CENTER FOR
ECONOMIC
BEHAVIOR &
INEQUALITY

CEBI WORKING PAPER SERIES

Working Paper 06/22

EARLY CAREER SETBACKS AND WOMEN'S CAREER-FAMILY TRADE-OFF

Itzik Fadlon

Frederik Plesner Lyngse

Torben Heien Nielsen

ISSN 2596-447X

CEBI
Department of Economics
University of Copenhagen
www.cebi.ku.dk

Early Career Setbacks and Women's Career-Family Trade-Off [*]

Itzik Fadlon, Frederik Plesner Lyngse, Torben Heien Nielsen

March 2024

Abstract: We study how early career setbacks—in the form of worse initial job matches—have permanent labor and marriage market impacts differentially for males and females. We analyze the Danish physician labor market and exploit a randomized lottery that determines sorting into internships, which differ in the bundle of location and career opportunities they provide. Using administrative data for over fifteen years after the lottery experiment, we find that initial labor market sorting has important long-run effects on occupational choice and career trajectories for women only, which increases the gender earnings gap by 10-15 percent over the decades after graduation from medical school. We show that the differential gender sensitivity to setbacks is driven by women's career-family trade-off, where women exhibit earlier and higher fertility and subsequently sort into more flexible but lower-paying jobs that facilitate their greater family responsibilities. Our findings have implications for policies aimed at gender equality, as they reveal how persistent gaps can arise even in settings with institutional equality of opportunity and they point to addressing family considerations and job flexibility as key channels.

[*] We thank Julie Cullen, Gordon Dahl, Claudia Goldin, Chinhui Juhn, Larry Katz, Henrik Kleven, Camille Landais, Petter Lundborg, Karthik Muralidharan, Claudia Olivetti, Linh Tô, Tom Vogl, Melanie Wasserman and seminar participants at UCSD, University of Copenhagen, Berlin Applied Micro Seminar, USC, McGill University/Boston College Labor Seminar, Harvard Medical School/Health Care Policy, UCLA's California Center for Population Research, UC Irvine, Norwegian School of Economics FAIR Bergen, University of Bergen, University of Aarhus, BI Norwegian Business School Oslo, Bristol University, Oxford University, Erasmus University Rotterdam, CEPR Public Economics, Essen Health Conference, IIPF Annual Meeting, the EEA Annual Meeting, the 2020 NBER Summer Institute Gender in the Economy Meeting, and the 2022 NBER Summer Institute Personnel Economics Meeting for helpful comments and discussions. We thank the Danish National Health Authority, De Regionale Råd for Lægers Videreuddannelse, and Dansk Telemedicin A/S for data access and helpful institutional knowledge. We also thank Carl Benjamin Simpson, Sarah Hamilton, and Philip Nye for excellent research assistance. Fadlon: University of California, San Diego, Department of Economics, and NBER, 9500 Gilman Drive #0508, La Jolla, CA 92093-0508, fadlon@ucsd.edu; Lyngse: University of Copenhagen, Department of Public Health, frederik.lyngse@sund.ku.dk; Nielsen: University of Copenhagen, Department of Economics, and CEBI, thn@econ.ku.dk. We gratefully acknowledge funding from the Independent Research Fund Denmark (grant no. 9061-00035B). The activities of CEBI are financed by the Danish National Research Foundation (grant no. DNRF134). Lyngse gratefully acknowledges support from the Novo Nordisk Foundation (grant no. NNF20OC0059309). This manuscript is a revised version of a previous working paper titled "Causal Effects of Early Career Sorting on Labor and Marriage Market Choices: A Foundation for Gender Disparities and Norms".

1. Introduction

Modern economies continue to struggle with gender disparities in the labor market. Gender gaps in performance and pay remain pervasive even among high-skilled professionals, such as lawyers, business and finance professionals, and physicians (Azmat and Ferrer 2017, Sarsons 2019, Zeltzer 2020, Wasserman 2022, Cullen and Perez-Truglia 2023). These gaps display a clear common phenomenon of a persistent widening pattern following the early-career stage and the subsequent family formation and arrival of children (Bertrand, Goldin, and Katz 2010, Goldin 2014, Goldin et al. 2017, Blau and Kahn 2017, Juhn and McCue 2017, Kleven, Landais, and Søgaard 2019).

In this paper, we ask whether this persistent pattern among the highly-skilled represents a causal link from the early-career stage to long-run gender gaps. In the context of the physician labor market, we study how early career setbacks in the form of worse internships differentially affect long-run labor market outcomes for men and women, with a focus on the family-career trade-off as a potential root cause.

Our analysis offers two main advantages. First, placement into medical internships—i.e., physicians' first jobs—is governed in Denmark by a purely randomized lottery that provides a clean source of idiosyncratic variation in entry-level labor market sorting. We illustrate that students with the best lottery ranks are effectively unrestricted in their choices and are assigned their highest priorities, whereas students with the worst lottery ranks face narrowed choice sets and are assigned their lowest priorities. This creates large exogenous variation in the location and career opportunities of a graduate's internship bundle, where graduates with the worst lottery numbers end up choosing lower-ranked positions in rural communities. Second, we exploit a novel dataset spanning the years 2001-2022 that combines the formal lottery data we have digitized with a range of administrative datasets on all medical doctors in Denmark, including medical registries on licenses and specialties, the Danish population-wide economic registers, and family linkages. These data allow us to conduct a comprehensive analysis of the broad potential causal effects of early careers—on work, family, and their trade-offs—over a long horizon of over 15 years after the treatment, along with a careful investigation of potential mechanisms.

With this setup and design, we provide two sets of novel findings. First, in terms of long-run treatment effects, we find that the internship placement has important impacts on career outcomes, where the findings clearly and systematically show that women are much more sensitive to the initial bad shock. Starting with geographic sorting, we find that by the end of the analysis horizon of 15 years—long after the internship itself that lasts for 1-1.5 years—women with the worst lottery ranks are 4.9 percentage points (pp) more likely to locate in a rural municipality (on a counterfactual of 9.9 pp). Accordingly, they are also 6.5 pp less likely (on a counterfactual of 48.1 pp) to hold the higher-paying positions at university hospitals that are located in the vicinity of the urban hubs. Moreover, female physicians in the treatment group of worst lotteries have significantly lower propensities to make important human capital investments, displaying a 5.2 pp lower likelihood of obtaining a medical PhD (on a counterfactual of 29.7 pp). Finally, and most importantly for our investigation of long-run careers, we find significant effects on occupational choice as reflected in medical specialties. Women in the treatment group specialize earlier, accordingly spending less time on increasing their human capital through training, and they end up sorting into lower-paying female-represented specialties at much higher rates. On all these margins we find no effects for males, despite the fact that males and females have the same preferences over internship positions and that they make similar choices conditional on lottery number. We aggregate our findings on gender differentials in causal impacts into long-run implications for earnings gaps using the "surrogate index" method (Athey et al. 2019), which allows predictions over the course of 35 years in our observational data. We find a growing gap in predicted earnings that can be causally attributed to the lottery, leading to gender gaps in earnings that are even larger for women in the treatment group. Our variation in physicians' very first jobs increases the gender earnings gap by 10-15 percent over the decades following graduation.

Second, in terms of the channels for the gender divergence in sensitivity to career setbacks, we find strong support for the career-family trade-off and women's greater family responsibilities. The initial conditions lead women to specialization in home production, which then leads to larger differences in long-term career outcomes across men and women. We show that single women in the treatment group begin having higher fertility compared to single women in the control group following graduation; and that they subsequently sort at higher rates into female-represented specialties, which offer more time flexible jobs but with significantly lower pay. Given the cultural norms of women as secondary movers, we also find that they are more likely to remain stuck in the original location of their partners. Men, however, do not display any of these family effects or career-family trade-offs. We further find evidence that even higher human capital investment may not be able to shield women from the career costs of family responsibilities. We show that women in the intermediate lottery group, who are also induced by their career setbacks to choose higher fertility, are still ultimately diverted back to lower-paying female specialties despite the fact they obtain medical PhDs at rates similar to the control group. Our findings provide novel empirical support for the view that closing the remaining gender gap would indeed require changes in the labor market that would facilitate jobs with more temporal flexibility (e.g., through greater substitutability of workers and teamwork) to accommodate family responsibilities (Goldin 2014, Goldin and Katz 2011).

Our advancement of the work on gender in modern high-skilled markets makes several contributions. Most directly, a key contribution of our analysis is to the long-standing work on gender inequality in economic outcomes and their underlying sources (see reviews and discussions in, e.g., Bertrand 2011, Goldin 2014, Olivetti and Petrongolo 2016, Blau and Kahn 2017, Lundberg and Stearns

2019).[1] We advance this literature by clearly establishing causality of early labor market sorting in initiating and perpetuating gender inequality and norms in long-run careers. We provide new evidence for how men and women differentially adapt to the tension between career and the family (Goldin and Katz 2008, Goldin 2021): early career setbacks lead women to specialize in home production—as reflected by earlier childbearing and higher fertility—and they subsequently sort into lower-paying female-represented occupations that are more flexible and family friendly. Our analysis reveals that significant gender inequality can still emerge in a randomized lottery setup with embedded early-stage equality of opportunity. It demonstrates that policies for outcome-based gender equality cannot merely rely on leveling the starting playing field, but they should also target the way in which opportunities and choices evolve over the formative stage of the early career. Our analysis of mechanisms offers important guidance in that direction, strongly supporting the notion that closing the gender gap would have to involve the promotion of job flexibility to prevent women's excessive career penalties from their disproportionate family responsibilities (Bertrand et al. 2010, Goldin and Katz 2008, 2011, Goldin 2014, Kleven, Landais, and Søgaard 2019).

We additionally contribute to the classic labor economics research that has highlighted the potential importance of early-career stages in determining long-run life cycle trajectories. This work has considered the role of search and job mobility, human capital investments, as well as on-the-job learning and skill accumulation (see, e.g., Topel and Ward 1992, and reviews in Weiss 1986 and Rubinstein and Weiss 2006). We contribute to this broad line of research by providing a novel, purely randomized source of idiosyncratic variation for identifying the causal effects of early-career sorting.[2] This type of variation can be useful in other important economic questions. For example, with a focus on market design, Arora, Goff, and Hjort (2021) study how shifting the Norwegian system of medical internship allocation from lottery-based to market-based has impacted employer-employee matches. In their analysis of earnings five years after graduation, they focus on identifying employer-specific value-added across categories of Norwegian hospitals with some evidence pointing to potential lower returns for women in that time frame. In comparison, the focus in our analysis of the Danish setting is to identify the far-reaching causal impacts of early-career setbacks on individuals' long-run career outcomes, family choices, and their interaction. We

[1] Recent important studies in this active research on underlying channels investigate the role of job search and labor market preferences, social interactions, personality characteristics, and family obligations. These include, among others, Gneezy et al. (2003), Niederle and Vesterlund (2007), Bertrand et al. (2010), Buser et al. (2014), Azmat et al. (2016), Card et al. (2016), Field et al. (2016), Azmat and Ferrer (2017), Bursztyn et al. (2017), Caliendo et al. (2017), Buser and Yuan (2019), Cai et al. (2019), Iriberri and Rey-Biel (2019), Kleven et al. (2019a), Kleven et al. (2019b), Cheng (2020), Porter and Serra (2020), Ginther et al. (2020), Le Barbanchon et al. (2021), Exley and Kessler (2022), Cullen and Perez-Truglia (2023), and Cortés et al. (2023).

[2] Related but distinct work had studied aggregate variation, in terms of entering the labor market in a recession (see von Wachter 2020 for a review), in contrast to the idiosyncratic variation that we study here. The former identifies the effects of changes to the choice set that come from bad economic times. In comparison, the latter, with variation at the individual level, identifies the causal effects of making different choices within a given distribution of options, i.e., a given choice set. As such, these effects form a key input in an individual's optimization problem of early-career choices. In that sense, our analysis resembles the economics of education literature that uses idiosyncratic exogenous variation (e.g., based on grade cutoffs) to analyze the returns to different choices of field of study (e.g., Kirkeboen et al. 2016).

provide novel evidence that internship placements have important impacts on women's long-run career outcomes and that the gender divergence in sensitivity to career setbacks likely operate through the career-family trade-off and women's greater share of family responsibilities.

Finally, we speak to the mounting recent evidence that highlights geographic location in determining life-cycle outcomes, from education, to economic well-being, to health.[3] We contribute to this strand of the literature first by finding a causal determinant of the household's choice of geographic location in the long run—namely, individuals' very early-career labor market sorting. This choice directly affects the local labor market in which the household operates and the amenities available to the family. Second, our findings identify a pathway by which location can shape behavior and welfare. We find that women with the worst lotteries—who differ relative to the other experimental groups in their increased likelihood to intern in rural areas far from where universities are located—end up displaying much lower rates of obtaining important further education (in the form of a medical PhD). This, in turn, highlights the potential role of geographic sorting in the early career in determining major life-cycle human capital investments.

The rest of the paper proceeds as follows. Section 2 describes the institutional setting of physician training in Denmark and the data sources we use. Section 3 lays out preliminary facts about the Danish labor market for physicians and their career-family trade-offs to set up our causal analysis, and Section 4 describes our empirical framework. Section 5 provides our main analysis of how the internship placement affects long-run career outcomes and choices differentially for males and females, and Section 6 analyzes key potential channels for the gender divergence in treatment effects. Section 7 concludes.

2. Institutional Setting and Data

2.1. Physician Training in Denmark

We begin by describing the context of our analysis of physicians in Denmark and their post-graduate professional training. We first provide a broad overview of the course of their post-graduate experience, which captures the early stages of their careers, and we then describe the process of matching to medical internships in Denmark, which provides the grounds for our causal analysis.

Broad Overview and Timeline. The timeline for Danish physicians' training process is generally typical of other OECD countries.[4] Following medical school, graduating physicians begin the period of their *residency*, which represents a lengthy period of on-the-job training. The ultimate goal of the residency, which marks its ending point, is the choice of medical specialty. Among physicians, the choice of medical specialty represents occupational choice that determines their careers, so we accordingly put a particular

[3] See, for example, Chetty and Hendren (2018) and Finkelstein et al. (2016) for the U.S., and Damm and Dustmann (2014), Laird and Nielsen (2016), and Eckert et al. (2022) for our context of Denmark.

[4] For the institutional structure in EU countries, for example, see EU Council Directive 75/363/EEC.

focus on specialty choice in our causal analysis as a main long-run outcome of physicians' careers. Upon the beginning of the residency, about 92 percent of senior medical students have not yet decided which specialty to pursue, where 75 percent remain undecided during their internship (Jensen et al. 2013).

Appendix Figure A.1 illustrates the timeline of the various stages of the residency period that we now describe. The initial stage of the residency is the *internship*, which typically lasts one to one and a half years. The internship represents the entry-level labor market for physicians. It stands as physicians' first effective medical experience, and it determines their initial exposure to practical knowledge and career opportunities. The key institutional feature, which we exploit as the basis for our identification, is that a random lottery underlies the placement to internships. Upon completion of the internship, physicians are allowed to practice medicine independently, that is, without the supervision of a senior physician. All positions thereafter are matched in a standard competitive labor market. Toward the end of the internship during its last part, physicians engage in a process of job search as well as human capital investments that will steer their career paths. Specifically, interns apply for different *introductory positions*, which typically last one year each. Physicians must complete at least one such position within their future specialty of interest in order to then qualify for a *main position* within that medical specialty. Main positions represent the process of medical specialization, which lasts four years for general medicine and up to five years for more specialized positions. Main positions can be highly competitive. Hence, physicians' success in this stage is influenced by their choices up to that point in terms of training via introductory positions and further academic education via obtaining a *medical PhD degree*, which can assist in accessing higher-paying and more prestigious specialties (as we show below). Physicians typically apply for PhD programs in the last part of the internship or while working in introductory positions, and they enroll in PhD programs between the internship and main specialty position. A medical PhD in Denmark typically lasts three years. The main position is the last stage of the residency, where the choice of specialty represents an absorbing state in terms of physicians' long-run careers. Upon completion of this last stage, physicians receive their specialty license and continue on to their future independent careers in their chosen occupation. Appendix A gives an account of the effective timing for the residency stages, plotting PhD completion rates and then medical specialty completion rates over time since graduation. We discuss these patterns in Section 3.2 in turn, when we investigate physicians' timing of key labor market choices.

The Internship Program. Internship positions provide hands-on work experience, allowing physicians to accumulate practical knowledge and skills through learning-by-doing. The purpose of the internship is to bring the theoretical knowledge from medical school into clinical practice by having the intern integrated into the daily work routines of hospital departments. Internship programs are required to ensure that physicians accumulate medical expertise in all aspects of medical care including diagnostics, examinations, implementation of procedures, treatment protocols, and treatment of acute and chronic conditions (Danish National Board of Health 2009). Internships consist of bundles of half-year primary positions in hospitals followed by secondary positions in primary care. Internships are institutionally tied to geographic regions and their hospitals. The healthcare system in Denmark is organized such that Danish counties (with a total of 16) act as local healthcare markets, similar to Hospital Referral Regions (HRRs) in the U.S. Spatial distribution of entry-level jobs for physicians is typical of post-graduate medical training positions in other developed countries as well, such as the U.S., and is a main dimension by which medical training programs are categorized (see, e.g., Brotherton and Etzel 2018).

The Internship Lottery as a Source of Variation. The internship provides our source of exogenous variation in initial job market sorting of physicians. After medical school, each graduate is assigned to an internship position by being matched with a hospital department—that in practice represents a workplace—which is responsible for facilitating the internship educational program. Internship positions are periodically created by the Danish National Health Authority (NHA) with respect to national demand for healthcare professionals and to accommodate all graduating students. Prior to each graduation round (twice a year), medical schools report how many students are planned to graduate and the NHA guarantees to create a number of internship positions of at least that amount. Finally, the positions are designed to distribute proportionally across the local labor markets (i.e., counties) based on their population shares.

The key institutional feature we exploit for identification is that a *randomized lottery* governs the placement into internships. For every graduating cohort (twice a year in every March and October), a public notary performs a lottery that allocates a random number to each graduating student, which sets the ordering of the matching process for that cohort. We capture a graduating physician's relative position in the matching order by mapping a lottery number to its rank relative to the lottery numbers of the graduate's cohort. We refer to it as the "lottery rank."[5] Historically, the lottery-based allocation was introduced in 1984 to mitigate the problem of excess supply of young physicians around university hubs and physician shortages in rural areas (Danish Ministry of Health 1989), which is a broader concern and a common policy target across the U.S. and many other OECD countries (OECD 2012, Ono et al. 2014).

2.2. Patterns in Assignment to Internships

We use information on interns' binding pre-placement rankings of all local labor markets, which have been solicited among the earlier cohorts (who graduated prior to 2008) as part of the allocation process. We study the mapping between choice rankings and placements as a function of the lottery by plotting individuals' pre-placement ranking of the local labor market they were assigned to (where 1 denotes the highest priority) against the percentile rank of their lottery number within their graduating cohort.

---

[5] This normalization permits a comparison between individuals with bad lottery numbers and individuals with good lottery numbers across cohorts of different sizes. Appendix Table H.2. provides estimates that include graduation round fixed effects for robustness with virtually similar findings as we would expect.

Panel A of Figure 1 shows a few key patterns in this relationship, which are similar across gender as shown in panel B. First, as expected by design, there is a clear gradient such that graduates with higher lottery ranks (worse lottery numbers) are assigned to their lower-ranked priorities. Second, there is a clear non-linearity in the market-clearing pattern of the available slots in terms of graduates' preferences: there is a virtually flat region in the vicinity of the best lottery ranks and a steep slope in the vicinity of the worst lottery ranks. By the nature of the assignment process, students with the best lottery ranks are effectively unrestricted in their choices. As they are the ones who make the choices first, their highest priority options are still available when they make a choice, and they end up being assigned to their first priority. Then, as the lottery rank increases (that is, the draws worsen), the set of available choices increasingly narrows. As a result, those with the worst lottery numbers are most restricted in their early-career choices, and they end up making choices that are lowest on their priority list. These patterns guide our choice of research design. We will compare outcomes of a "control" group of individuals with the best lottery ranks whose choices are essentially unaffected by the lottery; a "treatment" group of individuals with the worst lottery ranks who are the most affected by the lottery; and a "middle" group who are affected to an intermediate degree.

The exact implementation of assignment to internships based on the lottery has changed over the years, but it has been continuously designed so that a better lottery number (of lower rank) guarantees a student a more favorable position in the allocation process. We leverage this simple yet powerful feature and pool all graduating cohorts to maximize power. We show in Appendix B.1 that the patterns of allocation of students to internships remains very similar over time. To give context, prior to 2008, the NHA first allocated students to counties based on the order of their lottery numbers in the primary step of the placement process. Graduating students compiled a list of priority over all the Danish counties following their assignment of lottery numbers, and they were matched with their highest-ranked county among the counties with remaining open positions when their time to make a choice has come. Later, in the secondary step of the assignment process, each county matched its assigned graduates with the internship positions that were created in that round across the county's hospital departments, based on student choices in the order of their initial lottery number. In 2008, when the system was digitized, the process simplified into a one combined step, where interns make a single county-hospital department choice in the order of their lottery number from the positions available nationally at the time they choose (known as random serial dictatorship, see Abdulkadiroğlu and Sönmez 1998). In Appendix B.2, we discuss potential strategic choice considerations that could result from the incentives embedded in these choice processes, and we investigate how they play out in practice. It is important to note that strategic behavior is not going to affect the validity of our identification of the effects of initial labor market sorting because our choice of research design rests on reduced form effects of the randomized lottery numbers. Still, we describe these aspects in the appendix as they are potentially informative for a further understanding of the empirical context.

2.3. Data

We combine several administrative datasets that contain information on all medical doctors in Denmark and their households. We use the education registers starting in 1980 to identify all students ever enrolled in a Danish medical school through 2020. Our analysis population for the experiment is identified using information starting from 2001 on the internship lotteries, which we obtained from the physical archives at the Danish National Health Authority and digitized. We link these records using person-level identifiers with the following register datasets on the data servers at Statistics Denmark as follows.

The authorization register provides us with information through 2022 on registrations of medical licenses and specializations, which capture occupational choice in our setting. The economic registers include administrative information on geographic location (through 2021), employers and employer-employee linkages (through 2020), total work compensation including wage earnings, employer contribution to retirement accounts, and self-employment income (through 2021), demographics, including age and gender (through 2021), and education registers (through 2022), including high-school GPA and information on higher education completion (through 2020). We are able to link households using spousal and parent-child linkages (through 2021) to study marriage and fertility choices. We note that, in the demographic registers, partnership in the cohabitation case is measured as two individuals of the opposite sex who live in the same address. We reduce noise in measuring partnership from situations in which cohabiting couples temporarily live apart (e.g., due to work or school) in the following way: if two individuals are registered as partners in both years $t-1$ and $t+1$ but not in year t, we recode them as being partners in year t as well. While the end points for the various datasets differ, our time frame allows us to look at causal effects across all outcomes with precision over a long horizon of at least 15 years following the lottery. When possible given the dataset for particular outcomes, we stretch the analysis horizon further out.

In addition, we obtained confidential information from the internship exit surveys, which are processed at restricted research servers at the University of Copenhagen. With permission from the official governmental body, the Regional Councils for Physicians' Post-Graduate Education (De Regionale Råd for Lægers Videreuddannelse), these data were obtained from a private IT company, Dansk Telemedicin A/S, which administers the data on all post-graduate educational positions for physicians in Denmark. From 2008, due to the digitization of the internship selection process, we are able to link physicians and their lottery numbers to their exit surveys and to the exact positions they held at the specific hospital departments. In these surveys, which are externally conducted by the government, interning physicians provide an assessment of their workplace experience. We provide details about the survey in Appendix E.

3. Preliminary Facts

3.1. Initial Conditions: Preferences and the First Stage

We begin the analysis by illustrating the comparability of initial conditions across gender, which is crucial for the interpretation of our causal analysis. We show that males and females have the same preferences over internship positions and that, conditional on lottery number, they make similar choices. Importantly, this implies that potential differences in longer-run effects across gender could not be attributed back to either differential preferences over entry-level positions or to differential first stages.

Preferences over Entry-Level Positions. We analyze preferences across gender by leveraging information on graduates' rankings of entry-level local labor markets and internship specialties. We proceed in two steps. First, we analyze a measure for market desirability that reveals students' location preferences through their lottery-based choices. We construct market rankings based on the average lottery rank of the interns who sort into it, separately for males and females, and compare across them. Panel A of Figure 2 provides this comparison. Each dot represents a local labor market, where the x-axis denotes male rankings and the y-axis denotes female rankings. We plot the fitted line and report its slope, where the benchmark of non-differential ranking by gender is one (the 45-degree line). Overall, the estimation is notably close to the benchmark case under the null that males and females have similar average priorities over the entry-level markets.[6] Second, we investigate graduates' occupational preferences in their entry-level jobs. We base our analysis on revealed preferences for specialty choices in internships. Within the primary positions in hospitals, interns can broadly choose between internal medicine and surgery, and, within the secondary positions in primary care, interns can choose between general medicine and psychiatry. For each gender, we split the sample by deciles according to lottery ranks. Then, for each of the two types of positions, we calculate over deciles the gender-specific cumulative probability of making a particular choice of specialty over the other. We plot in panels B and C of Figure 2 the gender-specific CDFs against one another, where the 45-degree line again serves as a benchmark when preferences are similar across gender. We do not find any systematic differences across gender in these choices either. Put together, the evidence strongly suggests that preferences over entry-level positions are similar across males and females and cannot be a source of gender differentials in long-run outcomes. Of course, it is still possible that preferences could change ex-post in response to the treatment.

Choices Conditional on Lottery Rank. How do the lottery ranks translate to exposure to internship characteristics? As is natural in real-life quasi-experiments, the initial match to internships represents a bundle, in our case in the form of location and career opportunities. We characterize treatment

[6] We reach a similar conclusion if we instead use the information we have for the earlier cohorts about students' binding pre-placement rankings of all local labor markets as reported in their priority lists (see panel C of Appendix Figure B.2).

aspects of the internship bundle based on two key underlying dimensions—geographic location and program quality—which we now discuss in turn. We use later cohorts from after the digitization of the system for whom we have detailed information on both of these dimensions of the internship allocation.

The first dimension we focus on—geographic location—comes from the motivation of the lottery-based policy to counteract students' reluctance to intern in rural areas. We calculate for each student the distance between their location of residence at the time of the lottery and their location of work at the time of the internship, which captures their "relocation distance." To put it in context, graduating students reside near the major university cities in which medical schools are located in Denmark (Aarhus, Copenhagen, and Odense). Hence, short relocation distances imply staying in the vicinity of the urban labor market where the student was educated, and long relocation distances typically imply placement in rural areas. Panel A of Figure 3 plots a graduating student's relocation distance against the student's lottery rank split by gender. The figure reveals a clear gradient: relocation distance for those with better lottery numbers (lower ranks) is significantly shorter than for those with worse lottery numbers (higher ranks). This mirrors the underlying motivation for the lottery-based system, revealing interns' distaste for locating in rural labor markets.

To better understand the features of the geographic local labor markets that the interns with worse lottery numbers sort into, Appendix Table B.2 compares rural versus urban localities. Rural municipalities have populations that are less educated, sicker, and more reliant on welfare. These municipalities have worse economic conditions and amenities, as manifested by income, home prices, tax revenues, and local recreational expenditure. In terms of traditional gender family roles, in rural areas females are much more likely to take more parental leave with the opposite pattern for males. In terms of local representation that could further indicate gender norms, the share of elected officials who are female is lower in rural areas. These are consistent with general priors that rural areas may be more gender stereotypical overall. Finally, important for their future careers, rural labor markets are much less likely to offer interns work experience within a university (or teaching) hospital. University hospitals, typically located in the vicinity of larger urban areas, are well known to be the institutions where skill-intensive and highly-specialized procedures are performed, state-of-the-art technologies are first adopted, and innovative medical research is conducted. By definition, university hospitals aim to provide the highest quality of on-the-job training to new physicians. Moreover, since key players in the medical field often hold positions in these hospitals, interning in a university hospital could provide more favorable exposure to professional networks. Appendix Table B.2 uses the administrative registers to illustrate these points of the differences between the two types of hospitals.[7]

[7] We have highlighted two features of the market composite that seem to us to stand out—degree of rurality and affiliation with university hospitals. Of course, other features beyond these two could be (and likely are) a part of the local labor market composite.

The second dimension of the treatment that we focus on is the ranked quality of the educational program itself, i.e., the specific workplace/position, as reported by interns in the exit surveys. We use the ranking of a position's overall assessment, where graduates are asked about their evaluation of the educational experience in terms of the program's effort, quality of training, and their own professional development (see Appendix E.1).[8] The value of the ranked quality via the reported experiences of interns is that it captures measures that are not directly observed. Still, it is useful to corroborate these assessments against external measures. We have data from external inspections that the NHA conducts to assess the quality of the educational programs, which are carried out by independent inspectors. Appendix E.2 shows that these external rankings are highly predictive of exit-survey rankings. In panel B of Figure 3, we then investigate the quality of the internship a graduating student has been assigned to against the student's lottery rank split by gender. The figure reveals a clear gradient, where graduating students with better lottery ranks are significantly more likely to intern in the highest quality positions.

Aggregate First Stage and Comparability across Gender. We proceed with our sample split into the three groups of best lottery ranks (the "control" group), middle range lottery ranks (the "middle" group), and the worst lottery ranks (the "treatment" group). Panel C of Figure 3 plots the averages of relocation distance and internship quality together, for each of our experimental groups. This figure bears similarities to an "offer curve" if the internship bundle is to be thought of as a consumption bundle: the curve maps individuals' choice of a two-dimensional bundle for an increasingly narrow choice set.[9]

We first see, as expected, that interns in the "control" group, for whom there are virtually no restrictions on choice, choose internships that are closest to their medical school's urban hub and in positions that are higher ranked. Interns in the "treatment" group, who are most restricted, suffer on both margins. They are forced by the lottery to end up choosing remaining positions that are both located in remote geographic locations and are of lower-ranked quality. Finally, choices made by interns in the "middle" group reveal further information on preferences. On average, the "middle" group is on par with the "control" group in terms of distance, whereas the "middle" group is on par with the "treatment" group in terms of a position's ranked quality. This further illustrates that physicians prefer to remain closer to their university's hub. Moreover, the fact that the "treatment" and "middle" groups are affected differently in terms of the career setbacks they experience will turn out important in the investigation of mechanisms.

Panel D of Figure 3 visualizes that the first stage patterns are similar across gender by showing that the gender-specific offer curves are of the same shape. Appendix Table B.1 formalizes this statement by testing differentials across male and female interns on both dimensions across all three experimental groups,

[8] We use the leave-one-out mean (to avoid mechanical correlations with one's own experiences) of the overall evaluation of graduates who interned in a given workplace, which we normalize by the mean and standard deviation to create a z-score.
[9] Corresponding figures for a finer split of lottery ranks are provided in Appendix Figure B.2.

systemically showing no differentials. This implies that potential differences in longer-run effects across gender could not be attributed back to differential first stages along the studied dimensions.

3.2. Descriptives: Choice Timing and Returns to Investment

Before proceeding to our causal analysis, we take a final descriptive step and investigate the patterns that underlie the main choices and outcomes that we study in both the labor market and the marriage market. This establishes the setting of our analysis of the labor market for physicians in Denmark. It will accordingly help guide the causal analysis of the effects of early careers on long-run outcomes and will provide the motivation for the later investigation of the potential mechanisms that are at play—specifically, women's career-family trade-off.

Labor Market Choices. We look at choices in the chronological order of the residency period (as illustrated in Appendix Figure A.1), starting with human capital investments and then transitioning to the long-run career choice of medical specialty. We display these outcomes as a function of years since medical school, running from year 0 to year 25 and split by gender.

For our outcome of human capital investment—pursuing a medical PhD— we see that, by year 15 after graduation, 27 percent of women and 29 percent of men have obtained a medical PhD (Appendix Figure A.2). Importantly for our analysis, we find that completion rates stabilize by that time, allowing us to identify "steady-state" causal effects in our analysis horizon of 15 year after graduation. For the central choice of long-run careers, Appendix Figure A.3 shows completion rates of medical specialties. The steep gradient over the years after graduation slows down at the vicinity of year 15, and it stabilizes a few years later than PhD completion (as expected by the post-graduation timeline in Appendix A.1 and as those who pursue a PhD consequently have longer residencies). For this reason, in our causal analysis of specialty completion rates, we extend the horizon as allowed by our data up to year 18. For both genders, we see that 10 years after graduating from medical school around 30 percent have completed their specialization, with completion rates that steeply rise to 80 percent by year 15. By year 20, about 90 percent have completed their specialty, with the residual 10 percent never specializing and typically working in the industry.

Next, we look at the association between obtaining a medical PhD and the type of specialties physicians sort into. Indeed, 42 percent of medical students who report considering pursuing a PhD state qualifying for their desired specialty as the main reason (Jensen et al. 2013). Given our focus on gender, we classify medical specialties—which represent occupations in our setting—based on the share of females within a specialty relative to their overall proportion. "Female-represented" specialties are defined as specialties with a female share that is higher than this proportion, and "male-represented" specialties are defined as specialties with a female share that is lower than this proportion (see Appendix Table D.1 for a list of specialties and their groupings). Appendix Figure A.4 plots time to specialization for both classes of specialties, split by whether the physician holds a medical PhD. It clearly shows that, in the long run, obtaining a PhD is associated with higher rates of specialization in male-represented specialties. Importantly, male-represented specialties are higher-paying and considered more prestigious, as we show and discuss below. As for how further education of a PhD ultimately translates to higher earnings via occupational and labor supply choices, Appendix Figure F.1 uses the population-level register to show associations for the potential economic returns from this investment. The investment pattern bears a classic labor economics shape: obtaining a medical PhD is associated with early lifetime investments in terms of foregone earnings and with high returns later in the life cycle. This is a lengthy investment whose returns manifest only late, implying that earnings differentials from the choice to obtain a medical PhD require analysis over an extended time horizon, which will guide our analysis of the effects on earnings.

Family Choices. Family-related choices and career-related choices naturally intertwine. The post-graduation stage represents formative years with respect to family formation (Goldin and Katz 2008), which is reflected in our setting by the graduating physicians' average age of 27.5 at baseline (shown in Appendix Table C.2). Appendix Figure A.5 first plots the evolution of partnership in levels and changes. We see that about half of physicians graduate being partnered (in panel A), and that the highest rates of changes to partnership status post-graduation occur in the immediate years around the internship period (in panel B).

Central to our setting, we look at the evolution of fertility choices over the post-graduation years. Here, we naturally split our sample into two groups of partnered individuals and single individuals based on partnership status in the baseline period. Partnered individuals enter the post-graduation period as a joint unit of two partners who can immediately make family planning choices, whereas single individuals still face the marriage market phase before they can start their families.

Appendix Figure A.6 plots fertility timing for both males and females, splitting individuals by their initial partnership status within each gender. Several important patterns stand out. First, partnered physicians display clear timing of fertility already in the few years leading to graduation, with clear large spikes in the few years post-graduation. Second, among single physicians, we see the intuitive delay in fertility as they search for partners, where their fertility rates start spiking 3 years out for both genders. This underscores that our causal analysis should separately consider single and partnered graduates as they can be fundamentally differentially affected by their early career experiences in terms of both their family outcomes and their labor market outcomes. Third, fertility decisions reach their plateau at the vicinity of 15 years post-graduation. This allows us to investigate both fertility timing effects and the long-run steady-state family choices of physicians in our causal analysis that runs up to year 15 and above.

The Career-Family Tradeoff. Key to our analysis is the important work on gender gaps that points out the high value placed on long hours and job continuity as a central determinant of gender differentials in high-powered professions (Goldin 2014, Goldin and Katz 2011). In such professions—specifically among physicians—earnings are non-linear in hours worked, which manifests, e.g., in higher compensation outside of regular hours (overtime and weekends). This structure of remuneration, which similarly applies to physicians in Denmark (Lorentzen 2024), imposes heavy penalties on employees who place a higher value on job flexibility (in terms of number of hours, precise times, predictability, and ability to schedule one's own hours). The costs of workplace flexibility include economic penalties to labor supply behavior that is more compatible with having a family, such as job interruptions, short hours, and work flexibility during the day and over the week. Given biological differences in childbearing and social norms, women have a greater desire for workplace flexibility from greater family responsibilities (Goldin and Katz 2011). This notion has been causally corroborated among physicians (first by Wasserman 2023 in the U.S. and then by Lorentzen 2024 in our context of Denmark), showing that reforms that limit work-week hours lead women to choose specialties that previously had greater time demands at higher rates. Wasserman (2023) further finds in survey data on stated preferences of U.S. medical students that female graduates shifted their tastes toward previously time-intensive specialties in response to the reform. For our context of Denmark, we complement the analysis with a publicly available survey conducted among Nordic physicians that shows similar attitudes. In Appendix Figure D.1, we find that physicians of both genders believe that lower shift burden is an important determinant of why some specialties are more female-represented.

These underlying conditions—gender differentials in the value of job flexibility and physicians' remuneration structure—imply two patterns that lead to gender pay gaps: first, women will sort at lower rates into specialties that have more excessive time demands; and, second, specialties with more excessive time demands will have meaningfully higher pay. Indeed, the two patters hold closely in our setting.[10] First, as an indicator for the time demand of physicians, we follow Goldin and Katz (2011) and characterize whether the specialty has irregular, on-call, or emergency hours on a more regular basis. We classify specialties with our prior grouping of female-represented and male-represented specialties, and, for each class of specialties, we calculate the probability that a medical encounter occurs during weekend hours. Appendix Figure D.2 clearly shows the probability a medical encounter takes place over the weekend is an order of magnitude higher in male-represented specialties compared to female-represented specialties (16 pp compared to 3 pp). Given the way we classified specialties, it is indeed the case that women sort at lower rates into specialties that have more excessive time demands in our setting. Second, it is also the case that these specialties with more excessive time demands have higher pay. Appendix Figure F.1 plots the life-cycle earnings trajectories for the two classes of occupations, and it provides a clear visualization that male specialties provide higher pay than female specialties for both men and women.

[10] Wasserman (2023) shows how these patterns hold for the U.S. case.

As an implication of these patterns, we find an important career-family trade-off between specialty choice and family responsibilities in Appendix Figure A.7. In panel A, we plot the evolution of the number of children over years since graduation, separately for physicians who specialize in female-represented specialties and in male-represented specialties. We find a visually clear pattern that those who specialize in male-represented specialties trade off family choices, as reflected in both delayed childbearing and a lower number of children in the long run. Splitting this plot by gender (in panels B-C), we find that this trade-off is indeed accrued almost in its entirety to women. This is naturally in line with the norm that men have less family responsibilities, so they can incur the disamenity of excessive time demand that they care little about regardless of their family structure (Goldin and Katz 2011). This underlying gender discrepancy in the interplay between career choices and the family will therefore motivate our central hypothesis of the career-family trade-off as a mechanism for the gender differentials that we find in the effects of early career setbacks. Our hypothesis is furthermore motivated by the important work that has shown that family responsibilities, specifically from the arrival of children, impose excessively higher penalties on the career advancement of women relative to men (e.g., Goldin and Katz 2008, Bertrand et al. 2010, Kleven, Landais, and Søgaard 2019).

4. Empirical Framework

4.1. Verification of Lottery

As the basis for our empirical framework, we first establish the validity of the lottery in terms of random assignment. In Appendix Table C.1, we run specifications that regress the graduating physicians' lottery rank on baseline characteristics available in our data. These include gender, age, an indicator for having a partner, number of children, high school GPA rank, an indicator for residing in a rural municipality, and an indicator for having an employment at a university hospital. Consistent with random assignment, we find that these regressions have no predictive power. This is the case whether we test the significance of the coefficients individually or jointly. We also run the corresponding specifications separately for males and females with similar conclusions. This sets the grounds for our research design that we turn to next.

4.2. Research Design

We employ a straightforward design based on the randomized lottery, where we compare outcomes of a treatment group to outcomes of a control group. As natural experimental groups, we define the "control" group to be individuals with the best lottery ranks (below a certain lower cutoff rank), as we have shown they are essentially unaffected by the lottery; and we define the "treatment" group to be individuals with the worst lottery ranks (above a certain upper cutoff rank), as we have shown they are the most affected by the lottery. Our choice of research design provides an intuitive empirical framework with treatment effect coefficients that are directly interpretable. It also maximizes the differential treatment intensity across the differentially affected experimental groups since it compares individuals who are most restricted to those who are least restricted in their choices. It also does not impose functional form assumptions on the underlying relationship between outcomes and lottery ranks. Specifically, it does not use the common linear in rank specification, where linearity seems less appropriate in our setting (given the patterns in Figure 1).

In constructing our experimental groups, we need to make a choice of upper and lower lottery rank thresholds, which we do in the following way. First, to keep the experimental groups balanced with similar size, we use symmetric thresholds from above and below. Second, we pivot the analysis around the 30 percent most treated and least treated, i.e., with cutoff ranks 0.30 from below and 0.70 from above (as illustrated by the vertical lines in Figure 1), and we vary this bandwidth from 20 to 40 percent in Appendix Table H.1. This choice trades off increased power from higher treatment intensity with decreased precision from reducing sample sizes, which is the reason we investigate a broad range of 20 pp in lottery ranks from above and below in our robustness analysis of the bandwidth choice.

While we discuss our main results as the comparison between the treatment group and the control group, we systematically report estimates for the "middle" group of graduates with lottery ranks in the intermediate range for completeness and for shedding light on mechanisms. We provide plots for the full dynamics in outcomes from year 0 to year 15 and above for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). At the bottom of each plot, we report the full vector of estimates for the treatment effects along with their standard errors and with counterfactual levels that pertain to the control group. On top of clearly displaying the treatment effects, these plots allow us to assess baseline gender gaps among the control groups of graduates who start with the best shock. In the appendix, we also provide an alternative specification for the long-run effects of the lottery for our main outcomes. Appendix Figure H.1 provides figures that non-parametrically plot the year 10 and year 15 outcomes against lottery rank deciles. We additionally report on these appendix figures the coefficients from the corresponding linear in rank regressions on the underlying individual-level data.

Estimating Equation. With this design, we identify the causal effects of the internship lottery using the following dynamic estimating equation:

$$y_{it} = \sum_{\tau=0}^{15} I_\tau \times \alpha_\tau + \sum_{\tau=0}^{15} I_\tau \times Treat_i \times \beta_\tau + \varepsilon_{it}. \tag{1}$$

In this specification, y_{it} is the outcome of interest for individual i at time t; τ is the year relative to the baseline period of the last full calendar year in medical school, normalized to year 0 (where year 1 represents the calendar year when the lottery takes place); and I_τ is a vector of indicators of time relative to the baseline period in years. The variable $Treat_i$ is an indicator for being in the treatment group or in the control group.[11] We cluster standard errors at the individual level. Our parameters of interest are the elements of β_τ, which estimate the dynamic causal effects of the lottery over 15 years after graduation. We summarize average treatment effects using β from the following specification run over later periods (years 6-15):

(2) $$y_{it} = \alpha + \beta \times Treat_i + \varepsilon_{it}.$$

Analysis Sample. Appendix Table C.2 describes our analysis sample and provides summary statistics for our treatment, middle, and control groups. Overall, the groups together consist of 10,073 physicians. At the baseline period, our subjects are on average 27.5 years old and about half of them have a partner. Approximately 60 percent are female, with 3,970 male physicians and 6,103 females. Summary statistics that split the sample by gender are also provided in the same Appendix Table.

5. Treatment Effects

We now turn to our main analysis and investigate how the internship placement affects long-run career outcomes and choices differentially for males and females.

5.1. Location

We first consider the dynamics in a household's geographic sorting given that the internship allocation system is strongly governed by location. Geographic sorting has direct effects on the amenities available to individuals and their families as well as on the local labor market in which the physicians operate. In the context of aggregate amenities, we will analyze sorting into a rural community, and, in the context of features of the labor market, we will analyze physicians' affiliation with a university hospital.

Panel A of Figure 4 illustrates the dynamic effects of the lottery on the probability of sorting into a rural municipality throughout our entire analysis window. To reiterate, this and subsequent figures provide the full dynamics of a given outcome from year 0 to at least year 15 after graduation for our three experimental groups, and we report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors and counterfactual levels from the control group. With location outcomes, the early years mechanically capture the first stage effect on the internship placement, particularly in year 2, which is the period when the internship placement is in full effect in the annual registers (that report location at the end of the calendar year). We see that receiving the worst lottery ranks leads to a large increase in the probability of interning in a rural locality across gender (10.9 pp for men on a baseline of 8.8 pp and 8.4 pp for women on a baseline of 10.2 pp). However, focusing on the longer run, we find lingering effects only among women. Specifically, 15 years after the treatment—long after the internship itself—women in the

[11] The treatment effects on the middle group for our main outcomes are reported in Appendix Table H.1, where we extend specification (1) to include the middle group as an additional category of treatment.

treatment group are 4.9 pp more likely to locate in a rural municipality relative to a counterfactual of 9.9 pp among the control group. We find no significant effects for males.

A major characteristic of the physician local labor market is whether there is a university hospital nearby as a potential employer, since prestigious and higher-paying positions in the medical field are often attached to university hospitals. Appendix Figure F.1 descriptively illustrates the large economic returns from having a job that is affiliated with a university hospital in the long run among both men and women. Panel B of Figure 4 then studies causal effects on working at a university hospital as an outcome over our analysis horizon. Again, the similar differential across gender in the immediate years (-35.4 pp for males and -37.8 pp for females in year 2) reflects that graduates who receive the worst lottery ranks are much less likely to complete an internship at a university hospital. In the long run, however, only females again are less likely to hold positions at university hospitals along with the returns that they bring. In year 15, the effect amounts to a decrease of 6.5 pp on a counterfactual level of 48.1 pp.

5.2. Human Capital Investment

Next, we study the classic human capital investment of obtaining a PhD in Figure 5. The analysis for this outcome begins in year 5 as it is the period at which PhD completion begins to materialize following graduation from medical school. The findings reveal stark gender differentials. Males do not experience any adverse effects as a result of the treatment. However, females in the treatment group have significantly lower propensity to make this human capital investment. By year 15 (when investments reach the vicinity of their plateau as we have shown in Section 3.2), females' lower investment rate amounts to a large decline of 5.2 pp in obtaining a PhD on a counterfactual of 29.7 pp.

With our focus on occupational choice, this finding also speaks to gender-biased sorting into scientific careers. Indeed, gender inequality in science is a well-known phenomenon in the developed world (Holman et al. 2018, Huang et al. 2020). We calculate across our entire experimental sample that the male-female gap in holding a medical PhD 15 years after graduation is 5.1 pp. This implies that the treatment effect can account for 31 percent of the observed gap.[12] These effects are attributed to variation in the short internship period alone (out of the lengthy process of becoming a physician), underscoring just how important labor market sorting in the very early stages can be for women's long-run careers.

5.3. Occupational Choice

Finally, we investigate our central outcome of occupational choice that determines physicians' long-term careers and work-life balance. Figure 6 first investigates general patterns of time to medical specialization by plotting an indicator for whether the graduate has completed specialty requirements (in

[12] These assessments use the fact that those are individuals with the worst lottery numbers included in our treatment group who are adversely affected on this margin (as shown in Figure 5). As they compose 30 percent of the sample, our calculation is performed as follows: $(5.2 \times 0.3)/5.1 = 0.31$.

panel A). We find no discernable effects among men but clear effects on specialization timing for women. Whereas women in all experimental groups converge to the same specialization rate by the end of our analysis window, women in the treatment group specialize earlier (as compared to both the middle and the treatment groups), and they accordingly spend less time accumulating human capital through training.

Importantly, panel B of Figure 7 shows that the specialties that treated women sort into at higher rates are exactly the lower-paying female-represented specialties. Indeed, by year 15, women in the treatment group are 5.7 pp more likely to specialize in female-represented medical specialties on a counterfactual of 57.5 pp among women in the control group. As we have seen that overall completion rates converge by the end of the analysis horizon among women in all experimental groups, we see a corresponding lower sorting into higher-paying occupations. Panel C of Figure 7 shows that, by year 15, women in the treatment group are 5.3 pp less likely to sort into a male-represented occupations compared to the control group, whose baseline level is 16.6 pp. Our data on occupational choice allow us to push the horizon up to year 18. The long-run career sorting patterns are clear: women who experience early labor market setbacks are diverted away from the higher-paying male-represented specialties and show large increased propensities to follow careers in the lower-paying female-represented specialties.

5.4. Long-Run Earnings Gap

We have found that women in the treatment group, as opposed to men, end up forgoing important human capital investments and sort into lower-paying stereotypical career paths at higher rates. We conclude this section by aggregating our findings on labor market choices into the long-run implications for earnings and gender gaps.

In our context, while we have experimental data over a long horizon, we expect the effects on the career-defining labor market choices that we studied to translate into effects on earnings only later. As we have seen in Appendix Figure F.1, the returns to major human capital investments, specifically obtaining a medical PhD, materialize only in the very long run (in fact, starting just after year 15). Current earnings are therefore insufficient for studying physicians' long-run relative positions in the labor market in the analysis of their early careers, which is commonly known as the "life-cycle bias" (see, e.g., Black and Devereux 2011). Indeed, Figure 7 displays no treatment effects on earnings over an analysis horizon of 17 years.

Accordingly, we use the "surrogate index" method (Athey et al. 2019) as a solution to the common challenge in estimating longer-term impacts of treatments in cases where outcomes of interest are observed with a long delay. The method combines several intermediate outcomes into the "surrogate index," which is the predicted value of the longer-term outcome given the intermediate outcomes (the "surrogates") based on long-run observational data.[13] We study the average treatment effect on the surrogate index for earnings

[13] Athey et al. (2019) show that the average treatment effect on the surrogate index equals the treatment effect on the long-term outcome under the assumption that the long-term outcome is independent of the treatment conditional on the surrogate index, which

predictions over the course of 35 years, comparing across our experimental groups. Appendix F.1 describes the full details of the earnings predictions implementation.

Figure 8 presents the results in several ways. First, in panel A, we provide the full predicted earnings trajectories separately for men and women, where in each plot we provide three lines: one for the control group, one for the middle group, and one for the treatment group. Second, in panel B, we provide the predicted treatment effects on earnings for both men and women, capturing the level gaps from panel A, along with their 90-percent and 95-percent confidence intervals (to account for the decline in precision in our analysis over decades).[14] Third, in panel C, we plot the predicted baseline gender earnings gap (estimated using the control group) as well as the additional earnings gap among physicians in the treatment group caused by the experiment as a share of the baseline earnings gap (calculated using the predicted treatment effects).

Two key patterns stand out. First, we find that there is a gender gap that grows over time, echoing the typical pattern in high-skill professions (Bertrand, Goldin, and Katz 2010, Goldin 2014, Blau and Kahn 2017, Goldin et al. 2017, Juhn and McCue 2017, Kleven, Landais, and Søgaard 2019). Second, in terms of treatment effects, there are no predicted long-run impacts on earnings for males as expected, reflecting the fact that they displayed no effects on the entire array of outcomes. In contrast, we find a widening effect for females. As there is no effect for males, the effects on females that grow over time translate into a growing gender gap in earnings that is causally attributable to the lottery, leading to gender gaps in earnings that are even larger for women in the treatment group. Our variation in physicians' very first jobs increases the gender earnings gap by 10-15 percent over the decades after graduation.

6. Mechanisms

Our findings clearly and systematically show that women are much more sensitive to the initial bad shock of early labor market sorting. To understand why, we now explore key potential channels for this gender divergence. Analyzing fertility choices, along with search and mobility patterns, we find as a key explanation that the initial job-match bundle of location and career opportunities is important for women's career-family trade-off. Being at the age at which they get married and form their families, the initial setback leads women to prioritize the family, as mirrored by earlier childbearing and increased completed fertility. Subsequently, they end up sorting at higher rates into careers in female-represented specialties, which offer more time flexible jobs but with significantly lower pay. Given the cultural norms of women as secondary movers, we find further evidence that affected women prioritize their families over job search in that they

forms the "surrogacy condition." This method improves on previously suggested methods that use only one surrogate outcome in that it weakens the standard surrogacy condition for a single variable, with the notion that there is a greater likelihood that a set of intermediate outcomes could together satisfy the surrogacy condition.

[14] Predicted effects for the middle group compared to the control group are reported in Appendix Figure F.3.

are more likely to remain stuck in the original location of their partners. Men, however, do not display any of these career-family trade-offs. Overall, the initial conditions lead women to specialize in home production, which then leads to larger differences in long-term career outcomes across men and women.

6.1. Marriage and Fertility

As we described in Section 3.2, we split the sample in the analysis of family choices by whether the individual was partnered or single at baseline. We begin by investigating marriage patterns. Across all subsamples, we find no systematic effects of the lottery on the likelihood of partnership (see Figure 9).

Next, we investigate our central margin of fertility choices by studying the evolution of the number of children over the course of 15 years after the initial job placement (in Figure 10). We first note that there are no effects on the time path of the number of children for both genders among individuals who start their careers when they are already partnered. Single men also do not display any discernable effects.

For women who are single at baseline, however, we find a clear and systematic pattern of impacts on fertility. First, single women in the treatment group start showing increased fertility compared to the control group in the years after graduation, which widens over several years and stabilizes thereafter. This reveals meaningful retiming patterns, so that single women who experience early career setbacks display clear effects of earlier childbearing along with their lower engagement in career enhancing activities (of further education or on-the-job specialty training).

Second, we see that over the entire course of our analysis, fertility of women in the control group is systematically below and never catches up to fertility of women in the treatment group. Thus, the differential patterns seem to represent not only earlier childbearing but also differential completed fertility within our analysis horizon. To test the robustness of this hypothesis, we want to alleviate concerns that the long-run effect may be driven by a few large families among women in the treatment group (specifically as in later years we lose precision from narrowing sample sizes). We report in Appendix Figure G.3 the dynamics of fertility when we winsorize the number of children above their 99th percentile of 3 children (that is, at having 4 children or more). Our conclusions remain the same, as we indeed see that the persistent gaps go through all the way. Moreover, we have seen that by the end of our analysis period of 17 years fertility rates are very low (in Appendix Figure A.6), and we accordingly find that all our experimental groups reach a fertility plateau by then (in Figure 10 and Appendix Figure G.3). Finally, we have seen that women in the control group are more likely to sort into male-represented specialties where women tend to have fewer children, so that their number of children seems even less likely to catch up after our long analysis horizon. To put the magnitudes in context, the causal estimate in year 15 implies that at least 13 percent of women in the control group would need to have 1 additional child each in order to catch up with the treatment group. These patterns all provide strong support for a considerable impact of the lottery on completed fertility. We see similar effects on fertility among women in the middle group, which reveals important information that we come back to below when we describe their more nuanced dynamics.

With the differential effect on fertility across partnered and single individuals, we rerun our entire analysis of labor market outcomes from Section 5, where we now split individuals by their baseline partnership status. The results are all reported in Appendix Figure G.1. As before, there are no effects on men, whether they were partnered or not. For women, we find that the bulk of the negative effects on labor market outcomes is concentrated among women who have entered the labor market single and have also faced the process of forming a new household. The time patterns show that as they begin having higher fertility compared to single women in the control group, single women in the treatment group then sort into the more flexible jobs in female-represented specialties at higher rates. The predicted impacts on long-run earnings are accordingly larger among them. This bolsters the hypothesis of the career-family trade-off as a main channel for the long-run career outcomes, as it is those who exhibit effects on family related choices who consequently bear the majority of the adverse career effects of the initial placement.

The dynamics of women in the middle group provides further support. Recall that women in this group suffer career setbacks in terms of placements in lower-quality positions, but they remain close to their university hospital of origin during the internship (Figure 3). Earlier we found that they obtain PhD degrees at rates similar to the control group (Figure 5), likely facilitated by their proximity to their university of origin. Yet, we see now that their career setbacks still lead them to home specialization, as reflected in earlier and higher fertility (Figure 10), which ultimately diverts them back to the lower-paying female-represented specialties at higher rates (Figure 6). Their specialization patterns are delayed by their further education (as illustrated in panel A of Appendix Figure A.4), so their occupational choices materialize only later when they converge to those of women in the treatment group (while women in the control group go in the direction of catching up with men as seen in panel C of Figure 6). These patterns point to the notion that even human capital investment does not manage to shield female physicians from the career costs of family responsibilities. Consistent with this hypothesis, Appendix Figure A.8 shows that among women with a medical PhD, the likelihood of going into male-represented specialty is still much lower for those with higher fertility.

Finally, we find further support for this view in our predictions for earnings trajectories among women in the middle group (in panel H of Appendix Figure G.1). Consider first women in the middle group who are single at baseline, who are likewise the ones driving the fertility and occupational choice effects for this group. Despite their PhD completion rates that are similar to women in the control group, their predicted earnings converge to those of women in the treatment group with worst lottery ranks. In contrast, the earnings predictions of women who were partnered at baseline—and do not display changes in fertility—converge to those of women in the control group of best lottery ranks. In that sense, it raises the possibility that women with greater family responsibilities might "overinvest" in their education (Chen and Chevalier 2012). Overall, the dynamics in outcomes among the middle group further points out that what crucially governs female professionals' careers in the impacts of setbacks is indeed family responsibilities, even in the presence of further lengthy human capital investments.

6.2. Search and Mobility

The lack of any long run effects on men suggests they may engage in career-oriented actions in response to the lottery that may allow them to mitigate the potential adverse effects. A particular course of action we consider is search and mobility, which have been suggested in the classic work of Topel and Ward (1992) as a key driver of males' careers. Accordingly, we are interested in testing the conjecture that males and females may display different search behavior in the labor market in response to initial placements. Recently studied margins in the context of differential search behavior by gender—in relation to women's higher value of time flexibility, greater family responsibilities, and career penalties—are commuting and geographic relocation (Le Barbanchon et al. 2021; Caldwell and Danieli 2024).

Accordingly, in panel A of Table 1 we first analyze the average effect of the lottery on commuting distances. The evidence shows that, in contrast to women, men in the treatment group commute further, thus widening their scope of labor market opportunities.[15] This is consistent with the notion that higher willingness to commute further away from home could help mitigate males' potential adverse effects from the lottery through differential search behavior. The second margin we investigate in relation to search is migration across labor markets. In our context of the tension between career and the family, we are particularly interested in studying whether physicians move their families around. To do so, we study as an outcome the physician's propensity to reside within the baseline location of their spouse. In panel B of Table 1, we find a meaningfully decreased propensity among households of male physicians, revealing that men are more mobile and more likely to migrate their households across labor markets. There is no detectable effect among women.[16] This is consistent with the cultural norm that women are secondary movers and that men, as primary movers, do not have the same career-family trade-offs and can move their families. These search patterns overall again point to the notion that, following the early setback, women prioritize their family and men prioritize their careers in their choices.

[15] The effect on men is on the order of 9 percent, with a treatment effect of 2.3 kilometers (km) on a baseline of 26.3 km. In their French setting, Le Barbanchon et al. (2021) report an average commuting distance of 21.3 km for men, which is 12 percent lower for women. In our setting of Danish physicians, commuting distances are of a similar order of magnitude as in the French setting, but with men and women in the control group displaying comparability (26.3 km for men and 26 km for women, see Table 1).

[16] It is worth noting that it is not the case that women are less likely to move from the exact location of their initial placement (see Appendix Table G.1). Instead, along with the findings that they are more likely to live in rural communities in the long run and that they do not migrate their families, the patterns suggest women in the treatment group remain stuck in the rural community in which they met their husbands. Our findings are then consistent with the fact that they formed their households in communities that display a higher degree of traditional gender norms as compared to urban locations (as seen in Appendix Table B.2).

6.3. Alternative Explanations

Sensitivity to Workplace Characteristics and Potential Unequal Treatment. We assess the role of employer-side factors by investigating the potential role for workplace characteristics. The motivation for this analysis is the hypothesis that being unequally treated by the same entry-level employers could alter the career course of graduates differentially for males and females. We note that discrimination could manifest in this part of the analysis, but we refrain from giving such interpretation to our findings in the absence of variation and data that could speak to discrimination explicitly.

We analyze whether men and women display differential sensitivity—in terms of their own post-internship market-based placements—when they are exposed to internship employers who do better or worse in placing their interns in subsequent positions. The specific way in which we do this is by identifying hospital departments as unique employers and linking interns within a current employer to the interns' next employer. An important dimension of the next position (as described earlier and for which we have consistent information to link across individuals and employers) is whether the next position is held at a university hospital. For each intern of a given gender, we calculate the leave-out-mean of how well their internship employer places its interns of the same gender in a university hospital later on. We refer to this measure as "employer intensity." We then study interns' "sensitivity" to "employer intensity" by regressing one's own probability of being employed at a university hospital in their next position on the intensity of their internship employer. The benchmark for the slope of full passthrough is one.

Table 2 reports the results. We first show, as before, that males and females are similarly affected in the first stage (with an effect that if anything is slightly higher for males), here in terms of how the lottery leads them to intern for employers who do worse in future placements (panel A). In comparison, however, we find that women's outcomes display a higher sensitivity than men's outcomes to their employer's placement track record (economically and statistically), with an effect that is 30 percent larger for women (panel B). That is, we find similar exposure but differential sensitivity, so that the market penalizes women more than it does men for having interned in similarly bad placements. These patterns suggest that the gender divergence could be linked to unequal treatment by the market and illustrate how similar opportunities can still lead in practice to gender disparities in the labor market. Motivated by the literature,[17] one particular workplace characteristic in our context of gender, which could act as a mediator in the current analysis, is exposure to female role models. Indeed, we find that graduates with worse lottery ranks end up in internships that are much less likely to have supervisors or program department chairs who are female (panel C of Table 2).

[17] See, for example, Bettinger and Long (2005), Carrell et al. (2010), Blau et al. (2010), Dennehy and Dasgupta (2017), Kofoed and McGovney (2019), Porter and Serra (2020), and Ginther et al. (2020).

Differential Returns to Labor Market Investments. We have seen that the lottery leads to increases in the cost of career-related choices, specifically as it leads to earlier and higher fertility which has been shown to significantly penalize women in the labor market. An additional channel that could play a role in the effect differentials that we find is that men and women may have differential returns from similar labor market investments and choices.

We refer back to our descriptive analysis of how earnings vary by whether the physician holds a medical PhD, specializes in a male-represented specialty, or works at a university hospital from Appendix Figure F.1. As we have seen, there are clear patterns of positive returns for both men and women that are on the same order of magnitude, but they indeed seem somewhat lower for women (see the long-run averages across years 16-35 reported on the plots). This needs, however, to be interpreted with caution for two reasons. First, within gender, the plots capture associations since they rely on comparisons across those who choose one path over the other. Second, across gender, the plots do not compare compensation schedules that hold labor supply profiles constant (either in total hours or in hours over the day/week). Since women tend to have higher costs of labor supply and time inflexibility as we discussed earlier, they may choose to work less or to work less hours in the higher-compensation times outside of regular schedules (Chen and Chevalier 2012, Jolly et al. 2014). Hence, the observed gender differentials in pay could still reflect cost differentials of labor supply to some degree. Overall, it seems that differential returns may play a role but future work that addresses the two caveats is needed for its investigation.

7. Conclusion

Using a randomized lottery that determines Danish physicians' internship placements, we find that early career setbacks have far-reaching impacts on the long-run labor market outcomes of women. Our finding of women's higher sensitivity to the initial bad shock offers a novel route that initiates and perpetuates gender inequality and gender-biased labor market norms. Our analysis highlights how persistent gender inequality can arise even in an institutionally equitable setting. As such, our findings imply that policies that aim to achieve outcome-based gender equality cannot only rely on leveling the starting playing field. Rather, such policies should target the ways in which these opportunities play out in practice and shape into gender-differential choices over the formative years of the early career. In this direction, we find strong evidence for differences in the career-family trade-off and women's greater family responsibilities as a key channel that drives the gender differential in sensitivity to career setbacks. The patterns also suggest that, even in the presence of greater human capital investments, women's careers are still ultimately governed by their family choices and responsibilities. Our findings provide novel empirical support for the view that closing the remaining gender gaps must involve structural changes in the labor market that facilitate flexible jobs for a more even accommodation of family responsibilities (Goldin 2014, Goldin and Katz 2011).

References

Abdulkadiroğlu, A., and Sönmez, T. (1998). Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems. *Econometrica*, 66(3), 689–701.

Alsan, Marcella, Owen Garrick, and Grant Graziani (2019). Does Diversity Matter for Health? Experimental Evidence from Oakland. *American Economic Review*, 109 (12): 4071-4111.

Arora, A., Goff, L., and Hjort, J. (2021). Pure-Chance Jobs versus a Labor Market: The Impact on Careers of a Random Serial Dictatorship for First Job Seekers. *AEA Papers and Proceedings*, 111: pp. 470-75.

Athey, S., Chetty, R., Imbens, G.W., and Kang, H. (2019). The Surrogate Index: Combining Short-Term Proxies to Estimate Long-Term Treatment Effects More Rapidly and Precisely. NBER Working Paper 26463.

Azmat, G., Calsamiglia, C., and Iriberri, N. (2016). Gender Differences in Response to Big Stakes. *Journal of the European Economic Association*, 14(6), 1372–1400.

Azmat, G., and Ferrer, R. (2017). Gender Gaps in Performance: Evidence from Young Lawyers. *Journal of Political Economy*, 125(5), 1306–1355.

Bhattacharya, J., Hyde, T., and Tu, P. (2013). *Health Economics*. Macmillan International Higher Education.

Bettinger, E.P., and Long, B.T. (2005). Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students. *American Economic Review*, *95*(2), 152–157.

Bertrand, M. (2011). New Perspectives on Gender. In *Handbook of Labor Economics*, 4, 1543–1590.

Bertrand, M. (2020). Gender in the Twenty-First Century. *AEA Papers and Proceedings*, 110, 1-24.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2(3), 228–255.

Black, S.E., and Devereux, P.J. (2011). Recent Developments in Intergenerational Mobility. In *Handbook of Labor Economics*, 4, 1487–1541.

Blau, F. D., and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), 789–865.

Blau, F. D., Currie, J. M., Croson, R. T., and Ginther, D. K. (2010). Can Mentoring Help Female Assistant Professors? Interim Results from a Randomized Trial. *American Economic Review*, 100(2), 348–52.

Brotherton, S. E., and Etzel, S. I. (2018). Graduate Medical Education, 2017-2018. *JAMA*, 320(10), 1051–1070.

Bursztyn, L., Fujiwara, T., and Pallais, A. (2017). 'Acting Wife': Marriage Market Incentives and Labor Market Investments. *American Economic Review*, 107(11), 3288–3319.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices. *Quarterly Journal of Economics*, 129(3), 1409–1447.

Buser, T., and Yuan, H. (2019). Do Women Give Up Competing More Easily? Evidence from the Lab and the Dutch Math Olympiad. *American Economic Journal: Applied Economics*, 11(3), 225–252.

Cai, X., Lu, Y., Pan, J., and Zhong, S. (2019). Gender Gap under Pressure: Evidence from China's National College Entrance Examination. *Review of Economics and Statistics*, 101(2), 249–263.

Caldwell, S. and Danieli, O. (2024). "Outside Options in the Labor Market." *Review of Economic Studies*.

Caliendo, M., Lee, W. S., and Mahlstedt, R. (2017). The Gender Wage Gap and the Role of Reservation Wages: New Evidence for Unemployed Workers. *Journal of Economic Behavior and Organization*, 136, 161–173.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *Quarterly Journal of Economics*, 131(2), 633–686.

Carrell, S.E., Page, M.E., and West, J.E. (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap. *Quarterly Journal of Economics*, 125(3), 1101–1144.

Chen, M.K. and Chevalier, J.A. (2012). "Are Women Overinvesting in Education? Evidence from The Medical Profession." *Journal of Human Capital*, 6(2): pp.124-149.

Cheng, Stephanie D. (2020). Careers Versus Children: How Childcare Affects the Academic Tenure-Track Gender Gap.

Chetty, R., and Hendren, N. (2018). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *Quarterly Journal of Economics*, 133(3), 1107–1162.

Cortés, P., Pan, J., Reuben, E., Pilossoph, L. and Zafar, B. (2023). Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab. *Quarterly Journal of Economics*, 138(4): 2069-2126.

Cullen, Z. and Perez-Truglia, R. (2023). The Old Boys' Club Schmoozing and the Gender Gap. *American Economic Review*, 113(7), pp.1703-1740.

Damm, A. P., and Dustmann, C. (2014). Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior? *American Economic Review*, 104(6), 1806–32.

Danish Economic Councils (2015). Dansk Økonomi, Forår 2015.

Danish Ministry of Health (1989). Lægers Kliniske Videreuddannelse og arbejdstilrettelæggelse, betænkning nr. 1183, September 1989.

DellaVigna, S., Heining, J., Schmieder, J.F. and Trenkle, S. (2022). Evidence on Job Search Models from a Survey of Unemployed Workers in Germany. *Quarterly Journal of Economics*, 137(2): pp. 1181-1232.

Dennehy, T.C., and Dasgupta, N. (2017). Female Peer Mentors Early in College Increase Women's Positive Academic Experiences and Retention in Engineering. *PNAS*, 114(23), 5964–5969.

Eckert, F., Walsh, C., and Hejlesen, M. (2022). The Return to Big City Experience: Evidence from Danish Refugees. *Journal of Urban Economics*, 130, 103454.

Exley, C. L., and Kessler, J. B. (2022). The Gender Gap in Self-Promotion. *Quarterly Journal of Economics*, 137(3), pp.1345-1381.

Field, E., Jayachandran, S., Pande, R., and Rigol, N. (2016). Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship? *American Economic Journal: Economic Policy*, 8(2), 125–53.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of Geographic Variation in Health Care: Evidence from Patient Migration. *Quarterly Journal of Economics*, 131(4), 1681–1726.

Ginther, D. K., Currie, J. M., Blau, F. D., and Croson, R. T. (2020). Can Mentoring Help Female Assistant Professors in Economics? An Evaluation by Randomized Trial. *AEA Papers and Proceedings*, 110, 205–09.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *Quarterly Journal of Economics*, 118(3), 1049–1074.

Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4), 1091–1119.

Goldin, C. (2021). Career and Family: Women's Century-Long Journey toward Equity. Princeton University Press.

Goldin, C., and Katz, L. F. (2008). Transitions: Career and Family Life Cycles of the Educational Elite. *American Economic Review*, 98(2), 363–69.

Goldin, C. and Katz, L. F. (2011). "The Cost of Workplace Flexibility for High-Powered Professionals." *The Annals of the American Academy of Political and Social Science*, 638(1): pp. 45-67.

Goldin, C., Kerr, S.P., Olivetti, C., and Barth, E., (2017). "The Expanding Gender Earnings Gap: Evidence from The LEHD-2000 Census." *American Economic Review*, 107(5): pp.110-114.

Holman, L., Stuart-Fox, D., and Hauser, C. E. (2018). The Gender Gap in Science: How Long until Women are Equally Represented? *PLoS biology*, 16(4), e2004956.

Huang, J., Gates, A. J., Sinatra, R., and Barabási, A. L. (2020). Historical Comparison of Gender Inequality in Scientific Careers across Countries and Disciplines. *PNAS*, 117(9), 4609–4616.

Iriberri, N., and Rey-Biel, P. (2019). Competitive Pressure Widens the Gender Gap in Performance: Evidence from a Two-stage Competition in Mathematics. *Economic Journal*, 129(620), 1863–1893.

Jensen, M. L., Hangaard, S., Wolf, A. L. A., and Munch, K. (2013). "Medicinstuderende og yngre lægers speciale-og karrierevalg." KORA.

Jolly, S., Griffith, K.A., DeCastro, R., Stewart, A., Ubel, P. and Jagsi, R. (2014). "Gender Differences in Time Spent on Parenting and Domestic Responsibilities By High-Achieving Young Physician-Researchers." *Annals of Internal Medicine*, 160(5): pp.344-353.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of Study, Earnings, and Self-Selection. *Quarterly Journal of Economics*, 131(3), 1057-1111.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimuller, J. (2019). Child Penalties across Countries: Evidence and Explanations. *AEA Papers and Proceedings*, 109, 122–26.

Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4), 181–209.

Kofoed, M. S., and McGovney, E. (2019). The Effect of Same-Gender or Same-Race Role Models on Occupation Choice: Evidence from Randomly Assigned Mentors at West Point. *Journal of Human Resources*, 54(2), 430–467.

Laird, J. and Nielsen, T. (2016). The Effects of Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply: Evidence from Movers in Denmark.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2021). Gender Differences in Job Search: Trading off Commute Against Wage. *Quarterly Journal of Economics*, 136(1), 381-426.

Lorentzen, Kathrine Aaby (2024). "Family-Friendly Jobs and Occupational Sorting across Gender: Evidence from Introduction of a Maximum 40-hour Workweek." CEBI Working Paper Series, CEBI WP 03-24.

Lundberg, S., and Stearns, J. (2019). Women in Economics: Stalled Progress. *Journal of Economic Perspectives*, 33(1), 3–22.

National Board of Health (2009). Sundhedsstyrelsens målbeskrivelse for den kliniske basisuddannelse.

Niederle, M. and Vesterlund, L. (2007). Do Women Shy Away from Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, 122(3), 1067–1101.

OECD (2012). Health at a Glance 2011: OECD Indicators 2011. OECD Publishing, Paris.

Olivetti, C. and Petrongolo, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, 8(1), 405–434.

Ono, T., Schoenstein, M., and Buchan, J. (2014). Geographic Imbalances in Doctor Supply and Policy Responses. OECD Health Working Paper.

Porter, C., and Serra, D. (2020). Gender Differences in the Choice of Major: The Importance of Female Role Models. *American Economic Journal: Applied Economics*, 12, 226–254.

Rubinstein, Y., and Weiss, Y. (2006). Post Schooling Wage Growth: Investment, Search and Learning. *Handbook of the Economics of Education*, 1, 1–67.

Sarsons, H. (2019). Interpreting Signals in the Labor Market: Evidence from Medical Referrals.

Simoens, S., and Hurst, J. (2006). The Supply of Physician Services in OECD Countries. OECD Health Working Paper.

Sloan, F.A., and Edmunds, M. eds. (2012). Geographic Adjustment in Medicare Payment: Phase I: Improving Accuracy. *National Academies Press.*

Topel, R. H., and Ward, M. P. (1992). Job Mobility and the Careers of Young Men. *Quarterly Journal of Economics*, 107(2), 439–479.

von Wachter, Till. (2020). The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources. *Journal of Economic Perspectives*, 34(4), 168–194.

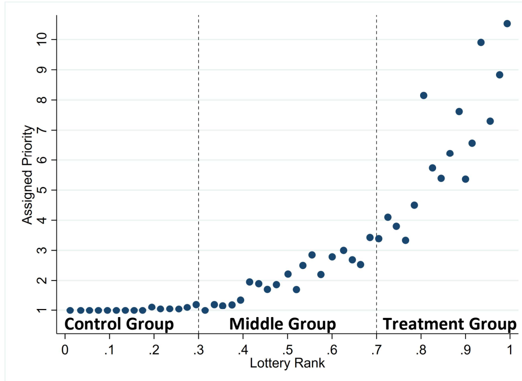
Wasserman, M. (2023). Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents. *Review of Economic Studies*, 90(3): 1535-1568

Weiss, Y. (1986). The Determination of Life Cycle Earnings: A Survey. *Handbook of Labor Economics*, 1, 603–640.

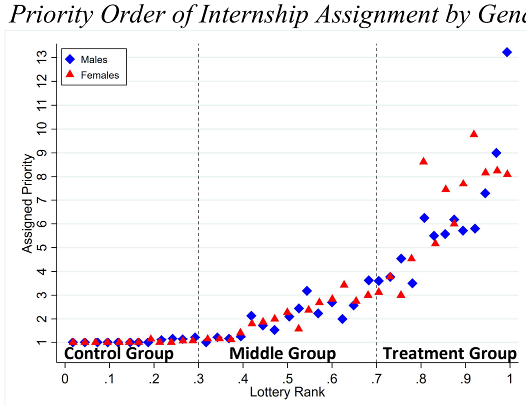
Zeltzer, D. (2020). Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap. *American Economic Journal: Applied Economics*, 12, 169–197.

Figure 1: Patterns in Assignment to Internships

A. Priority Order of Internship Assignment

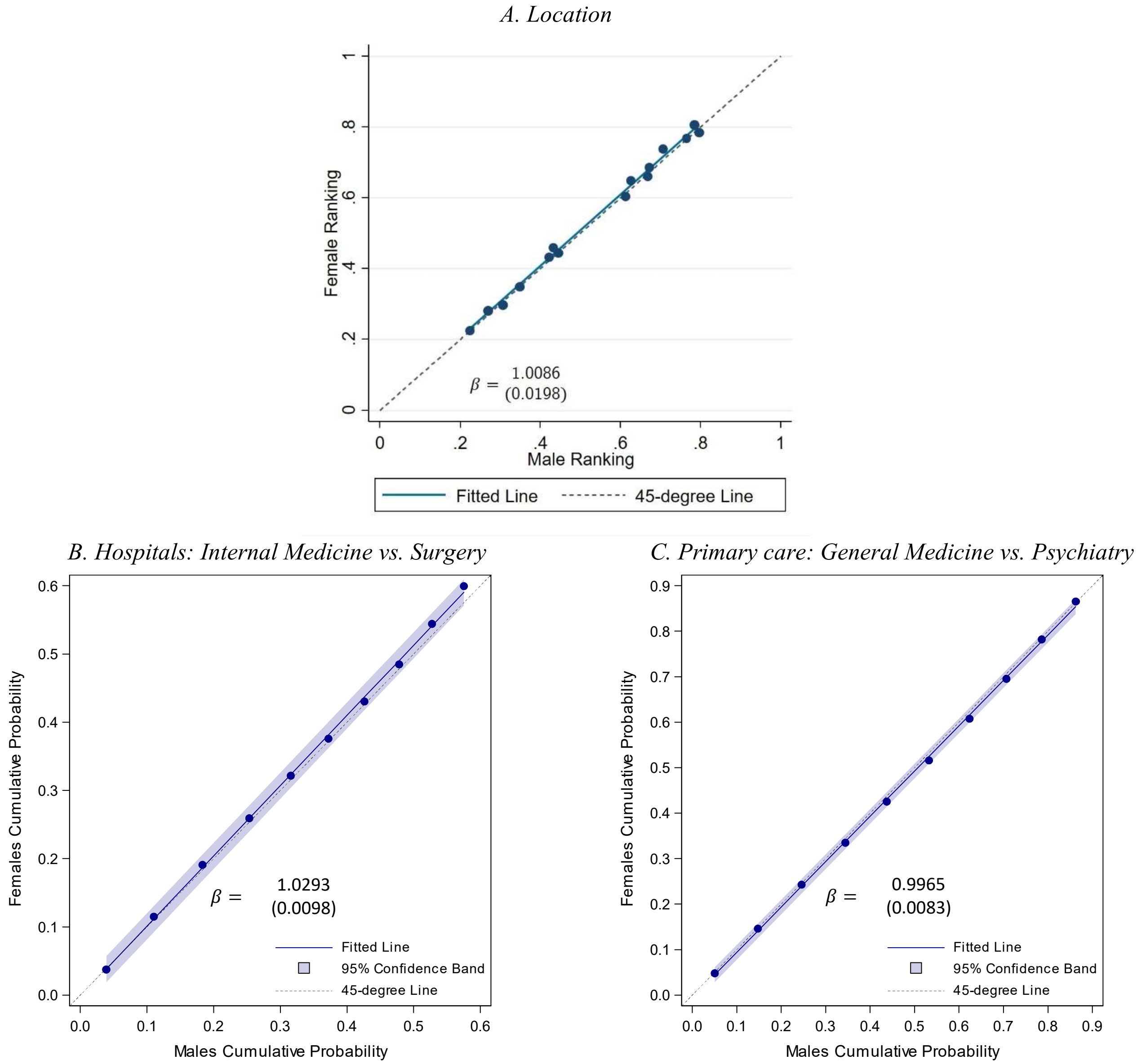


B. Priority Order of Internship Assignment by Gender



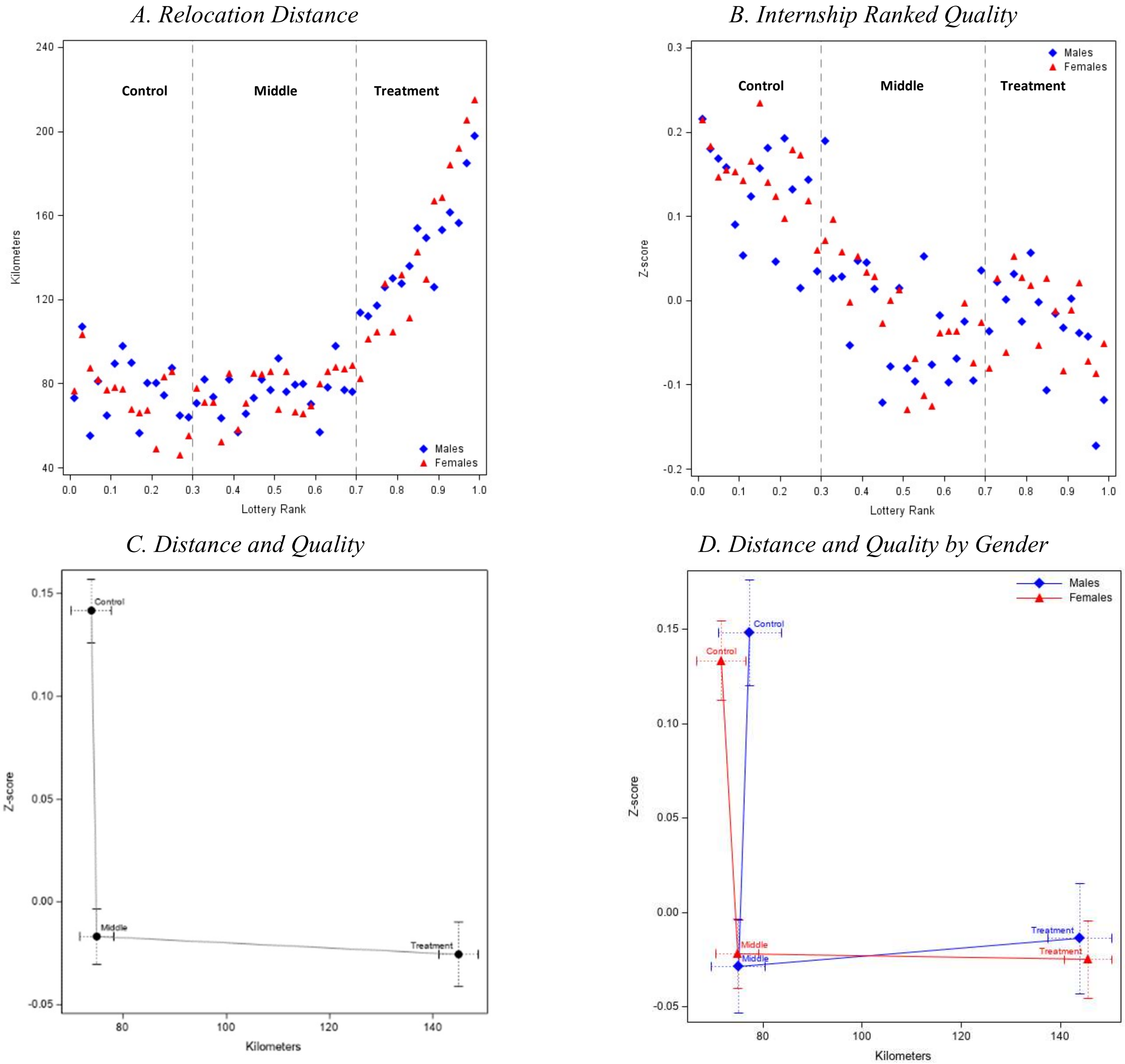
Notes: This figure studies the mapping between choice rankings and placements as a function of the lottery. We plot individuals' pre-placement ranking of the local labor market they were assigned to (where 1 denotes the highest priority) against the percentile rank of their lottery number within their graduating cohort. We use information on interns' binding pre-placement rankings of all local labor markets, which have been solicited among the earlier cohorts (who graduated prior to 2008) as part of the allocation process. Panel A includes all graduates, and panel B splits graduates by gender.

Figure 2: Preferences over Entry-Level Positions by Gender



Notes: This figure compares male and female graduates' revealed preferences over entry-level positions. Panel A compares preferences over local labor markets, where we analyze a measure for market desirability that reveals students' location preferences through their lottery-based choices. We construct market rankings based on the average lottery rank of the interns who sort into it, separately for males and females. Each dot represents a local labor market, where the x-axis denotes male rankings and the y-axis denotes female rankings. We plot the fitted line and report its slope, where the benchmark of non-differential ranking by gender is 1 (the 45-degree line). Panels B-C compare preferences over internship specialties. Within the primary positions in hospitals, interns can broadly choose between internal medicine and surgery, and, within the secondary positions in primary care, interns can choose between general medicine and psychiatry. For each gender, we split the sample by deciles according to lottery ranks. Then, for each of the two types of positions, we calculate over deciles the gender-specific cumulative probability of making a particular choice of specialty over the other. We plot the gender-specific CDFs against one another, where the 45-degree line again serves as a benchmark when preferences are similar across gender. We also plot the fitted line along with 95-percent confidence intervals and report its slope.

Figure 3: Choices Conditional on Lottery Rank and the First Stage

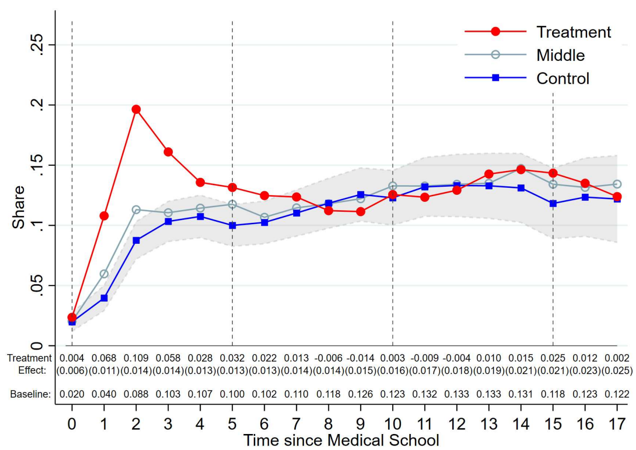


Notes: These figures study patterns in internship choices conditional on lottery rank. Panel A plots a graduating student's relocation distance against the student's lottery rank split by gender. We calculate for each student the distance between their ZIP Code of residence at the time of the lottery and their ZIP Code of work at the time of the internship, which captures their "relocation distance." Panel B investigates the quality of the internship a graduating student has been assigned to against the student's lottery rank split by gender. We use ranked quality of the educational program (at the hospital-department level) as reported by interns in the exit surveys. We use the leave-one-out mean of the overall evaluation normalized by the overall mean and standard deviation of this measure (to create a z-score). Panel C plots the averages of relocation distance and internship quality together, for each of our experimental groups: best lottery ranks (the "control" group), middle range lottery ranks (the "middle" group), and the worst lottery ranks (the "treatment" group). Panel D splits panel C by gender. Panels C-D also display the corresponding 95-percent confidence intervals.

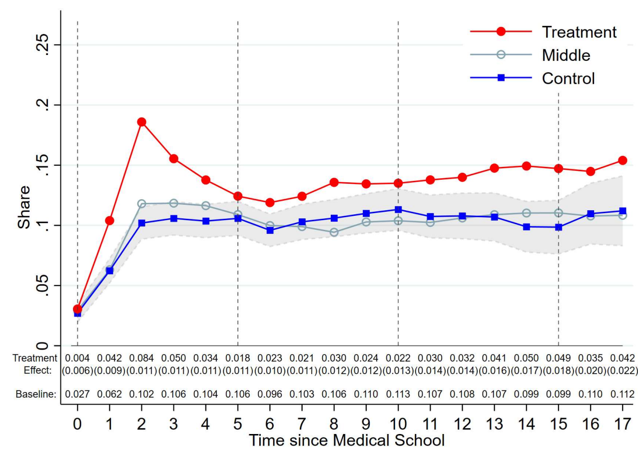
Figure 4: Geographic Sorting

A. Sorting into a Rural Labor Market

Males

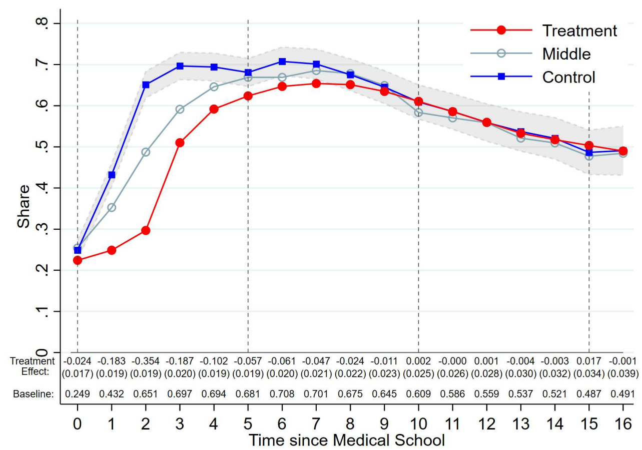


Females

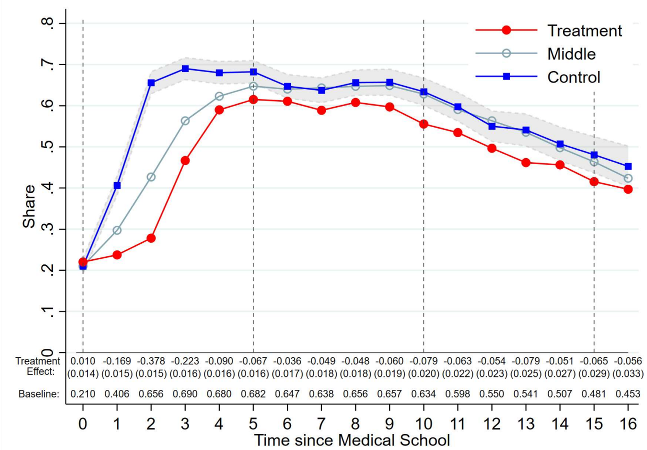


B. Affiliation with a University Hospital

Males

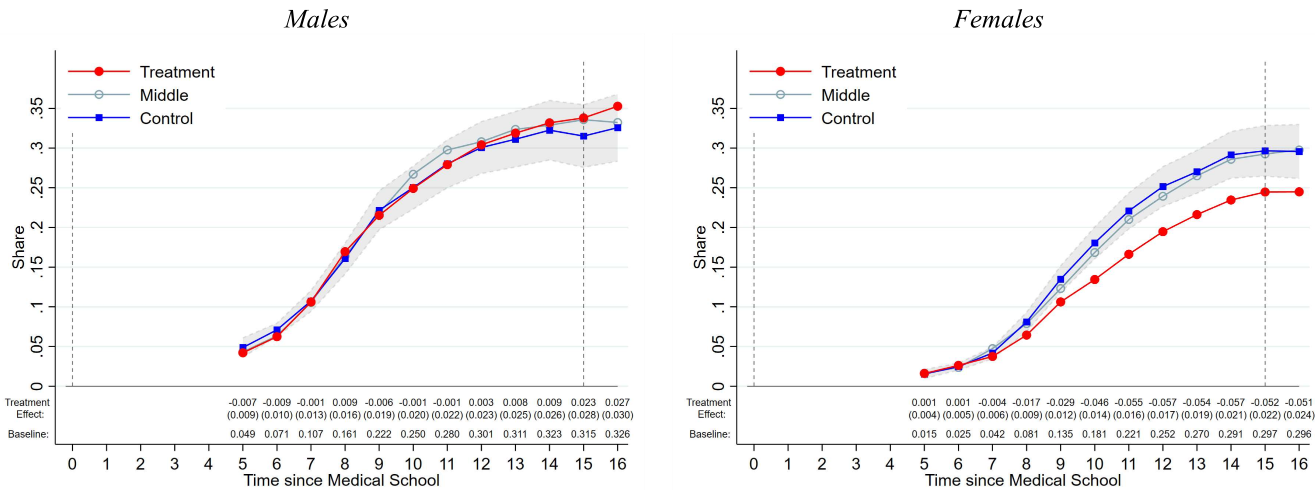


Females



Notes: This figure plots the impact of the lottery on geographic sorting. The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the full dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. Panel A plots the probability of sorting into a rural local labor market. The outcome is an indicator for whether a physician resides in a rural municipality by the end of December in a given year. Estimations run to year 17 since 2021 is our last available calendar year in the demographic data. Panel B plots the probability of having an affiliation with a university hospital in November in a given year. Estimations run to year 16 since 2020 is the last available year in the employment data.

Figure 5: Human Capital Investment—Obtaining a Medical PhD

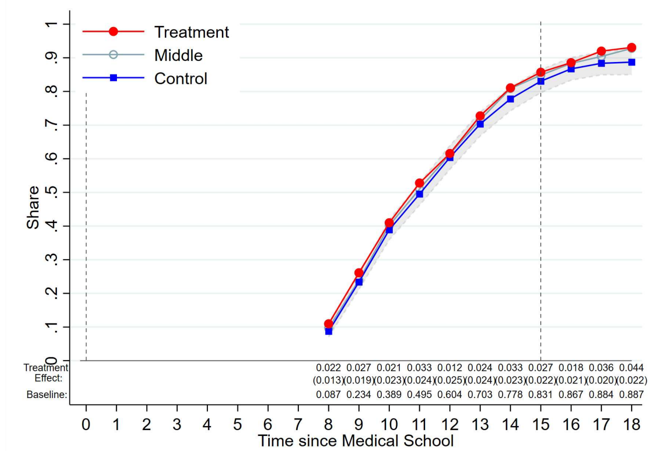


Notes: This figure plots the impact of the lottery on human capital investment. The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the full dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. The outcome is an indicator for having completed a medical PhD by the given year. Estimations run from when the outcome begins to materialize up to year 16 since 2020 is our last available completion year in the education registers.

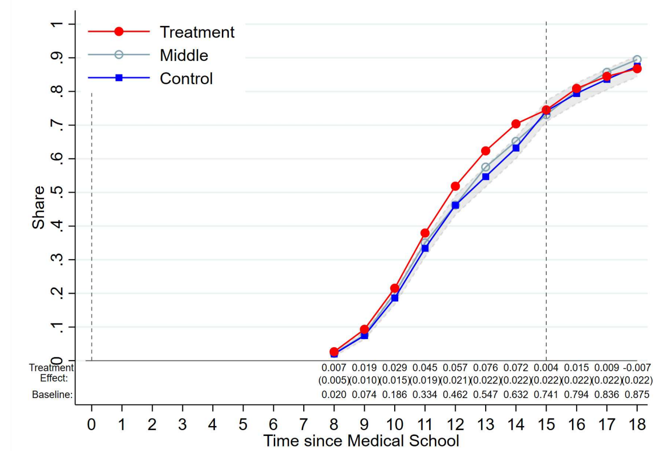
Figure 6: Timing of Occupational Choice

A. Completion of a Medical Specialty (of Any Type)

Males

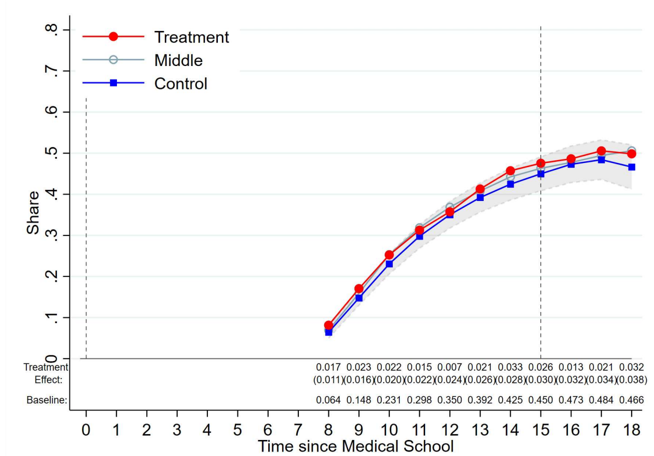


Females

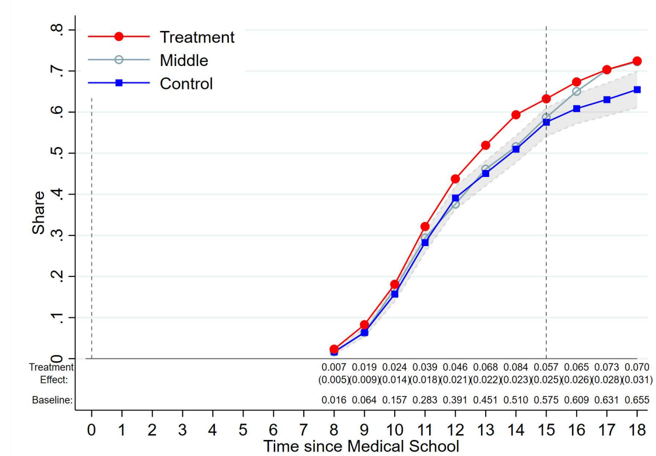


B. Completion of a Female-Represented Specialty

Males

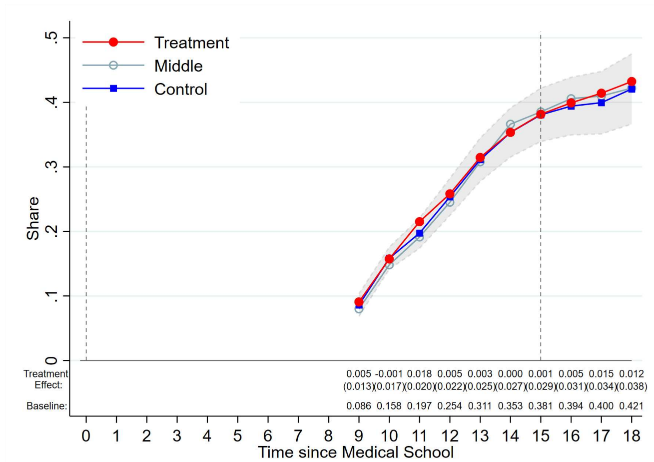


Females

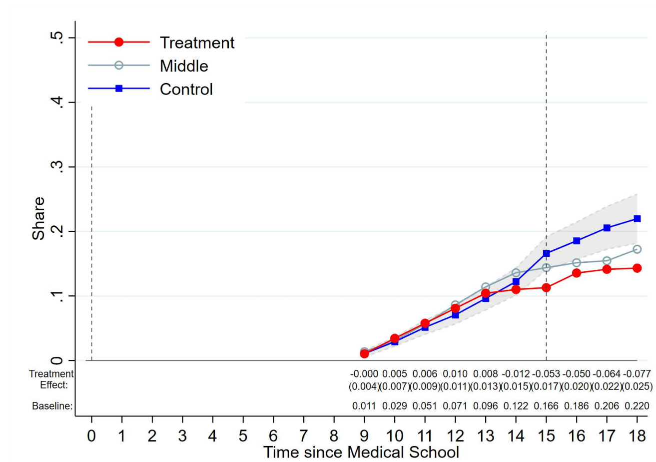


C. Completion of a Male-Represented Specialty

Males

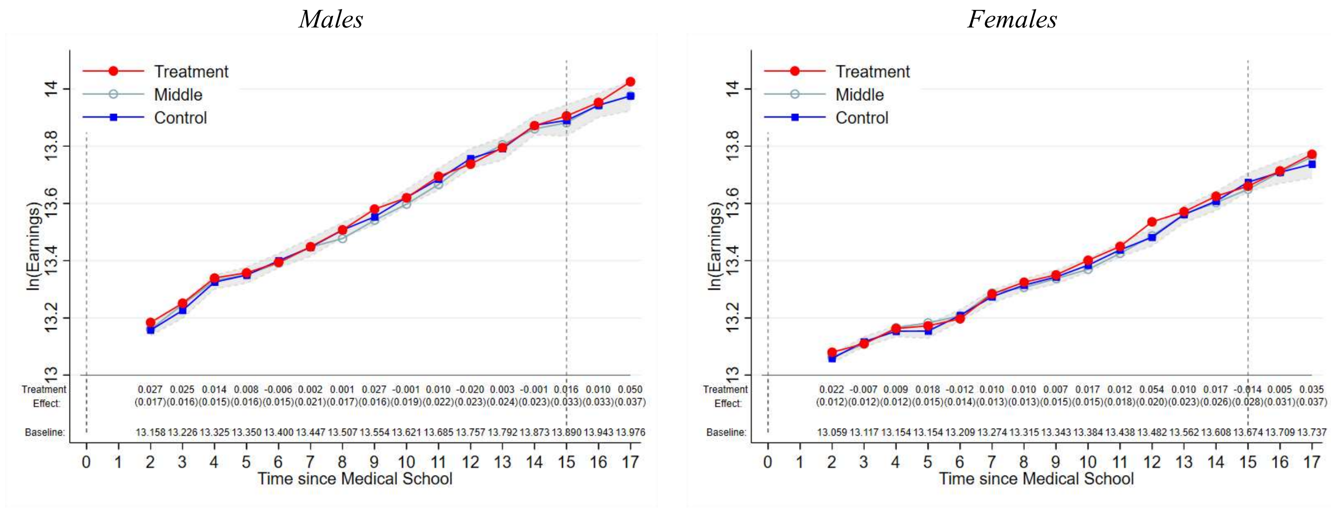


Females



Notes: This figure plots the impact of the lottery on occupational choice. We plot the dynamics in medical specialty completion from when the outcome begins to materialize up to year 18 after graduation (as 2022 is our last available year in the specialization register) for our three experimental groups. We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. Panel A studies completion of any medical specialty, and panels B-C split specialties into female-represented specialties and male-represented specialties.

Figure 7: Earnings Profiles

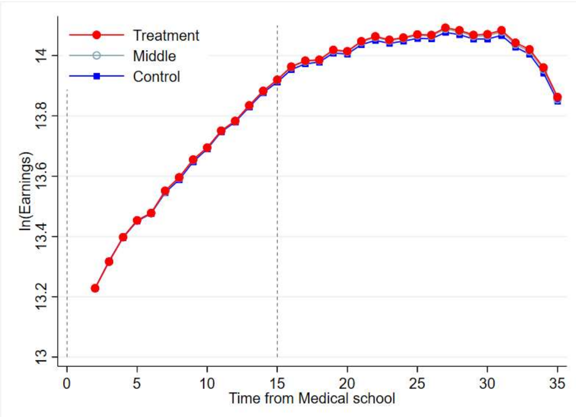


Notes: This figure plots the impact of the lottery on log earnings. The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the full dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. The outcome incorporates total compensation including annual wage earnings, net income from self-employment, and labor market pension contributions (analogous to employer contributions to 401(k)s). Compensation is measured pre-tax in 2017-prices (deflated by the Danish regions' wage index). Estimations run from year 2 (which is the first full calendar year in the labor market given that the lottery takes place during year 1) up to year 17 since 2021 is our last available calendar year in the income register.

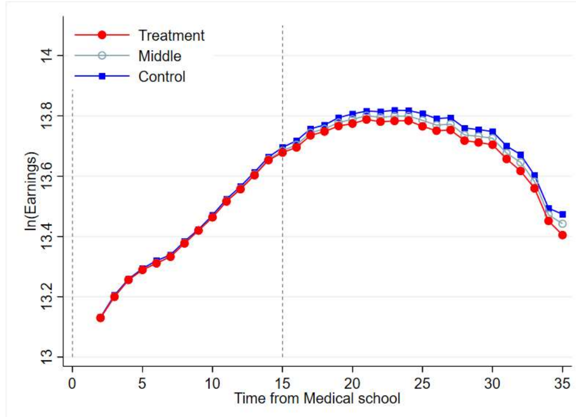
Figure 8: Predicted Long-Run Earnings

A. Predicted Long-Run Earnings Profiles

Males

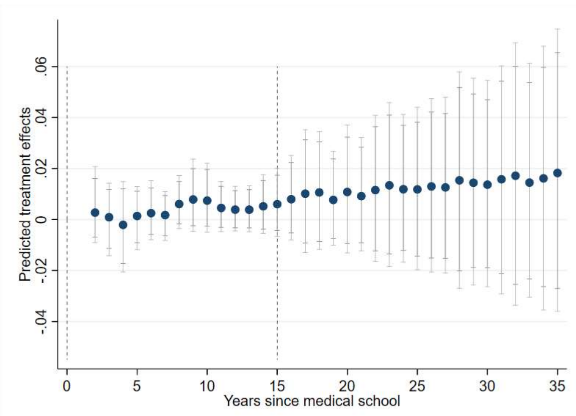


Females

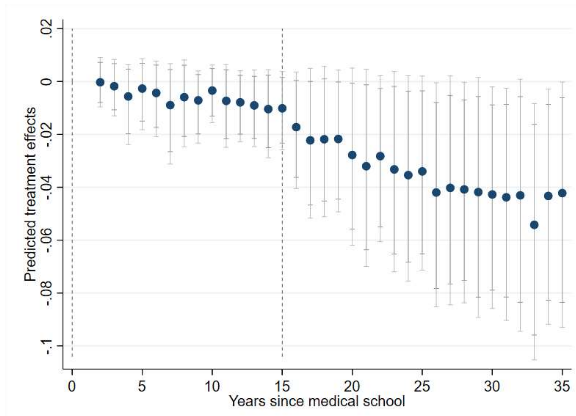


B. Predicted Long-Run Treatment Effects (Treatment vs. Control Group)

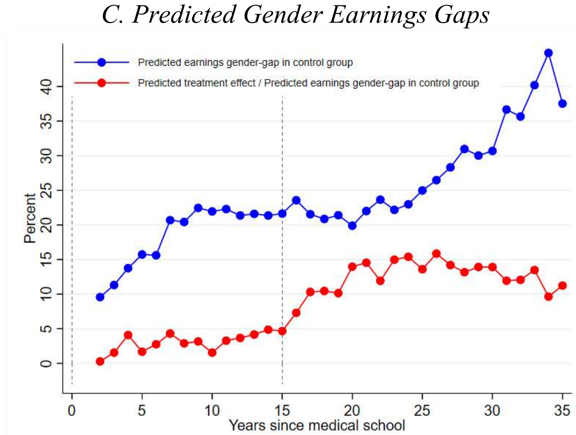
Males



Females



C. Predicted Gender Earnings Gaps

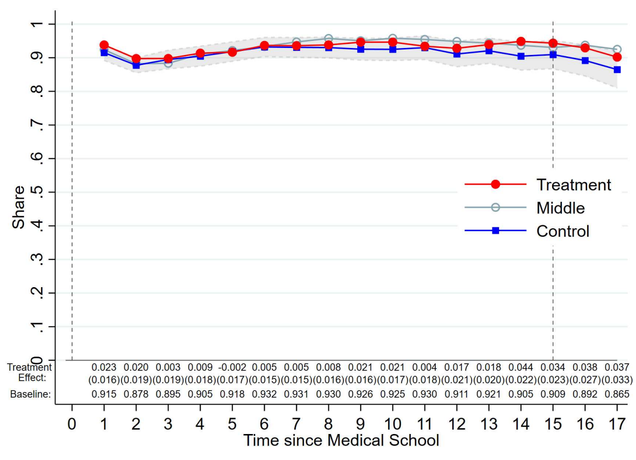


Notes: This figure investigates the long-run effects on total work compensation based on earnings predictions that run from year 2 (the first full year of employment) to year 35 after graduation. See Appendix F.1 for details on our prediction procedure. Earnings incorporate total compensation, including annual wage earnings, net income from self-employment, and labor market pension contributions (analogous to employer contributions to 401(k)s). Panel A plots the predicted long-run earnings profiles for our three experimental groups split by gender. Panel B provides the predicted treatment effects on earnings for males and females, capturing the level gaps from panel A along with their 90-percent and 95-percent confidence intervals. In panel C, the blue line plots the predicted baseline gender earnings gap (which is estimated using the control group), and the red line plots the additional earnings gap among physicians in the treatment group caused by the experiment as a share of the baseline earnings gap (which is calculated using predicted treatment effects). Standard errors are bootstrapped to account for estimation error from the two steps of the surrogate index analysis.

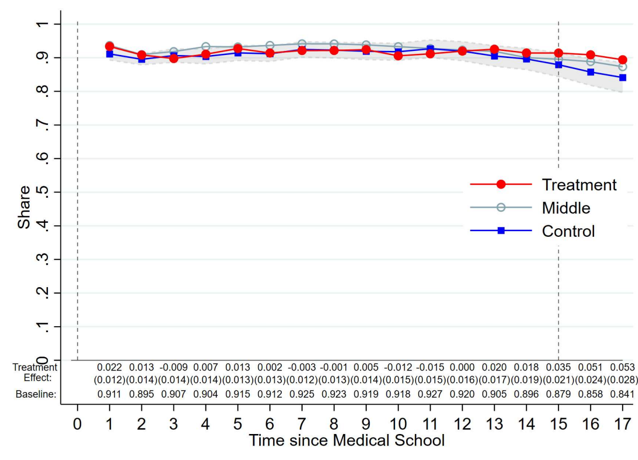
Figure 9: Likelihood of Partnership

A. Partnered at Baseline

Males

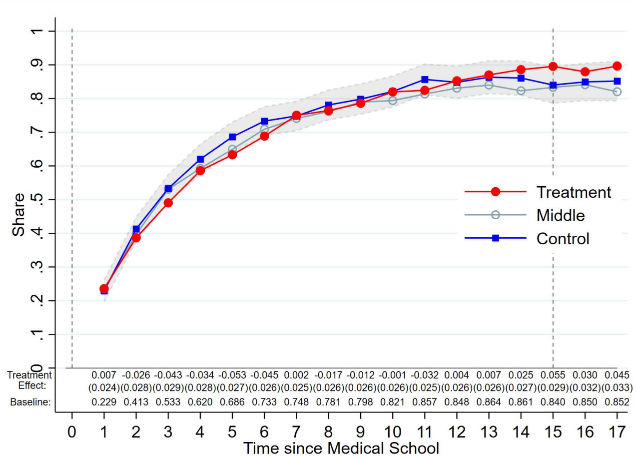


Females

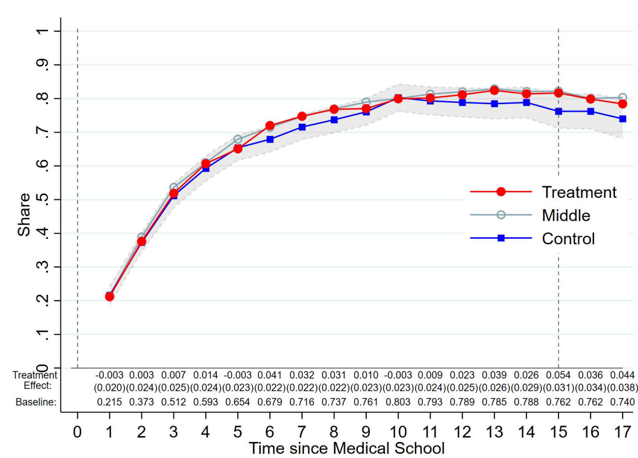


B. Single at Baseline

Males



Females

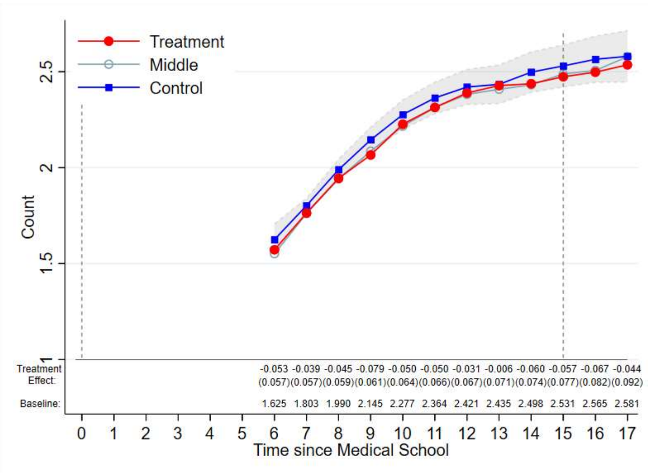


Notes: This figure plots the impact of the lottery on partnership. The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the full dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. The outcome is an indicator for having a registered partner (married or cohabiting). Estimations run to year 17 since 2021 is our last available calendar year in the demographic registers. We split the sample by partnership status at baseline, studying physicians who were in a partnership in panel A and physicians who were single in panel B.

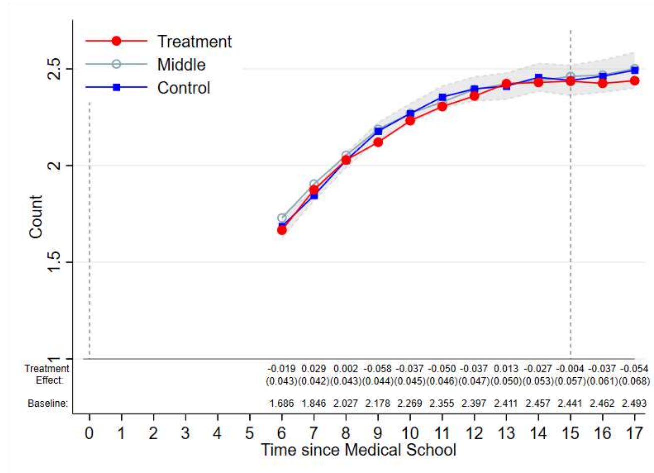
Figure 10: Fertility

A. Partnered at Baseline

Males

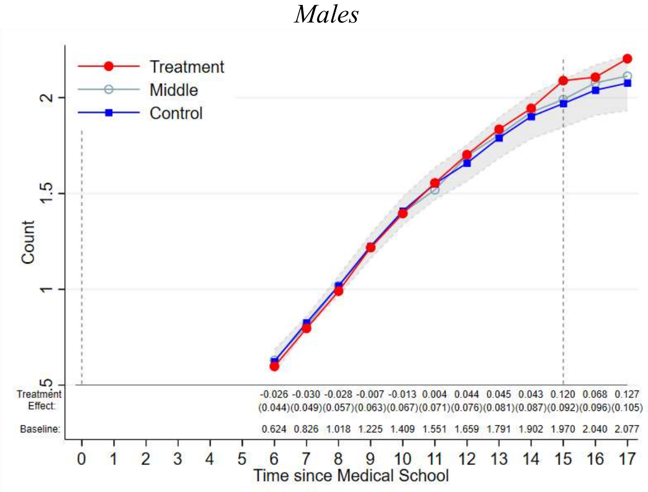


Females

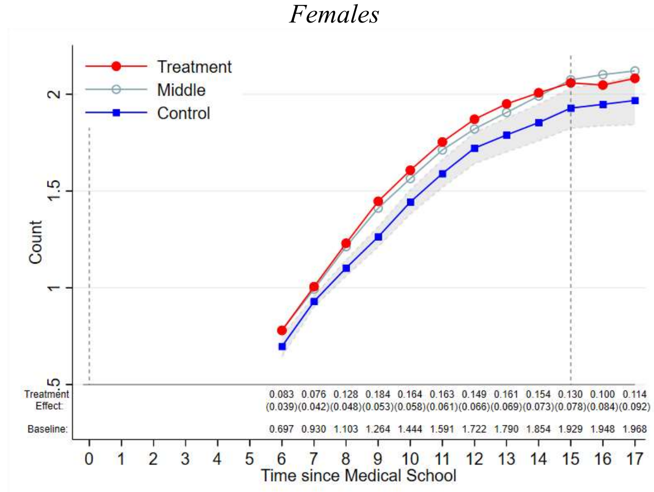


B. Single at Baseline

Males



Females



Notes: This figure plots the impact of the lottery on fertility. The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. The outcome is the number of children of whom the physician is registered as a parent. Estimations run from year 6 (after the spikes in fertility among those single at baseline shown in Appendix Figure A.6) to year 17 since 2021 is our last available calendar year in the demographic registers. We split the sample by partnership status at baseline, studying physicians who were in a partnership in panel A and physicians who were single in panel B.

Table 1: Search and Mobility

	A. Commuting Distance		*B. Living in Baseline Location of Partner*	
	Males	Females	Males	Females
Treatment Effect	2.2535	-0.5025	-0.0577	-0.0178
	(1.0320)	(0.7398)	(0.0193)	(0.0158)
Effect on Middle Group	0.6129	0.2210	-0.0391	0.0026
	(0.9163)	(0.7579)	(0.0182)	(0.0149)
Constant (Control Group)	26.3233	26.0107	0.4924	0.4559
	(0.6906)	(0.5644)	(0.0137)	(0.0112)
Observations	23,777	35,279	29,300	44,555
Individuals	3,565	5,470	3,970	6,103

Notes: This table studies search and mobility. In panel A we analyze the average effect of the lottery on commuting distances (in kilometers). In panel B we study a physician's propensity to reside within the pre-lottery location of their current spouse. The table provides estimates of β from equation (2) using years 6-15 after graduation. Robust standard errors clustered at the individual level are reported in parentheses.

Table 2: Sensitivity to Workplace Characteristics

A. Degree of Exposure to Employer Intensity		
	Males	Females
Treatment Effect	-0.1673	-0.1432
	(0.0081)	(0.0061)
Effect on Middle Group	-0.0675	-0.0542
	(0.0075)	(0.0058)
Constant (Control Group)	0.4713	0.4460
	(0.0056)	(0.0044)
Individuals	3,005	5,155

B. Sensitivity to Employer Intensity		
	Males	Females
Employer Intensity	0.4324	0.5785
	(0.0525)	(0.0407)
Constant	0.2113	0.1555
	(0.0225)	(0.0168)
Individuals	2,490	4,272

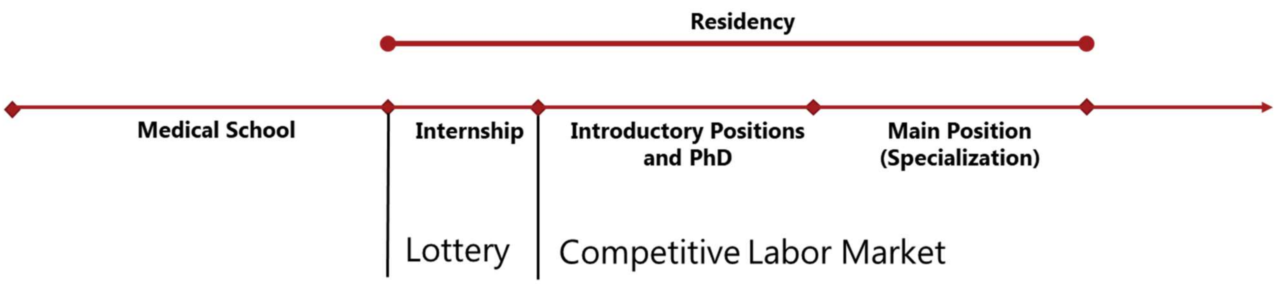
C. Exposure to Female Role Models		
	Males	Females
Treatment Effect	-0.1985	-0.1596
	(0.0236)	(0.0179)
Effect on Middle Group	-0.1066	-0.0706
	(0.0216)	(0.0170)
Constant (Control Group)	0.5965	0.5936
	(0.0162)	(0.0128)
Individuals	2,960	5,046

Notes: This table investigates the role of employer-side factors as an explanation for the gender divergence in treatment effects. Employer intensity is defined as the leave-one-out mean of a hospital department's propensity to place their interns in a university hospital in their subsequent positions by gender. Panel A provides estimates for interns' exposure to employer intensity. Panel B provides estimates for interns' sensitivity to employer intensity by regressing one's own probability of being employed at a university hospital in their next position on the intensity of their internship employer. Panel C uses the internship exit surveys to investigate the effect of the lottery on the probability of exposure to female role models during the internship, which include the assigned supervisor and the program department chair. Robust standard errors are reported in parentheses.

Online Appendix

Appendix A: Physician Training and Choice Timing

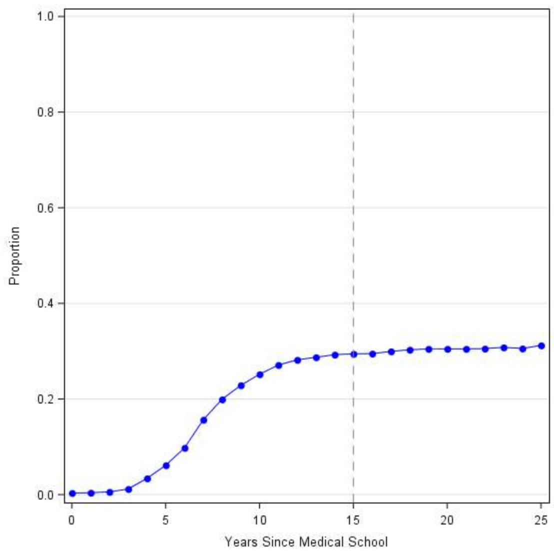
Appendix Figure A.1: Timeline



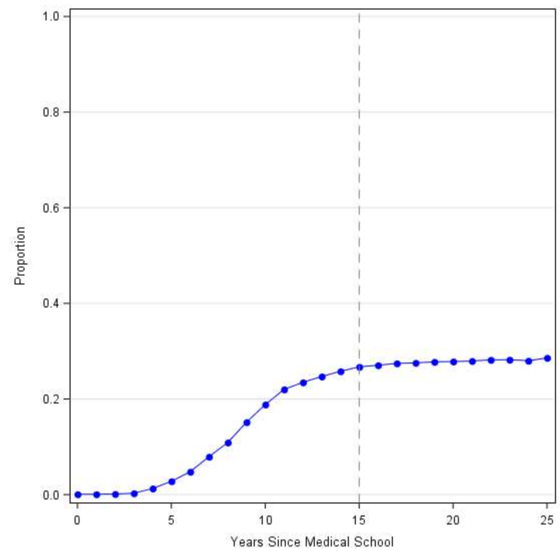
Notes: This figure summarizes the timeline of Danish physicians' training, which captures the early stages of their careers.

Appendix Figure A.2: Time to Obtaining a Medical PhD

Males



Females

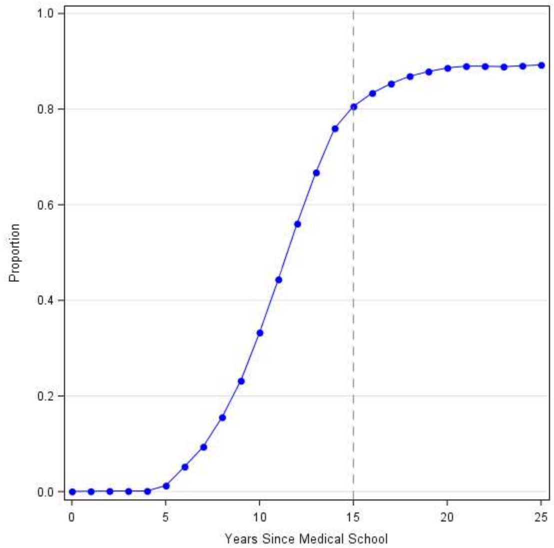


Notes: This figure plots the timing of obtaining a medical PhD. The sample includes physicians who graduated from medical school in years 1980-2000.

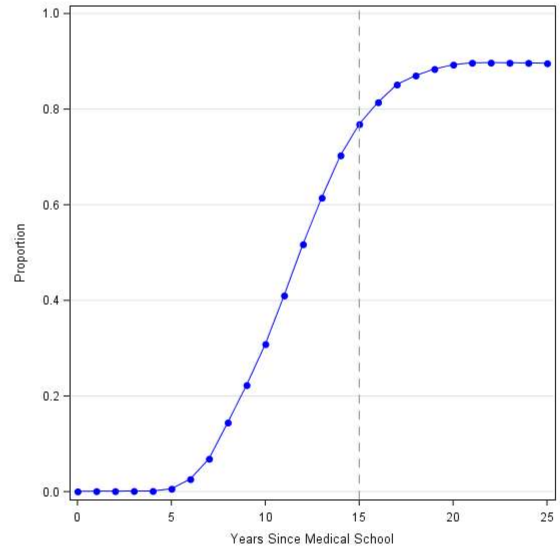
Appendix Figure A.3: Time to Completion of Medical Specialty

A. Any specialty

Males

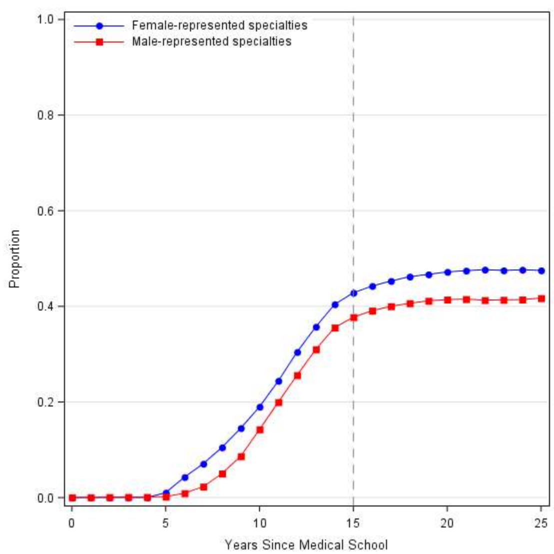


Females

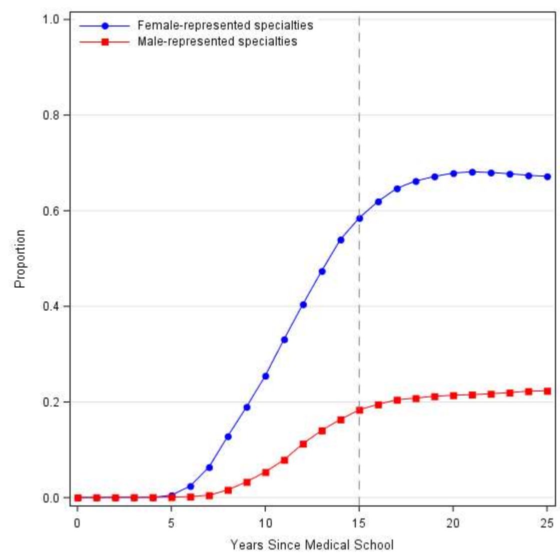


B. by Specialty Type

Males



Females

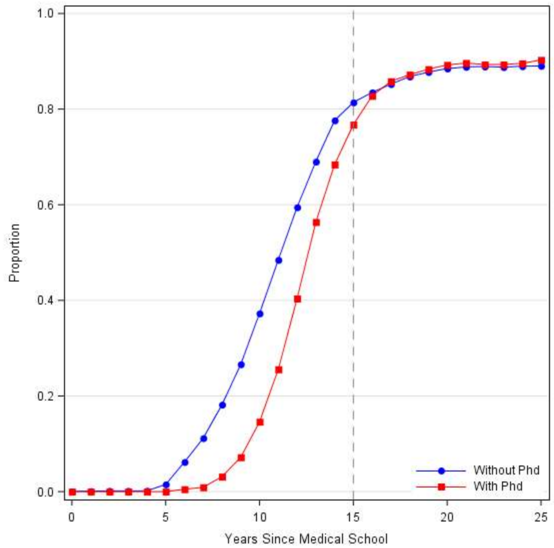


Notes: Panel A plots the timing of completing a medical specialty, i.e., physicians' occupational choice, where panel B splits specialties into female-represented specialties and male-represented specialties. The sample includes physicians who graduated from medical school in years 1980-2000.

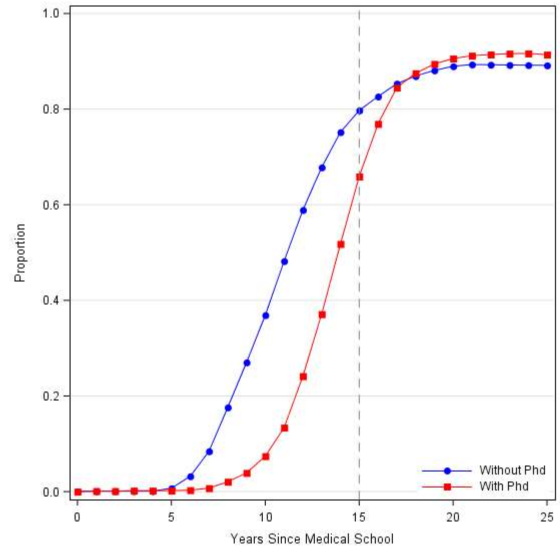
Appendix Figure A.4: Time to Medical Specialty by PhD Status and Specialty Type

A. Any specialty

Males

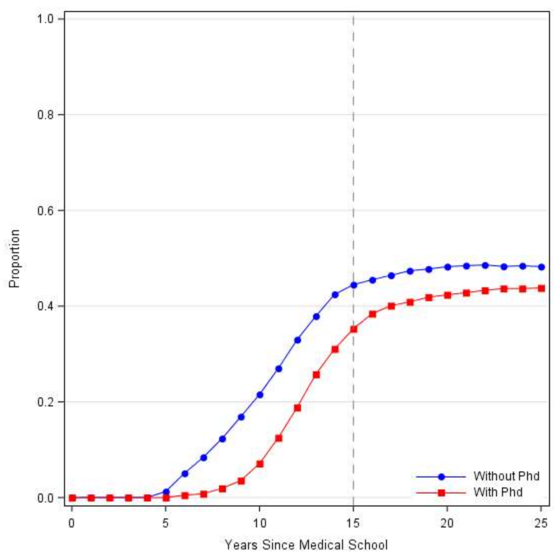


Females

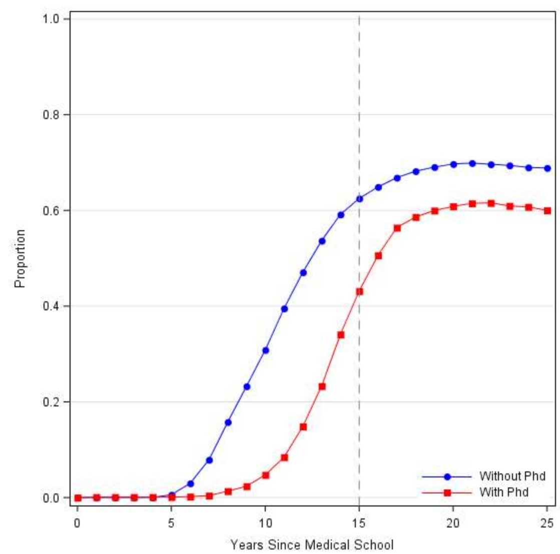


B. Female-Represented Specialties

Males

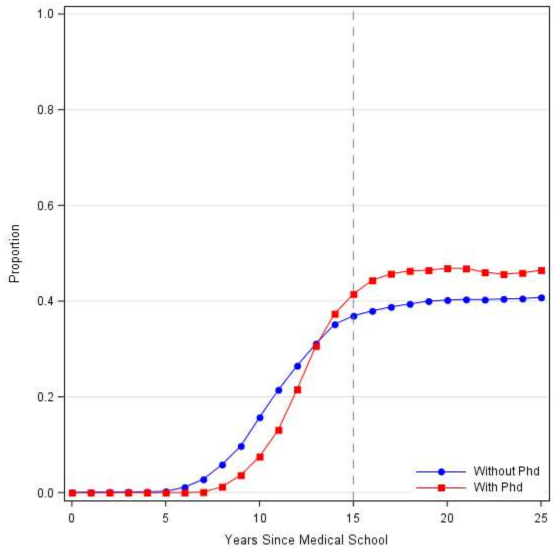


Females

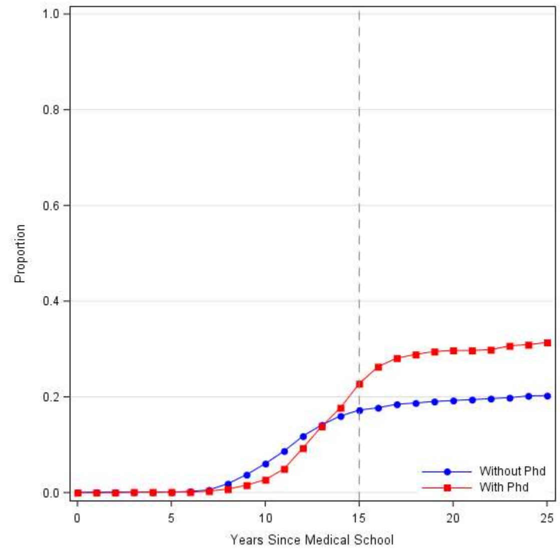


C. Male-Represented Specialties

Males



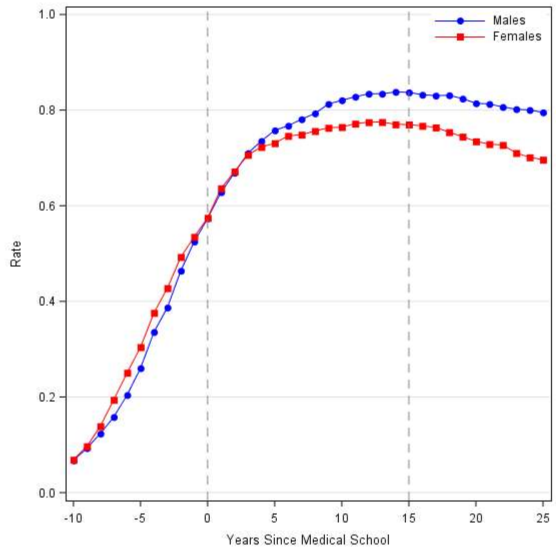
Females



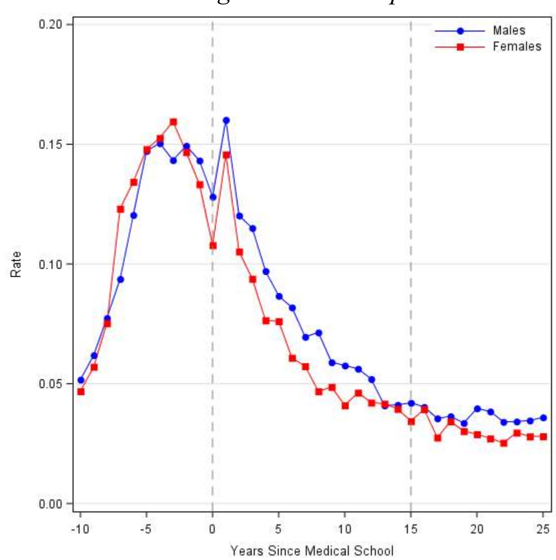
Notes: Panel A plots the timing of medical specialty completion split by whether a physician has obtained a medical PhD at any point in time, where panels B and C provide a split into female-represented specialties and male-represented specialties. The sample includes physicians who graduated from medical school in years 1980-2000.

Appendix Figure A.5: Partnership

A. Partnership Status

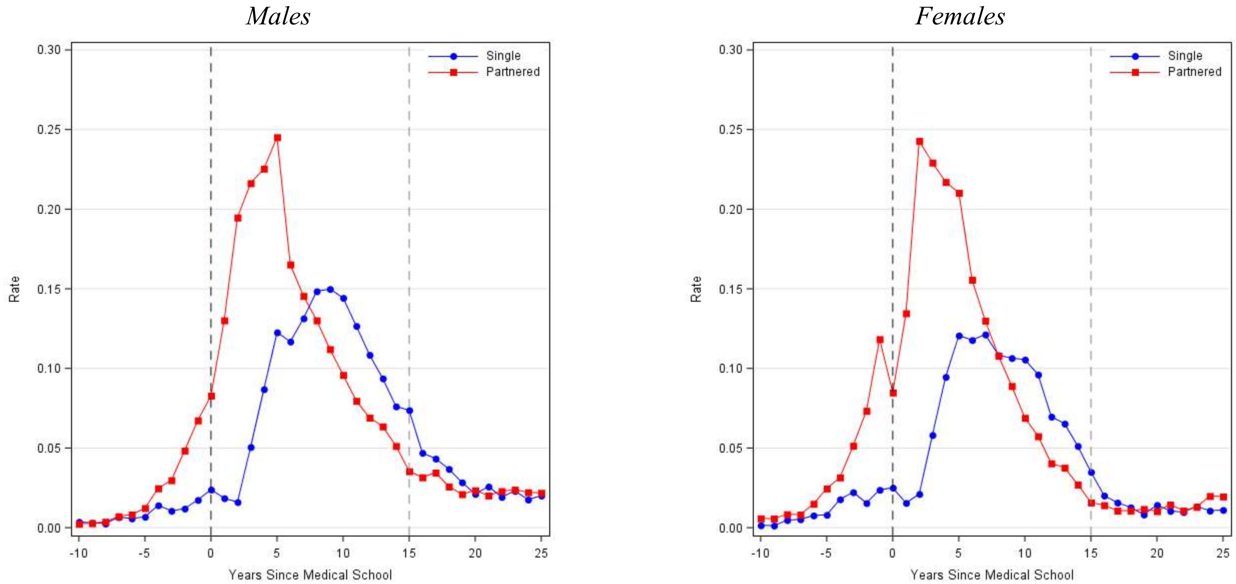


B. Change in Partnership Status



Notes: Panel A plots the probability of having a partner, and panel B plots an indicator for a change in partnership status between consecutive periods. The sample includes physicians who graduated from medical school in years 1980-2000.

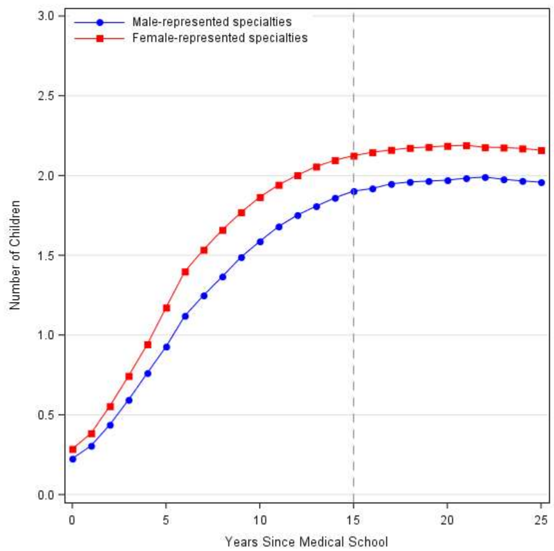
Appendix Figure A.6: Timing of fertility



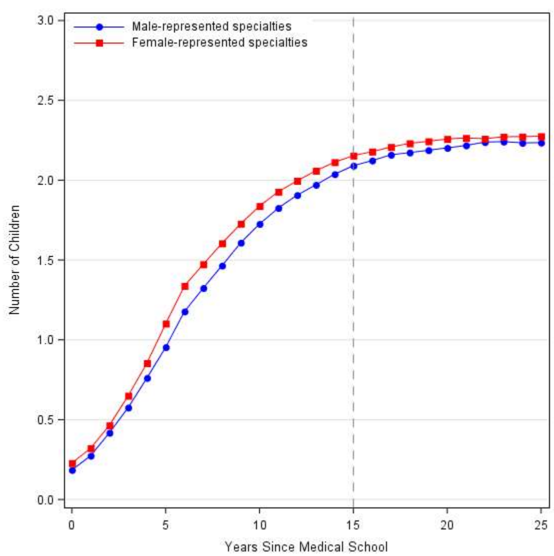
Notes: This figure plots the timing of fertility split by whether the physician was partnered or single at baseline. Fertility is defined as having a child in a given year. The sample includes physicians who graduated from medical school in years 1980-2000.

Appendix Figure A.7: Number of Children by Specialty Type

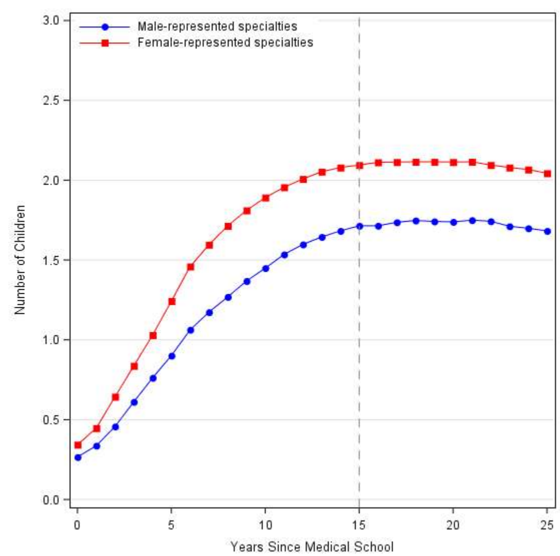
All Physicians



Males



Females

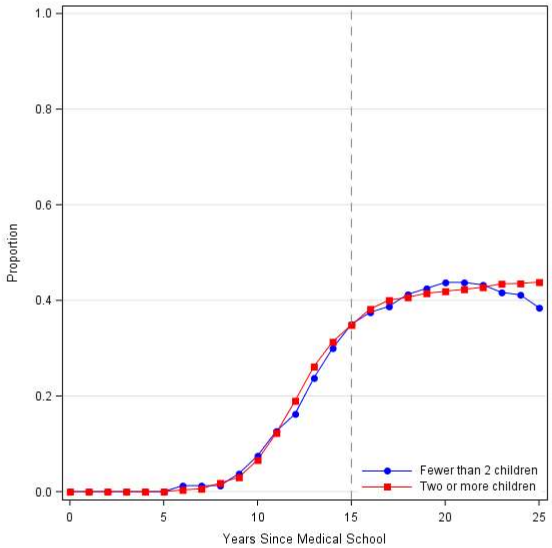


Notes: This figure plots the evolution of the number of children by whether the physician chose to specialize in a female-represented specialty or a male-represented specialty. The sample includes physicians who graduated from medical school in years 1980-2000.

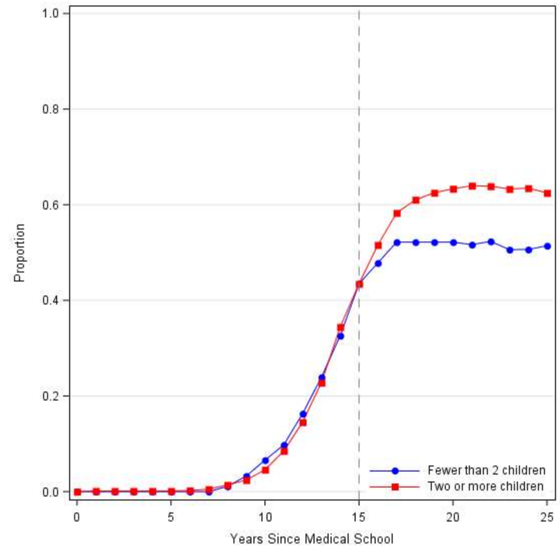
Appendix Figure A.8: Timing of Specialty by Number of Children for Those with a PhD

A. Female-Represented Specialties

Males

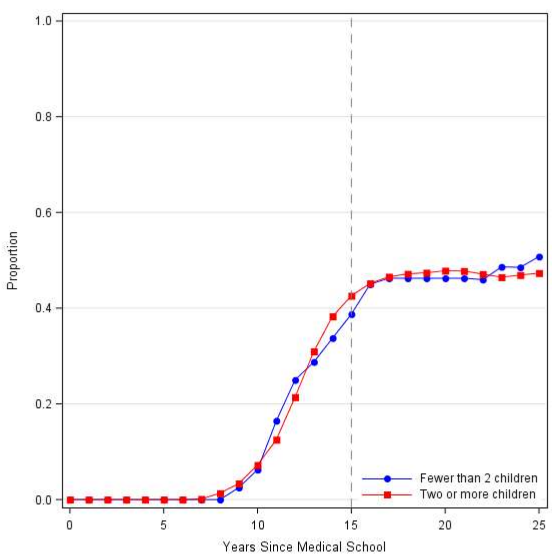


Females

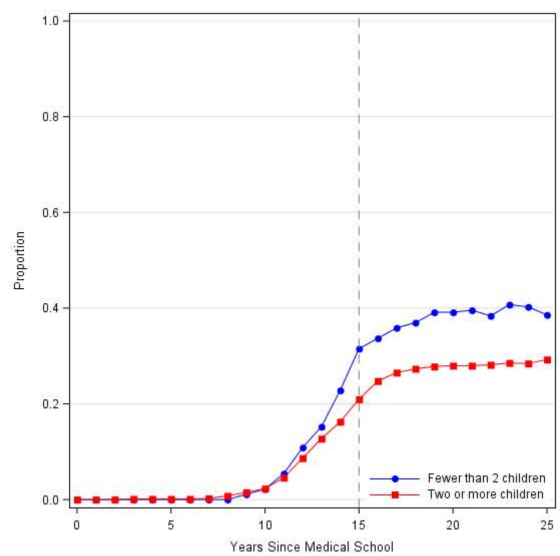


B. Male-Represented Specialties

C. Males



Females



Notes: This figure plots the timing of medical specialty completion among physicians who obtained a medical PhD, split by whether the physician has high or low fertility relative to the sample median of 2 children. Panel A studies female-represented specialties, and panel B studies male-represented specialties. The sample includes physicians who graduated from medical school in years 1980-2000 and obtained a medical PhD at any point in time.

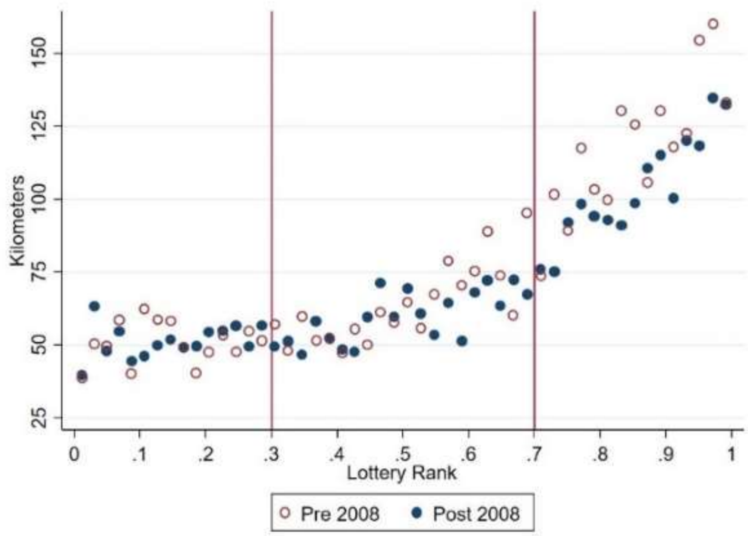
Appendix B: Patterns in Assignment to Internships and the First Stage

Appendix B.1: Internship Location Choices Over Time

The effective choice and assignment patterns of the medical internships display close similarities over the years, which is consistent with students' reluctance to intern in remote and rural areas. Throughout the years, geographic dispersion and relocation of graduating students have been a key dimension of variation that the lottery has created across the lottery rank distribution. We calculate, for each student, the distance between their municipality of residence at the time of the lottery and their municipality of work at the time of the internship, which captures their "relocation distance." Appendix Figure B.1 plots a graduating student's relocation distance against the student's lottery rank, where we split cohorts around 2008 (when the exact allocation process changed due to digitization). There is a clear gradient such that the relocation distance for those with better lottery numbers (lower ranks) is significantly shorter than for those with worse lottery numbers (higher ranks). This mirrors the underlying motivation for the lottery-based system, as it reveals interns' distaste for locating in rural labor markets when they get to choose.

The persistence of location preferences over the years, as they are revealed through choices, can be also shown in the following way. Let us characterize the desirability of a labor market (i.e., a county) based on the average lottery rank of the interns who choose to sort into it. We construct these desirability rankings for both earlier and later cohorts and compare across them in panel A of Appendix Figure B.2. Locations are effectively valued over the years to a similar extent.

Appendix Figure B.1: Relocation Distance



Notes: In this figure, we calculate, the distance (in kilometers) between graduating students' municipality of residence at the time of the lottery and their municipality of the internship, which captures their "relocation distances." We plot a graduating student's relocation distance against the student's lottery rank, where we split cohorts around 2008, when the process was digitized. Internship location is based on the physician's workplace in year one following the lottery (at the municipality level), as reported in annual employment registers as of the month of November.

Appendix B.2: Strategic Choice Considerations

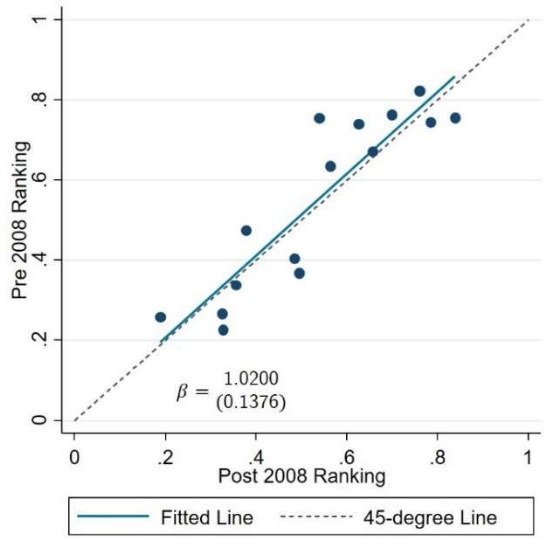
It is useful to discuss some potential choice and prioritization considerations, which could result from the incentives embedded in the choice processes we described, and to investigate how they play out in practice. That said, it is important to note that strategic behavior is not going to affect the validity of our identification of the effects of initial labor market sorting. This is because our choice for the main research design rests on reduced-form effects of the randomized lottery numbers. Still, describing these aspects is potentially informative for understanding the empirical context and for interpreting our findings. We use the information on the full formal rankings of local labor markets provided by the earlier cohorts in these investigations.

Given the structure of the matching process, individuals' equilibrium best-response strategy at each stage is to choose the option that maximizes their expected utility payoff based on their individual preferences and their expectations of other students' equilibrium play. For the later cohorts, this simply implies choosing their most preferred option among the options that are still available at the time they make their choice. For the earlier cohorts, there are additional potential considerations to take into account. To the extent that differential job aspects within a county play a role in ranking preferences (that is, aspects that go beyond the local labor market and its typical internship-related characteristics), the process implies that, at the first step of ranking counties, some consideration may be given to one's place in line for making a choice. For example, it may be preferable (along some job dimension) to be first in line in a worse labor market than last in line in a better labor market.

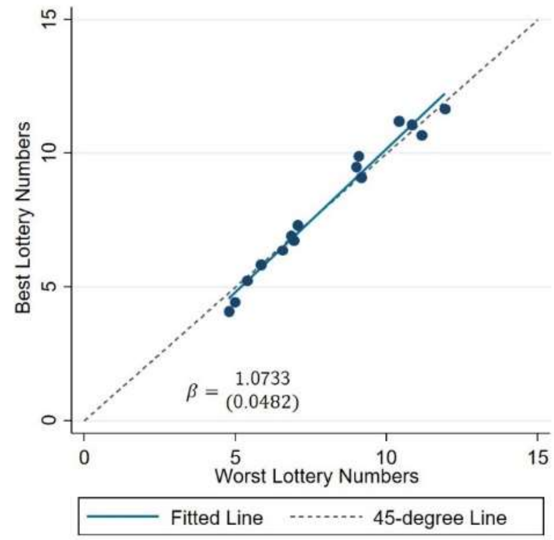
To test how this conjecture may play out in practice, we consider the rankings by those with the best lottery numbers as compared to the rankings by those with the worst lottery numbers. Specifically, we use graduates with lottery ranks in the highest 30 percent and the lowest 30 percent to be consistent with our main analysis. To the degree that students view their position in line for making a choice within a market as important, we would expect systematic differences in rankings over labor markets across the two groups. If, on the other hand, the choice of a local labor market is what dominates students' preferences regarding where to intern, we would expect similarities in their overall rankings. Panel B of Appendix Figure B.2 compares the average rankings of labor markets across the two groups. Each dot represents a local market, and we plot the fitted line as well as the 45-degree line, which is the benchmark under non-differential rankings. We also report the slope of the fitted line, where the benchmark null of non-differential rankings is one. The figure is consistent with the second hypothesis, i.e., that the choice of labor market itself leads students' rankings in the first step of the allocation process. The average rankings of markets across the two groups line up around the 45-degree line, and we cannot reject the benchmark null of a coefficient of one. The importance of location in students' preferences and choices is further underscored when we analyze the quasi-experiment's first stage across lottery rank groups.

Appendix Figure B.2: Labor Market Rankings

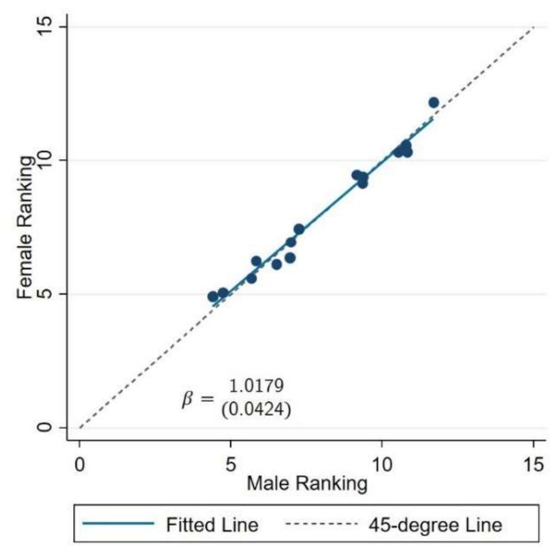
A. Over Time



B. Best vs. Worst Lottery Numbers



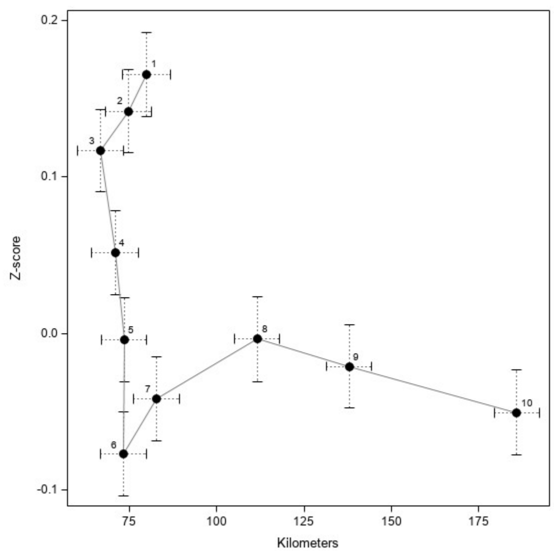
C. By Gender



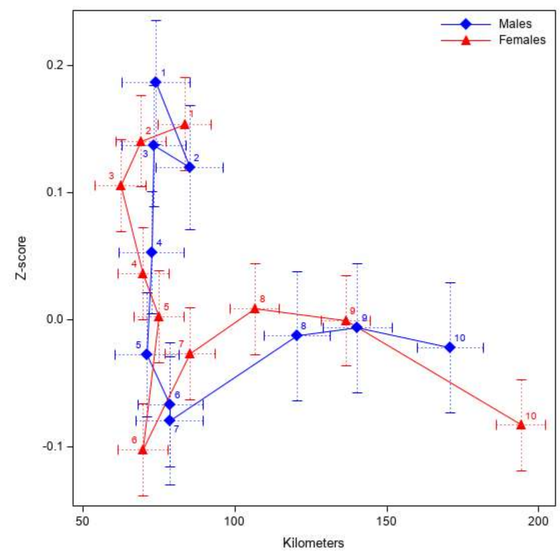
Notes: This figure makes several comparisons of the effective rankings of local labor markets. In panel A, we characterize the desirability of a labor market (i.e., a county) based on the average lottery rank of the interns who sort into it. We then compare the average rankings across earlier cohorts and later cohorts. In panels B-C, we use the information we have for earlier cohorts about students' binding pre-placement rankings of all local labor markets, as reported in priority lists. Panel B compares the average rankings of those with the best lottery numbers (the bottom 30 percent) with the average rankings of those with the worst lottery numbers (the top 30 percent). Panel C compares females' and males' priority rankings over entry-level local labor markets. We assign each local labor market its average priority by gender, and we then compare these priority rankings across males and females. In all panels, each dot represents a local labor market. We plot the fitted line, as well as the 45-degree line, which is the benchmark under non-differential rankings by gender. We also report the slope of the fitted line, where the benchmark of non-differential ranking is one.

Appendix Figure B.2: Distance and Quality

A. Overall Sample



B. Split by Gender



Notes: This figure replicates panels C-D of Figure 3, but when we group subjects into ten equal-sized bins based on their lottery ranks. Each dot represents a decile (whose number is displayed in the figure), and it plots the average values within that decile for the internship characteristics of relocation distance (on the x-axis) and a z-score of quality (on the y-axis), along with their corresponding 95-percent confidence intervals.

Appendix Table B.1: Distance and Quality—Comparability across Gender

	Distance (Km.) (1)	Quality (Z-score) (2)
Control	78.1	0.148
	(3.3)	(0.014)
Middle Group	75.2	-0.029
	(2.8)	(0.012)
Treatment	142.9	-0.014
	(3.3)	(0.015)
Female x Control	-5.117	-0.014
	(4.2)	(0.018)
Female x Middle Group	0.726	0.007
	(3.6)	(0.016)
Female x Treatment	0.983	-0.012
	(4.2)	(0.018)
Individuals	6,689	8,063

Notes: This table shows there are no differential first stage patterns across gender along the dimensions of relocation distance and internship quality, as displayed in panel D of Figure 3. Robust standard errors are reported in parentheses.

Appendix Table B.2: Characteristics of University Hospitals and Rural Locations

A. Characteristics of University Hospitals

		Non-University	University	Difference	*p*-value
Scale	Unique Patients	42,403	86,437	44,034	<0.0001
	Admissions	82,741	160,380	77,639	<0.0001
	Procedures	28,947	64,450	35,503	0.0001
Technology	Unique Procedures	816	1,485	669	<0.0001
	CT Scanner (probability)	0.75	0.98	0.23	0.0002
	CT Scans	12,839	41,472	28,633	0.0202
	MRI Scanner (probability)	0.58	0.98	0.41	<0.0001
	MRI Scans	6,380	24,678	18,298	0.0018
Human Capital	Medical Specialties	9.9	16.5	6.6	<0.0001
	Specialists with PhD (share)	0.069	0.156	0.087	<0.0001

B. Characteristics of Rural Locations

		Urban	Rural	Difference	*p*-value
Demographics	Population density (capita per sq km)	1,681	83	-1,598	0.0045
	Population size (capita)	165,284	53,849	-111,435	0.0027
	College degree (%, ages 25-64)	32.8	20.3	-12.5	<0.0001
	DI recipients (%, ages 17-64/66)	5.9	8.4	2.5	<0.0001
	Annual income (DKK, ages 25-59)	396,200	349,271	-46,929	<0.0001
Health/Healthcare	Interning in a University Hospital	0.3833	0.2090	-0.1743	0.0110
	Primary care expenditure per capita (DKK)	450	617	167	<0.0001
	Hospital visits per capita	0.84	0.97	0.13	<0.0001
	Daily smokers, %	16.3	18.7	2.4	<0.0001
Amenities/Norms	Home prices per square meter (DKK)	15,674	7,484	-8,190	<0.0001
	Revenue from income tax per capita (DKK)	39,352	36,087	-3,265	0.0041
	Places in daycare (%, ages 0-2)	40.0	22.7	-17.4	<0.0001
	School GPA (9th Grade Mandatory Finishing Exam)	7.41	7.41	-0.01	0.9264
	Expenditure on culture, sports, and leisure (per capita)	1,693	1,477	-216	0.0019
	Women elected officials (%)	34.2	27.8	-6.3	0.0003
	Parental leave, males (z-score)	0.023	-0.066	-0.089	0.0003
	Parental leave, females (z-score)	-0.017	0.054	0.071	<0.0001

Notes: Panel A provides characteristics of hospitals (with a total of 51 nationally), split by whether they are non-university or university hospitals. We use data from the national patient register, the registries for income and education, and the authorization register. Panel B provides characteristics of rural versus urban municipalities, where the classification follows the formal definitions used by the Danish Economic Councils (2015). We use data from: "Municipal Key Figures," Ministry of Interior Affairs and Housing ("Kommunale Nøgletal," Indenrigs- og boligministeriet); "Housing Market Statistics," Finance Denmark ("Boligmarkedsstatistikken," Finans Danmark); "National Goals," Ministry of Health ("Nationale mål," Sundhedsministeriet); "Municipality Report - Grade Point Averages of Mandatory Finishing Exams in 9th Grade," Ministry of Children and Education ("Kommunerapport - Gennemsnit i bundne prøver - 9. Klasse," Børne- og Undervisningsministeriet); and the absence, income, population, education, national health insurance, and national patient registers.

Appendix C: Lottery Verification and Summary Statistics

Appendix Table C.1: Verification of Lottery

	All	Males	Females
	(1)	(2)	(3)
Gender	0.0071		
	(0.0060)		
Age	0.0004	-0.0002	0.0010
	(0.0014)	(0.0020)	(0.0018)
Partnered	0.0079	0.0091	0.0074
	(0.0061)	(0.0097)	(0.0079)
Number of Children	-0.0048	-0.0133	-0.0011
	(0.0068)	(0.0118)	(0.0083)
GPA rank	0.0032	0.0012	0.0053
	(0.0105)	(0.0163)	(0.0137)
Rural	0.0099	0.0329	-0.0018
	(0.0210)	(0.0373)	(0.0254)
University Hospital	-0.0024	-0.0096	0.0029
	(0.0070)	(0.0108)	(0.0092)
Individuals	9,920	3,908	6,012
R-squared	0.0004	0.0009	0.0002
F	0.52	0.55	0.25
Prob > F	0.8187	0.7673	0.9613

Notes: This table tests the validity of the lottery in terms of random assignment. We run specifications that regress the graduating physicians' lottery rank on baseline characteristics available in our data. These include gender, age, an indicator for having a registered partner, number of children in the household, high school GPA rank, an indicator for residing in a rural municipality, and an indicator for having an employment at a university hospital. Robust standard errors are reported in parentheses, and we also report the p-value of the F-test for the joint predictive power of the specifications we run.

Appendix Table C.2: Analysis Sample Summary Statistics

	Control	Middle	Treatment	Difference Control-Middle	*p*-value Control-Middle	Difference Control-Treatment	*p*-value Control-Treatment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Overall Sample							
Female	0.5999	0.6062	0.6114	-0.0063	0.5909	-0.0115	0.3576
Partnered	0.4964	0.4971	0.5079	-0.0008	0.9497	-0.0115	0.3700
Age	27.5096	27.5144	27.5206	-0.0048	0.9345	-0.0111	0.8606
GPA Rank	0.5025	0.4986	0.5043	0.0039	0.5740	-0.0018	0.8082
Number of Children	0.1877	0.1902	0.1818	-0.0025	0.8385	0.0058	0.6519
Rural	0.0241	0.0258	0.0279	-0.0016	0.6655	-0.0037	0.3635
University Hospital	0.2255	0.2277	0.2218	-0.0021	0.8319	0.0037	0.7288
Number of Individuals	3,024	3,997	3,052				
B. Males							
Partnered	0.4636	0.4644	0.4696	-0.0008	0.9671	-0.0060	0.7681
Age	27.6455	27.5222	27.5995	0.1232	0.2019	0.0460	0.6665
GPA Rank	0.5064	0.5026	0.4983	0.0038	0.7380	0.0081	0.5041
Number of Children	0.1551	0.1352	0.1352	0.0199	0.2503	0.0199	0.2727
Rural	0.0198	0.0203	0.0236	-0.0005	0.9264	-0.0038	0.5262
University Hospital	0.2488	0.2541	0.2243	-0.0054	0.7463	0.0245	0.1587
Number of Individuals	1,210	1,574	1,186				
C. Females							
Partnered	0.5182	0.5184	0.5322	-0.0002	0.9911	-0.0140	0.3964
Age	27.4190	27.5093	27.4705	-0.0903	0.2161	-0.0516	0.5047
GPA Rank	0.5000	0.4960	0.5082	0.0039	0.6562	-0.0082	0.3848
Number of Children	0.2094	0.2259	0.2115	-0.0165	0.3278	-0.0021	0.9032
Rural	0.0270	0.0293	0.0305	-0.0023	0.6565	-0.0035	0.5216
University Hospital	0.2100	0.2105	0.2203	-0.0004	0.9716	-0.0102	0.4505
Number of Individuals	1,814	2,423	1,866				

Notes: This table provides summary statistics for the analysis sample in the last full year in medical school, defined as the baseline year, which is the calendar year prior to the internship lottery. Panel A provides statistics for the entire sample, and panels B and C split the sample by gender. Characteristics include gender, age, an indicator for having a registered partner (in cohabitation or marriage), number of children, high school GPA rank, an indicator of residing in a rural area, and an indicator for holding employment at a university hospital. Column 1 displays means for our control group, column 2 displays means for our middle group, and column 3 displays means for our treatment group. Column 4 provides the differences between column 1 and column 2. Column 5 reports the *p*-values of the test statistics of the differences in column 4 (*t*-statistics for continuous variables and *z*-statistics for binary variables). Column 6 provides the differences between column 1 and column 3. Column 7 reports the *p*-values of the test statistics of the differences in column 6 (*t*-statistics for continuous variables and *z*-statistics for binary variables).

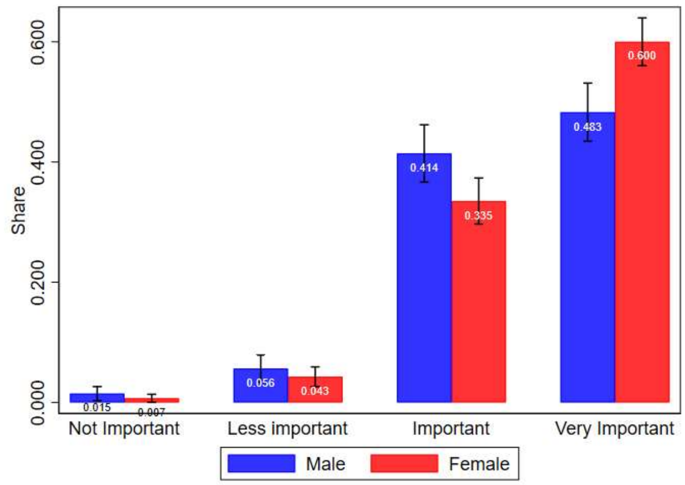
Appendix D: Male-Represented and Female-Represented Specialties

Appendix Table D.1: Medical Specialties Grouping

Specialty	Specialty Group
Panel A: Male-Represented	
Thorax Surgery	Surgery
Orthopedic Surgery	Surgery
General Surgery	Surgery
Neurosurgery	Surgery
Internal Medicine	Internal medicine
Clinical Biochemistry	Transverse specialties
Otorhinolaryngology	Surgery
Internal Medicine: Cardiology	Internal medicine
Ophthalmology	Surgery
Vascular Surgery	Surgery
Anesthesiology	Transverse specialties
Internal Medicine: Gastroenterology and Hepatology	Internal medicine
Urology	Surgery
Panel B: Female-Represented	
Internal Medicine: Hematology	Internal medicine
Clinical Microbiology	Transverse specialties
Neuro Medicine	Other
Clinical Immunology	Transverse specialties
Clinical Physiology and Nuclear Medicine	Transverse specialties
Occupational Medicine	Other
General Medicine	General medicine
Internal Medicine: Rheumatology	Internal medicine
Internal Medicine: Pulmonary Diseases	Internal medicine
Radiology	Transverse specialties
Internal Medicine: Endocrinology	Internal medicine
Plastic Surgery	Surgery
Psychiatry	Psychiatry
Internal Medicine: Nephrology	Internal medicine
Dermato-Venerology	Other
Clinical Pharmacology	Transverse specialties
Internal Medicine: Infectious Diseases	Internal medicine
Gynecology and Obstetrics	Surgery
Pathological Anatomy and Cytology	Transverse specialties
Public Medicine	Other
Pediatrics	Other
Clinical Oncology	Other
Internal Medicine: Geriatrics	Internal medicine
Forensic medicine	Other
Clinical Genetics	Transverse specialties
Child and Youth Psychiatry	Psychiatry

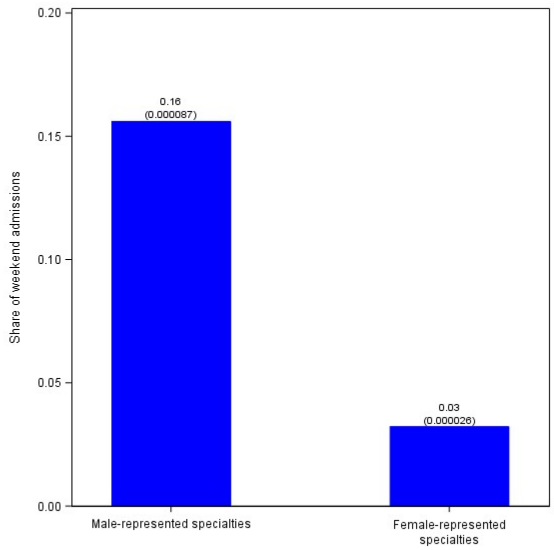
Notes: This table classifies medical specialties by gender representativeness based on the share of females within a specialty relative to their overall proportion. “Female-represented specialties” are specialties with a female share that is higher than this proportion, and “male-represented specialties” are specialties with a female share that is lower than this proportion.

Appendix Figure D.1: Importance of Shift Burden in Female Representation in Medical Specialties



Notes: This figure builds on survey data collected in Korreman (1993). The underlying micro data and documentation is publicly available via the Danish National Archives (https://digidata.rigsarkivet.dk/aflevering/36609 - Kønsaspekter i Lægers Karriereforløb 1993 – Gender Aspects in the Careers of Doctors, 1993). The data sample consists of 1,000 physicians who were interviewed in 1993 ten years after graduating medical school (i.e., senior physicians in our main analysis sample that runs from 2001-2022) and across four Nordic countries (Denmark, Sweden, Norway, and Finland). The figure shows the distribution of answers to the following question (question 50): "There is not an equal distribution of males and females across medical specialties. How important do you believe the feature 'lower shift burden' is for explaining why some specialties have a large female share?"

Appendix Figure D.2: Weekend Admissions by Specialty Type



Notes: This figure shows the share of admissions and patient contacts that take place during weekends in male-represented specialties versus female-represented specialties. Robust standard errors are reported in parentheses.

Appendix E: Exit Surveys

Appendix E.1: Physician Exit Surveys—Details

This appendix provides background information on the exit surveys. The questions in the surveys are grouped into seven overall categories. The survey questions changed in 2016, but the seven categories remained similar. Appendix Tables E.1 and E.2 show the groupings of the individual questions from the old and new questionnaires into the seven overall categories. The individual questions are provided in Appendix Tables E.3-E.6 in Danish (original) and English (translated). To provide numerical scoring of a department, interns also report the names of their assigned supervisor and program department chair. We use the names to deduct their gender. To do so, we construct an algorithm based on first names, which works as follows. We construct a gender probability using the first names of all doctors in the authorization register, which includes their names and gender. A first name is defined as "male" if more than 70 percent of the individuals with the given first name are males, and, accordingly, a first name is defined as "female" if less than 30 percent of the individuals with the given first name are males. We extract the names of supervisors and program department chairs from the exit surveys and match their first names to the gender proxy constructed from the authorization register.

Appendix Table E.1: Survey Evaluation Categories until 2015

Group	English (translated)	Danish (original)	Questions
1	Introduction	Introduktion	1-2
2	Supervision	Uddannelsesprogram	3-6
3	Daily guidance	Vejleder (Praksistutor)	7-11
4	Work organization	Arbejdstilrettelæggelse	12-17
5	Education	Øvrige forhold	18-22
6	Education	Samlet vurdering	23
7	Overall Assessment	Samlet vurdering	24

Notes: The evaluation scales range from 1 to 9. The individual questions are reported in Appendix Tables E.3 and E.4.

Appendix Table E.2: Survey Evaluation Categories from 2016

Group	English (translated)	Danish (original)	Questions
1	Introduction	Introduktion	1-3
2	Supervision	Uddannelsesvejledning	1-7
3	Daily guidance	Daglig vejledning	8-13
4	Work organization	Arbejdstilrettelæggelse	12-17
5	Education	Konference/undervisning	18-20
6	Work climate	Arbejdsklima	21-24
7	Overall Assessment	Øvrige	25-26

Notes: The evaluation scales range from 1 to 6. The individual questions are reported in Appendix Tables E.5 and E.6.

Appendix Table E.3: Questions in Evaluations until 2015, Danish

1	Hvordan vurderer du kvaliteten af introduktionen på uddannelsesstedet?
2	Fulgte du introduktionsprogrammet?
3	Hvordan vurderer du kvaliteten af uddannelsesprogrammet?
4	Svarer indholdet til målbeskrivelsens krav?
5	Svarede uddannelsesforløbet til uddannelsesprogrammet?
6	Har du indfriet checklistens delpunkter?
7	Hvordan var kvaliteten af vejlederens indsats i forhold til din uddannelse?
8	Anvendtes samtaleindholdet (og uddannelsesplanen) i praksis?
9	Hvordan var graden af supervision?
10	Var vejlederen tilstede i tilstrækkeligt omfang?
11	Anviste vejlederen dig uddannelsesrelevante arbejdsområder?
12	Hvordan vurderer du graden af selvstændighed i det kliniske arbejde?
13	Hvordan vurderer du arbejdsbyrden?
14	Var arbejdet tilrettelagt med rimeligt hensyntagen til uddannelsen?
15	Hvordan var vagthyppigheden i forhold til vagtens uddannelsesværdi?
16	Hvordan vurderer du uddannelsesværdien af vagtarbejdet?
17	Hvordan vurderer du uddannelsesværdien af dagarbejdet?
18	Deltog du i forskning/kvalitetsudviklingsarbejde?
19	Deltog du i administrativt arbejde?
20	Deltog du i afdelingens formaliserede undervisning?
21	Underviste du selv?
22	Hvordan vurderer du afdelingens uddannelsesmiljø/prioritering?
23	Hvordan vurderer du uddannelsesstedets samlede uddannelsesindsats?
24	Hvordan vurderer du dit samlede uddannelsesudbytte under ansættelsen?
Text	Vejleder
Text	Uddannelsesansvarlig

Appendix Table E.4: Questions in Evaluations until 2015, English

1	How do you assess the quality of the introduction at the place of education?
2	Did you follow the introductory program?
3	How do you rate the quality of the training program?
4	Does the content correspond to the requirements of the goal description?
5	Did the training course correspond to the training program?
6	Have you met the checklist sub-items?
7	How was the quality of the supervisor's efforts in relation to your education?
8	Was the interview content (and the training plan) used in practice?
9	How was the degree of supervision?
10	Was the supervisor present to a sufficient extent?
11	Did the supervisor instruct you in training-relevant work areas?
12	How do you assess the degree of independence in the clinical work?
13	How do you assess the workload?
14	Was the work organized with reasonable consideration for the education?
15	How was the shift frequency in relation to the shift's educational value?
16	How do you assess the educational value of the shift work?
17	How do you assess the educational value of day work?
18	Did you participate in research/quality development work?
19	Did you participate in administrative work?
20	Did you participate in the department's formalized teaching?
21	Did you teach yourself?
22	How do you assess the department's educational environment/priorities?
23	How do you assess the educational institution's overall educational efforts?
24	How do you assess your overall educational output during employment?
Text	Mentor
Text	Head of Educational Program

Appendix Table E.5: Questions in Evaluations from 2016, Danish

1	Uddannelsesstedet og jeg har afstemt forventninger til uddannelseselementet ved introduktionen.
2	Jeg blev introduceret til de opgaver, jeg skulle varetage.
3	Min hovedvejleder og jeg samarbejdede om at udarbejde min individuelle uddannelsesplan.
4	Mit behov for uddannelsesvejledning er blevet opfyldt.
5	De planlagte kompetencevurderinger er blevet gennemført.
6	Kompetencevurderinger er blevet efterfulgt af feedback.
7	Jeg er blevet tilbudt karrierevejledning svarende til mit behov.
8	Jeg har fået feedback i forhold til min evne til at samarbejde med sundhedsprofessionelle.
9	Jeg har fået feedback i forhold til min evne til at agere professionelt.
10	Jeg har fået feedback i forhold til min evne til at kommunikere.
11	Jeg har fået mulighed for at udvikle mig som leder/administrator og organisator.
12	Jeg har fået supervision svarende til mit behov i det daglige arbejde.
13	De daglige læringsmuligheder er blevet udnyttet.
14	De daglige vejledere har været til at få fat på, når jeg havde behov for det.
15	Arbejdstilrettelæggelsen har tilgodeset, at jeg også har varetaget opgaver, der er relevante for, at jeg har kunnet opnå kompetencerne som angivet i uddannelsesprogrammet.
16	I arbejdstilrettelæggelsen er det blevet prioriteret, at der har været progression i min kompetenceudvikling.
17	I arbejdstilrettelæggelsen er vejledersamtaler blevet prioriteret.
18	Jeg har fået mulighed for at udvikle mig som underviser.
19	Jeg har haft mulighed for at deltage i uddannelsesstedets undervisningstilbud.
20	Jeg har haft udbytte af uddannelsesstedets konferencer.
21	Jeg har oplevet, at der er en gensidigt respektfuld omgangstone på uddannelsesstedet.
22	Jeg har været tryg ved at stille spørgsmål til kollegaer.
23	Jeg har kunnet diskutere svære problemstillinger med mine kollegaer.
24	Jeg har oplevet, at jeg har arbejdet som del af et arbejdsfællesskab.
25	Samlet set har uddannelsesstedets indsats været tilfredsstillende.
26	Mit samlede uddannelsesmæssige udbytte har været tilfredsstillende.
Text	Vejleder
Text	Uddannelsesansvarlig

Appendix Table E.6: Questions in Evaluations from 2016, English

1	The place of education and I have reconciled expectations of the educational element at the time of the introduction.
2	I was introduced to the tasks I had to undertake.
3	My main supervisor and I collaborated on preparing my individual education plan.
4	My need for educational guidance has been met.
5	The planned competency assessments have been carried out.
6	Competence assessments have been followed by feedback.
7	I have been offered career guidance according to my needs.
8	I have received feedback regarding my ability to collaborate with health professionals.
9	I have received feedback in relation to my ability to act professionally.
10	I have received feedback in relation to my ability to communicate.
11	I have had the opportunity to develop as a leader / administrator and organizer.
12	I have received supervision according to my needs in the daily work.
13	The daily learning opportunities have been utilized.
14	The daily tutors have been available when I needed it.
15	The work organization has taken into account that I have also handled tasks that are relevant for me to have been able to achieve the competencies as stated in the training program.
16	In the work organization, it has been prioritized that there has been progression in my competence development.
17	In the work organization, supervisor feedback has been prioritized.
18	I have had the opportunity to develop as a teacher.
19	I have had the opportunity to participate in the educational offer of the educational institution.
20	I have benefited from the conferences of the educational institution.
21	I have experienced that there is a mutually respectful tone of voice at the place of education.
22	I have been comfortable asking questions to colleagues.
23	I have been able to discuss difficult issues with my colleagues.
24	I have experienced that I have worked as part of a working community.
25	Overall, the educational institution's efforts have been satisfactory.
26	My overall educational output has been satisfactory.
Text	Mentor
Text	Head of Educational Program

Appendix E.2: Exit Surveys and Inspector Evaluations

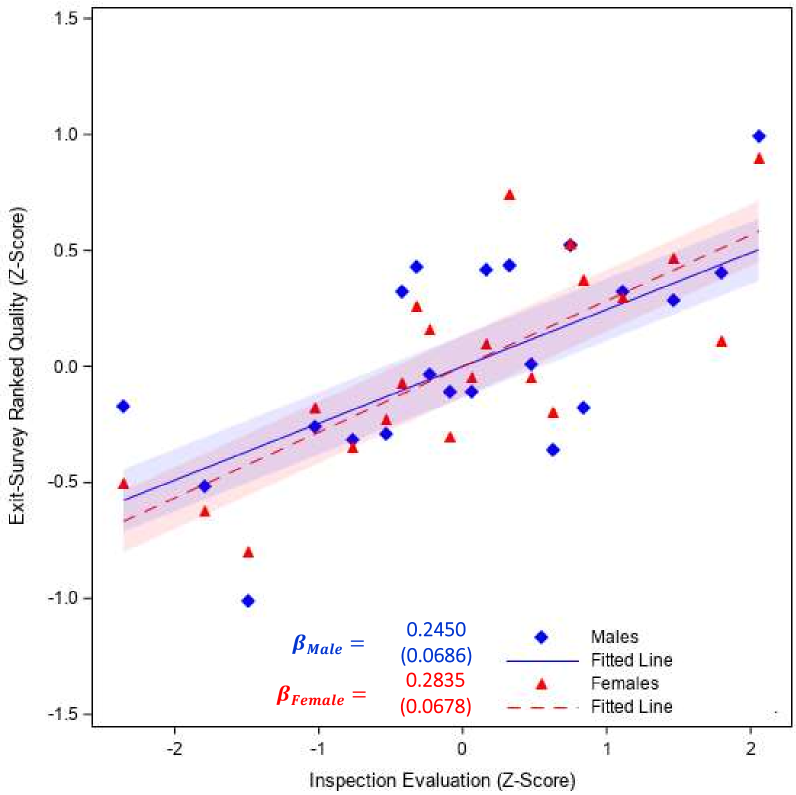
We use data from external inspections that the National Health Authority (NHA) conducts to assess the quality of the educational programs in hospital departments. In this appendix, we provide details about these assessments and study their correlation with the exit survey rankings.

The NHA has been conducting external inspections since 1997. Appointed by the NHA, the group of inspectors consists of impartial senior and junior physicians. The inspectors score the hospital department's performance in 16 categories (see panel A of Appendix Table E.7), and each category is scored on a 4-point scale (see panel B of Appendix Table E.7). For our purposes, we use inspectors' overall assessments of a hospital department's internship by summing over all categories. For more details, see: Inspektorordningen Håndbog, Sundhedsstyrelsen, 2016, https://www.sst.dk/da/Udgivelser/2016/Inspektorordningen-Haandbog.

The reports are publicly available on the NHA's website: https://www.sst.dk/da/inspektorrapporter. The NHA servers include inspections from 2013-2022 (where data from 1997-2012 have been erroneously deleted). We hand-code the hospital department IDs for each inspector report in order to link them to our data on the ranked quality from the interns' exit surveys. This provides us with inspector quality assessments of 202 hospital departments (61 percent of the internship positions).

In Appendix Figure E.1, we study the degree to which inspection assessments are predictive of how interns rank the quality of their internships in the exit surveys. We split the sample into 20 equal-sized bins based on the z-score of the external inspections, where the mean z-score of each bin is displayed on the x-axis. We then plot on the y-axis the average ranked quality from the exit surveys for each bin, split by gender.

Appendix Figure E.1: Associations between Exit Survey Evaluations and Inspector Evaluations



Notes: This figure displays the association between inspection assessments and interns' ranked quality of their internships. We split the sample into 20 equal-sized bins based on the z-score of the external inspections, where the mean z-score of each bin is displayed on the x-axis. We then plot on the y-axis the average ranked quality from the exit surveys for each bin, split by gender. We also plot the fitted lines along with 95-percent confidence intervals and report their slopes.

Appendix Table E.7: Inspector Evaluations

A. Performance Categories for Inspector Assessment

Category	Danish (original)	English (translated)
1	Introduktion til afdelingen	Introduction to the department
2	Uddannelsesprogram	Educational program
3	Uddannelsesplan	Education plan
4	Medicinsk ekspert - Læring i rollen som medicinsk ekspert	Medical expert - Learning the physician's role as a medical expert
5	Kommunikator - Læring i rollen kommunikator	Communicator - Learning the physician's role as a communicator
6	Samarbejder - Læring i rollen som samarbejder	Collaborator - Learning the physician's role as a collaborator
7	Leder/administrator - Læring i rollen som leder/administrator	Leader/administrator - Learning the physician's role as a leader/administrator
8	Sundhedsfremmer - Læring i rollen som sundhedsfremmer	Health promoter - Learning the physician's role as a health promoter
9	Akademiker - Læring i rollen som akademiker	Academic - Learning the physician's role as an academic
10	Professionel - Læring i rollen som professionel	Professional - Learning the physician's role as a professional
11	Forskning - Uddannelsessøgende lægers deltagelse i forskning	Research - Participation in research
12	Undervisning - som afdelingen giver	Teaching - provided by the department
13	Konferencernes - læringsværdi	The learning value of morning reports
14	Læring og kompetencevurdering	Learning and competence assessment
15	Arbejdstilrettelæggelse - Tilrettelæggelsen tager hensyn til videreuddannelsen af læger	Work organization - The organization takes postgraduate training of doctors into account
16	Læringsmiljøet på afdelingen	The learning environment in the department

B. Assessment Scoring Scale

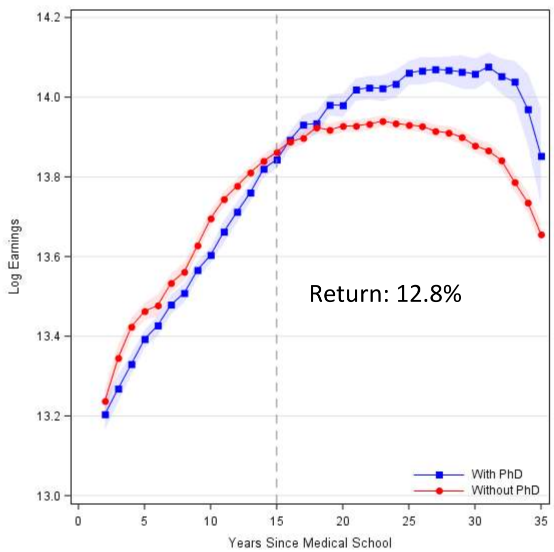
Score	Danish (original)	English (translated)
1	Særdeles problematisk	Extremely problematic
2	Utilstrækkelig	Inadequate
3	Tilstrækkelig	Adequate
4	Særdeles god	Extremely good

Appendix F: Earnings Profiles and Predictions

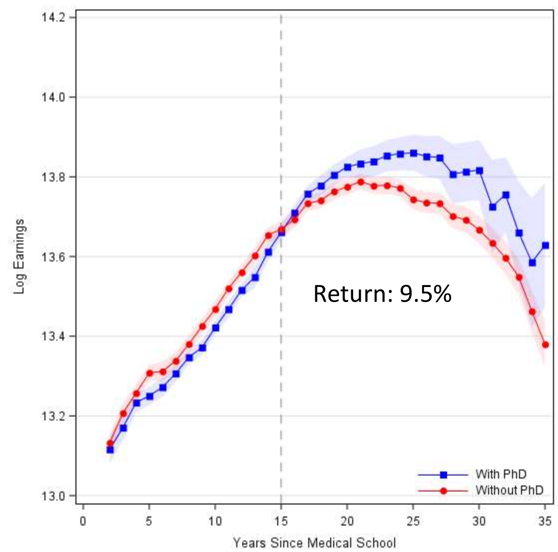
Appendix Figure F.1: Life-Cycle Log Earnings Trajectories

A. By Holding a Medical PhD

Males

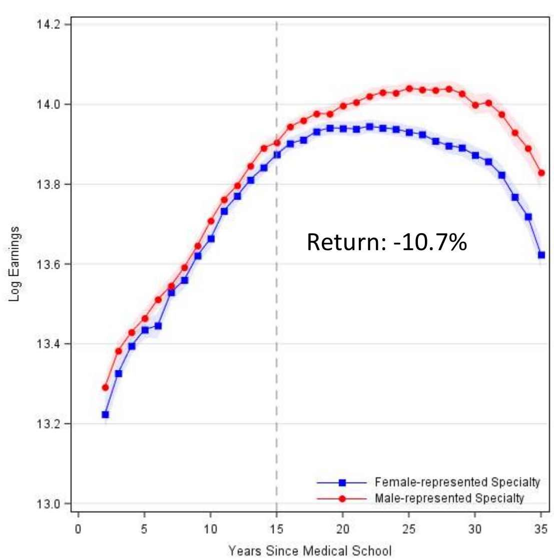


Females

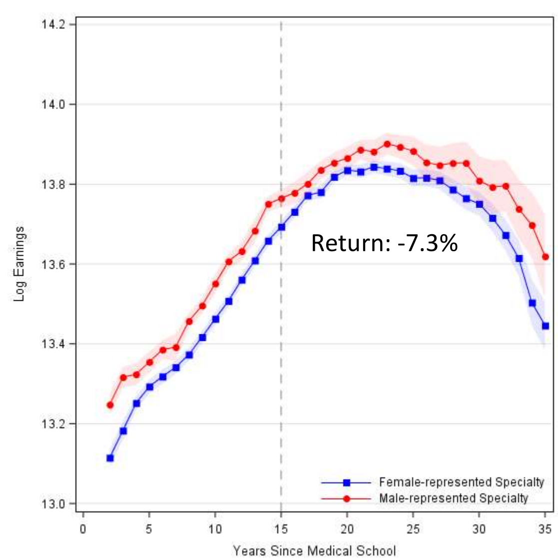


B. By Specialization in Male-Represented Specialties or Female-Represented Specialties

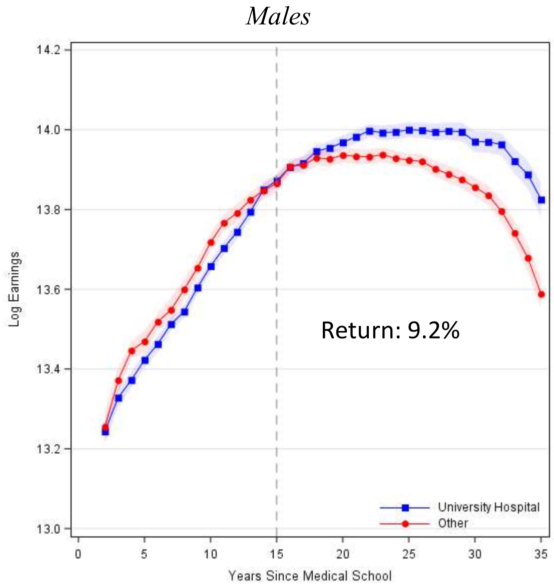
Males

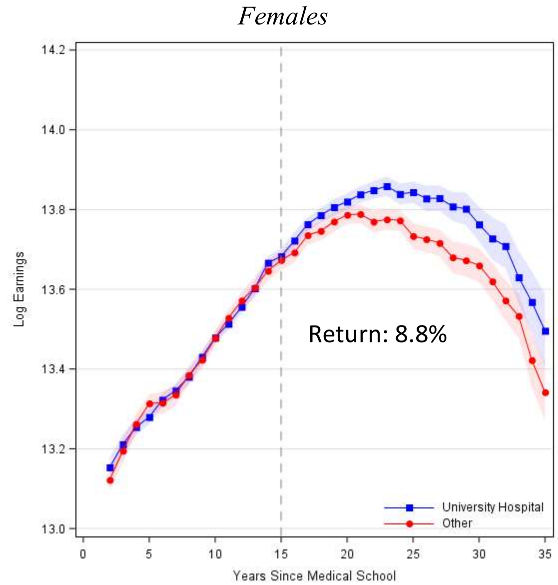


Females



C. By Employment at a University Hospital





Notes: This figure provides log earnings profiles from the first full year after medical school and up to 35 years after medical school using information from cohorts who graduated in the years 1970-2000. Earnings incorporate total compensation, including annual wage earnings, net income from self-employment, and labor market pension contributions (which are part of work compensation packages and analogous to employer contributions to 401(k)s). All compensation components are measured pre-taxes, which we consistently measure from 1995-2021. Earnings are measured in 2017 prices (deflated by the Danish regions' wage-index, see https://www.regioner.dk/aftaler-og-oekonomi/oekonomisk-vejledning/oekonomisk-vejledning-2024/) and we include observations with earnings above 1 DKK. Panel A splits physicians by whether they ever hold a medical PhD, panel B splits physicians by whether they chose a male-represented specialty or a female-represented specialty, panel C splits physicians by whether they were employed by a university hospital in year 15. On each plot we report the returns associated with making the labor market choice, averaged across the long-run of years 16-35.

Appendix F.1: Long-Run Earnings Predictions

To estimate the predicted long-run effects on earnings, we use the "surrogate index" method (Athey et al. 2019). This method was proposed as a solution to the common challenge in estimating longer-term impacts of treatments, where outcomes of interest are observed with a long delay. The idea is to combine several shorter-term outcomes into the "surrogate index," which is the predicted value of the longer-term outcome given the shorter-term outcomes (the "surrogates") based on long-run observational data. Athey et al. (2019) show that the average treatment effect on the surrogate index equals the treatment effect on the long-term outcome. This is the case under the assumption that the long-term outcome is independent of the treatment conditional on the surrogate index, which forms the "surrogacy condition." Our outcome is log earnings which incorporates total compensation including annual wage earnings, net income from self-employment, and labor market pension contributions (analogous to employer contributions to 401(k)s). Compensation is measured pre-tax in 2017-prices (deflated by the Danish regions' wage index) and we include observations with earnings above 1 DKK. We use information from the years 1995-2021 on an extended set of cohorts, who graduated in the years 1970-2000, so we could push predictions up to year 35 following graduation from medical school. Appendix Figure F.2 plots the earnings profiles for each cohort as well as their aggregation. The different cohorts have a typical overlapping pattern that together produces the usual inverse U-shape profile of earnings over the life cycle.

We follow the implementation in Athey et al. (2019) and estimate the following statistical models in two steps. First, let y_{it} represent the long-run outcome of interest and let $s_i = (s_{1i}, \dots, s_{mi})$ be the vector of intermediate outcomes. To construct the surrogate index estimator, we estimate the following OLS specification, separately for each year t after medical school and for each gender:

$$y_{it} = \delta_{0t} + \sum_{j=1}^{m} \delta_{jt} \times s_{ji} + \omega_{it}.$$

Specifically, we run in the first step:

$$\begin{aligned} Log(earnings)_{it} \quad = \quad & \delta_{0t} + \delta_{1t} PhD + \delta_{2t} FemaleRepSpecialty \; + \delta_{3t} MaleRepSpecialty \\ & + \delta_{4t} FemaleRepSpecialty \times PhD + \delta_{5t} MaleRepSpecialty \times PhD \\ & + \delta_{6t} Rural + \delta_{7t} University\ Hospital + \omega_{it}. \end{aligned}$$

We use as surrogates individual characteristics in year 15, when the main labor market choices in our application have reached their steady state (see Appendix A). For outcomes of medical specialty choice, we use the position in year 20. Recall from Section 3.2 that, for this central choice of long-run careers in the form of medical specialties, completion rates stabilize a few years later than PhD completion (see Appendix Figure A.3). While most individuals have completed their medical specialization by year 15, an additional 5-10 percent complete their specialization by year 20. This is particularly true for individuals who obtain a medical PhD, which is also correlated with specializing at higher rates in male-represented specialties relative to female-represented specialties. Second, the surrogate index for the long-run outcome y_{it} is calculated as the predicted value from these regressions, which we denote by $\hat{y}_{it}$. We construct this index for each individual in our experimental sample by calculating: $\hat{y}_{it} = \hat{\delta}_{0t} + \sum_{j=1}^{m} \hat{\delta}_{jt} \times s_{ji}$.

Practically, in the second step, we restrict the experimental sample to characteristics in year 15 to capture the intermediate outcomes s_{ji}. To account for the pattern from the observational data that 5-10 percent of the sample (particularly those with a medical PhD) is expected to complete their medical specialty between year 15 and 20, we adjust the calculation of $\hat{y}_{it}$ accordingly. Specifically, for physicians in the observational data, who have not yet specialized by year 15, we compute the probabilities of having specialized by year 20 in a male-represented specialty, a female-represented specialty, or remain without a specialty. We calculate these transition probabilities stratified by PhD status in year 15 and by gender. For instance, in the observational data we find that among women without a specialty and without a PhD by year 15, 36 percent obtain a female-represented specialty, 11 percent obtain a male-represented specialty, and 53 have no specialty by year 20. These shares capture the transition probabilities, which we use as weights to calculate the expected value of medical specialties $\hat{\delta}_{jt}$ in the earnings predictions $\hat{y}_{it}$ that we assign to individuals who have not yet obtained a specialty by year 15 in the experimental data.

Additionally, for 26 percent of the experimental sample we have missing information on whether they were employed at a university hospital or lived in a rural area in year 15. The missing information is coming particularly from the fact the employment data have two calendar years fewer than the other registers, and the demographic registers similarly having one fewer year. To fill in these missing values, we use the individual employment affiliation and indicators for living in rural areas in year 14 and, if that observation is missing too, we use employment affiliation and location information from year 13.

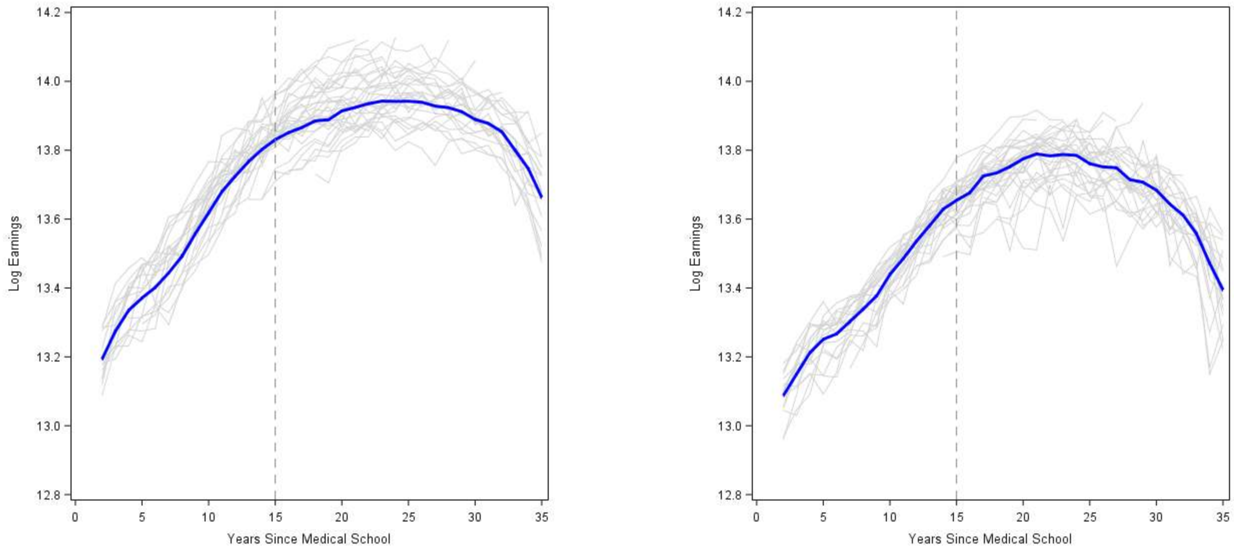
Finally, using the experimental sample, the average treatment effect on the long-term outcome is then estimated as the treatment effect of the quasi-experiment on the surrogate index, i.e., on the predicted value of the long-term outcome. Specifically, we estimate the average treatment effect on outcome y_{it} based on the following regression:

$$\hat{y}_{it} = \alpha_t^s + \beta_t^s \times Treat_i + \epsilon_{it},$$

where β_t^s is the parameter of interest. We bootstrap standard errors (with 5,000 iterations) to account for estimation error from both steps of the surrogacy analysis. We provide predictions of log earnings for years 2-35 after medical school (where year 2 is the first full year of employment).

Appendix Figure F.2: Earnings Profiles of Overlapping Cohorts by Gender

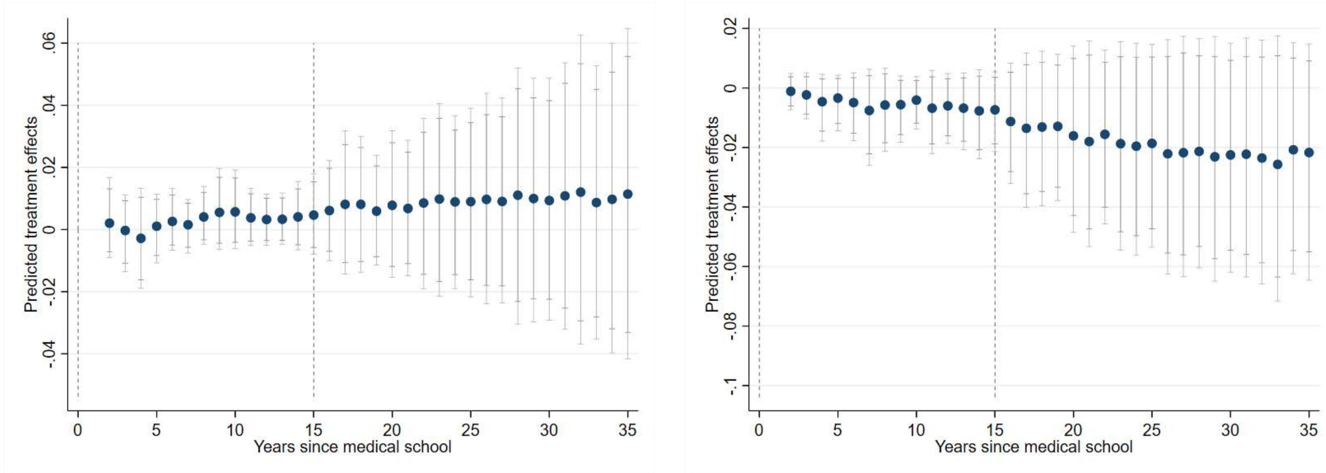
Males *Females*



Notes: This figure provides log earnings profiles from year 2 after medical school (the first full year of employment) up to 35 years after medical school using information from cohorts who graduated in the years 1970-2000. It displays the overlapping earnings profiles for each cohort (in gray) and their aggregation (in blue). Earnings incorporate total compensation, including annual wage earnings, net income from self-employment, and labor market pension contributions (which are part of work compensation packages and analogous to employer contributions to 401(k)s). All compensation components are measured pre-taxes, which we consistently measure from 1995-2021. Earnings are measured in 2017 prices (deflated by the Danish regions' wage-index) and we include observations with earnings above 1 DKK.

Appendix Figure F.3: Predicted Long-Run Treatment Effects—Middle vs. Control Groups

Males *Females*



Notes: This figure supplements panel B of Figure 8 and plots the estimated effects for the middle group on predicted earnings (the difference between the middle and the control group) and their 90-percent and 95-percent confidence intervals. Standard errors are bootstrapped to account for estimation error from the two steps of the surrogate index analysis.

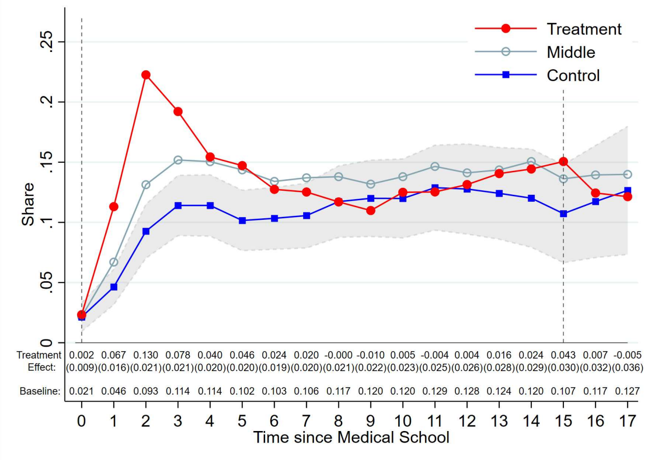
Appendix G: Additional Analysis

Appendix Figure G.1: Long-Run Treatment Effects by Baseline Partnership Status

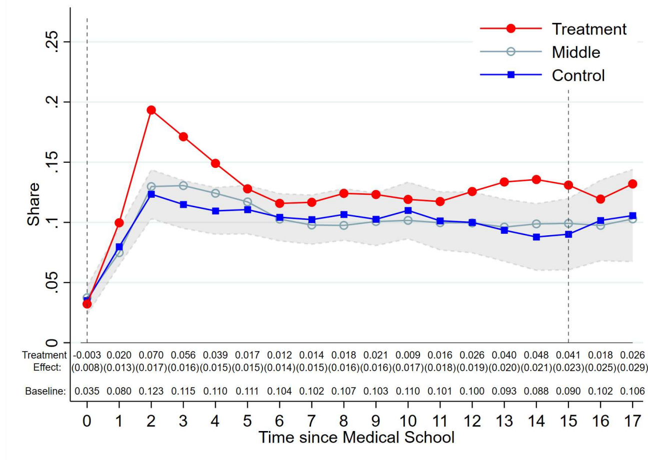
A. Sorting into Rural Labor Market

Partnered at Baseline

Males

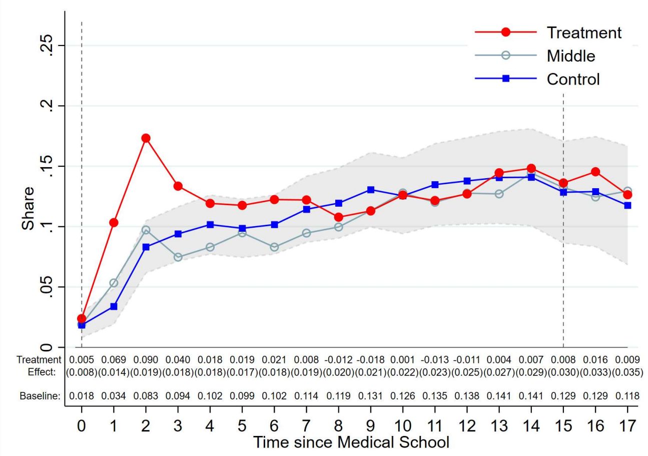


Females

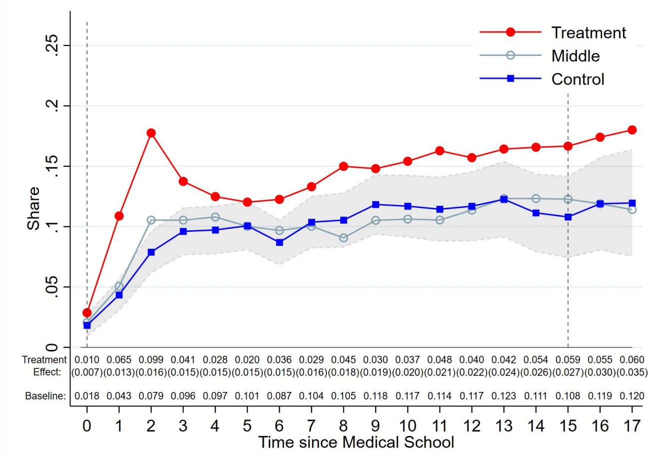


Single at Baseline

Males



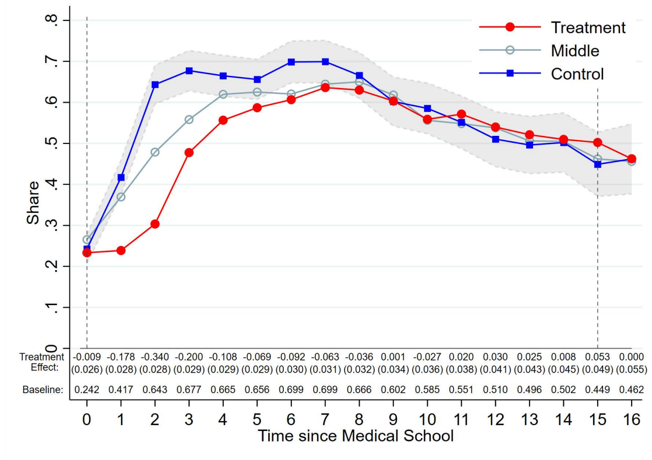
Females



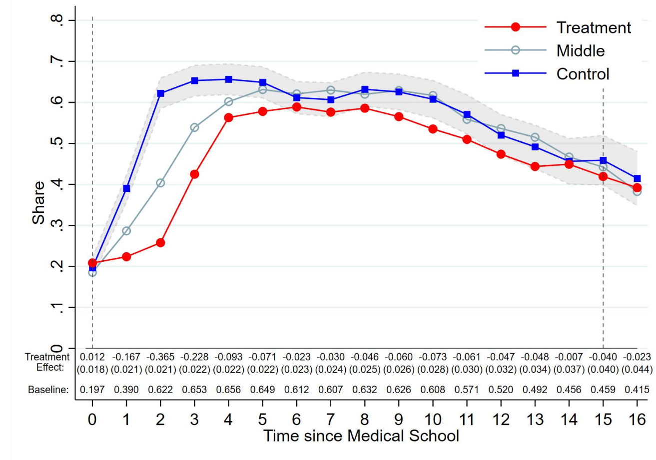
B. Affiliation with a University Hospital

Partnered at Baseline

Males

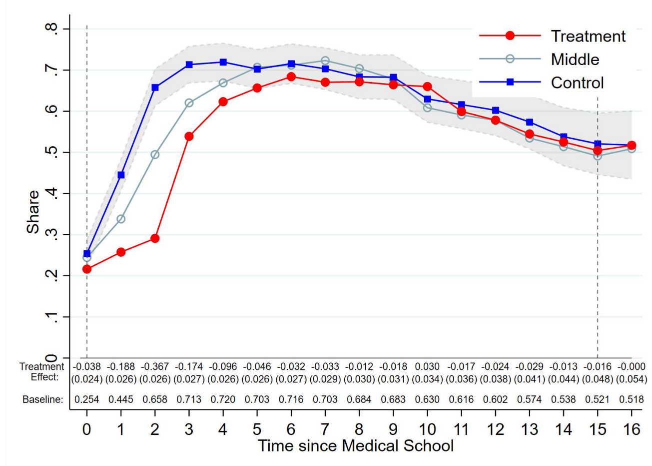


Females

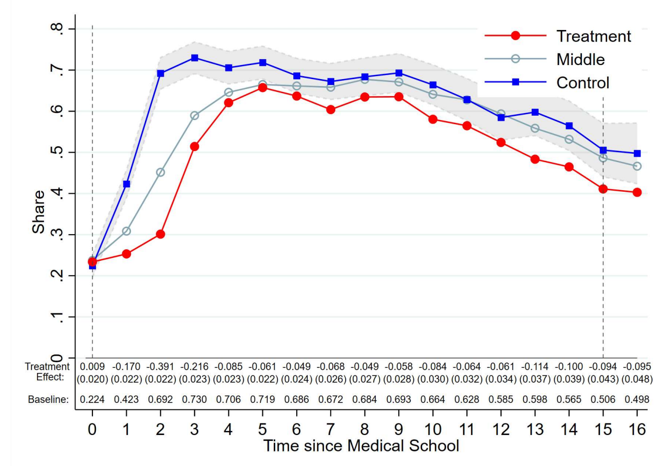


Single at Baseline

Males



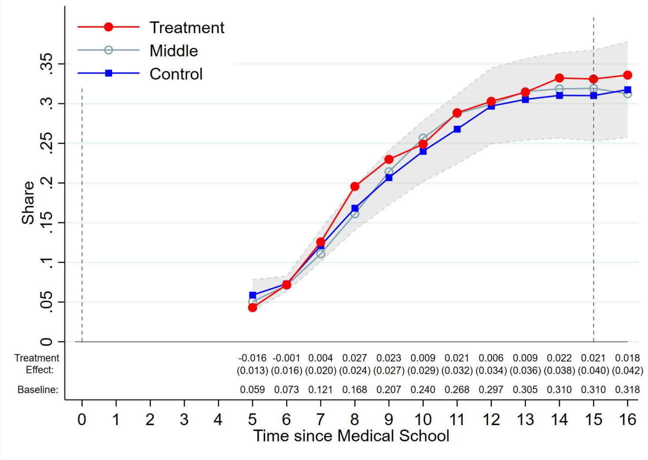
Females



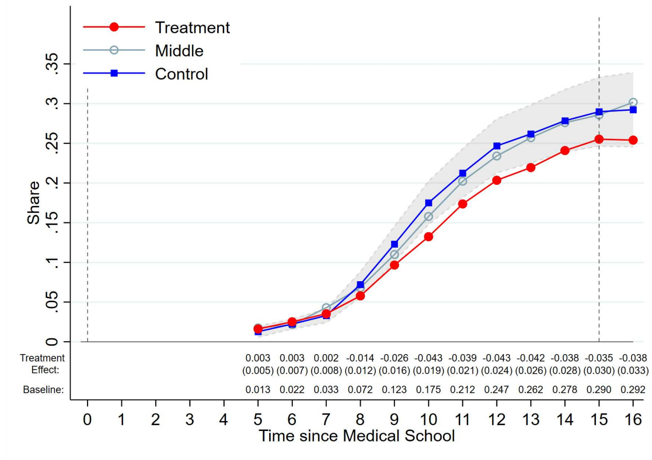
C. Obtaining a Medical PhD

Partnered at Baseline

Males

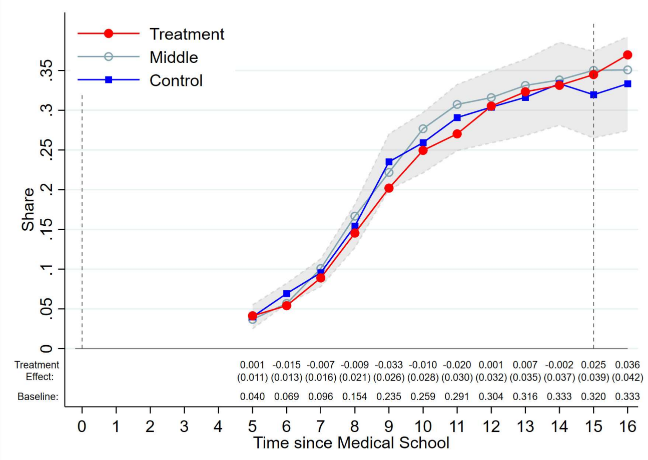


Females

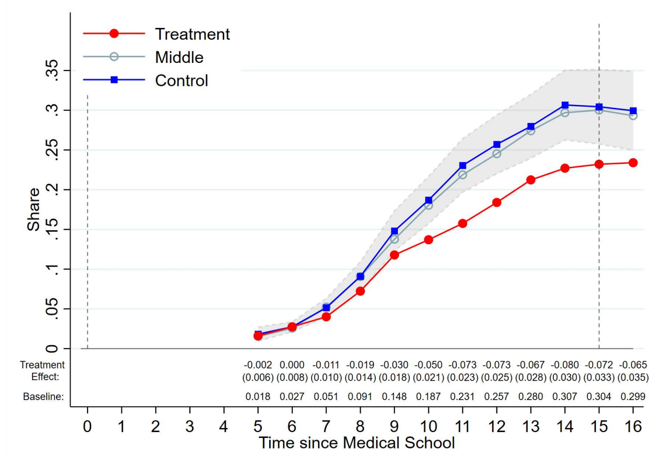


Single at Baseline

Males



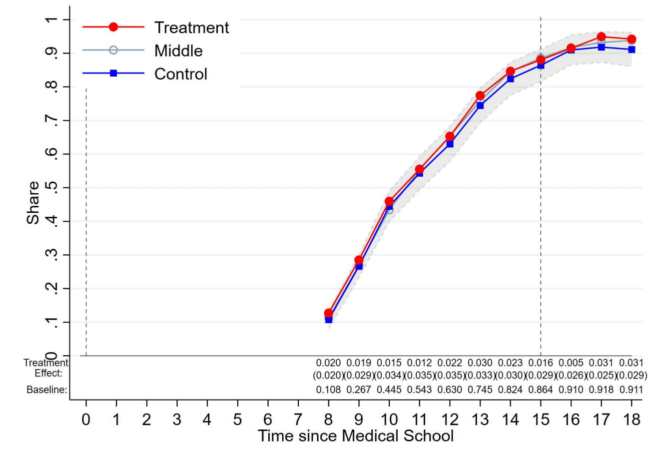
Females



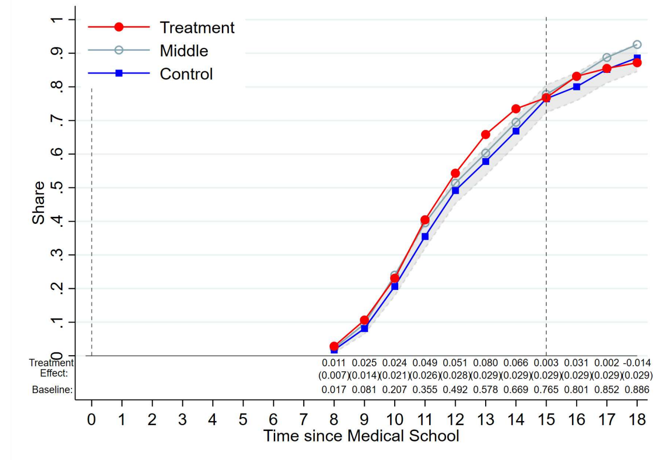
D. Completion of a Medical Specialty

Partnered at Baseline

Males

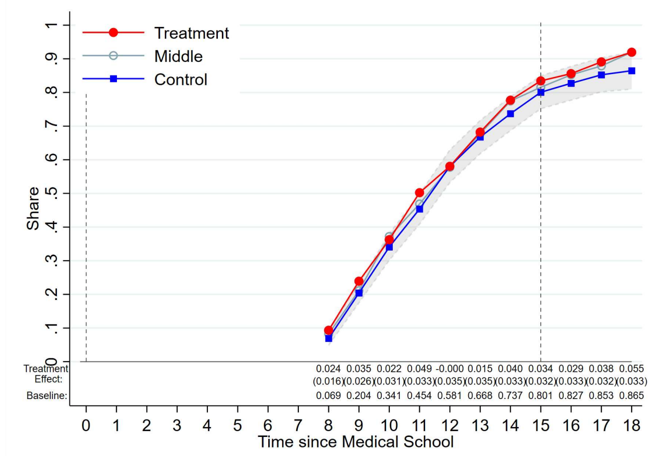


Females

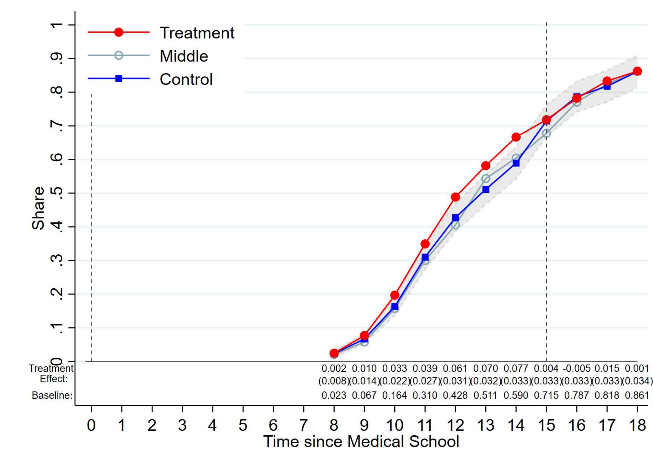


Single at Baseline

Males



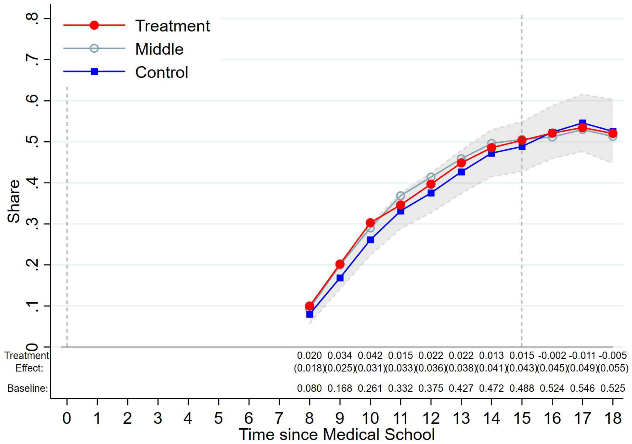
Females



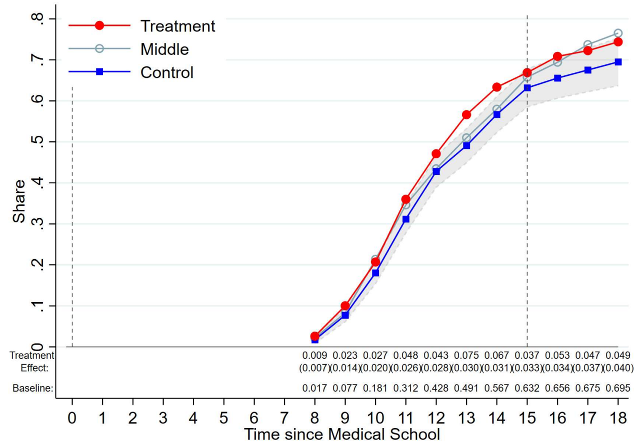
E. Completion of a Female-Represented Specialty

Partnered at Baseline

Males

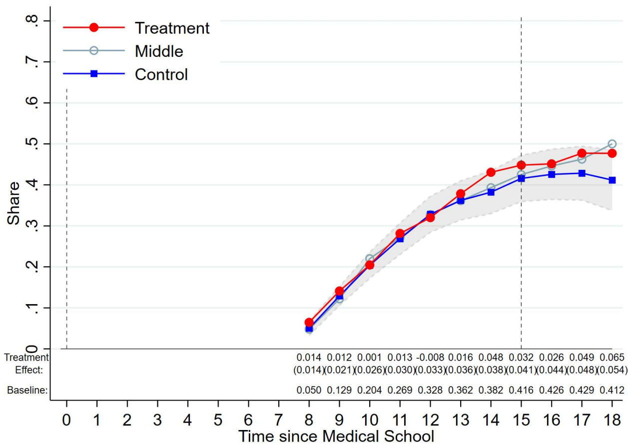


Females

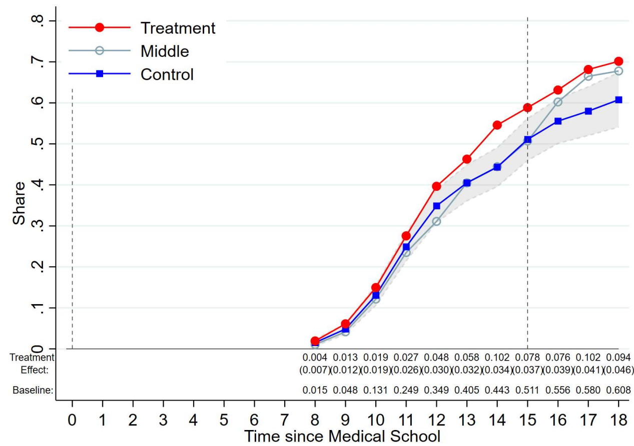


Single at Baseline

Males



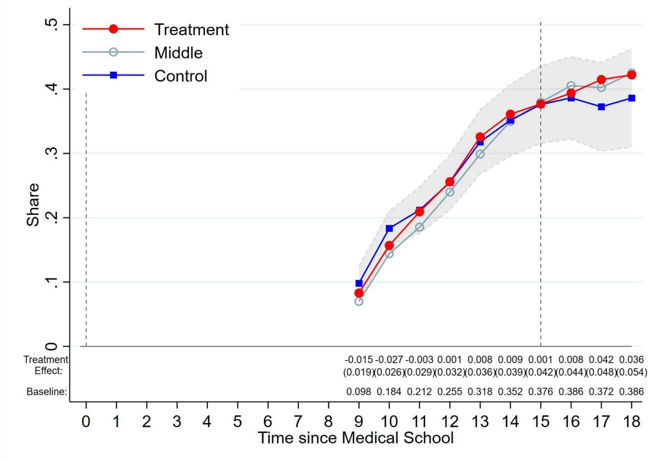
Females



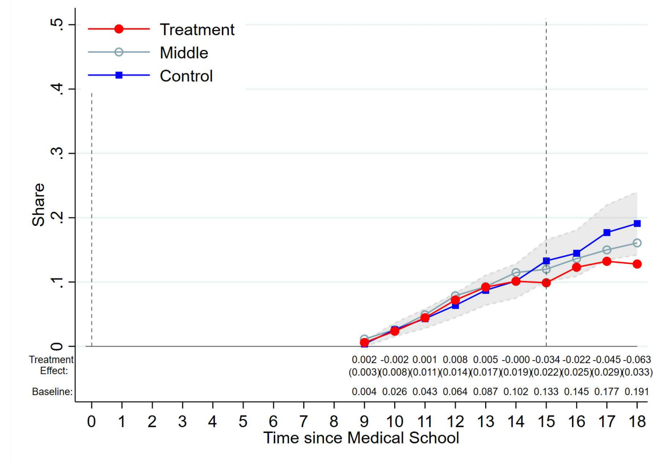
F. Completion of a Male-Represented Specialty

Partnered at Baseline

Males

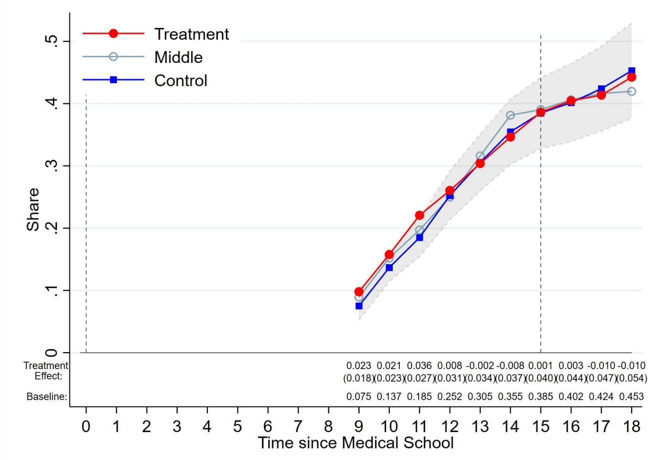


Females

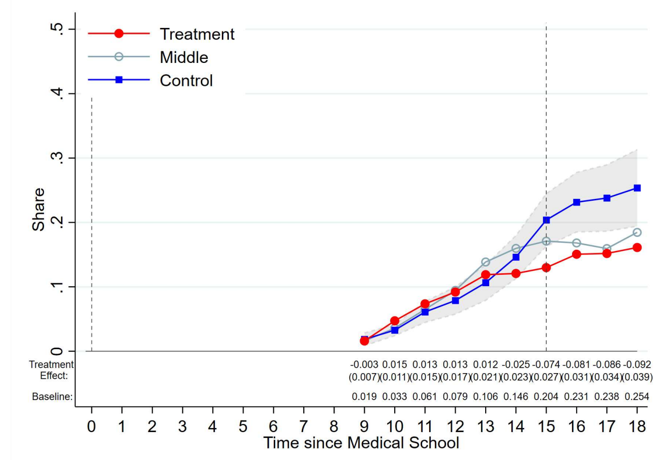


Single at Baseline

Males



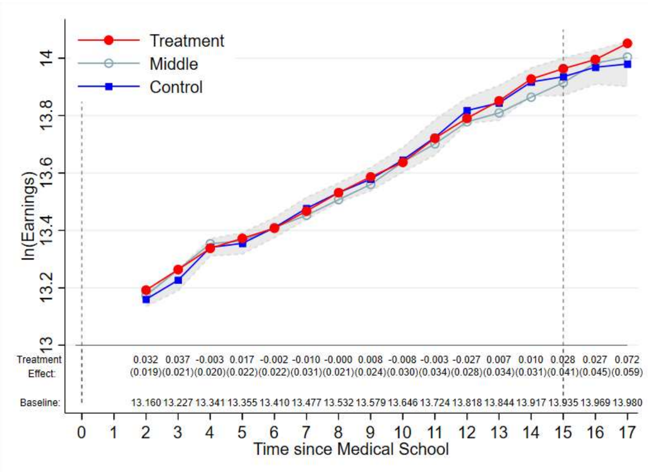
Females



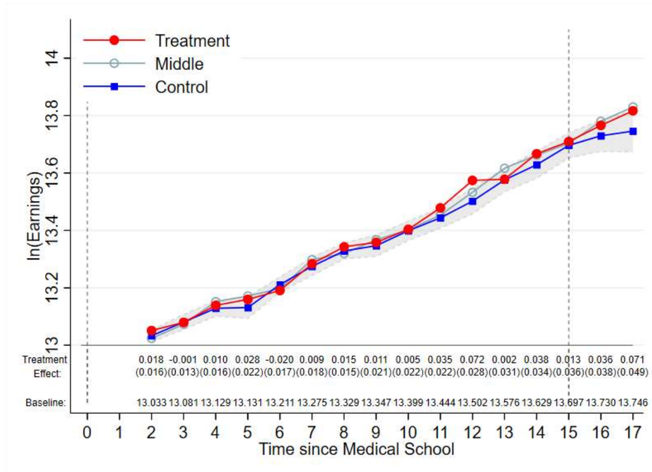
G. Log Earnings

Partnered at Baseline

Males

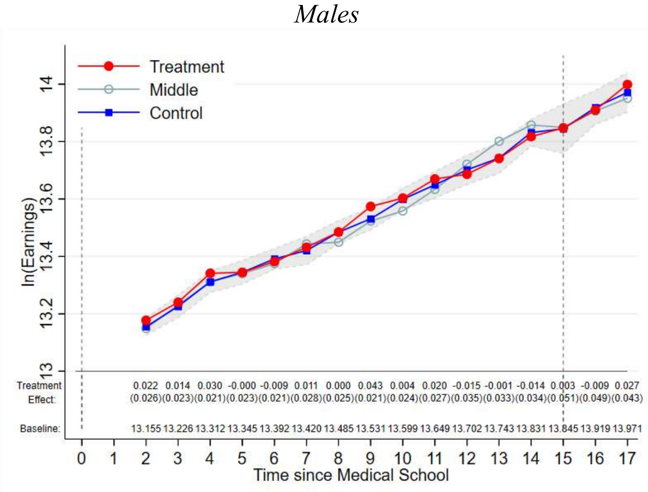


Females

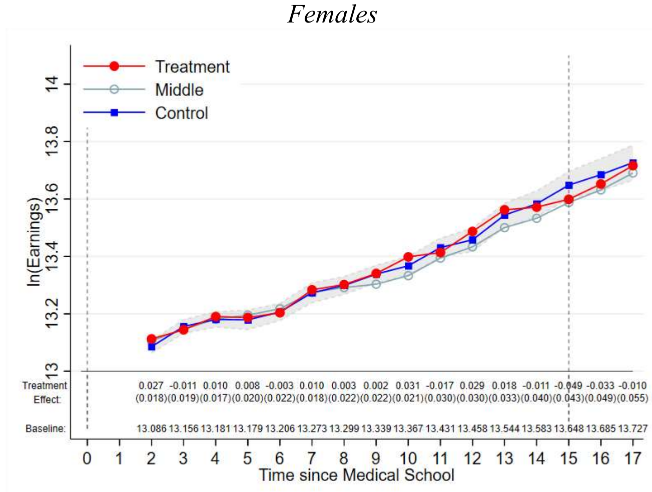


Single at Baseline

Males



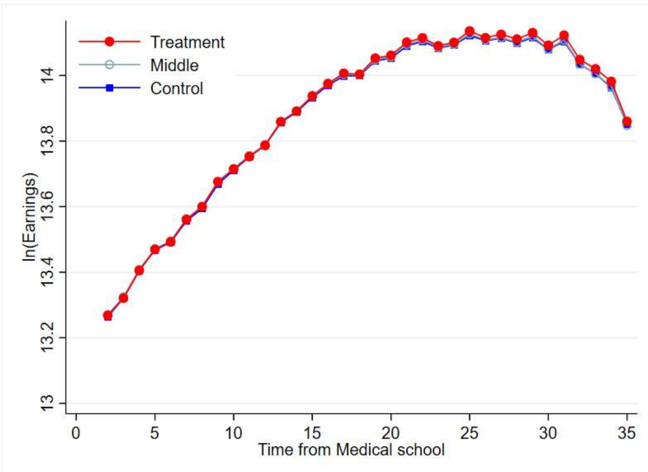
Females



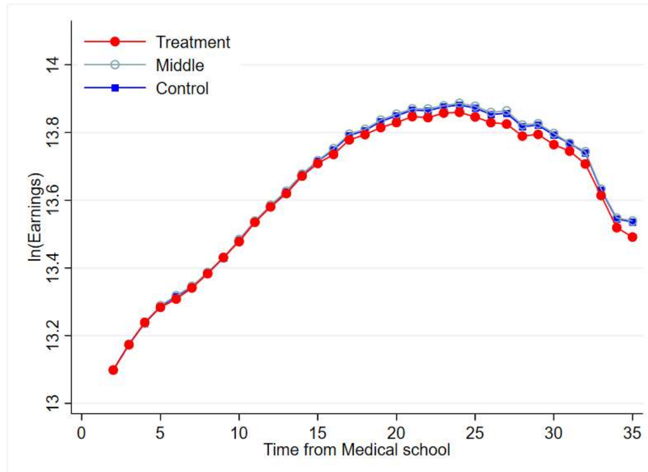
H. Predicted Earnings

Partnered at Baseline

Males

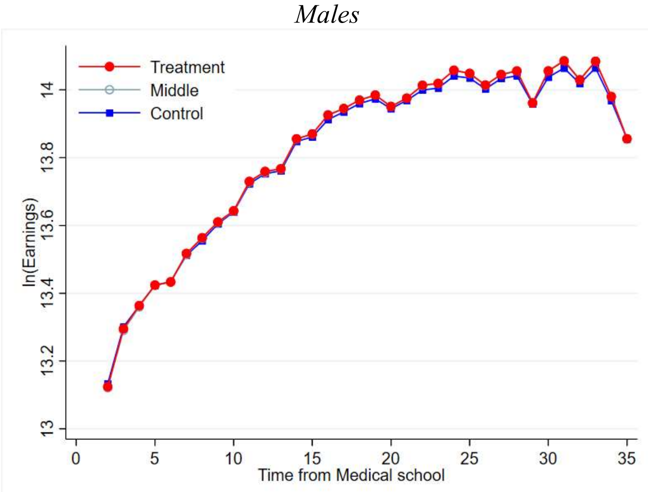


Females

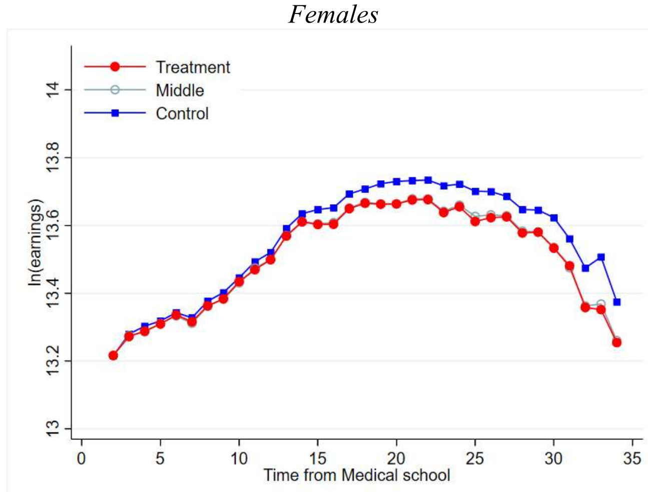


Single at Baseline

Males



Females

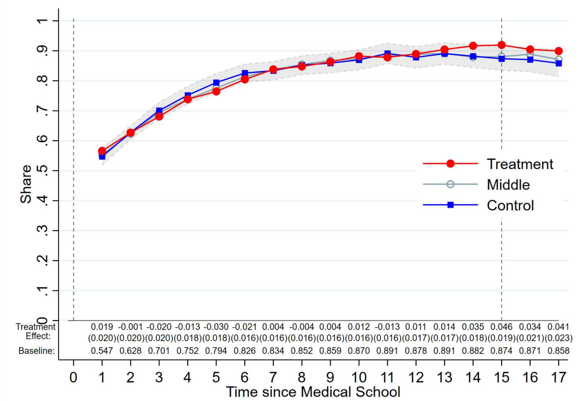


Notes: In this figure, we rerun our entire analysis of labor market outcomes from Section 5, where we split individuals by their baseline partnership status.

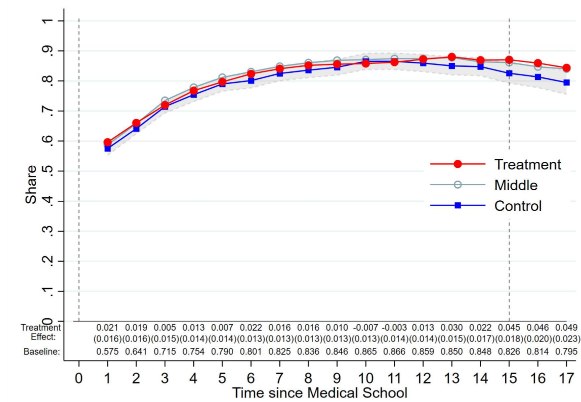
Appendix Figure G.2: Family Choices—All Physicians

A. Partnership

Males

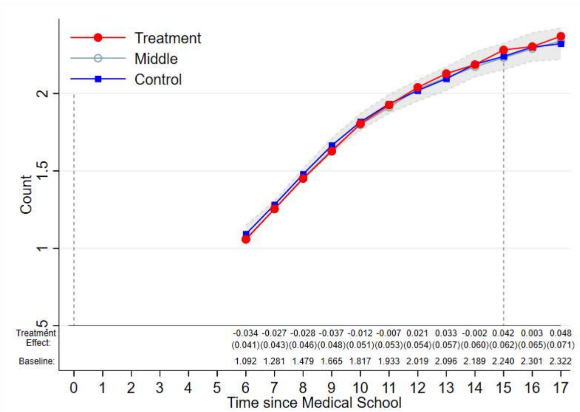


Females

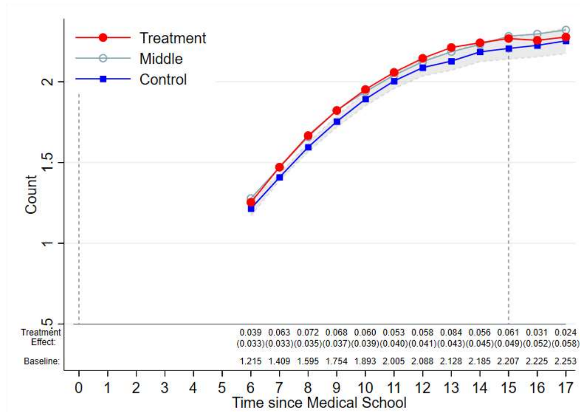


B. Fertility

Males



Females

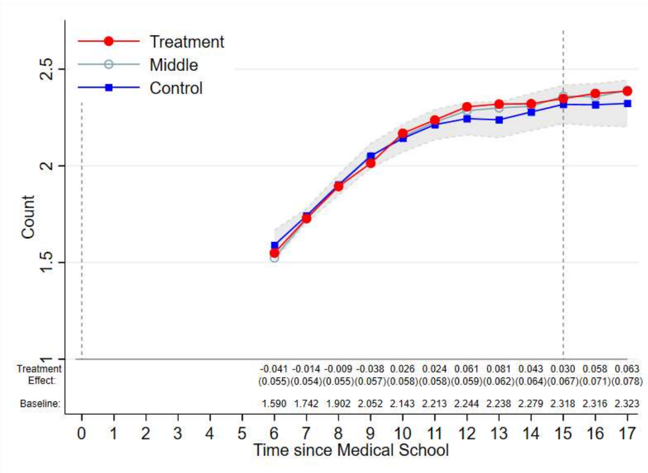


Notes: This figure plots the impact of the lottery on partnership (in panel A) and fertility (in panel B) among all physicians (who were either in a partnership or single prior to the lottery). The x-axis denotes the year relative to the last full calendar year in medical school as the baseline period. We provide plots for the dynamics of an outcome following graduation from medical school for the treatment group, the middle group, and the control group (along with the control group's 95-percent confidence intervals). We report at the bottom of each plot estimates for β_τ from equation (1) along with their standard errors in parentheses and counterfactual levels from the control group. Estimations run to year 17 since 2021 is our last available calendar year in the demographic registers. In panel A, the outcome is an indicator for having a registered partner (married or cohabiting). In panel B, the outcome is the number of children of whom the physician is registered as a parent. Estimations for this outcome run from year 6 after the spikes in fertility among physicians single at baseline (shown in Appendix Figure A.6).

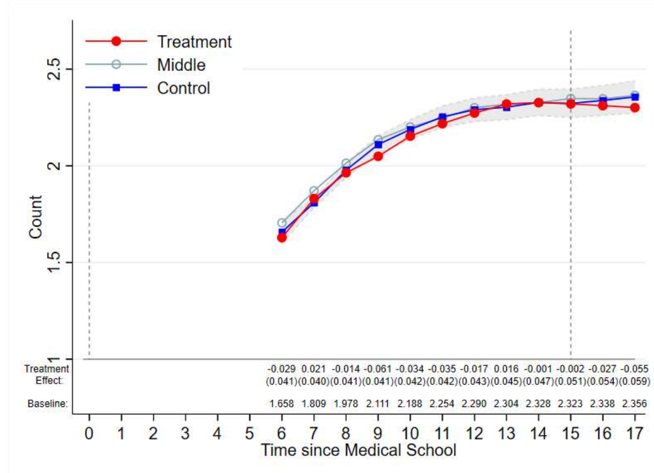
Appendix Figure G.3: Fertility—Robustness

A. Partnered at Baseline

Males

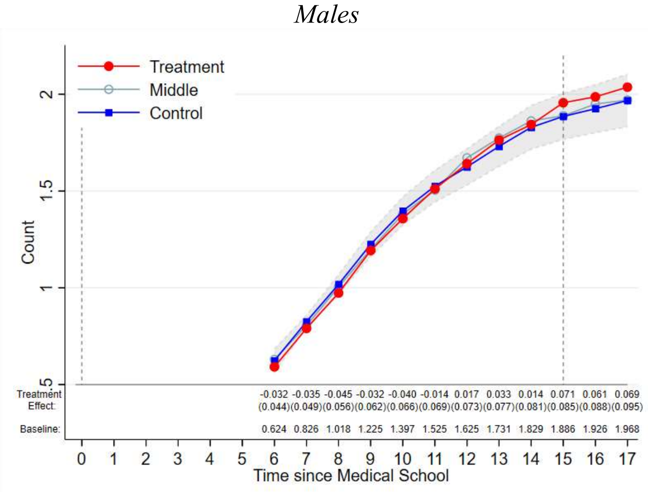


Females

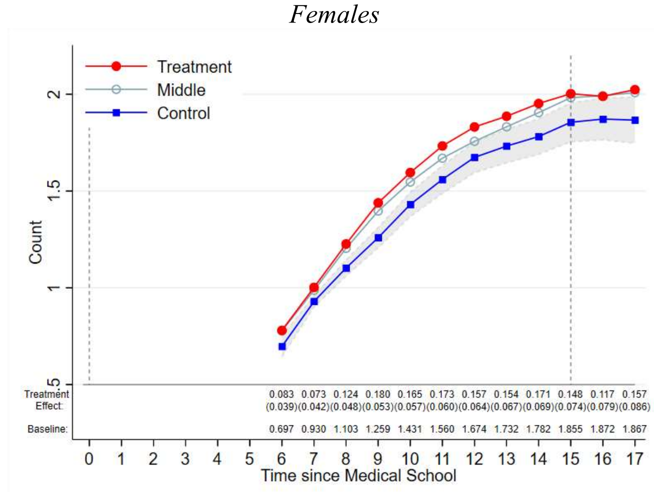


B. Single at Baseline

Males



Females



Notes: This figure replicates Figure 10, where we winsorize the number of children above their 99th percentile of 3 children (that is, at having 4 children or more).

Appendix Table G.1: Residing in the Internship's County

	Males	Females
	(1)	(2)
Treatment Effect	-0.0755	-0.0832
	(0.0196)	(0.0157)
Effect on Middle Group	-0.0216	-0.0129
	(0.0179)	(0.0145)
Constant (Control Group)	0.5610	0.5904
	(0.0134)	(0.0109)
Observations	27,303	41,535
Individuals	3,743	5,770

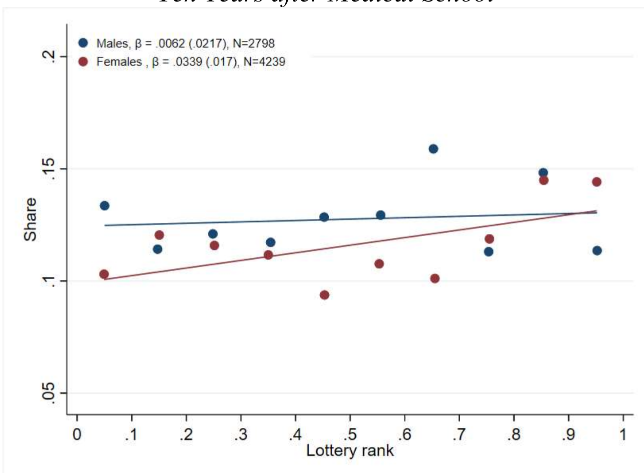
Notes: This table studies the propensity to reside in the county of the initial internship placement. We provide estimates for β from equation (2) using years 6-15. Robust standard errors clustered at the individual level are reported in parentheses.

Appendix H: Alternative Specifications and Robustness Checks

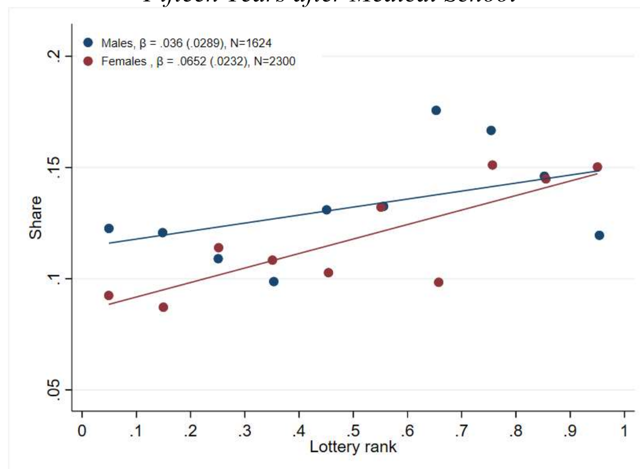
Appendix Figure H.1: Long-Run Outcomes by Lottery Rank

A. Sorting into a Rural Labor Market

Ten Years after Medical School

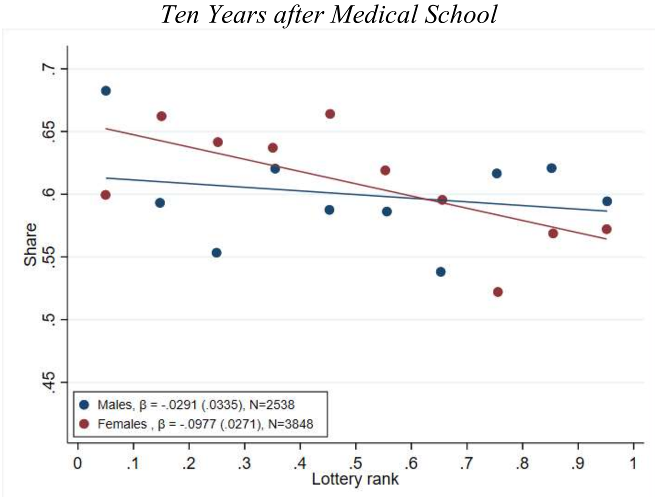


Fifteen Years after Medical School

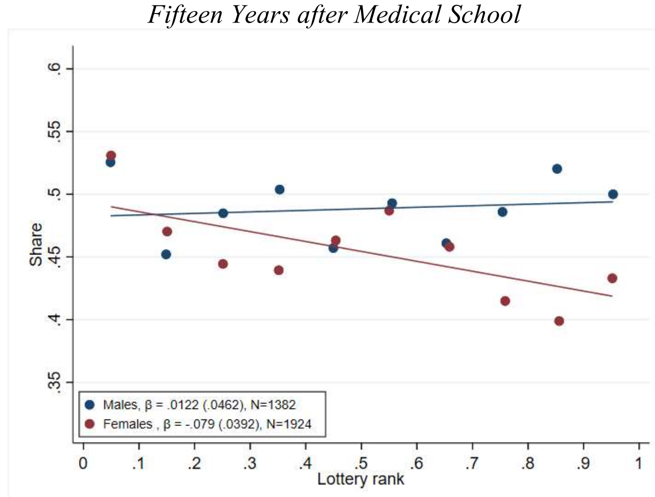


B. Affiliation with a University Hospital

Ten Years after Medical School

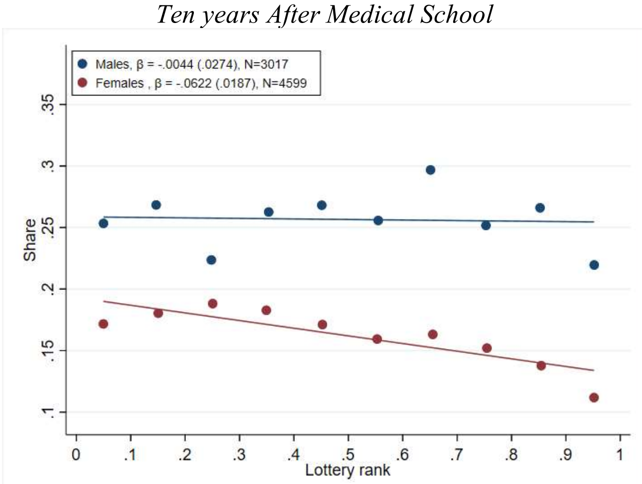


Fifteen Years after Medical School

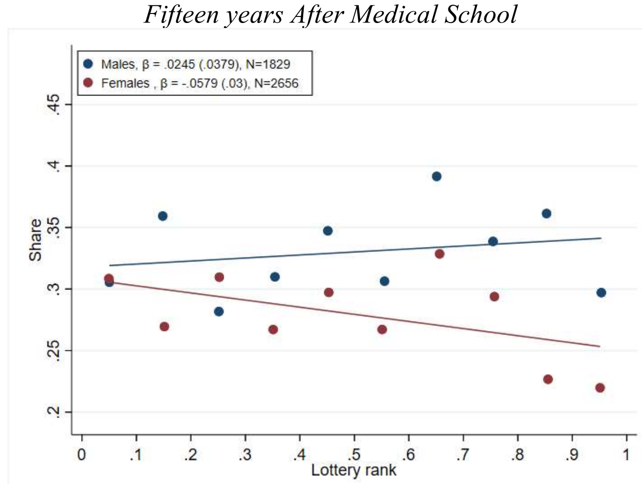


C. Obtaining a Medical PhD

Ten years After Medical School

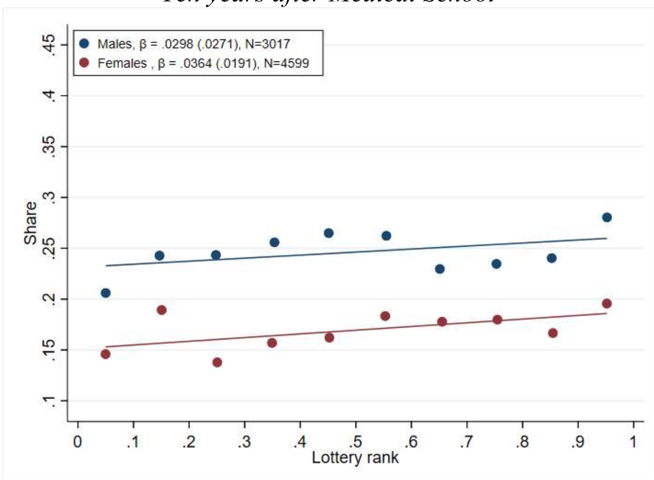


Fifteen years After Medical School

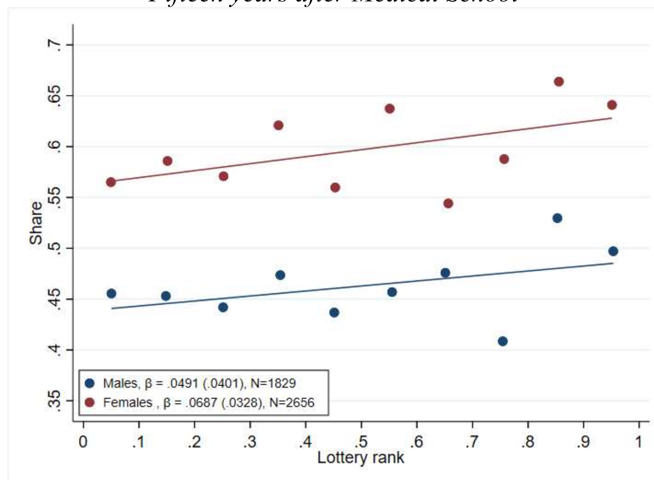


D. Completion of a Female-Represented Specialty

Ten years after Medical School

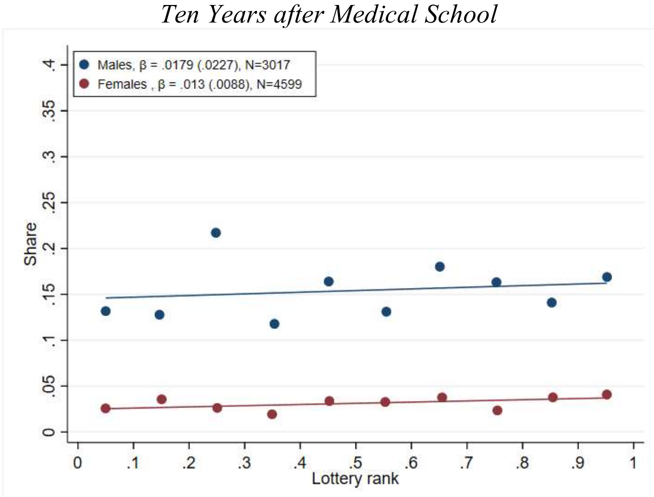


Fifteen years after Medical School

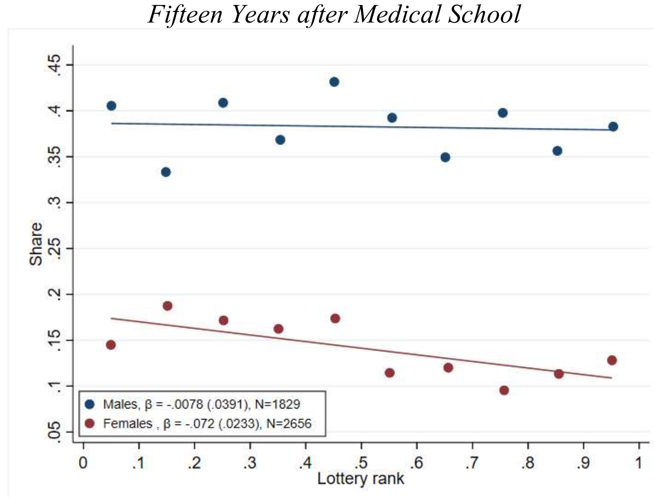


E. Completion of a Male-Represented Specialty

Ten Years after Medical School



Fifteen Years after Medical School



Notes: This figure provides non-parametric plots of outcomes in year 10 and year 15 after graduation from medical school against lottery rank deciles. We also report in the plots the coefficients from the corresponding linear in rank regressions estimated on the underlying individual-level data.

Appendix Table H.1: Robustness Checks

A. Sorting into a Rural Labor Market

Outcome: Rural Local Labor Markets. Males (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=0	0.0030	0.0034	0.0038	0.0055	0.0048	0.0035
	(0.0070)	(0.0067)	(0.0060)	(0.0056)	(0.0051)	(0.0057)
Treatment # t=1	0.0743	0.0688	0.0683	0.0602	0.0544	0.0636
	(0.0131)	(0.0117)	(0.0106)	(0.0097)	(0.0090)	(0.0101)
Treatment # t=2	0.1210	0.1126	0.1089	0.1022	0.0967	0.1048
	(0.0175)	(0.0155)	(0.0141)	(0.0130)	(0.0121)	(0.0135)
Treatment # t=3	0.0567	0.0568	0.0577	0.0505	0.0453	0.0502
	(0.0167)	(0.0151)	(0.0138)	(0.0127)	(0.0118)	(0.0132)
Treatment # t=4	0.0203	0.0238	0.0283	0.0214	0.0192	0.0216
	(0.0165)	(0.0147)	(0.0134)	(0.0124)	(0.0115)	(0.0128)
Treatment # t=5	0.0313	0.0317	0.0315	0.0234	0.0204	0.0253
	(0.0158)	(0.0144)	(0.0131)	(0.0122)	(0.0113)	(0.0126)
Treatment # t=6	0.0212	0.0197	0.0223	0.0212	0.0159	0.0215
	(0.0156)	(0.0142)	(0.0130)	(0.0121)	(0.0111)	(0.0125)
Treatment # t=7	0.0108	0.0109	0.0133	0.0149	0.0143	0.0121
	(0.0166)	(0.0151)	(0.0139)	(0.0130)	(0.0120)	(0.0134)
Treatment # t=8	-0.0049	-0.0055	-0.0062	-0.0011	0.0020	-0.0056
	(0.0175)	(0.0158)	(0.0144)	(0.0136)	(0.0127)	(0.0140)
Treatment # t=9	-0.0084	-0.0136	-0.0142	-0.0081	-0.0041	-0.0140
	(0.0183)	(0.0167)	(0.0151)	(0.0143)	(0.0134)	(0.0148)
Treatment # t=10	0.0066	0.0034	0.0026	0.0059	0.0112	0.0005
	(0.0197)	(0.0177)	(0.0160)	(0.0150)	(0.0142)	(0.0156)
Treatment # t=11	-0.0110	-0.0109	-0.0086	-0.0004	0.0022	-0.0055
	(0.0205)	(0.0186)	(0.0170)	(0.0159)	(0.0150)	(0.0165)
Treatment # t=12	-0.0059	-0.0038	-0.0039	-0.0018	0.0032	-0.0057
	(0.0218)	(0.0197)	(0.0180)	(0.0170)	(0.0159)	(0.0176)
Treatment # t=13	0.0046	0.0068	0.0097	0.0130	0.0166	0.0101
	(0.0234)	(0.0211)	(0.0194)	(0.0183)	(0.0170)	(0.0189)
Treatment # t=14	0.0048	0.0148	0.0152	0.0105	0.0182	0.0089
	(0.0246)	(0.0225)	(0.0206)	(0.0197)	(0.0183)	(0.0202)
Treatment # t=15	0.0104	0.0233	0.0252	0.0256	0.0382	0.0263
	(0.0256)	(0.0232)	(0.0213)	(0.0202)	(0.0188)	(0.0208)

Outcome: Rural Local Labor Markets. Males continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=0	0.0034	-0.0028	0.0005	-0.0004	0.0023	0.0008
	(0.0056)	(0.0055)	(0.0054)	(0.0054)	(0.0061)	(0.0054)
Middle group # t=1	0.0249	0.0188	0.0201	0.0234	0.0194	0.0185
	(0.0084)	(0.0083)	(0.0082)	(0.0088)	(0.0099)	(0.0084)
Middle group # t=2	0.0327	0.0267	0.0255	0.0295	0.0244	0.0213
	(0.0119)	(0.0115)	(0.0114)	(0.0119)	(0.0131)	(0.0116)
Middle group # t=3	0.0199	0.0094	0.0072	0.0090	0.0137	0.0024
	(0.0125)	(0.0120)	(0.0118)	(0.0122)	(0.0136)	(0.0120)
Middle group # t=4	0.0012	0.0025	0.0069	0.0050	0.0099	0.0044
	(0.0130)	(0.0122)	(0.0120)	(0.0124)	(0.0138)	(0.0122)
Middle group # t=5	0.0228	0.0138	0.0175	0.0164	0.0180	0.0120
	(0.0123)	(0.0119)	(0.0119)	(0.0124)	(0.0139)	(0.0121)
Middle group # t=6	0.0127	0.0027	0.0043	0.0031	0.0119	-0.0005
	(0.0123)	(0.0119)	(0.0117)	(0.0120)	(0.0135)	(0.0118)
Middle group # t=7	0.0117	0.0016	0.0043	-0.0031	0.0046	-0.0057
	(0.0134)	(0.0129)	(0.0127)	(0.0130)	(0.0144)	(0.0128)
Middle group # t=8	0.0024	-0.0052	-0.0004	-0.0045	-0.0027	-0.0092
	(0.0144)	(0.0138)	(0.0137)	(0.0139)	(0.0151)	(0.0137)
Middle group # t=9	0.0037	-0.0111	-0.0035	-0.0075	-0.0121	-0.0166
	(0.0153)	(0.0147)	(0.0145)	(0.0147)	(0.0159)	(0.0146)
Middle group # t=10	0.0051	-0.0023	0.0099	0.0090	0.0024	-0.0013
	(0.0161)	(0.0154)	(0.0153)	(0.0157)	(0.0170)	(0.0154)
Middle group # t=11	0.0039	-0.0054	0.0007	0.0025	-0.0093	-0.0081
	(0.0172)	(0.0165)	(0.0163)	(0.0167)	(0.0180)	(0.0164)
Middle group # t=12	0.0049	-0.0015	0.0010	-0.0041	-0.0086	-0.0142
	(0.0182)	(0.0174)	(0.0173)	(0.0176)	(0.0191)	(0.0173)
Middle group # t=13	0.0077	0.0037	0.0021	-0.0020	-0.0025	-0.0116
	(0.0193)	(0.0185)	(0.0183)	(0.0186)	(0.0202)	(0.0183)
Middle group # t=14	0.0212	0.0161	0.0162	0.0004	0.0063	-0.0046
	(0.0205)	(0.0197)	(0.0196)	(0.0201)	(0.0220)	(0.0198)
Middle group # t=15	0.0153	0.0165	0.0159	0.0076	0.0141	0.0052
	(0.0211)	(0.0201)	(0.0201)	(0.0205)	(0.0224)	(0.0202)

Outcome: Rural Local Labor Markets. Males continued (3/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=1	0.0198	0.0197	0.0198	0.0208	0.0238	0.0213
	(0.0065)	(0.0059)	(0.0053)	(0.0050)	(0.0047)	(0.0052)
t=2	0.0681	0.0671	0.0678	0.0667	0.0683	0.0691
	(0.0095)	(0.0087)	(0.0080)	(0.0073)	(0.0069)	(0.0077)
t=3	0.0817	0.0828	0.0835	0.0839	0.0840	0.0866
	(0.0101)	(0.0093)	(0.0085)	(0.0080)	(0.0074)	(0.0083)
t=4	0.0953	0.0897	0.0876	0.0904	0.0902	0.0904
	(0.0113)	(0.0100)	(0.0090)	(0.0085)	(0.0079)	(0.0087)
t=5	0.0780	0.0799	0.0802	0.0839	0.0858	0.0843
	(0.0107)	(0.0098)	(0.0089)	(0.0083)	(0.0079)	(0.0087)
t=6	0.0804	0.0828	0.0826	0.0832	0.0833	0.0843
	(0.0109)	(0.0100)	(0.0090)	(0.0084)	(0.0078)	(0.0088)
t=7	0.0883	0.0907	0.0905	0.0925	0.0906	0.0943
	(0.0117)	(0.0108)	(0.0098)	(0.0093)	(0.0085)	(0.0096)
t=8	0.0974	0.0986	0.0986	0.0988	0.0974	0.1016
	(0.0128)	(0.0117)	(0.0106)	(0.0099)	(0.0092)	(0.0103)
t=9	0.1009	0.1072	0.1058	0.1058	0.1053	0.1106
	(0.0136)	(0.0125)	(0.0113)	(0.0106)	(0.0099)	(0.0111)
t=10	0.1047	0.1061	0.1031	0.1034	0.1038	0.1081
	(0.0142)	(0.0129)	(0.0116)	(0.0109)	(0.0102)	(0.0114)
t=11	0.1110	0.1134	0.1122	0.1096	0.1118	0.1145
	(0.0151)	(0.0138)	(0.0125)	(0.0116)	(0.0109)	(0.0121)
t=12	0.1121	0.1123	0.1133	0.1148	0.1139	0.1193
	(0.0160)	(0.0145)	(0.0132)	(0.0124)	(0.0116)	(0.0130)
t=13	0.1127	0.1115	0.1131	0.1134	0.1117	0.1176
	(0.0167)	(0.0151)	(0.0138)	(0.0129)	(0.0120)	(0.0135)
t=14	0.1101	0.1087	0.1113	0.1189	0.1147	0.1210
	(0.0178)	(0.0160)	(0.0146)	(0.0140)	(0.0129)	(0.0145)
t=15	0.1026	0.0965	0.0984	0.1016	0.0951	0.1019
	(0.0183)	(0.0164)	(0.0150)	(0.0141)	(0.0129)	(0.0145)
Constant	0.0186	0.0217	0.0198	0.0194	0.0188	0.0197
	(0.0048)	(0.0046)	(0.0040)	(0.0037)	(0.0034)	(0.0038)
Observations	50774	50774	50774	50774	50774	50774
Individuals	3970	3970	3970	3970	3970	3970

Outcome: Rural Local Labor Markets. Females (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=0	0.0065	0.0034	0.0035	0.0047	0.0058	0.0065
	(0.0065)	(0.0059)	(0.0055)	(0.0052)	(0.0048)	(0.0053)
Treatment # t=1	0.0572	0.0472	0.0417	0.0348	0.0319	0.0398
	(0.0115)	(0.0102)	(0.0091)	(0.0084)	(0.0077)	(0.0086)
Treatment # t=2	0.1105	0.0993	0.0840	0.0671	0.0581	0.0768
	(0.0142)	(0.0127)	(0.0115)	(0.0106)	(0.0098)	(0.0109)
Treatment # t=3	0.0599	0.0595	0.0496	0.0394	0.0376	0.0446
	(0.0139)	(0.0123)	(0.0111)	(0.0103)	(0.0096)	(0.0106)
Treatment # t=4	0.0385	0.0405	0.0341	0.0276	0.0256	0.0316
	(0.0132)	(0.0119)	(0.0107)	(0.0101)	(0.0094)	(0.0103)
Treatment # t=5	0.0233	0.0199	0.0185	0.0120	0.0101	0.0151
	(0.0129)	(0.0117)	(0.0105)	(0.0098)	(0.0091)	(0.0101)
Treatment # t=6	0.0283	0.0267	0.0231	0.0183	0.0141	0.0213
	(0.0127)	(0.0113)	(0.0102)	(0.0094)	(0.0088)	(0.0097)
Treatment # t=7	0.0233	0.0254	0.0212	0.0129	0.0104	0.0178
	(0.0138)	(0.0122)	(0.0110)	(0.0101)	(0.0094)	(0.0104)
Treatment # t=8	0.0319	0.0340	0.0297	0.0211	0.0160	0.0274
	(0.0147)	(0.0132)	(0.0119)	(0.0109)	(0.0100)	(0.0112)
Treatment # t=9	0.0292	0.0296	0.0245	0.0166	0.0149	0.0211
	(0.0154)	(0.0138)	(0.0124)	(0.0115)	(0.0106)	(0.0118)
Treatment # t=10	0.0303	0.0316	0.0218	0.0145	0.0132	0.0200
	(0.0161)	(0.0145)	(0.0131)	(0.0120)	(0.0111)	(0.0124)
Treatment # t=11	0.0329	0.0385	0.0304	0.0246	0.0198	0.0290
	(0.0168)	(0.0152)	(0.0136)	(0.0126)	(0.0116)	(0.0130)
Treatment # t=12	0.0335	0.0413	0.0321	0.0220	0.0174	0.0280
	(0.0175)	(0.0159)	(0.0145)	(0.0134)	(0.0123)	(0.0138)
Treatment # t=13	0.0440	0.0516	0.0406	0.0284	0.0232	0.0346
	(0.0188)	(0.0170)	(0.0156)	(0.0144)	(0.0132)	(0.0148)
Treatment # t=14	0.0516	0.0636	0.0505	0.0413	0.0340	0.0479
	(0.0197)	(0.0180)	(0.0165)	(0.0152)	(0.0140)	(0.0157)
Treatment # t=15	0.0556	0.0638	0.0487	0.0416	0.0332	0.0439
	(0.0212)	(0.0193)	(0.0177)	(0.0162)	(0.0150)	(0.0167)

Outcome: Rural Local Labor Markets. Females continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=0	0.0065	0.0044	0.0023	-0.0001	0.0044	0.0025
	(0.0052)	(0.0051)	(0.0051)	(0.0052)	(0.0059)	(0.0051)
Middle group # t=1	0.0007	-0.0024	0.0009	-0.0072	-0.0024	0.0000
	(0.0081)	(0.0077)	(0.0075)	(0.0077)	(0.0085)	(0.0076)
Middle group # t=2	0.0284	0.0191	0.0161	0.0056	0.0089	0.0126
	(0.0100)	(0.0097)	(0.0097)	(0.0101)	(0.0113)	(0.0099)
Middle group # t=3	0.0076	0.0102	0.0126	-0.0041	0.0001	0.0032
	(0.0104)	(0.0098)	(0.0098)	(0.0101)	(0.0111)	(0.0099)
Middle group # t=4	0.0117	0.0102	0.0127	-0.0051	-0.0079	0.0003
	(0.0102)	(0.0098)	(0.0097)	(0.0100)	(0.0109)	(0.0098)
Middle group # t=5	0.0034	-0.0027	0.0031	-0.0094	-0.0079	-0.0052
	(0.0102)	(0.0098)	(0.0096)	(0.0099)	(0.0109)	(0.0097)
Middle group # t=6	0.0017	0.0027	0.0040	-0.0017	0.0033	-0.0001
	(0.0099)	(0.0093)	(0.0092)	(0.0095)	(0.0106)	(0.0093)
Middle group # t=7	-0.0099	-0.0077	-0.0040	-0.0096	-0.0067	-0.0072
	(0.0107)	(0.0101)	(0.0099)	(0.0102)	(0.0112)	(0.0100)
Middle group # t=8	-0.0074	-0.0120	-0.0117	-0.0175	-0.0135	-0.0131
	(0.0113)	(0.0106)	(0.0103)	(0.0106)	(0.0116)	(0.0104)
Middle group # t=9	-0.0066	-0.0087	-0.0071	-0.0167	-0.0082	-0.0115
	(0.0120)	(0.0113)	(0.0110)	(0.0113)	(0.0125)	(0.0111)
Middle group # t=10	-0.0034	-0.0074	-0.0094	-0.0195	-0.0113	-0.0154
	(0.0125)	(0.0118)	(0.0116)	(0.0118)	(0.0130)	(0.0117)
Middle group # t=11	-0.0024	-0.0054	-0.0050	-0.0141	-0.0133	-0.0120
	(0.0130)	(0.0122)	(0.0120)	(0.0122)	(0.0134)	(0.0120)
Middle group # t=12	0.0092	0.0057	-0.0017	-0.0119	-0.0077	-0.0095
	(0.0136)	(0.0129)	(0.0127)	(0.0131)	(0.0145)	(0.0128)
Middle group # t=13	0.0137	0.0107	0.0019	-0.0097	0.0035	-0.0069
	(0.0144)	(0.0136)	(0.0136)	(0.0140)	(0.0157)	(0.0137)
Middle group # t=14	0.0256	0.0213	0.0114	0.0029	0.0113	0.0063
	(0.0150)	(0.0143)	(0.0143)	(0.0149)	(0.0168)	(0.0145)
Middle group # t=15	0.0256	0.0212	0.0118	0.0076	0.0144	0.0076
	(0.0160)	(0.0152)	(0.0153)	(0.0160)	(0.0181)	(0.0156)

Outcome: Rural Local Labor Markets. Females continued (3/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=1	0.0394	0.0385	0.0353	0.0377	0.0370	0.0359
	(0.0065)	(0.0058)	(0.0052)	(0.0048)	(0.0044)	(0.0049)
t=2	0.0706	0.0731	0.0750	0.0811	0.0827	0.0779
	(0.0080)	(0.0074)	(0.0068)	(0.0065)	(0.0061)	(0.0065)
t=3	0.0854	0.0797	0.0788	0.0858	0.0848	0.0839
	(0.0087)	(0.0078)	(0.0071)	(0.0067)	(0.0063)	(0.0069)
t=4	0.0805	0.0777	0.0766	0.0834	0.0844	0.0824
	(0.0086)	(0.0077)	(0.0071)	(0.0067)	(0.0063)	(0.0069)
t=5	0.0821	0.0831	0.0788	0.0839	0.0844	0.0834
	(0.0088)	(0.0080)	(0.0072)	(0.0068)	(0.0064)	(0.0070)
t=6	0.0739	0.0704	0.0689	0.0712	0.0723	0.0714
	(0.0084)	(0.0075)	(0.0069)	(0.0065)	(0.0061)	(0.0066)
t=7	0.0852	0.0793	0.0760	0.0788	0.0792	0.0783
	(0.0094)	(0.0083)	(0.0075)	(0.0071)	(0.0066)	(0.0072)
t=8	0.0845	0.0818	0.0790	0.0807	0.0807	0.0796
	(0.0097)	(0.0087)	(0.0079)	(0.0074)	(0.0069)	(0.0075)
t=9	0.0887	0.0855	0.0830	0.0862	0.0843	0.0853
	(0.0103)	(0.0092)	(0.0083)	(0.0078)	(0.0073)	(0.0080)
t=10	0.0881	0.0858	0.0862	0.0893	0.0871	0.0885
	(0.0108)	(0.0097)	(0.0088)	(0.0083)	(0.0077)	(0.0085)
t=11	0.0856	0.0817	0.0804	0.0827	0.0834	0.0830
	(0.0113)	(0.0101)	(0.0091)	(0.0085)	(0.0080)	(0.0087)
t=12	0.0807	0.0776	0.0809	0.0853	0.0855	0.0847
	(0.0118)	(0.0105)	(0.0097)	(0.0091)	(0.0085)	(0.0093)
t=13	0.0790	0.0757	0.0800	0.0854	0.0841	0.0847
	(0.0124)	(0.0110)	(0.0102)	(0.0097)	(0.0090)	(0.0099)
t=14	0.0690	0.0660	0.0718	0.0758	0.0769	0.0745
	(0.0129)	(0.0115)	(0.0108)	(0.0101)	(0.0094)	(0.0103)
t=15	0.0675	0.0653	0.0715	0.0735	0.0759	0.0747
	(0.0136)	(0.0121)	(0.0114)	(0.0107)	(0.0100)	(0.0109)
Constant	0.0238	0.0259	0.0270	0.0273	0.0258	0.0260
	(0.0044)	(0.0041)	(0.0038)	(0.0035)	(0.0032)	(0.0036)
Observations	77370	77370	77370	77370	77370	77370
Individuals	6103	6103	6103	6103	6103	6103

B. Affiliation with a University Hospital

Outcome: Affiliation with a University Hospital. Males (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=0	-0.0270	-0.0210	-0.0245	-0.0095	-0.0102	-0.0136
	(0.0216)	(0.0191)	(0.0174)	(0.0163)	(0.0151)	(0.0168)
Treatment # t=1	-0.2064	-0.2003	-0.1835	-0.1725	-0.1607	-0.1703
	(0.0231)	(0.0207)	(0.0190)	(0.0178)	(0.0166)	(0.0183)
Treatment # t=2	-0.3808	-0.3642	-0.3544	-0.3429	-0.3244	-0.3478
	(0.0231)	(0.0207)	(0.0191)	(0.0179)	(0.0168)	(0.0183)
Treatment # t=3	-0.2190	-0.2071	-0.1866	-0.1881	-0.1824	-0.1943
	(0.0238)	(0.0213)	(0.0196)	(0.0183)	(0.0171)	(0.0188)
Treatment # t=4	-0.1096	-0.1062	-0.1023	-0.1003	-0.0899	-0.1091
	(0.0239)	(0.0214)	(0.0195)	(0.0181)	(0.0169)	(0.0187)
Treatment # t=5	-0.0629	-0.0494	-0.0570	-0.0580	-0.0543	-0.0606
	(0.0238)	(0.0213)	(0.0194)	(0.0181)	(0.0168)	(0.0186)
Treatment # t=6	-0.0430	-0.0475	-0.0607	-0.0503	-0.0442	-0.0566
	(0.0244)	(0.0219)	(0.0201)	(0.0187)	(0.0175)	(0.0192)
Treatment # t=7	-0.0262	-0.0184	-0.0473	-0.0424	-0.0390	-0.0514
	(0.0258)	(0.0230)	(0.0211)	(0.0197)	(0.0183)	(0.0202)
Treatment # t=8	-0.0289	-0.0138	-0.0240	-0.0233	-0.0236	-0.0246
	(0.0270)	(0.0241)	(0.0221)	(0.0206)	(0.0192)	(0.0212)
Treatment # t=9	-0.0251	-0.0163	-0.0105	-0.0199	-0.0249	-0.0149
	(0.0283)	(0.0253)	(0.0233)	(0.0217)	(0.0203)	(0.0224)
Treatment # t=10	-0.0303	-0.0073	0.0015	-0.0067	-0.0164	-0.0074
	(0.0301)	(0.0271)	(0.0248)	(0.0232)	(0.0218)	(0.0239)
Treatment # t=11	-0.0267	-0.0131	-0.0004	-0.0131	-0.0157	-0.0092
	(0.0321)	(0.0288)	(0.0263)	(0.0247)	(0.0231)	(0.0253)
Treatment # t=12	-0.0105	-0.0164	0.0005	-0.0050	-0.0053	-0.0063
	(0.0342)	(0.0306)	(0.0280)	(0.0263)	(0.0245)	(0.0269)
Treatment # t=13	-0.0153	-0.0225	-0.0042	-0.0007	0.0013	-0.0043
	(0.0363)	(0.0326)	(0.0298)	(0.0280)	(0.0262)	(0.0287)
Treatment # t=14	0.0080	-0.0069	-0.0033	0.0046	-0.0008	-0.0010
	(0.0386)	(0.0346)	(0.0317)	(0.0297)	(0.0278)	(0.0304)
Treatment # t=15	0.0211	0.0101	0.0167	0.0118	0.0056	0.0091
	(0.0419)	(0.0375)	(0.0343)	(0.0322)	(0.0301)	(0.0330)

Outcome: Affiliation with a University Hospital. Males continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=0	-0.0183	-0.0053	0.0054	0.0096	0.0314	0.0014
	(0.0178)	(0.0168)	(0.0166)	(0.0170)	(0.0190)	(0.0167)
Middle group # t=1	-0.0919	-0.0925	-0.0796	-0.0851	-0.0487	-0.0810
	(0.0200)	(0.0190)	(0.0187)	(0.0191)	(0.0210)	(0.0188)
Middle group # t=2	-0.2002	-0.1721	-0.1639	-0.1764	-0.1633	-0.1736
	(0.0194)	(0.0187)	(0.0186)	(0.0192)	(0.0213)	(0.0189)
Middle group # t=3	-0.1338	-0.1192	-0.1052	-0.1070	-0.1100	-0.1075
	(0.0187)	(0.0181)	(0.0181)	(0.0188)	(0.0209)	(0.0184)
Middle group # t=4	-0.0426	-0.0346	-0.0481	-0.0503	-0.0573	-0.0472
	(0.0190)	(0.0181)	(0.0179)	(0.0185)	(0.0205)	(0.0181)
Middle group # t=5	-0.0192	0.0000	-0.0120	-0.0219	-0.0393	-0.0175
	(0.0190)	(0.0182)	(0.0179)	(0.0184)	(0.0204)	(0.0181)
Middle group # t=6	-0.0403	-0.0365	-0.0387	-0.0502	-0.0665	-0.0490
	(0.0197)	(0.0188)	(0.0186)	(0.0193)	(0.0215)	(0.0189)
Middle group # t=7	-0.0010	-0.0007	-0.0159	-0.0181	-0.0292	-0.0292
	(0.0209)	(0.0199)	(0.0195)	(0.0200)	(0.0221)	(0.0197)
Middle group # t=8	-0.0042	0.0044	0.0028	0.0003	-0.0041	-0.0074
	(0.0219)	(0.0209)	(0.0206)	(0.0211)	(0.0232)	(0.0208)
Middle group # t=9	-0.0192	-0.0078	0.0039	-0.0018	-0.0227	-0.0017
	(0.0231)	(0.0221)	(0.0219)	(0.0224)	(0.0247)	(0.0221)
Middle group # t=10	-0.0549	-0.0210	-0.0256	-0.0250	-0.0192	-0.0285
	(0.0247)	(0.0238)	(0.0236)	(0.0243)	(0.0267)	(0.0239)
Middle group # t=11	-0.0276	0.0089	-0.0156	-0.0176	-0.0057	-0.0233
	(0.0264)	(0.0253)	(0.0251)	(0.0257)	(0.0283)	(0.0253)
Middle group # t=12	0.0008	0.0154	-0.0004	0.0099	0.0109	0.0002
	(0.0281)	(0.0269)	(0.0266)	(0.0272)	(0.0299)	(0.0268)
Middle group # t=13	-0.0100	-0.0070	-0.0161	0.0021	0.0037	-0.0141
	(0.0300)	(0.0287)	(0.0284)	(0.0291)	(0.0319)	(0.0287)
Middle group # t=14	-0.0128	-0.0051	-0.0112	0.0037	-0.0083	-0.0098
	(0.0319)	(0.0306)	(0.0303)	(0.0310)	(0.0340)	(0.0306)
Middle group # t=15	-0.0061	-0.0085	-0.0092	0.0071	-0.0189	0.0059
	(0.0345)	(0.0331)	(0.0328)	(0.0336)	(0.0369)	(0.0331)

Outcome: Affiliation with a University Hospital. Males continued (3/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=1	0.1819	0.1903	0.1835	0.1872	0.1779	0.1815
	(0.0178)	(0.0159)	(0.0145)	(0.0136)	(0.0126)	(0.0140)
t=2	0.4158	0.4053	0.4025	0.4082	0.4004	0.4055
	(0.0200)	(0.0180)	(0.0166)	(0.0154)	(0.0144)	(0.0160)
t=3	0.4629	0.4586	0.4479	0.4527	0.4524	0.4518
	(0.0203)	(0.0183)	(0.0169)	(0.0157)	(0.0147)	(0.0163)
t=4	0.4319	0.4369	0.4455	0.4505	0.4505	0.4487
	(0.0215)	(0.0191)	(0.0172)	(0.0160)	(0.0149)	(0.0166)
t=5	0.4233	0.4201	0.4322	0.4419	0.4474	0.4374
	(0.0213)	(0.0192)	(0.0174)	(0.0160)	(0.0149)	(0.0166)
t=6	0.4469	0.4527	0.4590	0.4630	0.4640	0.4619
	(0.0220)	(0.0196)	(0.0179)	(0.0164)	(0.0153)	(0.0170)
t=7	0.4268	0.4344	0.4527	0.4572	0.4611	0.4602
	(0.0230)	(0.0204)	(0.0185)	(0.0171)	(0.0159)	(0.0176)
t=8	0.4177	0.4191	0.4266	0.4333	0.4382	0.4323
	(0.0242)	(0.0216)	(0.0197)	(0.0183)	(0.0170)	(0.0189)
t=9	0.4003	0.4002	0.3966	0.4073	0.4170	0.4016
	(0.0252)	(0.0225)	(0.0205)	(0.0190)	(0.0176)	(0.0196)
t=10	0.3784	0.3603	0.3604	0.3657	0.3689	0.3643
	(0.0256)	(0.0231)	(0.0213)	(0.0199)	(0.0186)	(0.0205)
t=11	0.3418	0.3276	0.3373	0.3460	0.3462	0.3434
	(0.0270)	(0.0242)	(0.0221)	(0.0206)	(0.0193)	(0.0212)
t=12	0.3010	0.3045	0.3104	0.3140	0.3179	0.3137
	(0.0286)	(0.0255)	(0.0232)	(0.0216)	(0.0203)	(0.0223)
t=13	0.2790	0.2877	0.2886	0.2855	0.2873	0.2885
	(0.0300)	(0.0268)	(0.0244)	(0.0229)	(0.0214)	(0.0235)
t=14	0.2613	0.2682	0.2720	0.2687	0.2762	0.2715
	(0.0314)	(0.0282)	(0.0258)	(0.0241)	(0.0225)	(0.0248)
t=15	0.2277	0.2384	0.2379	0.2382	0.2487	0.2358
	(0.0335)	(0.0301)	(0.0276)	(0.0258)	(0.0240)	(0.0265)
Constant	0.2599	0.2515	0.2488	0.2439	0.2412	0.2475
	(0.0154)	(0.0136)	(0.0124)	(0.0115)	(0.0107)	(0.0119)
Observations	48186	48186	48186	48186	48186	48186
Individuals	3970	3970	3970	3970	3970	3970

Outcome: Affiliation with a University Hospital. Females (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=0	-0.0112	-0.0031	0.0102	0.0061	0.0030	0.0078
	(0.0167)	(0.0149)	(0.0136)	(0.0125)	(0.0117)	(0.0128)
Treatment # t=1	-0.1850	-0.1839	-0.1689	-0.1580	-0.1507	-0.1640
	(0.0185)	(0.0166)	(0.0152)	(0.0139)	(0.0131)	(0.0144)
Treatment # t=2	-0.4033	-0.3930	-0.3779	-0.3554	-0.3294	-0.3661
	(0.0184)	(0.0166)	(0.0152)	(0.0142)	(0.0134)	(0.0146)
Treatment # t=3	-0.2421	-0.2354	-0.2234	-0.2095	-0.1977	-0.2206
	(0.0192)	(0.0173)	(0.0159)	(0.0147)	(0.0138)	(0.0151)
Treatment # t=4	-0.0922	-0.1036	-0.0902	-0.0829	-0.0756	-0.0895
	(0.0192)	(0.0173)	(0.0158)	(0.0147)	(0.0137)	(0.0151)
Treatment # t=5	-0.0664	-0.0720	-0.0672	-0.0559	-0.0525	-0.0589
	(0.0192)	(0.0172)	(0.0157)	(0.0146)	(0.0137)	(0.0150)
Treatment # t=6	-0.0363	-0.0466	-0.0363	-0.0334	-0.0307	-0.0336
	(0.0205)	(0.0184)	(0.0167)	(0.0155)	(0.0145)	(0.0159)
Treatment # t=7	-0.0358	-0.0467	-0.0487	-0.0444	-0.0372	-0.0419
	(0.0217)	(0.0194)	(0.0177)	(0.0164)	(0.0153)	(0.0168)
Treatment # t=8	-0.0309	-0.0465	-0.0482	-0.0498	-0.0433	-0.0457
	(0.0224)	(0.0201)	(0.0183)	(0.0169)	(0.0158)	(0.0174)
Treatment # t=9	-0.0432	-0.0621	-0.0600	-0.0621	-0.0587	-0.0550
	(0.0233)	(0.0210)	(0.0191)	(0.0177)	(0.0165)	(0.0182)
Treatment # t=10	-0.0584	-0.0774	-0.0785	-0.0745	-0.0673	-0.0704
	(0.0248)	(0.0223)	(0.0204)	(0.0188)	(0.0176)	(0.0194)
Treatment # t=11	-0.0618	-0.0655	-0.0629	-0.0628	-0.0525	-0.0623
	(0.0265)	(0.0239)	(0.0218)	(0.0201)	(0.0188)	(0.0207)
Treatment # t=12	-0.0628	-0.0672	-0.0535	-0.0573	-0.0521	-0.0494
	(0.0284)	(0.0256)	(0.0234)	(0.0215)	(0.0202)	(0.0222)
Treatment # t=13	-0.0993	-0.0998	-0.0795	-0.0763	-0.0702	-0.0747
	(0.0305)	(0.0274)	(0.0250)	(0.0230)	(0.0216)	(0.0237)
Treatment # t=14	-0.0725	-0.0715	-0.0510	-0.0362	-0.0378	-0.0423
	(0.0328)	(0.0295)	(0.0269)	(0.0248)	(0.0232)	(0.0255)
Treatment # t=15	-0.0915	-0.0830	-0.0653	-0.0484	-0.0441	-0.0565
	(0.0356)	(0.0319)	(0.0292)	(0.0269)	(0.0253)	(0.0277)

Outcome: Affiliation with a University Hospital. Females continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=0	-0.0157	-0.0117	0.0004	0.0081	0.0135	0.0105
	(0.0137)	(0.0130)	(0.0127)	(0.0131)	(0.0146)	(0.0129)
Middle group # t=1	-0.1152	-0.1222	-0.1091	-0.0807	-0.0877	-0.0932
	(0.0161)	(0.0152)	(0.0148)	(0.0152)	(0.0166)	(0.0150)
Middle group # t=2	-0.2603	-0.2417	-0.2293	-0.2055	-0.1859	-0.2165
	(0.0156)	(0.0151)	(0.0150)	(0.0157)	(0.0174)	(0.0153)
Middle group # t=3	-0.1536	-0.1390	-0.1268	-0.1107	-0.1121	-0.1179
	(0.0154)	(0.0149)	(0.0148)	(0.0155)	(0.0173)	(0.0151)
Middle group # t=4	-0.0601	-0.0639	-0.0571	-0.0355	-0.0288	-0.0396
	(0.0156)	(0.0148)	(0.0147)	(0.0153)	(0.0169)	(0.0150)
Middle group # t=5	-0.0235	-0.0279	-0.0349	-0.0151	0.0064	-0.0091
	(0.0156)	(0.0148)	(0.0146)	(0.0151)	(0.0166)	(0.0148)
Middle group # t=6	-0.0038	-0.0075	-0.0069	0.0035	0.0133	0.0082
	(0.0167)	(0.0159)	(0.0156)	(0.0162)	(0.0178)	(0.0159)
Middle group # t=7	0.0008	0.0060	0.0059	0.0169	0.0194	0.0166
	(0.0176)	(0.0168)	(0.0165)	(0.0170)	(0.0187)	(0.0167)
Middle group # t=8	0.0060	0.0049	-0.0091	0.0129	0.0007	0.0071
	(0.0182)	(0.0173)	(0.0170)	(0.0175)	(0.0194)	(0.0172)
Middle group # t=9	-0.0074	-0.0055	-0.0082	0.0083	-0.0084	0.0058
	(0.0190)	(0.0180)	(0.0177)	(0.0182)	(0.0202)	(0.0179)
Middle group # t=10	-0.0193	-0.0168	-0.0057	0.0189	0.0078	0.0269
	(0.0202)	(0.0192)	(0.0188)	(0.0194)	(0.0214)	(0.0190)
Middle group # t=11	-0.0280	-0.0104	-0.0071	0.0043	-0.0179	0.0150
	(0.0216)	(0.0206)	(0.0202)	(0.0208)	(0.0231)	(0.0205)
Middle group # t=12	-0.0176	-0.0032	0.0133	0.0207	0.0035	0.0344
	(0.0232)	(0.0221)	(0.0217)	(0.0224)	(0.0247)	(0.0220)
Middle group # t=13	-0.0297	-0.0130	-0.0057	0.0208	0.0107	0.0073
	(0.0250)	(0.0238)	(0.0234)	(0.0241)	(0.0267)	(0.0237)
Middle group # t=14	-0.0388	-0.0192	-0.0094	0.0184	-0.0016	0.0052
	(0.0269)	(0.0256)	(0.0252)	(0.0261)	(0.0288)	(0.0256)
Middle group # t=15	-0.0497	-0.0286	-0.0172	0.0113	-0.0030	-0.0068
	(0.0295)	(0.0280)	(0.0275)	(0.0285)	(0.0313)	(0.0280)

Outcome: Affiliation with a University Hospital. Females continued (3/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=1	0.1929	0.1993	0.1963	0.1828	0.1808	0.1907
	(0.0150)	(0.0134)	(0.0122)	(0.0112)	(0.0105)	(0.0116)
t=2	0.4622	0.4505	0.4460	0.4284	0.4115	0.4383
	(0.0165)	(0.0151)	(0.0137)	(0.0127)	(0.0119)	(0.0131)
t=3	0.4877	0.4811	0.4802	0.4703	0.4651	0.4783
	(0.0166)	(0.0150)	(0.0137)	(0.0127)	(0.0119)	(0.0130)
t=4	0.4598	0.4684	0.4702	0.4614	0.4572	0.4663
	(0.0172)	(0.0154)	(0.0140)	(0.0129)	(0.0120)	(0.0132)
t=5	0.4507	0.4605	0.4724	0.4637	0.4589	0.4638
	(0.0173)	(0.0155)	(0.0141)	(0.0129)	(0.0120)	(0.0133)
t=6	0.4181	0.4291	0.4371	0.4355	0.4339	0.4348
	(0.0179)	(0.0161)	(0.0147)	(0.0136)	(0.0128)	(0.0140)
t=7	0.4069	0.4141	0.4276	0.4272	0.4271	0.4265
	(0.0189)	(0.0170)	(0.0154)	(0.0142)	(0.0133)	(0.0147)
t=8	0.4157	0.4274	0.4462	0.4432	0.4461	0.4439
	(0.0194)	(0.0175)	(0.0158)	(0.0146)	(0.0137)	(0.0150)
t=9	0.4240	0.4343	0.4471	0.4467	0.4518	0.4452
	(0.0201)	(0.0181)	(0.0163)	(0.0150)	(0.0141)	(0.0155)
t=10	0.4066	0.4161	0.4240	0.4202	0.4245	0.4155
	(0.0208)	(0.0188)	(0.0171)	(0.0157)	(0.0147)	(0.0163)
t=11	0.3803	0.3777	0.3876	0.3881	0.3913	0.3846
	(0.0218)	(0.0198)	(0.0179)	(0.0165)	(0.0155)	(0.0171)
t=12	0.3378	0.3381	0.3402	0.3448	0.3504	0.3373
	(0.0233)	(0.0209)	(0.0190)	(0.0174)	(0.0163)	(0.0180)
t=13	0.3281	0.3268	0.3312	0.3272	0.3320	0.3305
	(0.0244)	(0.0218)	(0.0199)	(0.0183)	(0.0172)	(0.0189)
t=14	0.3011	0.2959	0.2972	0.2869	0.2943	0.2935
	(0.0257)	(0.0230)	(0.0210)	(0.0193)	(0.0182)	(0.0199)
t=15	0.2777	0.2696	0.2708	0.2595	0.2632	0.2684
	(0.0277)	(0.0248)	(0.0226)	(0.0208)	(0.0195)	(0.0214)
Constant	0.2250	0.2199	0.2100	0.2088	0.2095	0.2072
	(0.0120)	(0.0107)	(0.0096)	(0.0088)	(0.0083)	(0.0091)
Observations	73191	73191	73191	73191	73191	73191
Individuals	6103	6103	6103	6103	6103	6103

C. Obtaining a Medical PhD

Outcome: Obtaining a Medical PhD. Males (1/2 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=5	-0.0052	-0.0062	-0.0066	-0.0034	-0.0004	-0.0029
	(0.0107)	(0.0094)	(0.0085)	(0.0080)	(0.0074)	(0.0082)
Treatment # t=6	-0.0134	-0.0099	-0.0087	-0.0065	-0.0038	-0.0048
	(0.0129)	(0.0112)	(0.0102)	(0.0096)	(0.0089)	(0.0098)
Treatment # t=7	-0.0087	-0.0013	-0.0012	-0.0006	0.0026	0.0014
	(0.0158)	(0.0138)	(0.0126)	(0.0119)	(0.0109)	(0.0122)
Treatment # t=8	-0.0045	0.0095	0.0086	0.0120	0.0159	0.0114
	(0.0199)	(0.0173)	(0.0160)	(0.0150)	(0.0139)	(0.0154)
Treatment # t=9	-0.0219	-0.0038	-0.0064	-0.0024	0.0023	-0.0028
	(0.0229)	(0.0202)	(0.0186)	(0.0175)	(0.0162)	(0.0179)
Treatment # t=10	-0.0200	0.0000	-0.0011	0.0058	0.0077	0.0043
	(0.0249)	(0.0219)	(0.0202)	(0.0191)	(0.0178)	(0.0195)
Treatment # t=11	-0.0133	0.0024	-0.0011	0.0084	0.0113	0.0050
	(0.0267)	(0.0236)	(0.0217)	(0.0205)	(0.0192)	(0.0210)
Treatment # t=12	-0.0074	0.0080	0.0034	0.0141	0.0153	0.0107
	(0.0287)	(0.0254)	(0.0233)	(0.0219)	(0.0205)	(0.0225)
Treatment # t=13	-0.0041	0.0114	0.0077	0.0154	0.0138	0.0101
	(0.0305)	(0.0270)	(0.0248)	(0.0233)	(0.0218)	(0.0239)
Treatment # t=14	-0.0063	0.0092	0.0093	0.0220	0.0195	0.0174
	(0.0324)	(0.0288)	(0.0264)	(0.0248)	(0.0232)	(0.0254)
Treatment # t=15	-0.0053	0.0164	0.0229	0.0388	0.0366	0.0332
	(0.0343)	(0.0307)	(0.0280)	(0.0263)	(0.0247)	(0.0269)
Middle group # t=5	-0.0086	-0.0064	-0.0056	-0.0076	-0.0048	-0.0045
	(0.0088)	(0.0082)	(0.0080)	(0.0081)	(0.0087)	(0.0080)
Middle group # t=6	-0.0164	-0.0082	-0.0075	-0.0131	-0.0063	-0.0082
	(0.0107)	(0.0098)	(0.0096)	(0.0097)	(0.0106)	(0.0096)
Middle group # t=7	-0.0139	-0.0018	-0.0020	-0.0083	0.0063	-0.0035
	(0.0129)	(0.0120)	(0.0118)	(0.0120)	(0.0134)	(0.0119)
Middle group # t=8	-0.0140	0.0082	0.0032	-0.0051	0.0184	0.0016
	(0.0161)	(0.0150)	(0.0149)	(0.0152)	(0.0170)	(0.0150)
Middle group # t=9	-0.0179	0.0031	-0.0035	-0.0099	0.0100	-0.0058
	(0.0189)	(0.0177)	(0.0175)	(0.0179)	(0.0198)	(0.0177)
Middle group # t=10	-0.0049	0.0240	0.0168	0.0047	0.0099	0.0135
	(0.0206)	(0.0193)	(0.0192)	(0.0197)	(0.0217)	(0.0194)
Middle group # t=11	0.0010	0.0249	0.0175	0.0065	0.0039	0.0184
	(0.0220)	(0.0208)	(0.0207)	(0.0212)	(0.0233)	(0.0209)
Middle group # t=12	-0.0123	0.0122	0.0074	-0.0021	-0.0085	0.0104
	(0.0236)	(0.0223)	(0.0222)	(0.0226)	(0.0248)	(0.0224)
Middle group # t=13	-0.0070	0.0197	0.0124	0.0092	-0.0070	0.0142
	(0.0251)	(0.0237)	(0.0236)	(0.0242)	(0.0265)	(0.0238)
Middle group # t=14	-0.0125	0.0102	0.0062	0.0122	-0.0014	0.0164
	(0.0267)	(0.0253)	(0.0251)	(0.0257)	(0.0282)	(0.0253)
Middle group # t=15	-0.0064	0.0086	0.0206	0.0215	0.0056	0.0272
	(0.0284)	(0.0269)	(0.0266)	(0.0272)	(0.0298)	(0.0269)

Outcome: Obtaining a Medical PhD. Males (2/2)

	(1) 20	(2) 25	(3) 30	(4) 35	(5) 40	(6) Tercile(33)
t=6	0.0272	0.0227	0.0223	0.0237	0.0226	0.0228
	(0.0057)	(0.0047)	(0.0042)	(0.0041)	(0.0037)	(0.0041)
t=7	0.0656	0.0582	0.0587	0.0610	0.0583	0.0600
	(0.0087)	(0.0074)	(0.0068)	(0.0064)	(0.0059)	(0.0065)
t=8	0.1232	0.1090	0.1121	0.1141	0.1090	0.1134
	(0.0124)	(0.0107)	(0.0099)	(0.0093)	(0.0085)	(0.0095)
t=9	0.1827	0.1685	0.1729	0.1742	0.1696	0.1742
	(0.0154)	(0.0136)	(0.0125)	(0.0118)	(0.0109)	(0.0121)
t=10	0.2128	0.1955	0.2015	0.2050	0.2057	0.2034
	(0.0171)	(0.0150)	(0.0138)	(0.0131)	(0.0122)	(0.0133)
t=11	0.2380	0.2246	0.2314	0.2337	0.2356	0.2316
	(0.0190)	(0.0168)	(0.0154)	(0.0145)	(0.0136)	(0.0148)
t=12	0.2626	0.2473	0.2519	0.2522	0.2544	0.2504
	(0.0204)	(0.0182)	(0.0167)	(0.0157)	(0.0146)	(0.0160)
t=13	0.2726	0.2565	0.2624	0.2621	0.2684	0.2631
	(0.0220)	(0.0195)	(0.0179)	(0.0167)	(0.0157)	(0.0171)
t=14	0.2858	0.2712	0.2738	0.2683	0.2745	0.2693
	(0.0234)	(0.0208)	(0.0191)	(0.0178)	(0.0167)	(0.0183)
t=15	0.2844	0.2724	0.2664	0.2620	0.2686	0.2628
	(0.0248)	(0.0221)	(0.0201)	(0.0188)	(0.0176)	(0.0192)
Constant	0.0507	0.0493	0.0488	0.0481	0.0457	0.0471
	(0.0077)	(0.0068)	(0.0062)	(0.0057)	(0.0052)	(0.0058)
Observations	33270	33270	33270	33270	33270	33270
Individuals	3970	3970	3970	3970	3970	3970

Outcome: Obtaining a Medical PhD. Females (1/2 continues on next page)

	(1) 20	(2) 25	(3) 30	(4) 35	(5) 40	(6) Tercile(33)
Treatment # t=5	-0.0004	0.0007	0.0006	-0.0009	-0.0023	-0.0010
	(0.0050)	(0.0044)	(0.0041)	(0.0038)	(0.0036)	(0.0040)
Treatment # t=6	0.0025	0.0009	0.0015	0.0002	-0.0010	0.0002
	(0.0065)	(0.0056)	(0.0052)	(0.0048)	(0.0045)	(0.0050)
Treatment # t=7	-0.0036	-0.0049	-0.0044	-0.0081	-0.0074	-0.0091
	(0.0080)	(0.0071)	(0.0064)	(0.0060)	(0.0057)	(0.0062)
Treatment # t=8	-0.0156	-0.0152	-0.0166	-0.0120	-0.0106	-0.0177
	(0.0110)	(0.0098)	(0.0090)	(0.0083)	(0.0079)	(0.0086)
Treatment # t=9	-0.0294	-0.0326	-0.0287	-0.0185	-0.0180	-0.0259
	(0.0142)	(0.0129)	(0.0119)	(0.0109)	(0.0103)	(0.0113)
Treatment # t=10	-0.0463	-0.0499	-0.0461	-0.0381	-0.0387	-0.0461
	(0.0166)	(0.0150)	(0.0138)	(0.0127)	(0.0121)	(0.0131)
Treatment # t=11	-0.0630	-0.0614	-0.0546	-0.0416	-0.0405	-0.0537
	(0.0188)	(0.0170)	(0.0156)	(0.0144)	(0.0137)	(0.0149)
Treatment # t=12	-0.0641	-0.0627	-0.0569	-0.0405	-0.0377	-0.0535
	(0.0209)	(0.0188)	(0.0173)	(0.0160)	(0.0150)	(0.0165)
Treatment # t=13	-0.0680	-0.0625	-0.0539	-0.0376	-0.0324	-0.0512
	(0.0227)	(0.0205)	(0.0189)	(0.0174)	(0.0164)	(0.0179)
Treatment # t=14	-0.0726	-0.0664	-0.0569	-0.0357	-0.0289	-0.0504
	(0.0246)	(0.0223)	(0.0206)	(0.0190)	(0.0180)	(0.0196)
Treatment # t=15	-0.0620	-0.0638	-0.0519	-0.0310	-0.0177	-0.0451
	(0.0268)	(0.0242)	(0.0223)	(0.0206)	(0.0194)	(0.0212)
Middle group # t=5	0.0009	0.0025	0.0011	-0.0004	-0.0024	-0.0017
	(0.0041)	(0.0039)	(0.0039)	(0.0041)	(0.0044)	(0.0040)
Middle group # t=6	-0.0020	-0.0002	-0.0013	-0.0011	-0.0031	-0.0024
	(0.0052)	(0.0049)	(0.0048)	(0.0049)	(0.0054)	(0.0048)
Middle group # t=7	0.0000	0.0030	0.0056	0.0062	0.0004	0.0027
	(0.0068)	(0.0064)	(0.0064)	(0.0068)	(0.0074)	(0.0066)
Middle group # t=8	-0.0043	-0.0010	-0.0022	0.0030	-0.0057	-0.0016
	(0.0094)	(0.0090)	(0.0089)	(0.0092)	(0.0099)	(0.0090)
Middle group # t=9	-0.0049	-0.0093	-0.0119	0.0003	-0.0124	-0.0096
	(0.0123)	(0.0118)	(0.0115)	(0.0119)	(0.0129)	(0.0116)
Middle group # t=10	-0.0069	-0.0108	-0.0121	0.0015	-0.0161	-0.0111
	(0.0144)	(0.0138)	(0.0136)	(0.0141)	(0.0153)	(0.0138)
Middle group # t=11	-0.0104	-0.0082	-0.0109	0.0037	-0.0137	-0.0087
	(0.0163)	(0.0156)	(0.0153)	(0.0158)	(0.0173)	(0.0155)
Middle group # t=12	-0.0104	-0.0072	-0.0122	0.0025	-0.0047	-0.0083
	(0.0180)	(0.0171)	(0.0168)	(0.0173)	(0.0191)	(0.0170)
Middle group # t=13	-0.0071	-0.0012	-0.0052	0.0093	0.0015	-0.0042
	(0.0195)	(0.0186)	(0.0182)	(0.0188)	(0.0208)	(0.0185)
Middle group # t=14	-0.0001	-0.0000	-0.0057	0.0081	-0.0028	-0.0045
	(0.0211)	(0.0202)	(0.0198)	(0.0204)	(0.0225)	(0.0201)
Middle group # t=15	0.0037	0.0028	-0.0041	0.0012	-0.0082	-0.0103
	(0.0228)	(0.0218)	(0.0214)	(0.0220)	(0.0241)	(0.0217)

Outcome: Obtaining a Medical PhD. Females continued (2/2)

	(1) 20	(2) 25	(3) 30	(4) 35	(5) 40	(6) Tercile(33)
t=6	0.0099	0.0100	0.0094	0.0085	0.0083	0.0085
	(0.0028)	(0.0026)	(0.0023)	(0.0020)	(0.0019)	(0.0021)
t=7	0.0279	0.0279	0.0265	0.0273	0.0283	0.0280
	(0.0047)	(0.0042)	(0.0038)	(0.0035)	(0.0034)	(0.0037)
t=8	0.0652	0.0649	0.0656	0.0620	0.0631	0.0647
	(0.0076)	(0.0068)	(0.0062)	(0.0057)	(0.0054)	(0.0059)
t=9	0.1148	0.1198	0.1195	0.1115	0.1138	0.1164
	(0.0104)	(0.0094)	(0.0086)	(0.0077)	(0.0074)	(0.0081)
t=10	0.1597	0.1652	0.1652	0.1584	0.1633	0.1641
	(0.0123)	(0.0113)	(0.0103)	(0.0094)	(0.0089)	(0.0098)
t=11	0.2035	0.2051	0.2055	0.1973	0.2018	0.2041
	(0.0141)	(0.0127)	(0.0116)	(0.0106)	(0.0101)	(0.0111)
t=12	0.2331	0.2343	0.2361	0.2266	0.2283	0.2332
	(0.0155)	(0.0140)	(0.0128)	(0.0117)	(0.0111)	(0.0122)
t=13	0.2544	0.2536	0.2547	0.2459	0.2473	0.2534
	(0.0168)	(0.0151)	(0.0138)	(0.0126)	(0.0119)	(0.0131)
t=14	0.2715	0.2743	0.2760	0.2659	0.2670	0.2735
	(0.0182)	(0.0164)	(0.0150)	(0.0137)	(0.0130)	(0.0143)
t=15	0.2742	0.2794	0.2811	0.2735	0.2707	0.2810
	(0.0198)	(0.0179)	(0.0163)	(0.0149)	(0.0140)	(0.0155)
Constant	0.0156	0.0146	0.0154	0.0165	0.0175	0.0170
	(0.0036)	(0.0031)	(0.0029)	(0.0028)	(0.0027)	(0.0029)
Observations	50658	50658	50658	50658	50658	50658
Individuals	6103	6103	6103	6103	6103	6103

D. Occupational Choice—Female-Represented Specialty

Outcome: Obtaining a Female-Represented Specialty. Males (1/2 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=8	0.0154	0.0134	0.0171	0.0114	0.0140	0.0153
	(0.0135)	(0.0123)	(0.0112)	(0.0103)	(0.0096)	(0.0107)
Treatment # t=9	0.0365	0.0261	0.0226	0.0128	0.0164	0.0154
	(0.0198)	(0.0179)	(0.0165)	(0.0154)	(0.0143)	(0.0159)
Treatment # t=10	0.0356	0.0247	0.0219	0.0114	0.0102	0.0134
	(0.0245)	(0.0219)	(0.0200)	(0.0187)	(0.0175)	(0.0192)
Treatment # t=11	0.0253	0.0185	0.0149	0.0030	0.0038	0.0090
	(0.0274)	(0.0244)	(0.0223)	(0.0209)	(0.0196)	(0.0215)
Treatment # t=12	0.0112	0.0127	0.0072	-0.0027	0.0015	0.0011
	(0.0298)	(0.0266)	(0.0243)	(0.0227)	(0.0213)	(0.0234)
Treatment # t=13	0.0437	0.0357	0.0207	0.0122	0.0205	0.0147
	(0.0321)	(0.0286)	(0.0262)	(0.0245)	(0.0229)	(0.0252)
Treatment # t=14	0.0622	0.0502	0.0327	0.0218	0.0283	0.0254
	(0.0342)	(0.0306)	(0.0280)	(0.0262)	(0.0245)	(0.0269)
Treatment # t=15	0.0573	0.0480	0.0257	0.0150	0.0262	0.0208
	(0.0363)	(0.0325)	(0.0298)	(0.0280)	(0.0262)	(0.0287)
Middle group # t=8	0.0090	0.0007	0.0050	0.0067	0.0156	0.0067
	(0.0106)	(0.0103)	(0.0101)	(0.0105)	(0.0118)	(0.0102)
Middle group # t=9	0.0272	0.0140	0.0114	0.0098	0.0239	0.0054
	(0.0157)	(0.0153)	(0.0152)	(0.0158)	(0.0176)	(0.0155)
Middle group # t=10	0.0249	0.0140	0.0230	0.0302	0.0281	0.0219
	(0.0197)	(0.0190)	(0.0188)	(0.0195)	(0.0216)	(0.0191)
Middle group # t=11	0.0163	0.0107	0.0208	0.0167	0.0032	0.0142
	(0.0223)	(0.0213)	(0.0211)	(0.0218)	(0.0239)	(0.0214)
Middle group # t=12	0.0026	0.0113	0.0185	0.0187	0.0208	0.0114
	(0.0244)	(0.0233)	(0.0231)	(0.0238)	(0.0262)	(0.0234)
Middle group # t=13	0.0103	0.0112	0.0159	0.0184	0.0191	0.0123
	(0.0262)	(0.0250)	(0.0249)	(0.0256)	(0.0282)	(0.0251)
Middle group # t=14	-0.0052	0.0000	0.0179	0.0110	0.0026	0.0145
	(0.0281)	(0.0268)	(0.0266)	(0.0272)	(0.0299)	(0.0268)
Middle group # t=15	-0.0079	0.0054	0.0133	-0.0054	-0.0112	0.0014
	(0.0299)	(0.0286)	(0.0284)	(0.0290)	(0.0318)	(0.0286)

Outcome: Obtaining a Female-Represented Specialty. Males continued (2/2)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=9	0.0725	0.0779	0.0834	0.0862	0.0850	0.0880
	(0.0104)	(0.0096)	(0.0090)	(0.0085)	(0.0079)	(0.0088)
t=10	0.1612	0.1654	0.1663	0.1676	0.1738	0.1702
	(0.0152)	(0.0137)	(0.0125)	(0.0117)	(0.0111)	(0.0121)
t=11	0.2328	0.2329	0.2335	0.2389	0.2457	0.2385
	(0.0180)	(0.0162)	(0.0147)	(0.0138)	(0.0131)	(0.0142)
t=12	0.2929	0.2832	0.2860	0.2895	0.2922	0.2913
	(0.0203)	(0.0182)	(0.0166)	(0.0155)	(0.0145)	(0.0160)
t=13	0.3267	0.3223	0.3278	0.3294	0.3300	0.3315
	(0.0220)	(0.0197)	(0.0181)	(0.0169)	(0.0158)	(0.0174)
t=14	0.3690	0.3609	0.3603	0.3653	0.3671	0.3642
	(0.0240)	(0.0215)	(0.0198)	(0.0183)	(0.0172)	(0.0189)
t=15	0.3928	0.3801	0.3856	0.3940	0.3921	0.3914
	(0.0256)	(0.0230)	(0.0211)	(0.0197)	(0.0184)	(0.0202)
Constant	0.0630	0.0677	0.0643	0.0655	0.0627	0.0642
	(0.0090)	(0.0083)	(0.0074)	(0.0070)	(0.0064)	(0.0071)
Observations	21360	21360	21360	21360	21360	21360
Individuals	3570	3570	3570	3570	3570	3570

Outcome: Obtaining a Female-Represented Specialty. Females

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=9	0.0112	0.0237	0.0189	0.0142	0.0155	0.0172
	(0.0118)	(0.0106)	(0.0095)	(0.0087)	(0.0081)	(0.0090)
Treatment # t=10	0.0146	0.0283	0.0237	0.0188	0.0232	0.0265
	(0.0175)	(0.0158)	(0.0142)	(0.0132)	(0.0123)	(0.0136)
Treatment # t=11	0.0457	0.0496	0.0390	0.0286	0.0347	0.0374
	(0.0222)	(0.0201)	(0.0182)	(0.0168)	(0.0157)	(0.0173)
Treatment # t=12	0.0632	0.0622	0.0463	0.0287	0.0343	0.0392
	(0.0251)	(0.0225)	(0.0205)	(0.0189)	(0.0176)	(0.0194)
Treatment # t=13	0.0906	0.0803	0.0684	0.0492	0.0386	0.0600
	(0.0267)	(0.0240)	(0.0220)	(0.0203)	(0.0190)	(0.0208)
Treatment # t=14	0.1023	0.0902	0.0838	0.0603	0.0452	0.0741
	(0.0281)	(0.0253)	(0.0232)	(0.0215)	(0.0201)	(0.0221)
Treatment # t=15	0.0724	0.0640	0.0570	0.0378	0.0231	0.0528
	(0.0298)	(0.0268)	(0.0245)	(0.0226)	(0.0212)	(0.0233)
Middle group # t=9	-0.0051	-0.0028	-0.0002	-0.0008	-0.0027	-0.0021
	(0.0093)	(0.0085)	(0.0084)	(0.0087)	(0.0094)	(0.0085)
Middle group # t=10	0.0018	0.0020	0.0128	0.0055	0.0181	0.0100
	(0.0142)	(0.0134)	(0.0132)	(0.0137)	(0.0153)	(0.0134)
Middle group # t=11	0.0155	0.0034	0.0111	-0.0111	0.0068	0.0048
	(0.0179)	(0.0171)	(0.0168)	(0.0174)	(0.0192)	(0.0171)
Middle group # t=12	-0.0115	-0.0242	-0.0151	-0.0248	-0.0021	-0.0093
	(0.0203)	(0.0193)	(0.0190)	(0.0196)	(0.0217)	(0.0193)
Middle group # t=13	0.0097	0.0002	0.0104	-0.0011	0.0019	0.0207
	(0.0219)	(0.0209)	(0.0205)	(0.0212)	(0.0234)	(0.0208)
Middle group # t=14	0.0043	-0.0024	0.0061	0.0024	0.0095	0.0200
	(0.0234)	(0.0222)	(0.0219)	(0.0226)	(0.0249)	(0.0222)
Middle group # t=15	0.0111	0.0104	0.0115	0.0053	0.0156	0.0259
	(0.0248)	(0.0236)	(0.0232)	(0.0239)	(0.0263)	(0.0235)
t=10	0.0954	0.0966	0.0936	0.0967	0.0930	0.0931
	(0.0096)	(0.0087)	(0.0077)	(0.0072)	(0.0067)	(0.0073)
t=11	0.2101	0.2198	0.2189	0.2274	0.2198	0.2204
	(0.0145)	(0.0131)	(0.0118)	(0.0110)	(0.0102)	(0.0112)
t=12	0.3230	0.3308	0.3275	0.3317	0.3220	0.3248
	(0.0172)	(0.0155)	(0.0140)	(0.0130)	(0.0121)	(0.0133)
t=13	0.3812	0.3905	0.3872	0.3940	0.3960	0.3844
	(0.0191)	(0.0171)	(0.0155)	(0.0144)	(0.0134)	(0.0147)
t=14	0.4437	0.4511	0.4461	0.4508	0.4536	0.4416
	(0.0207)	(0.0186)	(0.0168)	(0.0155)	(0.0146)	(0.0160)
t=15	0.5055	0.5110	0.5117	0.5176	0.5211	0.5065
	(0.0223)	(0.0200)	(0.0181)	(0.0167)	(0.0156)	(0.0172)
Constant	0.0702	0.0648	0.0638	0.0646	0.0636	0.0643
	(0.0081)	(0.0070)	(0.0063)	(0.0059)	(0.0055)	(0.0060)
Observations	26835	26835	26835	26835	26835	26835
Individuals	4998	4998	4998	4998	4998	4998

E. Occupational Choice—Male-Represented Specialty

Outcome: Obtaining a Male-Represented Specialty. Males (1/2 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=8	0.0116	0.0081	0.0051	0.0071	0.0062	0.0065
	(0.0082)	(0.0071)	(0.0068)	(0.0063)	(0.0061)	(0.0065)
Treatment # t=9	0.0190	0.0125	0.0048	0.0145	0.0133	0.0120
	(0.0150)	(0.0136)	(0.0128)	(0.0119)	(0.0110)	(0.0122)
Treatment # t=10	0.0259	0.0070	-0.0011	0.0102	0.0118	0.0042
	(0.0200)	(0.0183)	(0.0171)	(0.0159)	(0.0148)	(0.0164)
Treatment # t=11	0.0380	0.0189	0.0177	0.0226	0.0185	0.0191
	(0.0234)	(0.0212)	(0.0196)	(0.0182)	(0.0169)	(0.0187)
Treatment # t=12	0.0241	0.0045	0.0047	0.0107	0.0017	0.0083
	(0.0266)	(0.0241)	(0.0222)	(0.0206)	(0.0192)	(0.0212)
Treatment # t=13	0.0005	-0.0073	0.0035	0.0106	0.0010	0.0103
	(0.0299)	(0.0270)	(0.0248)	(0.0231)	(0.0216)	(0.0237)
Treatment # t=14	-0.0017	-0.0145	0.0002	0.0075	-0.0005	0.0078
	(0.0326)	(0.0293)	(0.0269)	(0.0251)	(0.0235)	(0.0259)
Treatment # t=15	0.0018	-0.0091	0.0008	0.0032	-0.0070	0.0070
	(0.0351)	(0.0316)	(0.0290)	(0.0271)	(0.0254)	(0.0278)
Middle group # t=8	0.0081	0.0094	0.0046	0.0077	0.0010	0.0065
	(0.0062)	(0.0061)	(0.0063)	(0.0065)	(0.0070)	(0.0064)
Middle group # t=9	0.0176	0.0100	-0.0061	-0.0005	0.0029	-0.0002
	(0.0119)	(0.0116)	(0.0117)	(0.0118)	(0.0130)	(0.0117)
Middle group # t=10	0.0329	0.0080	-0.0102	-0.0041	0.0019	-0.0052
	(0.0161)	(0.0159)	(0.0159)	(0.0161)	(0.0177)	(0.0160)
Middle group # t=11	0.0311	0.0066	-0.0060	0.0089	0.0140	0.0053
	(0.0188)	(0.0183)	(0.0182)	(0.0186)	(0.0206)	(0.0183)
Middle group # t=12	0.0282	-0.0020	-0.0084	0.0085	0.0140	0.0033
	(0.0217)	(0.0211)	(0.0209)	(0.0214)	(0.0238)	(0.0211)
Middle group # t=13	0.0224	-0.0033	-0.0034	0.0195	0.0149	0.0158
	(0.0247)	(0.0238)	(0.0235)	(0.0241)	(0.0267)	(0.0237)
Middle group # t=14	0.0336	0.0144	0.0129	0.0411	0.0379	0.0280
	(0.0270)	(0.0260)	(0.0257)	(0.0264)	(0.0293)	(0.0259)
Middle group # t=15	0.0256	0.0096	0.0046	0.0368	0.0357	0.0252
	(0.0291)	(0.0280)	(0.0277)	(0.0284)	(0.0313)	(0.0279)

Outcome: Obtaining a Male-Represented Specialty. Males continued (2/2)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=9	0.0516	0.0574	0.0631	0.0587	0.0555	0.0593
	(0.0084)	(0.0079)	(0.0076)	(0.0069)	(0.0062)	(0.0071)
t=10	0.1101	0.1288	0.1355	0.1303	0.1253	0.1326
	(0.0126)	(0.0120)	(0.0114)	(0.0105)	(0.0096)	(0.0109)
t=11	0.1551	0.1728	0.1744	0.1683	0.1666	0.1704
	(0.0154)	(0.0143)	(0.0134)	(0.0123)	(0.0115)	(0.0128)
t=12	0.2110	0.2319	0.2306	0.2239	0.2246	0.2259
	(0.0183)	(0.0169)	(0.0157)	(0.0145)	(0.0136)	(0.0150)
t=13	0.2785	0.2948	0.2882	0.2797	0.2838	0.2802
	(0.0210)	(0.0191)	(0.0176)	(0.0163)	(0.0153)	(0.0168)
t=14	0.3197	0.3356	0.3303	0.3217	0.3273	0.3243
	(0.0231)	(0.0211)	(0.0194)	(0.0180)	(0.0168)	(0.0185)
t=15	0.3481	0.3608	0.3577	0.3489	0.3546	0.3499
	(0.0251)	(0.0229)	(0.0210)	(0.0195)	(0.0183)	(0.0201)
Constant	0.0192	0.0197	0.0230	0.0216	0.0237	0.0220
	(0.0051)	(0.0046)	(0.0045)	(0.0041)	(0.0040)	(0.0043)
Observations	21360	21360	21360	21360	21360	21360
Individuals	3570	3570	3570	3570	3570	3570

Outcome: Obtaining a Male-Represented Specialty. Females

	(1) 20	(2) 25	(3) 30	(4) 35	(5) 40	(6) Tercile(33)
Treatment # t=9	-0.0002	-0.0027	-0.0002	-0.0002	0.0016	0.0003
	(0.0046)	(0.0040)	(0.0037)	(0.0035)	(0.0033)	(0.0036)
Treatment # t=10	0.0081	0.0035	0.0052	0.0074	0.0088	0.0090
	(0.0084)	(0.0073)	(0.0067)	(0.0061)	(0.0057)	(0.0064)
Treatment # t=11	0.0059	0.0033	0.0063	0.0084	0.0096	0.0087
	(0.0115)	(0.0099)	(0.0090)	(0.0084)	(0.0079)	(0.0087)
Treatment # t=12	0.0104	0.0033	0.0102	0.0113	0.0123	0.0110
	(0.0137)	(0.0120)	(0.0110)	(0.0103)	(0.0098)	(0.0106)
Treatment # t=13	0.0109	0.0035	0.0080	0.0081	0.0097	0.0091
	(0.0163)	(0.0144)	(0.0132)	(0.0124)	(0.0117)	(0.0127)
Treatment # t=14	-0.0144	-0.0193	-0.0121	-0.0127	-0.0128	-0.0138
	(0.0187)	(0.0165)	(0.0150)	(0.0141)	(0.0133)	(0.0145)
Treatment # t=15	-0.0444	-0.0579	-0.0531	-0.0508	-0.0514	-0.0532
	(0.0214)	(0.0191)	(0.0174)	(0.0161)	(0.0150)	(0.0167)
Middle group # t=9	0.0014	0.0023	0.0029	0.0035	0.0028	0.0024
	(0.0039)	(0.0038)	(0.0037)	(0.0040)	(0.0044)	(0.0038)
Middle group # t=10	-0.0011	0.0012	0.0015	0.0026	0.0045	0.0019
	(0.0065)	(0.0062)	(0.0061)	(0.0062)	(0.0070)	(0.0060)
Middle group # t=11	-0.0058	0.0022	0.0052	0.0041	-0.0009	0.0009
	(0.0091)	(0.0086)	(0.0084)	(0.0087)	(0.0094)	(0.0084)
Middle group # t=12	0.0062	0.0128	0.0155	0.0133	0.0027	0.0111
	(0.0110)	(0.0106)	(0.0104)	(0.0109)	(0.0118)	(0.0106)
Middle group # t=13	0.0106	0.0145	0.0179	0.0121	0.0039	0.0132
	(0.0132)	(0.0127)	(0.0126)	(0.0131)	(0.0143)	(0.0128)
Middle group # t=14	-0.0058	0.0047	0.0137	0.0031	-0.0053	0.0038
	(0.0157)	(0.0149)	(0.0146)	(0.0152)	(0.0166)	(0.0149)
Middle group # t=15	-0.0243	-0.0283	-0.0221	-0.0284	-0.0238	-0.0330
	(0.0184)	(0.0175)	(0.0171)	(0.0174)	(0.0192)	(0.0172)
t=10	0.0193	0.0185	0.0184	0.0171	0.0162	0.0167
	(0.0044)	(0.0039)	(0.0035)	(0.0031)	(0.0029)	(0.0032)
t=11	0.0467	0.0422	0.0407	0.0404	0.0411	0.0413
	(0.0073)	(0.0063)	(0.0056)	(0.0051)	(0.0049)	(0.0054)
t=12	0.0632	0.0614	0.0600	0.0613	0.0639	0.0617
	(0.0091)	(0.0081)	(0.0073)	(0.0067)	(0.0064)	(0.0069)
t=13	0.0862	0.0863	0.0855	0.0885	0.0905	0.0874
	(0.0112)	(0.0100)	(0.0091)	(0.0085)	(0.0080)	(0.0087)
t=14	0.1195	0.1153	0.1115	0.1168	0.1197	0.1166
	(0.0135)	(0.0120)	(0.0108)	(0.0101)	(0.0095)	(0.0104)
t=15	0.1537	0.1586	0.1553	0.1567	0.1561	0.1592
	(0.0160)	(0.0146)	(0.0132)	(0.0121)	(0.0113)	(0.0126)
Constant	0.0110	0.0113	0.0107	0.0109	0.0106	0.0109
	(0.0033)	(0.0030)	(0.0027)	(0.0025)	(0.0023)	(0.0026)
Observations	26835	26835	26835	26835	26835	26835
Individuals	4998	4998	4998	4998	4998	4998

F. Total Compensation—Log-Earnings (Wages, Self-Employment Income, and Labor Market Pensions)

Outcome: Log-Earnings. Males (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=2	0.0209	0.0281	0.0269	0.0311	0.0270	0.0353
	(0.0216)	(0.0187)	(0.0166)	(0.0157)	(0.0143)	(0.0162)
Treatment # t=3	0.0170	0.0192	0.0249	0.0197	0.0247	0.0253
	(0.0197)	(0.0165)	(0.0157)	(0.0148)	(0.0141)	(0.0149)
Treatment # t=4	0.0274	0.0178	0.0143	0.0087	0.0099	0.0170
	(0.0181)	(0.0158)	(0.0146)	(0.0135)	(0.0131)	(0.0138)
Treatment # t=5	0.0127	0.0084	0.0060	0.0084	0.0066	0.0090
	(0.0203)	(0.0175)	(0.0159)	(0.0149)	(0.0137)	(0.0152)
Treatment # t=6	-0.0062	-0.0053	-0.0058	-0.0075	-0.0106	-0.0076
	(0.0196)	(0.0167)	(0.0150)	(0.0143)	(0.0134)	(0.0146)
Treatment # t=7	0.0028	0.0079	0.0041	-0.0017	-0.0057	0.0017
	(0.0268)	(0.0227)	(0.0206)	(0.0192)	(0.0181)	(0.0200)
Treatment # t=8	0.0190	0.0077	-0.0007	-0.0061	-0.0135	-0.0040
	(0.0186)	(0.0165)	(0.0165)	(0.0157)	(0.0162)	(0.0164)
Treatment # t=9	0.0387	0.0264	0.0281	0.0218	0.0141	0.0218
	(0.0184)	(0.0166)	(0.0158)	(0.0150)	(0.0146)	(0.0156)
Treatment # t=10	0.0230	0.0111	-0.0001	-0.0018	-0.0128	0.0008
	(0.0195)	(0.0184)	(0.0186)	(0.0171)	(0.0190)	(0.0178)
Treatment # t=11	0.0348	0.0144	0.0132	0.0137	0.0152	0.0124
	(0.0264)	(0.0229)	(0.0206)	(0.0188)	(0.0190)	(0.0197)
Treatment # t=12	0.0231	-0.0031	-0.0176	-0.0055	-0.0035	-0.0088
	(0.0254)	(0.0224)	(0.0209)	(0.0194)	(0.0185)	(0.0202)
Treatment # t=13	0.0181	0.0079	0.0028	0.0057	0.0028	0.0009
	(0.0267)	(0.0233)	(0.0223)	(0.0206)	(0.0194)	(0.0214)
Treatment # t=14	-0.0140	-0.0046	-0.0059	0.0025	-0.0016	-0.0003
	(0.0247)	(0.0219)	(0.0204)	(0.0190)	(0.0184)	(0.0197)
Treatment # t=15	0.0091	0.0027	0.0018	0.0101	0.0176	0.0088
	(0.0403)	(0.0350)	(0.0310)	(0.0281)	(0.0253)	(0.0292)

Outcome: Log-Earnings. Males continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=2	0.0016	0.0043	0.0030	0.0210	0.0254	0.0217
	(0.0138)	(0.0142)	(0.0139)	(0.0145)	(0.0161)	(0.0144)
Middle group # t=3	0.0217	0.0128	0.0199	0.0270	0.0301	0.0226
	(0.0159)	(0.0153)	(0.0161)	(0.0170)	(0.0200)	(0.0167)
Middle group # t=4	0.0174	0.0111	0.0057	0.0118	0.0268	0.0109
	(0.0147)	(0.0141)	(0.0140)	(0.0150)	(0.0168)	(0.0145)
Middle group # t=5	0.0098	0.0022	0.0017	0.0182	0.0179	0.0125
	(0.0176)	(0.0162)	(0.0153)	(0.0155)	(0.0171)	(0.0154)
Middle group # t=6	-0.0061	-0.0114	-0.0125	0.0052	0.0125	-0.0005
	(0.0147)	(0.0139)	(0.0135)	(0.0136)	(0.0138)	(0.0134)
Middle group # t=7	-0.0022	-0.0011	0.0021	0.0140	0.0364	0.0146
	(0.0188)	(0.0180)	(0.0171)	(0.0168)	(0.0152)	(0.0168)
Middle group # t=8	-0.0268	-0.0248	-0.0302	-0.0244	0.0005	-0.0214
	(0.0165)	(0.0163)	(0.0165)	(0.0178)	(0.0170)	(0.0169)
Middle group # t=9	0.0011	-0.0015	-0.0109	-0.0004	0.0219	-0.0032
	(0.0164)	(0.0159)	(0.0161)	(0.0170)	(0.0179)	(0.0163)
Middle group # t=10	-0.0094	-0.0096	-0.0235	-0.0254	0.0068	-0.0270
	(0.0193)	(0.0196)	(0.0203)	(0.0229)	(0.0212)	(0.0215)
Middle group # t=11	0.0200	0.0034	-0.0197	-0.0196	0.0158	-0.0185
	(0.0244)	(0.0223)	(0.0220)	(0.0232)	(0.0235)	(0.0224)
Middle group # t=12	0.0106	0.0062	-0.0047	0.0107	0.0383	0.0012
	(0.0210)	(0.0197)	(0.0193)	(0.0201)	(0.0207)	(0.0196)
Middle group # t=13	-0.0008	0.0014	0.0086	0.0218	0.0282	0.0115
	(0.0250)	(0.0236)	(0.0240)	(0.0249)	(0.0287)	(0.0242)
Middle group # t=14	-0.0117	-0.0081	-0.0052	0.0110	0.0237	0.0027
	(0.0190)	(0.0187)	(0.0189)	(0.0201)	(0.0213)	(0.0194)
Middle group # t=15	0.0043	0.0039	-0.0040	0.0148	0.0146	0.0070
	(0.0382)	(0.0332)	(0.0308)	(0.0305)	(0.0342)	(0.0304)

Outcome: Log-Earnings. Males continued (3/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
t=3	0.0638	0.0731	0.0690	0.0772	0.0751	0.0780
	(0.0160)	(0.0147)	(0.0139)	(0.0133)	(0.0129)	(0.0139)
t=4	0.1544	0.1644	0.1679	0.1757	0.1717	0.1749
	(0.0151)	(0.0144)	(0.0125)	(0.0121)	(0.0118)	(0.0126)
t=5	0.1820	0.1912	0.1920	0.1940	0.1949	0.1970
	(0.0180)	(0.0165)	(0.0144)	(0.0142)	(0.0131)	(0.0145)
t=6	0.2368	0.2430	0.2427	0.2449	0.2443	0.2484
	(0.0158)	(0.0148)	(0.0129)	(0.0129)	(0.0119)	(0.0131)
t=7	0.2861	0.2880	0.2874	0.2936	0.2910	0.2936
	(0.0194)	(0.0183)	(0.0162)	(0.0153)	(0.0140)	(0.0160)
t=8	0.3441	0.3463	0.3483	0.3536	0.3480	0.3546
	(0.0165)	(0.0150)	(0.0135)	(0.0129)	(0.0119)	(0.0135)
t=9	0.3825	0.3892	0.3909	0.3956	0.3918	0.3989
	(0.0170)	(0.0157)	(0.0142)	(0.0136)	(0.0125)	(0.0142)
t=10	0.4370	0.4420	0.4494	0.4565	0.4507	0.4590
	(0.0179)	(0.0155)	(0.0142)	(0.0138)	(0.0129)	(0.0145)
t=11	0.4803	0.4980	0.5069	0.5123	0.5008	0.5153
	(0.0244)	(0.0215)	(0.0191)	(0.0178)	(0.0163)	(0.0187)
t=12	0.5421	0.5552	0.5647	0.5640	0.5578	0.5697
	(0.0209)	(0.0184)	(0.0166)	(0.0160)	(0.0149)	(0.0166)
t=13	0.5998	0.6044	0.6030	0.6065	0.6071	0.6127
	(0.0245)	(0.0213)	(0.0200)	(0.0185)	(0.0167)	(0.0193)
t=14	0.6683	0.6678	0.6665	0.6662	0.6651	0.6715
	(0.0190)	(0.0172)	(0.0157)	(0.0153)	(0.0142)	(0.0159)
t=15	0.6704	0.6763	0.6800	0.6788	0.6755	0.6834
	(0.0375)	(0.0313)	(0.0272)	(0.0249)	(0.0224)	(0.0260)
Constant	13.1616	13.1575	13.1575	13.1496	13.1509	13.1477
	(0.0115)	(0.0115)	(0.0100)	(0.0101)	(0.0093)	(0.0105)
Observations	40395	40395	40395	40395	40395	40395
Individuals	3953	3953	3953	3953	3953	3953

Outcome: Log-Earnings. Females (1/3 continues on next page)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Treatment # t=2	0.0238	0.0266	0.0216	0.0214	0.0151	0.0218
	(0.0139)	(0.0126)	(0.0118)	(0.0112)	(0.0104)	(0.0115)
Treatment # t=3	-0.0140	-0.0201	-0.0070	-0.0085	-0.0077	-0.0088
	(0.0145)	(0.0125)	(0.0116)	(0.0106)	(0.0097)	(0.0110)
Treatment # t=4	0.0105	0.0071	0.0092	0.0074	0.0049	0.0128
	(0.0141)	(0.0125)	(0.0117)	(0.0118)	(0.0108)	(0.0117)
Treatment # t=5	0.0300	0.0219	0.0182	0.0151	0.0150	0.0165
	(0.0203)	(0.0170)	(0.0150)	(0.0134)	(0.0122)	(0.0140)
Treatment # t=6	-0.0227	-0.0199	-0.0129	-0.0048	0.0009	-0.0020
	(0.0175)	(0.0149)	(0.0137)	(0.0134)	(0.0122)	(0.0140)
Treatment # t=7	0.0286	0.0200	0.0100	0.0040	0.0046	0.0061
	(0.0150)	(0.0130)	(0.0129)	(0.0118)	(0.0109)	(0.0123)
Treatment # t=8	0.0086	0.0101	0.0073	-0.0025	-0.0017	0.0014
	(0.0142)	(0.0139)	(0.0128)	(0.0122)	(0.0110)	(0.0127)
Treatment # t=9	0.0132	0.0103	0.0037	-0.0121	-0.0135	0.0010
	(0.0186)	(0.0165)	(0.0149)	(0.0154)	(0.0141)	(0.0146)
Treatment # t=10	0.0094	0.0063	-0.0008	-0.0111	-0.0144	-0.0036
	(0.0188)	(0.0162)	(0.0148)	(0.0135)	(0.0128)	(0.0140)
Treatment # t=11	0.0060	0.0017	-0.0033	-0.0019	0.0059	0.0023
	(0.0200)	(0.0192)	(0.0173)	(0.0163)	(0.0156)	(0.0169)
Treatment # t=12	0.0443	0.0516	0.0372	0.0178	0.0136	0.0245
	(0.0238)	(0.0215)	(0.0195)	(0.0185)	(0.0167)	(0.0194)
Treatment # t=13	-0.0122	-0.0024	-0.0066	-0.0238	-0.0204	-0.0202
	(0.0279)	(0.0239)	(0.0219)	(0.0206)	(0.0189)	(0.0216)
Treatment # t=14	-0.0064	0.0082	0.0076	-0.0082	-0.0096	-0.0016
	(0.0308)	(0.0267)	(0.0251)	(0.0236)	(0.0212)	(0.0248)
Treatment # t=15	-0.0213	-0.0176	-0.0144	-0.0216	-0.0216	-0.0153
	(0.0353)	(0.0299)	(0.0270)	(0.0263)	(0.0239)	(0.0275)

Outcome: Log-Earnings. Females continued (2/3)

	(1)	(2)	(3)	(4)	(5)	(6)
	20	25	30	35	40	Tercile(33)
Middle group # t=2	-0.0127	-0.0019	0.0051	0.0112	0.0089	0.0115
	(0.0103)	(0.0104)	(0.0105)	(0.0109)	(0.0116)	(0.0109)
Middle group # t=3	-0.0091	-0.0162	-0.0085	-0.0091	-0.0156	-0.0096
	(0.0099)	(0.0095)	(0.0099)	(0.0103)	(0.0120)	(0.0100)
Middle group # t=4	0.0084	0.0071	0.0117	0.0211	0.0066	0.0199
	(0.0116)	(0.0111)	(0.0113)	(0.0113)	(0.0126)	(0.0118)
Middle group # t=5	0.0343	0.0241	0.0293	0.0192	0.0235	0.0213
	(0.0171)	(0.0148)	(0.0134)	(0.0130)	(0.0134)	(0.0130)
Middle group # t=6	-0.0077	-0.0163	-0.0019	0.0156	0.0129	0.0151
	(0.0106)	(0.0105)	(0.0112)	(0.0118)	(0.0126)	(0.0118)
Middle group # t=7	0.0080	0.0034	0.0142	0.0112	0.0169	0.0068
	(0.0143)	(0.0128)	(0.0117)	(0.0113)	(0.0113)	(0.0113)
Middle group # t=8	-0.0093	-0.0081	-0.0074	-0.0126	-0.0179	-0.0079
	(0.0125)	(0.0119)	(0.0119)	(0.0124)	(0.0153)	(0.0119)
Middle group # t=9	0.0041	-0.0026	-0.0048	-0.0031	-0.0008	-0.0044
	(0.0160)	(0.0147)	(0.0144)	(0.0130)	(0.0134)	(0.0147)
Middle group # t=10	0.0008	-0.0102	-0.0188	-0.0321	-0.0328	-0.0272
	(0.0150)	(0.0139)	(0.0137)	(0.0144)	(0.0163)	(0.0139)
Middle group # t=11	-0.0082	-0.0125	-0.0157	-0.0177	0.0020	-0.0145
	(0.0156)	(0.0147)	(0.0149)	(0.0159)	(0.0172)	(0.0154)
Middle group # t=12	0.0134	0.0147	0.0072	-0.0028	-0.0173	0.0005
	(0.0188)	(0.0185)	(0.0177)	(0.0170)	(0.0193)	(0.0168)
Middle group # t=13	-0.0053	-0.0118	-0.0100	-0.0258	-0.0342	-0.0208
	(0.0182)	(0.0177)	(0.0176)	(0.0174)	(0.0201)	(0.0168)
Middle group # t=14	-0.0212	-0.0201	-0.0029	-0.0153	-0.0149	-0.0050
	(0.0192)	(0.0192)	(0.0192)	(0.0183)	(0.0206)	(0.0180)
Middle group # t=15	-0.0086	-0.0274	-0.0197	-0.0182	-0.0178	-0.0015
	(0.0213)	(0.0206)	(0.0206)	(0.0199)	(0.0221)	(0.0196)

Outcome: Log-Earnings. Females continued (3/3)

	(1) 20	(2) 25	(3) 30	(4) 35	(5) 40	(6) Tercile(33)
t=3	0.0499	0.0634	0.0584	0.0610	0.0585	0.0617
	(0.0108)	(0.0100)	(0.0091)	(0.0089)	(0.0082)	(0.0092)
t=4	0.0836	0.0939	0.0945	0.0955	0.0980	0.0936
	(0.0114)	(0.0106)	(0.0098)	(0.0103)	(0.0094)	(0.0107)
t=5	0.0744	0.0919	0.0950	0.1035	0.1008	0.1022
	(0.0174)	(0.0150)	(0.0131)	(0.0122)	(0.0111)	(0.0127)
t=6	0.1437	0.1562	0.1505	0.1452	0.1422	0.1440
	(0.0111)	(0.0104)	(0.0102)	(0.0113)	(0.0104)	(0.0118)
t=7	0.2012	0.2136	0.2144	0.2207	0.2172	0.2214
	(0.0152)	(0.0133)	(0.0120)	(0.0113)	(0.0102)	(0.0117)
t=8	0.2473	0.2535	0.2555	0.2617	0.2582	0.2595
	(0.0128)	(0.0117)	(0.0107)	(0.0104)	(0.0095)	(0.0109)
t=9	0.2645	0.2769	0.2817	0.2885	0.2859	0.2846
	(0.0161)	(0.0141)	(0.0126)	(0.0121)	(0.0110)	(0.0127)
t=10	0.3007	0.3152	0.3222	0.3301	0.3259	0.3273
	(0.0154)	(0.0136)	(0.0121)	(0.0113)	(0.0105)	(0.0118)
t=11	0.3529	0.3636	0.3678	0.3687	0.3570	0.3672
	(0.0149)	(0.0135)	(0.0123)	(0.0120)	(0.0123)	(0.0125)
t=12	0.3857	0.3911	0.4000	0.4109	0.4112	0.4083
	(0.0182)	(0.0171)	(0.0152)	(0.0138)	(0.0127)	(0.0145)
t=13	0.4616	0.4711	0.4732	0.4855	0.4814	0.4835
	(0.0172)	(0.0159)	(0.0143)	(0.0131)	(0.0126)	(0.0138)
t=14	0.5095	0.5123	0.5058	0.5165	0.5129	0.5117
	(0.0180)	(0.0166)	(0.0159)	(0.0149)	(0.0137)	(0.0156)
t=15	0.5470	0.5642	0.5610	0.5641	0.5602	0.5570
	(0.0192)	(0.0169)	(0.0154)	(0.0154)	(0.0145)	(0.0161)
Constant	13.0698	13.0613	13.0585	13.0562	13.0592	13.0559
	(0.0084)	(0.0081)	(0.0077)	(0.0077)	(0.0070)	(0.0080)
Observations	62414	62414	62414	62414	62414	62414
Individuals	6088	6088	6088	6088	6088	6088

Notes: These tables investigate the robustness of our design by studying the effects on our main outcomes when we vary the percentiles that define the experimental groups. We estimate versions of specification (1), and we report estimates for β_τ (denoted by "Treatment" in the tables) and for α_τ (which capture baselines among the control group). We also report estimates for the effects on the middle group (relative to the control group) from a simple extension to specification (1) (denoted by "Middle group" in the tables). Columns 1-5 report estimates for thresholds that vary in five percentage-point increments, where column 3 corresponds to our main specification. Column 6 reports estimates where the treatment, control, and middle groups are split at the 33rd and 67th percentiles (as a potentially natural benchmark). Robust standard errors clustered at the individual level are reported in parentheses.

Appendix Table H.2: Graduation Round Fixed Effects

Males (1/3 continues on next page)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
Treatment # t=0	0.0035	-0.0193				
	(0.0062)	(0.0168)				
Treatment # t=1	0.0680	-0.1784				
	(0.0106)	(0.0184)				
Treatment # t=2	0.1086	-0.3493				0.0261
	(0.0141)	(0.0184)				(0.0165)
Treatment # t=3	0.0574	-0.1814				0.0236
	(0.0137)	(0.0193)				(0.0157)
Treatment # t=4	0.0280	-0.0972				0.0131
	(0.0133)	(0.0193)				(0.0145)
Treatment # t=5	0.0312	-0.0519	-0.0042			0.0044
	(0.0131)	(0.0194)	(0.0086)			(0.0158)
Treatment # t=6	0.0220	-0.0555	-0.0063			-0.0073
	(0.0130)	(0.0201)	(0.0101)			(0.0150)
Treatment # t=7	0.0130	-0.0426	0.0012			0.0024
	(0.0139)	(0.0209)	(0.0125)			(0.0206)
Treatment # t=8	-0.0062	-0.0202	0.0113	0.0156	0.0070	-0.0023
	(0.0144)	(0.0220)	(0.0159)	(0.0112)	(0.0069)	(0.0165)
Treatment # t=9	-0.0140	-0.0086	-0.0033	0.0218	0.0069	0.0274
	(0.0151)	(0.0232)	(0.0186)	(0.0165)	(0.0127)	(0.0158)
Treatment # t=10	0.0032	0.0025	0.0016	0.0210	0.0014	-0.0007
	(0.0159)	(0.0247)	(0.0202)	(0.0200)	(0.0170)	(0.0185)
Treatment # t=11	-0.0077	0.0006	0.0000	0.0145	0.0195	0.0129
	(0.0169)	(0.0262)	(0.0217)	(0.0223)	(0.0196)	(0.0205)
Treatment # t=12	-0.0026	0.0007	0.0040	0.0068	0.0063	-0.0176
	(0.0180)	(0.0279)	(0.0233)	(0.0242)	(0.0222)	(0.0209)
Treatment # t=13	0.0120	-0.0046	0.0082	0.0208	0.0050	0.0029
	(0.0194)	(0.0297)	(0.0248)	(0.0262)	(0.0248)	(0.0223)
Treatment # t=14	0.0181	-0.0044	0.0095	0.0324	0.0018	-0.0053
	(0.0206)	(0.0316)	(0.0264)	(0.0279)	(0.0269)	(0.0205)
Treatment # t=15	0.0266	0.0155	0.0231	0.0263	0.0025	0.0015
	(0.0214)	(0.0342)	-0.0042	(0.0297)	(0.0290)	(0.0309)

Males (2/3)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
Middle group # t=0	0.0009	0.0052				
	(0.0055)	(0.0160)				
Middle group # t=1	0.0205	-0.0798				
	(0.0082)	(0.0181)				
Middle group # t=2	0.0259	-0.1641				0.0040
	(0.0114)	(0.0182)				(0.0138)
Middle group # t=3	0.0076	-0.1054				0.0206
	(0.0117)	(0.0179)				(0.0161)
Middle group # t=4	0.0073	-0.0483				0.0066
	(0.0120)	(0.0178)				(0.0140)
Middle group # t=5	0.0179	-0.0122	(0.0280)			0.0025
	(0.0118)	(0.0179)	-0.0057			(0.0153)
Middle group # t=6	0.0047	-0.0396	(0.0081)			-0.0119
	(0.0117)	(0.0186)	-0.0077			(0.0135)
Middle group # t=7	0.0047	-0.0170	(0.0095)			0.0025
	(0.0127)	(0.0193)	-0.0021			(0.0171)
Middle group # t=8	0.0003	0.0022	(0.0117)	0.0035	0.0060	-0.0296
	(0.0137)	(0.0206)	0.0030	(0.0101)	(0.0063)	(0.0164)
Middle group # t=9	-0.0030	0.0037	(0.0148)	0.0102	-0.0044	-0.0105
	(0.0145)	(0.0219)	-0.0036	(0.0152)	(0.0116)	(0.0161)
Middle group # t=10	0.0104	-0.0257	(0.0175)	0.0217	-0.0087	-0.0234
	(0.0153)	(0.0236)	0.0170	(0.0188)	(0.0158)	(0.0203)
Middle group # t=11	0.0012	-0.0150	(0.0192)	0.0195	-0.0039	-0.0201
	(0.0163)	(0.0249)	0.0176	(0.0211)	(0.0182)	(0.0220)
Middle group # t=12	0.0017	-0.0003	(0.0207)	0.0169	-0.0061	-0.0045
	(0.0173)	(0.0265)	0.0076	(0.0230)	(0.0209)	(0.0194)
Middle group # t=13	0.0031	-0.0162	(0.0221)	0.0155	-0.0013	0.0088
	(0.0183)	(0.0284)	0.0131	(0.0248)	(0.0235)	(0.0240)
Middle group # t=14	0.0175	-0.0115	(0.0236)	0.0166	0.0150	-0.0047
	(0.0197)	(0.0303)	0.0065	(0.0266)	(0.0257)	(0.0191)
Middle group # t=15	0.0168	-0.0090	(0.0251)	0.0133	0.0069	-0.0038
	(0.0202)	(0.0328)	0.0210	(0.0284)	(0.0277)	(0.0307)

Males (3/3)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
t=1	0.0198	0.1835				
	(0.0053)	(0.0145)				
t=2	0.0678	0.4025				
	(0.0080)	(0.0166)				
t=3	0.0835	0.4479				0.0693
	(0.0085)	(0.0169)				(0.0139)
t=4	0.0876	0.4455				0.1680
	(0.0090)	(0.0173)				(0.0125)
t=5	0.0802	0.4322				0.1923
	(0.0089)	(0.0174)				(0.0144)
t=6	0.0826	0.4697	0.0223			0.2430
	(0.0091)	(0.0178)	(0.0043)			(0.0129)
t=7	0.0890	0.4731	0.0587			0.2821
	(0.0097)	(0.0183)	(0.0068)			(0.0163)
t=8	0.0971	0.4576	0.1101			0.3419
	(0.0105)	(0.0195)	(0.0098)			(0.0135)
t=9	0.1018	0.4368	0.1719	0.0815	0.0625	0.3814
	(0.0111)	(0.0202)	(0.0123)	(0.0089)	(0.0076)	(0.0143)
t=10	0.0977	0.4107	0.2054	0.1645	0.1349	0.4385
	(0.0114)	(0.0209)	(0.0136)	(0.0124)	(0.0114)	(0.0142)
t=11	0.1043	0.3964	0.2414	0.2302	0.1785	0.4955
	(0.0122)	(0.0217)	(0.0150)	(0.0146)	(0.0133)	(0.0192)
t=12	0.1014	0.3758	0.2663	0.2825	0.2384	0.5552
	(0.0128)	(0.0228)	(0.0161)	(0.0163)	(0.0154)	(0.0166)
t=13	0.0966	0.3630	0.2818	0.3275	0.2957	0.5940
	(0.0133)	(0.0238)	(0.0172)	(0.0177)	(0.0173)	(0.0201)
t=14	0.0914	0.3510	0.2968	0.3657	0.3390	0.6556
	(0.0141)	(0.0250)	(0.0183)	(0.0192)	(0.0189)	(0.0160)
t=15	0.0822	0.3227	0.2924	0.3918	0.3697	0.6710
	(0.0146)	(0.0267)	(0.0191)	(0.0204)	(0.0204)	(0.0269)
Constant	0.0231	0.2252	0.0420	0.0653	0.0177	13.1621
	(0.0043)	(0.0119)	(0.0061)	(0.0078)	(0.0047)	(0.0101)
Observations	50774	48186	33270	21360	21360	40395
Individuals	3970	3970	3970	3570	3570	3953

Females (1/3 continues on next page)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
Treatment # t=0	0.0037	0.0089				
	(0.0056)	(0.0133)				
Treatment # t=1	0.0418	-0.1702				
	(0.0090)	(0.0148)				
Treatment # t=2	0.0841	-0.3792				0.0219
	(0.0114)	(0.0149)				(0.0117)
Treatment # t=3	0.0497	-0.2247				-0.0068
	(0.0110)	(0.0156)				(0.0116)
Treatment # t=4	0.0342	-0.0916				0.0095
	(0.0106)	(0.0157)				(0.0118)
Treatment # t=5	0.0186	-0.0686	0.0008			0.0184
	(0.0104)	(0.0156)	(0.0041)			(0.0150)
Treatment # t=6	0.0232	-0.0378	0.0016			-0.0126
	(0.0102)	(0.0166)	(0.0052)			(0.0136)
Treatment # t=7	0.0214	-0.0496	-0.0042			0.0101
	(0.0110)	(0.0177)	(0.0064)			(0.0128)
Treatment # t=8	0.0296	-0.0485	-0.0164			0.0072
	(0.0119)	(0.0182)	(0.0089)			(0.0128)
Treatment # t=9	0.0240	-0.0599	-0.0285	0.0179	-0.0004	0.0031
	(0.0124)	(0.0191)	(0.0118)	(0.0095)	(0.0038)	(0.0149)
Treatment # t=10	0.0213	-0.0784	-0.0459	0.0222	0.0050	-0.0013
	(0.0131)	(0.0203)	(0.0138)	(0.0142)	(0.0067)	(0.0148)
Treatment # t=11	0.0298	-0.0623	-0.0543	0.0374	0.0061	-0.0038
	(0.0136)	(0.0217)	(0.0156)	(0.0181)	(0.0090)	(0.0173)
Treatment # t=12	0.0312	-0.0521	-0.0565	0.0445	0.0100	0.0370
	(0.0145)	(0.0233)	(0.0172)	(0.0204)	(0.0111)	(0.0196)
Treatment # t=13	0.0390	-0.0782	-0.0536	0.0665	0.0078	-0.0075
	(0.0156)	(0.0250)	(0.0188)	(0.0218)	(0.0132)	(0.0219)
Treatment # t=14	0.0490	-0.0491	-0.0566	0.0817	-0.0120	0.0064
	(0.0165)	(0.0269)	(0.0205)	(0.0232)	(0.0150)	(0.0252)
Treatment # t=15	0.0467	-0.0630	-0.0520	0.0533	-0.0534	-0.0152
	(0.0176)	(0.0292)	(0.0222)	(0.0245)	(0.0174)	(0.0271)

Females (2/3)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
Middle group # t=0	0.0016	0.0008				
	(0.0052)	(0.0124)				
Middle group # t=1	0.0001	-0.1088				
	(0.0075)	(0.0145)				
Middle group # t=2	0.0153	-0.2289				0.0041
	(0.0096)	(0.0147)				(0.0105)
Middle group # t=3	0.0119	-0.1265				-0.0096
	(0.0097)	(0.0146)				(0.0099)
Middle group # t=4	0.0120	-0.0567				0.0105
	(0.0096)	(0.0146)				(0.0113)
Middle group # t=5	0.0024	-0.0346	0.0016			0.0283
	(0.0096)	(0.0146)	(0.0039)			(0.0134)
Middle group # t=6	0.0032	-0.0055	-0.0008			-0.0028
	(0.0092)	(0.0156)	(0.0048)			(0.0111)
Middle group # t=7	-0.0048	0.0072	0.0060			0.0129
	(0.0099)	(0.0165)	(0.0064)			(0.0116)
Middle group # t=8	-0.0127	-0.0077	-0.0018			-0.0087
	(0.0103)	(0.0170)	(0.0088)			(0.0119)
Middle group # t=9	-0.0081	-0.0074	-0.0114	-0.0024	0.0028	-0.0062
	(0.0111)	(0.0177)	(0.0115)	(0.0084)	(0.0037)	(0.0144)
Middle group # t=10	-0.0102	-0.0051	-0.0116	0.0105	0.0014	-0.0198
	(0.0116)	(0.0188)	(0.0135)	(0.0131)	(0.0061)	(0.0136)
Middle group # t=11	-0.0058	-0.0064	-0.0106	0.0090	0.0051	-0.0166
	(0.0120)	(0.0202)	(0.0152)	(0.0167)	(0.0084)	(0.0148)
Middle group # t=12	-0.0026	0.0145	-0.0119	-0.0174	0.0154	0.0066
	(0.0128)	(0.0217)	(0.0167)	(0.0189)	(0.0104)	(0.0177)
Middle group # t=13	0.0007	-0.0047	-0.0048	0.0080	0.0178	-0.0110
	(0.0136)	(0.0234)	(0.0181)	(0.0204)	(0.0126)	(0.0175)
Middle group # t=14	0.0103	-0.0079	-0.0051	0.0037	0.0138	-0.0047
	(0.0143)	(0.0252)	(0.0197)	(0.0218)	(0.0146)	(0.0191)
Middle group # t=15	0.0103	-0.0153	-0.0038	0.0075	-0.0223	-0.0213
	(0.0153)	(0.0276)	(0.0213)	(0.0232)	(0.0170)	(0.0205)

Females (3/3)

	Rural	University Hospital	PhD	Female Specialty	Male Specialty	Log Earnings
t=1	0.0353	0.1963				
	(0.0052)	(0.0122)				
t=2	0.0750	0.4460				
	(0.0068)	(0.0137)				
t=3	0.0788	0.4802				0.0587
	(0.0071)	(0.0137)				(0.0091)
t=4	0.0766	0.4702				0.0948
	(0.0071)	(0.0140)				(0.0098)
t=5	0.0788	0.4724				0.0953
	(0.0072)	(0.0141)				(0.0131)
t=6	0.0689	0.4472	0.0094			0.1505
	(0.0069)	(0.0147)	(0.0023)			(0.0102)
t=7	0.0733	0.4504	0.0265			0.2122
	(0.0075)	(0.0153)	(0.0038)			(0.0121)
t=8	0.0735	0.4783	0.0661			0.2491
	(0.0078)	(0.0156)	(0.0062)			(0.0108)
t=9	0.0749	0.4870	0.1202			0.2730
	(0.0082)	(0.0162)	(0.0085)			(0.0126)
t=10	0.0771	0.4724	0.1679	0.0903	0.0178	0.3121
	(0.0086)	(0.0169)	(0.0101)	(0.0077)	(0.0036)	(0.0121)
t=11	0.0692	0.4422	0.2097	0.2138	0.0404	0.3563
	(0.0088)	(0.0177)	(0.0114)	(0.0117)	(0.0056)	(0.0123)
t=12	0.0679	0.4004	0.2416	0.3216	0.0610	0.3897
	(0.0094)	(0.0188)	(0.0125)	(0.0139)	(0.0072)	(0.0152)
t=13	0.0634	0.3934	0.2612	0.3835	0.0874	0.4637
	(0.0098)	(0.0197)	(0.0134)	(0.0152)	(0.0090)	(0.0144)
t=14	0.0521	0.3633	0.2834	0.4468	0.1163	0.5001
	(0.0102)	(0.0207)	(0.0145)	(0.0164)	(0.0106)	(0.0160)
t=15	0.0498	0.3423	0.2926	0.5206	0.1632	0.5591
	(0.0107)	(0.0224)	(0.0156)	(0.0176)	(0.0129)	(0.0156)
Constant	0.0320	0.1894	0.0126	0.0670	0.0093	13.0628
	(0.0040)	(0.0094)	(0.0029)	(0.0064)	(0.0027)	(0.0078)
Observations	77370	73191	50658	26835	26835	62414
Individuals	6103	6103	6103	4998	4998	6088

Notes: These tables investigate the robustness of our results to the inclusion of graduation round fixed effects. We estimate versions of specification (1), and we report estimates for β_τ (denoted by "Treatment" in table) and for α_τ (which capture baselines among the control group). We also report estimates for the effects on the middle group (relative to the control group) from a simple extension to specification (1) (denoted by "Middle group" in the table). Robust standard errors clustered at the individual level are reported in parentheses.