

NBER WORKING PAPERS SERIES

DO DOCTORAL STUDENTS' FINANCIAL SUPPORT PATTERNS AFFECT THEIR
TIMES-TO-DEGREE AND COMPLETION PROBABILITIES

Ronald G. Ehrenberg

Panagiotis G. Mavros

Working Paper No. 4070

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1992

Irving M. Ives Professor of Industrial and Labor Relations and Economics at Cornell University and Research Associate at the National Bureau of Economic Research, and Ph.D. Candidate in Economics at Cornell University, respectively. We are grateful to the Alfred P. Sloan Foundation and to Cornell University for their financial support of our research and to numerous colleagues at Cornell, the University of Rochester, and the National Bureau of Economic Research Higher Education Working group, especially Mark An, George Jakobson, Nick Kiefer, and Martin Wells, for their comments. We are also grateful to the Graduate Dean at the university whose data we use for granting us access to the data. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of any of the above-mentioned institutions or individuals or of the National Bureau of Economic Research.

DO DOCTORAL STUDENTS' FINANCIAL SUPPORT PATTERNS AFFECT
THEIR TIMES-TO-DEGREE AND COMPLETION PROBABILITIES

ABSTRACT

Projections of forthcoming shortages of Ph.D.s abound. Part of the reason is that American college graduates are much less likely to receive doctorates today than they were 20 years ago. Two important factors in this decline may be the increase in the length of time necessary for doctorate students to complete their programs that occurred over the period and the low completion rates of entrants into doctoral programs.

Among the policies urged to prevent future Ph.D. shortages are increasing support for graduate students. Surprisingly little empirical evidence is available on how different types of support (fellowships, research assistantships, teaching assistantships) are likely to influence times-to-degree and completion rates.

Our paper uses data on all graduate students who entered Ph.D. programs in four fields during a 25-year period at a single major doctorate producing university to estimate how graduate student financial support patterns influence these outcomes. We find that completion rates and mean durations of times-to-completion are sensitive to the types of financial support the students received. Other things held constant, students who receive fellowships or research assistantships have higher completion rates and shorter times-to-degree than students who receive teaching assistantships or tuition waivers, or who are totally self-supporting. A major finding is that the impact of financial support patterns on the fraction of students who complete programs is much larger than its impact on mean durations of times-to-degree.

Ronald G. Ehrenberg
School of Industrial and Labor Relations
Cornell University
Ithaca, NY 14853-0953
and NBER

Panagiotis G. Mavros
Department of Economics
Cornell University
Ithaca, NY 14853-7601

I. Introduction

Projections of forthcoming shortages of Ph.D.s, and thus new faculty for the academic sector, abound (William Bowen and Julie Ann Sosa 1989; National Science Foundation 1989; National Research Council 1990; Richard Atkinson 1990). Indeed, one major book projected at least a 43 percent underproduction of new doctorates in the arts and sciences as a whole during the 1997-2002 period (Bowen and Sosa, Table 8.5).¹ Part of the reason for these projections is that American college graduates are much less likely to receive doctorates today than they were 20 years ago. While the ratio of doctorates granted by American universities to bachelors' degrees granted by American colleges and universities six years earlier was .064 in 1970-71, it fell to .035 in 1978-79 and has remained roughly constant at the lower level since then (Ronald G. Ehrenberg 1991, Table 6.4).

Numerous factors probably contribute to the decline in the propensity of American college graduates to receive doctorates, however one important factor may well be the increase in the length of time necessary for doctorate students to complete their programs. The median registered time to degree for new Ph.D.s granted in the United States in 1968 was 5.5 years. By 1988, this figure has risen to 6.9 years. The increase has been even more dramatic in some fields; for example, in the social sciences median registered time to degree rose from 5.1 to 7.4 years and in the humanities from 5.5 to 8.5 years during the same period (National Research Council 1989, Table I).²

National data on completion rates for entrants into doctoral programs are not systematically collected. However, data were collected for a set of selected major research universities during the 1970s and early 1980s. These data suggest that completion rates, while varying widely across fields and institutions, tend to lie in the 40 to 70 percent range (Ehrenberg 1991, Table

7.6). Even the very best science graduate students, those who win prestigious National Science Foundation Graduate fellowships, had completion rates of 80 percent or less during the 1962-1976 period (Lindsey Harmon 1977; Joan Snyder 1988). These completion rates should be contrasted with completion rates of over 98 percent in top 20 American law schools, of over 90 percent in major American medical schools, and of 80 to 95 percent for top MBA programs in the United States.³ Doctoral study is considerably riskier than its alternatives and this fact surely also discourages potential students from entering doctoral programs.

Among the policies urged to prevent future Ph.D. shortages is increased federal, foundation and corporate support for graduate students. Such a policy would reduce the private costs of doctoral study and thus hopefully should increase the number of college graduates willing to undertake graduate study. To the extent that financial support reduces the time students need to complete degrees and increases their probability of completing doctoral programs, the future supply of Ph.D.s should further increase. While conceptually these roles of financial support on the supply of doctorates are clear, empirical evidence on the effects of financial support on doctoral production is actually quite scanty (Ehrenberg 1991, Chapter 8).

Two recent studies of the determinants of time-to-degree used regression models and either aggregate annual time-series data by field, or data on doctoral recipients from a single institution over a ten-year period (Howard Tuckman, et al. 1990; J. Abedi and E. Benkin 1987). These studies, both for methodological and data reasons, do not permit one to identify the role that changing student abilities, changing job market opportunities and changing

financial support for graduate students may have played in the lengthening of times-to-degree.⁴

These studies also focused only on doctorate recipients and ignored the possibility that students' ability levels, job market opportunities, and financial support patterns may also influence completion rates. Failure to take account of these relationships, when analyzing data on times-to-degree of doctorate recipients, will lead to biased estimates of the effect of these variables on times-to-degree due to sample selection problems (James Heckman 1979). An analysis of the response of doctoral completion rates to these variables is also important, in itself, as it also will directly provide information needed to improve projections of the flow of doctorates and to evaluate possibly policy changes.

Two other recent studies did analyze the behavior of recipients and dropouts (William Bowen and Neil Rudenstine 1992; Allison L. Booth and Stephen E. Satchell 1991). The former utilized data on all entrants to graduate programs in six fields at ten major research universities over a 25-year period and showed that there were differences in time-to-degree and completion rates associated with differences in the primary type of financial support students received during their graduate student years. However, multivariate behavioral models were not estimated; most of their analyses were restricted to two-way comparisons of means.

The latter used British data on about 480 entrants to Ph.D. programs in 1980. Their data did not permit them to analyze the effects of different types of funding (e.g., fellowships, research assistantships, teaching assistants), nor to analyze how the time path of funding types influences durations. Because their data set was nationally-based and included all

fields of study, they also could not estimate whether financial support patterns had differential effects across fields and whether what they called the effects of financial support were really institutional effects (individual's institutions of doctoral study were not identified in the data and support patterns may differ across institutions). Finally, because their data came from a single entering doctorate class, they could not attempt to estimate how changing labor market conditions influence degree times and completion rates.

To isolate the effects of changes in financial support, changes in student ability, and changes in academic labor market conditions on doctoral students' time-to-degrees and dropout rates, the research we report below analyzes individual-level data from a single major doctorate producing university (University X) on all graduate students who entered Ph.D. programs in four fields (Economics, English, Mathematics and Physics) during the 1962-1986 period. As described below, data are available on students' ability (GRE scores), time-to-degree (or drop out from the program), and type of financial support (if any) received by each student in each of the (up to) first six years that the student was enrolled in the program. Coupled with aggregate data from other sources on job market opportunities by fields and year, these data permit the estimation of competing risk "duration", or "hazard function" models of times-to-degree or to drop-out (Nicholas Kiefer 1988; Tony Lancaster 1990).⁵

The plan of this paper is as follows. In the next section, we provide some background data, both nationally and at University X, on times-to-degree and completion probabilities in the four fields. Time patterns in key explanatory variables at the university level (financial aid patterns and

student ability are also examined). Section III sketches our analytical model (details are presented in the appendix), discusses the data used to estimate the model and then presents the estimates of the model itself. Based on the results of the estimation, policy simulations are conducted in the next section to illustrate the effects of changing the patterns of graduate student financial support. A brief concluding section discusses the implications of our findings for public policy and directions for future research.

II. Background Data

Our analyses seek to explain the pattern of doctoral students' completion rates and times-to-degree for all students who enrolled in doctoral programs in economics, English, physics and mathematics at University X between 1962-1986. University X is one of the Research I institutions that participated in Bowen and Rudenstine's (1992) study of graduate education. It was chosen for this study because of the willingness of its graduate dean to provide us with more detailed information on graduate students' financial support patterns than were provided for the Bowen and Rudenstine study.

A weakness of most previous studies of times-to-degree was their failure to adequately control for changes in field-specific academic labor market conditions. The fields of economics, physics, and mathematics were included in this study because of the availability of historical data on starting assistant professor salaries that are collected annually by professional associations in each field. While historical salary data are not available for English, this field was included because national median times-to-degree have increased most rapidly over the last two decades in the humanities.

Nationwide data on median registered time to degree for doctoral recipients in these four fields, grouped by year of degree for the 1970 to 1988 period, are presented in the left-hand panel of Table 1. While median registered time-to-degree increased by about one-half a year in physics and slightly over one-year in economics and mathematics, median registered time-to-degree grew by over two years in English during the period.

Of course, nationwide patterns reflect both changes in degree times in individual institutions and changes in the shares of degrees produced by different institutions. Degree times tend to be longer at lower "quality" programs and during the period the share of doctorates in these fields produced by lower quality programs increased (Bowen and Rudenstine, Chapters 4 and 7). It is possible that the nationwide patterns reflect only the change in the share of degrees being granted by the lower quality programs and that degree times may not have increased during the period at high-quality Research I universities, such as University X.⁶

The right-hand panel of Table 1 presents median and mean degree time data for the four fields at University X for the 1970-88 period, with the data again grouped by year of degree. These data suggest that similar patterns of increases were in fact observed at University X. Physics exhibited either no (median) or a small (mean) increase, economics and mathematics larger increases, and English the largest increase in median (3 year) and mean (2.4 year) registered time-to-degree.

An important qualification about time-to-degree data grouped by year of completion is that even if the distribution of times-to-degree in each entering cohort remains constant across years, reported average times-to-degree by year of completion will change if the size of entering cohorts is

systematically changing over time (William Bowen, Graham Lord, and Julie Ann Sosa 1991). In particular, if entering cohorts are declining in size, average time-to-degree by year of completion will spuriously appear to increase since, as time proceeds, those completing degrees in a given year will increasingly come from "slow" completers from relatively large cohorts.⁷ During the 1972-1988 period the annual number of doctorates conferred by top graduate programs in the four fields under study did in fact trend downward (which suggests either that entering cohorts were declining or that completion rates were falling) and entering cohorts did decline in the English departments Bowen and Rudenstine studied (Bowen and Rudenstine, Table 4.5, Figure 5.4). Hence, it is of interest to ascertain how registered time-to-degree has varied over time for the four fields in University X, when the data are grouped by year-of-entry.

The answers are found in Table 2, where median and mean registered times-to-degree, as well as completion rates, are presented for entering classes from 1963 to 1979. Three year moving averages are used to smooth year-to-year fluctuations; as a result, the series run from 1964 to 1978.⁸

These data, grouped by year-of-entering class, convey a different picture than the data grouped by year-of-degree. There are no discernible trends in mean and median registered time-to-degree for mathematics and physics students, and, at most, a slight positive trend for economics students at University X. Times-to-degree have increased significantly for English students at University X, but only for students who enrolled after 1970. Finally, although completion rates fluctuate over time, again there are no discernible trends. Hence, the econometric work that follows will be directed as much at explaining differences in degree times and completion probabilities

across students in a given entering class, as it will be in explaining trends over time.

Prior to undertaking this econometric research, it seems prudent to ask whether variables that are known to, or thought to, have changed systematically over time nationally, and that also may be postulated to affect degree times and completion rates, have also systematically changed over time for the four fields at University X. Three obvious candidates that come to mind are the quality of entering graduate students, the proportion of new graduate students that are foreign, and the proportions of students receiving various types of financial support.⁹

Data on entering student quality, as measured by Graduate Record Examination scores, and on the proportion of entering students who are U.S. citizens and permanent residents are reported by year and field in Table 3. Only for the field of economics is there evidence that University X's graduate programs are increasingly attracting foreign students. In spite of well-known fears of test score declines, the quality of University X's graduate students (as measured by GRE scores) in these four fields has not declined over time. Only for the field of economics has either GRE score declined appreciably over the 25-year period and this decline in the verbal score is undoubtedly due to the increasing proportion of foreign students in the field's graduate program.

While the measured quality of University X's graduate students in the four fields did not decline over the period, the nature of their financial aid packages did change significantly. University X provided us with data on the major source of support received by each student who entered graduate study in these fields during the period, for each of the individual's (up to) first six years of graduate study. This enabled us to compute the proportion of

students in each program receiving fellowship, research assistantship, and teaching assistantship support during each year after program entry and these proportions appear in Table 4 for each of the first four years of study, grouped over time in three-year intervals.

On balance one observes fellowship support declining over time in each year of study for economics and mathematics graduate students, and teaching assistant support increasing (at least during students' first three years in the program). First-year English graduate students have tended increasingly to be supported by fellowships and decreasingly by teaching assistantships. However, in their second, third, and fourth year of studies these patterns were reversed, as the probability of advanced students receiving fellowship support decreased and their probability of receiving teaching assistantship support increased over time.

Physics is the only one of the four fields at University X in which research assistantships provided support for a substantial proportion of advanced graduate students. These proportions tended to increase over time, while the proportions of second, third, and fourth year physics graduate students receiving teaching assistantships correspondingly decreased. Finally, the proportion of physics graduate students who received fellowships in each year of study declined between the mid-1960s and mid-1970s, but then increased thereafter.

III. Econometric Estimation of Competing Risk Models

An economic model of doctoral students' times-to-degree and completion probabilities was developed by David Breneman (1976). At the risk of simplifying, the model focuses on the effects of academic labor market

opportunities and financial support for graduate students. Other things held constant, improved labor market opportunities are postulated to lead students to speed up degree progress and thus shorten times-to-degree. Similarly, the dollar levels and types (fellowship, research assistantship, teaching assistantship, etc.) of financial support graduate students receive are also postulated to affect degree-times and completion probabilities, because financial support levels influence opportunity costs and the types of support may directly influence degree progress. For example, teaching assistantship responsibilities may take time away from studies while research assistant responsibilities may (or may not) be complementary to students' dissertation research.

A formal dynamic model of graduate students enrolled in doctoral programs that incorporates job market opportunities and financial support variables is presented in the appendix. Since students can exit from doctoral programs either by receiving their doctoral degrees or by dropping out, this dynamic model leads naturally to the specification of a competing risks duration model.

Table 5 presents data on the frequencies of "failure" times, that is durations of years until degree completion or program dropout in the sample, as well as data on program durations for individuals whose programs were still in progress at the time the data were collected (1988-89). Over 70 percent of the completers in the fields of economics, English and physics and over 90 percent of the completers in mathematics did so in six years or less. Between 89 and 97 percent of the completers, depending on the field, did so in eight years or less.

Approximately 50 percent of all people who dropped out of doctoral programs in the four fields did so within their first two years of graduate school. While it has been alleged that graduate students in the humanities who drop out of doctoral programs do so at later stages of their programs than do graduate students in the sciences and social sciences (Breneman 1976), one does not observe this for graduate students in English at University X.¹⁰ Finally, depending upon the field, between 8 to 15 percent of individuals in the sample were still enrolled in graduate school at the time the data were collected. These individuals are primarily people who first enrolled in graduate programs in 1985 and 1986; they are treated as censored observations in the estimation that follows.

The data in Table 5 are used to estimate discrete time duration models, in which the two risks (receive a degree, dropout) are assumed to have independent error terms.¹¹ Individuals still enrolled in a program are treated as censored observations. Furthermore since no individual completed his or her degree in two years or less in the sample, we restrict our attention to durations of at least three years for completers. The estimation utilizes the complete sample of completers, dropouts, and individuals still in progress.

We assume that the hazard function for each risk (the conditional probability of leaving the program during the period given that the individual remained in the program up until the period) is of a proportional hazard form (Nicholas Kiefer 1988). We specify a flexible form for the baseline hazard. It is allowed to differ between time periods, but is assumed to be constant during each period.¹² The length of a period is naturally set equal to a

year. Thus, the hazard function for each risk, λ , is specified to be of the form

$$\lambda(t, X_t, \beta) = \lambda_0(t) \exp(X_t' \beta)$$

where X is the set of explanatory variables, and $\lambda_0(t)$ and β are a set of unknown parameters to be estimated. As discussed below, both time varying and fixed explanatory variables are included in the model. The likelihood function we maximize and its derivation are found in the appendix.

The explanatory variables included in the model are intended to capture the affects of student ability, financial support for graduate students, and new doctorates' labor market conditions, as well as a vector of other control variables. The latter include whether the student had a masters degree at the time of first registration in the program (MA), whether the student is a male (SEX), whether the student is a U.S. citizen or permanent resident (CTZN), and a time trend term (YR) to capture the effects of any omitted variables.¹³

Student ability is proxied by each student's verbal and mathematics graduate record examination test scores (GREV, GREM). While it would have been preferable to include other measures, such as the quality of each student's undergraduate institution, his or her rank in class, and the graduate admission committee rankings for him or her, such information was not available to us.

Labor market conditions for new doctorates are captured by both a supply and a salary variable. The former is the percentage of new doctorates in the field seeking employment in the academic sector in the given academic year (PEED). This variable (which varies each year the individual is enrolled in

graduate school) will decline as the academic job market weakens and thus more new doctorates are forced to seek either nonacademic employment opportunities or postdoctorate positions; the latter in an attempt to enhance their academic employability in subsequent years.¹⁴ The salary variable (for economics, physics, and mathematics) is the mean starting salary for new assistant professors in the field in the current academic year deflated by the consumer price index (SLRY); this also is a time-varying variable. The average graduate student stipend in University X was very highly correlated with the consumer price index and thus attempts to also include it in the model proved fruitless.

The financial support that the individual has received is captured each period by the proportion of years in the program that his or her major source of support came from receiving a teaching assistantship (PTA), receiving a research or a graduate research assistantship (PRA), or from tuition waivers, loans, or other means (POTH).¹⁵ The omitted category here is the proportion of years the student's major source of support was from fellowships.

These source of support variables will vary across years if the individual does not receive the same source of support each year. For example, if the individual received a research assistantship in his first year (PRA(1), PTA(1), POA(1)) would equal (1,0,0). If he or she then received a teaching assistantship in the second year (PRA(2), PTA(2), POA(2)) would equal (0,1,0). Finally, if a teaching assistantship was again received in the third year (PRA(3), PTA(3), POA(3)) would equal (0,1/3,2/3,0).

The impact of financial support on the hazards in each period are specified in this framework to depend only on the fraction of periods to date in which support of different types is received. So, in our example, if the

research assistantship had been awarded in the second rather than the first year, this would have altered the hazards in the first year but it would not have altered the hazards in years two or three. This is a restrictive assumption. A less restrictive approach would allow both the current period's type of support and the shares of different types of support in previous periods to influence the hazards each period. Unfortunately, due to collinearity, such an approach could not be implemented.

These major sources of support variables are available to us only for (up to) the first six years each individual was enrolled in the program. This creates a problem because, as indicated in Table 5, only 72 to 91 percent of all completers (depending on the field) actually complete their programs in six years or less (although 95 percent of all dropouts occur during this time frame).

There are two ways to handle this problem. First, one may simply treat individuals who complete, or dropout, in more than six years as censored observations. This is the approach used in the estimation of the models reported in Table 6. Second, one can assume that no student receives financial support from fellowships, teaching assistantships, or research assistantships after year six and then include completers and dropouts who leave the program after their sixth year in the analysis by updating their financial support variables in later years accordingly. This is the approach used in the estimation of the models reported in Table 7, although we truncate the analyses underlying this table after eight years because of the small numbers of dropouts and completions that occur after this duration.

The pattern of results in Tables 6 and 7 are quite similar. Turning first to the control variables, individuals who had masters degrees prior to

entering their doctoral programs tend to be more likely to complete their programs and less likely to dropout in any period than individuals without masters degrees. For the most part, gender does not appear to matter, although males do tend to complete programs more rapidly than females in English and to have lower dropout rates than females in physics. U.S. citizens and permanent residents tend to have smaller completion hazards (longer durations) and larger dropout hazards (save for English) than do foreign residents. Finally, after controlling for these variables, and the measures of student ability, new doctorate labor market conditions, and financial support patterns, time trends are observed in a number of hazards. In particular, completion and dropout hazards declined over the period in both English and physics, while the dropout hazard declined in economics and the completion hazard declined in mathematics.

Somewhat surprisingly, for the most part students' ability, as measured by their graduate record examination scores, is not associated with completion and dropout probabilities. The only exceptions are in physics, where higher verbal scores lead to lower dropout probabilities, and in English, where higher verbal scores perversely are associated with lower completion probabilities. Labor market conditions for new Ph.D.s, as measured by starting academic salaries or the proportion of new doctorates in the field seeking academic appointments rarely influence completion probabilities.¹⁶ These results (or nonresults) may reflect the incomplete nature of the student ability measures, the lack of data on nonacademic salaries, and the limited applicability of the crude national new doctorate labor market conditions measures to the doctorate students from this one elite institution.

Of primary interest to us is the role of financial support patterns for graduate students and here we find statistically significant results. Relative to the omitted group (fellowship holders), students with teaching assistantships are less likely to complete degrees in any period in economics, English, physics, and mathematics and are more likely to dropout in any period in physics and mathematics. Similarly, students in most fields with other forms of support (primarily loans, tuition waivers, and self support) are less likely to complete degrees and more likely to dropout in any period. Finally, although the result is statistically significant for only one field (mathematics), students with research assistantships are less likely to dropout in any period, than students with fellowships. This result may reflect that research assistants tend to be closely tied to individual faculty members who may provide the student with more direction, and thus, more attachment to the program.¹⁷

Of course, some may argue that, within a single program, financial support is allocated by ability. Thus, if the "better" students receive fellowships and the "weaker" students receive teaching assistantships and we observe students with fellowships having higher completion hazards and lower dropout hazards than those with teaching assistantships, this may tell us nothing about the affects of support type on times-to-degree and completion probabilities. Rather, it may simply reflect that support type is a better proxy for student ability than are graduate record examination scores.

We might find such an argument compelling if the distribution of graduate students by type of financial support remained constant for a field over time. However, relatively small changes in average GRE scores of entering students in the four fields at University X over time (Table 3) have

been accompanied by relatively large changes in the distribution of graduate students by types of support (Table 4). The latter have been caused by variations over time in the proportions of graduate students supported by federal funds at the national level. As institutional funds are substituted for external funds, the types of support provided students have changed (Ehrenberg 1991, Chapter 7; Ehrenberg, Rees and Brewer 1992, forthcoming). Hence, in what follows, we shall treat the estimated effects of these financial support variables as truly reflecting financial support patterns, not differences in ability.¹⁸

IV. Policy Simulations

One can use the estimates presented in Tables 6 and 7 to simulate what the effects of differing graduate student financial support patterns are on the proportion of an entering graduate student class that will complete doctoral degrees, the proportion that will drop out of the program, and the mean durations of time-to-degree and dropout.¹⁹ Initially, we present simulations for individuals who are assumed to either always receive fellowships, to always receive teaching assistantships, to always receive research assistantships, or to always receive other forms of support. In each case, estimated hazard rates are computed each period for an individual who is assumed to have the mean value for graduate students in the field of all of the other variables in the model.

The simulations presented in Table 8, which are based on the estimates in Table 6, focus on completers and dropouts within the first six years of study. Individuals who might complete, or drop out after six years, are treated as still enrolled in these simulations. Similarly, the simulations in

Table 9, which are based on the estimates in Table 7 focus on completers and dropouts within the first eight years of graduate study. Individuals who might complete or drop out after eight years are treated as still enrolled in these simulations.²⁰

Turning first to Table 8, the simulations highlight the importance of utilizing a competing risks model that includes completers, dropouts, and students still enrolled in the analyses. Differences in mean times to degree, or to drop out, across financial support patterns in these simulations are small, with the range between the longest and the shortest estimated durations for a field typically being between .2 and .5 years. What is remarkable, however, is how much the different financial support patterns appear to influence the distribution of individuals between completion and dropout status.

For example, out of 100 individuals who receive fellowship support each year in physics, 78 are predicted to complete their degrees within six years, with a mean duration of 5.28 years, while 16 are predicted to have dropped out within six years, with a mean duration until dropout of 2.75 years. In contrast, out of 100 physics graduate students who receive teaching assistantships each year, only 32 are predicted to complete their degrees within six years, with a mean duration of 5.38 years, while 34 are predicted to dropout within six years, with a mean duration of 2.79 years.²¹ The differences in mean durations that occur between the two different financial support patterns are very small, but the changes in the number of completers and dropouts that occur are dramatically large.

Of course, the change in the mean duration of completers is limited by the truncation in Table 8 at durations of six years.²² When we extend the

simulations in Table 9 to consider individuals who complete degrees or drop out in eight years or less, slightly larger changes in mean durations do occur. For example, when we move from all fellowship support to all teaching assistantship support for an entering class of 100 students in physics, mean duration for completers rises from 5.46 to 5.80 years and mean duration for dropouts rises from 2.75 to 3.14 years. These changes are again dwarfed, however, by the vast reduction in the number of completers (from 82 to 44) and increase in the number of dropouts (from 16 to 37) that occur.²³

In general, the simulations in both Tables 8 and 9 suggest that fellowships and research assistantships increase completion rates and decrease dropout rates relative to teaching assistantships and to all other forms of support. While changes in financial support packages are associated with changes in mean durations for both types of exit from graduate study, we emphasize again that the effects of financial support are felt primarily through their influences on the shares of students who complete their degrees or drop out.

In addition to performing "all or nothing" simulations, one can also use the estimates in Tables 6 and 7 to simulate how changes in the time patterns of different types of support will influence mean durations and completion and dropout probabilities. For example, Table 10 presents simulations of the numbers of completers and dropouts within six years (and their respective mean durations) for an entering class of 100 mathematics graduate students in which these students are assumed to receive fellowships for one year in the program and teaching assistantships in all other years. The question asked in the table is whether the year that fellowship support is received matters?

The results suggest that the earlier that the year of fellowship support is received, the higher the completion rate and the lower the dropout rate will be. For example, providing the students with fellowships in the first year rather than in the fourth year is predicted to raise the number of completers from 54 to 60 and reduce the number of dropouts from 41 to 35.²⁴ The mean duration for dropouts is correspondingly predicted to rise by .6 years (with fellowships in the initial year dropouts occur later), while the mean duration of completers is predicted to fall by .06.²⁵

V. Concluding Remarks

We have provided evidence for doctoral students in four fields at University X that their completion rates, their dropout rates, and the mean durations of their times-to-completion and to dropout are all sensitive to the types of financial support the students received. The impact of financial support patterns on the fractions of students who complete and drop out of the programs is much larger than its impact on mean durations of times-to-degree or to dropout. Previous studies that have focused solely on times-to-degree for completers may thus have missed the more important role that financial support patterns play in facilitating the production of doctorates from a cohort of entering graduate students.

Our study is, of course, a study of only one institution's experiences in selected graduate fields. Would similar results be found for other fields and institutions? Do differences in the sensitivity of outcomes to financial support patterns across fields at this one institution reflect differences in the nature of graduate programs and the types of students who enroll in each program, or do they reflect the fact that the "quality" rankings of these four

fields at University X vary substantially? To answer such questions requires that our analyses be replicated at other institutions and for other fields.²⁶

While we have successfully identified the effects of financial support patterns on these outcomes, we have not found much evidence that either student ability or labor market conditions for new doctorates influence doctoral student completion rates or degree times. Better measures of the labor market conditions facing new doctorates at individual institutions are obviously required. While it is conceivable that graduate student quality within a field at an elite institution may vary too little to enable one to isolate "quality effects", our sense is that efforts to obtain additional quality measures, such as undergraduate institutional quality, undergraduate grade point averages, and admissions committee ranking scores might prove fruitful.

Given the projections that Bowen and Sosa and others have made that there is likely to be a shortage of new American citizen and permanent resident doctorates starting in the mid-1990s, the issue arises as to whether the federal government, corporations, and private foundations should be encouraged to increase their funding of fellowships and research assistants to help avert such a shortage. Of course, as long as salaries are free to rise, shortages will eventually be eliminated. Concern over potential shortages of doctorates in academe occurs both because academic institutions may not possess the resources to increase faculty salaries substantially, and because, even if they do, the time it takes graduate students to complete doctoral degrees is sufficiently long that an increase in graduate enrollments in response to a salary increase would increase the supply of new doctorates only many years later. Thus, if shortages do materialize in the future, they may

persist for a number of years. Hence, taking actions now to increase the flow of doctorates may be required.

Evidence presented by one of us elsewhere suggest that an increase in externally provided funds for graduate student support would not induce universities to reduce their own support for graduate students (Ehrenberg, Brewer, and Rees 1992, forthcoming). That is, the additional funds would be used for their intended purposes. The evidence we have presented here suggests, at least for graduate students in the four fields at University X, that increased fellowship and research assistant support would lead to increased Ph.D. completion rates, lower dropout rates, and some reductions in durations of time-to-degree.²⁷

While these reactions, in themselves, would help increase the flow of new doctorates that would derive from a given size entering cohort of graduate students, increases in the number of fellowships and research assistantships provided for graduate students should also lead to an increase in the number of students that enroll in doctoral programs. This should occur both because of the direct impact of increased availability of these types of financial support on the opportunity cost of graduate study and because of their indirect impact on opportunity cost via their reducing times to degree and increasing completion rates. Unfortunately, empirical evidence on the magnitude of these responses is virtually nonexistent (Ehrenberg 1991, Table 8.1). Research on these responses should be high on the agenda for those concerned with academic labor supply issues (Ehrenberg 1991, Chapter 10).

Finally, we must stress that our research has addressed the importance of the type of financial support students receive. It has not addressed how the dollar level of support, measured either in real terms or relative to the

salary level of new doctorates influences the propensity of people to apply to and enter doctoral programs, and also doctoral students' times-to-degree and completion rates. As noted above, the average graduate student stipend in University X was very highly correlated with the consumer price index during the period our data span and thus attempts to include the former in our model proved fruitless.

While higher stipend levels should encourage more students to enroll in doctoral programs, their effects on times-to-degree and completion rates are uncertain. On the one hand, higher stipend levels should serve to reduce students' financial worries and, by eliminating their need to turn to part-time nonacademic employment for additional support, should speed up degree progress. On the other hand, higher stipend levels relative to new doctorate salaries reduce the incentive students have to rapidly complete their programs, and thus may actually slow down degree progress. Hence, the effect of higher stipend levels on the supply of new doctorates is a priori uncertain and empirical research on this issue is also required.

References

- Abedi, J. and E. Benkin. 1987. "The Effects of Students, Academic, Financial and Demographic Variables on Time to Doctorate." Research in Higher Education 17(1): 3-14.
- American Medical Association. Undergraduate Medical Education. 1988. Journal of the American Medical Association 260(26 August): 1063-71.
- An, M.Y. 1991. "A Structural Model of Labor Market History: The Event Planning Approach." Ithaca, NY: Cornell University Economics Department, mimeo.
- Atkinson, Richard C. 1990. "Supply and Demand for Scientists and Engineers: A National Crisis in the Making." Presidential address delivered to the American Association for the Advancement of Science, New Orleans.
- Barron's Guide to Law Schools. 1986. 7th ed. New York: Barron's Educational Services.
- Bhattacharya, R.N. and M. Majumdar. 1989. "Controlling Semi-Markov Models - The Discounted Case." Journal of Statistical Planning and Inference 21: 365-381.
- Booth, Alison L. and Stephen E. Satchell. 1991. "The Hazards of Doing a Ph.D.: An Analysis of Completion and Withdrawal Rates of British Ph.D.s in the 1980s." London: Birbeck College Discussion Paper in Economics.
- Bowen, William G., Graham Lord, and Julie Ann Sosa. 1991. "Measuring Time to the Doctorate." Proceedings of the National Academy of Sciences 88(3): 713-17.
- Bowen, William G. and Neil L. Rudenstine. 1992. In Pursuit of the Ph.D. Princeton, NJ: Princeton University Press.
- Bowen, William G. and Julie Ann Sosa. 1989. Prospects for Faculty in the

- Arts and Sciences. Princeton, NJ: Princeton University Press.
- Breneman, David W. 1976. "The Ph.D. Production Process." In Education as an Industry, ed. J.T. Fromkin, D.T. Jamison, and R. Radner. Cambridge, MA: Ballinger.
- Ehrenberg, Ronald G. 1991. "Academic Labor Supply." Part II of Charles T. Clotfelter, Ronald G. Ehrenberg, Malcolm Getz, and John J. Siegfried, Economic Challenges in Higher Education. Chicago, IL: University of Chicago Press.
- Ehrenberg, Ronald G., Daniel I. Rees, and Dominic J. Brewer. 1992. "Institutional Responses to Increased External Support for Graduate Students." Ithaca, NY: Cornell University, mimeo.
- Ehrenberg, Ronald G., Daniel I. Rees, and Dominic J. Brewer. Forthcoming. "How Would Universities Respond to Increased Financial Support for Graduate Students." In C. Clotfelter and M. Rothschild, Eds., The Economics of Higher Education. Chicago, IL: University of Chicago Press.
- Han, A. and J.A. Hausman. 1990. "Flexible Parametric Estimation of Duration and Competing Risk Models." Journal of Applied Econometrics 5(): 1-28.
- Harmon, Lindsey. 1977. Career Achievement of NSF Graduate Fellows: The Awardees of 1952-72. Washington, DC: National Research Council.
- Heckman, James. 1979. "Sample Bias as Specification Error." Econometrica 47(1): 153-62.
- Jones, Lyle V., Gardner Lindzey, and Peter E. Coggeshall, Eds. 1982. An Assessment of Research-Doctorate Programs in the United States. 5 vols. Washington, DC: National Academy Press.
- Kiefer, Nicholas. 1988. "Economic Duration Data and Hazard Functions."

Journal of Economic Literature XXVI(2): 646-79.

Lancaster, Tony. 1990. The Econometric Analysis of Transition Data.

Cambridge, England: Cambridge University Press.

National Research Council. 1989. Summary Report, 1988: Doctorate Recipients from United States Universities. Washington, DC: National Academy

Press.

_____. 1990. Biomedical and Behavioral Research Scientists: Their Training and Supply. Vol. 1, Findings. Washington, DC: National

Academy Press.

National Science Foundation. 1989. "Future Scarcities of Scientists and Engineers: Problems and Solutions." Washington, DC: National Science Foundation, Division of Policy Research and Analysis, Directorate for Scientific, Technological, and International Affairs, mimeo.

Snyder, Joan. 1988. "Recent Trends in Higher Education Finance, 1976-77 to 1985-86." In Higher Education Administrative Costs: Continuing the Study, ed. Thomas P. Snyder and Eva C. Galambos. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Tuckman, Howard, Susan Coyle, and Yupin Bae. 1990. On Time to the Doctorate. Washington, DC: National Academy Press.

Footnotes

1. See Ronald G. Ehrenberg (1991) for a less pessimistic view of the future.
2. Nationally, median total time-to-degree (the length of time between an individual's receipt of a bachelor's degree and his or her receipt of a doctoral degree) has risen by even more than median registered time-to-degree in most fields (Ehrenberg 1991, Table 7.4). Among the factors response for this increase is that individuals are increasingly delaying entry to graduate school (Ehrenberg 1991, Table 7.5).
3. The law school data come from Barron's (1986) and the medical school data from American Medical Association (1988). The MBA data are "guestimates" provided by James Schmotter, Associate Dean at Cornell's Johnson School of Management.
4. See Ehrenberg (1991), pp. 190-194, for a complete description and critique of these studies.
5. While time-to-degree equations can be estimated using linear regression models, or generalizations such as the Tobit model which permits the specification of lower (or upper) bounds on completion time, these are not the appropriate statistical methods to use for two reasons. First, such analyses of time-to-degree for completers from a given entry cohort ignore the experience of individuals from the cohort who are still enrolled in the program and who ultimately may complete their degrees. Other things held constant, the latter group will obviously have longer times-to-degree and eliminating them from the sample will understate average time-to-degree and may lead to biased coefficient estimates.

Second, to the extent that the whole time path of types of financial support (e.g., fellowship, research assistantship, teaching assistantship) that a student receives influences his or her progress through doctoral study and current and expected future labor market conditions influence a student's decision as to how rapidly to begin, and complete his or her dissertation, times-to-degree will depend on entire sequences of financial support and labor market condition variables. The number of years of data in the sequence "relevant" to any person will depend on the number of years it takes the individual to complete his or her degree. However, the latter is the outcome one is trying to explain so that obvious simultaneity problems will exist.

A simple numerical example illustrates this point. Suppose that financial support does not influence time-to-degree, that all students receive assistantships for their first 4 years of graduate study and no support thereafter, and that times-to-degree randomly vary across students and are either 4, 5, 6, or 7 years. Given the way support is allocated, the proportion of years support is received would be 1.0, .8, .67, and .57, respectively, for students who receive degrees in 4, 5, 6, and 7 years. Regressing time-to-degree on proportion of years support was received would lead to a negative relationship even though by assumption support has no affect on time-to-degree.

To avoid these problems, one can estimate "duration" or "hazard function" models. Rather than directly estimating time-to-completion equations, one estimates the probability that an individual will complete his or her degree in a given year and the probability that he or she will drop out of the program in a given year, both conditional on the student's having "survived" in the program up until that year. Because the focus is on

conditional probabilities in a year, both completers and dropouts in the year, as well as those individuals with degree programs still in progress, are included in the analysis. Because each individual/year observation is "treated" as a separate observation in the analysis, this method permits time varying covariates and the researcher can "update" the financial support and labor market condition variables that a student faces each year.

Booth and Satchell have also estimated a competing risk model. However, their financial support variables are not permitted to be time varying and they use a different econometric approach than we do.

6. All four of these graduate programs at University X were ranked in the top 20 nationwide, in terms of the quality of their faculty, in the last national ranking of graduate program quality (Lyle Jones, Gardner Lindzey and Peter Coggeshall 1982).

7. A simple numerical example illustrates this point. Suppose that all entering students receive degrees, that (unrealistically) half of each year's entering doctoral cohort complete in one year, and the other half complete in two years. Average time to degree by year of entering cohort is thus constant at 1.5 years. Suppose that, in years 0 and 1 (and all previous years), entering cohort size is 100. The table below shows that reported time to degree by year of completion will increase from 1.5 to 1.526 years if starting in year 3, entering cohort size decreases by 10 percent per year.

Year	Entering Cohort Size	No. Who Will Complete in $t + 1$	No. Who Will Complete in $t + 2$	Average Time to Degree of Completers in the Year
0	100	50	50	1.5 ($\{50 \times 1\} + \{50 \times 2\}$)
1	100	50	50	1.5 ($\{50 \times 1\} + \{50 \times 2\}$)
2	90	45	45	1.5 ($\{50 \times 1\} + \{50 \times 2\}$)
3	81	40.5	40.5	1.526 ($\{45 \times 1\} + \{50 \times 2\}$)
4	72.9	36.45	36.45	1.526 ($\{40.5 \times 1\} + \{45 \times 2\}$)

8. The analyses end with the entering class of 1978 to avoid truncation problems. The data were collected at University X in 1988-89, hence only individuals who had received their degrees in 10 years or less would be included as completers from the class of 1978. As noted below, over 99 percent of degree recipients in each field received their degrees in 10 years or less.

9. See Ehrenberg 1991, Chapters 7 and 8, for a discussion of how these variables have trended nationally.

10. The comparisons across fields in Table 5 are confounded by variations in quality, or prestige, across the four programs. In recent years, English has been the most highly related program of the four.

11. Although we do not do so here, it is straightforward conceptually (but computationally more difficult) to allow the two risks to have correlated error terms.

12. A. Han and J.A. Hausman (1990) present a similar specification in which the baseline hazard is not necessarily constant within each period, but rather to be linearly changing with time within each period.

13. Data limitations preclude us from including other variables that might be postulated to also influence completion rates and degree times. Examples of such variables are marital status, number of dependents, age, and whether the individual's undergraduate field of study is the same as his or her doctorate field.

14. See Ehrenberg (1991, Chapter 7) for a discussion of the role of postdoctorate positions in the academic marketplace. Other possible measures of academic labor market tightness exist. These include the actual sectors of employment of new doctorate recipients with definite employment commitments,

the share of new doctorates with definite postdoctorate positions, and the unemployment rates of new doctorates who sought employment. Like the measure we employ, these all might well be considered endogenous; they all depend upon the decisions of potential new doctorates whether to postpone completion of their program. Moreover, none of these alternative measures is available back through the mid-1960s, as is PEED.

15. Graduate research assistantships differ from research assistantships in that the former, found in some government programs, are intended primarily to benefit the recipient and not primarily to enhance a faculty member's productivity.

16. Preliminary analyses suggested that the included labor market conditions variables never influenced the dropout hazards and they are omitted from the specifications reported here accordingly. Attempts to include an alternative earnings measure in the dropout hazards (starting salaries for MBA's) in the economics field equation also failed to yield significant results.

17. Faculty with research grants typically "hand-pick" their research assistants and, since it is desirable to employ an assistant over a number of years, they are likely to choose people who they believe have (unobservably to us) low probabilities of dropping out.

18. One extension warrants being briefly reported here. Requirements for the Ph.D. often change over time for a given field. For example, during the period the field of economics introduced core requirements in micro-economics, macroeconomics, econometrics, and mathematics, eliminated its language requirement, and reduced the number of subject matter areas (e.g., labor, development, trade) students were required to study from three to two.

Such changes potentially may affect completion rates and degree times. However, when attempts were made to model such changes by using dichotomous variables to indicate "regime" switches, no significant effects were found.

19. One can also see how well the estimated models predict within the sample from which they are obtained. The predicted frequencies of "failure times" in the sample correspond very closely to the actual frequencies that are reported in Table 5. The models fail to predict well only for the long durations (for which the actual number of observations are quite small).

20. As Table 5 indicates, about 99 percent of all dropouts and 89 to 97 percent of all completers in the sample did so within eight years.

21. In the former simulation, six students are censored (still enrolled), while in the latter 34 are censored.

22. As Table 5 indicates, between 9 to 18 percent of completers, by field, have durations of greater than six years.

23. In the former simulation, two students are censored, while in the latter 19 are censored.

24. In the former simulation, no students are censored, while in the latter two are censored.

25. One can also use the estimates presented in Tables 6 and 7 to simulate the effects of changes in new doctorate labor market conditions on times-to-degree and completion rates. However, since the labor market variables never proved significant in the tables, we do not present such simulations here.

26. We intend in subsequent research to pursue similar analyses for two engineering fields at University X.

27. Concern over projected shortages typically relates to projections of shortages of American citizen and permanent resident doctorates. The estimates reported in the text are for all graduate students. However, when one restricts the analyses to graduate students who are American citizens and permanent residents, one obtains very similar results (for example, contrast the estimates in Table 6 with those in Table A1).

Table 1

Median and Mean Registered Time to Degree:
Data Grouped by Year of Degree

	Median Nationwide ^a			Median (Mean) University X ^c				
	Economics	English	Physics ^b	Mathematics	Economics	English	Physics	Mathematics
1970	5.2	6.0	5.8	5.2	5 (4.9)	4 (4.6)	6 (5.5)	4 (4.3)
1971	5.3	6.0	6.0	5.4	5 (5.1)	4 (4.5)	6 (5.6)	5 (4.9)
1972	5.5	5.9	6.0	5.5	5 (4.9)	4 (4.6)	6 (5.6)	5 (5.2)
1973	5.5	6.3	6.1	5.6	5 (5.0)	4 (4.8)	6 (5.8)	5 (5.2)
1974	5.4	6.4	6.1	5.5	5 (5.1)	4 (5.4)	6 (6.1)	5 (4.7)
1975	5.3	6.4	6.2	5.5	5 (5.4)	5 (5.7)	6 (6.0)	5 (4.6)
1976	5.5	6.7	6.1	5.5	5 (5.3)	5 (5.8)	6 (5.8)	5 (4.8)
1977	5.7	7.0	6.2	5.8	5 (5.2)	5 (6.0)	6 (5.7)	5 (4.8)
1978	5.7	7.4	6.2	5.8	5 (5.2)	5 (6.5)	6 (5.8)	5 (4.8)
1979	5.9	7.5	6.1	5.9	5.5 (5.5)	5 (6.7)	6 (6.0)	5 (5.0)
1980	6.0	7.8	6.3	5.9	5 (5.7)	6 (6.4)	6 (6.2)	5 (5.3)
1981	6.1	7.7	6.2	5.9	5.5 (5.9)	6 (6.7)	6 (6.4)	5 (6.1)
1982	6.1	8.1	6.4	5.9	6 (6.1)	6 (6.7)	6 (6.6)	5 (6.0)
1983	6.3	7.8	6.4	6.1	6 (6.5)	6 (6.5)	6 (6.4)	5 (5.9)
1984	6.3	8.1	6.5	6.1	6 (6.4)	6 (6.3)	6 (6.2)	5 (5.1)
1985	6.5	8.3	6.5	6.3	5.5 (6.3)	6 (6.5)	6 (6.0)	5 (5.0)
1986	6.3	8.2	6.3	6.0	6 (7.0)	7 (7.0)	6 (6.0)	5 (5.2)
1987	6.5	8.5	6.3	6.3	6 (6.7)	7 (7.1)	6 (6.2)	5 (5.9)
1988	6.4	8.1	6.3	6.3	6 (7.1)	7 (7.0)	6 (6.1)	5 (5.7)

^aNationwide data are from National Research Council, Doctorate Recipients From United States Universities: Summary Report (Washington, DC: National Academy Press, various years), Appendix A, Table 2.

^bPhysics and Astronomy nationwide, Physics only in University X.

^cAuthors' calculations from data provided by the Graduate School of University X. While the nationwide calculations measure time to degree in calendar time, the calculations for University X assign a duration of n years to an individual if he or she entered graduate school in September of year t and received a degree anytime between May of year $t+n$ and January of year $t+n+1$. The calculations use three year moving averages to smooth fluctuations caused by small sample sizes.

Table 2

Median and Mean Registered Time to Degree and Completion Rates
Data Grouped by Year of Entrance, for University X
(3 Year Moving Averages)

	Economics		English		Physics		Mathematics	
	Median	(Mean) [CR]	Median	(Mean) [CR]	Median	(Mean) [CR]	Median	(Mean) [CR]
1964	5	(5.5) [.65]	5	(5.1) [.67]	6	(5.7) [.71]	5	(5.1) [.66]
1965	5	(5.4) [.81]	4.5	(5.1) [.66]	6	(5.7) [.70]	4.5	(4.7) [.63]
1966	5	(5.5) [.83]	4	(5.1) [.62]	6	(5.6) [.65]	5	(4.7) [.60]
1967	5	(5.6) [.77]	4	(4.7) [.61]	6	(5.5) [.59]	4	(4.4) [.58]
1968	5	(5.3) [.59]	4	(4.4) [.62]	6	(5.5) [.58]	5	(4.7) [.55]
1969	4	(4.6) [.56]	4	(4.5) [.63]	6	(5.7) [.60]	5	(4.6) [.57]
1970	4	(4.8) [.63]	4	(4.6) [.62]	6	(6.2) [.63]	5	(5.0) [.63]
1971	4	(4.8) [.69]	5	(5.2) [.57]	6	(6.1) [.63]	5	(5.1) [.69]
1972	5	(5.3) [.62]	5	(5.1) [.59]	6	(6.1) [.65]	5	(5.1) [.69]
1973	5	(5.4) [.60]	5	(5.6) [.57]	6	(5.8) [.73]	5	(4.9) [.60]
1974	6	(6.0) [.56]	6	(6.1) [.53]	6	(6.1) [.71]	5	(4.7) [.59]
1975	5	(5.7) [.58]	7	(6.7) [.59]	6	(6.0) [.71]	5	(4.8) [.60]
1976	6	(5.8) [.53]	7	(6.6) [.67]	6	(6.0) [.63]	5	(5.0) [.62]
1977	5	(5.7) [.52]	7	(6.6) [.78]	6	(5.9) [.67]	5	(5.0) [.56]
1978	6	(6.3) [.55]	6	(6.4) [.73]	6	(6.1) [.65]	5	(5.0) [.54]

Source: Authors' calculations from data provided by the Graduate School of University X.

Where CR = completion rate.

Table 3

Mean Graduate Record Examination Scores and Proportion of
 Entering Classes That are U.S. Citizens or Permanent
 Residents, by Year of Entrance and Field for University X
 (3 Year Moving Averages)

Year	Economics			English			Physics			Mathematics		
	GREV	GREM	CIT	GREV	GREM	CIT	GREV	GREM	CIT	GREV	GREM	CIT
1963	-	-	.71	703	574	.92	658	752	.79	-	-	.75
1964	-	-	.76	716	596	.94	661	749	.83	-	-	.71
1965	661	681	.72	732	613	.96	662	753	.86	640	744	.71
1966	628	672	.69	727	610	.97	643	750	.91	652	755	.74
1967	643	678	.66	724	604	.95	639	750	.91	652	755	.75
1968	630	675	.66	717	597	.88	634	746	.90	661	766	.68
1969	630	670	.52	715	601	.84	643	749	.90	661	765	.64
1970	618	669	.51	720	605	.85	659	760	.88	675	780	.67
1971	617	675	.54	713	620	.92	653	761	.85	679	782	.71
1972	617	686	.65	713	615	.95	645	760	.83	648	778	.78
1973	581	684	.53	704	610	.95	617	754	.77	620	761	.80
1974	550	667	.44	705	600	.92	629	765	.77	626	764	.77
1975	561	668	.51	716	609	.93	631	772	.75	631	776	.68
1976	577	680	.65	716	590	.98	651	769	.83	645	779	.51
1977	610	693	.69	731	572	.96	637	764	.80	621	777	.50
1978	587	707	.64	735	576	.94	639	765	.83	578	767	.43
1979	598	709	.61	741	576	.91	632	765	.82	555	772	.54
1980	541	713	.55	738	593	.89	638	767	.78	567	767	.56
1981	523	693	.54	741	590	.78	645	764	.78	612	768	.65
1982	514	693	.52	729	600	.76	651	764	.70	607	755	.71
1983	547	701	.53	716	594	.74	659	760	.76	571	742	.76
1984	557	721	.39	714	599	.84	654	764	.71	593	741	.78
1985	567	726	.34	714	608	.90	652	767	.74	601	754	.73
1986	570	723	.32	722	624	.97	649	772	.71	635	764	.75
1987	611	721	.38	707	610	.99	649	771	.78	620	774	.80

Where GREV = mean GRE verbal scores
 GREM = mean GRE mathematics scores
 CIT = proportion of entrants who are citizens or permanent residents

Table 4
Distribution of University X's Graduate Students
by Major Source of Support

Academic Years	FL1	FL2	FL3	FL4	TA1	TA2	TA3	TA4	RA1	RA2	RA3	RA4
A) Economics												
65-67	.61	.56	.35	.35	.33	.39	.47	.47	.00	.00	.06	.00
68-70	.71	.60	.67	.50	.24	.33	.21	.25	.00	.00	.03	.00
71-73	.55	.49	.35	.39	.35	.49	.59	.39	.00	.00	.00	.00
74-76	.47	.23	.28	.13	.38	.74	.65	.52	.00	.00	.07	.09
77-79	.40	.27	.18	.26	.45	.54	.55	.54	.00	.10	.11	.06
80-82	.43	.24	.09	.28	.41	.63	.80	.55	.00	.02	.00	.00
83-85	.35	.23	.26	.10	.28	.52	.59	.70	.00	.02	.03	.10
86-88	.42	.36	.17	.09	.47	.57	.58	.45	.00	.00	.00	.09
B) English												
65-67	.69	.49	.57	.60	.23	.41	.31	.13	.00	.00	.00	.00
68-70	.87	.67	.75	.72	.09	.30	.21	.17	.00	.00	.00	.00
71-73	.76	.56	.43	.29	.14	.38	.47	.55	.00	.02	.00	.00
74-76	.67	.14	.17	.32	.13	.83	.63	.61	.00	.00	.00	.00
77-79	.66	.10	.12	.21	.05	.82	.82	.68	.00	.00	.00	.00
80-82	.72	.11	.09	.24	.19	.89	.85	.61	.00	.00	.00	.03
83-85	.84	.22	.14	.26	.14	.71	.76	.63	.00	.02	.00	.00
86-88	.80	.27	.43	.31	.13	.67	.57	.62	.00	.00	.00	.00
C) Mathematics												
65-67	.46	.37	.41	.42	.49	.49	.36	.31	.04	.06	.16	.21
68-70	.26	.37	.34	.09	.69	.59	.50	.63	.00	.00	.05	.13
71-73	.19	.17	.18	.07	.76	.69	.62	.68	.02	.03	.06	.07
74-76	.14	.23	.17	.11	.86	.71	.77	.82	.00	.00	.00	.00
77-79	.09	.14	.17	.10	.91	.78	.83	.65	.00	.00	.00	.20
80-82	.14	.15	.07	.23	.84	.76	.75	.55	.00	.00	.07	.18
83-85	.11	.06	.06	.19	.89	.85	.88	.55	.00	.00	.06	.19
86-88	.23	.00	.00	.00	.77	1.00	.91	.00	.00	.00	.00	.00
D) Physics												
65-67	.38	.30	.34	.29	.59	.61	.34	.16	.01	.05	.30	.51
68-70	.29	.18	.11	.13	.69	.76	.55	.23	.01	.07	.33	.59
71-73	.19	.19	.21	.25	.79	.66	.28	.09	.02	.15	.47	.61
74-76	.06	.09	.11	.09	.87	.66	.25	.07	.06	.25	.59	.76
77-79	.18	.20	.17	.06	.75	.55	.18	.04	.04	.25	.65	.84
80-82	.22	.23	.18	.10	.78	.54	.16	.11	.00	.24	.66	.79
83-85	.26	.23	.20	.12	.73	.52	.15	.05	.01	.24	.64	.77
86-88	.32	.35	.30	.32	.64	.38	.04	.05	.04	.27	.61	.45

where FLj - proportion of students in their jth year of study with fellowships as their major source of support that year

TAj - proportion of students in their jth year of study with teaching assistantships as their major source of support that year

RAj - proportion of students in their jth year of study with research assistantships as their major source of support that year

Source: Authors' calculations from data provided by the Graduate School of University X. FLj, TAj, and RAj may sum to less than unity in each year because of the presence of students whose major sources of support are loans, tuition waivers, and other means.

Table 5

Frequencies of "Failure" Times in the Sample

	Duration	Economics	English	Physics	Mathematics
COMPLETERS	3	12 (.08)	19 (.09)	7 (.02)	15 (.09)
	4	41 (.35)	50 (.34)	33 (.10)	43 (.34)
	5	31 (.55)	49 (.58)	122 (.41)	74 (.79)
	6	27 (.73)	29 (.72)	123 (.72)	21 (.91)
	7	17 (.84)	24 (.84)	56 (.86)	12 (.93)
	8	8 (.89)	17 (.92)	32 (.95)	4 (.97)
	9	7 (.94)	10 (.97)	9 (.97)	1 (.98)
	10	3 (.96)	1 (.98)	6 (.98)	1 (.99)
	>10	6 (1.00)	4 (1.00)	6 (1.00)	2 (1.00)
		152 [.48]	203 [.56]	394 [.59]	173 [.54]
DROPOUTS	1	42 (.30)	25 (.21)	43 (.24)	31 (.27)
	2	33 (.54)	29 (.45)	56 (.56)	32 (.55)
	3	24 (.71)	18 (.60)	39 (.77)	17 (.70)
	4	16 (.82)	23 (.79)	17 (.87)	17 (.84)
	5	12 (.91)	11 (.88)	8 (.92)	8 (.90)
	6	9 (.97)	7 (.94)	8 (.96)	6 (.95)
	7	3 (.99)	4 (.97)	6 (.99)	3 (.98)
	8	1 (1.00)	1 (.98)		1 (.99)
	9				
	10		1 (.99)	1 (1.00)	1 (1.00)
>10		1 (1.00)			
	140 [.44]	120 [.33]	178 [.26]	116 [.36]	
CENSORED	3			1	
	4	11	13	23	10
	5	11	11	34	13
	6	3	11	18	4
	7	1	3	16	2
	8		3	6	
	>8			3	1
		26 [.08]	41 [.11]	101 [.15]	30 [.09]
Total	318	364	673	319	

where () cumulative share of the category's observations for the field
[] share of the category in total observations for the field

Source: Authors' calculations from data provided by the Graduate School of University X.

Table 6
Estimated Duration Models

	Economics			English			Physics			Mathematics		
	(C)	(D)	(E)	(C)	(D)	(E)	(C)	(D)	(E)	(C)	(D)	(E)
$\lambda_0(1)$.000	.517 (1.150)**	.000	.428 (1.127)**	.000	.495 (1.121)**	.000	.463 (1.123)**	.000	.463 (1.123)**	.000	.463 (1.123)**
$\lambda_0(2)$.000	.483 (1.411)*	.000	.566 (1.150)**	.000	.767 (1.144)**	.000	.767 (1.144)**	.000	.767 (1.144)**	.000	.767 (1.144)**
$\lambda_0(3)$.089 (0.039)*	.409 (1.126)*	.098 (0.030)*	.401 (1.124)*	.022 (0.008)*	.701 (1.144)**	.061 (0.019)*	.399 (1.104)*	.061 (0.019)*	.399 (1.104)*	.061 (0.019)*	.399 (1.104)*
$\lambda_0(4)$.428 (1.133)*	.339 (1.112)*	.408 (0.089)*	.685 (1.171)**	.112 (0.022)*	.385 (1.119)**	.242 (0.052)*	.399 (1.135)*	.242 (0.052)*	.399 (1.135)*	.242 (0.052)*	.399 (1.135)*
$\lambda_0(5)$.528 (1.161)*	.391 (1.137)**	.833 (1.130)**	.547 (1.088)*	.526 (0.068)*	.255 (1.101)**	.893 (0.060)*	.360 (1.145)*	.893 (0.060)*	.360 (1.145)*	.893 (0.060)*	.360 (1.145)*
$\lambda_0(6)$.650 (1.114)*	.545 (1.196)**	.892 (1.114)*	.570 (1.221)**	.958 (0.202)**	.444 (1.192)**	.639 (1.169)**	.619 (1.216)*	.639 (1.169)**	.619 (1.216)*	.639 (1.169)**	.619 (1.216)*
CON	-2.359 (2.183)	-2.385 (2.853)*	3.542 (2.646)	1.649 (1.941)**	1.018 (1.113)	-2.848 (1.118)*	4.766 (3.009)	-1.277 (1.254)	4.766 (3.009)	-1.277 (1.254)	4.766 (3.009)	-1.277 (1.254)
YR	.006 (0.029)	.046 (0.016)**	-.100 (0.029)**	-.030 (0.017)**	-.084 (0.029)**	-.036 (0.013)**	-.068 (0.036)**	-.014 (0.016)	-.068 (0.036)**	-.014 (0.016)	-.068 (0.036)**	-.014 (0.016)
MA	1.98 (2.110)	-.478 (2.429)**	.289 (2.03)	-.699 (2.933)**	.585 (1.196)*	-.861 (3.16)*	.499 (2.233)*	-.701 (3.34)*	.499 (2.233)*	-.701 (3.34)*	.499 (2.233)*	-.701 (3.34)*
SEX	.400 (2.096)	-.205 (2.09)	.367 (2.055)**	.140 (2.113)	.410 (2.295)	-.629 (3.260)**	.239 (2.271)	-.021 (2.290)	.239 (2.271)	-.021 (2.290)	.239 (2.271)	-.021 (2.290)
CTZN	-.337 (2.229)	.513 (2.185)*	-1.008 (3.17)*	-.767 (3.116)**	-.328 (1.175)**	.994 (3.37)*	-.332 (1.198)**	.236 (2.230)	-.332 (1.198)**	.236 (2.230)	-.332 (1.198)**	.236 (2.230)
GREM	1.130 (1.088)	-.101 (1.081)	-.259 (1.132)*	-.198 (1.135)	-.061 (0.066)	-.170 (0.093)**	-.088 (0.097)	-.108 (1.110)	-.088 (0.097)	-.108 (1.110)	-.088 (0.097)	-.108 (1.110)
GREM	1.659 (1.110)	-.100 (1.053)	.007 (0.088)	-.132 (0.099)	-.024 (1.136)	.193 (1.160)	-.162 (1.213)	-.012 (1.184)	-.162 (1.213)	-.012 (1.184)	-.162 (1.213)	-.012 (1.184)
PTA	-1.144 (3.78)*	.274 (2.42)	-.696 (4.03)**	-.209 (3.51)	-.971 (2.69)*	.784 (2.42)*	-.121 (2.293)**	.657 (3.305)*	-.121 (2.293)**	.657 (3.305)*	-.121 (2.293)**	.657 (3.305)*
POA	-.374 (1.023)	-1.279 (1.166)	-.292 (2.316)	-.251 (1.445)	-.131 (1.270)	-.447 (1.440)	-.266 (1.515)	-6.601 (1.924)**	-.266 (1.515)	-6.601 (1.924)**	-.266 (1.515)	-6.601 (1.924)**
NRS	-.705 (3.469)		-1.884 (3.582)*			1.464 (3.859)**		-1.611 (3.584)*		-1.611 (3.584)*		-1.611 (3.584)*
PEED	4.435 (3.970)		-.145 (2.737)			-.324 (1.146)		2.347 (3.108)		2.347 (3.108)		2.347 (3.108)
DUM						1.274 (1.674)		-1.964 (1.571)		-1.964 (1.571)		-1.964 (1.571)
Ln L		-608.311		-713.073		-1199.26		-633.037		-633.037		-633.037
N		318		366		673		319		319		319

where $\lambda_0(t)$ is the baseline hazard in time period t , (C) completion hazards, (D) drop out hazards
 *indicates significant parameter
 ** indicates significant parameter
 (**) significantly different from zero at the .05 (1.10) level of significance

and
 YR time trend variable, = 1 if year of entrance is 1962, 2 if year of entrance is 1963, etc.
 MA = student had a master's degree at time of first registration, 0 = otherwise
 1 = male, 0 = female
 CTZN permanent resident, 0 = other
 GREM verbal aptitude GRE score/100, equals 0 if missing for mathematics students
 GREM mathematics aptitude GRE score/100, equals 0 if missing for mathematics students
 PTA proportion of student's registered time to date spent as a teaching assistant
 POA proportion of student's registered time to date financed by tuition waivers or student's own personal means (spent as a fellowship holder)
 NRS proportion of student's registered time to date financed by research assistant/committed category is proportion
 SIRY(t) mean starting salary for new assistant professors in the field (if not reported for Physics equals zero) in the current academic year deflated
 by the consumer price index (in ten thousands)
 0 = mean starting salary reported
 NRS(t) percentage of new graduates in the field seeking employment in academia in the current academic year
 PEED(t) percentage of new graduates for mathematics students, 0 = GRE scores reported
 DUM 1 = GRE scores not reported for mathematics students, 0 = GRE scores reported
 (t) these variables are time-varying covariate in the estimation of the duration model

Data Sources:
 (1) Author's calculations from data provided by the graduate school of University X (YR, MA, SEX, CTZN, GREM, PTA, PRA, POA).
 (2) National Research Council, Doctorate Recipients From United States Universities: Summary Report (various years) (Washington, DC: National
 American Institute of Physics, Graduate Student Survey (various years), American Economic Association, Annual Salary Survey (data prior to 1965
 provided by David Stapleton at Dartmouth), AHC-MAA Survey, Notices of the American Mathematical Society (various issues) - (SIRY).

Table 7
Estimated Duration Models: Eight-Year Truncation

	Economics			English			Physics			Mathematics		
	(C)	(D)	(E)	(C)	(D)	(E)	(C)	(D)	(E)	(C)	(D)	(E)
$\lambda_1(1)$.000	.583 (.147)*	.000	.000	.434 (.131)*	.000	.000	.490 (.162)*	.000	.000	.376 (.124)*	.000
$\lambda_1(2)$.000	.544 (.145)*	.000	.000	.575 (.154)*	.000	.000	.755 (.235)*	.000	.000	.438 (.142)*	.000
$\lambda_1(3)$.010 (.010)	.459 (.132)*	.072 (.026)*	.072 (.026)*	.408 (.129)*	.023 (.009)*	.023 (.009)*	.685 (.225)*	.063 (.021)*	.063 (.021)*	.281 (.103)*	.063 (.021)*
$\lambda_1(4)$	-.047 (.050)	.381 (.125)*	.292 (.083)*	.292 (.083)*	.695 (.180)*	.112 (.026)*	.112 (.026)*	.374 (.145)*	.250 (.058)*	.250 (.058)*	.376 (.135)*	.250 (.058)*
$\lambda_1(5)$.057 (.061)	.441 (.156)*	.570 (.148)*	.570 (.148)*	.556 (.189)*	.513 (.091)*	.513 (.091)*	.247 (.112)*	.932 (.125)*	.932 (.125)*	.338 (.147)*	.932 (.125)*
$\lambda_1(6)$.086 (.091)	.616 (.232)*	.601 (.164)*	.601 (.164)*	.581 (.227)*	.908 (.134)*	.908 (.134)*	.430 (.204)*	.673 (.169)*	.673 (.169)*	.584 (.230)*	.673 (.169)*
$\lambda_1(7)$.999 (.003)*	.408 (.240)**	.995 (.037)*	.995 (.037)*	.587 (.264)*	.996 (.017)*	.996 (.017)*	.540 (.278)**	.930 (.143)*	.930 (.143)*	.664 (.336)*	.930 (.143)*
$\lambda_1(8)$.000 (.001)	.000	.000 (.001)	.000 (.001)	.000	.000 (.001)	.000 (.001)	.000	.001 (.001)	.001 (.001)	.000	.001 (.001)
CON	.175 (1.429)	-2.347 (.765)*	3.637 (2.302)	3.637 (2.302)	1.534 (.943)	1.556 (1.370)	1.556 (1.370)	-2.765 (1.191)*	4.508 (2.892)	4.508 (2.892)	-.690 (1.655)	4.508 (2.892)
YR	-.013 (.019)	.051 (.016)*	-.090 (.025)*	-.090 (.025)*	-.026 (.017)	-.077 (.026)*	-.077 (.026)*	-.038 (.012)*	-.068 (.034)*	-.068 (.034)*	-.012 (.015)	-.068 (.034)*
MA	.213 (1.198)	-.413 (.237)**	.346 (.182)**	.346 (.182)**	-.667 (.285)*	.462 (.193)*	.462 (.193)*	-.872 (.511)**	.402 (.217)**	.402 (.217)**	-.728 (.353)*	.402 (.217)**
SEX	.418 (.305)	-.209 (.200)	.227 (.187)	.227 (.187)	.161 (.210)	.390 (.251)	.390 (.251)	-.580 (.263)*	.141 (.254)	.141 (.254)	.075 (.288)	.141 (.254)
CTZN	-.334 (.217)	.497 (.213)*	-.972 (.326)*	-.972 (.326)*	-.825 (.299)*	-.380 (.161)*	-.380 (.161)*	.995 (.334)*	-.255 (.191)	-.255 (.191)	.263 (.230)	-.255 (.191)
GREV	.128 (.082)	-.088 (.079)	-.272 (.125)*	-.272 (.125)*	-.182 (.134)	-.017 (.059)	-.017 (.059)	-.159 (.090)**	-.144 (.095)	-.144 (.095)	-.120 (.108)	-.144 (.095)
GREH	.106 (.103)	.056 (.098)	.084 (.080)	.084 (.080)	-.153 (.940)	-.040 (.119)	-.040 (.119)	.173 (.160)	-.112 (.212)	-.112 (.212)	-.058 (.235)	-.112 (.212)
PTA	-1.022 (.351)*	.289 (.245)	-.619 (.382)	-.619 (.382)	-.216 (.347)	-.866 (.258)*	-.866 (.258)*	.763 (.238)*	-1.223 (.296)*	-1.223 (.296)*	.647 (.308)*	-1.223 (.296)*
PPA	.287 (.999)	-1.416 (1.077)				-.015 (.262)	-.015 (.262)	-.398 (.426)	-1.741 (.564)*	-1.741 (.564)*	-5.970 (2.008)*	-1.741 (.564)*
POA	-.380 (.561)	.262 (.322)	-1.830 (.538)*	-1.830 (.538)*	-.336 (.448)	-1.294 (.669)*	-1.294 (.669)*	1.254 (.748)**	-1.741 (.564)*	-1.741 (.564)*	1.057 (.520)*	-1.741 (.564)*
SLRY	-.694 (.464)					-.140 (.472)	-.140 (.472)					
NRS						-.870 (1.038)	-.870 (1.038)					
PEED	.009 (.004)*		-.409 (2.350)	-.409 (2.350)		.966 (1.491)	.966 (1.491)					
DUM												
Ln L	-705.738		-775.934	-775.934	-1314.438	-1314.438	-1314.438		2.556 (2.884)	2.556 (2.884)	-1.129 (1.684)	2.556 (2.884)
n	318		364	364	673	673	673		-1.841 (1.596)	-1.841 (1.596)		-1.841 (1.596)

See table 5 for variable definitions.

Table 8

Policy Simulations: Completers and Dropouts After Six Years
Under Various Financial Assistance Scenarios for an
Initial Entering Class of 100 Students

	All Fellowship	All Teaching Asst.	All Research Asst.	All Other Support
<u>Economics</u>				
Completers				
Number	63	27	76	42
Mean Duration	4.54	4.89	4.76	4.83
Dropouts				
Number	37	52	12	54
Mean Duration	2.62	2.85	2.83	2.85
Truncated (Still Enrolled)				
Number	0	21	12	14

<u>English</u>				
Completers				
Number	63	47	-	18
Mean Duration	4.59	4.89	-	5
Dropouts				
Number	33	30	-	32
Mean Duration	2.91	3.23	-	3.38
Truncated (Still Enrolled)				
Number	4	23	-	50

<u>Physics</u>				
Completers				
Number	78	32	77	11
Mean Duration	5.28	5.38	5.30	5.45
Dropouts				
Number	16	34	11	59
Mean Duration	2.75	2.79	2.91	2.68
Truncated (Still Enrolled)				
Number	6	34	12	30

<u>Mathematics</u>				
Completers				
Number	80	46	100	25
Mean Duration	4.33	4.85	4.82	4.84
Dropouts				
Number	20	45	0	66
Mean Duration	2.30	2.84	-	2.80
Truncated (Still Enrolled)				
Number	0	9	0	9

Source: Authors' calculations based upon estimates reported in Table 6 and individual observations for University X.

Table 9

Policy Simulations: Completers and Dropouts After Eight Years
Under Various Financial Assistance Scenarios for an
Initial Entering Class of 100 Students

	All Fellowship	All Teaching Asst.	All Research Asst.	All Other Support
<u>Economics</u>				
Completers				
Number	48	15	87	28
Mean Duration	6.52	6.60	6.62	6.57
Dropouts				
Number	48	59	13	58
Mean Duration	3.44	3.22	3.31	3.29
Truncated (Still Enrolled)				
Number	4	23	0	14

<u>English</u>				
Completers				
Number	67	59	-	26
Mean Duration	4.75	5.29	-	5.46
Dropouts				
Number	33	34	-	34
Mean Duration	2.91	3.28	-	3.74
Truncated (Still Enrolled)				
Number	0	7	-	40

<u>Physics</u>				
Completers				
Number	82	44	87	24
Mean Duration	5.46	5.80	5.48	6.00
Dropouts				
Number	16	37	11	55
Mean Duration	2.75	3.14	2.91	3.24
Truncated (Still Enrolled)				
Number	2	19	2	21

<u>Mathematics</u>				
Completers				
Number	80	52	100	29
Mean Duration	4.29	5.02	4.51	5.24
Dropouts				
Number	20	46	0	67
Mean Duration	2.30	3.04	-	3.09
Truncated (Still Enrolled)				
Number	0	2	0	4

Source: Authors' calculations based upon estimates reported in Table 7 and individual observations for University X.

Table 10

Policy Simulations: Completers and Dropouts After Six Years
for an Initial Entering Class of 100 Mathematics Graduate
Students if All Students Are Awarded Fellowships in Their
jth Year of Study and Teaching Assistantships in All Other Years

	First Year	Second Year	Third Year	Fourth Year	Fifth Year	Sixth Year
Completers						
Number	60	58	55	54	52	48
Mean Duration	4.70	4.74	4.73	4.76	4.85	4.90
Dropouts						
Number	35	37	40	41	43	45
Mean Duration	3.06	2.76	2.65	2.66	2.72	2.84
Truncated (Still Enrolled)						
Number	5	5	5	5	5	7

Source: Authors' calculations based upon the estimates reported in Table 6 and individual observations for University X.

Table A1
Estimated Duration Models: Sample Restricted to
U.S. Citizens and Permanent Residents

	Economics		English		Physics		Mathematics	
	(C)	(D)	(C)	(D)	(C)	(D)	(C)	(D)
$\lambda_0(1)$.000	.587 (.165)**	.000	.432 (.133)*	.000	.393 (.113)*	.000	.335 (.119)*
$\lambda_0(2)$.000	.405 (.135)**	.000	.510 (.015)*	.000	.621 (.162)**	.000	.367 (.132)*
$\lambda_0(3)$.067 (.036)**	.277 (.105)*	.101 (.033)*	.402 (.127)*	.015 (.007)*	.547 (.146)**	.031 (.013)*	.329 (.123)*
$\lambda_0(4)$.462 (.160)*	.371 (.135)**	.387 (.087)*	.678 (.176)*	.094 (.021)*	.334 (.108)*	.243 (.061)*	.388 (.147)*
$\lambda_0(5)$.596 (.195)*	.537 (.182)**	.812 (.122)*	.450 (.169)*	.507 (.072)*	.277 (.106)*	.942 (.039)*	.427 (.189)*
$\lambda_0(6)$.826 (.142)**	.607 (.225)*	.857 (.129)*	.565 (.244)*	.965 (.017)*	.384 (.156)*	.548 (.186)*	.687 (.259)*
CON	-4.921 (3.356)	-3.329 (1.046)*	2.472 (2.646)	.615 (.978)	1.111 (1.100)	-2.268 (1.178)**	3.307 (3.660)	-704 (2.072)
YR	-.011 (.040)	-.060 (.030)**	-1.03 (.030)**	-.036 (.018)*	-.051 (.030)**	-.036 (.013)*	-.069 (.047)	-.017 (.018)
MA	-.227 (.333)	-.557 (.330)**	.200 (.213)	-.818 (.332)*	.648 (.255)*	-.790 (.566)	-.691 (.320)*	-.786 (.470)**
SEX	-.292 (.352)	-.029 (.236)	.343 (.211)	.135 (.225)	.512 (.367)	-.648 (.280)*	-.035 (.316)	-.498 (.400)
GREY	-.027 (.137)	-.045 (.111)	-.184 (.139)	-.146 (.151)	-.000 (.079)	-.186 (.101)**	-.136 (.146)	.080 (.148)
GREM	.203 (.156)	.109 (.123)	-.011 (.091)	-.164 (.105)	-.175 (.139)	.297 (.173)**	.099 (.220)	-.241 (.283)
PTA	-.811 (.467)**	.482 (.288)**	-.638 (.419)	.001 (.369)	-1.231 (.305)*	.705 (.245)*	-1.193 (.383)*	.674 (.348)*
FPA	-.929 (1.137)	.126 (1.437)			-.018 (.283)			-9.613(4.432)*
FOA	-.923 (.861)	.377 (.423)			-1.593 (.686)**	1.246 (.917)	-2.174 (.791)*	1.216 (.556)*
SLRY	.036 (.782)				-.031 (.428)		-.644 (.578)	
NRS					-.496 (.943)			
FRED	5.031 (5.432)		-.530 (2.794)		2.810 (1.759)		-.043 (3.808)	
DUM							-.393 (1.508)	-1.010 (2.078)
Ln L								
n	-409,770	-659,182	-1049,50				-431,806	
	191	337	578				219	

See Table 6 for variable definitions.

Appendix¹

I. The Structural Model

In what follows we concentrate on individuals enrolled in a graduate program, their effort levels each period and their decisions whether to complete the program or to quit before receiving their Ph.D. degrees. For the purposes of this study, we assume that the process terminates when one quits or receives his or her Ph.D. degree.

Time is taken to be discrete. We denote the time interval by $t=1,2,\dots$. At each time interval an individual can be characterized as (i) a graduate student, denoted by g , (ii) a drop out, that is, having quit without receiving a degree, denoted by q , and (iii) a Ph.D., that is, having successfully completed one's studies, denoted by p .

Let $Z(t)$ be a binary indicator denoting the status of an individual at time period t . We let $Z(t)=0$ if the person is a graduate student, and $Z(t)=1$ if not. Notice that if $Z(r)=1$ for some r , then $Z(t)=1$ for all $t \geq r$. That is, people who have dropped out of the program are assumed never to reenter.

We start by assuming that each individual seeks to maximize the expected present value of income, discounted to the present; at some rate $\beta \in (0,1)$, over an infinite horizon (A1).

While in the program the student has to choose a level of effort each period denoted by e_t . We assumed that there is a minimum level of effort, e , required each period for the person to remain in the program. Furthermore, we assume that to successfully complete one's studies and receive a Ph.D. degree, a cumulative effort level of at least \bar{E} is required. We also assume that the effort level at each time period lies in the interval $[0, E_n]$ where $2E_n < \bar{E}$. The last assumption assures that no individual finishes a Ph.D. in less than two years (which is the minimum length of degree program observed in our data).

Effort is assumed to be costly and the cost function associated with effort at time t is denoted by $C(e_t)$. This function is assumed to be a twice continuously differentiable function of effort level with $C' \geq 0$, $C'' \geq 0$, and $C(0)=0$

¹This appendix was written by Panagiotis Mavros and will appear in his forthcoming Cornell Ph.D. dissertation.

(A2). Note that a person who leaves the program, either by dropping out or receiving the Ph.D., exerts no additional effort from that time and thus reduces his or her future cost to zero.

At the end of each period the student knows the level of cumulative effort $E = \sum e_t$, extended over the periods he or she was enrolled where $E \in [0, \hat{E}]$ with $\hat{E} < \infty$. Furthermore at the end of each period, each individual is informed as to whether he or she is eligible to receive a Ph.D. degree. In other words, the student is informed about the value of the binary indicator d defined to be $d=1$ if the student is eligible for the Ph.D. and $d=0$ if not.

In addition, at the end of each period, each individual still enrolled in the program receives a pair of offers, f and R . We let f denote the amount of financial aid the student receives if he or she decides to continue in the program for another year, while R denotes the constant income the individual would receive each period if the individual does not continue in the program, either by quitting or completing graduate studies. Consequently $R = dR_p + (1-d)R_q$, where R_p and R_q represent the offers received in case the Ph.D. degree is received, or the individual decides to drop out, respectively. It is assumed that $R \in [0, \hat{R}]$ where $\hat{R} < \infty$.

The pair of offers f and R along with the realization of d will be called an event and be denoted by $x = (f, R, d)$. With respect to the occurrence of such events we assume that an event occurs with certainty to those continuing in the program at the end of each time period, while once a person has left the program, no event occurs (A3).

Let s be defined as $s = (x, E)$. Knowledge of s would imply knowledge of an event as well as of the cumulative effort level up to that time. With respect to s we assume that each $s = (x, E)$ is an independent random draw from a known distribution $Q(s|s', a)$ that defines the transition probabilities to state s given that the current state is s' and that action a was taken (A4). We write

$$Q(s|s', a) = P(1_{\{E \in [E', E]\}}) \cdot G(x|s', a) = G(x|s', a)$$

where we assume that $G(x|s', a)$ has the compact support

$$X = \{(f, R, d) \in \mathbb{R}^3 : f \in [0, f_n], R \in [0, \hat{R}], d \in (0, 1)\}$$

for all s' and a and $Q(s|s', a)$ has the compact support

$$S = \{(x, E) \in X \times \mathbb{R} : x \in X, E \in [0, \hat{E}]\}$$

for all s' and a .

Notice that $P(1_{(E \in E', e)})$ is always equal to one since E' represents the current level of cumulative effort, e the choice of effort to be put forth in the next period and therefore at the end of the next period $E = E' + e$ with certainty.

With regard to the structure of $G(x|\cdot)$ the following will be assumed. The distribution $G(x)$ can be written as $G(x) = G(f, R) = H(f) \cdot J(R)$ where $H(\cdot)$ and $J(\cdot)$ are the marginal distributions of f and R respectively. Furthermore let $P(d=1)$ denote the probability that a Ph.D. degree is awarded, which depends on the cumulative effort level. Then $J(R) = P(d=1)J_p(R_p) + (1-P(d=1))J_q(R_q)$ where $J_p(\cdot)$ and $J_q(\cdot)$ are the distributions of R_p and R_q respectively.

The decision process of an individual can now be described. The individual at the end of each period (while in a graduate program) observes the state $s = (x, E)$ of an event $x = (f, R, d)$ and the cumulative effort level. Given this he or she chooses an action denoted by $a = (Z, e) \in (0, 1) \times [0, E_n]$ and receives a reward $r(s, a)$ until a new event occurs. Consequently we assume that the action space is given by (A5)

$$A = \{(Z, e) \in \mathbb{R}^2 : Z = 0 \text{ \& } e \in [0, E_n] \text{ or } Z = 1 \text{ \& } e = 0\}.$$

and that the reward rate $r(s, a)$ that one receives in state s when action a is taken, is given by (A6)

$$r(s, a) = r(x, a) = (1-Z)(f - C(e)) + Z(dR_p + (1-d)R_q).$$

A policy π is a sequence of functions, that is $\pi = (\pi_0, \pi_1, \dots, \pi_n, \dots)$, where π_n is a Borel measurable map on $(S \times A)^n \times S$ into A . A policy π is Markovian if for each n , π_n is a map on S into A . Furthermore a policy is stationary if it is

Markovian and if in addition $\pi_n = \pi \forall n$. Given some policy π and an initial state s , the total expected discounted reward is given by

$$I^\pi(s) = \sum_{k=0}^{\infty} \beta^k E_t^\pi r(s_k, a_k) \\ - \sum_{k=0}^{\infty} \beta^k E_x^\pi r(X_k, a_k).$$

where E_t^π and E_x^π denotes expectation under π and s or x respectively.

With the assumptions we have made it is straightforward to show that

Proposition 1:

(i) the optimal discounted reward is a continuous function on S satisfying

$$V(s) = \max_{a \in A} \{ r(s, a) + \beta \int V(y) dG(y|s, a) \}. \quad (*)$$

(ii) There exists a Borel measurable function π^* on S into A such that $a = \pi^*(s)$ maximizes for each s the right side of (*). The stationary policy following π^* is optimal.

Proof: This proposition is a direct application of theorem 3.3 in Bhattacharya and Majumdar (1989). One only needs to check whether the required conditions are satisfied in the present case. The conditions imposed there are

C1: The state space S is a non-empty Borel subset of a complete separable metric space. The action space A is a compact metric space.

$$U(s) = \sum_{k=0}^{\infty} \beta^k r_k(s) < \infty \quad (s \in S)$$

where

$$r_0(s) = \sup_{a \in A} |r(s, a)|$$

$$r_{k+1}(s) = \sup_{a \in A} \int r_k(y) dG(y|s, a) \quad (k=0, 1, 2, \dots).$$

C3: $(s, a) \rightarrow \int \phi(y) dG(y|s, a)$ is continuous on $S \times A$ for all Borel measurable ϕ satisfying $|\phi(s)| \leq U(s)+1$ for all $s \in S$.

Condition C1 holds under A4 and A5. Condition C2 holds under A4, and A6. Condition C3 obviously holds.

□

Let $\varphi(f, R, d)$ be defined as

$$\varphi(f, R, d) = \max_e V(s|Z=0, e) - V(s|Z=0, e')$$

where the maximization in the second expression takes places over $e \in [0, E_0]$, and e' is the argument that maximizes the expression. This is the total discounted reward if the current state is $s=(x, E)$ and the individual takes action $Z=0$ now, while future actions will be taken according to the optimal strategy. Define in addition to this the function $\omega(x)$ to be

$$\omega(f, R, d) = R/\beta - \varphi(f, R, d).$$

Then the optimal policy can be characterized.

Proposition 2: The optimal policy has the following form. There exist the function $\omega(f, R, d)$ on S , such that the Borel measurable function π' with $a=\pi'(s)=\pi'(x)$ in proposition 1 (ii), is defined as

$$a = (z, e) = \pi'(f, R, d) = \begin{cases} (0, e) & \text{if } \omega(f, R, d) \leq 0 \\ (1, 0) & \text{if } \omega(f, R, d) > 0 \end{cases}$$

Proof: Notice that R/β is the total discounted reward that the individual will receive if the current state is $s=(x, E)$ where $x=(f, R, d)$ and he chooses an action $Z=1$, and $e=0$. That is $V(s|Z=1, e=0) = R/\beta$.

Similarly we construct the total discounted reward if the current state is $s=(x, E)$ where $x=(f, R, d)$ and the individual takes an action $Z=0$ now, while future

actions will be taken according to the optimal strategy. That was defined earlier as $\varphi(f,R,d)$.

Obviously the individual's optimal choice given a state $s=(x,E)$ at each time period will be to remain in the program if $R/\beta \leq \varphi(f,R,d)$, and to withdraw otherwise. Once withdrawal has occurred the individual remains out of the program for ever.

□

It would be of interest to consider the sample path and the transition probabilities from one point of the sample path to another. As mentioned in the beginning of the appendix an individual at each time period can be characterized in three different ways denoted as g , q , and p . The transition probabilities from one point to another are defined in the following proposition.

Proposition 3: The transition probabilities $p_{ij}(t)$ from state i to state j where $i,j=g,q,p$ at time t are given by

$$\begin{array}{c}
 \\
 \\
 \\
 \end{array}
 \begin{array}{c}
 \\
 \\
 \\
 \end{array}
 \begin{array}{ccc}
 g & p & q \\
 \\
 \\
 \end{array}
 \begin{array}{c}
 \\
 \\
 \\
 \end{array}
 \begin{array}{c}
 \\
 \\
 \\
 \end{array}
 \left[\begin{array}{ccc}
 P_{gg}(t) & P_{gp}(t) & P_{gq}(t) \\
 0 & 1 & 0 \\
 0 & 0 & 1
 \end{array} \right]$$

$$\begin{array}{l}
 \text{where } P_{gg}(t) = [1-P(d=1)]\Pr(\omega(x) \leq 0) \\
 P_{gp}(t) = P(d=1) \\
 P_{gq}(t) = [1-P(d=1)]\Pr(\omega(x) > 0)
 \end{array}$$

where $\Pr(d=1)$ is the probability that a Ph.D. degree is awarded and $\Pr(\omega(x) \leq 0)$ the probability that the individual will continue in the program given that the Ph.D degree was not awarded.

Proof: It is apparent under the above mentioned assumptions.

□

Our goal is to estimate the transition probabilities from one state to another on the optimal sample path. Notice that the model does not impose any restrictions on the functional forms of these probabilities. It does suggest though that both the probability of receiving the Ph.D, and the probability of continuing in a program conditional on $d=0$, depend on the cumulative effort level, and through this, on the time spent in a program. Keeping this in mind we proceed to estimate the transition probabilities, by making some a priori assumptions about their functional form. In the following section we specify the model and the estimation procedure that we follow.

II. Specification and Estimation of the Competing Risks Model

From the analysis so far, it should be clear that an individual might be out of the program either because the individual dropped out before receiving the Ph.D degree, or because the individual received a Ph.D degree. As such, it is appropriate to utilize a competing risks model.

Let $j=1,2$ denote the two causes for terminating the process. What is recorded with respect to a spell is whether failure or censoring occurred in any of a set of disjoint intervals of the form $[a_{t-1}, a_t)$ where $t=1,2,\dots,K$, $a_0=0$ and $a_K=\infty$. The index $t=1,2,3,\dots,K$ is also used to denote time period. We assume that the cause specific hazards are of the proportional hazard form. That is,

$$\begin{aligned} \Pr(T_j \in [a_{t-1}, a_t) | T_j \geq a_{t-1}) &= \lambda^j(t, X_{j,t}, \beta_j) \\ &= \lambda_0^j(t) \phi(X_{j,t}, \beta_j) = \lambda_0^j(t) \exp[X_{j,t} \beta_j] \end{aligned}$$

where $X_{j,t}$ denotes the vector of covariates for the j^{th} risk at time period t . The vector of covariates will be assumed to be constant within each time interval. With respect to the time path of the hazard we assume that the hazard is constant within each time interval but is allowed to vary among time periods.

Given the above assumptions the integrated hazard (Λ^j) and the survivor function (S^j) at time period t are defined as

$$\begin{aligned} \Lambda^1(t, X_{j,t}, \beta_j) &= \int_0^{a_t} \lambda_0^1(u) \exp[X_{j,u} \beta_j] du \\ &= \sum_{i=1}^t \int_{a_{i-1}}^{a_i} \lambda_0^1(u) \exp[X_{j,u} \beta_j] du \\ &= \sum_{i=1}^t [a_i - a_{i-1}] \lambda_0^1(s) \exp[X_{j,s} \beta_j] \end{aligned}$$

$$\text{and } S^1(t, X_{j,t}, \beta_j) = \exp[-\Lambda^1(t, X_{j,t}, \beta_j)].$$

Using the above and introducing the transformation $\epsilon_j = \ln \Lambda^1(t, X_{j,t}, \beta_j)$ we can see that $\Pr(T_j \geq t) = \Pr(\epsilon_j \geq \ln \Lambda^1(t, X_{j,t}, \beta_j))$ where ϵ_j has a type I extreme value distribution.

In the competing risks framework, one can think of an observed duration as the minimum of a number of different possible failure times. That is, if one observes a duration of t , time periods where failure was due to the first cause then we think of it as being $t = \min(t_1, t_2)$. If an observation is censored at time period t then we only know that both durations are larger than $t-1$ periods.

Consider now the case where a failure of type one is observed in the interval $[a_{t-1}, a_t)$. This implies that the unobserved failure of type two is greater than the observed failure time for every point in the specified interval. The probability term that such an event contributes to the likelihood is the $\Pr(a_{t-1} \leq T_1 < a_t \text{ and } T_2 > T_1)$ which is given by

$$P_1^1(\theta) = \frac{\ln \Lambda^1(t, X_{1,t}, \beta_1)}{\ln \Lambda^1(t-1, X_{1,t-1}, \beta_1)} \int_0^\infty f(\epsilon_1, \epsilon_2) d\epsilon_2 d\epsilon_1,$$

where $f(\epsilon_1, \epsilon_2)$ is the joint distribution of the errors ϵ_1 and ϵ_2 , which may be specified to allow for correlation among the disturbances. The term θ denotes the vector of the unknown parameters involved in the expression. The limit of the second integration $g(\epsilon_1)$ is such that the implied failure of type two is

larger than that of type one. It is of interest then to derive the exact form of $g(\cdot)$.

In the framework of discrete hazards with constant covariates and constant hazards within time intervals we know that for any $\tau \in [a_{t-1}, a_t)$ the following relations are true.

$$\lambda_0(\tau) \exp(X_t \beta) = \Lambda(t, X_t, \beta) - \Lambda(t-1, X_{t-1}, \beta)$$

$$\text{and } \Lambda(\tau, X_t, \beta) = \Lambda(t-1, X_{t-1}, \beta) + [\Lambda(t, X_t, \beta) - \Lambda(t-1, X_{t-1}, \beta)] [\tau - a_{t-1}]$$

Let's now return to our problem with two risks. Suppose that failure of type 1 occurs at $\tau \in [a_{t-1}, a_t)$. Then using the above specifications we have

$$\ln \Lambda^1(\tau, X_{1,t}, \beta_1) = \epsilon_1^1$$

$$\Leftrightarrow \epsilon_1^1 = \ln[\Lambda^1(t-1, X_{1,t-1}, \beta_1) + \{\Lambda^1(t, X_{1,t}, \beta_1) - \Lambda^1(t-1, X_{1,t-1}, \beta_1)\} (\tau - a_{t-1})]$$

$$\Leftrightarrow \tau - a_{t-1} = [\exp(\epsilon_1^1) - \Lambda^1(t-1, X_{1,t-1}, \beta_1)] / [\Lambda^1(t, X_{1,t}, \beta_1) - \Lambda^1(t-1, X_{1,t-1}, \beta_1)]$$

For the second type we can write similar expressions, however it must be such that the implied failure time has to be bigger than τ . This is equivalent to saying that

$$\epsilon_2 \geq \ln[\Lambda^2(t-1, X_{2,t-1}, \beta_2) + \{\exp(\epsilon_1^2) - \Lambda^1(t-1, X_{1,t-1}, \beta_1)\} \times$$

$$\times \{\Lambda^2(t, X_{2,t}, \beta_2) - \Lambda^2(t-1, X_{2,t-1}, \beta_2)\} / \{\Lambda^1(t, X_{1,t}, \beta_1) - \Lambda^1(t-1, X_{1,t-1}, \beta_1)\}].$$

Having found the support of ϵ_2 we can see that $g(\epsilon_1)$ is nothing but the expression given above for the support of ϵ_2 . Furthermore we can in a similar manner derive the function $g(\epsilon_2)$ involved when considering the likelihood of a failure of type two.

Consider now an observation that is being censored at time period t . We know that the duration of this observation is longer than $t-1$, and thus its contribution to the likelihood is given by

$$P_i^2(\theta) = \int_{-\infty}^{\infty} \ln \Lambda^1(t-1, X_{1,t-1}, \beta_1) \int_{-\infty}^{\infty} \ln \Lambda^2(t-1, X_{2,t-1}, \beta_2) f(\epsilon_1, \epsilon_2) d\epsilon_2 d\epsilon_1$$

Taking into account the above analysis we can now write the likelihood function. Let $i=1,2,\dots,N$ denote observations and $t=1,2,\dots,T$ discrete time periods. Let an observation for individual i be given by the quintuple $(t_i, d_i, q_i, X_{1,t_i}, X_{2,t_i})$ where t_i is the period where failure is observed, $d_i=1$ if failure of type 1 occurred and $=0$ otherwise, $q_i=1$ if the observation is not censored and $=0$ otherwise, and finally X_{1,t_i} and X_{2,t_i} the vector of covariates for risks 1 and 2 respectively. Let $y_{it}=1$ if individual i is observed to fail, or being censored at t , and $=0$ otherwise. Then the likelihood function is

$$\ln L(\theta) = \sum_{i=1}^N \sum_{t=1}^T y_{it} [q_i d_i \ln P_1^1(\theta) + q_i (1-d_i) \ln P_2^2(\theta) + (1-q_i) \ln P_1^2(\theta)].$$

By maximizing the likelihood function one can get parameter estimates of the baseline hazards for each time interval for each risk, as well as estimates of the β 's.