Research Report DFE-RR251

Department for Education

Financial incentives and working in the education sector

Professor Paul Dolan, Dr Robert Metcalfe &

Dr Daniel Navarro-Martinez

London School of Economics & University of Oxford

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

The Centre for the Understanding of Behaviour Change is an independent research centre with funding from the Department for Education. It is a partnership between leading researchers from the Universities of Bristol and Oxford, the Institute for Fiscal Studies, the National Centre for Social Research, the Institute of Education and the London School of Economics.

Executive Summary

Background

The Department for Education (DfE) is committed to improving the quality of applicants into initial teacher training (ITT). DfE and its predecessor Departments have offered bursaries to graduates wishing to train as teachers. Over time, these bursaries have been differentiated, by degree subject, with higher priority subjects attracting higher bursaries. From 2012/13, these financial incentives will be designed to make training to teach more attractive to the most talented graduates in the shortage subjects by increasing the level of this differentiation and also differentiating by degree class.

Whilst financial incentives are a way of encouraging applicants into ITT courses, there is little evidence on what types of students enter teaching, their preferences, and how they respond to financial incentives. This project uses an online experiment to address the two main questions that have not been fully addressed previously:

(a) What type of preferences and background characteristics are drivers of students choosing an ITT course?
(b) Are higher ability students impacted by higher endowments?

Methodology

Much of the research into why people teach has involved asking that question directly. Such approaches are at risk of being affected by conceptions of what is acceptable or desirable or by post hoc rationalisation. For this research, therefore, we adopted an experimental approach to try and eliminate these sources of bias.

For instance, to answer question a) we elicited students' risk and time preferences from choices between hypothetical gambles that we included in the experiment, and the students' pro-social inclinations by their willingness to donate to charity. Their career intentions, including likelihood of going into both primary and secondary school teaching after their undergraduate course we established through direct questioning. We also asked a range of standard trust, personality and ability questions.

To address participants' response to incentives as outlined in question (b), it was necessary to get them to engage in a task which required then to put in meaningful effort. We therefore drew on previous experiments on effort to design a task where people are paid a piece-rate system for exerting effort. We used an established approach where people have to place sliders on the mid-point of a series of lines. Subjects are rewarded for the number of correct

answers. The subject's “points score" in the task is the number of sliders positioned at 50 at the end of the allotted time, which was 120 seconds per screen. As the task proceeds, the screen displays the current point score and the time remaining. There were 42 sliders per screen and 20 screens (although subjects could stop at any time after screen 4). Each student received two pence per correct slider.

In addition to the effort-based reward, there was an initial up-front payment or “endowment” which students received. This was not conditional on how much effort participants put in to the experiment but on the subject and predicted classification of their degree, thus mimicking the incentives offered by ITT bursaries. To test the effect of this incentive, all students saw the whole endowment structure, just as they would observe different constraints on incentives for ITT for policy purposes. At the end of the experiment, the total winnings were the endowment plus the money earned from the slider task

To generate our sample, we emailed administrators in various University departments. We had a total of 1,574 participants in our experiment. The average age of the respondents was 21.34 years (standard deviation of 3.81). 38% were male and 62% were female. Compared with the whole undergraduate population (43% male v 57% female), women were slightly over-represented. There is no obvious reason why this should be the case, especially since the introduction to the experiment (see appendix C) should not induce any gender bias. Of the undergraduate population as a whole only 10% of undergraduates are in the social sciences (see www.hesa.ac.uk) compared with 20% of our sample. Compared with the population of actual ITT trainees our sample was not too dissimilar, with a slight bias towards the social sciences and (as we might expect) an underweighting of primary education specialists. Overall, our sample was more like the ITT trainee population than the overall undergraduate population.

Results

In relation to who wants to do ITT, we find:

1. There is a strong effect of gender on intention to become a teacher and to do ITT (women are considerably more likely to do this), but the effect seems to be driven completely by the intention to become a primary (not secondary) school teacher. This effect is well- documented in the research literature.
2. Third year students in the sample report a lower intention to become a secondary school teacher.
3. People in their third year report a lower intention to do ITT than people in their second year.

4. Those who trust others more have a stronger intention to become teachers and to do ITT, and this seems to be driven mainly by primary school teaching.
5. Giving importance to bursaries is strongly and positively associated with intentions to become a teacher and to do ITT, although the causality here might be in the opposite direction. It seems likely that having the intention to become a teacher leads to giving more importance to bursaries, instead of (or as well as) vice versa.

In relation to the effect of endowments and the incentive task, we find:

1. The size of the endowment assigned to participants does not have a significant impact on effort. In addition, the interactions of endowment with marks and with intention to do ITT are also insignificant, which indicates that the effect of the endowment is not significantly different for people with different marks of different intentions to do ITT.
2. The manipulations of the incentive framing (i.e. gain vs. loss and social information vs. example information vs. no information) had no significant effect on the effort exerted by participants.

Discussion

This research shows that the types of respondents who are most interested in becoming a teacher and in taking an ITT course are more likely to be female and care about bursaries. The gender effect is well-established in the literature, especially for primary school teaching. Given DfE interest in the potential for bursaries to attract high quality applicants to teaching, perhaps the most interesting of these findings given the motivation for this study is the fact that potential teachers care more about the potential bursaries than those not considering teaching. There are two possible effects at play here: (a) those that care more about bursaries are attracted to teaching by the bursaries; (b) those that are attracted to teaching use the bursaries to further justify their attraction to teaching.

The first effect is the standard ‘people respond to incentives’ mantra of economics. The presence of the bursary tips some of those who are interested in teaching ‘over the edge’. Without the bursary, some of these people would not be sufficiently incentivised to become a teacher. So, if they act as a sufficiently large incentive, bursaries serve to increase the supply of teachers.

The second effect draws on the psychological concept of ‘lay rationalism’. Lay rationalism posits that we need to provide reasons for our actions, ideally reasons that others can relate to. It is easier to explain a decision to go into teaching ‘because of the bursary’. It is therefore

important that the DfE seeks to understand the degree to which bursaries incentivise possible teachers *before the event* and the degree to which they provide rationalisations for going into teaching *after the event*.

Recommendations

In the meantime, there are some interesting findings here that can be used to inform the policy debates: firstly, about how bursaries impact on students; and secondly, about the importance of background and personality characteristics in influencing students' intentions to teach and the implications for marketing the profession to potential applicants. On the latter, there is much more which could be said about market segmentation and the marketing approaches which could be used to attract different groups, but this goes beyond the remit of this particular study and its data. This research offers some useful insight of its own:

1. In our experiment, offering higher endowments for priority subjects and degree classes did not impact negatively on the effort of other students. This would suggest that offering larger incentives for the most able students in high priority subjects and lower incentives for others would not affect applications for teaching in other, lower priority subjects. Unlike real bursaries, however, the endowments were small-scale and we cannot rule out the possibility that the differences were not large enough to find an effect, so this needs further testing. It is important to use the natural environment as a test bed for the impact of different types of bursaries – for instance by conducting natural field experiments where some people are randomly incentivised to go into teaching.

2. In our experiment, the ways the incentives are framed do not have a significant impact on effort. But again more research is needed to test whether this finding is unique to the parameters used in this study.

3. There is a clear gender effect, especially with regard to primary school, with women much more likely than men to express an intention to teach at primary level. This is a feature common to many countries, and there is no evidence to suggest that academic outcomes are adversely affected by it. There are other reasons, such as the provision of positive male role models for boys, why it may be desirable to have a more balanced teacher workforce. DfE has a programme to increase the number of male teachers in primary schools, and there is a body of literature on the reasons men do not become primary school teachers which is beyond the scope of this study to report, but there is still further investigation to be done in understanding how to get

men into teaching, the existing barriers which need to be overcome and what incentives might be most effective.

4. Again confirming the previous literature, our experiment shows that those with lower problem solving skills and 'A' level grades are more likely to express an intention become a teacher. As with gender, this finding is significant for primary teaching intentions only. While high grades do not guarantee teaching ability (or lower grades preclude it), such a finding is nonetheless bound to generate concern that teaching may be losing out to other professions when it comes to attracting high-achievers. While the Schools White Paper (DfE, 2010) notes that the average degree class of those entering initial teacher training has moved from below to above average in recent years it also notes that the status of the teaching profession has some way to go to catch up with higher-performing countries and there is still a case for DfE to consider targeted approaches to students with higher grades.

5. Those in the high priority subjects are more likely to be male, care more about the future, feel that they have a higher degree of autonomy, have higher problem-solving ability, have felt happy yesterday (at the time of answering the question), and report a weaker feeling that the things in their life are worthwhile. All of these factors should be taken into consideration when designing the marketing of ITT to attract students in the high priority areas.

1. Introduction

The Department for Education (DfE) is committed to improving the quality of applicants into initial teacher training (ITT). In their *The Importance of Teaching* White Paper (November 2010) and initial teacher training strategy (June 2011) and implementation plan (November 2011), the DfE set out a comprehensive programme of action to improve the recruitment, training and subsequent professional development of teachers. The three main areas in which the DfE attempted to do this were:

1. to raise the bar for entry to initial training: attracting more of the highest achieving graduates and having higher expectations of the academic and interpersonal skills of those funded to train to teach;
2. to refocus government investment in teacher training so that it is effective in attracting and retaining in teaching more of the best graduates, especially in shortage subjects;
3. to improve the routes through teacher training, so that it is easier to apply for teacher training and so that the nature and content of the training is more effective in preparing trainees to be successful in the classroom.

The second area is related to how to incentivise the best students into initial teacher training (ITT) courses. For those entering mainstream post-graduate ITT in 2012/13, the DfE are providing financial incentives (i.e. bursaries) designed to make training to teach more attractive to the most talented graduates, especially in shortage subjects. (DfE, 2011) This has led to different levels of bursaries for different subjects and the targeting of more money towards those they most want to attract (DfE, 2011).

		ITT specialism		
		Physics, mathematics, chemistry, modern languages	**Other priority secondary specialisms[1] and primary**	**General science and non-priority secondary specialisms[2]**
Training bursary 2012/13	**Trainee with first**	£20,000[3]	£9,000	£0
	2:1	£15,000[3]	£5,000	
	2:2	£12,000	£0	

Source: www.education.gov.uk/teachpgfunding

Whilst financial incentives are a way of encouraging applicants into ITT courses, there is little direct (behaviour-based) evidence on what types of students enter teaching, their preferences, and how they respond to financial incentives. This is the focus of this project. We seek to understand what types of students are in the high priority specialisms and the different ability levels, and how they respond to financial incentives. We will also seek to understand what types of students are interested in ITT courses, and how framing of the bursaries might impact on the motivation to become a teacher.

Examining the impact that financial incentives have on teacher recruitment is important, since getting the best teachers in to the classroom is a major concern. For instance, Sanders and Rivers (1996), Ferguson (1991), and Rivkin et al (1998) all suggest that teacher quality is one of the key determinants of educational attainment. Sanders and Rivers (1996) found that students who had strong teachers for three years in a row made reading gains over the period that were 54% higher than their fellow students who began at the same level but who had weak teachers for three consecutive years.

The structure of the report is as follows. Section 2 presents a brief review of relevant literature. Section 3 frames the research questions of the project and the methodology used to answer the questions. Section 4 states the experiment used to answer the questions, and section 5 presents the descriptive and multivariate analysis of the experiment. Section 6 presents the discussion, and frames the future questions that arise from the research.

2. What does the literature suggest?

There is some evidence on what types of people go into teaching. Women are more likely than men to enter teaching (Henke et al, 2000), although the proportion of female college graduates entering teaching has declined over time (Flyer and Rosen, 1997; Broughman and Rollefson, 2000; - both U.S. studies). There are many studies that find that students with the highest levels of measured ability tend not to go into teaching (e.g. Manski, 1987; Ballou, 1996; Gitomer et al, 1999; Henke et al, 2000; Stinebrickner, 2002; Podgursky et al, 2004), although some studies found that this holds primarily for early years school teachers rather than secondary school teachers. In addition, attitudinal characteristics are sometimes correlated with becoming a teacher e.g. Farkus et al (2000) found that teachers believed that loving the job was important, and that contributing to society was also important.

There are also studies that suggest that financial incentives, in terms of pay, may have a small positive impact on teacher recruitment. A few studies have attempted to disentangle

the impact of pay on teacher recruitment (Hanushek and Pace, 1995; Hounshell and Griffin, 1989). For instance, Evans (1987) and Rumberger (1987) found that increased salary for teachers led to an increased recruitment of mathematics and science teachers.

Dolton (1990) found that relative earnings in teaching (in comparison to non-teaching occupations) have a marked effect on graduates' choices to pursue teaching. What is crucial in his analysis is that these earnings effects operate on initial choices, and there is considerable inertia to remain in teaching (in terms of wages) because of the relative non-pecuniary rewards to teaching. In Dolton and Chung (2004), Chevalier et al (2006), and Dolton (2006), it is shown that the income stream from teaching relative to other professions has been declining for the last 25 years, meaning that now (on average) male teachers, compared to their counterparts in other professions, lose up to £40,000–67,000 in terms of the value of earnings over their lifetime whereas women gain approximately the same relative to an alternative occupation.

There are still some important gaps in the literature. These gaps include: (i) understanding the predictors of why people choose ITT courses in the UK; (ii) what type of students are affected by different levels of incentives; and (iii) can the financial incentives be framed in a particular way that it can improve people's effort during tasks.

3. Research questions and methodology

While research has been done on the characteristics motivations of people to teach (including their response to incentives), much of this has centred around asking direct questions. Such approaches are at risk of being affected by conceptions of what is acceptable or desirable, or by post hoc of rationalisation. This research therefore we use an innovative online experiment to help us address the following three main questions that have not been fully addressed previously:

(a) What type of preferences and background characteristics are drivers of students choosing an ITT course?
(b) Are higher ability students impacted by higher monetary rewards?
(c) Does the framing of the incentives impact on people's efforts?

For question (a), the research was based around eliciting students', we used a range of standard tests designed to elicit personality traits such as neuroticism, self-determination, and trust, while risk and time preferences were established using choices between hypothetical gambles that we included in the experiment. Students' pro-social inclinations

were elicited by their willingness to donate to charity. We also asked a number of direct questions about their likelihood of going into both primary and secondary school teaching after their undergraduate course, and their intentions about entering ITT and other career options. For (b) and (c), we used a range of different options which randomly allocated students to different "treatments" offering a range of different ability- and effort-related incentives. The latter drew on previous experiments on effort to design a task where people were paid a piece-rate system for exerting effort.

4. The experiment

4.1 The effort experiment

Real and financially incentivised "effort tasks" (i.e., tasks focussed on making participants exert a systematic and measurable effort) will increase the external validity of the results from the experiment, compared to hypothetical questions. We draw on previous attempts by economists to examine real effort in the lab. The approaches include: solving mazes (Gneezy et al, 2003), mathematical problems (Sutter and Weck-Hannemann, 2003) or word games (Burrows and Loomes, 1994); answering general knowledge questions (Hoffman et al, 1994); decoding (Chow, 1983) or entering (Dickinson, 1999) strings of characters; performing numerical optimisation (van Dijk et al, 2001); filling envelopes (Konow, 2000), cracking walnuts (Fahr and Irlenbusch, 2000) or other physical tasks. All of these tasks involve effort and/or abilty.

Two recent innovative papers have used slightly different tasks to understand effort. In both games, people are rewarded for the number of correct answers, so it represents a piece rate payment system – which is something that would suit our approach: Abeler et al (2010) use a "count the number of zeros" task from a screen containing both zeros and ones; Gill and Prowse (2012) use a slider task, where people have to place a slider on the mid-point of fifty lines. This task is probably the most effortful task that does not use much ability (although some hand-eye coordination is clearly needed).

We use the approach by Gill and Prowse (2012), where people have to place sliders on the mid-point of a series of lines. In this task, subjects are rewarded for the number of correct answers, so it represents a piece rate payment system (see appendix A for a more detailed description of the task).

The screen does not vary across experimental subjects or across repetitions of the task. A schematic representation of a single slider is shown in the figure below. When the screen containing the effort task is first displayed to the subject all of the sliders are positioned at 0, as shown for a single slider in the Figure below.

(a) Initial position. (b) Positioned at 50.

By using the mouse, the subject can position each slider at any integer location between 0 and 100 inclusive. Each slider can be adjusted and readjusted an unlimited number of times and the current position of each slider is displayed to the right of the slider. The subject's "points score" in the task, interpreted as effort exerted, is the number of sliders positioned at 50 at the end of the allotted time, which was 120 second per screen. Figure (b) shows a correctly positioned slider. As the task proceeds, the screen displays the subject's current point score and the amount of time remaining. There were 42 sliders per screen and a total of 20 screens (although participants could stop the task at any time after screen 4). Each student received two pence per correct slider. So if they obtained 42 correct sliders in one screen, they would receive £0.84.

This task has a number of desirable attributes which make it particularly suitable for our experiment. First, it is simple, and does not require pre-existing knowledge. Second, the task is identical across repetitions. Third, it involves little randomness, so the number of correctly positioned sliders corresponds closely to the effort exerted by the subject.

To ensure that participants were fully familiar with the requirements of the effort task, prior to the main task they undertook a practice session, which involved completing four screens of sliders.

4.2 The incentive structure

In addition to the effort task described above, there were other incentives in the form of "endowments". These were not conditional on how much effort participants put in to the experiment, but on subject of participants' current undergraduate course and predicted degree performance based on their scores in the examinations at the last university year

completed. This mimicked the incentives offered by ITT bursaries,[1] in that we used the same relative levels of reward (albeit at much lower absolute values) as those offered by the DfE bursaries (as shown below in endowment structure A).

Table A: Endowment structure A

	High priority specialisms	Medium priority specialisms	Non-priority specialisms
Outstanding potential (1st)	£10.00	£4.50	£0
Good potential (2.1)	£7.50	£2.50	£0
Satisfactory potential (2.2)	£6.00	£0	£0
Low potential (3rd)	£0	£0	£0

Table B: Endowment structure B

	High priority specialisms	Medium priority specialisms	Non-priority specialisms
Outstanding potential (1st)	£14.00	£4.50	£0
Good potential (2.1)	£8.50	£1.00	£0
Satisfactory potential (2.2)	£6.00	£0	£0
Low potential (3rd)	£0	£0	£0

We also constructed, in consultation with DfE, an additional endowment structure B to investigate whether a more polarised incentive structure, with greater difference in rewards for those in higher priority subjects and of higher ability compared with others, affected participant responses. In particular we sought to discover whether it adversely affected the responses of those receiving smaller endowments and whether these participants exerted less effort in the task as a result of the greater "distance" between their own reward and rewards at the top of the scale. To investigate this issue, all students saw the whole endowment structure, just as they would observe different constraints on incentives for ITT for policy purposes (see appendix B for the visual representation of this).

At the end of the experiment, the total winnings for each participant were their endowment plus the money earned from the slider task.

[1] http://www.education.gov.uk/schools/careers/traininganddevelopment/a0078019/training-outstanding-teachers. The high priority specialisms – physics, mathematics, chemistry, modern languages. Medium priority specialisms – Art and design, design and technology, economics, engineering, English, dance, drama, geography, history, information and communications technology, computer science, classics, music, biology, physical education, primary, and religious education. Non-priority specialisms – General science, business studies, citizenship, applied science, health and social care, leisure and tourism, media studies, psychology, social sciences (except economics).

4.3 The treatments, questions, and elicited preferences

In addition to testing the effect of incentives and their relative size, we looked at how they were framed, in particular the effects of loss aversion and social norms. These are well-established behavioural phenomena and which have been shown to affect behaviour in previous experiments. Loss aversion refers to our tendency to feel the pain of a loss more acutely than the pleasure of a gain of an equivalent size. Social norms refer to our tendency to want t0 be like other people who we consider to be like us.

To address loss aversion, some participants were told that they would lose money instead of gaining money. So notionally they were allocated the full money and money was taken away if they did not complete the experiment. To examine the effects of social norms, some participants were given a social prime or "peer" anchor (i.e. information on others' performance based on the results from a pilot study) and others an anchor presented simply as an example of the number of sliders that a participant could complete. To ensure comparability between the groups, the number given as the example was exactly the same as that given for the social anchor.
Combining these different experiments we had twelve different treatment groups to which students were randomly allocated as shown below:

Group	Endowment	Gain/loss frame	Peer/example anchor
1	Low (A)	Gain	None
2	Low (A)	Loss	None
3	High (B)	Gain	None
4	High (B)	Loss	None
5	Low (A)	Gain	Example
6	Low (A)	Loss	Example
7	High (B)	Gain	Example
8	High (B)	Loss	Example
9	Low (A)	Gain	Peer
10	Low (A)	Loss	Peer
11	High (B)	Gain	Peer
12	High (B)	Loss	Peer

We also asked participants a number of questions to establish their key characteristics so that we could investigate whether these had a relationship with intention to teach. In addition to demographic characteristics such as gender, age and type of school attended we asked questions about

- ability/achievement
- personality traits
- time, risk and social preferences
- employment plans

- attitudes to bursaries

Some of these questions were asked prior to and some after the effort task described above.

Ability (or achievement) was measured by previous educational attainment. We asked questions about the average mark obtained in the first or second year of undergraduate courses, which university participants attended, which course they were doing and number of A and B grades at A-level. We also asked participants to complete problems with a mathematical structure similar in style to those used in the Graduate Management Admissions Test (GMAT), which are widely used to assess problem solving ability.[2]

Other questions elicited personality traits such as neuroticism and extraversion (from the 'big 5' personality trait questionnaire – a tool commonly used in psychological research and independently validated by several research teams (e.g McCrae and Costa, 1987), subjective wellbeing (SWB), self-determination and trust. The SWB questions were based on the 'ONS Four' used by the Office for National Statistics (Dolan et al, 2011; Dolan and Metcalfe, 2012), that is: life satisfaction, happiness yesterday, anxiety yesterday, and worthwhile activities. The self-determination questions were based on Ryff (1989) and examined autonomy, competence/accomplished and relatedness.[3] The trust question was whether you thought other people can be trusted (taken from Glaeser et al, 2000).

We elicited time, risk and social preferences. Time preference is a concept used in economics to assess the extent to which someone values a positive event or outcome now or in the near future compared to one in the more remote future. While everyone values a positive event in the near future more than the same outcome in the more distant future, if they place a much higher value on a near-future outcome than a more distant one, they are said to discount the future more than someone who values near- and distant-future events more equally. Social preferences recognise that people do not always act out of self-interest alone, but have values of fairness, altruism and generosity.

To examine participants' time preference we used hypothetical questions of the type: "*If you had a choice of receiving £1,000 today or £1,200 in one years' time, which would you choose? If you had a choice of receiving £1,000 in one year's time or £1,200 in two years'*

[2] The four GMAT questions used were: (1) If half of the money in a mutual fund was invested in stocks, one fifth in bonds and the remaining $300,000 in cash, what was the total amount of the mutual fund?; (2) In a local intramural basketball league, there are 10 teams and each team plays every other team exactly one time. Assuming that each game is played by only two teams, how many games are played in total? (3) Solution x contains 75% water and 25% oil; how many more liters of water than liters of oil are in 200 liters of solution x?; (4) If you look at a clock and the time is 3.15, what is the angle between the hour and the minute hands? Respondents were not told the correct answer.

[3] These three are elicited as follows: "autonomy": response to "I am free to decide for myself how to live my life"; "competence": response to "most days I feel a sense of accomplishment from what I do"; "relatedness": response to "people in my life care about me".

time, which would you choose?" Attitudes to risk were elicited using questions of the form "*Imagine that you win £100,000 in a lottery. Before collecting your money, you are given the following option by a reputable bank: "You can invest a part of your lottery money in an investment plan in which you have a 50% chance of doubling the investment in one year, but there is also a 50% chance of losing half of the amount invested." What fraction of the £100,000 would you choose to invest?"* We gave students twelve pairs of these.

Social preferences were explored through the incentives offered by the experiment: we asked participants whether they would be willing to give either 10% or 30% of their winnings to one of two charities. This allowed us to understand their preferences for other people and organisations.

Importantly, we also asked about employment plans (from a list of standard professions), and the likelihood of going into an ITT course and pursuing primary and secondary school teaching (as three separate questions on a scale from 0 to 10). We also asked whether bursaries are important when thinking about careers. These were stand-alone questions toward the end of the experiment and were created after discussions with DfE.

To generate our sample, we emailed administrators in various University departments to recruit second or third year students. We did not skew the emails toward a particular subject area. The email came from the project team and stated that the DfE funded the research. We used the following diverse (in the sense of location, ability levels, and speciality) mix of English universities for our sample: Birkbeck College, Brunel University, Kingston University, Lancaster University, London South Bank University, London School of Economics, Oxford Brookes University, University College London, University of Birmingham, University of Buckingham, University of Cambridge, University of Kent, University of Manchester, University of Newcastle, University of Oxford, University of Reading, University of Sheffield. We did not approach more universities given resource constraints.

So to summarise the setup, the approach of the overall experiment was as follows:

1. Email administrator of University department for forward the experiment link;
2. Administrator offers students to take part in the experiment (via the link);
3. Student clicks on link and is directed to the online experiment;
4. Each student answers wellbeing questions and questions about ability, and then reads instructions about the slider task;
5. Each student completes a practice version of the slider task;

6. Each student is then told their endowment, and then undertakes the piece-rate slider task (some have been randomised in to receiving high/low endowments, gain/loss framing, and peer information/example anchor/no information);
7. At end of slider task, they are offered the possibility to give some of their winnings to charity (to elicit social preferences);
8. All students answer questions relating to hypothetical risk and time preferences, personality, demographics, and then questions relating to attitudes toward teaching and intentions about the future.

5. Analysis

5.1 Sample statistics

Tables D1-D4 in Appendix D summarise the sample. We had a total of 1,574 participants in our experiment. 38% were male and 62% were female. The average age of the respondents was 21.34 years (standard deviation of 3.81). The full age distribution can also be found in Appendix D. The majority of participants were white (74%), followed by Asian (17%), black (4%) and mixed race (3%). Clearly, there were more in the sample studying social science (20%) than any other subject.

These splits are similar to the overall statistics for the undergraduate population, although we have slightly more women in the experiment than we might expect. There is no obvious reason why this should be the case, especially since the introduction to the experiment (see appendix C) should not induce any gender bias. The whole UK undergraduate population had greater numbers of females than males, (43% male and 57% female) and only 10% of undergraduates were in the social sciences (see www.hesa.ac.uk) - these are the main two differences with our sample. Our sample was not too dissimilar to that of actual ITT trainees where, in the 2009/10 cohort, 86% were white and 72% were female (figures provided by DfE). In terms of subject priority we find that 9.5% of our sample is in high priority subjects, 57% is in medium priority subjects, and 33% is in low priority subjects. Overall we seem to have a relatively representative sample of UK university students, with a bias towards the social sciences.

We have a high proportion (29%) from London universities, but we have students from universities outside London, such as Lancaster, Birmingham, Kent, Manchester, Newcastle, Reading and Sheffield.

A large proportion (15%) of students want to become teachers. This is larger than any other single profession chosen but similar to financial services (accountant, actuary and financial analyst). The full distributions of responses to the intention of doing ITT and pursuing primary or secondary teaching as a career are shown in Appendix D, together with cross-tabulations of intentions to do ITT by priority of degree subject and by marks obtained. We also asked participants what subject they would want to specialise in if they did an ITT course. It seems that the most popular choice amongst people from our sample was to teach English, closely followed by Mathematics.

5.2 Characteristics of those with intentions to teach

We begin by conducting a regression analysis of the determinants of teaching and ITT intentions. Regression analysis is one of the most widely used statistical techniques and it is the appropriate tool to capture the effects of a series of explanatory variables or factors on a target variable (normally called the dependent variable). The regressions presented here capture the size of the effects of the explanatory variables on the dependent variable used and also whether the effects are statistically significant or not.

To begin with, we present four different regressions (captured in Table E1 in Appendix E, which presents detailed regression results) to capture the effects of the same group of explanatory variables on four different dependent variables. The first one is a logistic regression using "*intention to become a teacher (1) or not (0)*" as a dependent variable. This variable is constructed from the list of career intentions shown in the descriptive statistics. The other three regressions are standard linear regressions, but using robust standard errors. The dependent variables are "*score (from 0 to 10) on intention to become a primary school teacher*", "*score (from 0 to 10) on intention to become a secondary school teacher*" and "*score (from 0 to 10) on intention to do ITT*". So, overall, these four regressions capture the effects of the explanatory variables on intentions to become a teacher and to do ITT.

Before we provide the regressions, we note that we find strong and significant correlations between the four dependent measures (from 0.3 to 0.7). This shows that these variables are actually capturing related constructs and that participants responded consistently to these questions, which is very reassuring for our sample. Note also that two of these correlations involve correlating a binary variable with a variable with eleven categories, which is expected to decrease the correlations because of the very different variable structure.

Table E1 presents the full regression results for the four dependent variables measuring intentions to go into teaching or ITT. Among other things, these regressions allow us to show which preferences and background characteristics are drivers of students choosing to do an ITT course (question (a) in section 3). All the effects are conditional on controlling for the rest of the explanatory variables. The main results can be summarised as follows. The first thing to note is that many of the characteristics are not significant and it is likely that there are important characteristics associated with a desire to teach which our analysis does not capture. A number of important variables do emerge, however:

Demographic Characteristics

1. There is a strong effect of gender on intention to become a teacher and to do ITT (women are considerably more likely to do this), but the effect seems to be driven completely by the intention to become a primary (not secondary) school teacher.
2. Third year students in the sample report a lower intention to become a secondary school teacher.
3. White people appear as more likely to pursue teaching as a career, but only in the first regression (using the binary dependent variable derived from the career intention list), so that this result is somewhat inconclusive.
4. Those who vote for the Conservative Party report a lower intention to pursue teaching as a career and to do ITT, and this seems to be driven mainly by a lower intention to do primary school teaching. Those who vote for the Labour Party report a stronger intention to become secondary school teachers.
5. People in their third year report a lower intention to do ITT than people in their second year.

Personality Traits

6. Those who trust others more have a stronger intention to become teachers and to do ITT, and this seems to be driven mainly by primary school teaching.
7. People with a higher degree of neuroticism appear as less inclined to pursue teaching as a career, but only in the first regression (using the binary dependent variable derived from the career intention list), so that this result is somewhat inconclusive.
8. Those with a higher degree of reported autonomy seem less inclined towards teaching.
9. People's perceived competence is consistently and positively associated with intention to become a teacher (both primary and secondary) and to do ITT.

Ability

10. The scores from our problem-solving, GMAT-style questions are consistently and negatively associated with intentions to become a teacher and to do ITT (i.e. people with higher problem-solving abilities are less inclined to do this), but the effect seems to be driven mainly by the intention to become a primary (not secondary) school teacher.
11. Similarly, a higher number of A marks at A-level appears as negatively associated with the intention to become a teacher. As with problem-solving ability this seems to be driven mainly by the intention to become a primary school teacher.

Attitude to Bursaries

12. Giving importance to bursaries is strongly and positively associated with intentions to become a teacher and to do ITT, although the causality here might be in the opposite direction. It seems likely that having the intention to become a teacher leads to giving more importance to bursaries, instead of (or as well as) vice versa.

So, overall, people who express a stronger intention to pursue teaching as a career and do ITT are more likely to be female (only in primary), younger (only in secondary), more inclined to vote Labour than Conservative, trusting (only in primary), less autonomous, have a greater perceived competence, have lower problem solving abilities (only in primary), have lower marks at A-levels (only in primary) and give importance to bursaries,. Altogether, they contribute to provide a much clearer picture of the kinds of people pursuing teaching as a career and doing ITT.

5.3 Effort task results

We now turn to a regression analysis of the determinants of the effort made in the experimental effort task. This will help us to answer questions (b) and (c) from section 3: are higher ability students impacted by higher endowments; and does the framing of the incentives impact on people's efforts? This part contains the results of three linear regressions (shown together in one table), using as dependent variables "*total number of sliders done in the effort task*", "*average number sliders per screen in the effort task minus average number of sliders per screen in practice*", and "*number of screens done in the effort task*".

Table 1 presents the overall descriptive statistics relating to the effort task. Note that the task consisted of 20 screens with 42 sliders per screen, and participants could stop the task at any time after screen 4 (see explanation in Section 4.1 and in Appendix A). "Sliders done"

means sliders correctly set at 50; “screens attempted” means screens that the participants faced before stopping (or finishing) the task.

Table 1: Descriptive statistics related to the effort task

Average number of sliders done in the task	519.64
Average number of screens attempted in the task (out of 20)	16.24
Percentage of people attempting all the 20 slider screens	62.58%
Average number of sliders per screen in practice	27.21
Average number of sliders per screen in the task (screens 1 to 4)	31.98
Average number of sliders per screen in the task	32.65
Average number of sliders per screen in the task (screens 1 to 4) for those who do all screens	32.55
Average number of sliders per screen in the task for those who do all screens	33.04

Once again, before we turn to the regressions themselves, we consider whether the dependent variables are capturing similar constructs, which (all being effort-related measures) we would expect them to do. The correlations between the dependent variables (Spearman) are as follows: “total sliders” with “sliders task minus sliders practice”: 0.11; and “total sliders” with “screens”: 0.82. While both are statistically significant, we consider the first correlation very low and the second one very high. The low correlation may indicate that the variable “sliders task minus sliders practice” is affected by aspects of practice behaviour not necessarily related to effort. For example, certain types of people might simply do fewer sliders in the practice stage because they are inspecting the set-up and familiarizing themselves with how the sliders work – an activity which is largely unrelated to effort. We consider the total number of sliders in the actual task the main effort variable, noting that this also determines part of the overall reward which participants receive (in addition to the initial endowment).

Table E2 in the Appendix presents the linear regressions with the three different dependent variables.

The results obtained in the third regression (in which the dependent variable is the difference between practice and task) are quite different from the other two regressions, and they even contradict the other regressions in some cases. As we note in the discussion of the correlations above, these results might reflect differences in practice behaviour that are not related to effort so should be treated with caution. For this reason, the results of the third regression are not included in the summary of results given below (but see the regression table for more information).

The main results can be summarized as follows. Many of variables are insignificant and the amount of variance explained by the regressions is only around 10 per cent, so clearly this

analysis does not include every characteristic which might affect the effort exerted by participants. However, a number of important characteristics are associated with effort in the task (a shorter and more focused list is provided in the paragraph after this enumeration):

Characteristics of the Incentives

1. The size of the endowment assigned to participants does not have a significant impact on effort. In addition, the interactions of endowment with marks and with intention to do ITT are also insignificant, which indicates that the effect of the endowment is not significantly different for people with different marks of different intentions to do ITT.
2. The manipulations of the incentive framing (i.e. gain vs. loss and social information vs. example information vs. no information) had no significant effect on the effort exerted by participants.

Demographic Characteristics

3. Gender is significantly associated with the level of effort in the task. Women consistently exert less effort than men in the task.
4. Those in their third year exert more effort than second-year students, in terms of the total number of sliders done (the effect is not significant in terms of number of screens).

Personality Traits

5. Those who trust others more exert more effort in the task.
6. Those who discount the future more exert less effort in the task.
7. Relatedness is significantly and positively associated with effort.
8. Those who contribute to charity exert significantly less effort in the task. Note that causality here might be in the opposite direction. People who exert more effort might then be less willing to give part of the money to charity.

So, overall the relative size of the endowment and the framing of the incentives had no effect on the amount of effort exerted by participants – a result for which we provide further verification in section 5.5. In addition, people who exert more effort in our task are more likely to be male, trusting, discount the future less, have a greater degree of relatedness (i.e., inclination to interact with others), and be less pro-social (i.e., less inclined to give to charity).

5.4 Characteristics of high-performing students and those taking high priority subjects

We also performed an additional regression analysis on the determinants of the priority of the degree subject chosen by students and of the marks obtained so far at the undergraduate level. Two regressions capturing this analysis are shown together in Table

E3. Both of them are ordered logistic regressions using as dependent variables “priority of degree subject” (from 1 to 3, with 1 being low priority and 3 high-priority) and “mark obtained” (from 1 to 4, with 1 being 1st and 4 being 3rd). The factors related to having chosen a high-priority degree subject are found to be: being male, discounting the future less, having a higher degree of autonomy, having a greater problem-solving ability, having felt happy yesterday (at the time of answering the question), and reporting a weaker feeling that the things in their life are worthwhile.

The main determinants of having a better mark are found to be: having a lower degree of extraversion, a higher degree of competence, a greater problem-solving ability, and being less inclined to give to charity (in terms of the contributions made as part of our experiment). They also help to have a clearer picture of the determinants of choosing high-priority subjects and having a high achievement level.

5.5 Effect of different experimental conditions

We manipulated three main factors in the experiment: 1) different levels of monetary endowment, 2) the information given to participants (i.e. giving information about others’ behaviour in the effort task or giving that information in the form of an example or giving no information), and 3) using a gain or a loss framing to explain the money obtained in the effort task. The first manipulation is aimed at testing the effects of bursaries on the effort exerted by people; the second and third are intended to investigate the moderating effects of social information (i.e., information on others’ behaviour) and loss aversion on effort.

The details of the results are shown in Appendix F. Table F1 summarizes the effects of these manipulations on the effort variables and Table F2 shows the effects on the intention to teach and to do ITT variables. Note that manipulations 2 and 3 are largely unrelated to intention to teach or do ITT, so only the results of manipulation 1 are summarised in F2. Both tables contain p-values resulting from pair-wise, non-parametric, Mann-Whitney tests. In the case of the endowment manipulation, the comparison of the conditions is also done for specific cells of the endowment tables in which the difference between scenarios A and B is particularly salient.

The main conclusion of all these tests is that none of the main experimental manipulations implemented had significant effects on any of the variables analysed, which is in line with the results obtained in the regression analysis of Table E2 for the effort variables. More specifically, none of the p-values reported is below 0.05, which we consider the appropriate

significance level in this case. The note to Table F2 also shows the p-value resulting from a Kruskal-Wallis test comparing across the twelve different conditions, which is not significant.

Finally, Tables 2A, 2B and 3A, 3B provide descriptive information on the sample size, teaching and ITT intentions and effort exerted in the different cells of the endowment tables. The main numbers contained in the tables are averages, and in some cases standard deviations are provided in brackets for additional information. It is important to note that these tables have been included as additional illustrative descriptive statistics, but all the relevant results deriving from them have been discussed and statistically tested in the previous analyses. Note also that most of the differences between the cells are not statistically significant and should not be taken as evidence of any systematic effects.

Table 2A: Low-endowment condition – intentions to teach and do ITT

	High-priority	Medium-priority	Low-priority	TOTALS
1st	*Endow.: £10* N = 29 ITT = 3.83 (2.99) Prim.=1.93(2.63) Sec. = 3.28 (2.84)	*Endow.: £4.50* N = 116 ITT = 3.31 Prim. = 1.71 Sec. = 2.66	*Endow.: £0* N = 37 ITT = 3.89 Prim. = 2.89 Sec. = 2.59	N = 182 ITT = 3.51 Prim. = 1.98 Sec. = 2.75
2:1	*Endow.: £7.50* N = 20 ITT = 3.22 Prim. = 1.47 Sec. = 2.74	*Endow.: £2.50* N = 250 ITT = 4.24 Prim. = 2.67 Sec. = 3.12	*Endow.: £0* N = 138 ITT = 3.65 Prim. = 2.42 Sec. = 2.10	N = 408 ITT = 4.00 Prim. = 2.53 Sec. = 2.76
2:2	*Endow.: £6* N = 14 ITT = 4.07 Prim. = 1.14 Sec. = 3.36	*Endow.: £0* N = 55 ITT = 3.42 Prim. = 2.20 Sec. = 2.82	*Endow.: £0* N = 54 ITT = 3.22 Prim. = 2.59 Sec. = 2.30	N = 123 ITT = 3.41 Prim. = 2.25 Sec. = 2.65
3rd	*Endow.: £0* N = 3	*Endow.: £0* N = 4	*Endow.: £0* N = 7	N = 14
TOTALS	N = 66 ITT = 3.58 Prim. = 1.54 Sec. = 2.98	N = 425 ITT = 3.90 Prim. = 2.35 Sec. = 2.96	N = 236 ITT = 3.58 Prim. = 2.52 Sec. = 2.25	*N = 727* *ITT = 3.77* *Prim. = 2.33* *Sec. = 2.73*

Table 2B: High-endowment condition – intentions to teach and do ITT

	High-priority	Medium-priority	Low-priority	TOTALS
1st	*Endow.: £14* N = 40 ITT = 4.16 Prim. = 1.47 Sec. = 3.00	*Endow.: £4.50* N = 96 ITT = 3.19 Prim. = 1.71 Sec. = 2.69	*Endow.: £0* N = 43 ITT = 2.85 Prim. = 1.34 Sec. = 1.90	N = 179 ITT = 3.32 Prim. = 1.57 Sec. = 2.57
2:1	*Endow.: £8.50* N = 22 ITT = 4.41 Prim. = 1.68 Sec. = 3.55	*Endow.: £1* N = 248 ITT = 4.06 Prim. = 2.56 Sec. = 2.82	*Endow.: £0* N = 138 ITT = 3.84 Prim. = 2.63 Sec. = 2.27	N = 408 ITT = 4.00 Prim. = 2.53 Sec. = 2.67
2:2	*Endow.: £6* N = 18 ITT = 4.07 Prim. = 2.21 Sec. = 3.29	*Endow.: £0* N = 82 ITT = 4.37 Prim. = 2.37 Sec. = 3.18	*Endow.: £0* N = 69 ITT = 3.57 Prim. = 2.32 Sec. = 2.20	N = 169 ITT = 4.01 Prim. = 2.34 Sec. = 2.78
3rd	*Endow.: £0* N = 1	*Endow.: £0* N = 6	*Endow.: £0* N = 6	N = 13
TOTALS	N = 81 ITT = 4.29 Prim. = 1.65 Sec. = 3.31	N = 432 ITT = 3.91 Prim. = 2.33 Sec. = 2.85	N = 256 ITT = 3.58 Prim. = 2.31 Sec. = 2.18	*N = 769* *ITT = 3.85* *Prim. = 2.25* *Sec. = 2.68*

Table 3A: Low-endowment condition – effort

	High-priority	Medium-priority	Low-priority	TOTALS
1st	*Endow.: £10* N = 29 Sliders = 526.75 (235.18) Screens = 16.03 (5.83)	*Endow.: £4.50* N = 116 Sliders = 549.09 (217.47) Screens = 16.67 (5.25)	*Endow.: £0* N = 37 Sliders = 518.92 (220.63) Screens = 15.92 (5.66)	N = 182 Sliders = 539.47 Screens = 16.42
2:1	*Endow.: £7.50* N = 20 Sliders = 439.74 (251.39) Screens = 14.05 (6.48)	*Endow.: £2.50* N = 250 Sliders = 517.37 (221.76) Screens = 16.24 (5.57)	*Endow.: £0* N = 138 Sliders = 504.70 (219.50) Screens = 16.28 (5.70)	N = 408 Sliders = 509.44 Screens = 16.15
2:2	*Endow.: £6* N = 14 Sliders = 648.79 (184.09) Screens = 18.57 (4.04)	*Endow.: £0* N = 55 Sliders = 547.8 (203.72) Screens = 16.93 (4.88)	*Endow.: £0* N = 54 Sliders = 546.81 (212.99) Screens = 17.13 (5.02)	N = 123 Sliders = 558.86 Screens = 17.20
3rd	*Endow.: £0* N = 3	*Endow.: £0* N = 4	*Endow.: £0* N = 7	N = 14
TOTALS	N = 66 Sliders = 534.92 Screens = 16.09	N = 425 Sliders = 529.61 Screens = 16.42	N = 236 Sliders = 513.97 Screens = 16.41	*N = 727* *Sliders = 525.12* *Screens = 16.39*

Table 3B: High-endowment condition – effort

	High-priority	Medium-priority	Low-priority	TOTALS
1st	*Endow.: £14* N = 40 Sliders = 553.39 (250.29) Screens = 15.71 (6.12)	*Endow.: £4.50* N = 96 Sliders = 564.10 (235.76) Screens = 16.44 (5.63)	*Endow.: £0* N = 43 Sliders = 530.78 (226.04) Screens = 16.17 (5.77)	N = 179 Sliders = 553.85 Screens = 16.21
2:1	*Endow.: £8.50* N = 22 Sliders = 504.45 (245.04) Screens = 15.23 (6.51)	*Endow.: £1* N = 248 Sliders = 513.03 (222.80) Screens = 16.31 (5.64)	*Endow.: £0* N = 138 Sliders = 474.44 (218.83) Screens = 15.37 (5.54)	N = 408 Sliders = 499.52 Screens = 15.94
2:2	*Endow.: £6* N = 18 Sliders = 543.71 (274.47) Screens = 15.64 (5.92)	*Endow.: £0* N = 82 Sliders = 521.37 (224.32) Screens = 16.37 (5.39)	*Endow.: £0* N = 69 Sliders = 515.88 (208.73) Screens = 16.4 (5.53)	N = 169 Sliders = 521.09 Screens = 16.32
3rd	*Endow.: £0* N = 1	*Endow.: £0* N = 6	*Endow.: £0* N = 6	N = 13
TOTALS	N = 81 Sliders = 532.93 Screens = 15.45	N = 432 Sliders = 523.60 Screens = 16.31	N = 256 Sliders = 490.41 Screens = 15.70	*N = 769* *Sliders = 513.94* *Screens = 16.02*

6. Concluding remarks

6.1 Summary

One of the most interesting and relevant set of findings from this research relate to the types of respondents who are most interested in becoming a teacher and in taking an ITT course. Such people are more likely to:

1) be female (not for secondary teaching);
2) feel competent;
3) have lower problem solving abilities (only in the case of primary teaching);
4) care about bursaries.

The gender effect is well-established in the literature, especially for primary school teaching. The overall effect of higher feelings of competence and lower problem solving abilities is ambiguous in terms of its effect on teacher quality, though it is entirely possible that the former positive effect outweighs the latter negative one.

Perhaps the most interesting of these findings given the motivation for this study is the fact that potential teachers care more about the potential bursaries than those not considering teaching. There are two possible effects at play here:

1. Those who care more about bursaries are attracted to teaching by the bursaries.

2. Those who are attracted to teaching use the bursaries to further justify their attraction to teaching.

The first effect is the standard 'people respond to incentives' mantra of economics. The presence of the bursary tips some of those who are interested in teaching 'over the edge'. Without the bursary, some of these people would not be sufficiently incentivised to become a teacher. So, if they act as a sufficiently large incentive, bursaries serve to increase the supply of teachers.

The second effect draws on the psychological concepts of 'lay rationalism' and 'motivational crowding in'. Lay rationalism posits that we need to provide reasons for our actions, ideally reasons that others can relate to. It is easier to explain a decision to go into teaching 'because of the bursary'. Motivational crowding in operates in a similar way. There is an intrinsic motivation to go into teaching that is then 'crowded in' by the bursary. Under such conditions, it is possible that bursaries increase the supply of teachers, but they may not if these reasons are merely 'post-hoc' rationalisations for a decision that has already been made.

Indeed, the two main caveats to our research are that, we have used intentions to teach and not observed actual applicant behaviour, and that we have used an online experiment and not observed behaviour when people are not in an experiment (and hence we have some selection bias that limits our external validity of the research). It is important that the DfE seeks to understand the degree to which bursaries incentivise possible teachers *before the event* and the degree to which they provide rationalisations for going into teaching *after the event*. Some of our behaviour is influenced by situational contextual factors that often lie below conscious awareness. So much of what we do is not so much thought about; it simply comes about. As a consequence, intentions and recollections are notoriously poor guides to the real reasons for behaviour. Efforts should therefore be directed at testing the effects of bursaries on real behaviour; that is, to use the natural environment as a test bed for the impact of different types of bursaries. This requires us to think more about conducting field experiments where some people are randomly incentivised to go into teaching.

Ultimately, this will place DfE in a better situation to actually assess how effective scarce resources are in attracting and retaining the best graduates in to teaching, especially in shortage subjects. This lab experiment has provided some useful insights into where that future research might be directed.

6.2 Recommendations

There are some interesting findings here that can be used to inform the policy debates: firstly about how bursaries impact on students,; and secondly about the importance of background and personality characteristics in influencing students' intentions to teach and the implications for marketing the profession to potential applicants. On the latter, there is much more which could be said about market segmentation and the marketing approaches which could be used to attract different groups, but this goes beyond the remit of this particular study and its data. This research offers some useful insight.

1. In our experiment, offering higher endowments for priority subjects and degree classes did not impact negatively on the effort of other students. This would suggest that offering larger incentives for the most able students in high priority subjects and lower incentives for others would not affect applications for teaching in other, lower priority subjects. Unlike real bursaries, however, the endowments were small-scale and we cannot rule out the possibility that the differences were not large enough to find an effect, so this needs further testing. It is important to use the natural environment as a test bed for the impact of different types of bursaries – for instance by conducting natural field experiments where some people are randomly incentivised to go into teaching

2. In our experiment, the ways the incentives are framed do not have a significant impact on effort. But again more research is needed to test whether this finding is unique to the parameters used in this study.

3. There is a clear gender effect, especially with regard to primary school, with women much more likely than men to express an intention to teach at primary level. This is a feature common to many countries, and there is no evidence to suggest that academic outcomes are adversely affected by it. There are other reasons, such as the provision of positive male role models for boys, why it may be desirable to have a more balanced teacher workforce. DfE has a programme to increase the number of male teachers in primary schools, and there is a body of literature on the reasons men do not become primary school teachers which is beyond the scope of this study to report, but there is still further investigation to be done in understanding how to get men into teaching, the existing barriers which need to be overcome and what incentives might be most effective.

4. Again confirming the previous literature, our experiment shows that those with lower problem solving skills and 'A' level grades are more likely to express an intention become a teacher. As with gender, this finding is significant for primary teaching intentions only. While high grades do not guarantee teaching ability (or lower grades preclude it), such a finding is nonetheless bound to generate concern that teaching may be losing out to other professions when it comes to attracting high-achievers. While the Schools White Paper (DfE, 2010) notes that the average degree class of those entering postgraduate initial teacher training has moved from below to above average in recent years it also notes that the status of the teaching profession has some way to go to catch up with higher-performing countries and there is still a case for DfE to consider targeted approaches to students with higher grades.

5. Those in the high priority subjects are more likely to be male, care more about the future, feel that they have a higher degree of autonomy, have higher problem-solving ability, have felt happy yesterday (at the time of answering the question), and report a weaker feeling that the things in their life are worthwhile. All of these factors should be taken into consideration when designing the marketing of ITT to attract students in the high priority areas.

6. It seems that those who are more trusting of other people are more likely to become teachers. So marketing teaching as a trusting and trustworthy profession might elicit more applications into ITT.

7. People who feel competent and accomplished as an individual are more likely to become teachers. So stressing the opportunity to develop one's own personal competencies during teaching (i.e. to obtain skills that are important for all jobs – not only teaching) could be important for social marketing of the teaching profession. Attracting teaching entrants in this way should not increase turnover as we know from the previous literature that they will have inertia once they join teaching and will value intrinsic benefits of teaching over time and will not move.

References

Abeler, Johannes, Armin Falk, Lorenz Goette and David Huffman (2010). Reference Points and Effort Provision. *American Economic Review*, 101, 470-492.

Ballou, D. (1996). Do public schools hire the best applicants? *Quarterly Journal of Economics, 111*(1), 97–133.

Broughman, S., & Rollefson, M. (2000). Teacher supply in the United States: Sources of newly hired teachers in public and private schools: 1987–88 to 1993–94. *Education Statistics Quarterly, 2*(3), 28-32.

Burrows, P & Loomes, G. (1994). The impact of fairness on bargaining behaviour. *Empirical Economics*, 19, 201–221.

Cadsby, C. Bram, Fei Song and Francis Tapon (2009). The Impact of Risk Aversion and Stress on the Incentive Effect of Performance Pay. unpublished paper, University of Guelph, 2009.

Chevalier, A., Dolton, P., McIntosh, S. (2006). Recruiting and retaining teachers in the UK: An analysis of graduate occupation choice from the 1960s to the 1990s. Economica.

Department for Education (DfE), (2011). *Training our next generation of outstanding teachers: Implementation plan*. London.

Dickinson, David L. (1999). An Experimental Examination of Labor Supply and Work Intensities. *Journal of Labor Economics,* 17: 638-670.

Dolan, Paul, Richard Layard, and Robert Metcalfe (2011) "Measuring Subjective Well-Being for Public Policy" UK Office for National Statistics, available at http://www.statistics.gov.uk/articles/social trends/ measuring-subjective-wellbeing-for-public-policy.pdf

Dolan, Paul and Robert Metcalfe (2012). Measuring Subjective Wellbeing : Recommendations on Measures for use by National Governments. *Journal of Social Policy*, 41, 409-427.

Dolton, P. (1990). The Economics of UK Teacher Supply: The Graduate's Decision. *Economic Journal*, 100, 91-104.

Dolton, P. (2006). Teacher Supply. In Handbook of the Economics of Education, chapter 19.

Dolton, P.J., Chung, T.P. (2004). The rate of return to teaching. National Institute Economic Review 190, 89–104

Evans, R.H. (1987). Factors Which Deter Potential Science/Math Teachers from Teaching: Changes Necessary To Ameliorate Their Concerns. *Journal of Research in Science Teaching, 24*, 77-85.

Fahr, R. and Irlenbusch, B. (2000). Fairness as a constraint on trust and reciprocity: earned property rights in a reciprocal exchange experiment. *Economic Letters*, 66, 275– 282.

Farkas, S., Johnson, J. & Foleno, T. (2000). *A sense of calling: Who teaches and why.* New York: Public Agenda.

Ferguson, R.F. (1991). Paying for Public Education: New Evidence on How and Why Money Matters. *Harvard Journal on Legislation, 28,* 465-498.

Flyer, F., & Rosen, S. (1997). The new economics of teachers and education. *Journal of Labor Economics, 15*(1), S104–139.

Gill, David and Victoria L. Prowse. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *American Economic Review*, 102: 469-503

Gitomer, D., Latham, A., & Ziomek, R. (1999). *The academic quality of prospective teachers: The*

impact of admissions and licensure testing. Princeton, NJ: Educational Testing Service.

Glaeser, E. G., Laibson, D., Scheinkman, J. A., Soutter, C. L. (2000). Measuring Trust. *Quarterly Journal of Economics*, 65, 811-846.

Gneezy, Uri, Ernan Haruvy, and Alvin E. Roth (2003). Bargaining Under a Deadline: Evidence from the Reverse Ultimatum Game. *Games and Economic Behavior*, 45: 347-368.

Hanushek, E.A. and Pace, R.R. (1995). Who Chooses To Teach (and Why)? *Economics of Education Review, 14*, 101-117.

Henke, R., Chen, X., Geis, S., & Knepper, P. (2000). *Progress through the teacher pipeline: 1992–93 college graduates and elementary/secondary teaching as of 1997.* Washington, DC: National Center for Education Statistics.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. L. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7, 346–380.

Hounshell, P.B., and Griffin, S.S. (1989). "Science Teachers Who Left: A Survey Report." *Science Education, 73*, 433-443

Konow, J. (2000). Fair shares: accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90, 1072–1091.

Manski, C.F. (1987). Academic Ability, Earnings, and the Decision To Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972. In D.A. Wise (Ed.), *Public Sector Payrolls.* Chicago, IL: University of Chicago Press.

McCrae, R.R.; Costa, P.T.; Jr (1987). "Validation of the five-factor model of personality across instruments and observers". *Journal of Personality and Social Psychology* **52** (1): 81–90

Podgursky, M., Monroe, R., & Watson, D. (2004). The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review, 23,* 507–518.

Reed, D.F., and Busby, D.W. (1985). Teacher Incentives in Rural Schools. *Research in Rural Education, 3*, 69-73.

Rivkin, S.G., Hanushek, E.A. and Kain, J.F. (1998). *Teachers, Schools and Academic Achievement*. (Working Paper 6691). Cambridge, MA: National Bureau of Economic Research.

Rumberger, R.W. (1987). The Impact of Salary Differentials on Teacher Shortages and Turnover: The Case of Mathematics and Science Teachers. *Economics of Education Review, 6*, 389-399.

Ryff, C. D. (1989). *Journal of Perspectives on Social Psychology*, 57**,** 1069-1081.

Sanders, W. and Rivers, J. (1996). *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. Knoxville, TN: University of Tennessee, Value-Added Research and Assessment Center.

Stinebrickner, T.R. (2002). An Analysis of Occupational Change and Departure from the Labor Force. Evidence of the Reasons that Teachers Leave. *Journal of Human Resources, 37*, 192-216.

Sutter, M. and H. Weck-Hannemann (2003). Taxation and the Veil of Ignorance - a Real Effort Experiment of the Laffer Curve, *Public Choice,* 115.

U.S. Department of Education, National Center for Education Statistics. (2002). *The condition of education, 2002* (NCES 2002-025). Washington, DC: Author.

Van Dijk , Frans, Joep Sonnemans and Frans Van Winden (2001). Incentive Systems in a Real Effort Experiment. *European Economic Review* 45: 187-214.

Appendix A: The slider task

The slider task does not require any prior knowledge, performance is easily measurable, and there is little learning possibility. At the same time, the task is boring and pointless and we can thus be confident that the task entails a positive cost of effort for subjects. The task is also clearly artificial, and the output is of no real value to the experimenter. This eliminates any tendency for subjects to use effort in the experiment as a way to reciprocate for any incentives offered by the experimenter.

The screen will not vary across experimental subjects or across repetitions of the task. A schematic representation of a single slider is shown in the figure below. When the screen containing the effort task is first displayed to the subject all of the sliders are positioned at 0, as shown for a single slider in the Figure below.

(a) Initial position. (b) Positioned at 50.

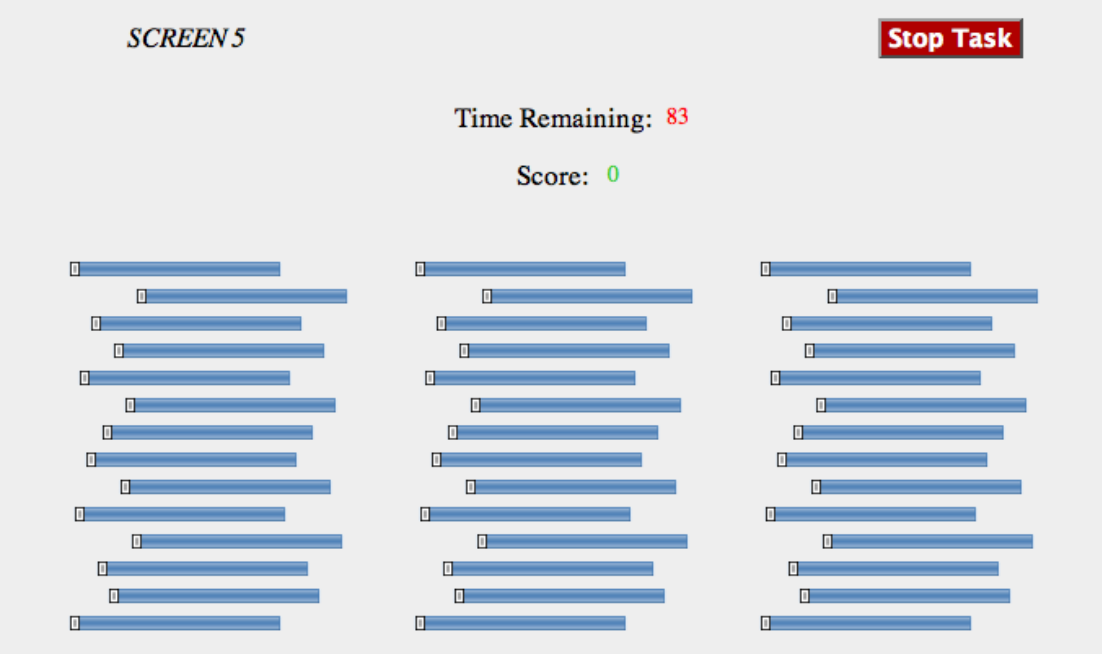
By using the mouse, the subject can position each slider at any integer location between 0 and 100 inclusive. Each slider can be adjusted and readjusted an unlimited number of times and the current position of each slider is displayed to the right of the slider.

The subject's “points score" in the task, interpreted as effort exerted, is the number of sliders positioned at 50 at the end of the allotted time. As the task proceeds, the screen displays the subject's current points score and the amount of time remaining. An incorrect position of the slider incurs a penalty, whether the slider is at 0 or 49 or 51.

To ensure that all the sliders are equally difficult to position correctly, the sliders are arranged on the screen such that no two sliders are aligned exactly one under the other. This prevents the subject being able to position the higher slider at 50 and then easily position the lower slider by copying the position of the higher slider.

For our experiment, each student will have two minutes to place the slider at the half way point in as many lines as possible per page. The task will last for 40 minutes (two minutes per page), and they will be shown 20 different screens/pages. They will have the opportunity to quit at any point during the task, but only after screen 4. Each person is incentivised by being rewarded two pence per correct slider.

The actual slider task looked like the figure below. It is clear that there is a stop button on the screen, which meant that participants could stop the slider task as soon as they wanted after the fourth screen.


SCREEN 5
Stop Task
Time Remaining: 83
Score: 0

Appendix B: The endowment frame

*You set a total of **5** sliders at 50 in the 2 practice screens. So, **5** would have been your total score if this would have been for real.*

Your payment:

You are about to start the slider task for real, and your performance in it will influence your final earnings in the experiment. So, first we need to clarify how your payment will be calculated. Your final earnings will depend on 2 separate elements:

1. A fixed sum given to you based on the subject you are studying and the marks you obtained.
2. An additional amount based on your performance in the slider task.

We now explain briefly these two elements in turn.

The fixed sum

Depending on the priority of the subject you are studying (according to the Department for Education) and the marks you obtained, you are assigned a fixed sum in the experiment. Your payment in the experiment will be this fixed sum plus the amount you earn in the slider task (explained in the next section). The table below shows the sums corresponding to the priority of the subject and the marks. The sum assigned to you is highlighted in red.

	High priority	*Medium priority*	*Low priority*
1st	£14	**£4.50**	£0
2:1	£8.50	£1	£0
2:2	£6	£0	£0
3rd	£0	£0	£0

Payment of the slider task

You will now start the real slider task. It consists of 20 consecutive screens with 42 sliders each, and you are given 2 minutes (120 seconds) per page, just like in the practice screens. You have to set as many sliders as you can at 50 (the middle position of the sliders). For every slider that you set at 50 you will **obtain** 2 pence (£0.02). So, your maximum earnings in this task are £16.80. Your payment in the experiment will be the amount that you earn in this task plus your fixed sum (explained in the previous section).

IMPORTANT NOTES:

- After the fourth screen, you will have a **"Stop Task"** button on every screen. If you click that button you will be asked whether you want to skip the rest of the slider screens, and if you confirm you will skip the rest of the screens and jump to the final part of the experiment. You can do that whenever you want, but remember that you will lose the possibility to increase your earnings with the rest of the sliders.
- You will see your current total earnings in the slider task at all times on the upper right part of the screens.

Appendix C: The invitation to University administrators

Dear ..

Apologies for the unsolicited email. I am a researcher at the University of Oxford working with a group of academics interested in the future career paths of second- and third-year students in xxx at the University of xxx.

We are running an experiment to find out what students intend to do in the future. If they participate in the experiment, they can earn up to £30 – the Department for Education has funded this exciting project.

Please may you pass the below script to ONLY second- and third-year students in your department?

Your help would be most appreciated.

Yours sincerely,

Dear student

A group of researchers are currently running an experiment to understand the career choices of second- and third-year xxx undergraduates at the University of xxx.

We would like to invite you to take part in this exciting experiment. It will take no more than 50 minutes to complete. You will receive a maximum of £30 for participation in the experiment.

The link to the experiment is www… You only have one week to take part in the experiment, and we advise you completing the experiment using a mouse with your computer.

It is important not pass on the link to other people – you have been specifically invited to take part in this experiment.

We would really appreciate your participation in our experiment.

Yours faithfully,

Appendix D: Descriptive statistics

Table D1: Degree specialism

Subject	%	% (actual 2009-10 ITT cohort)
Applied science	1.08%	0.25%
Art and design	1.84%	1.85%
Biology	9.59%	3.04%
Business studies	5.72%	1.55%
Chemistry	2.22%	2.20%
Citizenship	0.25%	0.72%
Classics	0.95%	0.10%
Computer science	1.02%	n/a%
Design and tech.	0.32%	2.90%
Drama	2.73%	1.00%
Economics	11.56%	0.03%
Engineering	4.00%	0.04%
English	8.96%	6.48%
General science	1.52%	n/a
Geography	1.65%	2.08%
Health and soc. Care	2.48%	0.16%
History	6.73%	2.04%
Info. and com. tech.	0.64%	2.72%
Mathematics	3.43%	7.50%
Media studies	1.27%	0.11%
Modern languages	0.64%	3.20%
Music	1.40%	2.09%
Physical education	0.13%	4.31%
Physics	4.13%	1.29%
Primary	3.49%	39.22%
Psychology	0.89%	0.29%
Religious education	0.83%	2.36%
Social sciences	20.52%	0.31%

Table D2: Universities represented

University	%
Birkbeck College	0.51%
Brunel University	0.95%
Kingston University	4.57%
Lancaster University	1.27%
London South Bank University	1.52%
LSE	8.70%
Oxford Brookes University	0.44%
UCL	13.02%
University of Birmingham	8.96%

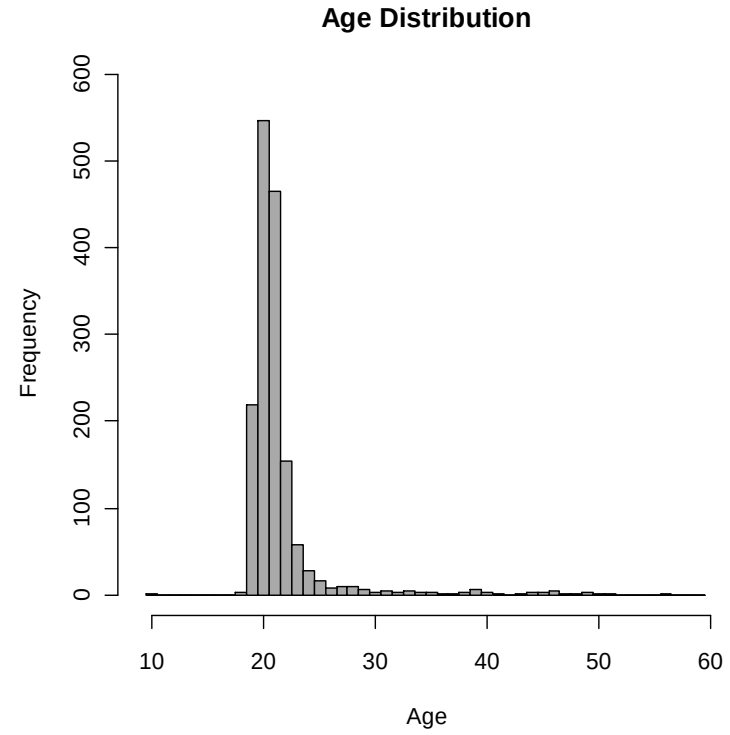
University of Buckingham	0.25%
University of Cambridge	1.21%
University of Kent	26.75%
University of Manchester	3.62%
University of Newcastle	11.18%
University of Reading	7.69%
University of Sheffield	7.31%
Other	2.03%

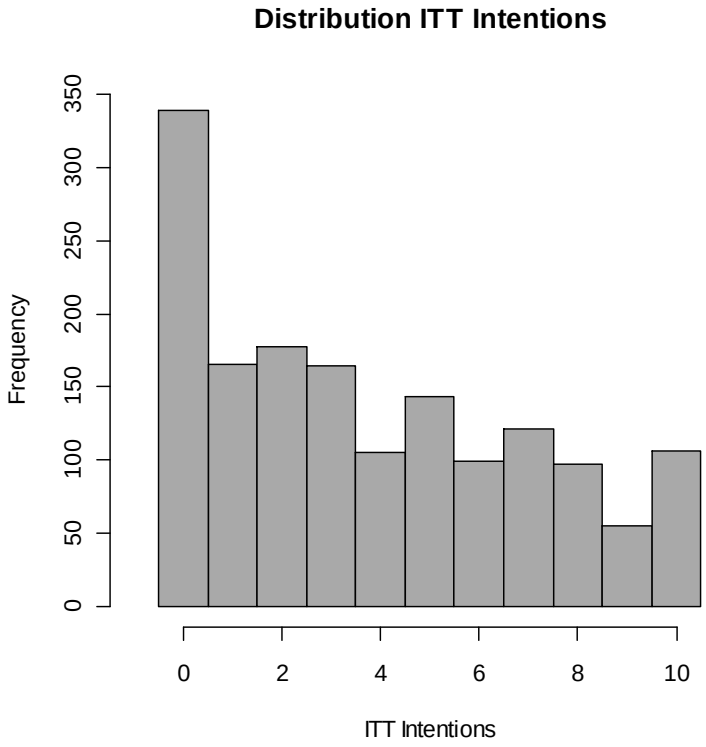
Table D3: Career intentions

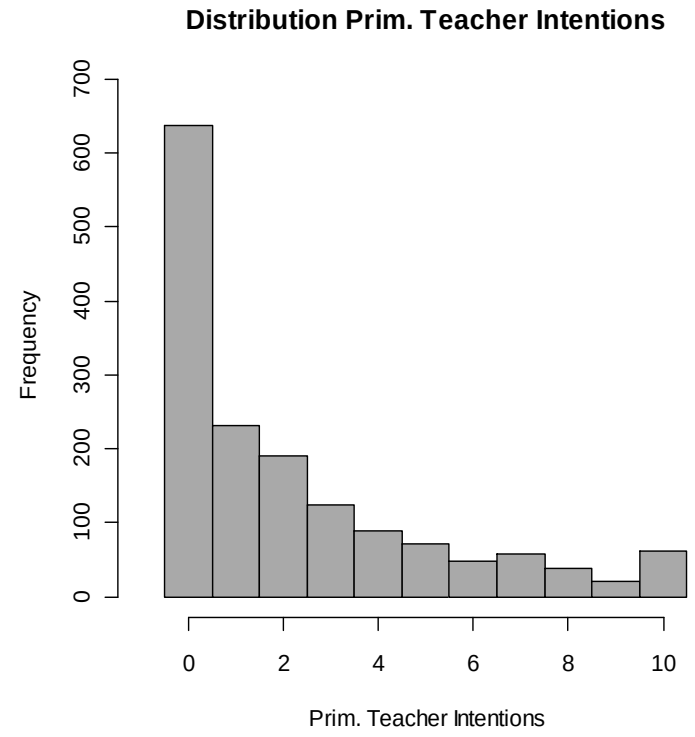
Career	**%**
Accountant	7.05%
Actuary	0.44%
Advocate	0.38%
Archaeologist	1.14%
Architect	0.19%
Chef	0.38%
Dentist	0.13%
Engineer	5.08%
Financial analyst	8.70%
Interpreter	0.44%
Journalist	5.40%
Lawyer	7.62%
Nurse	0.32%
Pharmacist	0.19%
Philosopher	0.76%
Physician	0.83%
Pilot	0.19%
Professor	2.73%
Psychologist	0.51%
Quantity surveyor	0.06%
Scientist	9.47%
Social worker	3.81%
Surgeon	0.57%
Teacher	14.93%
Vet	0.38%
Other	28.27%

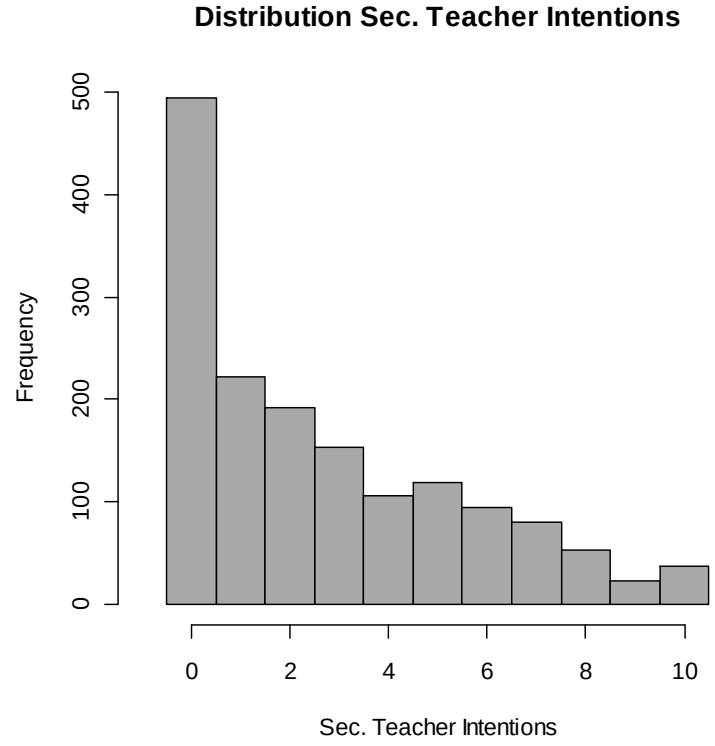
Table D4: Preferred ITT subject

ITT subject	%
Art	2.41%
Biology	7.81%
Business education	3.30%
Chemistry	2.54%
Classics	0.83%
Computer studies	1.59%
Design and tech.	0.70%
Drama/dance	3.11%
Earth sci./geology	0.13%
Economics	7.81%
Engineering	1.65%
English	12.90%
Environ. Science	0.57%
French	1.33%
Geography	2.03%
German	0.64%
Health and soc. Care	1.40%
History	7.43%
Home economics	0.38%
Humanities	2.80%
Italian	0.19%
Japanese	0.32%
Leisure and tourism	0.44%
Lower primary	1.97%
Mathematics	9.09%
Music	1.65%
Physical education	1.08%
Physics	3.56%
Primary	6.67%
Psychology	1.14%
Religious education	1.52%
Russian	0.13%
Science	3.11%
Spanish	0.70%
Upper primary	0.89%
Other	6.16%


Age Distribution
Frequency
0
100
200
300
400
500
600
10
20
30
40
50
60
Age


Distribution ITT Intentions
Frequency
0
50
100
150
200
250
300
350
0
2
4
6
8
10
ITT Intentions


Distribution Prim. Teacher Intentions
Frequency
0
100
200
300
400
500
600
700
0
2
4
6
8
10
Prim. Teacher Intentions


Distribution Sec. Teacher Intentions
Frequency
0
100
200
300
400
500
0
2
4
6
8
10
Sec. Teacher Intentions

Number of participants by ITT intention and priority of degree subject

ITT intention	High priority	Medium priority	Low priority	Totals
0	32	185	122	339
1	23	93	49	165
2	22	98	55	175
3	15	93	56	164
4	16	53	34	103
5	15	83	43	141
6	11	59	29	99
7	18	70	33	121
8	11	58	28	97
9	5	30	19	54
10	8	66	32	106
Totals	176	888	500	

Number of participants by ITT intention and marks obtained

ITT intention	1st	2:1	2:2	3rd	Totals
0	98	176	60	5	339
1	52	84	24	5	165
2	37	96	42	2	177
3	41	83	36	4	164
4	30	49	22	4	105
5	38	76	27	2	143
6	22	60	17	0	99
7	28	69	22	2	121
8	21	56	19	1	97
9	13	33	8	1	55
10	17	70	17	2	106
Totals	397	852	294	28	

Appendix E: Regression tables

Table E1: Teaching and ITT intentions regressions

	(I)	(II)	(III)	(IV)
	Intention Teacher (Binary, Logistic)	Intention Prim. Teacher (OLS)	Intention Sec. Teacher (OLS)	Intention ITT (OLS)
Constant		0.101 (0.998)	3.210 (0.998)**	-0.136 (1.092)
Gender	0.080 (0.019)**	0.982 (0.160)**	0.07 (0.167)	0.461 (0.192)*
Age	0.001 (0.003)	0.026 (0.028)	-0.083 (0.024)**	0.034 (0.026)
Marital status	0.043 (0.030)	0.285 (0.279)	0.123 (0.264)	0.246 (0.293)
Ethnicity	0.051 (0.018)**	0.081 (0.180)	0.024 (0.179)	0.103 (0.205)
Conservative	-0.045 (0.019)*	-0.389 (0.187)*	-0.092 (0.194)	-0.485 (0.215)*
Labour	-0.003 (0.018)	-0.242 (0.183)	0.373 (0.178)*	0.189 (0.202)
Study year	-0.008 (0.017)	-0.101 (0.156)	-0.172 (0.151)	-0.573 (0.172)**
Risk preference	-0.005 (0.004)	0.052 (0.034)	0.001 (0.036)	0.027 (0.037)
Trust	0.009 (0.005)#	0.093 (0.044)*	0.069 (0.043)	0.154 (0.050)**
Time preference	-0.008 (0.005)	-0.033 (0.047)	-0.056 (0.048)	-0.023 (0.051)
Extraversion	-0.012 (0.008)	-0.044 (0.068)	-0.022 (0.067)	-0.029 (0.075)
Neuroticism	-0.013 (0.006)*	-0.044 (0.059)	-0.019 (0.058)	0.018 (0.067)
Autonomy	-0.007 (0.007)	-0.236 (0.072)**	-0.137 (0.071)#	-0.091 (0.078)
Competence	0.017 (0.008)*	0.238 (0.073)**	0.149 (0.079)#	0.233 (0.085)**
Relatedness	-0.005 (0.010)	0.022 (0.078)	-0.150 (0.083)#	-0.103 (0.091)
GMAT score	-0.02 (0.010)*	-0.165 (0.075)*	-0.084 (0.077)	-0.231 (0.086)**
Charity preference	-0.016 (0.017)	0.278 (0.159)#	-0.048 (0.156)	0.128 (0.178)
Life satisfaction	-0.003 (0.007)	0.014 (0.065)	-0.038 (0.066)	-0.004 (0.07)
Happy yesterday	-0.001 (0.005)	0.013 (0.048)	0.079 (0.051)	0.092 (0.056)
Anxious yesterday	0.000 (0.004)	0.011 (0.035)	0.005 (0.035)	0.031 (0.040)
Worthwhile	-0.005 (0.006)	-0.035 (0.049)	-0.021 (0.049)	-0.031 (0.057)
Imp. bursary	0.019 (0.003)**	0.261 (0.023)**	0.376 (0.023)**	0.433 (0.027)**
Mark	0.003 (0.013)	0.036 (0.105)	-0.037 (0.108)	0.060 (0.122)
A at A-level	-0.011 (0.006)#	-0.132 (0.05)**	0.014 (0.052)	-0.007 (0.058)
Effort (total score)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
	Pseudo-R^2: 0.175	R^2: 0.227	R^2: 0.262	R^2: 0.307

Notes:
1) The main numbers in the table are regression coefficients, with the standard deviations in brackets (robust standard errors have been used in regressions II, III and IV). The results of the first regression (logistic) are in terms of marginal effects.
2) "**" means p-value ≤ 0.01, "*" p-value ≤ 0.05, "#" p-value ≤ 0.10.
3) The variables are as follows. "Gender": 0 for male, 1 for female; "age": age of the participants; "marital status": 0 for single, 1 for not single; "ethnicity": 0 for non-white, 1 for white; "conservative": 1 for intention to vote conservative, 0 otherwise; "labour": 1 for intention to vote labour, 0 otherwise; "study year": 0 for second year, 1 for third year (other cases excluded); "risk preference": degree of risk aversion, with higher values for less risk aversion; "trust": how much participants think that others can be trusted (on scale from 0 to 10); "time preference": degree of temporal discounting, with higher values for more temporal discounting; "extraversion": degree of extraversion (on scale from 0 to 6); "neuroticism": degree of neuroticism (on scale from 0 to 6); "autonomy": response to "I am free to decide for myself how to live my life" (on scale from 0 to 6); "competence": response to "most days I feel a sense of accomplishment from what I do" (on scale from 0 to 6); "relatedness": response to "people in my life care about me" (on scale from 0 to 6); "GMAT score": score obtained across four problem-solving questions, with higher scores for higher problem-solving ability; "charity preference": 0 for not contributing to charity, 1 for contributing; "life satisfaction": how satisfied people are with their lives (on scale from 0 to 10); "happy yesterday": how happy people felt yesterday (on scale from 0 to 10); "anxious yesterday": how anxious people felt yesterday (on scale from 0 to 10); "worthwhile": captures whether people think that things in their life are worthwhile (on scale from 0 to 10); "imp. bursary": importance given to bursaries to do ITT (on scale from 0 to 10); "mark": mark obtained in the latest university year completed; "A at A-level": number of A marks at A-level; "effort (total score)": captures the effort made in the effort task (slider task), in terms of total number of slider done.

Table E2: Effort regressions

	(I)	(II)	(III)
	Total Sliders (OLS)	Number of Screens (OLS)	Sliders Practice Minus Sliders Task (OLS)
Constant	536.939 (91.758)**	14.579 (2.451)**	6.329 (3.449)#
Gender	-78.321 (15.561)**	-1.051 (0.380)**	-1.199 (0.579)*
Age	-3.592 (2.202)	0.039 (0.065)	0.004 (0.065)
Marital status	-1.343 (21.196)	-0.052 (0.551)	-0.390 (0.758)
Ethnicity	4.890 (17.044)	0.253 (0.424)	-2.048 (0.730)**
Conservative	-27.974 (17.899)	-0.566 (0.456)	0.192 (0.740)
Labour	-6.401 (15.581)	0.079 (0.390)	-0.889 (0.568)
Study year	38.015 (13.757)**	0.481 (0.348)	-0.048 (0.504)
Risk preference	-2.446 (3.438)	-0.064 (0.081)	-0.087 (0.129)
Trust	9.856 (3.928)*	0.223 (0.099)*	-0.357 (0.177)*
Time preference	-9.724 (4.19)*	-0.307 (0.107)**	0.061 (0.156)
Extraversion	-11.539 (6.333)#	-0.253 (0.158)	-0.003 (0.269)
Neuroticism	-0.393 (5.125)	0.032 (0.128)	0.031 (0.183)
Autonomy	0.057 (5.891)	-0.005 (0.151)	0.068 (0.226)
Competence	9.203 (6.976)	0.230 (0.176)	0.431 (0.266)
Relatedness	26.818 (7.878)**	0.629 (0.198)**	-0.023 (0.303)
GMAT score	12.137 (7.132)#	0.071 (0.177)	-0.480 (0.299)
Charity preference	-76.097 (14.133)**	-2.226 (0.358)**	-0.392 (0.506)
Life satisfaction	-3.017 (5.715)	-0.130 (0.146)	0.381 (0.239)
Happy yesterday	-0.694 (4.502)	-0.056 (0.114)	-0.226 (0.164)
Anxious yesterday	-3.630 (2.992)	-0.130 (0.074)#	0.005 (0.123)
Worthwhile	-0.962 (4.745)	0.041 (0.119)	-0.184 (0.177)
Imp. Bursary	3.049 (2.297)	0.056 (0.059)	0.053 (0.086)
Mark	-1.692 (12.171)	0.074 (0.312)	0.413 (0.500)
A at A-level	3.977 (4.635)	0.004 (0.117)	0.231 (0.171)
Endowment	-5.508 (6.886)	-0.128 (0.166)	0.390 (0.289)
ITT intentions	-2.500 (2.947)	-0.047 (0.076)	0.035 (0.106)
Gain/loss framing	-3.961 (13.534)	-0.168 (0.340)	-0.336 (0.499)
Low/high endow.	-8.481 (13.666)	-0.348 (0.345)	0.025 (0.504)
Social information	8.687 (8.243)	0.245 (0.207)	0.270 (0.316)
Mark*Endow.	4.681 (3.438)	0.078 (0.082)	-0.207 (0.138)
Endow.*ITT	0.021 (0.788)	0.000 (0.020)	0.017 (0.036)
	R^2: 0.113	R^2: 0.085	R^2: 0.044

Notes:
1) The main numbers in the table are regression coefficients, with the standard errors in brackets (robust standard errors have been used).
2) "**" means p-value ≤ 0.01, "*" p-value ≤ 0.05, "#" p-value ≤ 0.10.
3) The variables are as follows. "Gender": 0 for male, 1 for female; "age": age of the participants; "marital status": 0 for single, 1 for not single; "ethnicity": 0 for non-white, 1 for white; "conservative": 1 for intention to vote conservative, 0 otherwise; "labour": 1 for intention to vote labour, 0 otherwise; "study year": 0 for second year, 1 for third year (other cases excluded); "risk preference": degree of risk aversion, with higher values for less risk aversion; "trust": how much participants think that others can be trusted (on scale from 0 to 10); "time preference": degree of temporal discounting, with higher values for more temporal discounting; "extraversion": degree of extraversion (on scale from 0 to 6); "neuroticism": degree of neuroticism (on scale from 0 to 6); "autonomy": response to "I am free to decide for myself how to live my life" (on scale from 0 to 6); "competence": response to "most days I feel a sense of accomplishment from what I do" (on scale from 0 to 6); "relatedness": response to "people in my life care about me" (on scale from 0 to 6); "GMAT score": score obtained across four problem-solving questions, with higher scores for higher problem-solving ability; "charity preference": 0 for not contributing to charity, 1 for contributing; "life satisfaction": how satisfied people are with their lives (on scale from 0 to 10); "happy yesterday": how happy people felt yesterday (on scale from 0 to 10); "anxious yesterday": how anxious people felt yesterday (on scale from 0 to 10); "worthwhile": captures whether people think that things in their life are worthwhile (on scale from 0 to 10); "imp. bursary": importance given to bursaries to do ITT (on scale from 0 to 10); "mark": mark obtained in the latest university year completed; "A at A-level": number of A marks at A-level; "endowment": monetary endowment assigned in the experiment, before doing the effort task; "ITT intentions": intention to do ITT (on scale from 0 to 10); "gain/loss framing": 0 for gain framing, 1 for loss framing; "low/high endow.": 0

for endowment structure A, 1 for endowment structure B; “social information”: 0 for no information, 1 for example anchor, 2 for peer anchor.

Table E3: Priority and Ability (marks obtained) regressions

	(I)	(II)
	Priority of Subject (Ordered Logistic)	Marks Obtained (Ordered Logistic)
Gender	-0.276 (0.118)*	0.116 (0.111)
Ethnicity	0.065 (0.126)	-0.176 (0.118)
Conservative	-0.260 (0.139)#	0.070 (0.130)
Labour	-0.028 (0.120)	-0.027 (0.111)
Risk preference	-0.013 (0.024)	0.026 (0.022)
Trust	0.039 (0.030)	-0.023 (0.028)
Time preference	-0.103 (0.031)**	0.007 (0.029)
Extraversion	-0.059 (0.047)	0.222 (0.044)**
Neuroticism	-0.067 (0.038)#	-0.024 (0.036)
Autonomy	0.103 (0.048)*	-0.016 (0.045)
Competence	0.054 (0.051)	-0.291 (0.049)**
Relatedness	0.091 (0.057)	-0.007 (0.054)
GMAT score	0.335 (0.055)**	-0.264 (0.051)**
Charity preference	-0.015 (0.107)	0.209 (0.100)*
Life satisfaction	0.003 (0.047)	-0.053 (0.045)
Happy yesterday	0.069 (0.034)*	0.008 (0.032)
Anxious yesterday	0.043 (0.023)#	-0.021 (0.021)
Worthwhile	-0.074 (0.037)*	0.036 (0.034)
	Pseudo-R^2: 0.042	Pseudo-R^2: 0.032

Notes:
1) The main numbers in the table are regression coefficients, with the standard errors in brackets.
2) “**” means p-value ≤ 0.01, “*” p-value ≤ 0.05, “#” p-value ≤ 0.10.
3) The variables are as follows. “Gender”: 0 for male, 1 for female; “ethnicity”: 0 for non-white, 1 for white; “conservative”: 1 for intention to vote conservative, 0 otherwise; “labour”: 1 for intention to vote labour, 0 otherwise; “risk preference”: degree of risk aversion, with higher values for less risk aversion; “trust”: how much participants think that others can be trusted (on scale from 0 to 10); “time preference”: degree of temporal discounting, with higher values for more temporal discounting; “extraversion”: degree of extraversion (on scale from 0 to 6); “neuroticism”: degree of neuroticism (on scale from 0 to 6); “autonomy”: response to “I am free to decide for myself how to live my life” (on scale from 0 to 6); “competence”: response to “most days I feel a sense of accomplishment from what I do” (on scale from 0 to 6); “relatedness”: response to “people in my life care about me” (on scale from 0 to 6); “GMAT score”: score obtained across four problem-solving questions, with higher scores for higher problem-solving ability; “charity preference”: 0 for not contributing to charity, 1 for contributing; “life satisfaction”: how satisfied people are with their lives (on scale from 0 to 10); “happy yesterday”: how happy people felt yesterday (on scale from 0 to 10); “anxious yesterday”: how anxious people felt yesterday (on scale from 0 to 10); “worthwhile”: captures whether people think that things in their life are worthwhile (on scale from 0 to 10).

Appendix F: Effects of different conditions

Table F1: Comparing experimental conditions: effort variables

	Total number of sliders done	**Number of screens done**	**Difference in sliders per screen in task minus practice**
High vs low endowment condition	0.49	0.40	0.57
High vs low endowment condition, only 1st and high priority students	0.25	0.71	0.80
High vs low endowment condition, only 2:2 and high priority	0.36	0.15	0.43
High vs low endowment condition, only 1st and medium priority	0.21	0.79	0.85
No information vs information on others' behaviour	0.88		0.54
No information vs giving an example	0.49		0.15
Gain framing vs loss framing	0.09		0.15

Note: *p*-value of Kruskal-Wallis test comparing all twelve conditions in terms of total sliders: 0.82.

Table F2: Comparing experimental conditions: ITT variables

	Intention to do ITT	**Intention to become a primary school teacher**	**Intention to become a secondary school teacher**
High vs low endowment condition	0.52	0.85	0.93
High vs low endowment condition, only 1st and high priority students	0.21	0.57	0.73
High vs low endowment condition, only 2:2 and high priority	0.97	0.18	0.81
High vs low endowment condition, only 1st and medium priority	0.78	0.98	0.60

Ref: DFE-RR251

ISBN: 978-1-78105-182-5

October 2012