Swiss Leading House
Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 144

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

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

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 144

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

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

March 2021 (first version: January 2018)

Published as: “Do preferences and biases predict life outcomes? Evidence from education and labor market entry decisions.“ *European Economic Review*, 134(2021). By Uschi Backes-Gellner, Holger Herz, Michael Kosfeld and Yvonne Oswald.

DOI: https://doi.org/10.1016/j.euroecorev.2021.103709

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

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

March 2021

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 significant predictor 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. 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.

Keywords: Economic preferences, education, dropout, human capital, job search
JEL Classification Codes: D01, D91, I21, J64

*Uschi Backes Gellner: Department of Business Administration, University of Zurich, Plattenstr. 14, CH-8032 Zürich, backes-gellner@business.uzh.ch; Holger Herz: Department of Economics, University of Fribourg, Bd. de Perolles 90, CH-1700 Fribourg, holger.herz@unifr.ch; Michael Kosfeld: Faculty of Economics and Business Admnistration, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, D-60323 Frankfurt am Main, kosfeld@econ.uni-frankfurt.de; Yvonne Oswald: Department of Business Administration, University of Zurich, Plattenstr. 14, CH-8032 Zürich, yvonne.oswald@oec.uzh.ch. 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 and to Leander Kopp for research assistance.

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. Figlio et al. (2019) provide complementary evidence on the association between long-term orientation and educational attainment using administrative data from immigrant students in the U.S., and Castillo et al. (2019) show that discount rates, elicited via incentivized experiments, of 8th graders in U.S. public middle schools predict high school graduation.

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 behave inconsistently over time, which can be modelled by 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 empirical evidence on the relationship between dropouts, economic time preferences, and behavioral biases.[2]

We analyze a unique data set that combines experimentally 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. Completing post-secondary education is associ-

[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.

[2] In principle, a third reason for dropping out of voluntary upper-secondary education could be diminishing authority and influence of parents, the idea being that parents "make" students' enrollment decision (as the latter are still minors), while students (when they become older) decide to dropout. However, most students in our context are still minors at the time of dropping out and therefore still need their parents' consent to terminate the contract. Formally, parents' influence is therefore the same at both points in time. Given the results we find, we consider their influence to be rather weak and therefore do not explore this channel further in our analysis. Also, policy implications are very similar to the ones for incomplete information.

ated with clear positive outcomes, both economically and psychologically, as Fritschi et al. (2009) show.[3] 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). Our results thus rely on a rich set of individual measures allowing us to differentiate between long-run patience and present bias as well as a large set of covariates to control for potential confounding factors.

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. The results corroborate the findings from Castillo et al. (2019) documenting the role of (long-run) patience in the context of high-school dropouts.

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 the motivation to invest into job search, implying a negative effect on the transition from unemployment to employment. Empirically, they find an association between a measure of impatience and the length of unemployment spells in the NLSY.[4] 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.

[3] E.g., the annual salary is estimated to be more than 23.000 CHF higher on average and physical or mental problems about 50 percent less prevalent among individuals with completed post-secondary education compared to those without.

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

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. This is because apprentices are employed by host companies yielding non-negligible wage earnings, but their contracts expire at the end of the program.[5]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 neither. We were able to collect survey responses from 181 students (out of 223 students who indeed finished their program in the year of the survey).[6] Hence, we received responses from 81% of all students in the initial sample that transitioned to the labor market in the year of the survey. 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 are able to investigate the predictions of these job search theories empirically.

Consistent with DellaVigna and Paserman (2005), we find that 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 increase in the estimated present bias 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. With respect to the impact of loss aversion, our results are weaker. While we do find that higher loss aversion is positively associated with a higher probability to have a definite job offer, which is consistent with DellaVigna et al. (2017), the association does not always reach standard thresholds of statistical significance.

Our 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; Alan and Ertac, 2018), 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.[7] Second, such commitment devices may

[5]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.

[6]The rest either dropped out or did not finish the program in time.

[7]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

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 experimentally elicited preference measures are for a variety of lifetime outcomes. Castillo et al. (2011, 2018) find that impatience as well as risk preferences correlate 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 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 in the important context of vocational schooling.

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 beta-delta model (Laibson, 1997). 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 $t = 2, \ldots$.[8] 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

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

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.

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

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,

$$-\beta\delta c + \beta \sum_{t=2}^{\infty} \delta^t b \quad > \quad 0 \tag{2}$$

$$\frac{\delta}{1-\delta} \quad > \quad \frac{c}{b}. \tag{3}$$

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

$$-c + \beta \sum_{t=1}^{\infty} \delta^t b \quad > \quad 0 \tag{4}$$

$$\frac{\beta\delta}{1-\delta} \quad > \quad \frac{c}{b}. \tag{5}$$

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:

$$\frac{\beta\delta}{1-\delta} > \frac{\tilde{c}}{\tilde{b}}, \tag{6}$$

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 DellaVigna and Paserman (2005). Extending a classic job search model to account for present-biased time preferences, the authors show that present bias (β) primarily affects an individual's search effort while long-run patience (δ) primarily affects an individual's reservation wage. Intuitively, present biased individuals search less, because the costs of search effort accrue immediately. Individuals, who are impatient in the long run, however, have a lower reservation wage as they compare wage earnings that accrue in the future. In consequence, the results in DellaVigna and Paserman (2005) predict that time inconsistency (i.e., stronger present bias) is associated with a lower probability to secure a job and a higher probability to be unemployed, whereas impatience of individuals, who are time consistent, is ambiguous or even positive. We refer to DellaVigna and Paserman (2005) for details.

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. Again, we refer to the original paper for details.

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.[9]

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

[9] 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.

companies. The students are employed at the host company for the duration of the education program and earn a non-negligible wage.[10] 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. The Swiss vocational education model is very similar to models in Germany and Austria and considered to be a prime example for apprenticeship training programs by a number of other countries in the world (Kelsall, 2015).

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. All students of the incoming class that were present on the day of the experiment participated, i.e., there was no self-selection in or out of the experiment. In total, 265 students from 14 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 and represent about 20 percent of the overall student population (OPET, 2011). The training program for students in the commercial sector lasts three years and 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. We chose this experimental design as it is straightforward to implement in the field and furthermore, our school set-up did not allow us to implement more complicated and

[10] Average monthly wages are between CHF 600–700 in the first year and increase to CHF 1250–1450 in the final year. Financial constraints are hence an unlikely cause of dropouts, in contrast to college dropouts in the US.

time-consuming procedures.[11]

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 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 $\beta < 1$.[12]

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

[11] Similar designs 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. See also Halevy (2014) and Cohen et al. (2020). 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).

[12] Similarly, a student reveals future bias, if $X_0 > X_1$ or, equivalently, $\beta > 1$. While measurement errors in the case of β might be larger compared to δ, as the measure depends on the ratio of the two switch points, we perform various robustness checks in our empirical analysis (e.g., dummy model) to rule out that these drive our results. See details below.

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.[13] 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 $riskaversion = (10 - Y)$.[14]

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.

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 $lossaversion = (7 - X)$.[15]

[13] Similar designs have been used, e.g., in Burks et al. (2009), Gächter et al. (2010), and Abeler et al. (2011).

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

[15] 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.

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 19 out of our 265 students (7.2%). Moreover, it should be noted that our loss aversion measure is also impacted by students' risk aversion. As a logical consequence, our measures of risk and loss aversion are highly correlated ($\rho = 0.35$, $p < 0.01$). To isolate the impact of loss aversion on behavior, it is hence important to control for risk aversion in our regression framework, and vice versa.

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.[16]

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.[17]

In our regression analysis, we group our control variables as follows: The first set of controls includes all socioeconomic variables, i.e., gender, 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 second 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.

[16] Our individual measures of risk aversion also allow us to assess whether our measure of present bias is confounded by uncertainty. Not choosing the delayed option in the first price list meant receiving the payment immediately, which eliminated uncertainty in case a student did not find the promise of the delayed payment credible. Hence, displayed present-biasedness (captured by a lower β) could potentially reflect risk aversion. However, we find that risk aversion and β are not correlated ($\rho = -0.01$, $p = 0.93$), which does not support the hypothesis that uncertainty is confounding our measure of present-bias.

[17] 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.

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.[18]

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 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).

[18]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.

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.[19] 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 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 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

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

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

CHF 5. However, the standard error is relatively large, implying considerable heterogeneity in risk aversion.[20]

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	0.68
	(1.69)	(1.62)	(2.17)		(1.67)	(1.76)	
Loss Aversion	4.93	4.95	4.77	0.40	4.95	4.84	0.61
	(1.05)	(1.05)	(1.07)		(1.03)	(1.17)	
Number of Obs.	227	201	26		182	45	

Note: Standard Deviations in parentheses. Column p shows the p-value of two-sided Wilcoxon rank-sum tests 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 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.[21]

[20]While other studies have found risk aversion on average (Holt and Laury, 2002; Dohmen et al., 2011), Herz et al. (2020) find similar risk neutrality in a sample of 496 students at a lower secondary school in Switzerland (i.e., the educational stage just before students start apprenticeships). Average risk neutrality hence seems common for this type of population.

[21]While the average annual discount factors implied by our three-months discount rates are rather small (.38), they

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. Columns (1)-(12) 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. Columns (1)-(6) include class fixed effects, to control for potential unobserved explanatory variables at the class level. In columns (1)-(4), the four preference measures are successively included as explanatory variables. Column (5) adds socioeconomic controls, and column (6) adds our personality measures. In all regression specifications, a higher δ, which implies less discounting of the long-run future, is associated with a lower dropout probability. As soon as other preferences are controlled for and socioeconomic and personality measures are added to the regression specification (columns (3-6), the association becomes statistically significant.[22]

A student's present bias may represent an additional source of discounting to future payments. However, the coefficient on β is not significantly different from zero in all regression specifications in columns (1)-(6) of Table 3. Risk and loss aversion, which are added in columns (4) to (6) also do not show any consistent and significant predictive power with regard to dropout.[23]

Because students self-select into an occupation (electrician, polytechnician or business), and different occupations are sampled from different schools, we report the same regressions specifications as in columns (1)-(6) again in columns (7)-(12), except that we include school fixed effects instead of class fixed effects. The pattern remains remarkably robust: A higher δ is associated with

are not unusual compared to other studies in which annual discount rates are inferred from discount rates over a rather short time horizon. See Frederick et al. (2002) and Ericson and Laibson (2019) for reviews.

[22]Table B.1 in Appendix B presents the coefficients for all regressors added in columns (5) and (6).

[23]The exception is one significant association between loss aversion and dropout in column (6).

Table 3: Marginal effects of logit regressions on dropout

	Class F.E.'s						School F.E.'s						Lasso
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Patience (δ)	–0.17		–0.18*	–0.17*	–0.21*	–0.22***	–0.12		–0.13*	–0.12*	–0.16**	–0.15**	–0.13*
	(0.14)		(0.11)	(0.09)	(0.11)	(0.08)	(0.09)		(0.08)	(0.07)	(0.08)	(0.07)	(0.08)
	[0.37]		[0.17]	[0.11]	[0.09]	[0.03]	[0.24]		[0.17]	[0.13]	[0.06]	[0.03]	[0.06]
Present Bias (β)		–0.05	–0.08	–0.07	–0.09	–0.09		0.01	–0.01	–0.00	–0.03	–0.04	–0.02
		(0.11)	(0.10)	(0.10)	(0.10)	(0.06)		(0.08)	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)
Risk aversion				0.01	0.01	0.01				0.01	0.01	0.01	
				(0.01)	(0.01)	(0.01)				(0.01)	(0.01)	(0.01)	
Loss aversion				–0.02	–0.03	–0.03**				–0.02	–0.02	–0.02	
				(0.02)	(0.02)	(0.01)				(0.01)	(0.01)	(0.01)	
Class FE's	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	
School FE's	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	
Soc. Ec. Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes	
BIG 5, GRIT	No	No	No	No	No	Yes	No	No	No	No	No	Yes	
Pseudo R^2	0.07	0.06	0.07	0.08	0.10	0.18	0.04	0.04	0.04	0.05	0.08	0.13	0.01
Observations	188	188	188	188	173	170	227	227	227	227	212	209	227

Logit Regressions on Dropout. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: female, age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Columns (1)-(6) all include class fixed effects. Columns (7)-(12) include school fixed effects. Because no dropouts are observed in two of the classes, the number of observations is reduced in columns (1)-(6), where class fixed effects are included. Since some students did not report grades or did not completely fill out the survey, a few more observations are dropped in regression specifications (5) and (6) as well as (11) and (12). Column (13) reports estimates from a double selection lasso logit estimation that endogenously selects controls from our full set of controls used in column (6). Std. Errors are clustered at the class level and are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1. p-values of hypotheses tests of the regression coefficient on δ against 0 using a score bootstrap (Kline and Santos, 2012) are reported in square brackets.

a lower dropout probability, and as soon as other preferences are controlled for and socioeconomic and personality measures are added to the regression specification (columns (9-12), the association becomes statistically significant. With respect to β, risk aversion and loss aversion, the results from the regressions with class fixed effects are also confirmed when using school fixed effects. Neither of these preference measures has significant predictive power with respect to dropout.

Taken together, we find evidence that patience is meaningfully associated with dropout behavior. A one standard-deviation increase in our measure of δ, which is equal to 0.16, is associated with a 1.9 to 3.5 percent lower probability of dropping out of the vocational training program.

One might be worried that our results are biased due to a small number of clusters, since we sampled only from a maximum of 14 classes. To address this concern, we report p-values of hypotheses tests of the regression coefficient on δ against 0 using the score bootstrap method (Kline and Santos, 2012) in square brackets in Table 3. As one would expect, statistical significance is slightly reduced in these tests. However, the association between patience and dropout remains significant as soon as socioeconomic variable are controlled for.

Finally, in column (13), we apply a double selection lasso logit regression (Belloni et al., 2014, 2016) in which control variables are selected in a data-driven way for inclusion from our full set of control variables. The aim is to select those control variables that actually have predictive power for our measures of time preference, δ and β, or for the dropout outcome. To determine inclusion of controls based on having predictive power for the time preference measures, linear regressions are used. Then, a logit regression is run including those variables selected in the initial lasso stage, clustering standard errors at the class level. This procedure ensures that control variables that are non-negligibly related with either a key predictor or the dependent variable are accounted for.

The lasso estimates confirm our previous results. δ is significantly associated with dropout, whereas the association between dropout and β is small and insignificant. Applying the score bootstrap (Kline and Santos, 2012) using the lasso estimates also confirms that the estimate on δ is marginally significantly different from zero.[24]

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, columns (1)-(6) report coefficients of logit regressions with class fixed effects, columns (7)-(12) report coefficients of logit regressions with school fixed effects,

[24]The lasso specification in column (13) does not pre-select school fixed effects or class fixed effects. If we pre-select school fixed effects, the lasso specification is equivalent to column (3) in table 3. If we pre-select class fixed effects, the lasso specification is equivalent to column (9) in table 3. Note that we use a linear probability specification when pre-selecting class fixed effects due to non-convergence of the penalty term when using double selection logit. Results can be obtained from the authors upon request.

and column (13) reports coefficients of a double selection lasso logit regression. Columns (1)-(4) as well as (7)-(10) subsequently add our preference measures to the regression specification, and in columns (5) and (6) as well as (11) and (12), socioeconomic controls and personality measures are added.

The data using class fixed effects in columns (1) to (6) show that a higher δ is associated with a higher probability of finishing in time. Once all preference measures are controlled for in the regression, the association becomes marginally significant. If socioeconomic controls and personality measures are additionally controlled for in columns (5) and (6), the association becomes significant at the 5% level. Controlling for school fixed effects instead of class fixed effects in columns (7)-(12) reveals a very similar picture. Once socioeconomic controls and personality measures are controlled for, the association between patience and finishing in time becomes significant at the 5% level.

Taken together, we find evidence that patience is meaningfully associated with timely program completion. A one standard deviation increase in our measure of δ is associated with a 2.2-4.5 percent larger probability of finishing the training program in time.

To control for potential bias due to the small number of clusters, we again perform hypotheses tests of the regression coefficient on δ against 0 using the score bootstrap method (Kline and Santos (2012)) in square brackets in Table 4. Again, statistical significance is slightly reduced in these tests. However, the association between patience and finishing in time remains marginally significant in both the regressions with class fixed effects and school fixed effects as soon as socioeconomic controls are controlled for (columns (5)-(6) and (11)-(12)).

Looking at the estimated coefficient on β in Table 4, we also see a positive estimate, but much smaller and not statistically significant. Moreover, our measures of risk and loss aversion have, again, no predictive power with regard to finishing in time.[25]

Finally, in column (13) of Table 4, we again apply a double selection lasso logit regression (Belloni et al., 2014, 2016) in which control variables are selected in a data-driven way for inclusion from our full set of control variables. The lasso estimates confirm our previous results. δ is marginally significantly associated with finishing in time, whereas the association between finishing in time and β is small and insignificant. Applying the score bootstrap (Kline and Santos, 2012) using the lasso estimates also confirms that the estimate on δ is marginally significantly different from zero[26]

In Appendix B and Appendix C, we provide further robustness checks for the above results. First, 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 for achieving good grades in school, as well as on the size of the company. Columns (1) and (2) in Table B.2 and C.2 report regression results

[25] All regression coefficients for columns (5)-(6) and (11)-(12) are reported in Table C.1 in Appendix C.

[26] The lasso specification in column (13) does not pre-select school fixed effects or class fixed effects. If we pre-select school fixed effects, the lasso specification is equivalent to column (3) in table 4. If we pre-select class fixed effects, the lasso specification is equivalent to column (9) in table 4. Note that we use a linear probability specification when pre-selecting class fixed effects due to non-convergence of the penalty term when using double selection logit. Results can be obtained from the authors upon request.

Table 4: Marginal effects of logit regressions on finished in time

	Class F.E.'s						School F.E.'s						Lasso
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Patience (δ)	0.14		0.17	0.18*	0.26**	0.28**	0.14		0.15	0.14	0.20*	0.22**	0.24*
	(0.13)		(0.12)	(0.11)	(0.13)	(0.13)	(0.12)		(0.12)	(0.11)	(0.11)	(0.11)	(0.13)
	[0.34]		[0.22]	[0.15]	[0.09]	[0.095]	[0.31]		[0.26]	[0.25]	[0.10]	[0.09]	[0.10]
Present Bias (β)		0.04	0.08	0.07	0.19	0.23		0.02	0.05	0.03	0.09	0.11	0.06
		(0.13)	(0.12)	(0.12)	(0.13)	(0.16)		(0.10)	(0.09)	(0.09)	(0.08)	(0.08)	(0.09)
Risk aversion				–0.01	–0.00	0.00				–0.01	–0.00	–0.01	
				(0.02)	(0.01)	(0.02)				(0.02)	(0.02)	(0.02)	
Loss aversion				0.03	0.04	0.02				0.02	0.03	0.02	
				(0.03)	(0.03)	(0.03)				(0.02)	(0.02)	(0.02)	
Class FE's	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	
School FE's	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	
Soc. Ec. Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes	
BIG 5, GRIT	No	No	No	No	No	Yes	No	No	No	No	No	Yes	
Pseudo R^2	0.10	0.09	0.10	0.10	0.22	0.24	0.07	0.06	0.07	0.07	0.17	0.18	0.01
Observations	200	200	200	200	186	183	227	227	227	227	212	209	227

Logit Regressions on finished in time. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: female, age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Columns (1)-(6) all include class fixed effects. Columns (7)-(12) include school fixed effects. Because all students finished in time in one of the classes, the number of observations is reduced to a maximum of 200 in columns (1)-(6), where class fixed effects are included. Since some students did not report grades or did not completely fill out the survey, a few more observations are dropped in regression specifications (5) and (6) as well as (11) and (12). Column (13) reports estimates from a double selection lasso logit estimation that endogenously selects controls from our full set of controls. Std. Errors are clustered at the class level and are shown in parentheses. The p-values of hypotheses tests of the regression coefficient on δ against 0 using a score bootstrap (Kline and Santos, 2012) are reported in square brackets. Significance levels: *** p<.01, ** p<.05, * p<.1.

using the regression specification with the full set of controls as well as with class and school fixed effects, additionally adding indicators for performance pay for achieving good grades in school. In columns (3) and (4), indicators for firm size are added. Table B.2 shows that estimates for the association between long-run patience and dropout remain similar in magnitude and significant. Table C.2 shows that the association between long-run patience and finishing in time remains similar in magnitude, but only reaches standard significance thresholds when controlling for firm size.

Columns (5) and (6) control for a student's potential present bias by modeling β as a dummy variable that equals one if $\beta < 1$ and zero otherwise. In columns (7) and (8), we exclude all students who are present biased and re-run the regression only with students with $\beta \geq 1$. Both robustness checks show significant associations between long-run patience and dropout, respectively finishing in time.

Finally, since we assess the impact of long-run patience on two different outcomes, we report p-values for our main regression specifications after controlling for the false discovery rate using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) in Table F.1 in Appendix F. It can be seen that, even after the adjustment on false discovery rates as well as adjusting for the small number of clusters using the score bootstrap, the association between long-run patience and dropout remains at least marginally significant in almost all specifications.

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. 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. 223 out of the initial 261 students were expected to finish in total. The response rate hence corresponds to an 81% of the relevant subsample.

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).[27] 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: $p = 0.89$).

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. Only students with consistent answers in the preference measures and who are finishing their apprenticeship are included. *No Plan* contains those students without continuation plans. *Job* contains those students with definite job offers. *Education* contains those students with continuing education plans.

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. 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.[28] One complication of our data is that some student's have not finished accumulating human capital and are planning to continue their education, for

[27] 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.

[28] 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.

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 marginal effects of logit regressions on the dummy variable "continuation plan" introduced above. Standard errors are clustered at the class level. Columns (1) to (3) control for class fixed effects, and columns (4) to (6) control for school fixed effects. Columns (1) and (4) additionally control for all four preference measures. Columns (2) and (5) add socioeconomic control variables. In columns (3) and (6), our personality measures are added.[29]

As can be seen, present bias β has a highly significant effect in all regression specifications.[30] A one standard deviation increase in β (which is equal to 0.23) increases the probability of having a continuation plan by 6.4 to 11.5 percentage points.

To control for the small number of clusters in our regression specification, we also perform hypotheses tests of the regression coefficient on β against 0 using the score bootstrap method (Kline and Santos, 2012) in square brackets in Table 7. It can be seen that the association between β and having a continuation plan remains statistically significant.

Moreover, column (7) of table 7 shows coefficient estimates of a double selection lasso logit regression on having continuation plans, in which control variables are again selected in a data-driven way for inclusion from our full set of control variables. Our dependent variables of interest are β and loss aversion, as well as the continuation plan outcome. It can be seen that the association between β and having a continuation plan is again highly significant in this specification, also when applying the score bootstrap test.[31]

[29]Because we are interested in both the effect of loss aversion and the effect of present bias, and since patience and risk aversion are important controls to capture the effects of loss aversion and present bias, we do not report regression specifications between the outcome variable and individual preference measures only in Table 7. In results that can be obtained from the authors, we find a highly significant association between present bias and continuation plans, and no significant association between loss aversion and continuation plans when individually including those preference measures in the regression (in addition to class or school fixed effects).

[30]See Table D.1 in Appendix D for coefficients of all control variables in columns (2, 3, 5 and 6).

[31]The lasso specification in column (7) of table 7 does not pre-select school fixed effects or class fixed effects. If either class fixed effects or school fixed effects are pre-selected, the association between β and having a continuation plan remains highly significant. Note that we use a linear probability specification when pre-selecting class fixed effects due to non-convergence of the penalty term when using double selection logit. Results can be obtained from the authors upon request.

Table 7: Marginal effects of logit regressions on having a continuation plan

	Class F.E.'s			School F.E.'s			Lasso
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Patience (δ)	–0.04	–0.15	–0.13	–0.02	–0.10	–0.07	
	(0.20)	(0.26)	(0.28)	(0.18)	(0.25)	(0.25)	
Present Bias (β)	0.32***	0.39***	0.50***	0.28***	0.36***	0.40***	0.28***
	(0.12)	(0.11)	(0.17)	(0.10)	(0.09)	(0.12)	(0.09)
	[< 0.01]	[0.01]	[0.03]	[< 0.01]	[< 0.01]	[0.01]	[0.02]
Risk aversion	–0.05*	–0.05	–0.03	–0.05*	–0.05*	–0.04	
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
	[0.09]	[0.19]	[0.34]	[0.08]	[0.11]	[0.17]	
Loss aversion	0.04	0.06*	0.05	0.04	0.05*	0.04	0.04
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
	[0.37]	[0.09]	[0.26]	[0.29]	[0.09]	[0.22]	[0.26]
Class FE's	Yes	Yes	Yes	No	No	No	
School FE's	No	No	No	Yes	Yes	Yes	
Socioeconomic Controls	No	Yes	Yes	No	Yes	Yes	
BIG 5, GRIT	No	No	Yes	No	No	Yes	
Double Selection Lasso							Yes
Pseudo R^2	0.10	0.19	0.25	0.06	0.17	0.21	
Observations	151	140	139	157	146	145	139

Logit Regressions on having a continuation plan. 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 shown. Regressions only contain subjects with consistent answers in the choice experiments. Socio-economic controls include: female, age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Columns (1)-(3) include class fixed effects. Columns (4)-(6) include school fixed effects. Because continuation plans have no variance in one of the classes, the number of observations is reduced by 6 in columns (1)-(3). Column (7) reports estimates from a double selection lasso logit estimation that endogenously selects controls from our full set of controls. Std. Errors are clustered at the class level and are shown in parentheses. The p-values of hypotheses tests of the regression coefficient on β, riskaversion and lossaversion against 0 using a score bootstrap (Kline and Santos, 2012) are reported in square brackets. Significance levels: *** p<.01, ** p<.05, * p<.1.

Finally, we again report p-values for our main regression specifications after controlling for the false discovery rate using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) in Table F.1 in Appendix F. It can be seen that, even after the adjustment on false discovery rates as well as adjusting for the small number of clusters using the score bootstrap, the association between present-bias and having continuation plans remains significant in all specifications.

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. Similar to before, specifications (1)-(3) include class fixed effects, and specifications (4)-(6) school fixed effects. Standard errors are always clustered at the class level. Specifications (1) and (4) only include the four preference measures. Socioeconomic controls are added in columns (2) and (5), personality measures are added in columns (3) and (6).

Lower present bias (higher β) is positively associated with both, having a job or planning continued education, relative to the baseline of having neither of them. For having a job, the association is significant for all regression specifications. For pursuing continued education, the association is significant as soon as socioeconomic controls are added to the regression specification.[32] 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 11-19 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 actually exerted any job search effort themselves. In Table 9 we therefore provide exploratory analyses from a sub-sample of students who self-reported search activity, excluding all students who have a definite job offer *and* report that they have not actively searched themselves. 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.[33] 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. Here, we no longer control for class fixed effects, because some classes show no variation in the outcome variable, and hence inclusion would further reduce the sample size. Column (1) only controls for our preference measures, socioeconomic controls are added in column (2), and personality measures are added in column (3). Standard errors are clustered at the class level.

As can be seen in Table 9, the results confirm our findings from above: students with a higher

[32] Again, to control for the small number of clusters in our sample, we also estimate p-values using pairs cluster bootstrapped t-statistics for multinomial logit models (Cameron et al., 2008), using the R-package "clusterSEs" (Esarey and Menger, 2019). Because of substantial collinearity between class fixed effects, we were only able to obtain cluster bootstrapped t-statistics for our model specifications with school fixed effects. Present bias remains significantly associated with having a job or planning continued education with at least $p < 0.1$, with one exception in column (4), where the p-value for the association between present bias and continued education drops to $p = 0.14$.

[33] As a consequence of these restrictions, our sample size for this analysis becomes rather small. While we regard the analysis as interesting, one should remain cautious regarding the statistical power of the analysis.

Table 8: Multinomial logit regression on Job offer and Continued Education Plans

	Class F.E.'s			School F.E.'s		
	(1)	(2)	(3)	(4)	(5)	(6)
OUTCOME: JOB						
Patience (δ)	–0.00	–1.02	–0.95	0.18	–0.56	–0.22
	(1.36)	(2.09)	(2.56)	(1.30)	(2.02)	(2.21)
Present Bias (β)	2.02***	2.87***	4.04*	1.75***	2.68***	3.19**
	(0.77)	(1.11)	(2.12)	(0.55)	(0.77)	(1.28)
Loss aversion	0.30	0.53**	0.49	0.31	0.49**	0.45*
	(0.28)	(0.26)	(0.32)	(0.24)	(0.22)	(0.26)
Risk aversion	–0.25	–0.30	–0.23	–0.27	–0.33	–0.28
	(0.19)	(0.26)	(0.32)	(0.17)	(0.23)	(0.25)
Constant	–1.09	–9.29	–14.59	–0.54	–8.96	–13.08*
	(1.60)	(7.15)	(9.23)	(1.47)	(5.97)	(7.17)
OUTCOME: EDUCATION						
Patience (δ)	–0.93	–1.64	–1.44	–0.66	–1.02	–1.16
	(1.38)	(2.11)	(2.31)	(1.16)	(1.90)	(1.95)
Present Bias (β)	1.79	2.94**	4.12**	1.65*	2.95***	3.51**
	(1.09)	(1.18)	(1.84)	(0.88)	(0.87)	(1.39)
Loss aversion	0.14	0.28	0.21	0.05	0.15	0.09
	(0.36)	(0.37)	(0.39)	(0.28)	(0.32)	(0.34)
Risk aversion	–0.37*	–0.47*	–0.44	–0.31**	–0.42**	–0.38**
	(0.20)	(0.27)	(0.32)	(0.15)	(0.18)	(0.18)
Constant	1.11	1.86	1.64	0.30	–3.26	–6.95
	(1.31)	(10.15)	(10.07)	(1.39)	(8.87)	(9.34)
Class FE's	Yes	Yes	Yes	No	No	No
School FE's	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	Yes	Yes	No	Yes	Yes
BIG 5, GRIT	No	No	Yes	No	No	Yes
Pseudo R^2	0.16	0.25	0.31	0.05	0.13	0.18
Observations	157	146	145	157	146	145

Multinomial Logit Regressions on *plans after graduation*. The upper panel reports coefficient estimates on the outcome "Job". The bottom panel reports coefficient estimates on the outcome "continued education". 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. Columns (1)-(3) include class fixed effects. Columns (4-6) include school fixed effects. Socio-economic controls include: female, 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 (2) and (3), as well as (5) and (6). Standard Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

β are more likely to have searched actively. As socioeconomic controls are added, the association becomes highly statistically significant. A one standard deviation increase in β (a decrease in present bias) makes it 9 to 10 percentage points more likely that a student has already actively searched

Table 9: Marginal effects of logit regressions on job search

	School F.E.'s			Lasso
	(1)	(2)	(3)	(4)
Patience (δ)	0.021	0.182	0.153	
	(0.270)	(0.284)	(0.266)	
Present Bias (β)	0.383*	0.431***	0.405**	0.429**
	(0.209)	(0.165)	(0.197)	(0.209)
	[0.19]	[0.09]	[0.17]	[0.13]
Risk aversion	–0.030	–0.057*	–0.042*	
	(0.027)	(0.033)	(0.024)	
	[0.22]	[0.10]	[0.01]	
Loss aversion	0.078	0.102	0.047	0.107*
	(0.049)	(0.072)	(0.071)	(0.060)
	[0.11]	[0.17]	[0.48]	[0.09]
School FE's	Yes	Yes	Yes	
Socioeconomic Controls	No	Yes	Yes	
BIG 5, GRIT	No	No	Yes	
Double Selection Lasso				Yes
Pseudo R^2	0.090	0.167	0.298	0.093
Observations	90	84	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 shown. Regressions only contain subjects with consistent answers in the choice experiments. Socioeconomic controls include: female, age, CRT score and high school grades, a dummy for native German speakers, difficulty to raise CHF 100, and parental educational background. Table E.1 in Appendix E.1 presents the coefficients for all regressors added in column (2) and (3). Columns (1)-(3) include school fixed effects. Regressions with class fixed effects generate unreliable results due to the small number of observations in this sample. Column (4) reports estimates from a double selection lasso logit estimation that endogenously selects controls from our full set of controls. Std. Errors are clustered at the class level and are shown in parentheses. The p-values of hypotheses tests of the regression coefficients on β, riskaversion and lossaversion against 0 using a score bootstrap (Kline and Santos (2012)) are reported in square brackets. Significance levels: *** p<.01, ** p<.05, * p<.1.

for a job, one month prior to the termination of the vocational training program.[34]

Finally, column (4) of Table 9 shows coefficient estimates of a double selection lasso logit regression on searching actively, in which control variables are selected in a data-driven way for inclusion from our full set of control variables. The aim is again to select those control variables that have predictive power for our dependent variables of interest, which are β and loss aversion, or for actively searching for a job. It can be seen that the association between beta and searching for a job

[34] To control for the small number of clusters in our regression specification, we also perform hypotheses tests of the regression coefficient on β against 0 using the score bootstrap method (Kline and Santos, 2012) and report the resulting p-values in square brackets in Table 9. It can be seen that statistical significance of the association between β and having searched for a job becomes smaller in these tests, only reaching marginal significance in regression specification (2), in which socioeconomic controls are included. However, the small number of observations in these exploratory analyses has to be kept in mind.

is again significant in this specification.[35]

Additional results reported in Appendix D include robustness checks for our regressions on having a continuation plan, in particular including firm characteristics and average school grades (Table D.2). The results show that present bias remains a significant predictor for students having a continuation plan. Similar robustness checks are conducted for job search. Table E.2 includes firm characteristics and grade point averages. Again, results on present bias remain robust and significant.

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. However, the statistical significance of this association is weak.*

First, looking again at Table 7, we see that loss aversion is positively associated with having a continuation plan, but the association is only marginally significant in the specification in which socioeconomic controls are added (columns (2) and (5)). This is partially unsurprising as the outcome variable continuation plan includes both having a job and planning to continue education. Continuing education, however, exposes the student to additional risk — future employment remains unclear — and may also come at an immediate loss as remuneration is usually considerably lower during education programs compared to employment. Contrary to having a job, it should therefore not be positively associated with loss aversion.

Our multinomial regressions in Table 8 partly confirm this. There, higher loss aversion is associated with a higher probability of having a definite job offer. The association is significant once socioeconomic characteristics are controlled for.[36] At the same time, loss aversion is not significantly associated with plans for continuing education.[37] Rather, the data suggest that less risk averse students are actually more likely to continue education. This is consistent with the above view that continued education plans expose students to additional risk, which the less risk averse students are more likely to take.

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 the double selection lasso logit specification.[38]

[35]The lasso specification in column (4) of table 9 does not pre-select school fixed effects. If school fixed effects are pre-selected, the association between β and active job search remains significant at $p = 0.06$. Results can be obtained from the authors upon request.

[36]If we control also for average school grades in the final year (Table D.3), the effect increases and becomes significant in all specifications.

[37]Again, to control for the small number of clusters in our sample, we also estimate p-values using pairs cluster bootstrapped t-statistics for multinomial logit models (Cameron et al., 2008), using the R-package "clusterSEs" (Esarey and Menger, 2019). Because of substantial collinearity between class fixed effects, we were only able to obtain cluster bootstrapped t-statistics for our model specifications with school fixed effects. Loss Aversion remains significantly associated ($p = 0.05$) with having a job in column (5), but becomes insignificant ($p = 0.16$) in column (6).

[38]If school fixed effects are pre-selected as controls for the double selection lasso logit regressions, the association between loss aversion and active job search again becomes insignificant. Results can be obtained from the authors upon request.

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), the likely cause of these differences is due to reduced search activity among present-biased students. Second, we also find some evidence that more loss averse students are more likely to have a definite job offer. This is consistent with DellaVigna et al. (2017) who show that loss aversion increases effort in job search when individuals have reference-dependent preferences. However, the statistical significance of our loss aversion results remains 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; Alan and Ertac, 2018). 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 is significantly associated with the probability of having a definite job offer about a month prior to completion of the education program, consistent with theories of impatience and job search. Second, we also see a positive association between loss aversion and a higher probability of having a definite job offer, consistent with theories of reference-dependent preferences and job search. However, the statistical significance of this association remains weak. Our 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

Abeler, Johannes, Armin Falk, Lorenz Götte, and David Huffman, "Reference Points and Effort Provision," *American Economic Review*, April 2011, *101* (2), 470–492.

Alan, S and Seda Ertac, "Fostering patience in the classroom: Results from randomized educational intervention," *Journal of Political Economy*, October 2018, *126* (5), 1865–1911.

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.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 2014, *81* (2), 608–650.

_ , _ , and Ying Wei, "Post-selection inference for generalized linear models with many controls," *Journal of Business & Economic Statistics*, 2016, *34* (4), 606–619.

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

Benjamini, Yoav and Yosef Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B*, 1995, *57* (1), 289–300.

Burks, Stephen, Jeffrey Carpenter, Lorenz Götte, and Aldo Rustichini, "Cognitive skills affect economic preferences, strategic behavior, and job attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7750.

_ , _ , _ , and _ , "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, Jeffrey L. Jordan, and Ragan Petrie, "Children's rationality, risk attitudes and field behavior," *European Economic Review*, 2018, *102*, 62–81.

_ , Jeffrey L Jordan, and Ragan Petrie, "Discount rates of children and high school graduation," *The Economic Journal*, 2019, *129* (619), 1153–1181.

_ , 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.

Cohen, Jonathan, Keith Marzilli Ericson, David I Laibson, and John Myles White, "Measuring Time Preferences," *Journal of Economic Literature*, June 2020, *58* (2), 299–347.

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.

— , — , — , — , **Jürgen Schupp, and Gert G Wagner**, "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

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.

Ericson, Keith M. and David Laibson, "Intertemporal choice," *Handbook of Behavioral Economics*, 2019.

Esarey, Justin and Andrew Menger, "Practical and effective approaches to dealing with clustered data," *Political Science Research and Methods*, 2019, *7* (3), 541–559.

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

Figlio, David, Paola Giuliano, Umut Özek, and Paola Sapienza, "Long-term orientation and educational performance," *American Economic Journal: Economic Policy*, 2019, *11* (4), 272–309.

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

— , **George Loewenstein, and Ted O'Donoghue**, "Time discounting and time preference: A critical review," *Journal of Economic Literature*, 2002, *40* (2), 351–401.

Fritschi, Tobias, Thomas Oesch, and Ben Jann, "Gesellschaftliche Kosten der Ausbildungslosigkeit in der Schweiz," *Büro für Arbeits- und Sozialpolitische Studien BASS*, 2009.

Gächter, Simon, Eric J Johnson, and Andreas Hermann, "Individual-level loss aversion in riskless and risky choices," *Working Paper*, November 2010, pp. 1–29.

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.

Halevy, Yoram, "Somme comments on the use of monetary and primary rewards in the measurement of time preferences," *Working Paper*, 2014.

Herz, Holger, Martin Huber, Tjasa Maillard-Bjedov, and Svitlana Tyahlo, "Time Preferences across Language Groups: Evidence on Intertemporal Choices from the Swiss Language Border," *unpublished manuscript*, 2020.

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American economic review*, 2002, *92* (5), 1644–1655.

Kelsall, Richard, "Cheap and cheerless," *The Economist*, March 5 2015, *https://www.economist.com/news/britain/21645758-britain-needs-better-apprenticeships-not-more-them-cheap-and-cheerless.*

Kline, Patrick and Andres Santos, "A score based approach to wild bootstrap inference," *Journal of Econometric Methods*, 2012, *1* (1), 23–41.

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 indicating female gender. *Business*, *Polytechnician* and *Electrician* are dummies indicating the field of study. *Native German speaker* 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 CHF 100* 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)

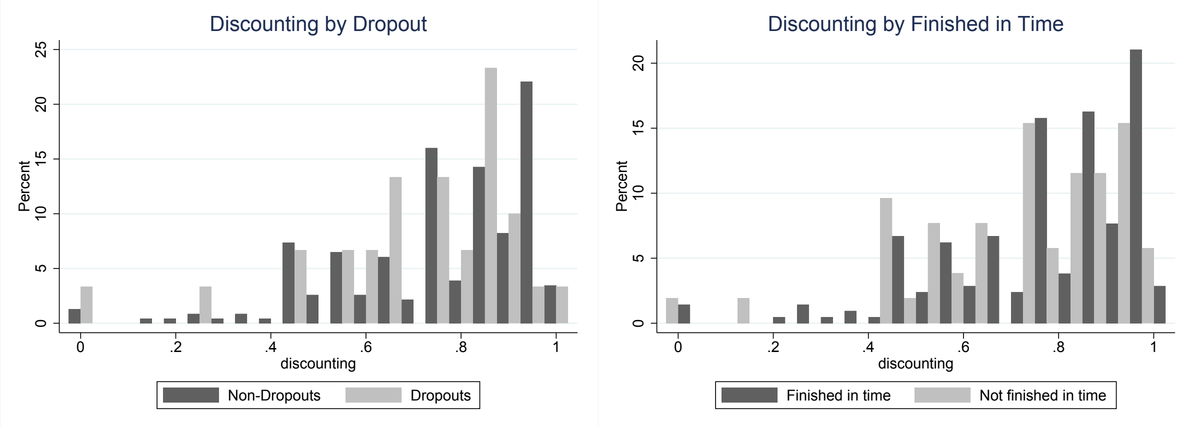

Figure A.2: Distribution of Present Bias (β)

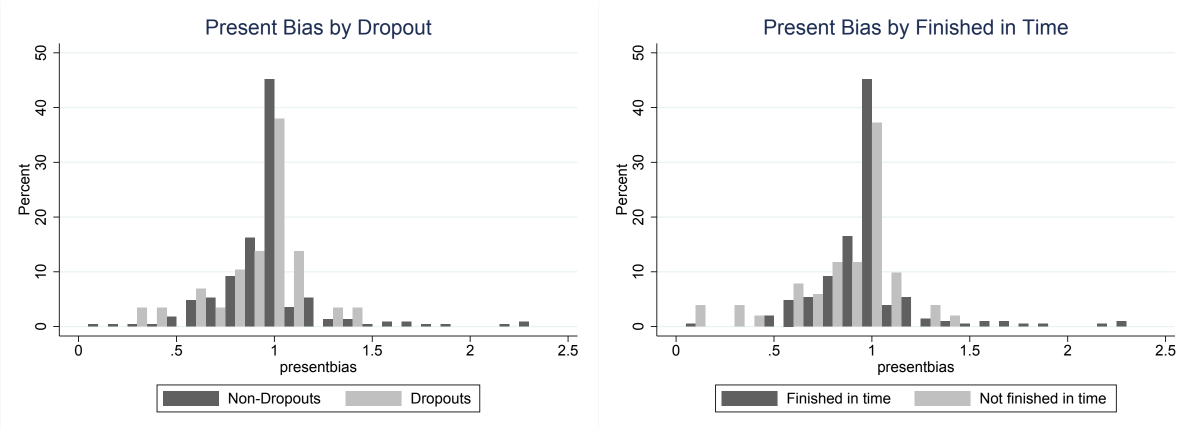

Figure A.3: Distribution of Risk Aversion

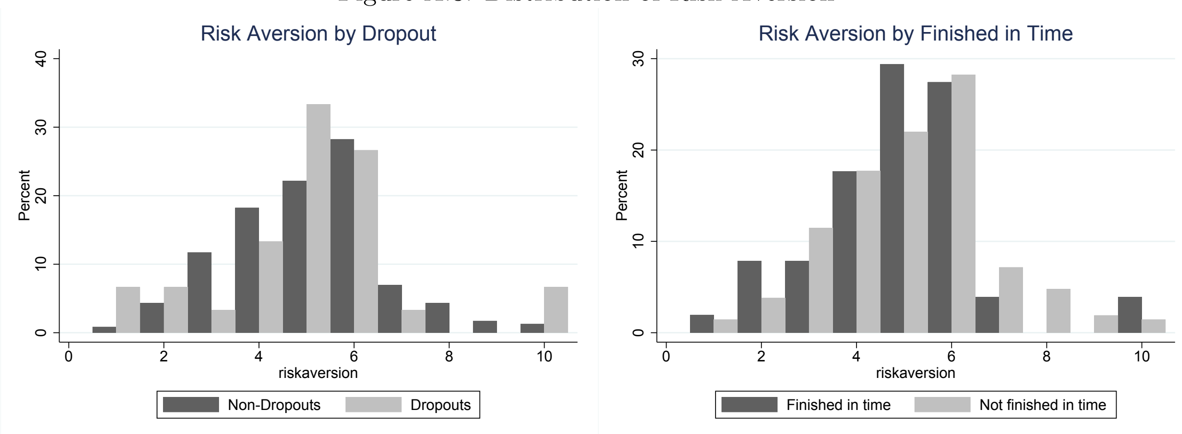

Figure A.4: Distribution of Loss Aversion

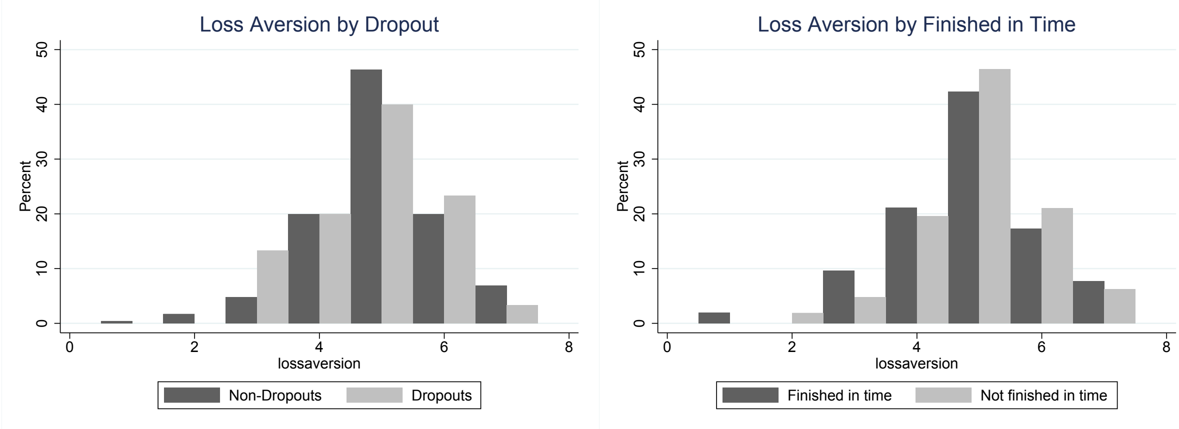

Table A.2: Summary Statistics and Selection in the Labor Market Sample

	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)	(4)
Patience (δ)	–0.21*	–0.22***	–0.16**	–0.15**
	(0.11)	(0.08)	(0.08)	(0.07)
Present Bias (β)	–0.09	–0.09	–0.03	–0.04
	(0.10)	(0.06)	(0.08)	(0.06)
Risk aversion	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Loss aversion	–0.03	–0.03**	–0.02	–0.02
	(0.02)	(0.01)	(0.01)	(0.01)
female (d)	0.01	–0.03	0.01	–0.02
	(0.07)	(0.06)	(0.05)	(0.05)
german (d)	–0.04	–0.05	–0.02	–0.02
	(0.06)	(0.09)	(0.04)	(0.05)
age	–0.01	–0.02	–0.01	–0.01
	(0.03)	(0.03)	(0.03)	(0.02)
Math grade	–0.03	–0.02	–0.02	–0.01
	(0.03)	(0.04)	(0.03)	(0.03)
German grade	–0.03	–0.04	–0.01	–0.00
	(0.06)	(0.05)	(0.04)	(0.04)
English grade	–0.03	–0.02	–0.03	–0.02
	(0.04)	(0.03)	(0.03)	(0.02)
borrowing difficulty	0.01	0.01	0.01	0.01
	(0.03)	(0.02)	(0.02)	(0.02)
CRT score	0.02	0.02	0.03	0.03**
	(0.02)	(0.02)	(0.02)	(0.01)
Education Mother	–0.01	–0.01	–0.01	–0.01
	(0.03)	(0.02)	(0.02)	(0.02)
Education Father	0.02	0.02	0.02	0.01
	(0.03)	(0.02)	(0.02)	(0.02)
Openness		0.02**		0.01
		(0.01)		(0.01)
Conscientousness		–0.01		–0.01
		(0.01)		(0.01)
Extroversion		–0.00		–0.00
		(0.01)		(0.00)
Agreeableness		0.01		0.00
		(0.01)		(0.01)
Neuroticism		0.00		0.00
		(0.01)		(0.00)
GRIT		–0.01		0.00
		(0.03)		(0.03)
Business (d)			–0.07	–0.06
			(0.06)	(0.07)
Polytechnician (d)			–0.08***	–0.08***
			(0.02)	(0.02)
Class Fixed Effects?	Yes	Yes	No	No
School Fixed Effects?	No	No	Yes	Yes
Pseudo R^2	0.10	0.18	0.08	0.13
Observations	173	170	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 (4). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table B.2: Marginal effects of a Logit regression on dropout with alternative specifications

	Performance Pay		Firm Size		Present Bias Dummy		dropping present biased subjects	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (δ)	–0.22***	–0.14*	–0.17***	–0.13*	–0.17**	–0.13**	–0.20	–0.23**
	(0.07)	(0.08)	(0.06)	(0.07)	(0.08)	(0.06)	(0.20)	(0.10)
Present Bias (β)	–0.08	–0.03	–0.07	–0.03				
	(0.06)	(0.07)	(0.05)	(0.07)				
Risk aversion	0.00	0.00	0.00	0.01	0.01	0.01	–0.01	–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*	–0.03**	–0.02*	–0.01	–0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
no pfp (d)	–0.18**	–0.09**						
	(0.09)	(0.05)						
pfp (d)	–0.07***	–0.07**						
	(0.02)	(0.03)						
$<$100 Emp. (d)			–0.15***	–0.09***				
			(0.04)	(0.03)				
$>$100 Emp. (d)			–0.14***	–0.09***				
			(0.05)	(0.03)				
Present Bias (d)					0.01	0.00		
					(0.04)	(0.03)		
Class FE's	Yes	No	Yes	No	Yes	No	Yes	No
School FE's	No	Yes	No	Yes	No	Yes	No	Yes
Socioec. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BIG5 / GRIT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.21	0.15	0.24	0.16	0.17	0.13	0.40	0.35
Observations	170	209	170	209	170	209	92	98

Logit Regressions on Dropout. Marginal Effects shown. Regressions only contain subjects with consistent answers in the choice experiments. Columns (1) and (2) add whether a student received performance pay for good grades in school. Columns (3) and (4) add firm size controls[a] Columns (5) and (6) include a binary indicator of present bias (= 1 if $\beta < 1$ instead of the estimated value of β). Columns (7) and (8) only include non-present biased subjects ($\beta \geq 1$). Std. Errors clustered at the class level shown in parentheses. 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)	(4)
Patience (δ)	0.26**	0.28**	0.20*	0.22**
	(0.13)	(0.13)	(0.11)	(0.11)
Present Bias (β)	0.19	0.23	0.09	0.11
	(0.13)	(0.16)	(0.08)	(0.08)
Risk aversion	–0.00	0.00	–0.00	–0.01
	(0.01)	(0.02)	(0.02)	(0.02)
Loss aversion	0.04	0.02	0.03	0.02
	(0.03)	(0.03)	(0.02)	(0.02)
female (d)	0.08	0.08	0.09	0.08
	(0.08)	(0.10)	(0.06)	(0.07)
german (d)	–0.14***	–0.12**	–0.12**	–0.10*
	(0.05)	(0.05)	(0.05)	(0.05)
age	–0.01	–0.01	–0.01	–0.01
	(0.03)	(0.03)	(0.03)	(0.03)
Math grade	–0.04	–0.07	–0.02	–0.04
	(0.05)	(0.06)	(0.04)	(0.05)
German grade	–0.06	–0.07	–0.03	–0.03
	(0.09)	(0.08)	(0.07)	(0.06)
English grade	0.14**	0.13**	0.13***	0.13**
	(0.06)	(0.06)	(0.05)	(0.05)
borrowing difficulty	–0.04	–0.04	–0.01	–0.01
	(0.03)	(0.03)	(0.02)	(0.02)
CRT score	–0.02	–0.03	–0.01	–0.02
	(0.03)	(0.03)	(0.02)	(0.02)
Education Mother	0.05	0.05*	0.03	0.04**
	(0.03)	(0.03)	(0.02)	(0.02)
Education Father	–0.01	–0.02	–0.01	–0.02
	(0.03)	(0.03)	(0.02)	(0.03)
Openness		0.00		0.00
		(0.01)		(0.01)
Conscientousness		0.01		0.00
		(0.01)		(0.01)
Extroversion		–0.00		–0.00
		(0.01)		(0.01)
Agreeableness		–0.00		0.01
		(0.01)		(0.01)
Neuroticism		–0.01		–0.00
		(0.01)		(0.01)
GRIT		0.07		0.06
		(0.07)		(0.04)
Business (d)			0.08	0.07
			(0.08)	(0.08)
Polytechnician (d)			0.16***	0.16***
			(0.03)	(0.03)
Class Fixed Effects?	Yes	Yes	No	No
School Fixed Effects?	No	No	Yes	Yes
Pseudo R^2	0.22	0.24	0.17	0.18
Observations	186	183	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 (4). Std. Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table C.2: Marginal effects of a Logit regression on finished in time with firm controls

	Performance Pay		Firm Size		Present Bias Dummy		dropping present biased subjects	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patience (δ)	0.28	0.21	0.28**	0.20*	0.22*	0.19*	0.30*	0.22**
	(0.17)	(0.13)	(0.13)	(0.11)	(0.12)	(0.10)	(0.16)	(0.10)
Present Bias (β)	0.25*	0.11	0.23	0.10				
	(0.15)	(0.08)	(0.15)	(0.09)				
Risk aversion	0.00	–0.00	0.00	–0.01	0.00	–0.00	0.01	0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Loss aversion	0.02	0.02	0.02	0.02	0.02	0.02	0.00	–0.01
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
no performance pay (d)	0.24*	0.10						
	(0.14)	(0.10)						
performance pay (d)	0.16***	0.10						
	(0.05)	(0.08)						
<100 Employees (d)			0.20**	0.09				
			(0.08)	(0.07)				
>100 Employees (d)			0.23***	0.13**				
			(0.08)	(0.06)				
Present Bias (d)					–0.08	–0.05		
					(0.05)	(0.04)		
Class FE's	Yes	No	Yes	No	Yes	No	Yes	No
School FE's	No	Yes	No	Yes	No	Yes	No	Yes
Socioec. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BIG5 / GRIT	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.26	0.19	0.27	0.19	0.23	0.18	0.47	0.41
Observations	183	209	183	209	183	209	94	98

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 whether a student received performance pay for good grades in school. Columns (3) and (4) add firm size controls[a] Columns (5) and (6) include a binary indicator of present bias (= 1 if $\beta < 1$ instead of the estimated value of β. Columns (7) and (8) only include non-present biased subjects ($\beta \geq 1$). Std. Errors clustered at the class level shown in parentheses. 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 D Additional Regressions on Continuation Plans

Table D.1: Marginal Effects of Logit Regressions on Continuation Plans: all coefficients

	(1)	(2)	(3)	(4)
Patience (δ)	–0.151	–0.133	–0.103	–0.072
	(0.261)	(0.276)	(0.248)	(0.248)
Present Bias (β)	0.389***	0.500***	0.363***	0.397***
	(0.114)	(0.167)	(0.087)	(0.121)
Risk Aversion	–0.046	–0.032	–0.048*	–0.039
	(0.030)	(0.032)	(0.029)	(0.029)
Loss Aversion	0.057*	0.046	0.048*	0.040
	(0.032)	(0.037)	(0.029)	(0.034)
Female (d)	–0.050	–0.043	–0.034	–0.021
	(0.049)	(0.081)	(0.050)	(0.066)
German (d)	–0.023	0.001	–0.006	0.013
	(0.103)	(0.089)	(0.091)	(0.084)
Age	0.031	0.026	0.033	0.028
	(0.044)	(0.042)	(0.042)	(0.042)
Math Grade	0.018	–0.021	0.027	0.013
	(0.047)	(0.049)	(0.045)	(0.047)
German Grade	–0.064	–0.075	–0.065	–0.068
	(0.086)	(0.082)	(0.073)	(0.070)
English Grade	0.070	0.073	0.073	0.071
	(0.052)	(0.063)	(0.048)	(0.055)
borrowing difficulty	–0.012	0.006	–0.009	0.001
	(0.026)	(0.025)	(0.022)	(0.023)
CRT Score	–0.094**	–0.081**	–0.094***	–0.084**
	(0.038)	(0.040)	(0.032)	(0.036)
Education mother	0.039	0.022	0.056	0.041
	(0.039)	(0.041)	(0.037)	(0.040)
Education father	0.032	0.025	0.024	0.017
	(0.034)	(0.029)	(0.032)	(0.030)
Openness		0.006		0.006
		(0.011)		(0.010)
Conscientiousness		0.024*		0.016
		(0.015)		(0.014)
Extraversion		0.013		0.015
		(0.011)		(0.010)
Agreeableness		–0.016		–0.010
		(0.022)		(0.016)
Neuroticism		0.001		0.002
		(0.014)		(0.011)
GRIT		0.032		0.024
		(0.071)		(0.069)
Business (d)			0.053	0.040
			(0.077)	(0.095)
Polytechnician (d)			–0.019	0.001
			(0.079)	(0.095)
Class Fixed Effects?	Yes	Yes	No	No
School Fixed Effects?	No	No	Yes	Yes
Pseudo R^2	0.194	0.254	0.165	0.211
Observations	140	139	146	145

Logit Regressions on various continuation plans. Columns (1)-(2) refer to columns (2)-(3) and columns (3)-(4) to columns (5)-(6) in Table 7. 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 further observations are dropped in columns (2) and (4). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table D.2: Marginal effects of a Logit regressions on continuation plans with additional controls

	Performance Pay		Firm Size		Final GPA	
	(1)	(2)	(3)	(4)	(5)	(6)
Patience (δ)	–0.03	0.00	–0.14	–0.08	–0.20	–0.16
	(0.30)	(0.25)	(0.28)	(0.25)	(0.33)	(0.29)
Present Bias (β)	0.52***	0.38***	0.50***	0.40***	0.63**	0.41**
	(0.16)	(0.12)	(0.16)	(0.12)	(0.26)	(0.19)
Risk aversion	–0.04	–0.04	–0.03	–0.03	–0.02	–0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
Loss aversion	0.04	0.03	0.04	0.04	0.10*	0.08**
	(0.04)	(0.04)	(0.04)	(0.03)	(0.06)	(0.04)
no performance pay (d)	0.21	0.12				
	(0.17)	(0.13)				
performance pay (d)	–0.04	–0.08				
	(0.13)	(0.14)				
<100 Employees (d)			0.08	0.06		
			(0.14)	(0.10)		
>100 Employees (d)			0.00	–0.01		
			(0.13)	(0.10)		
GPA					0.20	0.07
					(0.16)	(0.11)
Class FE's	Yes	No	Yes	No	Yes	No
School FE's	No	Yes	No	Yes	No	Yes
Socioec. Controls	Yes	Yes	Yes	Yes	Yes	Yes
BIG5 / GRIT	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.30	0.24	0.26	0.22	0.31	0.25
Observations	139	145	139	145	112	128

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 whether a student received performance pay. Columns (3) and (4) add firm size controls[a] Columns (5) and (6) additionally control for the grade point average in the final year. Std. Errors clustered at the class level shown in parentheses. 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.

Table D.3: Multinomial Logit Regressions on Job Offer and Continued Education Plans including GPA in final year

	Class F.E.'s			School F.E.'s		
	(1)	(2)	(3)	(4)	(5)	(6)
OUTCOME: JOB						
Patience (δ)	–0.37	–1.68	–1.23	–0.08	–0.89	–0.83
	(1.78)	(2.41)	(3.23)	(1.57)	(2.25)	(2.60)
Present Bias (β)	3.21***	4.02***	4.72**	2.37***	3.15***	3.31***
	(1.24)	(1.44)	(2.39)	(0.71)	(0.76)	(1.22)
Loss aversion	0.55*	0.88**	0.90**	0.45*	0.70***	0.71**
	(0.31)	(0.36)	(0.44)	(0.25)	(0.24)	(0.29)
Risk aversion	–0.22	–0.18	–0.17	–0.22	–0.23	–0.20
	(0.21)	(0.27)	(0.31)	(0.17)	(0.22)	(0.24)
GPA	1.45	1.32	1.55	0.92	0.48	0.15
	(0.97)	(1.25)	(1.28)	(0.83)	(1.01)	(0.96)
Constant	–9.92**	–21.93**	–28.89**	–6.06*	–16.37***	–20.64**
	(4.42)	(8.99)	(11.80)	(3.55)	(5.89)	(8.29)
OUTCOME: EDUCATION						
Patience (δ)	–1.17	–2.80	–1.99	–0.82	–1.70	–2.63
	(1.55)	(2.21)	(2.47)	(1.40)	(1.99)	(1.93)
Present Bias (β)	3.13**	4.23**	5.45**	2.00**	3.04***	3.24**
	(1.56)	(1.72)	(2.22)	(0.91)	(0.80)	(1.31)
Loss aversion	0.54*	0.69	0.65	0.32	0.47	0.51
	(0.32)	(0.59)	(0.61)	(0.27)	(0.36)	(0.39)
Risk aversion	–0.49*	–0.53*	–0.56*	–0.41**	–0.45**	–0.43**
	(0.26)	(0.30)	(0.31)	(0.20)	(0.19)	(0.17)
GPA	2.01**	2.46**	3.38***	1.48*	1.37*	1.44**
	(0.90)	(1.19)	(1.19)	(0.79)	(0.74)	(0.65)
Constant	–10.49**	–12.77	–15.05	–7.58*	–12.35	–15.86
	(4.59)	(15.37)	(16.83)	(4.18)	(10.37)	(13.21)
Class FE's	Yes	Yes	Yes	No	No	No
School FE's	No	No	No	Yes	Yes	Yes
Socioeconomic Controls	No	Yes	Yes	No	Yes	Yes
BIG 5, GRIT	No	No	Yes	No	No	Yes
Pseudo R^2	0.21	0.30	0.38	0.08	0.16	0.23
Observations	137	129	128	137	129	128

Multinomial Logit Regressions on *plans after graduation.* The upper panel reports coefficient estimates on the outcome "Job". The bottom panel reports coefficient estimates on the outcome "continued education". The table shows the coefficients of the multinomial logit regressions. Regressions only contain students with consistent answers and students who are expected to finish their education program in the year of the survey. Columns (1)-(3) include class fixed effects. Columns (4-6) include school fixed effects. Socio-economic controls include: female, 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 (2) and (3), as well as (5) and (6). Standard Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Appendix E Additional Regression Specifications on Job Search

Table E.1: Marginal Effects of Logit Regressions on active Job Search

	(1)	(2)
Patience (δ)	0.182	0.153
	(0.284)	(0.266)
Present Bias (β)	0.431***	0.405**
	(0.165)	(0.197)
Risk aversion	–0.057*	–0.042*
	(0.033)	(0.024)
Loss aversion	0.102	0.047
	(0.072)	(0.071)
Business (d)	0.106	0.090
	(0.137)	(0.137)
Polytechnician (d)	0.037	0.024
	(0.085)	(0.067)
female (d)	0.128**	0.100*
	(0.064)	(0.059)
german (d)	0.063	0.119
	(0.114)	(0.132)
age	–0.014	–0.027
	(0.042)	(0.034)
Math Grade	0.072	0.035
	(0.073)	(0.054)
German Grade	0.021	0.072
	(0.117)	(0.076)
English Grade	0.085	0.048
	(0.065)	(0.051)
borrowing difficulty	0.013	0.005
	(0.036)	(0.019)
CRT Score	–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		0.048
		(0.063)
Pseudo R^2	0.167	0.298
Observations	84	84

Logit Regressions on active job search. Columns (1)-(2) are equivalent to columns (2)-(3) in Table 9. 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). Std. Errors clustered at the class level are shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Table E.2: Marginal effects of a Logit regression on job search with additional controls

	Performance Pay	Firm Size	Final GPA
	(1)	(2)	(3)
Patience (δ)	0.23	0.17	0.00
	(0.28)	(0.29)	(0.32)
Present Bias (β)	0.36**	0.36**	0.45
	(0.16)	(0.17)	(0.34)
Risk aversion	–0.05*	–0.04*	–0.03*
	(0.02)	(0.02)	(0.02)
Loss aversion	0.03	0.04	0.07
	(0.06)	(0.07)	(0.09)
no performance pay (d)	–0.05		
	(0.14)		
performance pay (d)	–0.26		
	(0.28)		
<100 Employees (d)		0.08	
		(0.16)	
>100 Employees (d)		0.00	
		(0.15)	
GPA			–0.01
			(0.11)
Constant			
School FE's	Yes	Yes	Yes
Socioec. Controls	Yes	Yes	Yes
BIG5 / GRIT	Yes	Yes	Yes
Pseudo R^2	0.32	0.31	0.33
Observations	84	84	74

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) adds controls for performance pay; column (2) adds firm size controls, and column (3) controls for the grade point average in the final year. Std. Errors clustered at the class level shown in parentheses. Significance levels: *** p<.01, ** p<.05, * p<.1.

Appendix F Multiple Hypothesis Testing Corrections

Table F.1: P-Values of key coefficient estimates after multiple hypothesis correction using the Benjamini-Hochberg (1995) method

	Class F.E.'s			School F.E.'s			Lasso
	(3)	(4)	(5)	(8)	(9)	(10)	(11)
Patience in Table 3	0.09	0.06	< 0.01	0.11	0.06	0.03	0.03
	[0.15]	[0.09]	[0.05]	[0.20]	[0.09]	[0.05]	[0.1]
Patience in Table 4	0.09	0.06	0.03	0.19	0.06	0.04	0.06
	[0.15]	[0.09]	[0.09]	[0.25]	[0.10]	[0.09]	[0.1]
Present-bias in Table 7	0.02	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	
	[0.03]	[0.04]	[0.05]	[0.03]	[0.02]	[0.04]	[0.05]

Note: This table reports adjusted p-values of coefficient estimates in our main regression specifications after applying the Benjamini-Hochberg method to correct for false positive discovery rates (Benjamini and Hochberg, 1995). Values without brackets are adjusted p-values of the corresponding regression estimates after clustering at the class level. Values in square brackets are adjusted p-values after first using a score bootstrap (Kline and Santos, 2012)). Column labels (3)-(11) refer to the respective columns in Tables (3) and (4). For Table 7, the columns refer to columns (1)-(7), exactly in the order as presented here.