dropout without further controls. The coefficient on β is basically zero, and far from being statistically significant. This is also the case in columns (3)-(5), when we additionally control for our other preference measures. Risk and loss aversion are added in columns (4) and (5), respectively. None of them has any predictive power with regard to dropout. Finally, In columns (6)-(8), our additional controls, as explained in Section 3.2.3, are subsequently introduced into the regression to provide robustness checks for Result 1. The effect

Table 3: Marginal Effects of Logit regressions on Dropout

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (δ)	–0.17*		–0.17**	–0.18**	–0.17**	–0.12**	–0.16**	–0.15**
	(0.09)		(0.08)	(0.07)	(0.08)	(0.06)	(0.08)	(0.07)
Present Bias (β)		0.02	–0.02	–0.02	–0.02	–0.00	–0.03	–0.04
		(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.06)
Risk aversion				0.00	0.01	0.01	0.01	0.01
				(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Loss aversion					–0.02	–0.02	–0.02	–0.02
					(0.01)	(0.01)	(0.01)	(0.01)
Field of study, gender	No	No	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	No	No	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	No	No	No	No	Yes
Pseudo R^2	0.01	0.00	0.01	0.01	0.02	0.05	0.08	0.13
Observations	227	227	227	227	227	227	212	209

Logit Regressions on Dropout. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. A regression table including all coefficients is included in Table B.1. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (7) and (8). Std. Errors are clustered at the class level and are shown in parentheses. Significance levels: *** $p<.01$, ** $p<.05$, * $p<.1$.

of δ remains robust and significant in all these regressions.[16]

Our second outcome measure, whether or not students finish their program within the standard time, corroborates the above empirical pattern as the next result shows.

Result 2 (Finish in Time) *The more a student discounts the long-run future, the less likely he or she is to finish the vocational training program in time.*

Evidence for Result 2 is provided in Table 4. Again, all columns show marginal effects of logit regressions to explain whether a student finished his or her program in time, and standard errors are clustered at the class level. As before, we subsequently add our four preference measures in columns (1)-(5) as regressors. It can be seen that a higher δ is significantly associated with a higher probability of finishing in time.[17] A one standard deviation increase in our measure of δ is associated with a roughly 3.5 percent larger probability of finishing the training program in time. The coefficient on β is also positive, but much smaller and not statistically significant. When adding our additional sets of controls in columns (6)-(8), the coefficient estimates on δ remain constant and significant. Our measures of risk and loss aversion have, again, no predictive power with regard to finishing in time.[18]

In Appendix B and Appendix C, we provide a number of robustness checks for the above results. First, one might be worried that our results are biased due to a small number of clusters, since we sampled only from 14 classes. In Table B.2 and C.2, we provide results from linear probability models where we calculate p-values using the wild-t bootstrap procedure for linear regressions proposed by Cameron et al. (2008) to correct for the small number of clusters. As can be seen our results remain largely the same. Second, we consider firm characteristics. Unfortunately, we only possess self-reported data on firm characteristics collected in the second year of the program. Here, we were not able to get responses from all students, in particular from those who already dropped out. In the survey, we obtained data on whether students receive performance pay as well as on the size of the company. Table B.3 and C.3 report regression results using our initial regression specification (5) from Table 3 and 4, respectively, that contains all preference measures and add, one by one, indicators for performance pay as well as firm size. It can be seen that the effect of long-run patience on dropout remains constant and significant at the ten percent level. Specification (3) in Table B.3 and C.3 shows results using class fixed effects and robust standard errors. Again, the coefficient estimates remain roughly unchanged. Standard errors do get slightly bigger though, implying that the effect loses its significance in this specification. Next, we control for a student's potential present bias by modeling β as a dummy variable that equals one if $\beta < 1$ and zero otherwise (see model (4) in Table B.3 and model (5) in Table C.3). Alternatively, we

[16] Table B.1 in Appendix B presents the coefficients for all regressors added in columns (6)-(8). Results show that polytechnicians are less likely to drop out, which can be explained by selection into this particular field of study attracting on average better qualified students than the commercial or electrician program. Personality measures are, somewhat surprisingly, non-predictive for dropouts, while a higher CRT score actually is.

[17] The p-value on the marginal effect of δ in regression (1) is $p = 0.101$. The p-value on the respective logit coefficient is $p = 0.082$

[18] All regression coefficients for columns (6)-(8) are reported in Table C.1 in Appendix C.

Table 4: Marginal effects of Logit regressions on Finished in Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Long-Run Patience (δ)	0.22		0.24*	0.24**	0.23**	0.19**	0.20*	0.22**
	(0.13)		(0.13)	(0.12)	(0.12)	(0.09)	(0.11)	(0.11)
Present Bias (β)		0.01	0.06	0.06	0.05	0.04	0.09	0.11
		(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.08)	(0.08)
Risk aversion				0.00	–0.00	–0.01	–0.00	–0.01
				(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Loss aversion					0.01	0.02	0.03	0.02
					(0.03)	(0.02)	(0.02)	(0.02)
Field of study, gender	No	No	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	No	No	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	No	No	No	No	Yes
Pseudo R^2	0.01	0.00	0.01	0.01	0.01	0.08	0.17	0.18
Observations	227	227	227	227	227	227	212	209

Logit Regressions on Finished in Time. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Socioeconomic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. A regression table including all coefficients is included in Table C.1. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (7) and (8). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

exclude all students who are present biased and re-run the regression only with students with $\beta \geq 1$ (see model (5) in Table B.3 and model (6) in Table C.3). Neither of this affects our results with regard to long-run patience. Finally, as explained in Section 4.2, finished in time and dropout are outcome variables that vary in important dimensions, but nonetheless most dropouts are also coded as not finishing in time (cf. Table 1). Column (4) in table C.3 therefore reports results on finished in time when all dropouts are excluded. It can be seen that the coefficient on δ remains positive, but becomes smaller in size and as a result looses significance.

In summary, we find that an individual's long-run time preference is predictive of human capital acquisition such as completing or dropping out of an important post-secondary education program. A stronger discounting leads to an underestimation of the future benefits of acquiring human capital thereby increasing the likelihood of quitting education. In contrast, our results show that present bias and time inconsistency are only insignificant factors in explaining dropout behavior.

4.3 Labor Market Transition

Besides our register data on dropout and finishing in time, we collected survey data on students' labor market transition plans prior to the completion of their vocational training program. For this survey, we were able to collect replies from 196 students, of whom 181 were expected to finish, which corresponds to an 81% response rate.

Table 5: Job Market Outcomes and Plans

	Number	Percent
Definite Job Offer	92	51
Planning continued education	47	26
No Job Offer and no education plans	42	23
Number of Observations	181	100

Table 5 shows the distribution of answers to our survey. Table 6 provides information on our preference measures conditional on survey answers, excluding the 24 (out of the 181) students who gave inconsistent answers in our preference measures. As the first column in Table 6 shows, students in the reduced sample are similar to the overall sample in all four preference measures (cf. Table 2).[19] Further, similar to before students with inconsistent answers do not differ significantly from those with consistent answers in labor market transition outcomes (no plan: 22.2% vs. 22.5%; job offer: 48.2% vs. 52.1%; education: 29.6% vs. 25.4%; Fisher's Exact test: $p = 0.89$).

Different theories on job search make predictions as to how preferences should affect the probability to look for jobs and correspondingly to secure a job offer. First, DellaVigna and Paserman (2005) predict that present-biased individuals should be less likely to have definite job offers.

[19] In table A.2 in Appendix A, we directly compare the average preference measures of those students who responded to the labor market transition survey with those students who did not. As can be seen, while those who did respond on average have slightly higher patience and slightly less present bias, none of the observed differences are significant. Given the high response rate and the similarity of those who responded and those who did not, it appears that there is no significant selection into the labor market transition survey based on preference characteristics.

Table 6: Average Preference Measures, by Labor Market Transition

	All	No Plan	Job	Education
Patience (δ)	0.78	0.80	0.79	0.76
	(0.17)	(0.16)	(0.17)	(0.17)
Present Bias (β)	0.96	0.90	0.97	0.99
	(0.23)	(0.14)	(0.26)	(0.24)
Risk Aversion	5.09	5.55	5.06	4.73
	(1.69)	(1.66)	(1.71)	(1.63)
Loss Aversion	4.94	4.94	5.09	4.66
	(1.06)	(0.83)	(0.98)	(1.32)
Number of Observations	157	36	80	41

Note: Standard Errors in parentheses.

Present bias unambiguously leads individuals to postpone the costly activity of looking for jobs. Long-run discounting, on the other hand, has an ambiguous (or even positive) effect on the probability of having a job offer, since it primarily affects the reservation wage. Second, DellaVigna et al. (2017) propose a model of job search with reference-dependent preferences. This model predicts that loss averse individuals search harder when they face potential losses. In our setting, students face the potential of unemployment if they do not secure a job. Hence, if students have reference-dependent preferences, one would expect that more loss averse students invest more effort in job search and consequently are more likely to have a definite job offer.

Our preference measures, taken 3-4 years prior to the survey on labor market transition, provide an opportunity to test these predictions.[20] One complication of our data is that some student's have not finished accumulating human capital and are planning to continue their education, for example at a university. These students naturally do not have a job or a job offer, and are not searching for one. However, this is for obvious reasons that are not captured by the theories of job search mentioned before. To deal with this issue, we perform various types of analyses. First, we construct a dummy variable for having a "continuation plan", which equals 1 in case a student has either a job offer *or* plans to acquire additional education. Second, we conduct a multinomial logit regression that allows for multiple categorical outcomes. Finally, we look at students' explicit search activities as reported in the survey.

Result 3 (Labor Market Transition – Present Bias) *Present biased students are less likely to have a job offer or plans to continue education one month prior to finishing their vocational training program.*

Evidence for Result 3 is presented in Table 7 and Table 8. Table 7 shows logit regressions on

[20]Unfortunately, we do not have register data on whether or not students ultimately had jobs, and what kind of jobs, and we do not have register data on the continued education programs the students enrolled in, if they did so. Nonetheless, we believe that this cross-section close to the termination of the programs is informative with respect to the predictions of job search theories.

the dummy variable "continuation plan" introduced above. As can be seen, present bias β has a highly significant effect. A one standard deviation increase in β increases the probability of having a continuation plan by 6.2 to 8.7 percentage points. As columns (2)-(4) show, the effect of present bias is robust to the inclusion of our various sets of controls.[21]

Table 7: Logit Regressions on having a Continuation Plan

	(1)	(2)	(3)	(4)
Patience (δ)	–0.027	–0.003	–0.103	–0.072
	(0.168)	(0.186)	(0.248)	(0.248)
Present Bias (β)	0.297***	0.283***	0.363***	0.397***
	(0.091)	(0.091)	(0.087)	(0.121)
Risk aversion	–0.044*	–0.048*	–0.048*	–0.039
	(0.024)	(0.026)	(0.029)	(0.029)
Loss aversion	0.032	0.032	0.048*	0.040
	(0.027)	(0.033)	(0.029)	(0.034)
Field of study, gender	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	Yes
Pseudo R^2	0.046	0.057	0.165	0.211
Observations	157	157	146	145

Logit Regressions on students' continuation plans. The outcome variable is a dummy that takes value 1 in case a student has a job offer or indicated that he/she plans to attend a continued education program. Marginal Effects are shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. A regression table including all coefficients is included in table D.1 in Appendix D. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (3) and (4). Std. Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Next, Table 8 provides results from multinomial logit regressions with the outcomes (i) having a definite job offer, (ii) planning continued education or (iii) having neither. The baseline outcome in all regressions is having neither a job offer nor plans for continued education. Specification (1) only includes the four preference measures and standard errors are clustered at the class level. Columns (2)-(4) again subsequently add our various sets of controls. As can be seen, in all regression models lower present bias (higher β) significantly increases the probability of both, having a job or planning continued education, relative to the baseline of having neither of them. The marginal effect of present bias on having a job offer is sizeable. A one standard deviation increase in β (i.e. a reduction in present bias) makes it between 13-18 percentage points more likely for a student to have a definite job offer rather than having no continuation plan. Discounting, on the other hand, has no significant effect in either direction in our data.

Since some of the students with a definite job offer may have received the offer from the company they did the vocational training program with, it may be unclear to what extent these students

[21] See Table D.1 for coefficients of all control variables.

Table 8: Multinomial Logit Regressions on Job Offer and Continued Education Plans

	(1)		(2)		(3)		(4)	
	Job	Educ	Job	Educ	Job	Educ	Job	Educ
Patience (δ)	0.128	–0.660	0.210	–0.379	–0.565	–1.015	–0.224	–1.163
	(1.226)	(1.076)	(1.348)	(1.216)	(2.021)	(1.897)	(2.209)	(1.951)
Present Bias (β)	1.805***	1.770**	1.740***	1.741**	2.684***	2.955***	3.191**	3.505**
	(0.547)	(0.846)	(0.530)	(0.846)	(0.768)	(0.868)	(1.284)	(1.393)
Loss aversion	0.328	–0.038	0.312	0.004	0.494**	0.148	0.450*	0.089
	(0.214)	(0.252)	(0.234)	(0.305)	(0.220)	(0.317)	(0.261)	(0.340)
Risk aversion	–0.261	–0.276**	–0.269	–0.331**	–0.331	–0.415**	–0.279	–0.379**
	(0.164)	(0.137)	(0.178)	(0.144)	(0.230)	(0.184)	(0.249)	(0.177)
Constant	–1.252	0.584	–0.531	0.323	–8.961	–3.262	–13.078*	–6.951
	(1.203)	(0.865)	(1.493)	(1.421)	(5.965)	(8.868)	(7.172)	(9.345)
Field of study, gender	No		Yes		Yes		Yes	
Socioeconomic Controls	No		No		Yes		Yes	
BIG 5, GRIT	No		No		No		Yes	
Pseudo R^2	0.040		0.059		0.133		0.182	
Observations	157		157		146		145	

Multinomial Logit Regressions on *plans after graduation*. Job denotes having a definite job offer. Educ indicates a plan to enroll in a continued education program. The table shows the coefficients of the multinomial logit regressions. Regressions only contains students with consistent answers and students who are expected to finish their education program in the year of the survey. Socio-economic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in regression specifications (3) and (4). Standard Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

actually exerted any job search effort themselves. In Table 9 we therefore provide results from a sub-sample analysis of students' self-reported search activity excluding all students who have a definite job offer *and* report that they have not actively searched themselves. The outcome variable in this logit regression is not whether a student has a job offer, but whether a student has searched actively or not. The sample is further reduced because we exclude those students who plan to continue their education, which similarly eliminates the need to search for a job. Our sample is therefore reduced to those students who actually have incentives to engage in job search. As can be seen in Table 9, the results confirm our findings from above: students with a higher β are significantly more likely to have searched actively, the result being robust to the inclusion of controls (columns (1)-(4)). A one standard deviation increase in β (a decrease in present bias) makes is around 10 percentage points more likely that a student has already actively searched for a job, one month prior to the termination of the vocational training program.

Table 9: Logit Regressions on active Job Search

	(1)	(2)	(3)	(4)
Patience (δ)	0.000	0.095	0.182	0.153
	(0.295)	(0.267)	(0.284)	(0.266)
Present Bias (β)	0.358**	0.397**	0.431***	0.405**
	(0.182)	(0.172)	(0.165)	(0.197)
Risk aversion	–0.025	–0.033	–0.057*	–0.042*
	(0.027)	(0.027)	(0.033)	(0.024)
Loss aversion	0.044	0.078*	0.102	0.047
	(0.040)	(0.045)	(0.072)	(0.071)
Field of study, gender	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	Yes
Pseudo R^2	0.045	0.121	0.167	0.298
Observations	90	90	84	84

Logit Regressions on whether a student actively searched for a job, either internally or externally. Students who indicated that they have a job offer but haven't actively searched for it are excluded from the analysis, since they had no need to search. For the same reason, students who indicate that they continue their education are excluded. Marginal Effects are shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (3) and (4). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Finally, turning to the second hypothesis relating to loss aversion and reference-dependent preferences, we find the following:

Result 4 (Labor Market Transition – Loss Aversion) *Loss averse students have a higher likelihood of having secured a job offer.*

First, looking again at Table 7, we see that loss aversion is indeed positively associated with having a continuation plan. This effect is also statistically significant at the ten percent level once we control for field of study and socioeconomic variables (column (3)). The multinomial logit regressions in table 8 show that loss aversion in particular predicts a higher probability of having a definite job offer. Again, once controlling for socio-economic characteristics (columns (3) and (4)), the effect becomes significant. If we control also for average school grades in the final year (Table D.4), the effect even increases. The implied magnitude of the effect is again large. A one standard deviation increase in loss aversion makes it between 11-15 percentage points more likely for a student to have a definite job offer rather than having no continuation plan. Interestingly, loss aversion does not predict continued education plans. This is not surprising and consistent with the theory, since continued education exposes the student to additional risk — future employment remains unclear — and may come at an immediate loss, since remuneration is usually considerably lower during education programs compared to employment. Consistent with this view, it is noteworthy that less risk averse students are actually significantly more likely to plan continued education, relative to the baseline.

Finally, we can assess the effect of loss aversion on reported job search activities in table 9. The marginal effect is positive in all four regression specifications though reaching only marginal significance in one of the specifications.

In Appendix D, we again provide additional robustness checks for all results presented in this section. First, we replicate the result on continuation plans based on a linear probability model using the wild-t bootstrap procedure (Table D.2). Next, in table D.3 we include firm characteristics (column 1 and 2) and use class fixed effects instead of clustering at the class level to control for potential dependencies (column 3). The results show that present bias remains a significant predictor of having a continuation plan. An additional important control for labor market transition are the final grades in the education program. Since students had not completed the program, their final grades were not yet definite, but we obtained their current grade point average (this was also the best available information for potential employers at that time) and add it as a control in column (4). It can be seen that, while better grades are associated with a higher probability of having a continuation plan, their inclusion in the model does not change the predictive power of present bias for having a definite job offer. Finally, we also add the average school grades in the final year to the multinomial logit regression (Table D.4), and our results remain robust.[22]

In summary, we find strong evidence that present biased students are less likely to have a definite job offer one month prior to the termination of their vocational training program. In line with the theory of DellaVigna and Paserman (2005), we also find that the likely cause of these differences is due to reduces search activity among present-biased students. Second, we find evidence that more loss averse students are more likely to have a definite job offer. This is consistent with DellaVigna et

[22]The same type of robustness checks are also performed for our regressions on job search. Table D.6 reports results of logit regressions on job search based on a linear probability model using the wild-t bootstrap procedure. Table D.7 includes firm characteristics (columns 1 and 2), uses class fixed effects instead of clustering at the class level to control for potential dependencies (column 3) and includes the grade point average in the final year as an additional control (column 4). Again, results on present bias remain robust and significant.

al. (2017) who show that loss aversion increases effort in job search when individuals have reference-dependent preferences. However, when considering the direct effect of loss aversion on job search activity in our (at that point admittedly reduced) sample, the effect is directionally consistent with the theory, but statistical significance is weak.

5 Conclusion

In this paper, we analyze a dataset that is unique in it's combination of incentivized elicited preference data, real world register data on termination of upper-secondary education programs and survey data on labor market transition. We find that students' long-run patience is a significant predictor of dropout decisions, whereas the impact of present bias is small and insignificant. This finding has important implications for policies aiming at increasing human capital and reducing the costs associated with dropping out of voluntary education programs. The result suggests that any effective policy with this aim should focus primarily on factors that influence students' long-run patience rather than potential present bias and time inconsistency. Policies that have proven successful in this respect emphasize the importance of early childhood interventions (Cunha and Heckman, 2007; Cunha et al., 2010; Falk and Kosse, 2016). To the extent that students acquire important new information also over the course of the program whether the chosen education fits their ability and preferences (Manski, 1989; Stinebrickner and Stinebrickner, 2012), limiting the possibility to change education after students have become enrolled (in order to make dropouts more difficult) may actually be harmful. Instead reducing informational frictions prior to enrollment, e.g. by offering trial courses and open house days, should prove beneficial (Arcidiacono et al., 2016). Further, our data sheds new light on recent theories of job search (DellaVigna and Paserman, 2005; DellaVigna et al., 2017). We find that present bias as well as loss aversion are significantly associated with the probability of having a definite job offer about a month prior to completion of the education program. First, stronger present bias is associated with a lower probability of having a job, consistent with theories of impatience and job search. Second, a stronger degree of loss aversion is associated with a higher probability of having a definite job offer, consistent with theories of reference-dependent preferences and job search. These results are informative for policies aiming at the transition from education to employment and reducing youth-unemployment. In particular, they suggest that towards the end of the education program commitment devices that increase the difficulty to procrastinate on securing future employment may well be effective and also welfare improving.

References

Altonji, Joseph G., "The Demand for and return to education when education outcomes are uncertain," *Journal of Labor Economics*, 1993, pp. 48–83.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Eliciting risk and time preferences," *Econometrica*, May 2008, *76* (3), 583–618.

Andreoni, James and Charles D. Sprenger, "Estimating time preferences from convex budgets," *American Economic Review*, December 2012, *102* (7), 3333–3356.

_ and _ , "Risk preferences are not time preferences," *American Economic Review*, December 2012, *102* (7), 3357–3376.

_ , Michael A. Kuhn, and Charles D. Sprenger, "Measuring time preferences: A comparison of experimental methods," *Journal of Economic Behavior & Organization*, August 2015, *116*, 451–464.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College attrition and the dynamics of information revelation," *Working Paper*, 2016.

Augenblick, Ned, Muriel Niederle, and Charles D. Sprenger, "Working over time: Dynamic inconsistency in real effort tasks," *Quarterly Journal of Economics*, July 2015, *130* (3), 1067–1115.

Balakrishnan, Uttara, Johannes Haushofer, and Pamela Jakiela, "How soon is now? Evidence of present bias from convex time budget experiments," *IZA Discussion Paper*, December 2015.

Becker, Gary S., "Investment in human capital: A theoretical analysis," *Journal of Political Economy*, 1962, pp. 9–49.

Burks, Stephen, Jeffrey Carpenter, Lorenz Götte, and Aldo Rustichini, "Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment," *Journal of Economic Behavior & Organization*, 2012, *84* (1), 308–320.

Cadena, Brian C. and Benjamin J. Keys, "Human capital and the lifetime costs of impatience," *American Economic Journal: Economic Policy*, 2015, *7* (3), 126–153.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-based improvements for inference with clustered errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Castillo, Marco, Paul J. Ferraro, Jeffrey L. Jordan, and Ragan Petrie, "The today and tomorrow of kids: Time preferences and educational outcomes of children," *Journal of Public Economics*, 2011, *95* (11), 1377–1385.

Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathon P. Schuldt, and Dmitry Taubinsky, "Individual laboratory-measured discount rates predict field behavior," *Journal of Risk and Uncertainty*, 2008, *37* (2-3), 237–269.

Cunha, Flavio and James J. Heckman, "The technology of skill formation," *American Economic Review*, April 2007, *97* (2), 31–47.

_ , James J Heckman, and Susanne M Schennach, "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 2010, *78* (3), 883–931.

DellaVigna, Stefano and M. Daniele Paserman, "Job search and impatience," *Journal of Labor Economics*, 2005, *23* (3), 527–588.

_ , Attila Lindner, Balázs Reizer, and Johannes F. Schmieder, "Reference-dependent job search: Evidence from Hungary," *Quarterly Journal of Economics*, 2017, *132* (4), 1969–2018.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Are risk aversion and impatience related to cognitive ability?," *American Economic Review*, 2010, *100*, 1238–1260.

_ , _ , _ , and _ , "The robustness and pervasiveness of sub-additivity in intertemporal choice," *Working Paper*, 2017.

Duckworth, Angela L., Christopher Peterson, Michael D. Matthews, and Dennis R. Kelly, "Grit: perseverance and passion for long-term goals.," *Journal of personality and social psychology*, 2007, *92* (6), 1087.

Eckstein, Zvi and Kenneth I. Wolpin, "Why youths drop out of high school: The impact of preferences, opportunities, and abilities," *Econometrica*, 1999, *67* (6), 1295–1339.

Falk, Armin and Fabian Kosse, "Early childhood environment, breastfeeding and the formation of preferences," *Working Paper*, December 2016.

Frederick, Shane, "Cognitive reflection and decision making," *Journal of Economic perspectives*, 2005, pp. 25–42.

Gerlitz, Jean-Yves and Jürgen Schupp, "Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP," *DIW Research Notes*, 2005, *4*.

Golsteyn, Bart H.H., Hans Grönqvist, and Lena Lindahl, "Adolescent time preferences predict lifetime outcomes," *Economic Journal*, 2013, pp. 1–23.

Halima, Bassem Ben and Mohamed Ali Ben Halima, "Time preferences and job search: Evidence from France," *Labour*, 2009, *23* (3), 535–558.

Laibson, David, "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 1997, pp. 443–477.

Lindahl, Mikael and Alan B. Krueger, "Education for growth: Why and for whom?," *Journal of Economic Literature*, 2001, *39* (4), 1101–1136.

Maghsoodi, Eiman and Irene Kriesi, "Wiedereinstieg und Anschlusslösung nach einer Lehrvertragsauflösung im Kanton Zürich: Analyse der Lehrvertragsauflösungen der Jahre 2008 und 2009," *Eidgenössisches Hochschulinstitut für Berufsbildung*, 2013.

Manski, Charles F., "Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon," *Economics of Education Review*, 1989, *8* (4), 305–312.

Meier, Stephan and Charles D. Sprenger, "Present-biased preferences and credit card borrowing," *American Economic Journal: Applied Economics*, 2010, pp. 193–210.

_ **and** _ , "Time discounting predicts creditworthiness," *Psychological Science*, 2012, *23* (1), 56–58.

_ **and** _ , "Temporal stability of time preferences," *Review of Economics and Statistics*, May 2015, *97* (2), 273–286.

Mincer, Jacob, "Investment in human capital and personal income distribution," *Journal of political economy*, 1958, *66* (4), 281–302.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

OPET, "Vocational and Professional Education and Training in Switzerland," *Federal Office for Professional Education and Technology Bern*, 2011.

Oreopoulos, Philip, "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling," *Journal of public Economics*, 2007, *91* (11), 2213–2229.

Oswald, Yvonne and Uschi Backes-Gellner, "Learning for a bonus: How financial incentives interact with preferences," *Journal of Public Economics*, 2014, *118*, 52–61.

Schultz, Theodore W., "Investment in human capital," *American economic review*, 1961, pp. 1–17.

Stinebrickner, Ralph and Todd Stinebrickner, "A major in science? Initial beliefs and final outcomes for college major and dropout," *Review of Economic Studies*, 2013, p. rdt025.

Stinebrickner, Todd and Ralph Stinebrickner, "Learning about academic ability and the college dropout decision," *Journal of Labor Economics*, 2012, *30* (4), 707–748.

Sutter, Matthias, Martin G. Kocher, Daniela Glatzle-Rutzler, and Stefan Trautmann, "Impatience and uncertainty: Experimental decisions predict adolescents' field behavior," *American Economic Review*, 2013, *103* (1), 510–531.

Zafar, Basit, "How do college students form expectations?," *Journal of Labor Economics*, 2011, *29* (2), 301–348.

Appendix A Summary Statistics and Distributions of Preference Parameters

Table A.1: Summary statistics of the control variables

Variable	Mean	Std. Dev.	N
Female	0.40	0.49	231
Business	0.62	0.49	231
Polytechnician	0.18	0.38	231
Electrician	0.20	0.40	231
native German speaker	0.84	0.36	231
Age	16.36	0.93	231
Math grade	4.83	0.64	222
German Grade	4.77	0.43	223
English Grade	4.90	0.61	222
Difficulty to borrow CHF 100	3.94	0.96	231
CRT score	0.84	0.94	231
Openness	13.92	3.26	230
Conscientiousness	14.88	3.22	230
Extraversion	15.48	3.76	230
Agreeableness	10.40	2.36	231
Neuroticism	11.98	3.56	231
GRIT	3.36	0.51	231
Education Mother	2.31	1.12	229
Education Father	2.75	1.23	227
$>$ 100 Employees	0.53	0.50	209
Performance pay	0.27	0.44	211

Note: $Female$ is a dummy indication female gender. *Business*, *Polytechnician* and *Electrician* are dummies indicating the field of study. *nativeGermanspeaker* is a dummy for native German speakers. $Math$, *German* and *English* Grades are measures on a scale from 1 (worst) to 6 (best). $difficulty\ to\ borrow\ CHF100$ is measured on a 5 point Likert scale, 5 indicating the least difficulty. *Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* and $Neuroticism$ are the values from the 15-item version of the BIG 5 questionnaire. *Education mother* and $Education\ father$ indicate the respective education levels, 1 indicating "compulsory school or less", and 5 indicating "university". > 100 *Employees* is a dummy for firms having more than 100 employees. $Performance\ pay$ is a dummy indicating performance pay.

Figure A.1: Distribution of Patience (δ)

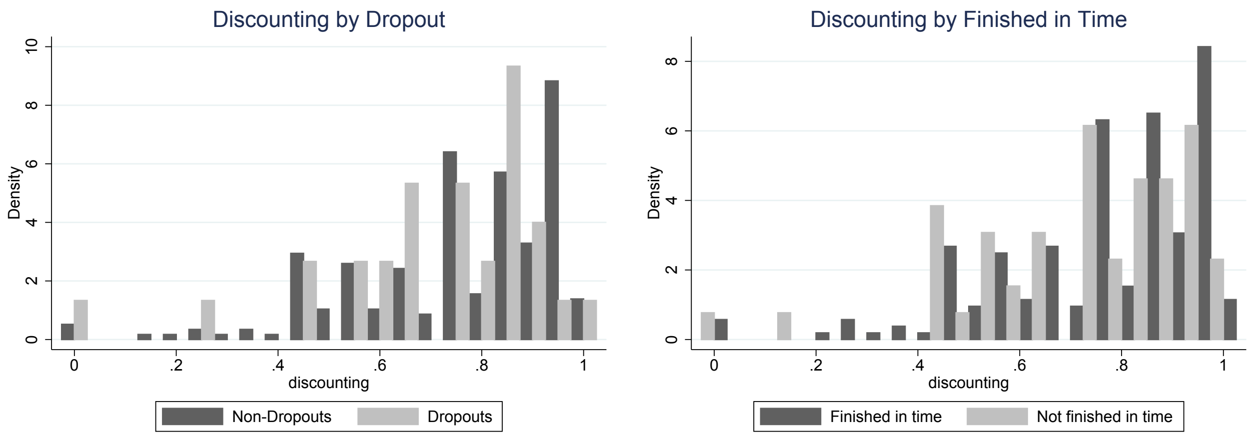

Figure A.2: Distribution of Present Bias (β)

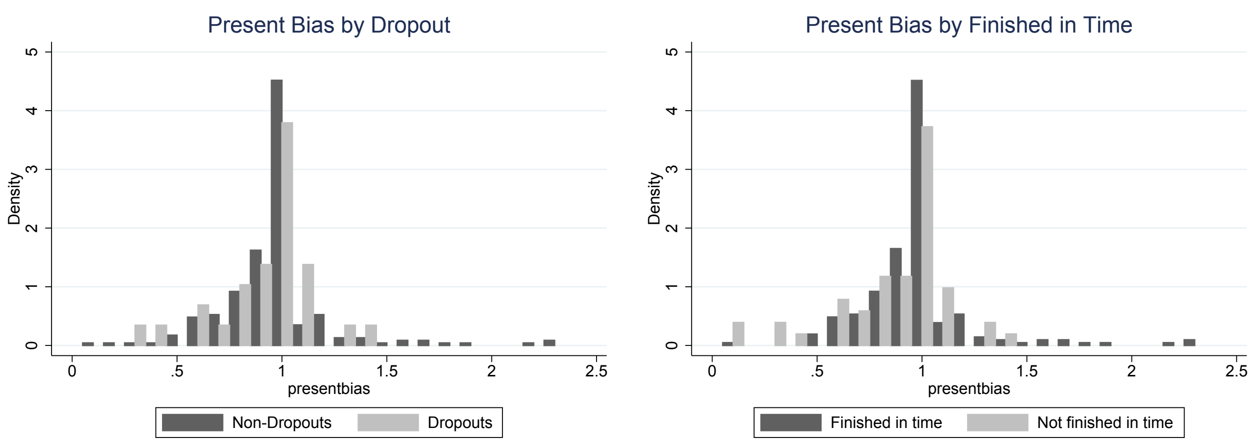

Figure A.3: Distribution of Risk Aversion

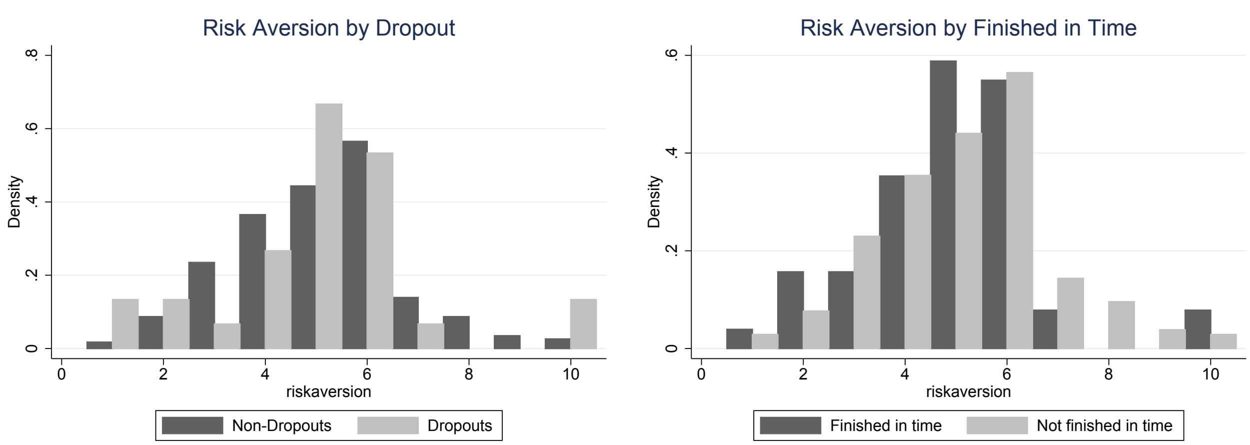

Figure A.4: Distribution of Loss Aversion

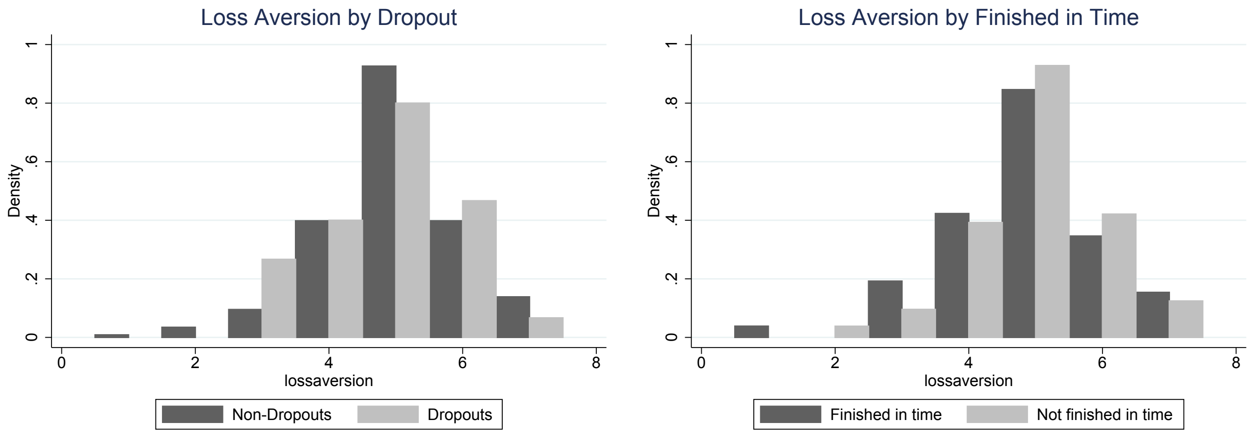

Table A.2: Summary Statistics and Selection in the Labor Market Sample

	Labor Market Sample	Non-Replies	p-value
Long-run Patience (δ)	0.786	0.766	0.49
	(0.162)	(0.172)	
Present Bias (β)	0.961	0.912	0.53
	(0.230)	(0.203)	
Risk Aversion	5.036	4.967	0.67
	(1.675)	(1.737)	
Loss Aversion	4.922	4.950	0.92
	(1.070)	(1.016)	
Number of Obs.	167	60	

Note: Only students with consistent answers in the preference measurements are included. Standard Errors in parentheses. Column p shows the p-value of a Wilcoxon rank-sum test comparing preference values of those included in the labor market sample with those who did not reply.

Appendix B Additional Regressions on Dropout

Table B.1: Marginal effects of a logit regression on Dropout

	(1)	(2)	(3)
Patience (δ)	–0.12**	–0.16**	–0.15**
	(0.06)	(0.08)	(0.07)
Present Bias (β)	–0.00	–0.03	–0.04
	(0.07)	(0.08)	(0.06)
Risk aversion	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)
Loss aversion	–0.02	–0.02	–0.02
	(0.01)	(0.01)	(0.01)
female (d)	–0.00	0.01	–0.02
	(0.05)	(0.05)	(0.05)
Business (d)	–0.11	–0.07	–0.06
	(0.07)	(0.06)	(0.07)
Polytechnician (d)	–0.09***	–0.08***	–0.08***
	(0.03)	(0.02)	(0.02)
german (d)		–0.02	–0.02
		(0.04)	(0.05)
age		–0.01	–0.01
		(0.03)	(0.02)
Math grade		–0.02	–0.01
		(0.03)	(0.03)
German grade		–0.01	–0.00
		(0.04)	(0.04)
English grade		–0.03	–0.02
		(0.03)	(0.02)
borrowing difficulty		0.01	0.01
		(0.02)	(0.02)
CRT score		0.03	0.03**
		(0.02)	(0.01)
Education Mother		–0.01	–0.01
		(0.02)	(0.02)
Education Father		0.02	0.01
		(0.02)	(0.02)
Openness			0.01
			(0.01)
Conscientiousness			–0.01
			(0.01)
Extroversion			–0.00
			(0.00)
Agreeableness			0.00
			(0.01)
Neuroticism			0.00
			(0.00)
GRIT			0.00
			(0.03)
Pseudo R^2	0.05	0.08	0.13
Observations	227	212	209

Logit Regressions on Dropout. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (2) and (3). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table B.2: OLS Regressions on Dropout: Wild Cluster Bootstrap

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (δ)	–0.18		–0.19*	–0.19*	–0.19*	–0.14*	–0.18*	–0.19**
	(0.11)		(0.06)	(0.06)	(0.06)	(0.08)	(0.07)	(0.04)
Present Bias (β)		0.02	–0.02	–0.02	–0.02	–0.01	–0.04	–0.05
		(0.84)	(0.80)	(0.80)	(0.84)	(0.91)	(0.68)	(0.53)
Risk aversion					0.01	0.01	0.01	0.01
					(0.69)	(0.51)	(0.38)	(0.46)
Loss aversion					–0.02	–0.02	–0.02	–0.02
					(0.30)	(0.22)	(0.26)	(0.27)
Constant	0.26**	0.10	0.28**	0.28**	0.33**	0.41***	0.73	0.62
	(0.01)	(0.32)	(0.01)	(0.01)	(0.03)	(0.01)	(0.12)	(0.23)
Field of study, gender	No	No	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	No	No	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	No	No	No	No	Yes
R^2	0.01	0.00	0.01	0.01	0.01	0.04	0.05	0.09
Observations	227	227	227	227	227	227	212	209

OLS Regressions on Dropout. Regressions only contains subjects with consistent answers in the choice experiments. The sets of controls are explained in more detail in section 3.2.3. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (7) and (8). The p-values clustering on 14 classes are shown in parentheses. To account for the low number of clusters we apply a wild cluster bootstrap-t procedure (Cameron et al., 2008) to determine the p-values, with 5000 repetitions. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table B.3: Marginal effects of Logit regressions on Dropout

	Firm Controls		Class F.E.	Alternative β	
	(1)	(2)	(3)	(4)	(5)
Patience (δ)	–0.16*	–0.16**	–0.17	–0.15*	–0.26**
	(0.08)	(0.07)	(0.13)	(0.09)	(0.12)
Present Bias (β)	–0.02	–0.01	–0.07		
	(0.07)	(0.07)	(0.09)		
Present Bias (d)				–0.03	
				(0.04)	
Risk aversion	0.00	0.01	0.01	0.01	–0.00
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Loss aversion	–0.01	–0.02	–0.02	–0.02	–0.03
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
no pay-for-performance (d)	–0.07				
	(0.07)				
pay-for-performance (d)	–0.08**				
	(0.04)				
<100 Emp. (d)		–0.07*			
		(0.04)			
>100 Emp. (d)		–0.11**			
		(0.05)			
Field of study	No	No	No	No	No
Socioeconomic Controls	No	No	No	No	No
BIG 5, GRIT	No	No	No	No	No
Pseudo R^2	0.03	0.04	0.08	0.02	0.05
Observations	227	227	188	227	127

Logit Regressions on Dropout. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Columns (1) and (2) add different firm size effects.[a] Column (3) controls for class fixed effects instead of clustering SE's at the class level, which also capture the field of study. Column (4) includes a binary indicator of present bias (= 1 if $\beta < 1$ instead of the estimated value of β. Column (5) only includes non-present biased subjects ($\beta \geq 1$). Std. Errors clustered at the class level shown in parentheses, except for column (3) which shows robust standard errors. Significance levels: *** p<.01, ** p<.05, * p<.1.

[a]In order not to lose observations, we create categorical variables that indicate whether a student receives performance pay, does not receive performance pay, or did not provide any information; similarly, whether the firm has more than 100 employees, less than 100 employees, or no information. The no-information categories are obviously highly correlated, though not perfect since some students answered one but not both questions. However, the correlation of regressors increases standard errors. We therefore include one of the two firm characteristics at a time.

Appendix C Additional Regressions on Finished in Time

Table C.1: Marginal effects of Logit regressions on Finished in Time

	(1)	(2)	(3)
Patience (δ)	0.19**	0.20*	0.22**
	(0.09)	(0.11)	(0.11)
Present Bias (β)	0.04	0.09	0.11
	(0.09)	(0.08)	(0.08)
Risk aversion	–0.01	–0.00	–0.01
	(0.02)	(0.02)	(0.02)
Loss aversion	0.02	0.03	0.02
	(0.02)	(0.02)	(0.02)
female (d)	0.10	0.09	0.08
	(0.07)	(0.06)	(0.07)
Business (d)	0.12**	0.08	0.07
	(0.06)	(0.08)	(0.08)
Polytechnician (d)	0.18***	0.16***	0.16***
	(0.04)	(0.03)	(0.03)
german (d)		–0.12**	–0.10*
		(0.05)	(0.05)
age		–0.01	–0.01
		(0.03)	(0.03)
Math grade		–0.02	–0.04
		(0.04)	(0.05)
German grade		–0.03	–0.03
		(0.07)	(0.06)
English grade		0.13***	0.13**
		(0.05)	(0.05)
borrowing difficulty		–0.01	–0.01
		(0.02)	(0.02)
CRT score		–0.01	–0.02
		(0.02)	(0.02)
Education Mother		0.03	0.04**
		(0.02)	(0.02)
Education Father		–0.01	–0.02
		(0.02)	(0.03)
Openness			0.00
			(0.01)
Conscientiousness			0.00
			(0.01)
Extroversion			–0.00
			(0.01)
Agreeableness			0.01
			(0.01)
Neuroticism			–0.00
			(0.01)
GRIT			0.06
			(0.04)
Pseudo R^2	0.08	0.17	0.18
Observations	227	212	209

Logit Regressions on Finished in Time. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (2) and (3). Std. Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table C.2: OLS Regressions on Finished in Time: Wild Cluster Bootstrap-t procedure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (δ)	0.23		0.25*	0.25*	0.24*	0.20*	0.23	0.25*
	(0.14)		(0.10)	(0.10)	(0.08)	(0.08)	(0.11)	(0.10)
Present Bias (β)		0.01	0.06	0.06	0.06	0.04	0.09	0.09
		(0.92)	(0.57)	(0.57)	(0.58)	(0.65)	(0.37)	(0.27)
Risk aversion					–0.00	–0.01	–0.01	–0.01
					(0.92)	(0.59)	(0.70)	(0.70)
Loss aversion					0.01	0.02	0.03	0.02
					(0.71)	(0.43)	(0.25)	(0.52)
Constant	0.62***	0.79***	0.55**	0.55**	0.51*	0.36*	0.20	0.17
	(< 0.00)	(0.00)	(0.02)	(0.02)	(0.07)	(0.08)	(0.73)	(0.75)
Field of study, gender	No	No	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	No	No	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	No	No	No	No	Yes
R^2	0.01	0.00	0.01	0.01	0.01	0.08	0.15	0.17
Observations	227	227	227	227	227	227	212	209

OLS Regressions on Finished in Time. Regressions only contains subjects with consistent answers in the choice experiments. The sets of controls are explained in more detail in section 3.2.3. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (7) and (8). The p-values clustering on 14 classes are shown in parentheses. To account for the low number of clusters we apply a wild cluster bootstrap-t procedure (Cameron et al., 2008) to determine the p-values, with 5000 repetitions. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table C.3: Marginal effects of Logit regressions on Finished in Time

	Firm Controls		Class F.E.	excl. dropouts	Alternative β	
	(1)	(2)	(3)	(4)	(5)	(6)
Patience (δ)	0.23*	0.22*	0.18	0.12	0.21*	0.22
	(0.13)	(0.11)	(0.19)	(0.09)	(0.12)	(0.18)
Present Bias (β)	0.05	0.05	0.07	0.05		
	(0.09)	(0.09)	(0.12)	(0.08)		
Present Bias (d)					0.01	
					(0.05)	
Risk aversion	–0.00	–0.00	–0.01	0.02	–0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Loss aversion	0.01	0.01	0.03	0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.04)
no pay-for-performance (d)	0.10					
	(0.11)					
pay-for-performance (d)	0.10					
	(0.09)					
<100 Emp. (d)		0.07				
		(0.08)				
>100 Emp. (d)		0.15*				
		(0.09)				
Field of study	No	No	No	No	No	No
Socioeconomic Controls	No	No	No	No	No	No
BIG 5, GRIT	No	No	No	No	No	No
Pseudo R^2	0.02	0.03	0.10	0.02	0.01	0.01
Observations	227	227	200	201	227	127

Logit Regressions on finished in time. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Columns (1) and (2) add different firm size effects. Column (3) controls for class fixed effects instead of clustering SE's at the class level. Column (4) excludes all observations that dropped out. Column (5) includes a binary indicator of present bias (= 1 if $\beta < 1$ instead of the estimated value of β. Column (6) only includes non-present biased subjects ($\beta \geq 1$). Std. Errors clustered at the class level shown in parentheses, except for column (3) which shows robust standard errors. Significance levels: *** p<.01, ** p<.05, * p<.1.

Appendix D Additional Regressions on Continuation Plans

Table D.1: Logit Regressions on Continuation Plans

	(1)	(2)	(3)
Patience (δ)	–0.003	–0.103	–0.072
	(0.186)	(0.248)	(0.248)
Present Bias (β)	0.283***	0.363***	0.397***
	(0.091)	(0.087)	(0.121)
Risk aversion	–0.048*	–0.048*	–0.039
	(0.026)	(0.029)	(0.029)
Loss aversion	0.032	0.048*	0.040
	(0.033)	(0.029)	(0.034)
business (d)	–0.090	0.053	0.040
	(0.103)	(0.077)	(0.095)
polytech (d)	–0.165	–0.019	0.001
	(0.140)	(0.079)	(0.095)
female (d)	0.046	–0.034	–0.021
	(0.069)	(0.050)	(0.066)
german (d)		–0.006	0.013
		(0.091)	(0.084)
age		0.033	0.028
		(0.042)	(0.042)
grade 2009 math		0.027	0.013
		(0.045)	(0.047)
grade 2009 german		–0.065	–0.068
		(0.073)	(0.070)
grade 2009 english		0.073	0.071
		(0.048)	(0.055)
difficulty to save 100 SFR		–0.009	0.001
		(0.022)	(0.023)
cognitive reflection test		–0.094***	–0.084**
		(0.032)	(0.036)
Education mother		0.056	0.041
		(0.037)	(0.040)
Education father		0.024	0.017
		(0.032)	(0.030)
Openness			0.006
			(0.010)
conscientiousness			0.016
			(0.014)
extraversion			0.015
			(0.010)
agreeableness			–0.010
			(0.016)
neuroticism			0.002
			(0.011)
GRIT			0.024
			(0.069)
Pseudo R^2	0.057	0.165	0.211
Observations	157	146	145

Logit Regressions on various continuation plans. Columns (1)-(3) are equivalent to columns (2)-(4) in table 7 and display regressions on a dummy that takes value 1 in case a student has a job offer or indicated that he/she plans to attend a continued education program. Marginal Effects are shown. Regressions only contain subjects with consistent answers in the choice experiments. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (2) and (3). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.2: OLS Regressions on Continuation Plans: Wild Cluster Bootstrap-t procedure

	(1)	(2)	(3)	(4)
Patience (δ)	0.01	0.03	–0.06	–0.02
	(0.95)	(0.86)	(0.86)	(0.95)
Present Bias (β)	0.25***	0.24***	0.25***	0.28***
	(< 0.01)	(< 0.01)	(0.01)	(< 0.01)
Risk aversion	–0.05*	–0.05*	–0.05	–0.04
	(0.06)	(0.06)	(0.15)	(0.29)
Loss aversion	0.03	0.03	0.05	0.04
	(0.32)	(0.39)	(0.12)	(0.33)
Constant	0.60***	0.67***	–0.28	–0.55
	(< 0.01)	(< 0.01)	(0.80)	(0.64)
Field of study, gender	No	Yes	Yes	Yes
Socioeconomic Controls	No	No	Yes	Yes
BIG 5, GRIT	No	No	No	Yes
R^2	0.05	0.06	0.15	0.20
Observations	157	157	146	145

OLS Regressions on various continuation plans. Columns (1)-(4) display regressions on a dummy that takes value 1 in case a student has a job offer or indicated that he/she plans to attend a continued education program. Regressions only contain subjects with consistent answers in the choice experiments. The sets of controls are explained in more detail in section 3.2.3. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (3)-(4). The p-values clustering on 14 classes are shown in parentheses. To account for the low number of clusters we apply a wild cluster bootstrap-t procedure (Cameron et al., 2008) to determine the p-values, with 5000 repetitions. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.3: Marginal effects of Logit regressions on Continuation Plans

	Firm Controls		Class F.E.	Final GPA
	(1)	(2)	(3)	(4)
Patience (δ)	0.05	–0.03	–0.04	–0.06
	(0.17)	(0.16)	(0.23)	(0.26)
Present Bias (β)	0.26**	0.30***	0.32**	0.39**
	(0.10)	(0.10)	(0.15)	(0.17)
Risk aversion	–0.05**	–0.04*	–0.05**	–0.04*
	(0.03)	(0.02)	(0.02)	(0.02)
Loss aversion	0.03	0.03	0.04	0.06
	(0.03)	(0.03)	(0.04)	(0.04)
no pay-for-performance (d)	0.06			
	(0.15)			
pay-for-performance (d)	–0.14			
	(0.15)			
<100 Emp. (d)		0.16		
		(0.11)		
>100 Emp. (d)		0.01		
		(0.10)		
Grade Point Average				0.16
				(0.11)
Field of study, gender	No	No	No	No
Socioeconomic Controls	No	No	No	No
BIG 5, GRIT	No	No	No	No
Pseudo R^2	0.09	0.08	0.10	0.08
Observations	157	157	151	137

Logit Regressions on a dummy that takes value 1 in case a student has a job offer or indicated that he/she plans to attend a continued education program. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Columns (1) and (2) add different firm size controls. Column (3) controls for class fixed effects instead of clustering SE's at the class level, which also capture the field of study. Column (4) additionally controls for the grade point average in the final year. Std. Errors clustered at the class level shown in parentheses, except for column (3) which shows robust standard errors. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.4: Multinomial Logit Regressions on Job Offer and Continued Education Plans including GPA in final year

	(1)		(2)		(3)		(4)	
	Job	Educ	Job	Educ	Job	Educ	Job	Educ
Patience (δ)	-0.12	–0.7	0.03	–0.58	–0.88	–1.70	–0.83	–2.63
	(1.50)	(1.37)	(1.58)	(1.43)	(2.25)	(1.99)	(2.60)	(1.92)
Present Bias (β)	2.45***	2.20**	2.45***	2.18**	3.15***	3.04***	3.31***	3.24**
	(0.71)	(0.94)	(0.71)	(0.86)	(0.76)	(0.80)	(1.22)	(1.30)
Loss aversion	0.46**	0.25	0.43*	0.27	0.70***	0.47	0.71**	0.51
	(0.23)	(0.27)	(0.24)	(0.28)	(0.24)	(0.36)	(0.29)	(0.39)
Risk aversion	–0.21	–0.37**	–0.22	–0.43**	–0.23	–0.45**	–0.20	–0.43**
	(0.16)	(0.19)	(0.18)	(0.19)	(0.22)	(0.19)	(0.24)	(0.17)
Grade Point Average	0.67	1.58*	0.89	1.41*	0.48	1.37*	0.15	1.44**
	(0.84)	(0.83)	(0.82)	(0.76)	(1.01)	(0.74)	(0.96)	(0.65)
Field of study, gender	No		Yes		Yes		Yes	
Socioeconomic Controls	No		No		Yes		Yes	
BIG 5, GRIT	No		No		No		Yes	
Pseudo R^2	0.040		0.059		0.133		0.182	
Observations	157		157		146		145	

Multinomial Logit Regressions on *plans after graduation.* Job denotes having a definite job offer. Educ indicates a plan to enroll in a continued education program. The table shows the coefficients of the multinomial logit regressions. Regressions only contains students with consistent answers and students who are expected to finish their education program in the year of the survey. Socio-economic controls include: age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in regression specifications (3) and (4). Standard Errors clustered at the class level. P-values shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.5: Logit Regressions on active Job Search

	(1)	(2)	(3)	(4)
Patience (δ)	0.000	0.095	0.182	0.153
	(0.295)	(0.267)	(0.284)	(0.266)
Present Bias (β)	0.358**	0.397**	0.431***	0.405**
	(0.182)	(0.172)	(0.165)	(0.197)
Risk aversion	–0.025	–0.033	–0.057*	–0.042*
	(0.027)	(0.027)	(0.033)	(0.024)
Loss aversion	0.044	0.078*	0.102	0.047
	(0.040)	(0.045)	(0.072)	(0.071)
business (d)		0.082	0.106	0.090
		(0.099)	(0.137)	(0.137)
polytech (d)		0.012	0.037	0.024
		(0.108)	(0.085)	(0.067)
female (d)		0.165***	0.128**	0.100*
		(0.054)	(0.064)	(0.059)
german (d)			0.063	0.119
			(0.114)	(0.132)
age			–0.014	–0.027
			(0.042)	(0.034)
grade 2009 math			0.072	0.035
			(0.073)	(0.054)
grade 2009 german			0.021	0.072
			(0.117)	(0.076)
grade 2009 english			0.085	0.048
			(0.065)	(0.051)
difficulty to save 100 SFR			0.013	0.005
			(0.036)	(0.019)
cognitive reflection test			–0.035	0.001
			(0.035)	(0.026)
Education mother			0.038	0.009
			(0.063)	(0.044)
Education father			0.004	0.036
			(0.047)	(0.033)
Openness				–0.021
				(0.015)
conscientiousness				0.041***
				(0.013)
extraversion				0.018**
				(0.008)
agreeableness				–0.021
				(0.014)
neuroticism				0.015**
				(0.007)
grit_grit				0.048
				(0.063)
Constant				
Pseudo R^2	0.045	0.121	0.167	0.298
Observations	90	90	84	84

Logit Regressions on active job search. Columns (1)-(3) are equivalent to columns (2)-(4) in table 9 and display regression coefficients from a regression on a dummy that takes value 1 in case a student has actively searched for a job. Marginal Effects are shown. Regressions only contain subjects with consistent answers in the choice experiments. Moreover, only students with an actual search motive are included (see section 4.3 for details). Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (2) and (3). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.6: OLS Regressions on active Job Search: Wild Cluster Bootstrap

	(1)	(2)	(3)	(4)
Patience (δ)	0.071	0.138	0.132	0.140
	(0.790)	(0.655)	(0.650)	(0.660)
Present Bias (β)	0.277**	0.284*	0.331***	0.285*
	(0.050)	(0.050)	(0.000)	(0.075)
Risk aversion	–0.027	–0.039	–0.058*	–0.052
	(0.325)	(0.205)	(0.085)	(0.150)
Loss aversion	0.046	0.079	0.103	0.072
	(0.290)	(0.135)	(0.170)	(0.475)
Constant	0.392	0.113	–0.696	–1.022
	(0.395)	(0.810)	(0.605)	(0.550)
R^2	0.039	0.103	0.155	0.255
Observations	90	90	84	84

OLS Regressions on active job search. Columns (1)-(4) display regressions on a dummy that takes value 1 in case a student has actively searched for a job. Regressions only contain subjects with consistent answers in the choice experiments. Moreover, only students with an actual search motive are included (see section 4.3 for details). The sets of controls are explained in more detail in section 3.2.3. Since some students did not report grades or did not completely fill out the survey, a few observations are dropped in columns (3)-(4). The p-values clustering on 14 classes are shown in parentheses. To account for the low number of clusters we apply a wild cluster bootstrap-t procedure (Cameron et al., 2008) to determine the p-values, with 5000 repetitions. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.7: Marginal effects of a Logit regression on active Job Search with firm controls

	Firm Controls		Class F.E.	Final GPA
	(1)	(2)	(3)	(4)
Patience (δ)	0.10	0.02	–0.31	–0.10
	(0.29)	(0.30)	(0.52)	(0.36)
Present Bias (β)	0.30*	0.36*	0.61**	0.43*
	(0.18)	(0.19)	(0.26)	(0.24)
Risk aversion	–0.04	–0.02	–0.02	–0.01
	(0.03)	(0.02)	(0.04)	(0.02)
Loss aversion	0.05	0.04	0.11	0.06
	(0.04)	(0.04)	(0.08)	(0.04)
no pfp (d)	0.00			
	(0.18)			
pfp (d)	–0.19			
	(0.23)			
$<$100 Emp. (d)		0.12		
		(0.14)		
$>$100 Emp. (d)		0.01		
		(0.10)		
Grade Point Average				0.06
				(0.15)
Constant				
Pseudo R^2	0.09	0.07	0.19	0.05
Observations	90	90	69	77

Logit Regressions on a dummy that takes value 1 in case a student has actively searched for a job. The number of observations is explained in section 4.3. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Moreover, only students with an actual search motive are included (see section 4.3 for details). Columns (1) and (2) add different firm size controls. Column (3) controls for class fixed effects instead of clustering SE's at the class level, which also capture the field of study. Column (4) additionally controls for the grade point average in the final year. Std. Errors clustered at the class level shown in parentheses, except for column (3) which shows robust standard errors. Significance levels: *** p<.01, ** p<.05, * p<.1.