Apply or Defer? Identifying the Most Important Factors for Economics PhD Admissions

Knickolas Laux

Michigan State University

December 2024

Abstract

I identify the most important elements of an applicant's profile in economics PhD admissions using 16 years of survey data from Urch.com. I divide factors into those that signal academic or research potential. I find no statistical evidence that working as a full-time predoctoral research assistant improves the median or best rank of the programs that admit an applicant, nor does it increase their chances of admission at a top-10 or top-30 program. Increases in the quantitative GRE score above 168 are not associated with an improvement in the median or best rank of the programs that accept them.

Keywords: Economics PhD, Graduate Admissions, Signaling

(*JEL* A22, I23, I21)

1 Introduction

When a prospective economics PhD student is looking for advice on the most important features they should have in their profile, there are only a few viable channels. The most common practice is consulting one's economics professors, who have personal experience with applying to graduate schools and being admitted. Though some advice may be consistent between professors—such as recommending proof-based math courses—these mentors may be years or even decades removed from the application process, and therefore might not understand the current trade-offs students are faced with when trying to obtain a competitive profile. As another resource, the literature on economics PhD admissions is growing. However, it only contains a few empirical studies, which estimate the marginal effects of specific elements of applicant profiles on the chance of admission.

In this study, I gather 16 years of anonymous application profiles and results from Urch.com, which includes detailed information on each applicant, the programs they applied to, and the ones that accepted them. Unobserved variables include the strength of an applicant's letters of recommendation and the prominence of their letter writers, but I add a series of controls to the models to minimize potential bias in the estimates. The results provide increased clarity for future applicants regarding the most important characteristics for improving the expected median and best rank of the programs that admit them. The results may also enhance the chances of their admission to a top-10 (T10) or top-30 (T30) program.

In line with the literature on PhD admissions, I model the decisions of admissions committees as maximizing the potential for success in the program for the incoming cohort as a function of observable characteristics. This allows for a division of applicant traits along two vectors: those that signal the potential for survival in the first year of core courses and those that signal the potential to propose and complete a dissertation. Traits that may signal the potential for survival in the first year include characteristics related to academic performance, such as an applicant's grade point average (GPA), their selection of math courses, whether they have already taken graduate courses in economics, and the strength of their letters of recommendation. Traits that may signal the potential to propose and complete a dissertation include completing a senior thesis, working as a part-time undergraduate research assistant, working as a full-time research assistant after graduation, completing a master's degree, and earning funding from the National Science Foundation Graduate Research Fellowship Program (NSF GRFP).

I estimate a model of the median rank of the programs that accept an applicant, as well as the rank of the best program that accepts an applicant with all of the above as regressors except for letters of recommendation. In addition, I estimate a probit model of the probability of acceptance to at least one T10 program and a probit model of the probability of acceptance to at least one rank-11 to rank-30 program.

One aspect of most applications that does not cleanly fit into either category is an applicant's quantitative Graduate Record Examinations (GRE) score. The quantitative portion of the GRE does not test an applicant's abstract mathematical knowledge or their ability to propose a compelling research question. For these reasons, the quantitative GRE score may be the purest signal in an applicant's portfolio—that is, it is a metric the applicant must acquire at a cost, provides no discernible human capital gains, and reveals little information about the applicant's ability for success in the program past a certain level of achievement. In fact, the quantitative GRE score almost certainly serves as a screening device for admissions committees to assess the applicant's match with the program or their analytical IQ before spending further resources reviewing the rest of their application. In this study I identify a cutoff score on the quantitative GRE, above which an increase in the score is not associated with a change in median or best acceptance rank. The cutoff in the rank models provides some evidence for the quantitative-GRE screening hypothesis.

Other notable findings primarily concern the most important courses and research-related activities for admissions, as well as the outcomes on which they have an effect. For example, I find that taking a course in real analysis is associated with an improvement in the best rank of the programs that admit an applicant—as well as an increase in an applicant's chance of admission to a T30 program—but does not have a statistically significant effect on the median rank of the programs that admit them or their chance of admission to a T10 program. Conversely, taking a course in multivariable calculus is associated with an improvement in the median rank of the programs that admit an applicant, but does not have a statistically significant effect on other admission outcomes. Unexpectedly, working as a full-time, predoctoral research assistant does not have a statistically significant effect on any of the admissions outcomes. In fact, the estimated coefficient on the predoc variable is both positive and within two-tenths of zero in both rank models. These and other findings may help advisors and professors when deciding which experiences and courses to recommend to prospective applicants.

2 Literature Review

This study contributes to the growing literature on the factors that influence an applicant's likelihood of acceptance to an economics PhD program. The theory of signaling and screening underlies nearly all studies on the most important traits for entry into an economics PhD program. This is especially true when it comes to the signaling power of education and how admissions committees screen applicants to discern their abilities and potential for success. Spence (1973) lays the foundation for signaling in the case of labor markets and restricts signals to the observable attributes of an individual they can alter—often at some cost—and which serve as parameters for an employer or signal-receiver's probabilistic estimation of the true abilities of the signaler. Signals in a graduate admissions setting include the quantitative and qualitative aspects of an applicant's profile, which clarify admissions committees' imperfect information about the applicant. Signal receivers—the admissions committees—can implement screening devices to incentivize applicants to reveal their true ability and desire to attend. Stiglitz (1975) discusses the theory of screening and the conditions under which the most capable people are incentivized to incur the cost of screening in order to distinguish themselves. In the case of PhD admissions, the benefit to society of the appropriate allocation of scarce educational resources may be worth the collective cost of screening.

The role of signaling and screening in PhD admissions may reflect the role the two mechanisms play in undergraduate admissions, an area with which the literature is familiar. Historically, the human capital model of education painted a much different picture of the value of educational attainment compared with the signaling model of education. However, it may be the case that education provides a student with both greater human capital and a signal of their innate ability (Becker, 1994). Bedard (2001) tests the educational sorting hypothesis, whereby greater university access will produce a greater number of high-school dropouts and university enrollees, against the human capital model of education and finds evidence for signaling as a component of the labor market, since the share of high school dropouts increased with increasing university access. As the educational sorting hypothesis proposes, the less often individuals of different abilities are incentivized to send the same signal, the more precise the information relayed in the signal. Rekus and Jiang (2021) discuss a model of college admissions and the subsequent allocation of scholarships, in which universities screen applicants using standardized test scores and interviews to improve their incomplete information on applicant ability and interest. Rekus and Jiang conclude that there exists a separating equilibrium that incentivizes every student to reveal their ability and interest. In addition to interviews, undergraduate admissions offices employ the frequency of an applicant's contact with the university as a signal of their desire to attend. Dearden et al. (2017) study the benefits of off-site and on-site contacts as signals of intent to matriculate from applicants to selective universities, and find that the more costly on-site contacts increase the likelihood of admission more than off-site contacts.

As the theory would predict, the institution an economics PhD graduates from has real consequences for their set of potential employers, and thus so does the set of institutions that offer admission to a prospective graduate student. Focusing on the academic labor market, the impact of the rank of a new PhD's graduating institution is apparent. Stock, Alston, and Milkman (2000) survey 897 job market candidates and search committees from the 1995-96 market and find that the majority of candidates had moved at least 50 ranks below that of their graduating department. While the trend may not be as strong in recent years, it still exists. Jones and Sloan (2020) find upward mobility in the top 96 economic departments to be minimal, with the majority of economics PhDs hired in an assistant professor position placing 10 to 20 ranks below the rank of their graduating department. Therefore, uncovering the most important characteristics for admission to economics PhD programs reveals the set of signals and screening devices admissions committees employ to not only allocate scarce educational resources, but also build cohorts that will carry the profession forward.

The literature on the factors that influence admissions to economics PhD programs is mostly confined to the last 20 years, but is already rich regarding its identification of important applicant characteristics. Not surprisingly, an applicant's GPA and their GRE score often correlate with their chances of admission to PhD programs, with a particular focus on the quantitative portion of the GRE. Attiyeh and Attiyeh (1997) analyze data on applications from 1990 and 1991 to PhD programs in mechanical engineering, mathematics, English, economics, and biochemistry, using a probit model for admissions based on talent and demographics. They find that applicants with an above-average quantitative GRE score or GPA have a higher probability of admission to one of the 48 leading universities in the sample. Probit models are often employed in the literature to test the marginal effects of improving the quantitative aspects of an applicant's profile on their chance of admission. However, they can also reveal relationships between demographic characteristics and likelihood of admission. Stock and Siegfried (2015) collect the undergraduate institutions and GRE scores of 586 students who started their PhD program in fall 2002 and create a probit model of enrollment in a top-15 economics PhD program. They find that quantitative and analytical GRE scores have positive marginal effects on chance of admission, as does being an international student from a country where English is the primary spoken language.

Another key component of an applicant's profile, which the literature has identified primarily through surveys, is evidence of taking a set of mathematics courses that are prerequisites for the theoretical courses of the first year of the PhD program. Milkman and Marjadi (2017) survey economics PhD graduate program directors at 67 universities to compile a list of the mathematics courses required for admission, finding over half require the calculus sequence, linear algebra, and introductory statistics. These results were replicated in another survey-based study, confirming the assertion that programs look for specific mathematical coursework. In that study, Jones et al. (2020) survey admissions coordinators at 69 of the top 132 PhD programs in economics on the importance of various profile attributes for admission, finding the majority of programs place high importance on recommendation letter strength, GPA, the quantitative portion of the GRE, and mathematical prerequisites such as the calculus sequence, linear algebra, and statistics. Notably, far more top-25 programs than lower-ranked programs indicate that real analysis is very important for admissions. These results are not surprising, since leaders in economics already steer prospective graduate students toward mathematical training. One example is the American Economic Association (2024), which lists real analysis, mathematical statistics, multivariable calculus, and linear algebra on its website as recommended courses for those seeking a PhD in economics.

Recent studies on economics PhD admissions have drawn specific attention to the recent trend of deferring applications until after working as a full-time, predoctoral research assistant, as well as potential bias in admissions decisions based on the rank of an applicant's undergraduate institution. Schlauch and Startz (2018) create a sample of all job market candidates from the top 50 (as defined by *U.S. News & World Report 2013* rankings) economics PhD programs in 2017 and find those who graduated from a high-ranked undergraduate institution and those who pursued a full-time research experience before entering graduate school graduated from a higher-ranked PhD program on average. Analysis of publicly available curriculum vitae (CV) has been a common method for determining the importance of predoctoral research and undergraduate institution rank. Wei (2022) gathers education and experience data from the CVs of job market candidates from the top 50 economics PhD programs, finding full-time research assistant experience and originating from western or eastern Europe to be associated with graduating from a higher-ranked program. In contrast, graduating from a relatively low-ranked or unranked undergraduate institution is associated with graduating from a lower-ranked program.

While full-time predoctoral research assistantships and high undergraduate rank are associated with graduating from a higher-ranked PhD program on average, it is important to note that another key component of applications—the strength of an applicant's recommendation letters—is likely correlated with both full-time research work and attending a top undergraduate university. Bai et al. (2022) weigh in on the subject with their study on 6,320 applications between 2013 and 2015, finding positive marginal effects for recommendation length and recommendation subjective score on the probability of admissions to a T10 program; also, recommendations from writers who were able to compare the applicant to past successful applicants have greater effect. Thus the inability to observe recommendation letters may bias the estimated effects of working as a predoctoral research assistant or attending a top undergraduate institution in ex post studies on the CVs of job market candidates. While letters of recommendation may explain much of the relationship between undergraduate rank and admissions, as Bai et al. (2022) hypothesize, it may also be the case that admissions committees believe that undergraduate admissions offices themselves accurately screen college applicants for their ability, and therefore undergraduate rank can be treated as an accurate signal of the ability of prospective graduate students.

In addition to the potential importance of predoctoral research assistantships, recent studies also comment on the impact of instead pursuing a master's degree. Schlauch and Startz (2018), in their study on 2017 job market candidates, find no relationship between doing a master's degree and PhD rank unless an applicant's undergraduate institution was unranked, in which case they graduated from a higher-ranked PhD program on average. Therefore, having a master's degree may not act as a signal on its own, but rather as a supplement to applicants from universities the admissions committee is unfamiliar with, regardless of whether the university is unranked or international. Perhaps less intuitively, Wei (2022) finds that having a previous graduate degree in any field before applying is associated with graduating from a lower-ranked program. When analyzing the potential effects of earning a master's degree or working as a full-time predoctoral research assistant on chances of admission, it is important to note the differences in terminal master's programs within and outside the United States, as well as the difference between the two in the availability of predoctoral research assistant positions.

Whereas a previous master's degree in economics and full-time research assistant work may serve as postgraduate signals of an applicant's exposure to and comfort with research, undergraduate research experience can potentially send a similar message. Boileau (2020) specifically highlights working as an undergraduate research assistant and writing a senior thesis as potential methods for applicants to distinguish themselves from the applicant pool. A senior thesis provides an undergraduate the opportunity to work on an original project, and working as a part-time undergraduate research assistant can expose a student to the research process in academia.

Grove, Dutkowsky, and Grodner (2007) examine the completion status of several cohorts of economics PhD candidates at Syracuse University and find positive marginal effects for having already completed a master's and the quantitative portion of the GRE on the probability of passing the theory comprehensive exam, as well as positive marginal effects for age and having written a thesis before graduate school on the probability of completing the dissertation. Thus previous original research may serve as a strong signal of an applicant's potential to propose and complete their dissertation in the latter part of the PhD program. Kartik (2020) argues that GRE scores and GPA are less distinguishable between candidates applying to top programs, and reinforces the importance of independent research for admission to top programs.

Another factor with which some applicants can distinguish themselves from other applicants is support from the NSF GRFP. The GRFP was established to support students in graduate school who are pursuing a STEM degree, and grants a $37,000 stipend and $16,000 education allowance for each of three years (U.S. National Science Foundation, 2023). While the literature has not historically touched on the GRFP, earning the fellowship indicates that an applicant is capable of writing a convincing research proposal, and fellowship funds significantly reduce the cost to graduate programs of admitting them. It is therefore highly likely that earning the GRFP increases an applicant's chances of admission; this study is the first to estimate the exact effect. Even without direct empirical evidence of the GRFP's effect on admissions, it is common practice in the discipline to recommend applying for the fellowship to those who are looking to go to graduate school. For example, Athey (2024), lists her recommendations for applicants on her website and claims that just applying to the NSF GRFP increases the probability of being accepted by a PhD program.

Yet another aspect of an applicant's profile the literature addresses is whether they took graduate courses in economics while completing their undergraduate degree. Olszewski (2020) claims that taking graduate economics courses as an undergraduate can send admission committees the clearest signal of an applicant's preparedness. Indeed, if admissions committees are looking to maximize the potential for the incoming cohort to survive the first year of courses, previous enrollment in the core sequence provides an unambiguous signal of such potential. In particular, success in the first graduate microeconomics course may be the most useful of the core sequence for demonstrating an applicant's readiness for graduate school. Berliant (2020) recommends taking the graduate microeconomics sequence as an undergraduate to prepare for applications, and argues that graduate microeconomics is the most standardized across universities and therefore provides the clearest signal of an applicant's ability to survive the first year of graduate courses. While the rewards of taking graduate courses are emphasized in the literature, the potential risk of earning a low grade in a graduate course is less clear. It is entirely possible that a low grade in a graduate course would have the opposite effect and lower an applicant's chances of admission.

Demographic information, such as an applicant's gender and race, is the final component on which the literature comments. Spence (1973) refers to these features of an individual as indices rather than signals, and views them as unalterable in the majority of cases. However, while an applicant's race and gender may not serve as signals associated with either potential for survival in the first year or potential to propose and finish a dissertation, the literature finds them to be important factors in admissions. Attiyeh and Attiyeh (1997) find that female applicants and underrepresented minority applicants have a higher probability of admission to one of the 48 universities in their sample, citing public policy as the impetus for the difference across demographics. Unsurprisingly, the trend of admissions committees' increased attention to demographic characteristics continues to today. Jones et al. (2020) note that the majority of their admissions committee respondents ascribe at least moderate importance to the race and gender of applicants.

This study contributes to the literature on economics PhD admissions through the inclusion of several unique explanatory variables and controls, the longest span of application years to date, and estimation of a median-rank and a best-rank model. In particular, the data include the specific course selections of each applicant, including math courses and graduate-level microeconomics, as well as whether the applicant received funding from the NSF GRFP. The data also detail the programs to which each applicant applied and their application set, which I include in the models to control for selection bias. In total, the sample spans 16 years of application cycles, from 2007 to 2022. In addition to probit models of the chance of admission to at least one T10 and T30 program, I model the median rank and best rank of the programs that admit an applicant—a unique advantage of observing the full set of programs which admit each applicant. The median-rank and best-rank models enable the separation of important characteristics into those that improve the best rank of the programs that admit an applicant and those that improve the median rank, if not both.

3 Data

3.1 Sample Source and Limitations

Applicant profiles were gathered from the Profiles and Results forums on Urch.com (Urch), which are created annually (see Appendix A for links). Urch hosts various forums that discuss the details of admissions for those interested in pursuing a range of degrees, from an MBA to a political science PhD, although the website is most popular for its content on applying for and earning an economics PhD. Profiles and Results forums are posted annually and include a header post that provides a template for applicant respondents to complete. The template includes prompts for the type of undergraduate institution the applicant attended, their undergraduate GPA, the math courses they took, a description of their relationship with their letter writers, their research experience, the schools that accepted them, the schools that rejected them, and other details. Each response is anonymous and each applicant's details are presented in the template format.

Therefore, the data share many similarities with survey data, including a set of specific limitations. In particular, the anonymity of the Profiles and Results responses could lead some respondents to be dishonest about the strength of their profile without fear of repercussions. While this self-reporting bias is possible, I do not see a clear incentive for respondents to be dishonest about their statistics or results, again due to the anonymity of their profiles. Moreover, I do not have any reason to suspect there was systematic self-reporting bias in the sample, since respondents' anonymity eliminated any reputation gains from overrepresenting the strength of a given university or program. Another potential limitation of the data due to the voluntary nature of the forum is selection bias. It is possible that the sample consists of a disproportionate number of successful applicants, but fortunately I observe factors such as the set of programs each respondent applied to and the approximate rank of the economics department at their university, which I implement as controls to minimize the effects of potential sample selection bias.

3.2 Variable Creation and Summary Statistics

As the four outcome variables, I chose the median rank of the programs that accepted each applicant, the best rank of the programs that accepted each applicant, whether each applicant was accepted to a T10 program conditional on having applied to one, and whether each applicant was accepted to a T30 program conditional on having applied to one. Respondents who were not accepted to any program, or who were only accepted to an agricultural economics program, were not included in the sample. To create the median-rank and best-rank variables, I collected the set of programs that accepted each applicant and assigned each program its corresponding rank according to *U.S. News & World Report 2024* to create each applicant's acceptance set. The median-rank variable, *MedRank*, was defined as the median of the acceptance set, and the best-rank variable, *BestRank*, was defined as the minimum of the acceptance set. I followed a similar process for the entire set of programs to which a respondent applied to create each applicant's application set. A dummy variable, *AdmitT10*, was created to take the value 1 if a number equal to or less than 10 was in the acceptance set. Similarly, a dummy variable, *AdmitT30*, was created to take the value 1 if a number greater than 10 but less than or equal to 30 was in the acceptance set.

For the majority of respondents, their GPA was reported on a 4.0 scale and their quantitative GRE score based on the current 170-point scale. However, some international applicants reported a GPA on a different scale, and some applicants from the earliest years in the sample reported a quantitative GRE score based on the previous 800-point scale. For those international applicants who reported their exact country of origin in their profile, their GPA was converted to a 4.0 scale based on the country-specific grading scales provided on the World Education Services website on Country Resources (World Education Services, 2024). For international students who did not include their country of origin, their GPA was converted to a 4.0 scale based on conversions performed on GPAs using the same scale from a known country (see Appendix B for exact conversion equations). There are no official conversion tables between the previous 800-point scale and the current 170-point scale, but unofficial conversion tables generate a steep drop in the equivalent 170-point score for a few missed points on the old 800-point scale. A simple linear function was used to make the conversions, as follows: $Quant_{170} = 10 + \frac{1}{5}Quant_{800}$ for $Quant_{800} \in [700, 800]$, which encompasses the entire domain of the scores reported using the old scale.

Table 1 presents summary statistics across the sample of 530 applicants for the four outcome variables and for the GPA and quantitative GRE (Quant) variables—the two continuous explanatory variables in the sample. Of particular note is the median of Quant, which is a near-perfect score, and that 73 percent of the sample was admitted to a T30 program.

Table 1: Applicant Characteristics: Outcome and Continuous Variables

Variable	**Min**	**Median**	**x̄**	**Max**	**s**
MedRank	1.00	22.00	28.39	110	20.79
BestRank	1.00	14.00	19.46	110	17.74
AdmitT10	0.00	0.00	0.34	1	0.47
AdmitT30	0.00	1.00	0.73	1	0.44
Quant	150.00	169.00	167.69	170	3.14
GPA	2.30	3.80	3.73	4	0.25

Respondents were prompted to list the math courses, economics courses, and research experiences they had completed before applying, which enabled the creation of a set of dummy variables, many of which have yet to be included in empirical studies on the most important factors of admission to an economics PhD. Among the math courses included as dummy variables were multivariable calculus, linear algebra, and mathematical statistics due to the high percentage of programs that indicated they were necessary courses for admission (Milkman and Marjadi, 2017) or placed high importance on them (Jones et al., 2020). A math-course dummy was also created for real analysis, due to its prominence in the literature as an effective way to distinguish oneself as a competitive applicant. To round off the course-selection dummies, I also created a dummy for the first of the microeconomics PhD courses, because it was also highlighted in the literature as a differentiating factor and enables admission committees to directly observe an applicant's ability to handle the rigor of first-year material. All course-related dummy variables allowed for the estimation of effects that have not previously been measured in the literature.

Under the category of prior research experience, respondents stated whether they had (1) completed a senior thesis as an undergraduate, (2) worked as a part-time research assistant as an undergraduate, (3) completed a master's degree in economics or a related discipline, and (4) had worked as a full-time predoctoral research assistant after graduating with their bachelor's degree. I created a dummy variable, *SThesis*, to indicate the completion of a senior thesis, and another, *RA*, to indicate work experience as a part-time undergraduate research assistant. I also created a dummy variable, *Masters*, to indicate the completion of a master's degree, and finally a dummy called *Predoc* to indicate work experience as a full-time predoctoral research assistant.

Respondents who received support from the NSF GRFP listed their award in the section for the programs that accepted them, which enabled the creation of an NSF dummy variable, *NSF*—a unique feature of this study. Finally, since most internatio-nal respondents indicated their country of origin—or alluded to it through the grading scale for their GPA—I created a dummy variable, *ITLE*, to indicate an applicant's status as an international applicant from an English-first-language country, as well as a dummy to indicate an applicant's status as an international applicant from an non-English-first-language country, *ITLNE*. My decision to divide the international student category along the English-speaking line was motivated by the findings of Stock and Siegfried (2015), who provide evidence of a higher acceptance rate among applicants from an English-first-language country.

Table 2 provides summary statistics for the dummy explanatory variables. Of particular note, 77% of applicants took a course in real analysis—a much higher percentage than one would expect, based on the low percentage of programs that indicated real analysis was necessary for admission in Milkmand and Marjadi's (2017) study. Also, 23% of applicants took the first microeconomics PhD course, and only 2% received support from the NSF GRFP.

Table 2: Applicant Characteristics: Dummy Variables

Variable	$\bar{\mathbf{x}}$	**s**
MulCal	0.86	0.35
RAnalysis	0.77	0.42
LinAlg	0.87	0.34
MathStat	0.68	0.47
PhDMicro	0.23	0.42
RA	0.40	0.49
Predoc	0.28	0.45
SThesis	0.43	0.50
Masters	0.35	0.48
NSF	0.02	0.15
ITLE	0.07	0.26
ITLNE	0.19	0.39

I created a set of dummy variables to control for different trends in admissions across years, different admission rates across tiers of undergraduate institutions, and at least some of the potential bias caused by the inability to observe each applicant's letters of recommendation. Each respondent was given the opportunity to write about their impression of the strength of their letters of recommendation, based on the prominence of their letter writers and their relationship with them. Using the Natural Language Toolkit (NLTK) library in Python, I followed the process Ali (2023) outlines to perform a bag-of-words sentiment analysis of each of the responses in the letters of recommendation section (see Appendix C for the exact code). The result was a dummy variable, *Sent*, that takes the value 1 if the respondent had a positive impression of their letters of recommendation—which means that the sentiment analysis assigned a positive decimal to their statement—and 0 if the respondent expressed a neutral or negative impression, which means that the sentiment analysis assigned a negative or value-0 score to their statement. While admittedly crude, *Sent* was created to act as a proxy for the strength of an applicant's letters of recommendation, given the theoretically positive, if weak, correlation between an applicant's impression of the strength of their letters and their true impact.

To control for different trends in admissions across years, a dummy variable was created for each year in the sample, resulting in 16 year dummies. Last, to control for differences in admissions across undergraduate institutions, I created a dummy variable for the approximate tier of the economics PhD program at each applicant's bachelor's university, which was self-reported. The resulting four dummy variables were based on four groupings: institutions in the T10, *T10*; institutions ranked 11 to 30, *T30*; institutions ranked 31 to 50, *T50*; and institutions ranked 51 to 100, *T100*. A value of 1 in any of the undergraduate rank dummies was mutually exclusive with a value of 1 in any of the other undergraduate rank dummies, as well as with a value of 1 in either of the two international applicant dummies. Table 3 presents summary statistics for the set of dummy control variables. Notably, 72% of applicants expressed a positive impression of their letters of recommendation, and fewer applicants chose to post their profiles and results on Urch over time.

Table 3: Applicant Characteristics: Dummy Control Variables

Variable	$\bar{x}$	s
Sent	0.72	0.45
T10	0.12	0.33
T30	0.18	0.38
T50	0.16	0.37
T100	0.28	0.45
y2007	0.08	0.28
y2008	0.08	0.27
y2009	0.12	0.32
y2010	0.12	0.32
y2011	0.12	0.32
y2012	0.10	0.30
y2013	0.09	0.29
y2014	0.08	0.27
y2015	0.05	0.21
y2016	0.04	0.19
y2017	0.03	0.18
y2018	0.02	0.14
y2019	0.02	0.15
y2020	0.01	0.11
y2021	0.02	0.15
y2022	0.02	0.14

The final control variables were constructed from each respondent's application set to minimize selection bias in the estimated parameters. The resulting four count variables are *AST10*, which indicates the number of T10 programs to which a respondent applied, *AST30*, the number of rank-11 to rank-30 programs; *AST50*, the number of rank-31 to rank-50 programs; and *AST100*, the number of rank-51 to rank-100 programs. Importantly, because business school PhD programs were assigned the same rank as the rank of the economics department at the same university, it was possible for a respondent to have applied to more than 10 T10 programs. Table 4 provides summary statistics for the four application-set count variables, as well as the variable *AppSet*, which is a count of the total number of programs to which each applicant applied.

Table 4: Applicant Characteristics: Application Set Control Variables

Variable	Min	Median	$\bar{x}$	Max	s
AppSet	1	12	11.77	31	5.05
AST10	0	4	4.60	15	3.96
AST30	0	5	5.20	19	3.51
AST50	0	1	1.30	11	1.74
AST100	0	0	0.67	10	1.38

3.3 Empirical Framework

This study focuses on the estimation of four models of economics PhD admissions:

$$\begin{aligned}MedRank_i &= \beta_0+\beta_1 Quant_i+\beta_2(Quant_i-\overline{Quant})^2+\beta_3 GPA_i+\beta_4 MulCal_i+\beta_5 RAnalysis_i \\ &+ \beta_6 LinAlg_i + \beta_7 MathStat_i + \beta_8 PhDMicro_i + \beta_9 RA_i + \beta_{10} Predoc_i \\ &+ \beta_{11} SThesis_i + \beta_{12} NSF_i + \beta_{13} Masters_i + \beta_{14} ITLE_i + \beta_{15} ITLNE_i \\ &+ \beta_{16} ITLE_i * Masters_i + \beta_{17} ITLNE_i * Masters_i + X_i\gamma + A_i\alpha + u_i \quad (1)\end{aligned}$$

$$\begin{aligned}BestRank_i &= \delta_0+\delta_1 Quant_i+\delta_2(Quant_i-\overline{Quant})^2+\delta_3 GPA_i+\delta_4 MulCal_i+\delta_5 RAnalysis_i \\ &+ \delta_6 LinAlg_i + \delta_7 MathStat_i + \delta_8 PhDMicro_i + \delta_9 RA_i + \delta_{10} Predoc_i \\ &+ \delta_{11} SThesis_i + \delta_{12} NSF_i + \delta_{13} Masters_i + \delta_{14} ITLE_i + \delta_{15} ITLNE_i \\ &+ \delta_{16} ITLE_i * Masters_i + \delta_{17} ITLNE_i * Masters_i + X_i\gamma + A_i\alpha + v_i \quad (2)\end{aligned}$$

$$\begin{aligned}AdmitT10_i &= \eta_0+\eta_1 Quant_i+\eta_2(Quant_i-\overline{Quant})^2+\eta_3 GPA_i+\eta_4 MulCal_i+\eta_5 RAnalysis_i \\ &+ \eta_6 LinAlg_i + \eta_7 MathStat_i + \eta_8 PhDMicro_i + \eta_9 RA_i + \eta_{10} Predoc_i \\ &+ \eta_{11} SThesis_i + \eta_{12} NSF_i + \eta_{13} Masters_i + \eta_{14} ITLE_i + \eta_{15} ITLNE_i \\ &+ \eta_{16} ITLE_i * Masters_i + \eta_{17} ITLNE_i * Masters_i + X_i\gamma + A_i\alpha + w_i \quad (3)\end{aligned}$$

$$AdmitT30_i = \lambda_0 + \lambda_1 Quant_i + \lambda_2 (Quant_i - \overline{Quant})^2 + \lambda_3 GPA_i + \lambda_4 MulCal_i + \lambda_5 RAnalysis_i$$
$$+ \lambda_6 LinAlg_i + \lambda_7 MathStat_i + \lambda_8 PhDMicro_i + \lambda_9 RA_i + \lambda_{10} Predoc_i$$
$$+ \lambda_{11} SThesis_i + \lambda_{12} NSF_i + \lambda_{13} Masters_i + \lambda_{14} ITLE_i + \lambda_{15} ITLNE_i$$
$$+ \lambda_{16} ITLE_i * Masters_i + \lambda_{17} ITLNE_i * Masters_i + X_i \gamma + A_i \alpha + e_i \quad (4)$$

where $MedRank_i$ measures the median of applicant i's acceptance set; $BestRank_i$ measures the minimum of applicant i's acceptance set; $AdmitT10_i$ is a 0 - 1 variable indicating whether applicant i was admitted to at least one T10 program, conditional on having applied to at least one; and $AdmitT30_i$ is a 0 - 1 variable indicating whether applicant i was admitted to at least one T30 program, conditional on having applied to at least one. In each equation, X_i is a row vector of dimension 19 containing the dummy control variables: $Sent_i$, the undergraduate tier dummies, and the year dummies. The $T10_i$ undergraduate tier dummy and the year dummy for 2007 were omitted from X_i as reference groups. Also in each equation, A_i is a row vector of dimension four containing the application set controls: $AST10_i$, $AST30_i$, $AST50_i$, and $AST100_i$. Last, the error term for individual i in equation (1) is u_i, and v_i, w_i, and e_i are error terms for individual i in equations (2), (3), and (4), respectively.

Taking the partial derivative of equations (1) and (2) with respect to $Quant_i$ and setting the result equal to 0 produces the following two equations:

$$0 = \beta_1 + 2\beta_2 (Quant_i - \overline{Quant}) \quad (5)$$

$$0 = \delta_1 + 2\delta_2 (Quant_i - \overline{Quant}) \quad (6)$$

whose solutions represent the quantitative GRE scores at which the associated decrease in $MedRank_i$ and $BestRank_i$ of a one-point increase in the quantitative score is 0. For equation (5), the turning point is given by $Quant_i = \overline{Quant} - \beta_1/2\beta_2$, and for equation (6) the turning point is given by $Quant_i = \overline{Quant} - \delta_1/2\delta_2$. A turning point within the scoring range of the GRE, [130, 170], would provide evidence of a screening cutoff in admissions. Interaction variables between $ITLE_i$ and $Masters_i$ and between $ITLNE_i$ and $Masters_i$ are included to allow for the relationship between completing a prior master's degree and the four outcome variables to differ for international applicants compared with American applicants.

4 Results

In the following subsections I present the results of estimating equations (1), (2), (3), and (4), as well as the effect of including the unique application set controls.

4.1 Rank Models and Effects of the Application Set Controls

Equations (1) and (2) were estimated by simple ordinary least squares estimation, with small-sample-bias-adjusted, White, heteroskedasticity-robust standard errors. As a note, negative coefficients are synonymous with an improvement in the median and best rank of the acceptance set, since a lower numerical rank corresponds to a better program. Table 5 presents the estimated coefficients for the two rank models, but with an increasing number of control variables, starting with no controls (NC). Dummy controls (DC) are Sent, the undergraduate tier control variables, and the year control variables. Application set controls are only present in the full models.

Table 5: Median- and Best-rank Models with Increasing Controls

	Median Rank			Best Rank		
	NC	DC	Full	NC	DC	Full
Quant	-1.752^{***}	-1.322^{***}	-0.064	-1.127^{***}	-0.794^{*}	0.069
	(0.351)	(0.336)	(0.302)	(0.330)	(0.324)	(0.326)
(Quant - $\overline{Quant}$)^2	0.052	$0.070^{\dagger}$	$0.099^{\dagger}$	$0.096^{\dagger}$	0.111^{*}	0.137^{*}
	(0.048)	(0.041)	(0.053)	(0.050)	(0.047)	(0.057)
GPA	-21.005^{***}	-21.848^{***}	-9.635^{**}	-16.379^{***}	-16.807^{***}	-8.471^{**}
	(3.393)	(3.175)	(2.964)	(2.942)	(2.841)	(2.775)
MulCal	-1.676	-2.431	$-3.504^{\dagger}$	0.927	0.044	-0.689
	(2.354)	(2.350)	(1.937)	(2.131)	(2.203)	(1.976)
RAnalysis	-7.327^{***}	-6.214^{***}	-2.242	-8.089^{***}	-7.223^{***}	-4.448^{**}
	(1.914)	(1.852)	(1.565)	(1.779)	(1.746)	(1.669)
LinAlg	2.438	1.606	0.639	1.222	0.892	0.510
	(2.535)	(2.575)	(2.027)	(2.423)	(2.442)	(2.079)
MathStat	2.906^{*}	$2.521^{\dagger}$	1.564	1.363	1.064	0.515
	(1.422)	(1.402)	(1.129)	(1.266)	(1.275)	(1.120)
PhDMicro	-7.063^{***}	-6.481^{***}	-2.851^{*}	-5.193^{***}	-5.276^{***}	-2.643^{*}
	(1.562)	(1.535)	(1.181)	(1.312)	(1.307)	(1.132)
RA	-3.101^{*}	$-2.358^{\dagger}$	-0.985	-3.129^{*}	-2.739^{*}	$-1.832^{\dagger}$
	(1.447)	(1.366)	(1.022)	(1.240)	(1.179)	(1.038)
Predoc	-5.621^{***}	-3.346^{*}	0.160	-3.728^{*}	-2.122	0.025
	(1.658)	(1.639)	(1.250)	(1.458)	(1.457)	(1.271)
SThesis	$-2.699^{\dagger}$	-1.969	-1.031	-1.824	-1.331	-0.606
	(1.426)	(1.377)	(1.087)	(1.242)	(1.221)	(1.088)
NSF	-5.493^{*}	-6.136^{**}	-2.413	-7.257^{***}	-7.240^{***}	-4.219^{**}
	(2.219)	(2.369)	(1.938)	(1.449)	(1.976)	(1.601)
Masters	6.504^{**}	3.231	1.466	4.640^{*}	2.221	1.150
	(2.270)	(2.294)	(1.892)	(1.906)	(1.993)	(1.929)
ITLE	-9.855^{***}	-3.071	-2.718	-6.857^{***}	-1.235	-1.336
	(2.079)	(2.413)	(2.308)	(1.908)	(2.340)	(2.298)
ITLNE	-2.034	4.694	0.498	-0.546	$5.390^{\dagger}$	2.107
	(3.466)	(3.453)	(2.144)	(3.078)	(2.982)	(2.170)
Masters*ITLE	-6.992	-2.601	2.383	-3.022	0.223	4.108
	(4.333)	(4.531)	(3.873)	(4.129)	(4.290)	(3.964)
Masters*ITLNE	-12.857^{**}	-8.636^{*}	$-5.153^{\dagger}$	-8.083^{*}	-4.851	-2.714
	(4.313)	(4.154)	(2.745)	(3.856)	(3.601)	(2.810)

	NC	DC	Full	NC	DC	Full
(Intercept)	409.459***	328.519***	79.120	276.678***	212.149***	43.260
	(57.155)	(54.977)	(50.467)	(53.451)	(52.516)	(53.849)
Num. obs.	530	530	530	530	530	530
R^2	0.432	0.516	0.705	0.392	0.470	0.586
Adj. R^2	0.413	0.481	0.681	0.372	0.431	0.552

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$

I will begin by presenting the results of estimating the full rank models, then discuss the effects of including the unique application set controls. Starting with the continuous variables, the estimated coefficient on $(Quant - \overline{Quant})^2$ is statistically significant in both rank models, although the coefficient on Quant is not statistically significant in either rank model. Substituting $\hat{\beta}_1$ and $\hat{\beta}_2$ for β_1 and β_2 in equation (5), given that $\overline{Quant}$ is about 167.69, results in a turning point of about 167.37—which means that once an applicant scores 168 on the quantitative portion of the GRE, further increases in the score are not associated with an improvement in the median rank of the programs that accept them, ceteris paribus. Similarly, substituting $\hat{\delta}_1$ and $\hat{\delta}_2$ for δ_1 and δ_2 in equation (6) results in a turning point of about 167.44, which means that once an applicant scores a 168 on the quantitative portion of the GRE, further increases in the score are not associated with an improvement in the best rank of the programs that accept them, ceteris paribus. I test the robustness of these findings to a dummy variable identification strategy in Appendix D. The estimated coefficient on GPA is statistically significant in both rank models. A 0.2 increase in GPA is associated with a 1.927 rank decrease in the median rank and a 1.694 rank decrease in the best rank of an applicant's acceptance set, ceteris paribus.

Interestingly, the estimated coefficients on the course dummy variables reveal some separation between factors that are important for the median rank and those that are important for the best rank of an applicant's acceptance set. In particular, the estimated coefficient on MulCal is negative in both models, but is only statistically significant, at the 10% level, in the median-rank model. Conversely, the estimated coefficient on RAnalysis is negative in both models, but is only statistically significant in the best-rank model. Both LinAlg and MathStat have statistically insignificant coefficients in both rank models, and the null hypothesis that the coefficients on LinAlg and MathStat are jointly zero in both models is not rejected. In contrast, the estimated coefficient on PhDMicro is negative and statistically significant in both models. Taking the first microeconomics PhD course is associated with a 2.851 rank decrease in the median rank and a 2.643 rank decrease in the best rank of an applicant's acceptance set, ceteris paribus.

The estimated coefficients on the research-related dummy variables provide an indication of the signaling power of the experiences meant to signal research potential. For example, the estimated coefficient on RA is negative in both models, but is only statistically significant, at the 10% level, in the best-rank model. Specifically, working as a part-time undergraduate research assistant is associated with a 1.832 rank decrease in the best rank of an applicant's acceptance set, ceteris paribus. Perhaps more surprising is the estimated coefficient on Predoc, which is not only statistically insignificant, but also less than two-tenths from zero in both models. The estimated coefficient on SThesis is negative but insignificant in both models, although the coefficient on SThesis in the median model is less than -1. The estimated coefficient on Masters is positive but insignificant in both models, although the coefficient is greater than 1 in both. Last, the estimated coefficient on NSF is negative in both models, but is only statistically significant in the best-rank model, in which earning the support of the NSF GRFP is associated with a 4.219 rank decrease in the best rank of an applicant's acceptance set, ceteris paribus. Although the coefficient on NSF is statistically insignificant in the median-rank model, it is less than -2.

When interpreting the estimated coefficients on ITLE and ITLNE, is it important to remember that the reference group for both variables is American applicants from a T10 undergraduate university. The estimated coefficient on ITLE is negative in both models, and the estimated coefficient on ITLNE is positive in both models, but none of the coefficients are statistically significant. Therefore, while there appears to be some difference in the expected acceptance sets of international applicants from English-first-language countries and non-English-first-language countries, statistically speaking, both types of international applicants are not expected to fare worse than an American applicant from a top economics department. The estimated coefficients on the interaction terms are generally statistically insignificant, except for the coefficient on Masters: ITLNE, which is negative and statistically significant at the 10% level.

For the control variables, it is worth noting that the estimated coefficient on T100 is positive and statistically significant in both models. Graduating from an American undergraduate university with a PhD program ranked 51 to 100 is associated with an acceptance set with a 6.952 higher numerical median and a 5.830 higher numerical minimum than an applicant from an American university with a T10 PhD program, ceteris paribus. The cause of the difference in expected median and best rank is most likely a combination of the prominence of the recommendation letters within each group, as well as graduate admissions committees' reliance on the screening work of undergraduate admissions.

Another notable estimated coefficient among the controls is on AST10, which is negative and statistically significant, and the coefficient on AST100, which is positive and statistically significant. The significance of the coefficient on AST10 and on AST100 provides evidence that there would have been selection bias in the estimated coefficients had the application set not been included. The use of each respondent's application set as a set of control variables is unique to this study, and since several of the application set controls have statistically significant coefficients, I will use the next few paragraphs to briefly investigate the effect of omitting them.

In the estimated median- and best-rank models without any controls, far more of the coefficients are statistically significant than in the models with full controls, including those on RAnalysis, MathStat, RA, Predoc, and Masters. Even after including the dummy controls in the median-rank model, only the estimated coefficient on Masters is statistically insignificant, while the estimated coefficients on RA and MathStat are statistically significant at the 10% level. Notably, including the dummy control variables decreases the magnitude of most of the estimates across both models, even if they remain statistically significant. This suggests that the majority of the independent variables are at least somewhat correlated with Sent, the tier of an applicant's undergraduate university, and the year an applicant applied, as one would expect. Furthermore, including the dummy control variables increases the fraction of the variance in the median rank of an applicant's acceptance set that the model can explain by 0.084, and the fraction of the variance in the best rank of an applicant's acceptance set that the model can explain by 0.078.

Comparing the estimated coefficients obtained from models with dummy controls with the estimated coefficients from models with full controls reveals a much larger difference. Inclusion of the application set controls shifts almost every estimated coefficient closer to zero, which renders the estimated coefficients on RAnalysis, Predoc, and NSF no longer statistically significant in the median-rank model. Thus, it appears that applying to a high number of T10 and T30 programs is at least moderately correlated with having taken real analysis, having worked as a full-time predoctoral research assistant, and having applied for and ultimately won the support of the NSF GRFP. This means that omitting the applicant set controls induces omitted-variable bias in the model. Since inclusion of the application set controls shifts most of the previously negative estimates closer to zero, the sign of the omitted-variable bias associated with the application set controls is negative.

More specifically, I assert that omission of the application set controls induces selection bias in the estimates. For example, without the application set controls, holding the other variables constant and comparing the expected median rank of the acceptance set for those applicants who worked as a predoctoral research assistant and those who did not enables comparison of applicants using an unbalanced fraction of each group who chose to apply to many or most T10 or T30 programs. Those who applied to few or no T10 and T30 programs automatically have a higher expected median rank, because there is no chance they were admitted to some of the top programs which the members of the other group have at least some chance of being admitted to based on simply having applied. In short, including the application set controls allows for a less biased comparison, since they allow for estimation of the relationship between working as a predoctoral research assistant and the median and best rank when not only other independent variables are held constant, but also the exact number of programs within each tier to which the applicant applied.

As a final note, including the application set controls increases the fraction of variance in the median rank of an applicant's acceptance set that the model can explain by 0.189 over the model with dummy controls. Similarly, including the application set controls increases the fraction of the variance in the best rank of an applicant's acceptance set that the model can explain by 0.116 over the model with dummy controls. These increases are large but not surprising, since the acceptance set is a subset of the application set.

4.2 Probit Models

Equations (3) and (4) were estimated by maximum likelihood estimation using a probit regression, again with small-sample-bias-adjusted, White, heteroskedasticity-robust standard errors. Table 6 presents the average marginal effects for the two models.

Table 6: T10 and T30 Probit Models with Full Controls

	T10 Probit	T30 Probit
Quant	0.020	−0.004
	(0.012)	(0.007)
(Quant - $\overline{Quant}$)^2	0.003	−0.001
	(0.002)	(0.001)
GPA	0.151	0.002
	(0.105)	(0.064)
MulCal	0.018	−0.043
	(0.053)	(0.046)
RAnalysis	0.037	0.086*
	(0.045)	(0.042)
LinAlg	−0.070	−0.041
	(0.051)	(0.041)
MathStat	−0.045	0.022
	(0.037)	(0.031)
PhDMicro	0.114*	0.007
	(0.051)	(0.036)
RA	0.029	0.004
	(0.038)	(0.034)
Predoc	−0.025	0.021
	(0.042)	(0.037)

	T10 Probit	T30 Probit
SThesis	-0.001	0.083^{**}
	(0.038)	(0.032)
NSF	0.341^{***}	0.203^{***}
	(0.070)	(0.014)
Masters	0.015	-0.055
	(0.071)	(0.046)
ITLE	-0.048	0.087
	(0.072)	(0.078)
ITLNE	-0.086	-0.089
	(0.059)	(0.080)
Masters*ITLE	-0.049	-0.236
	(0.115)	(0.176)
Masters*ITLNE	0.084	0.100
	(0.105)	(0.059)
AIC	361.511	389.802
BIC	524.955	561.605
Log Likelihood	-139.755	-153.901
Deviance	279.511	307.802
Num. obs.	398	488

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Compared with the rank models, there are some clear differences in the results of estimating the probit models. In contrast to the estimated coefficients in the rank models, the estimated average marginal effects of Quant, $(Quant - \overline{Quant})^2$, and GPA are not statistically significant in either probit model, perhaps reflecting the density of high GPAs and high quantitative GRE scores among the applicant pools at top universities. Of the math course dummy variables, the only statistically significant estimated average marginal effect is for RAnalysis, and only in the T30 probit model. Taking a course in real analysis is associated with a 0.086 increase in the probability of acceptance to at least one rank-11 to rank-30 program, conditional on having applied to at least one, ceteris paribus. While the estimated average marginal effect of RAnalysis is only statistically significant in the T30 probability model, the estimated average marginal effect of PhDMicro is only statistically significant in the T10 probit model, with an even larger magnitude. Taking the first microeconomics PhD course is associated with a 0.114 increase in the probability of acceptance to at least one T10 program, conditional on having applied to at least one, ceteris paribus.

When it comes to the research-related dummy variables, most of the estimated average marginal effects are statistically insignificant. The estimated average marginal effect of RA is positive in both models, but not statistically significant in either. For Predoc, the estimated average marginal effect is only positive in the T30 model, and it is statistically insignificant in both. This provides more evidence for the hypothesis that working as a full-time predoctoral research assistant does not improve an applicant's admission outcomes. Unlike RA and Predoc, the estimated average marginal effect of SThesis is statistically significant in the T30 probability model, although it is not statistically significant in the T10 model. Specifically, writing a senior thesis is associated with an 0.083 increase in the probability of acceptance to at least one rank-11 to rank-30 program, conditional on having applied to at least one, ceteris paribus. Similar to Predoc, the estimated average marginal effect of Masters switches signs across the two models, but is statistically insignificant in both cases.

Earning the support of the NSF GRFP appears to have the largest single-variable impact, since the estimated average marginal effect of NSF is positive, relatively large, and statistically significant in both probit models. Specifically, earning the support of the NSF GRFP is associated with a 0.341 increase in the probability of acceptance to at least one T10 program, conditional on having applied to at least one, and a 0.203 increase in the probability of acceptance to at least one rank-11 to rank-30 program, conditional on having applied to at least one, ceteris paribus. The estimated average marginal effects of the international dummies are not statistically significant, and neither are the average marginal effects of the interactions between the international dummies and Masters. Therefore, the hypothesis that international students are as likely to be admitted to a program in the top 30 as an American applicant from an undergraduate university with a T10 economics PhD program is not rejected.

Briefly analyzing the estimated average marginal effects of the control variables provides some insight into the other factors at play in admissions to T10 and T30 programs. In the T10 probit model, all three of the undergraduate tier dummy variables have negative and statistically significant estimated average marginal effects. Because letters of recommendation are unobserved and the prominence of letter writers is almost certainly negatively correlated with the tier of the undergraduate university, the statistically significant estimates on the tier dummies may be evidence of the negative impact of letters written from less-well-known writers. It may also be the case that T10 programs believe that undergraduate admissions perform their own screening, and therefore interpret graduating from a lower-tier program to be a signal of ability, as was discussed in the literature review. Across the year dummy variables, it is worth noting the large, negative, and statistically significant estimated average marginal effect of y2020. The estimate suggests that applying in the 2020 cycle is associated with a 0.612 decrease in the probability of acceptance by a rank-11 to rank-30 program, conditional on having applied to one, ceteris paribus. Since the COVID-19 pandemic was throwing a wrench into most programs during the 2019-2020 academic year, this estimate is not necessarily surprising.

A final consideration when interpreting the results of estimating the two probit models is the sample size for each. Since respondents were only included in the data for each model if they applied to at least one program in each tier, the sample sizes for the two models are smaller than for the rank models. In particular, while the full sample consisted of 530 respondents, only 398 applied to at least one T10 program, and only 488 applied to at least one rank-11 to rank-30 program.

5 Discussion

The results of this study offer insights into the most important factors for economics PhD admissions, which will likely motivate further research on the subject and may benefit advisors and professors when deciding which experiences and courses to recommend to students. Since the quantitative GRE score does not fit cleanly into either the set of signals related to the potential to survive the first year of a PhD program or the set of signals related to the potential to propose and complete a dissertation, I will begin with a summary of the results with respect to the test. I find evidence for a quantitative GRE screening hypothesis, because a score above 168 is not associated with an improvement in the median or best rank of the set of acceptances, ceteris paribus. In the probit models, the estimated average marginal effect of the quantitative GRE score variable is not statistically significant, which provides evidence for Kartik's (2020) claim regarding the lack of difference between applicants to top programs on the basis of the GRE and GPA alone. Therefore, when it comes to the quantitative GRE score, the results support a target score of 168 or above—but if achieving a 168 impedes one's ability to do well in math courses or complete a senior thesis, it may be worth settling for a slightly lower score. In short, it is best to view the quantitative GRE score as a screening device, which may only serve to shrink the pile of applications before the rest of the profile is reviewed.

When it comes to signals related to the potential to survive the first year of the PhD program, the results highlight a variety of important factors, some of which are only effective in specific contexts. Starting with GPA, both rank models reflect the importance of maintaining a relatively high GPA. However, the difference in the estimated median and best rank of the acceptance set between a 3.9 GPA and a 4.0 GPA is less than a single rank, and in general it is likely not worth avoiding math classes or missing out on research experiences in order to maintain a perfect or near-perfect GPA. In terms of math courses, taking multivariable calculus is found to be associated with a more-than-three-rank improvement in the median rank of an applicant's acceptance set. In contrast, neither linear algebra nor statistics is associated with an improvement in median or best rank or an increase in the probability of acceptance to a T10 or T30 program. Taken at face value, it might seem as though linear algebra is not worth taking at all—but importantly, linear algebra is often a prerequisite for real analysis, which is an important factor in the results.

Specifically, taking a course in real analysis is found to be associated with a nearly three-rank improvement in the best rank of an applicant's acceptance set, ceteris paribus, and with an 0.086 increase in the probability of being accepted to at least one T30 program, conditional on having applied to one, ceteris paribus. These findings are consistent with those of Jones et al. (2020), who find real analysis to have a much higher importance for admission to the top 25 programs compared with programs outside the top 50. Therefore, there is evidence in favor of recommending a course in real analysis and to explain its prerequisites. The final academic-related signal is taking the first microeconomics PhD course, which is found to be associated with a more-than-two-rank improvement in both the median and best rank of an applicant's acceptance set—as well a 0.114 increase in the probability of their admission to at least one T10 program, conditional on having applied to one, ceteris paribus. Thus, taking the first microeconomics PhD course has the single largest estimated average marginal effect for a course, which is consistent with the course's being the most direct signal of the potential to survive the first year of an economics PhD program. Therefore, if it is feasible for a prospective graduate student to take the necessary mathematical prerequisites and succeed in the first microeconomics PhD course at their undergraduate university, my results provide evidence to support that choice.

Moving to signals related to proposing and completing a dissertation, the results provide evidence for the importance of research experience as an undergraduate. They also cast doubt on the value of predoctoral assistantships and previous master's degrees for improving admissions outcomes. When it comes to undergraduate research experience, working as a part-time undergraduate research assistant is associated with a nearly two-rank improvement in the best rank of the programs that admit an applicant, ceteris paribus, and completing a senior thesis is associated with a 0.083 increase in the probability of admission to at least one T30 program, conditional on having applied to one, ceteris paribus. Neither of these experiences has a statistically significant estimated coefficient in the median-rank model or a statistically significant estimated average marginal effect in the T10 probit model. Thus there is evidence to suggest that they are not as universally beneficial as taking the first microeconomics PhD course. However, given the contexts in which working as a part-time research assistant and completing a senior thesis are found to improve admission outcomes, there is evidence to support the recommendation to pursue one or both.

Perhaps the most surprising finding in this study is the lack of statistical evidence for the claim that working as a full-time predoctoral research assistant improves the rank of programs that accept an applicant or increases the probability of being accepted by a T10 or T30 program. In fact, once I control for each applicant's choice of application set, the estimated coefficient on Predoc is positive and less than two-tenths from zero in both rank models, with a more precise estimate than without the controls. Similarly, the estimated average marginal effect of Predoc in both probabit models is less than 0.025, and not statistically significant in either. Although I use the tier of the PhD program at an applicant's undergraduate university and their sentiment about the strength of their letters as proxies for the strength of an applicant's letters of recommendation, it may be the case that not observing letters of recommendation biases the Predoc estimate. However, it is reasonable to assume that the sign of the bias on the Predoc estimate as a result of omitting letters of recommendation would be negative in the rank models and positive in the probability models, since one of the primary benefits of predoctoral assistantships is the generally positive letters of recommendation one is assumed to obtain as a product of close work with a professor on research projects. This suggests that the true value of the coefficient on Predoc in the rank models may be even more positive than the finding in this study, and the true average marginal effect of Predoc in the probit models may be even more negative—or closer to zero—depending on the model.

The question then becomes how to interpret the value of working as a predoctoral assistant and whether advisors and professors should recommend the experience to prospective applicants. In short, I find that when comparing two applicants with otherwise equal statistics and including the exact number of programs within each tier to which they applied, working as a predoctoral research assistant does not appear to serve as a strong differentiating signal. It may be the case, however, that a prospective applicant has yet to take the GRE and has yet to take real analysis, and working as a predoctoral research assistant would give them time to improve other aspects of their profile. At the same time, they would be able to earn an income, which may improve their quality of life during graduate school. In such a case, it may be worth recommending a full-time predoctoral research assistantship, especially compared with the alternative of recommending a master's degree. Even if a predoctoral research assistantship does not serve as a positive signal in admissions, it is possible that the experience provides an applicant with an increase in human capital, and is therefore beneficial for reasons outside the realm of admissions.

With respect to earning a master's degree, I find no statistical evidence that completing a master's degree prior to applying to PhD programs improves the median or best rank of an applicant's acceptance set or increases the probability of an applicant's being admitted to a T10 or T30 program, ceteris paribus. However, if an applicant is an international student from a non-English-first-language country, completing a master's is associated with a more-than-five-rank improvement in the median rank of the programs that accept them, ceteris paribus. This finding is consistent with Schlauch and Startz (2018), who find no relationship between doing a master's and PhD rank—unless the applicant graduated from a university less familiar to the admissions committee, in which case completing the master's degree was associated with graduating from a better PhD program on average. Thus, unless a prospective applicant is from a non-English-first-language country and will defer applying in favor of a postgraduate research experience, the results support recommending a predoctoral assistantship rather than a master's degree, if only for the difference in net income.

To round off the research-related dummy variables, I find that earning the support of the NSF GRFP is associated with a more-than-four-rank improvement in the best rank of an applicant's acceptance set, ceteris paribus. I also find that earning the support of the NSF GRFP is associated with a 0.341 increase in the probability of being admitted to at least one T10 program, conditional on having applied to one, and a 0.203 increase in the probability of being admitted to at least one T30 program, conditional on having applied to one, ceteris paribus. The lack of statistical significance for the estimated coefficient on the NSF in the median-rank model is consistent with the timing of the announcement of awardees—which is usually in early April—since most initial admission offers have been made at that time. These results, combined with the timing of the award announcement, provide evidence for the hypothesis that the main mechanism through which earning the support of the NSF GRFP acts is converting an applicant's position on the waitlist for a top program into an acceptance. Given the relatively large estimated boost in the probability of acceptance to at least one top program, there is strong evidence to support recommending applying for the fellowship.

To touch on the international dummy variables, the signs of their estimated coefficients are consistent with Stock and Siegfried (2015), who find that applicants from English-first-language countries have a higher probability of acceptance to top programs than those from non-English-first-language countries, but the coefficients are also not statistically significant. Since the reference group for both international dummy variables is American applicants from an undergraduate university with a T10 PhD program, I cannot reject the null hypothesis that international applicants fare no worse than American applicants from top-ranked universities in the U.S. For the last point regarding the estimated models, at least half of the application set controls have statistically significant estimated parameters in each model, which provides evidence for omitted-variable bias in models that do not account for each applicant's choice of application set. I argue that the bias in models without application set controls is a form of selection bias, whereby applicants with profile characteristics that are colloquially believed to improve their chances of admission, such as having done a predoctoral assistantship, select to apply to a greater number of T10 and T30 schools and therefore inherently have a higher chance of admission to at least one of those programs. Inclusion of the application set controls allows the model to hold the exact number of applications to each program tier equal in order to obtain less biased, and often more precise, estimates of the parameters.

Having discussed the results in full, I would now like to reiterate the limitations associated with the anonymous survey nature of the data source as well as the inability to observe letters of recommendation. As with any survey that prompts respondents to self-report their academic statistics, there is a risk of dishonesty in the reporting. However, given the anonymous nature of the forum, there is no immediately apparent incentive for respondents to misrepresent their performance and experiences, and thus there is not a strong reason to believe there is systematic misrepresentation in the data. The other limitation related to the voluntary nature of the data source is selection bias. It is possible that a disproportionate number of successful or confident applicants chose to post to the forum, but measures related to the confidence of each applicant, such as their choice of application set, are observable and included as controls. Finally, I was unable to observe each applicant's letters of recommendation, which might have introduced omitted-variable bias to the estimates. I attempted to minimize the potential for this bias through the creation and inclusion of a dummy variable for each applicant's sentiment about the strength of their letters of recommendation, as well as dummy variables indicating the tier of the PhD program at each applicant's undergraduate university.

My findings may motivate further research on and discussion of the most important factors in an applicant's profile for admission to economics PhD programs. Given the contrast between my findings related to working as a full-time predoctoral research assistant and the general consensus in the economics profession, further work is merited to explore its true benefits. While it is not easy to collect data on the application set of each applicant, this study highlights the importance of doing so in order to obtain unbiased estimates. Further research should continue to model signaling between applicants and admissions committees. Last, if the GRE is ever removed as a required element of graduate applications, future research should attempt to detect the emergence of a new screening device to take its place, if one emerges at all.

6 Acknowledgements

I acknowledge Dr. Hanzhe Zhang and especially Dr. Tim Vogelsang for their valuable suggestions.

7 References

Ali, M. 2003. *NLTK Sentiment Analysis Tutorial for Beginners.* https://www.datacamp.com/tutorial/text-analytics-beginners-nltk.

American Economic Association. 2024. *Recommended Mathematical Training to*

Prepare for Graduate School in Economics. https://www.aeaweb.org/resources/students/grad-prep/math-training.

Athey, S. 2024. *Professional Advice.* https://gsb-faculty.stanford.edu/susan-athey/professional-advice/#: :text=On%20your%20graduate%20school %20application,thesis%20or%20research%20assistant%20work.

Attiyeh, G., and R. Attiyeh. 1997. Testing for Bias in Graduate School Admissions. *Journal of Human Resources* 32 (3): 524-548. doi: 10.2307/146182.

Bai, J., M. Esche, W. B. MacLeod, and Y. Shi. 2022. Subjective Evaluations and Stratification in Graduate Education. NBER Working Paper 30677.

Becker, G. S. 1994. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

Bedard, K. 2001. Human Capital Versus Signaling Models: University Access and High School Dropouts. *Journal of Political Economy* 109 (4): 749-775. doi: 10.1086/322089.

Berliant, M. 2020. Preparing Undergraduates for Application to Graduate School: Comments on Jones et al (2020). *Journal of Economic Education* 51 (2): 206-208. doi: 10.1080/00220485.2020.1731395.

Boileau, M. 2020. Admissions to Economics PhD Program: Perspectives from a Large Public University. *Journal of Economic Education* 51 (2): 201-202. doi: 10.1080/00220485.2020.1731393.

Dearden, J. A., S. Li, C. D. Meyerhoefer, and M. Yang. 2017. Demonstrated Interest: Signaling Behavior in College Admissions. *Contemporary Economic Policy* 35 (4): 630-657.

Grove, W., D. Dutkowsky, and A. Grodner. 2007. Survive then Thrive: Determinants of Success in the Economics PhD Program. *Economic Inquiry* 45 (4):

863-871. doi: 10.1111/j.1465-7295.2007.00041.x

Jones, A., P. Schuhmann, D. Soques, and A. Witman. 2020. So You Want to go to Graduate School? Factors that Influence Admissions to Economics PhD Programs. *Journal of Economic Education* 51 (2): 177-190. doi: 10.1080/00220485.2020.1731385.

Jones, T. R. and A. Sloan. 2020. Staying at the Top: The Ph.D. Origins of Economics Faculty. EdWorkingPaper: 20-324.

Kartik, N. 2020. Comments for JEE Based on ASSA Panel Discussion "Preparing Undergraduates for Application to Graduate School". *Journal of Economic Education* 51 (2): 199-200. doi: 10.1080/00220485.2020.1731394.

Milkman, M., and R. Marjadi. 2017. Undergraduate Mathematics Courses Required and Recommended for Admission to Economics PhD Programs in the United States. *American Economist* 62 (1): 118-125. doi: 10.1177/0569434516672777.

Olszewski, W. 2020. Preparing Undergraduates for Application to Graduate School. *Journal of Economic Education* 51 (2): 203-205. doi: 10.1080/00220485.2020.1731396.

Rekus, J., and Y. Jiang. 2021. Modeling College Admission as a Signaling Game. *Applied Economics and Finance* 8 (6): 21-33.

Schlauch, J., and R. Startz. 2018. The Path to an Economics PhD. *Economics Bulletin* 38 (4): 1864-1876.

Spence, M. 1973. Job Market Signaling. *Quarterly Journal of Economics* 87 (3): 355-374. doi: 10.2307/1882010.

Stiglitz, J. E. 1975. The Theory of "Screening," Education, and the Distribution of Income. *American Economic Review* 65 (3): 283-300.

Stock, W. A., and J. J. Siegfried. 2015. The Undergraduate Origins of PhD

Economists revisited. *Journal of Economic Education* 46 (2): 150-165.

Stock, W. A., R. M. Alston, and M. Milkman. 2000. The Academic Labor Market for Economists: 1995-96. *Atlantic Economic Journal* 28 (2): 164-185.

U.S. National Science Foundation. 2023. NSF Graduate Research Fellowship Program (GRFP). https://new.nsf.gov/funding/opportunities/nsf-graduate-research-fellowship-program-grfp.

Wei, Y. 2022. Would Origin and Type of Undergraduate Universities Affect Ph.D. Ranking? Evidence from 2021-2022 Job Market. UC Santa Barbara: Department of Economics. Retrieved from https://escholarship.org/uc/item/4tv459hn#main.

World Education Services. 2024. Country Resources. https://applications.wes.org/country-resources/.

8 Appendices

8.1 Appendix A: Urch Profiles and Results Pages

Listed below are links to the 16 profiles and results pages from Urch.com.

2022: https://www.urch.com/forums/topic/155581-profile-results-2022/

2021: https://www.urch.com/forums/topic/155015-profile-results-2021/

2020: https://www.urch.com/forums/topic/154464-profiles-and-results-2020/

2019: https://www.urch.com/forums/topic/153908-profiles-and-results-2019/

2018: https://www.urch.com/forums/topic/153174-profiles-and-results-2018/

2017: https://www.urch.com/forums/topic/152277-profiles-and-results-2017/

2016: https://www.urch.com/forums/topic/151159-profiles-and-results-2016/

2015: https://www.urch.com/forums/topic/149304-profiles-and-results-2015/

2014: https://www.urch.com/forums/topic/146181-profiles-and-results-2014/

2013: https://www.urch.com/forums/topic/141770-profiles-and-results-2013/

2012: https://www.urch.com/forums/topic/135122-profiles-and-results-2012/

2011: https://www.urch.com/forums/topic/126469-profiles-and-results-2011/

2010: https://www.urch.com/forums/topic/113943-profiles-and-results-2010/

2009: https://www.urch.com/forums/topic/108744-profiles-and-results-2009/

2008: https://www.urch.com/forums/topic/83699-profiles-and-results-2008/

2007: https://www.urch.com/forums/topic/64922-profiles-and-results-2007/

8.2 Appendix B: International GPA Conversions

Below are the formulas used to convert the GPAs of international students from different scales to the 4.0 scale.

First Class: $GPA_{4.0} = 3.5$

First Class with Honors or Distinction: $GPA_{4.0} = 3.75$

Second Class with Honors: $GPA_{4.0} = 3.3$

7-point scale: $GPA_{4.0} = \frac{4}{7}GPA_7$

5-point scale: $GPA_{4.0} = min\{\frac{1}{2}GPA_5 + 1.5, 4.0\}$

Percent: $GPA_{4.0} = \frac{1}{25}GPA_{Per}$

20-point scale: $GPA_{4.0} = \frac{1}{12}GPA_{20} + \frac{7}{3}$

10-point scale: $GPA_{4.0} = min\{\frac{4}{9}GPA_{10}, 4.0\}$

8.3 Appendix C: Sentiment Analyzer Code

Below is the Python code used to create the Sent variable with the NLTK package.

```
import nltk

import pandas as pd
```

```
from nltk.sentiment import SentimentIntensityAnalyzer
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report

df = pd.read_csv('https://raw.githubusercontent.com/pycaret/pycaret/ma
ster/datasets/amazon.csv')
df

# create preprocess_text function
def preprocess_text(text):

    # Tokenize the text
    tokens = word_tokenize(text.lower())

    # Remove stop words
    filtered_tokens = [token for token in tokens if token not in stopw
    ords.words('english')]

    # Lemmatize the tokens
    lemmatizer = WordNetLemmatizer()
    lemmatized_tokens = [lemmatizer.lemmatize(token) for token in filte
    red_tokens]
```

```
    # Join the tokens back into a string
    processed_text = ' '.join(lemmatized_tokens)
    return processed_text

# apply the function df
df['reviewText'] = df['reviewText'].apply(preprocess_text)
df

# initialize NLTK sentiment analyzer
analyzer = SentimentIntensityAnalyzer()

# create get_sentiment function
def get_sentiment(text):
    scores = analyzer.polarity_scores(text)
    sentiment = 1 if scores['pos'] > 0 else (scores['neg'] * 0)
    return sentiment

# apply get_sentiment function
df['sentiment'] = df['reviewText'].apply(get_sentiment)
df

print(df.head())

df1 = pd.read_csv(r"C:\Users\Knickolas\Documents\RWork\UrchLOR.csv")
```

```
df1
```

```
df1['LORText'] = df1['LORText'].apply(preprocess_text)
df1
```

```
df1['sentiment'] = df1['LORText'].apply(get_sentiment)
df1
```

8.4 Appendix D: Screening Hypothesis Robustness Check

To test the robustness of findings in support of the quantitative GRE screening hypothesis, I construct a simplified binary variable for the GRE in place of the linear and centered quadratic terms of equations (1) and (2). I label the general version of the resulting variable QX_i, where X is one of six quantitative GRE scores from the set {164, 165, 166, 167, 168, 169}. The QX_i variable takes the value 1 if applicant i earned a score of X or greater on the quantitative portion of the GRE. Since 168 was found to be the score after which an increase in an applicant's quantitative GRE score is no longer associated with an increase in the median or best rank of the programs that accept them, an estimated coefficient on QX_i that is statistically indistinguishable from zero when X = 169, and is statistically significant and negative for the remaining values of the set, would provide further evidence in support of the quantitative GRE screening hypothesis. Table 7 and Table 8 present the results of estimating the simplified forms of equations (1) and (2), respectively, by simple ordinary least squares estimation with small-sample-bias-adjusted, White, heteroskedasticity-robust standard errors.

Table 7: GRE Cutoff Checks: Median-rank Model

	Q169+	Q168+	Q167+	Q166+	Q165+	Q164+
Q169	-0.582					
	(1.140)					
Q168		$-2.548^{\dagger}$				
		(1.317)				
Q167			-2.824^{*}			
			(1.392)			
Q166				-3.545^{*}		
				(1.711)		
Q165					-4.247^{*}	
					(1.900)	
Q164						-5.970^{*}
						(2.491)
Num. obs.	530	530	530	530	530	530
R^2 (full model)	0.694	0.696	0.696	0.697	0.698	0.699
Adj. R^2 (full model)	0.669	0.672	0.672	0.673	0.674	0.675

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$

Table 8: GRE Cutoff Checks: Best-rank Model

	Q169+	Q168+	Q167+	Q166+	Q165+	Q164+
Q169	-0.666					
	(1.198)					
Q168		-2.669^{*}				
		(1.334)				
Q167			$-2.767^{\dagger}$			
			(1.413)			
Q166				-3.618^{*}		
				(1.808)		
Q165					-4.066^{*}	
					(1.981)	
Q164						-7.438^{**}

	Q169+	Q168+	Q167+	Q166+	Q165+	Q164+
						(2.624)
Num. obs.	530	530	530	530	530	530
R^2	0.560	0.563	0.563	0.565	0.565	0.571
Adj. R^2	0.525	0.529	0.529	0.530	0.530	0.537

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$

8.5 Appendix E: Predoc Estimates by Year

In this subsection I present the results of estimating equations (1) and (2) with the modification of interacting Predoc with each year to produce estimates of the coefficient on Predoc in each year of the sample. In both models, 2007 is the reference group. Figure 1 and Figure 2 present the point estimates of the Predoc coefficients for each year from estimation of the modified versions of equations (1) and (2), respectively, as well as a 90 percent confidence interval around each estimate.

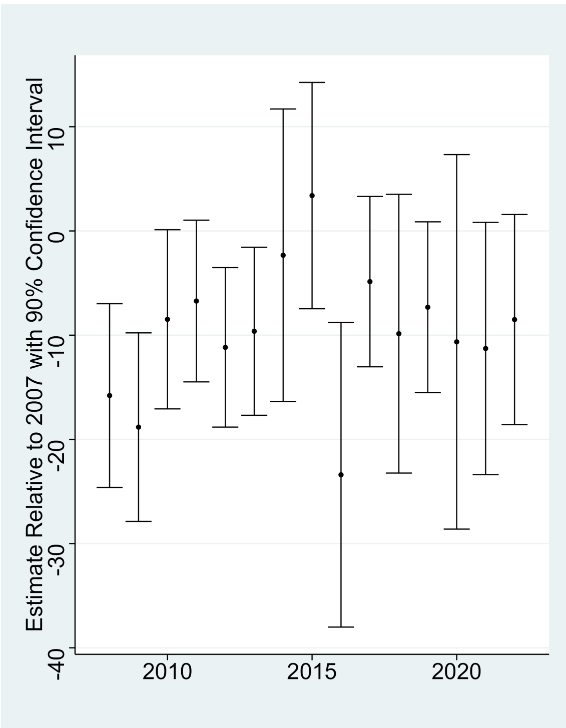

Figure 1: Median Rank: Yearly Predoc Esimates Relative to 2007 (9.265, se: 3.867)

In the case of the median rank of an applicant's acceptance set, the Predoc coefficient estimates in Figure 1 are mostly statistically indistinguishable from the reference estimate from 2007, which is about 9.625. Looking at the point estimates alone, the estimates from 2018 to 2022 are all relatively close to -9.625, which suggests that working as a full-time pre doctoral research assistant in those years is associated with little to no improvement in the median rank of the schools that accept an individual, ceteris paribus.

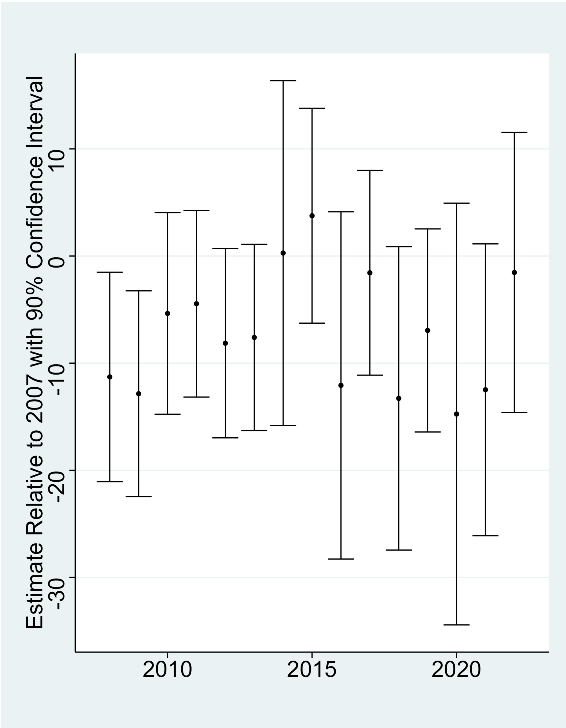

Figure 2: Best Rank: Yearly Predoc Esimates Relative to 2007 (6.328, se: 4.600)

In the case of the best rank of an applicant's acceptance set, the Predoc coefficient estimates in Figure 2 are again mostly statistically indistinguishable from the reference estimate from 2007, which is about 6.328. Looking at the point estimates alone, estimates from the most recent 5 years of the sample vary greatly, with some well below -6.328 and some closer to zero. Since the portion of the sample from the most recent years is significantly lower than in the first few years, a larger sample size would likely generate more precise estimates.