Teacher influence on student performance and selection in broad-spectrum tertiary education

Gigi Foster[*], University of South Australia

January 31, 2007

Working Paper

Abstract

Tertiary education is now accessible even to those who appear unlikely *ex ante* to succeed in jobs requiring post-high school education. Institutions that have broadened access to their programs must rely on two things to protect the quality of the degrees they award: selection mechanisms operating during students' tenure, and effective teaching. This paper explores the relative strength of these two forces in a broad-spectrum, first-year undergraduate setting. Using detailed data from the University of South Australia on student background, tutors, performance, and enrollment across 15 weeks in a first-year core course, I aim to illuminate the extent to which teachers impact upon the success of their classes directly (through effective teaching) and indirectly (through facilitating the dropping out of more poorly-prepared students). Results indicate that teachers vary widely in their influence on the success of poorly-prepared students taught face-to-face, and that only one is effective in facilitating the disproportionate out-selection of students with poor initial preparation. Performance at neither of these tasks seems strongly related to teachers' formal university affiliation.

[*]I thank UniSA's Division of Business and Centre for Regulation and Market Analysis for support, and Alan Muscardin and Matthew Giro for research assistance. All viewpoints and errors contained within are solely my own.

Introduction

Much effort in the educational policy arena over the last twenty to thirty years has sought to enable the broadened access to tertiary education now observed in most developed countries. As individuals with less academic promise have increasingly been viewed as potential recipients of higher degrees, increasing numbers of vocationally-oriented courses, degree programs, and universities have come into being in many countries to cater to their needs. In Australia, this trend has manifested itself in part in the birth and continuing existence of the Australian Technology Network (ATN) group of universities. These five universities (Curtin, UniSA, UTS, QUT, and RMIT), all established as such between 1987 and 1992, aim to carve out a niche in the tertiary marketplace that is less conventionally academic than the comparatively elite "group of eight" (Melbourne, the ANU, Adelaide, Monash, UNSW, Queensland, Sydney, and UWA). While ATN faculty do engage in research—satisfying the primary definitional criterion for the work of "university" faculty—the stated research focus of most ATN institutions is on applied, often multi-disciplinary research with clear industry and/or community links, rather than on basic research or discipline-based theory development.[1] The marketing campaigns of these new universities underscore their devotion to a similar role in the teaching arena, including their active promotion of broad access to a wide range of both traditional and non-traditional degree programs.[2]

This trend did not halt in the 1990s. Page 8 of Evans (1999) describes the then-current increase within the tertiary education system more broadly in alternative pathways to tertiary education, the broadening of credit transfer arrangements, and an emphasis on "lifelong learning" rather than primary degree completion. In addition to welcoming school-leavers, mature age students, transfer students from technical colleges (TAFE) and international students, UniSA has recently advertised yet another pathway to university:

[1]For example, UniSA's research strategy is expressed thus on its website, http://www.unisa.edu.au/community/: "The success of our research is in its real-world application." QUT's slogan is "A university for the real world."

[2]Curtin's website, http://about.curtin.edu.au/index.cfm, states that it offers "over 850 undergraduate and postgraduate courses in business, engineering, health sciences, humanities, science, mining and agriculture."

UniSA Foundation Studies program, beginning in 2006 and designed for people with no existing qualifications for university entry.[3]

These new universities face a challenge unique to their role in the market. If entry to their programs is easy for poorly-prepared students, but the degree awarded at the conclusion of study is to have enough market value to justify its cost, then one or both of two things must happen to poorly-prepared students during their tenure at the university: either they must leave before they complete their programs, or they must be taught up to a market-supportable standard.[4] If neither phenomenon occurs, then these newer universities will struggle to survive economically in the longer term. In the short- to medium-term, in the absence of value-added teaching or preparation-biased attrition, no counter-pressure is brought against the possibility that the initial sorting of some students into higher education may not be socially efficient. Prolonged mismatches of students to tertiary education are expensive, and can be painful for students and teachers alike.[5] If students below a certain threshold of preparation cannot be taught up to a market-supportable standard within the tertiary institutions available, then only one of two arguments can be made: that their attendance at university is not socially efficient (in the case that the available institutions accurately reflect social goals), or that a change in institutions is required (in the alternative case).

In another subfield of education literature, increased attention has been devoted recently, both in the field and in the academy, to teacher quality as a potential driver of educational parity.[6] A number of recent studies by education scholars and economists have addressed empirically the potential power of teachers to catalyze student retention and/or learning. Authors have examined the extent to which primary and secondary teachers' personal characteristics, test scores and qualifications (Goldhaber 2006, Rivkin,

[3]This new program is advertised at http://www.unisa.edu.au/future/foundation/default.asp.

[4]Note that neither of these phenomena is required for the long-term viability of universities whose entry requirements are sufficiently stringent: their viability would be possible even if the ultimate degree were only a signal of underlying quality.

[5]Recent research using student data from multiple Australian universities (Long, Ferrier & Heagney 2006) has indicated the prevalence of matching problems throughout the tertiary sector.

[6]See Rowe (2003) for an impassioned and thorough review of this issue with reference to the primary and secondary education of boys versus girls in Australia.

Hanushek & Kain 2005), demographic match to their students (Dee 2005), pathway into the profession (Boyd, Grossman, Lankford, Loeb & Wyckoff 2006), and experience (Rivkin et al. 2005) impact on student performance. In many studies, teachers' experience and qualifications have been found to be associated to some degree with student gains (Darling-Hammond 1999), as has the degree of positive demographic match between a student and his or her teacher (Dee 2005). Other studies (Rivkin et al. 2005, Kane, Rockoff & Staiger 2006, Rockoff 2004) find that while substantial heterogeneity in teacher quality exists, this heterogeneity is not well-captured by observable characteristics such as experience and education. In addition, with very few exceptions, teacher quality has not been examined at the tertiary level.[7] Finally, no existing papers known to the author examine the variance of teachers' associations with the success of their poorly-prepared students.

In this paper, these two topics—the market needs of broad-spectrum university education, and teacher influence on student outcomes—are conceptually and empirically linked. I explore in particular the extent to which different teachers are associated with different patterns in the attrition and performance of poorer students, whose trajectories are of particular importance in determining the long-term viability of broad-spectrum higher education. I track the progress of a diverse student body through a large, first-year core course at the University of South Australia, staffed (as are many courses) by a blend of casual academics, graduate students, and permanent staff members. A panel of weekly enrolment snapshots across the semester allows the observation of aggregate attrition patterns, and information on tutors teaching multiple tutorials allows me to ascertain the extent to which individual tutors are associated with both the attrition and the final examination performance of their less well-prepared students. Results demonstrate that teachers do display economically and statistically significantly different profiles with regard to both attrition and performance—implying that universities' selection of tutors based on teaching style may be a potential tool in shaping the survival strategies of players in

[7] One notable exception is Hoffmann & Oreopoulos (2006), who find a small degree of variation in tertiary teachers' value added with regard to student outcomes using a large student-level data set from a Canadian university. However, neither the association of this variance with teacher observables, nor teachers' potential contribution to preparation-biased retention within higher education, are examined.

the mass education marketplace.

Data

Empirical setting

The data for this study are drawn from the 2006 offering of Microeconomics at the University of South Australia (UniSA), a public ATN university with approximately 29,000 students and 2,000 staff, whose main campuses are located in Adelaide, South Australia. All students enrolled internally in programs offered by UniSA's Division of Business are required to take a core of eight business-related courses, one of which is Microeconomics. This course is taken by most students in their first semester, although it can be taken in the second semester or later years. It is a standard first-year microeconomics course whose principal text is an Australian adaptation of the introductory American textbook by Campbell McConnell.

Delivery of this course to internal students consists of a two-hour lecture and a one-hour tutorial each week, over the course of 15 weeks (inclusive of a two-week intervening teaching break). The course concludes with a final examination. Lectures are given by one individual (with the exception of a guest lecture in one week), but each tutorial is taught for the semester's duration by one of a team of people. In 2006, as in most years, this team was a mix of full-time staff members, graduate students, and "casual academics" hired on short-term contracts who had no other status within the university. Some of these individuals had extensive experience delivering and even developing the course, and some had comparatively little. None had a doctorate. The teacher with the highest level of experience with the course under study was officially a casual academic who had previously been employed by the University. This teacher taught two tutorials in 2006 as well as giving the lectures, and was functionally more similar to a permanent staff member than a casual academic (and therefore was coded as such). All but one teacher (Casual Tutor 4) taught more than one tutorial.[8]

[8]The number of tutorials taught by each tutor was as follows: Permanent Staff Tutor 1—2 tutes;

For each week of the semester, students' enrolment status in the course is observed, allowing the construction of both aggregated and week-by-week attrition measures. Other observable student-level information used here is as follows: tutorial assignment (day, time, and tutor); final examination performance; and an array of information regarding prior academic knowledge and background characteristics drawn from a questionnaire and online multiple choice test that students complete for course credit early in the semester.

The multiple choice component of this initial online assessment is based heavily on the Test of Economic Literacy (TEL), used elsewhere to measure basic competence in economic reasoning (Beck & Krumm 1991, McKenna 1994, Whitehead & Halil 1989). Students' results on this test are used here to construct an important control for a pre-existing basket of underlying knowledge and ability that is directly related to likely success in the course. Following dialog amongst education researchers regarding the importance of mastery of threshold concepts for student success (Davies 2006), this basket of knowledge and ability is constructed as the concept-weighted percent correct achieved by students on 19 questions on this test relating to the following fundamental discipline concepts: partial equilibrium; inter-related markets; public goods; opportunity cost; the fallacy of composition; the scientific model; zero sum games; gains from trade; comparative advantage; marginality; the economic system; the production possibility frontier; trade-offs; markets; and market failure. The ensuing analysis will focus heavily on this variable as a measure of students' prior preparation for the study of microeconomics. This measure is not only highly relevant to the specific discipline being taught, but it is in addition not based on socioeconomic status, race, ethnic origin, or other personal background characteristics. It has also been found elsewhere (Shanahan, Foster & Meyer 2006) to be strongly related to student performance in the University of South Australia's introductory microeconomics course, even controlling for a host of other student-specific and institutional factors. Students' success on this constructed measure may also be thought of as a dimension along which teachers of economics might be expected to exert productive change: course content

Permanent Staff Tutor 2—2 tutes; Permanent Staff Tutor 3—2 tutes; Graduate Student Tutor 1—5 tutes; Graduate Student Tutor 2—3 tutes; Graduate Student Tutor 3—3 tutes; Casual Tutor 1—3 tutes; Casual Tutor 2—2 tutes; Casual Tutor 3—4 tutes; and Casual Tutor 4—1 tute.

absorbed by students who initially perform poorly on this measure may directly aid their understanding of economics, ultimately leading to success in the course in spite of their initial poor preparation. In this way, this "initial discipline knowledge" measure serves as an objective measure of poor student preparation that teachers of economics might reasonably be expected to counteract. As well, a disproportionately high attrition rate of students with poor preparation by this measure may be the result of having teachers who intentionally encourage such students to leave a field in which they have little chance of success.

The initial online assessment also provides other student-level background data used as control variables in the present study. The comprehensive list of available covariates used from this assessment is as follows: students' stated expectations of success in the course, on an ordinal scale of 0 to 4; students' answers to the following questions, coded into dummy variables: "Is English your first language?," "Have you taken economics elsewhere?", "Did you attend a public school?", "Has anyone else in your family ever been to university?", and "Do you like maths?"; whether the student is female; which of two alternative lecture times the student was enrolled for; the number of days after February 1 that the student in question completed the initial online assessment (this is included as a generic measure of commitment, organization, and/or motivation to undertake the course); initial discipline knowledge, as described in detail above; and highest maths level taken previously. This final variable is coded into five categories, as follows: a course taken at a school outside Australia (the left-out category in all models); Specialist Mathematics (Stage 2); Mathematical Studies (Stage 2); Mathematical Methods and Applications (Stage 2); any mathematical course (but only to Stage 1); or either another maths course (not in Stage 1 or Stage 2) within South Australia, or a maths course taken at a school outside of South Australia.

Data description

Before analysis, the original data set of 751 students ever enrolled in this course in any week of the semester was restricted in the following ways. First, 113 external students were dropped, as they did not attend tutorials and therefore could not be subject to the same type of teacher effects as internal students. An additional 19 internal students whose tutorial could not be identified were also dropped. Next, in an effort to simplify the analysis, only students who were enrolled in the course in either or both of weeks 1 and 2 were included (a restriction which caused the loss of 38 students). Finally, to allow multivariate analysis and permit a focus on students identifiable as poorly prepared, those students whose background information was incomplete because they did not complete the initial online assessment were also dropped. This left a final analysis sample of 494 students. While not a large sample by any definition, the detail available in this data set allows the examination of attrition and performance patterns by tutor in a quasi-case study format, with the goal of providing tentative results that can be further developed using larger samples.

Figure 1 shows the overall pattern of attrition from the course over time, in the form of a Kaplan-Meier survival curve by week estimated for the sample of leavers. 107 students of the starting sample of 494 leave the course sometime before the final examination, with the bulk of these exits (79 of 107) occurring after the final week of classes but before the examination. A drop-off during the first few weeks of the course is also evident. Attrition is defined quite strictly here: a student must have begun the course and stayed continuously enrolled until a given week, after which he or she left permanently. This allows the analysis to be uncomplicated, although there is the possibility that important dynamics regarding student thrashing in and out of the course may be missed. For this reason, an alternative measure of attrition—whether a student failed to sit the final examination—is also used below.

Table 1 shows the distribution of students across tutors and key weeks during the course of the semester. Most tutor-specific attrition rates are close to the overall average of 78%, with the exception of Casual Tutor 4, who retains all but one of the 18 students appearing in Week 2. While a Cox proportional hazards model yields a wide range of point estimates for the effects of different teachers, these estimates (available upon request) are not statistically distinguishable; and a chi-squared test of the association between attrition and tutor across the whole sample fails to reject the null of no association.

Panel A of Table 2 shows sample-wide statistics at the student level for a variety of background characteristics, all of which were drawn from the initial questionnaire and multiple choice test. Panel B of this table shows the same statistics, and adds in examination performance information, for the subset of students who sit the final exam. The two samples appear observably to be virtually identical, indicating that the decision not to take the final exam is not well-predicted by these background variables.

Finally, Figure 2 shows the distribution of scores on the key initial discipline knowledge variable constructed from the initial online assessment for students who fail to sit the final examination, compared to the distribution of scores for students who do sit the final exam. It appears from this graph that attrition from the course is occurring throughout the distribution of initial knowledge: there is no clear difference between these distributions.

Nonparametric empirical results

The primary purpose of the ensuing analysis is to determine the extent to which teachers differ in the patterns of student attrition and performance with which they are associated, with particular focus on poorer students. Because this is a first-year course populated by mostly inexperienced students, and because the assignment of tutors to tutorials is not made far in advance, the ability of students to select tutors in any informed fashion is limited. While causation cannot be incontrovertibly proven, some evidence is given in Table 3 that the starting observable attributes of students in this class are not systematically related to the identity of their tutors. This table shows the tutor-specific means of a variety of student attributes, including expectations of success in the course, initial discipline knowledge, whether English is the student's first language, and whether anyone in the student's family previously attended university. The primary measure of students' preparation, initial discipline knowledge, is shown in the first column. The means on this variable for all permanent staff and graduate student tutors hover closely around the global mean. While no differences in this column are statistically distinguishable, the students of two casual tutors appear slightly better prepared than those of other tutors. This could be due to random chance, or—given that these particular tutors have tutored in the course previously, and are familiar with the content of the online assessment exercise—it could be that students in these tutorials were "prepped" for the assessment by their tutors. In the latter case, the initial knowledge of students will be mismeasured for these two tutors. Other measures of student background do vary across tutors, but none significantly, and there is no apparent pattern wherein a subset of tutors is receiving students who appear better-qualified for university.

Students' examination performance, however, does differ systematically by teacher, as well as by initial discipline knowledge. Table 4 shows mean examination scores by teacher and initial preparation quartile. Across the board, students with better initial knowledge of the discipline perform better on the final examination, regardless of tutor. Comparing the scores of each tutor's students with the overall mean reveals that Permanent Staff Tutor 1 was associated with scores at least seven points above the global mean in 3 out of the 4 quartiles (including the lowest), whereas Permanent Staff Tutor 2 was associated with lower-than-average scores in all but the topmost quartile. Permanent Staff Tutor 3's best-prepared students fell far short of their cohort's average, while Casual Tutor 2's worst-prepared students fell far short of theirs. Other comparisons are not remarkable. It is important to bear in mind that given the small sample sizes, these differences are not statistically significant; however, they are useful as part of an emerging picture of potential differences by tutor in students' performance across the initial preparation distribution.

Results so far have indicated that those students who fail to complete the course are not overall less well-prepared than those who persist (Figure 2); yet, conditional on sitting the examination, better-prepared students perform better on the examination (Table 4). It has also been shown that while students are not sorted into tutorials by initial level of preparation (in spite of some possibility of "prepping" students by two casual tutors in particular), there does appear to be an association between teacher identity, on the one hand, and performance patterns, on the other. In order to partial out the effect of teachers—as distinct from the effect of ability or any other institutional or student-level characteristic—on both the performance and attrition of poorer students, we move now to multivariate regression where a host of potential predictors of student behavior can be controlled.

Parametric empirical results

Attrition

Table 5 shows results of simple linear probability models of attrition from the course using a host of student-specific and institutional covariates, including tutor identity. Columns 1 and 2 model the probability of a student having been continuously enrolled and then having left the course at any point during the semester; Columns 3 through 6 model the probability of a student failing to sit the final examination, regardless of her enrollment pattern to that point. All variables discussed above are included in these regressions, although only a subset of estimates is shown in the table. No variables excluded from the table were significant in any model.

Column 1 of Table 5 reveals that the only student-specific background variable that significantly impacts a student's likelihood of attrition is the number of days after February 1 that she was observed to complete the initial online assessment. The later the completion of the assessment, the more likely the student is to leave at some point during the semester. A student's initial knowledge of economics, by contrast, is not significantly related to her likelihood of attrition, and the sign on this variable is the opposite of what we would expect if more poorly-prepared students have a stronger tendency to leave.[9]

[9] As discussed above, attrition is captured starting in week 2. This decision is partly to allow students

The lower rows of this table show the estimates of the association of tutor assignment with attrition, where the excluded tutor dummy is that for the first of the three graduate student tutors, who taught the largest number of tutorials. Point estimates are quite different for different tutors, but one estimate stands out and is statistically significant: Casual Tutor 4 has a much stronger relationship with student attrition than any other tutor. According to the estimates in Column 1 of Table 5, Casual Tutor 4 is 13 percentage points more likely to retain a given student than is the excluded graduate student tutor, conditional on these covariates.

Column 2 of Table 5 shows the results of specifying the same model, but only including students whose initial discipline knowledge is below the mean score for the entire analysis sample. In this column, Casual Tutor 4 continues to be associated with lower attrition than Graduate Student Tutor 1, and late completers of the online assessment continue to be more likely to leave. Amongst these more poorly-prepared students, women are significantly more likely to leave. We also see some sign and magnitude changes in the tutors' effects, although nothing that is statistically significant at these sample sizes. There is therefore only a vague indication that the teacher effects pattern may be different for students of differing levels of initial discipline knowledge.

The final four columns of Table 5 show the results of models where the dependent variable takes a value of one if the student failed to sit the final examination, and zero otherwise. The pattern of results is extremely similar to what we see when modelling attrition from a base of continuous enrollment. Students who complete their online assessment early are less likely to skip the final exam; women (particularly those who were initially poorly-prepared) are less likely to sit; better-prepared students are no more likely overall to sit the final exam than other students; and Casual Tutor 4 seems to be retaining students more than other tutors. The pattern in tutor effect coefficients does not change even when all observable variables are excluded from the models, as shown in the final two columns of the table, further substantiating the claim that students are not matching

who are awaiting credit for a previous economics course to de-enrol and not be counted as having dropped out.

to tutors by observable characteristics.

Finally, the first two columns of Table 6 show results from models of attrition including the an identical array of control variables but parameterizing the effect of teachers differently. In these specifications, the main effect for Graduate Student Tutor 1 is again excluded, and two terms are included for each tutor: a main effect and an interaction of the main effect with an indicator for whether the student scored below the mean of the sample (.58) on initial discipline knowledge.[10] If a tutor disproportionately influenced more poorly prepared students to leave, then the interaction term coefficient for that tutor should be higher on the number line than that tutor's main effect. In fact, the opposite is the case for most tutors, and only Graduate Student Tutor 3 observes a significant difference in the expected direction in student attrition by preparation level. Casual Tutor 4 does not seem to be keeping students in proportion to their ability, although the main effect of this tutor on attrition continues to be negative and significant; and from Permanent Staff Tutor 3's tutorials, it is significantly more likely for better-prepared students to leave.

In summary, these models have illustrated that even in a small sample, tertiary tutors vary significantly with respect to the total amount of attrition they experience, and with respect to the preparation-biased shift they experience in their tutorials as the semester wears on. Surprisingly, only one tutor—Graduate Student Tutor 3—is significantly associated with the disproportionate departure of poorly-prepared students. In addition, few student-level observables are significantly related to attrition, with the exception of the initial measure of organization and/or commitment to the course as embodied in the date students complete the initial online assessment.

Performance

Moving to the second arena on which teachers may exert influence, we now examine the contribution to performance made by both teachers and student-specific observable characteristics. Table 7 shows the results of both ordinary least squares and Heckman

[10]The interacted variable for Graduate Student Tutor 1 is also included in the model.

selection models of final examination performance. Column 1 displays results for the full analysis sample and with all covariates included. Here, unlike in the attrition models, prior student preparation makes an enormous impact on success. Going from no initial knowledge of the discipline to perfect initial knowledge yields 41 marks on the 80-mark final examination; at the mean, the marginal effect of a one-standard deviation increase in initial discipline knowledge yields an final examination score that is an estimated .4 of a standard deviation higher, all else equal. This is an enormous effect, in the same league as only one other statistically significant effect in the table: the effect of whether English is the student's first language. A change from "no" to "yes" on this variable is predicted to yield an examination score higher by approximately .35 of a standard deviation, all else equal. Women are statistically significantly likely to do worse on the final examination, as are those who dallied in completing their initial online assessment.

The teacher effects pattern shown in Column 1 of Table 7 is very different than that seen in the attrition models estimated in the preceding section. Casual Tutor 4, who was associated with significantly higher student retention, does not seem to be associated with more measured learning than other tutors. In Column 1, only one tutor's performance is distinguishable from others in terms of increasing examination performance: Permanent Staff Tutor 1. This is the tutor who was technically a casual staff member, but had previously been a permanent staff member and was also giving the lectures in the course. When all other control variables are excluded, in Column 2, this tutor's effect still exceeds that of all other tutors (though it is not quite significant at the 5% level).

Column 3 of Table 7 shows the same specification, but only includes students whose initial discipline knowledge is below the mean score for the entire analysis sample. Here we find that language and initial knowledge continue to be important to the examination performance of more poorly-prepared students, although the female dummy has lost both magnitude and significance for this group. The latter rows of the table also show that the excluded tutor, Graduate Student 1, was associated with higher final examination performance for more poorly-prepared students than all tutors with the exception of one permanent staff tutor—and this difference is statistically significant for roughly half the tutors. In terms of point estimates, the tutor associated with the worst performance of more poorly-prepared students in this linear model is Permanent Staff Tutor 2, followed by Graduate Student Tutor 2 and Casual Tutor 1.

Column 3 of Table 6 (also referenced in the previous section) shows the interacted specification for performance with both tutor-specific main effects and an interaction of each main effect with an indicator for whether the student scored below the mean on initial discipline knowledge. This table further confirms the differential positive effect of Permanent Staff Tutor 1 on student performance, and indicates that that effect is not confined only to well-prepared students. Casual Tutor 4, who had previously been shown to be associated with lower attrition across the board, is shown here to be associated with both better performance for better-prepared students, and worse performance for more poorly-prepared students. Permanent Staff Tutor 3 and Graduate Student Tutor 3, by contrast, appear to assist more poorly-prepared students, and statistically to be neither a help nor a hindrance to better-prepared students (compared to the effect of the excluded tutor, Graduate Student Tutor 1). There seems to be little association between the performance effects seen here and the official affiliation that tutors have with the university.

Finally, Columns 4 and 5 of Table 7 show the results of a Heckman model of selection and ensuing second-stage performance. Results from the examination performance equation are shown in Column 4, and results from the selection equation modeling the first stage (sitting the exam) are shown in Column 5. Comparing the results in Column 4 to the analogous OLS results in Column 1, we can see few changes in the estimated effects of tutors, indicating that tutors are not idiosyncratically driving students' decision about whether to sit the exam in a way correlated with those students' likely examination success. This is despite the fact that one tutor is again distinguished with regard to the likelihood of losing students using this two-stage model, as seen in Column 5: Casual Tutor 4 is less likely to see students drop out. A Wald test for independent equations does not come close to rejecting for this model, indicating once again that students are not selecting into the examination on the basis of their likelihood of success.

Discussion

Teaching prototypes

To summarize the most robust results found here regarding tertiary tutors' association with attrition and performance, three tutors' cases are now reviewed: one from each of the tutor groups.

First, only one tutor, Graduate Student Tutor 3, was associated with the disproportionate out-selection of those students with the least initial knowledge of the discipline. This tutor was also associated with higher performance of more poorly-prepared students, relative to other tutors (as seen in Table 6). In this sense, Graduate Student 3 seems to be the sole tutor who played the two-part role of the broad-spectrum tertiary teacher suggested at the start of this paper: to catalyze the departure of under-prepared students, and also to facilitate their learning disproportionately, should they choose to stay.

Casual Staff Tutor 4 was associated with significantly higher retention across the board. As this tutor only took one tutorial and only lost one student, it is difficult to generalize regarding this tutor's tendency or lack thereof toward initial knowledge-biased student attrition. However, it is clear from the models of examination performance shown in Table 6 that students under this tutor did even worse if they were poorly-prepared, and even better if they were better-prepared, than the sample averages. This tutor essentially appears to have exacerbated the inequality in tested knowledge across students which existed at the start of the course.

Permanent Staff Tutor 1, who lectured and had been heavily involved in the course in previous years, was associated with the best student performance overall. While this tutor exhibited no particular pattern with regard to student attrition, both well-prepared and poorly-prepared students benefitted from this tutor in the final examination.[11] This

[11] This tutor may have had more input into the marking guidelines for the final examination than other

tutor may be said to exemplify the classically good teacher: able to meet students where they are, and add value at each point in the distribution of initial knowledge. Whether this ability is related to course-specific experience, enhanced coherence of tutorials with lectures, or other factors remains uninvestigated.

Relevance to pre-university teaching

While the empirical analysis in this paper relies only on data at the university level, the concepts discussed here are potentially relevant to teaching in other contexts as well. Further exploration of these concepts may prove rewarding for future research into teaching at the primary, secondary or tertiary level.

First, there is broad applicability of the notion that teachers' service to students may lie not only in catalyzing an increase in their stock of discipline-specific knowledge, but in providing students with information about the discipline that helps them to match themselves to an appropriate career path—which may involve leaving their original course of study. This information provision may be accomplished through a variety of methods, such as the elucidation of the parameters of the field of inquiry; the inculcation in the student of his ultimate responsibility for his own learning; the upholding of discipline-appropriate standards in delivery and assessment (and thus the avoidance of "dumbing down" course material to meet perceived student ability levels); and the provision of ancillary information relating to how knowledge emanating from the discipline is used in various occupations. The present paper has not distinguished amongst these sorts of information, but receipt of such information by undergraduate students new to a discipline may well encourage less well-prepared students to withdraw and choose a path better matched to their level of preparation. Similar, age-appropriate direction and potential redirection would seem suitable in the secondary school environment as well. This redirection is not necessarily a phenomenon that should be avoided, as it has the potential to ease the burden on both students and teachers. In future work, it would be of inter-

tutors, due to prior involvement in the course; but this tutor's positive and significant effect on performance holds even when predicting students' performance on only the multiple-choice section of the examination, which is not subject to interpretation by markers.

est to explore what sorts of messages from teachers are particularly effective in helping students to determine whether they are studying in an area that is well-matched to their preparation levels, ability, and interests.

Second, it seems likely from the results found here that what makes a good teacher is relative to what sort of student is being taught. Teaching poorly-prepared students is likely to demand a different skill set from that required to teach well-prepared students, as judged from the finding in this paper that only one of ten teachers was statistically distinguishable from other teachers in adding value to students with both high and low levels of initial knowledge. Three out of ten teachers seemed to specialize in assisting either well-prepared or poorly-prepared students—but not both. It is likely that teachers have an effect on learning at all points in the spectrum, compared to the counterfactual of teacher absence; but what is more interesting is the diversity of teacher effects, and their concentration for several teachers in either the lower or the upper end of the initial preparation spectrum. This indicates the presence of distinguishable skill sets or pedagogical approaches for different types of students, illustrating that "teacher quality," even as reflected in crude, short-term performance effects as done in this paper, is not necessarily a homogenous concept measurable on an absolute scale. Instead, it may be better conceptualized as relative to the presenting level of knowledge of the particular student being taught. Not only does this call into question direct comparisons of the quality of teaching across universities with entering students of different preparation levels, but it also highlights the potential efficacy of targeted teaching approaches to assist students at *both* ends of the preparation spectrum. However, isolating the approaches or sets of teacher attributes that work best for students at different points on this spectrum is beyond the scope of this paper.

Finally, the point should be made that while an alternative to the attrition of poorly-prepared students would be to deny them entry in the first instance, this is not a clearly superior option if our society values educational opportunity. Students' prior preparation levels in a particular discipline, whether at university level or below, may or may not reflect their innate levels of aptitude for that discipline; many students may simply not have been exposed to the discipline previously, or may have suffered educational disadvantages earlier in life. To deny them a chance to succeed, even if that chance brings with it the requirement of significant effort on the part of both teacher and student, brings with it a social cost. At the same time, there are costs involved in offering this chance. The relative costs of both options should be weighed by any society wishing to decide how much to invest in the provision of educational opportunity.

Conclusion

Broadened access to tertiary education has meant that modern tertiary classrooms are increasingly diverse across a wide range of variables, including prior knowledge of the fundamental academic disciplines taught in first year. This poses a challenge not only for the students themselves, but for the institutions catering to them. If broad-spectrum-focused players in the tertiary education market are to remain economically viable in the long run, they must either witness the out-selection of more poorly-prepared students, or the effective teaching of more poorly-prepared students, before those students graduate.[12]

This paper has explored the relative contribution of teachers to attrition and performance in a broad-spectrum, first-year undergraduate setting. Using a detailed microcosm of student data across one semester-long course and simple statistical techniques, I have confirmed previous findings at the primary and secondary level that there exist meaningful differences in value-added amongst teachers. While it is difficult to generalize with the small sample at hand, the performance effects found here differ by initial student preparation and do not correspond well with the traditional personnel sources from which tertiary tutors are drawn (permanent staff, graduate students, and casual academics). Little evidence has been found of attrition based on prior knowledge, either overall or (with one exception) under particular tutors.

Theorizing about the source of differences in teacher effects in such a small sample

[12] An economically feasible alternative would be to publicly reduce the HECS fees associated with ATN universities' programs across the board, a move that would re-formalize a second rung of tertiary education.

without more detailed information about each teacher's style, background, and approach to teaching, not to mention idiosyncratic characteristics of the classes taught, is ill-advised. What can be said is that variance in teacher effectiveness for students at different stages of preparation is of interest not only to researchers, but to educational institutions interested in the strategic alignment of their teaching model with their position in the tertiary education marketplace. While traditional universities may not require professors to add value to students who have already surpassed an admissions threshold, those universities that allow access by poorly-prepared students face an economic mandate: either facilitate poorly-prepared students' out-selection before degree completion, and/or provide sufficient resources to bring them up to an economically sustainable standard of knowledge and skills during the course of their degree. It may be that enhancing universities' sources of student support other than tutors would provide the most cost-effective way to increase the learning capacity of more poorly-prepared students. In the diverse tertiary environment studied here, however, and even in a small sample, there appears to be heterogeneity in the attrition and performance outcomes with which tutors are associated—a phenomenon that demands further study.

References

Beck, K. & Krumm, V. (1991), 'Economic literacy in German speaking countries and the United States: First steps to a comparative study', *Economia* **1**(1), 17–23.

Boyd, D., Grossman, P., Lankford, H., Loeb, S. & Wyckoff, J. (2006), 'How changes in entry requirements alter the teacher workforce and affect student achievement', *Education Finance and Policy* **1**(2), 176–216.

Darling-Hammond, L. (1999), Teacher quality and student achievement: A review of state policy evidence, Technical report, University of Washington Center for the Study of Teaching and Policy, Document R-99-1.

Davies, P. (2006), Threshold concepts: how can we recognise them?, *in* J. Meyer & R. Land, eds, 'Overcoming barriers to student understanding: threshold concepts and troublesome knowledge', Routledge, pp. 70–84.

Dee, T. S. (2005), Teachers and the gender gaps in student achievement. NBER Working Paper 11660.

Evans, M. (1999), School-leavers' transition to tertiary study: A literature review, Technical report, Monash University Department of Econometrics and Business Statistics Working Paper 3/99.

Goldhaber, D. (2006), Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? University of Washington working paper.

Hoffmann, F. & Oreopoulos, P. (2006), Do better professors produce better students? the effects of objective and subjective measures of teacher quality on academic achievement. University of Toronto Department of Economics working paper.

Kane, T. J., Rockoff, J. E. & Staiger, D. (2006), 'What does certification tell us about teacher effectiveness? Evidence from New York City', *SSRN eLibrary* .

Long, M., Ferrier, F. & Heagney, M. (2006), Stay, play, or give it away? Students continuing, changing, or leaving university study in first year, Technical report, Monash University, ACER Centre for the Economics of Education and Training.

McKenna, K. (1994), High school economics in Australia, *in* W. Walstad, ed., 'An International Perspective on Economic Education', Kluwer Academic Publishers, pp. 219–232.

Rivkin, S., Hanushek, E. & Kain, J. (2005), 'Teachers, schools, and academic achievement', *Econometrica* **73**(2), 417–458.

Rockoff, J. (2004), 'The impact of individual teachers on student achievement: Evidence from panel data', *American Economic Review* **94**(2), 247–252.

Rowe, K. (2003), The importance of teacher quality as a key determinant of students' experiences and outcomes of schooling. Background paper to ACER Research Conference.

Shanahan, M., Foster, G. & Meyer, J. (2006), Associations among prior acquisition of threshold concepts, learning dimensions, and examination performance in first-year economics. University of South Australia working paper.

Whitehead, D. & Halil, T. (1989), *The Test of Economic Literacy: Standardisation in the U.K.*, Institute of Education, University of London, London. Research Papers in Economics Education.

Table 1: Weekly attrition by tutor

	Week 2	Week 3	Week 5	Week 15	Sat exam
Tutor and number of tutorials[a]					
Permanent Staff					
1 (2 tutorials)	37	36	34	34	30 (81%)
2 (2 tutorials)	39	39	39	38	32 (82%)
3 (2 tutorials)	37	36	35	35	28 (77%)
Graduate Students					
1 (5 tutorials)	95	95	90	90	77 (81%)
2 (3 tutorials)	50	49	49	47	34 (68%)
3 (3 tutorials)	57	57	53	53	44 (77%)
Casual Tutors					
1 (3 tutorials)	56	57	55	55	44 (79%)
2 (2 tutorials)	34	32	31	30	26 (76%)
3 (4 tutorials)	66	66	63	63	51 (77%)
4 (1 tutorial)	18	18	18	18	17 (94%)
ALL	489	485	467	463	383 (78%)

[a]Numbers in the first four columns show the count of students enrolled in tutorials taught by the relevant tutor at key points during the semester, beginning in Week 2. The last column shows the number of students who sat the final examination, by teacher, both in absolute figures and as a percent of all students enrolled under that teacher in Week 2.

Table 2: Summary statistics: Full sample versus examination sitters

		Percentile			
	Mean	25th	50th	75th	Standard deviation
Panel A: Full Analysis Sample: N = 494[a]					
Expectations of course success	2.67	2.00	3.00	3.00	.68
Initial Discipline Knowledge	.59	.49	.59	.72	.16
Maths Level > 1	.88	-	-	-	-
English 1st language	.75	-	-	-	-
Economics taken elsewhere	.47	-	-	-	-
Public school	.86	-	-	-	-
Family at university previously	.65	-	-	-	-
Like math	.69	-	-	-	-
Female	.50	-	-	-	-
Panel B: Examination Sitters Only: N = 383					
Expectations of course success	2.67	2.00	3.00	3.00	.68
Initial Discipline Knowledge	.59	.49	.59	.72	.16
Maths Level > 1	.87	-	-	-	-
English 1st language	.77	-	-	-	-
Economics taken elsewhere	.48	-	-	-	-
Public school	.85	-	-	-	-
Family at university previously	.65	-	-	-	-
Like math	.69	-	-	-	-
Female	.49	-	-	-	-
Final Exam Score	45.42	33	45.50	57	16.19
Exam - Part A	16.83	13	17	21	5.17
Exam - Part B	15.37	9.5	15	21.50	7.50
Exam - Part C	13.22	8	13	18	6.35

[a] Panel A includes all students in the analysis sample; Panel B includes only those students in the analysis sample who sat the final examination. The sample size exceeds that shown in the previous table due to the presence of 5 students who appear in Week 1 but drop the course before the enrolment snapshot in Week 2.

Table 3: Key summary statistics by tutor

Attribute:	Initial Disc. Knowledge	Expectations	English 1st Lang	Family at uni
Tutor[a]				
Permanent Staff				
1	**.56**	**2.62**	**.77**	**.77**
	(.19)	(.63)		
2	**.56**	**2.73**	**.78**	**.63**
	(.13)	(.71)		
3	**.57**	**2.81**	**.73**	**.65**
	(.16)	(.62)		
Graduate Students				
1	**.60**	**2.61**	**.83**	**.59**
	(.15)	(.69)		
2	**.60**	**2.58**	**.66**	**.68**
	(.17)	(.73)		
3	**.58**	**2.72**	**.65**	**.68**
	(.15)	(.62)		
Casual Tutors				
1	**.58**	**2.70**	**.68**	**.63**
	(.18)	(.78)		
2	**.70**	**2.79**	**.76**	**.79**
	(.16)	(.64)		
3	**.58**	**2.64**	**.83**	**.59**
	(.14)	(.65)		
4	**.67**	**2.61**	**.78**	**.50**
	(.15)	(.78)		
ALL	**.59**	**2.67**	**.75**	**.65**
	(.15)	(.68)		

[a]Figures in parentheses are standard deviations within the sample of students enrolled under the given tutors.

Table 4: Examination scores by tutor*preparation quartile

Quality of Discipline Knowledge:	Top 25%	25-50%	50-75%	Bottom 25%
Tutor[a]				
Permanent Staff				
1	**62.25**	**57.19**	**38.40**	**44.06**
	(13.74)	(12.08)	(15.12)	(14.59)
2	**56.50**	**40.35**	**35.21**	**30.82**
	(10.54)	(16.69)	(17.73)	(12.01)
3	**41.83**	**51.64**	**48.93**	**30.88**
	(15.62)	(15.26)	(13.37)	(10.37)
Graduate Students				
1	**56.53**	**44.22**	**52.20**	**33.46**
	(16.86)	(15.30)	(11.14)	(15.29)
2	**56.20**	**50.83**	**44.09**	**29.86**
	(15.03)	(16.75)	(13.21)	(11.79)
3	**47.68**	**47.63**	**48.23**	**37.15**
	(9.61)	(16.70)	(16.14)	(14.59)
Casual Tutors				
1	**57.58**	**47.54**	**38.94**	**36.13**
	(17.57)	(16.95)	(11.33)	(13.71)
2	**56.05**	**45.13**	**49.00**	**22.00**
	(12.11)	(27.03)	(8.49)	(15.13)
3	**54.29**	**41.21**	**44.09**	**35.54**
	(15.62)	(13.51)	(11.4)	(12.17)
4	**58.42**	**50.83**	**41.17**	**27.00**
	(13.92)	(12.61)	(15.70)	(2.83)
ALL	**55.12**	**46.51**	**44.34**	**34.18**
	(14.66)	(15.83)	(13.73)	(13.63)

[a]Note: Figures represent the mean raw examination scores (out of 80) of students enrolled in tutorials taught by the relevant tutor. Figures in parentheses are standard deviations within the sample of students enrolled under the given tutor.

Table 5: Attrition models

	(1)	(2)	(3)	(4)	(5)	(6)
English 1st language[a]	-.07	-.00	-.07	.00		
	(.05)	(.05)	(.05)	(.05)		
Female	.07	.14*	.07*	.15		
	(.03)	(.06)	(.03)	(.06)		
Days to assess	.02**	.01*	.02**	.01		
	(.00)	(.01)	(.00)	(.01)		
Initial disc. knowledge	.05	.12	.06	.17		
	(.14)	(.34)	(.14)	(.35)		
CT1	.06	.01	.06	.01	.04	.07
	(.11)	(.13)	(.11)	(.12)	(.08)	(.07)
CT2	.04	-.06	.04	-.06	.05	.12
	(.06)	(.12)	(.07)	(.13)	(.06)	(.10)
CT3	.03	-.07	.05	-.08	.04	-.08
	(.07)	(.09)	(.07)	(.08)	(.06)	(.08)
CT4	-.13**	-.12*	-.13**	-.12*	-.13**	-.08**
	(.02)	(.05)	(.03)	(.05)	(.01)	(.06)
PS1	-.01	-.04	.04	.01	.04	.09
	(.05)	(.10)	(.06)	(.10)	(.03)	(.07)
PS2	-.02	-.20	-.00	-.17	.03	-.03
	(.06)	(.14)	(.06)	(.12)	(.03)	(.07)
PS3	.05	-.02	.05	-.01	.05	-.05
	(.03)	(.15)	(.03)	(.14)	(.03)	(.12)
GS2	.14	.00	.14	.00	.13	.09
	(.08)	(.13)	(.08)	(.13)	(.03)	(.10)
GS3	.01	.02	.00	.02	.04	.11
	(.06)	(.12)	(.06)	(.12)	(.03)	(.06)
N	494	234	494	234	494	234
R^2	0.07	0.11	0.06	0.12	0.01	0.03

[a]Column (1) models attrition for the full analysis sample; (2) models attrition for only poorly-prepared students (those who scored below the mean on initial discipline knowledge); (3) models failure to sit the final exam for the full analysis sample; (4) models failure to sit the final exam for only poorly-prepared students; and Columns (5) and (6) model failure to sit the final examination for the full and poorly-prepared samples, respectively."Days to assess" is the number of days after February 1 that the student completed her initial online assessment. "CT1" is the dummy for Casual Tutor 1; "PS2" for Permanent Staff Tutor 2; "GS3" for Graduate Student Tutor 3; etc. Graduate Student 1 is the excluded tutor dummy. Standard errors are clustered at the tutorial level here and in all subsequent models. Additional control variables are included in every regression, as discussed in the text.

Table 6: Tutor*low preparation student interacted models

	Attrition from course[a]	Failure to sit exam	Final exam score/80
CT1	.04 (.15)	.04 (.15)	-2.30 (5.99)
CT1*poor prep	.13 (.10)	.13 (.10)	1.69 (6.23)
CT2	.06 (.06)	.07 (.07)	-1.09 (3.12)
CT2*poor prep	.04 (.09)	.03 (.09)	.97 (5.78)
CT3	.16 (.08)	.20 (.11)	-3.19 (3.27)
CT3*poor prep	-.17 (.10)	-.20 (.13)	6.15 (3.40)
CT4	-.10** (.04)	-.10** (.04)	4.41** (1.21)
CT4*poor prep	.02 (.08)	.02 (.08)	-6.42** (2.63)
PS1	-.06 (.07)	-.01 (.05)	9.40** (1.69)
PS1*poor prep	.16 (.14)	.17 (.09)	-.61 (4.62)
PS2	.05 (.06)	.04 (.06)	-2.31 (6.70)
PS2*poor prep	-.05 (.09)	-.01 (.07)	-2.51 (8.16)
PS3	.16* (.07)	.16* (.07)	-2.03 (2.75)
PS3*poor prep	-.15 (.19)	.09 (.10)	8.23** (2.68)
GS1*poor prep	.08 (.10)	-.14 (.19)	3.81 (3.90)
GS2	.19 (.09)	.19 (.10)	-.72 (4.47)
GS2*poor prep	-.02 (.14)	-.02 (.14)	.23 (3.29)
GS3	-.06 (.06)	-.06 (.06)	-2.23 (2.88)
GS3*poor prep	.19** (.07)	.20** (.06)	8.70** (2.34)
N	494	494	387
R^2	0.09	0.09	0.35

[a]Estimates shown are of models of attrition (columns 1 and 2) and performance (column 3) using the entire analysis sample, and including all control variables included on previous tables. Tutor main effects are abbreviated as on previous tables. "CT1*poor prep" is a dummy for the interaction of Casual Tutor 1 with a student who scored below the mean on initial discipline knowledge; other interacted variables are labelled analogously.

Table 7: Performance models

	(1)	(2)	(3)	(4)	(5)
English 1st language[a]	6.15**		7.09**	6.16**	.24
	(1.94)		(2.49)	(1.99)	(.16)
Female	-3.28		-1.77	-3.29*	-.25
	(1.69)		(2.40)	(1.65)	(.12)
Days to assess	-.29*		.07	-.29	-.06*
	(.13)		(.18)	(.16)	(.02)
Initial disc. knowledge	41.36**		45.74**	41.35**	-.34
	(4.42)		(10.87)	(4.31)	(.51)
CT1	-3.78	.52	-11.49*	-3.79	-.13
	(4.37)	(2.24)	(5.31)	(4.26)	(.37)
CT2	-2.34	5.26	-9.40*	-2.34	.17
	(2.49)	(4.17)	(3.90)	(2.39)	(.21)
CT3	-1.35	-2.69	-1.72	-1.36	-.16
	(2.35)	(3.31)	(4.21)	(2.25)	(.23)
CT4	1.26	3.92	-5.56	1.28	.76**
	(1.67)	(1.99)	(4.05)	(1.81)	(.11)
PS1	7.13**	6.39	-1.57	7.13**	-.15
	(2.23)	(3.25)	(3.81)	(2.21)	(.20)
PS2	-5.94	-7.19**	-15.47**	-5.94*	.15
	(3.07)	(1.99)	(4.57)	(2.94)	(.20)
PS3	.55	-2.15	.1.11	.53	-.22
	(2.67)	(4.25)	(4.08)	(2.60)	(.12)
GS2	-2.59	.83	-12.45**	-2.62	-.35
	(3.16)	(5.95)	(4.42)	(3.19)	(.25)
GS3	-.17	.33	-5.39	-.17	.04
	(3.51)	(2.65)	(4.55)	(3.38)	(.19)
N	387	387	181	494	494
R^2	0.33	0.04	0.30	-	-

[a]Columns (1) and (2) model final exam performance for the full analysis sample; (3) models exam performance for only poorly-prepared students (those who scored below the mean on initial discipline knowledge); and (4) and (5) display the results of a Heckman selection model for the full analysis sample, with Column (4) showing results for the examination performance equation and Column (5) the selection equation. Tutor dummies are labelled following conventions in previous tables, and Graduate Student Tutor 1 is again the left-out tutor dummy.

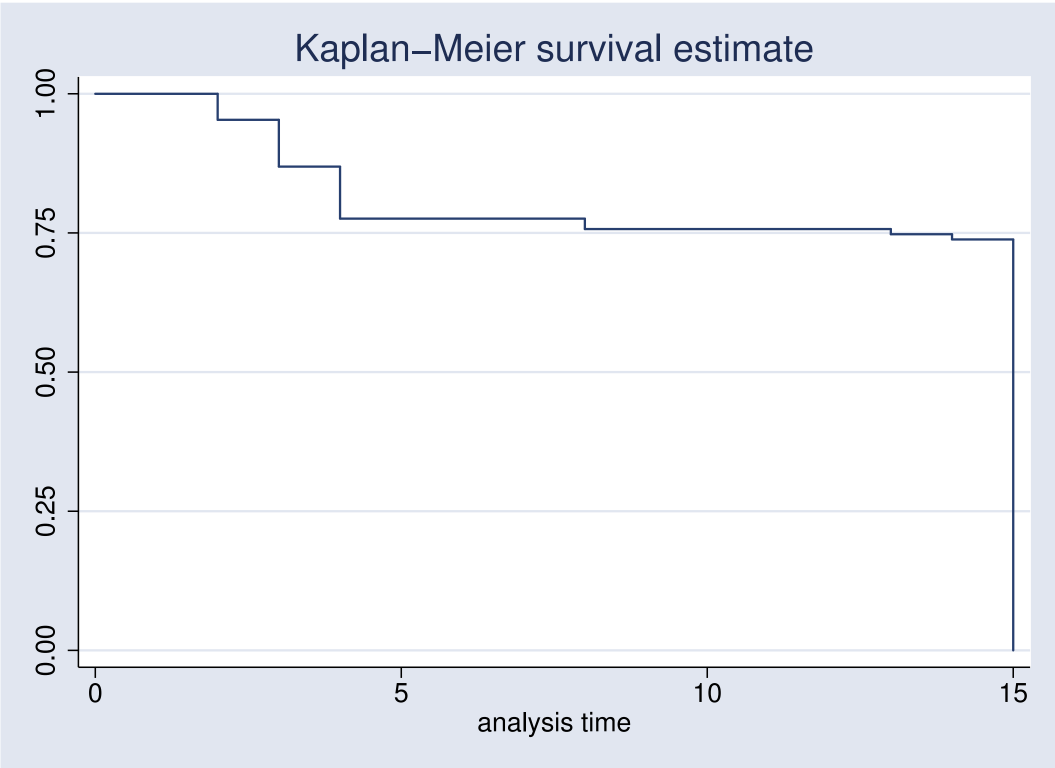

Figure 1: Attrition over time

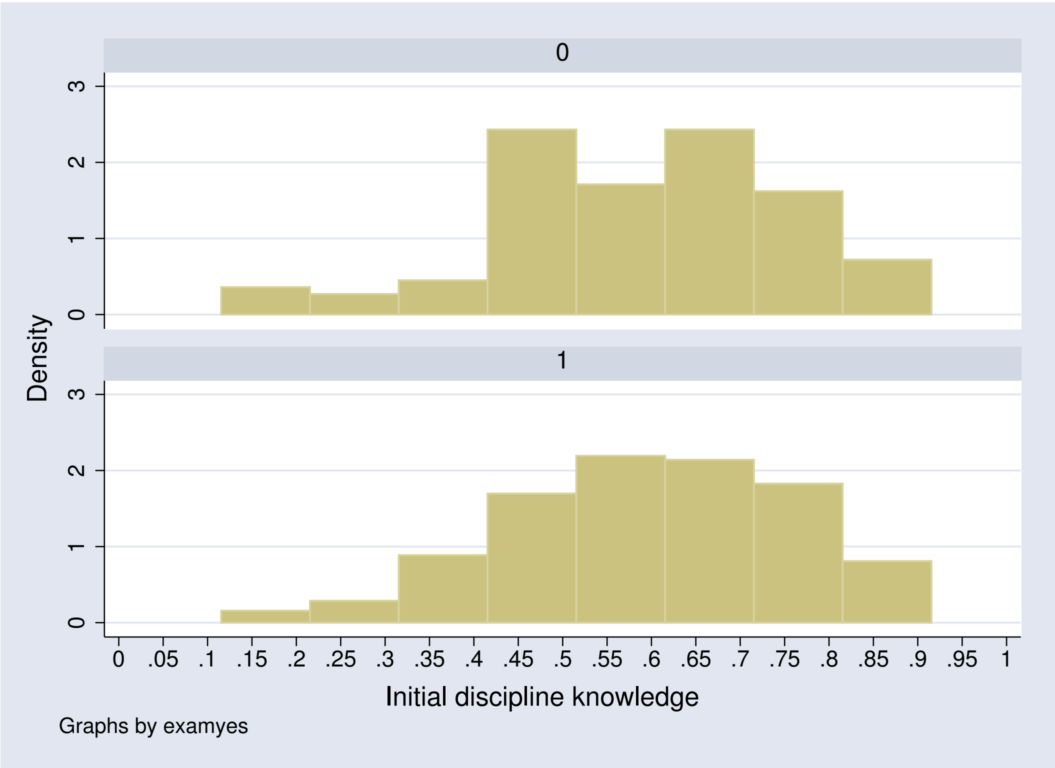

Figure 2: Initial discipline knowledge: Exam non-sitters (top) versus exam-sitters (bottom)