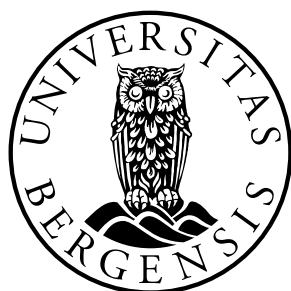


WORKING PAPERS IN ECONOMICS

No. 03/19

JULIE RIISE, BARTON WILLAGE AND
ALEXANDER WILLÉN

CAN FEMALE DOCTORS CURE
THE GENDER STEM GAP?
EVIDENCE FROM RANDOMLY
ASSIGNED GENERAL
PRACTITIONERS



Department of Economics
UNIVERSITY OF BERGEN

Can Female Doctors Cure the Gender STEMM Gap?

Evidence from Randomly Assigned General Practitioners

By JULIE RIISE, BARTON WILLAGE AND ALEXANDER WILLÉN*

Abstract

We use random assignment of general practitioners (GPs) to provide the first evidence on the effects of female role models in childhood on the long-run educational outcomes of girls. We find that girls who are exposed to female GPs in childhood are significantly more likely to sort into traditionally male-dominated education programs in high school, most notably STEMM. These effects persist as females enter college and select majors. We also find strong positive effects on educational performance throughout their academic careers, suggesting that female role models in childhood improve education matches of girls. The effects we identify are significantly larger for high-ability girls with low educated parents, suggesting that female role models may improve intergenerational mobility and narrow the gifted gap for disadvantaged girls.

JEL CODES: I20, I24, J24

KEYWORDS: Role Models, STEMM, Gender Gap

* Riise: Department of Economics, University of Bergen (e-mail: julie.riise@uib.no). Willage: Department of Economics, Louisiana State University (e-mail: bwillage@lsu.edu). Willén: Department of Economics, Norwegian School of Economics (e-mail: alexander.willen@nhh.no). Willén gratefully acknowledges financial support from the Research Council of Norway through its Centers of Excellence Scheme, FAIR project no. 262675. We thank Benjamin Castleman, Chloe East, Ariel Kalil, Katrine Løken, Marianne Page, Bertil Tunggodden, and Lise Vesterlund, as well as several other colleagues and seminar participants for valuable comments and suggestions.

1 Introduction

Females outperform males in educational attainment but remain significantly underrepresented in fields with high financial returns, most notably STEMM (Science, Technology, Engineering, Mathematics, and Medicine). This gender imbalance in educational choice can explain a large part of the gender wage gap (Carrell et al. 2010; Lavy and Sand 2015; Weinberger 1999), and a better understanding of the mechanisms underlying this phenomenon is imperative. As the conventional explanations of discrimination and differences in aptitude have largely been ruled out (e.g. Card and Payne 2017; Ceci et al. 2014; Hyde 2005), increasing attention has been placed on alternative mechanisms, such as the existence and availability of same-gender role models.

Exposure to same-gender role models may affect educational decisions through several channels. Role models may fuel higher aspirations, reduce “stereotype threats” (internalizing other peoples’ stereotypes), and be a source of important information (Breda et al. 2018).¹ A small but growing literature supports the role model hypothesis, showing that females who are exposed to female science teachers and professors in high school and college perform better in school, and are more likely to select into STEMM (Bettinger and Long 2005; Carrell et al. 2010; Dee 2004; Dee 2005; Eble and Hu 2017; Griffith 2014; Hoffman and Oreopoulos 2009; Lim and Meer 2017; Mansour et al. 2018).² However, several questions remain: are these effects temporary or permanent? Do they generalize to non-classroom settings? Can exposure at earlier ages produce similar effects?

The goal of this paper is to move beyond the existing role model literature by addressing each of these three questions, exploiting exogenous variation in general practitioner (GP) assignment for Norwegian children. Norwegian GPs act as gatekeepers to the country’s health care system, and are responsible for examining, diagnosing, and treating patients, as well as for hospital and specialist referrals. When individuals are no longer able to consult with their existing GPs, due to factors such as GP retirement, The Norwegian Health Economics Association reassigns them to new local GPs.³ We use GP reassignments due to GP retirement or other causes outside the patient’s control as a source of exogenous variation to test whether

¹ We define role models in its broadest sense, though we acknowledge that several definitions and types of role models exist. For a concise discussion on this topic, see Chung (2000).

² A few studies have also explored the effect of role models on occupational choice. For example, using conditional random assignment of officers to cadets at West Point, Kofoed and McGovney (2017) show that females and minorities assigned to same-gender and same-race mentors are more likely to choose the same army branches as their mentors.

³ The Norwegian Health Economics Administration (HELFO) is part of the Norwegian Directorate of Health.

childhood exposure to female GPs – a group of successful female STEMM role models – has a positive influence on the educational choice and performance of girls.⁴

Using doctor-patient interactions to test for same-gender role model effects has several benefits. First, the role model interactions we study take place in childhood before any educational decisions have been made, which stands in contrast to the existing literature. Second, by tracing the role model effects throughout the children’s education careers – from compulsory school to college – we are able to examine effect persistency as the children age. Third, the role model interactions we study take place outside the classroom. This is interesting since the classroom effects may be driven by gender differences in teaching practices, the student-teacher interaction may be very different from other types of social interactions, and the purpose of a classroom (to create an environment conducive to the academic development of children) may make children more receptive to role model influences.⁵ Additionally, while it is very difficult to disentangle same-gender role model effects from other potential confounders in the classroom (e.g. gender differences in teaching practices), our setting provides an opportunity to test for these types of cofounders (gender differences in health practices) directly.

To perform our analysis, we leverage rich matched doctor-patient administrative data on all children in Norway that were subject to an exogenous GP reassignment between 2002 and 2011. We link these data to detailed information on the educational performance and choices made by individuals throughout their academic careers, from 10th grade – the earliest age at which students have options regarding subject specialization – through college.

Exposure to a female GP has a large positive effect on both STEMM choice and educational performance among girls. Specifically, assignment to a female GP during childhood leads to an increase in the probability of choosing a STEMM program in high school by 4 percentage points (20 percent relative to the mean), and an increase in high school STEMM GPA by 9 percent of a standard deviation. These effects persist as the girls in our sample enter college: assignment to a female GP during childhood leads to an increase in the probability of

⁴ It should be noted that individuals have the right to independently change the GP they have been assigned twice per year. However, using information on the exact cause of the GP swap, we ignore endogenous swaps initiated by the patients, and use only swaps caused by factors outside the control of the patients (GPs’ moving, reducing their workload or retiring). See Section 2 for more details.

⁵ Breda et al. (2018) address some of these concerns by exposing students in a classroom to a one-hour presentation by female scientists designed to encourage girls to pursue STEM. Porter and Serra (2019) do a similar intervention, focusing on economics instead of STEM. Our paper complement theirs in several ways: (1) we use a natural experiment rather than a field experiment, (2) we look at interactions outside the classroom rather than inside the classroom, and (3) our treatment takes place at a much younger age. The relative contribution of our paper is thus to see whether the results they obtain from field experiments in the classroom generalize to naturally-occurring role model interactions at an earlier age outside the classroom.

choosing a STEMM college major by more than 2 percentage points. This suggests that female role models can close the gender gap in college STEMM choice by almost 20 percent. The effects we identify are large, but fall within the range of the effects that have been identified from shorter information interventions in the classroom (e.g. Breda et al. 2018; Porter and Serra 2019).

We find that the effects of same-gender role models are significantly larger for girls with low-educated mothers – a group of girls that may be relatively less exposed to same-gender STEMM role models in general. This suggests that same-gender role models may facilitate intergenerational occupational mobility, contributing to a long-standing debate on the intergenerational transmission of human capital and how to facilitate upward socioeconomic mobility (e.g. Black, Devereux and Salvanes 2005). Unconditional quantile regressions reveal that it is high ability children with low-educated mothers that are driving our results. This suggests that female role models may help narrow the gifted gap for disadvantaged girls.⁶ This is consistent with evidence suggesting that investments in early childhood, in particular among children from disadvantaged backgrounds, can reduce inequality and help level the playing field (e.g. Carneiro and Heckman 2003; Elango, Garcia, Heckman and Hojman 2015). Similar to existing papers on role model effects in the classroom (e.g. Carell et al. 2010), we do not find an effect of GP gender match on male educational choice or performance.

This paper provides novel insights into how same-gender role models in childhood outside the classroom may shape educational choices and career decisions. These results have important policy implications. Specifically, educational choices in high school and college are likely to have significant effects on future labor market opportunities and career decisions, and may help in closing the gender wage gap (Carrell et al. 2010; Lavy and Sand 2015). Further, female students with same-gender role models are not only selecting into traditionally male-dominated education programs, but they also are performing better in school, suggesting that same-gender role models improve education matches. Thus, intentionally matching girls to female role models (doctors, professors, supervisors, mentors), and establishing/scaling up same-gender mentorship programs, may be effective policy tools for narrowing the gender gap in educational choice and labor market outcomes.

The key assumption underlying our estimation strategy is that the gender of the exogenously provided GP that the government assigns to the individual is orthogonal to other individual characteristics that influence educational outcomes. We show extensive evidence

⁶ The gifted gap refers to the achievement gap between high-ability students from more and less advantaged socioeconomic backgrounds.

that our data are consistent with this assumption. First, we conduct a balance test showing that the gender of the exogenously assigned GP is uncorrelated with a rich set of observable characteristics. Second, we demonstrate that same-gender GP assignments after high school graduation has no impact on educational performance and choice in high school.⁷ Third, we conduct permutation tests in which we randomly assign treatment across individuals in our sample. Our true estimates are meaningfully larger than the distribution of these placebo estimates. Our robustness and falsification tests are inconsistent with the presence of biases due to non-random sorting of children to GPs of a specific gender, and the permutation tests demonstrate that we are not simply picking up random noise. These results support a causal interpretation of our findings.

In addition to a direct role model effect, some of our results could operate through health (same-gender GPs may impact how individuals interact with the health system) and the family (mothers' interactions with the GP might also expose the mother to a female STEM role model, which could impact the mothers' outcomes directly and their children's outcomes indirectly).⁸ While these alternative pathways do not pose a threat to our identification strategy, they would affect the interpretation of our results. Specifically, it would suggest that some of the effects we have identified are operating through a health-based non-role model channel, or through an indirect role-model channel via the impact of the GP on the girl's mother.

We find no evidence in favor of these alternative pathways: assignment to a same-gender GP has no impact on the number of diagnoses, the likelihood of being diagnosed with a mental health issue, the probability of visiting the GP for birth control reasons, or fertility. We also find very little evidence to suggest that the effects are operating through the mother as measured by her education, labor market and health outcomes. This suggests that the effects we identify likely are driven by direct role model influences between the GP and the child. However, we acknowledge that the assignment of a same-gender GP will unavoidably change the conversation at home from talking about the doctor as a "he" to talking about the doctor as a "she". To the extent that this drives some of our results, we consider it part of the treatment of being exposed to a visible role model.

⁷ The idea underlying this placebo test is that the overwhelming majority of individuals above age 20 have completed their high school education and entered either college or the labor market. Thus, GP gender matches after age 20 should not affect an individual's educational performance and choice in compulsory school and high school.

⁸ For example, girls may feel more comfortable with a female GP, and therefore be more likely to disclose health issues and receive treatment. It is also possible that GPs are better able to relate with same-gender patients. While it is unlikely that such potential health effects will impact educational choice, improved health could affect educational performance (Ickovics et al. 2014).

Our paper contributes to the existing literature in several ways. First, this is the first paper to study the effect of role model interactions that take place in childhood. This is an important contribution to the literature as the children in our sample have not made any educational decisions prior to exposure, such that the potential for role models to impact the children’s educational outcomes is much greater. Second, no other paper has examined the effects of same-gender role model interactions outside the classroom on educational choices. Shedding light on this question is important for understanding the extent to which the earlier research on teacher-student interactions generalize to non-classroom settings. Third, by examining the effect of GP gender match on health behavior and parental outcomes, we are able to rule out alternative pathways and better isolate the role model effect. This is an important finding, as one of the main concerns with the existing research on same-gender role models in the classroom has been that it may simply represent differences in teaching practices rather than true role model effects. Finally, our use of Norwegian registry data allows us to trace children throughout their educational careers and explore longer-term effects, something that existing role model studies have not been able to do. The persistency of the role model effect is interesting given the “leaking pipeline” phenomenon - a metaphor for the loss of women in STEMM fields at every step of the career ladder.

The rest of this paper proceeds as follows: In Section 2, we provide detailed information on the health care system as well as the education system in Norway; In Section 3, we introduce our data and empirical method; In Section 4, we present the main results from our analysis; In Section 5, we show results from falsification tests, robustness checks and sensitivity analysis; In Section 6, we conclude and provide policy recommendations.

2 Background

2.1 Health Care System and GP Assignment in Norway

The Norwegian public health care system is based on the principle of universal access, and enrollment is automatic for all Norwegian residents. As in many other European countries, the health care system is a two-part system, with primary care provided by the local municipalities and specialist care provided by larger health regions.⁹

Access to specialist care and hospitals can normally only be achieved through referrals from GPs in the primary care sector (except in emergencies). The GP is therefore the first point of contact for non-emergency and preventive care, and is responsible for initial examination,

⁹ There are currently 422 municipalities and 4 health regions in Norway.

diagnosis and treatment. When GPs deem it necessary, they refer patients to a specialist at the hospital or to private specialists outside the hospital. In other words, the Norwegian GPs act as gatekeepers to the health system of Norway.

Beginning in 2001, the government assigns every individual residing in Norway to a local GP.¹⁰ In most cases, patients interact with their assigned GP every time they engage with the healthcare system.¹¹ Prior to the introduction of this system, individuals were not tied to a specific GP, and had to find a GP every time they needed care. The system was introduced to improve doctor-patient relationships and ensure appropriate use of health care services. As of 2015, there were 4,500 GPs in Norway, and each GP had an average of 1,200 patients on their list.¹² The average GP was 47 years old, and 60% of GPs were male.

When GPs retire, move, or for some other reason decide to terminate/reduce their patient list, the patients on that GP's list are reassigned to a new GP in the municipality. We use GP reassignments due to GP retirement or other causes outside the patient's control as a source of exogenous variation to test whether exposure to female GPs during childhood has a positive influence on the educational choice and performance of females.¹³ It is important to note that children often have the same GP as their mother. However, we find no evidence that our estimated effects of same-gender GP on girls' education are operating through their mothers (as measured by the mothers' education, labor market, and health outcomes).

The majority of GPs are self-employed (less than 5 percent are salaried municipality employees), and municipalities contract with individual GPs to provide services to their residents by assigning them to a list of patients. GPs receive a combination of capitation from the municipalities (approximately 30 percent of their income), fee-for-service from the Norwegian Health Economics Administration (almost 70 percent of their income) and out-of-pocket payments from patients. GP financing is determined nationally through bargaining between the Ministry of Health and the Norwegian Medical Association.

Health care is not entirely free in Norway, and a one-time consultation with a GP is associated with a copayment of NOK 155 (\$18). The maximum annual amount that an individual can spend on health care (including medication) per year is NOK 2400 (\$280) – the

¹⁰ Specifically, The Norwegian Health Economics Administration (part of the Norwegian Directorate of Health), assigns individuals to local GPs on behalf of the government.

¹¹ There are a few exceptions to this, for example if the patient is brought in to the ER. In this case, there is not necessarily any interaction between the patient and the GP.

¹² Patient list refers to the roster of individual patients that the doctor is responsible for.

¹³ It should be noted that individuals have the right to independently change the GP they have been assigned twice per year. However, using information on the exact cause of the GP swap, we ignore endogenous swaps initiated by the patients, and use only swaps caused by factors outside the control of the patients (GPs' moving, reducing their workload or retiring). See Section 2 for more details.

government will cover any expenses exceeding that amount. Children under the age of 16 are exempted from copayments. Thus, there are no direct financial obstacles to GP access among the children included in our analysis.

2.2 The Norwegian Education System

The Norwegian education system consists of 10 years of tuition-free comprehensive compulsory education starting at age 6, with the curriculum set by the central government. During the first seven years, the children are taught a wide range of subjects but receive no official grades. In the last three years, students study a smaller set of subjects and begin to receive official grades. Following the successful completion of compulsory education, each individual has a statutory right to three years of tuition-free high school. The majority of Norwegian children pursue this option, but in contrast to many other countries, high school is not mandatory.

High school in Norway is very different from that in the US, and is comprised of 13 distinct education programs: 5 academic and 8 vocational. The academic programs consist of three years of classroom education, while the vocational programs consist of two years of classroom education followed by 1-2 years of practical training in the field. In the first year of high school, students enroll in one of the 13 education programs. Following the first year, students choose a specialization within their broad education program. Approximately 60 percent of students pursue one of the five academic tracks, while 40 percent choose one of the eight vocational programs.

High school education provides the student with university admission certification, vocational competence, or basic (craft) competence.¹⁴ Students in academic programs receive a university admission certification (and can apply to college), while students in vocational programs do not.¹⁵ Appendix Table A1 provides an overview of the education programs and tracks available at the high school level. To obtain a high school STEM degree with university admission certification, a student has to select the “Specialization in General Studies” education program in the first year of high school, and then specialize in “Natural Science and Mathematics” in the second and third years.

¹⁴ Vocational and craft competence permits individuals to engage in specific professions. University admission certification permits individuals to apply to, and enroll in, college.

¹⁵ However, individuals in vocational programs can take supplemental courses to attain this qualification

Students apply to high school through a centralized online system with the grades from their final year of compulsory education at the age of 15. The application consists of ranking three education programs in the county of residence. If the number of applications exceeds the number of available slots for a given education program, students will be assigned based exclusively on their grades in compulsory school. However, even though admission to specific education programs are based on the grades obtained in compulsory school, Norwegian law ensures that the student will gain admission to one of the three programs on their list.

Higher education in Norway is offered by a range of universities and colleges, the majority of which are tuition-free public institutions. Admission to higher education is coordinated through the Norwegian Universities and Colleges Admission Service. Students apply to specific programs at the different universities, and if the number of applications exceeds the number of available slots for a given program, students will be assigned based on their grades in high school. Admission to university is conditional on having graduated from high school with a university admission certification. In addition, certain programs – most notably STEMM – impose specific course pre-requisites such that only individuals who have taken certain high school courses are eligible for admission. This makes it difficult for individuals with non-STEMM high school tracks to enroll in STEMM university programs, and highlights the importance of the student’s choice of high school track.¹⁶

To summarize: For the purpose of our analysis, it is important to note that (1) the compulsory school GPA is imperative for admission to selective high schools and programs, (2) the choice of an academic high school program is important for obtaining a university admission certification and being able to access higher education, and (3) the GPA as well as the specific courses studied in high school (e.g. STEMM) are crucial for satisfying certain university major pre-requisites and gaining admission to these programs.

3 Data and Method

3.1 Data

Our data come from rich administrative records on the universe of Norwegian residents born between 1967 and 2014. In our analysis, we include cohorts that graduated from high school between 2006 and 2014 (born between 1988 and 1996). We focus on these cohorts because our health visit data begin in 2006, and the last year for which we have detailed

¹⁶ This importance is further exemplified by the correlation between high school STEMM and college STEMM, which is around 0.35 during our analysis period.

education data is 2014. Panel A of Table 1 shows demographic information for these individuals.

The main strength of our data is that we can link individuals across different longitudinal micro-level data registries through unique individual identifiers. This allows us to combine the demographic information displayed in Panel A of Table 1 with detailed information on GP interactions (through the GP and health care use registries), educational choice and performance (through the compulsory, high school and university education registries) and family characteristics (through unique intergenerational identifiers and the population registry).¹⁷

The GP registry provides us with information on the GP of every individual in our sample, for each year since the introduction of the GP list system. We use unique GP identifiers to combine this registry with information from the health care use registry, which provides us with the number of times an individual visited her GP, for each year since 2006. The GP registry also contains information on whether an individual changed GP during the year, and the reason for that change. For the purpose of our study, we are interested in GP swaps that are outside the control of the individual patient, which generates plausibly exogenous variation in the gender of the patients' new GPs.¹⁸ To this end, we focus exclusively on GP swaps that occurred due to the doctor deciding to terminate, or significantly reduce, her patient list. List reductions and terminations are unlikely to be correlated with the gender of the newly assigned GP, something we provide extensive evidence of in Section 5.

Panel B of Table 1 provides summary statistics for the GP visiting behavior of the individuals in our sample, as well as information on the frequency of GP swaps.¹⁹ The data show that the individuals in our sample met with their GP an average of 2 times per year conditional on going to the GP, amounting to approximately 30 times before deciding which high school education program to pursue. This means that the GP intervention we examine is less intense than the classroom experiments where students interact with teachers on a daily basis for a relatively long time (e.g. Carrell et al. 2010; Lim and Meer 2017), but more intense than the information interventions in which successful females have come to classrooms to provide a brief one-time presentation (e.g. Breda et al. 2018; Porter and Serra 2019).

¹⁷ Appendix Table A2 reproduces Panel A of Table 1 for individuals not subject to an exogenous swap before age 15 (individuals omitted from analysis). The table shows that the composition of individuals included in our analysis are very similar to those excluded from the analysis on all observable dimensions.

¹⁸ Individuals have the right to independently change the GP they have been assigned two times per calendar year. However, these GP swaps suffer from endogeneity issue, and we do not use these swaps in our analysis.

¹⁹ Frequency of GP swaps are shown in Appendix Table A3.

Panel B of Table 1 demonstrates that girls are more likely to visit their GPs than boys. While a non-negligible fraction of our sample experienced a second swap during the period we study, very few individuals swapped GPs more than twice (Appendix Table A3). The average number of years that the individual remained with their exogenously-assigned GP is 3, and approximately 60 percent of children who experienced an exogenous swap prior to age 15 still have the same GP at age 15. This is important for the interpretation of our results as we are identifying intent-to-treat effects based on the initial exogenous swap.²⁰

Our education data include detailed information on educational choice as well as academic performance. In terms of educational choice, we begin by using data from the high school registry to examine if GP gender match has an effect on the probability of choosing one of the five university-qualifying academic tracks discussed in Section 2. As discussed in Section 2, students in academic programs receive a university admission certification (and can apply to college), while students in vocational programs do not.²¹ Next, we focus on the primary research question of this paper: Does a same-gender GP role model encourage girls to enroll in STEM fields? We examine this question by estimating the effect of GP gender match on the probability of graduating with a high school STEM degree. As discussed in Section 2, this choice is important for STEM-eligibility at the university level (the correlation between high school STEM and college STEM is 0.35), and in supplemental analyses we use information from the university registry to explore if any potential educational choice effects in high school persist as individuals enter college.²²

After exploring the impact of GP gender match on educational choice, we examine if it also had an impact on educational performance. The performance outcomes we examine include both compulsory and high school GPA. Since we are primarily interested in the impact of same-gender role models on the STEM gender gap, we focus specifically on STEM

²⁰ Appendix Table A4 compares the background characteristics of girls who remain with their exogenously-assigned GPs until age 15 with the background characteristics of girls who do not remain with their exogenously assigned GP until age 15 (either due to an additional exogenous swap, or due to an endogenous swap). With the exception of the mother's years of education (which is economically small but marginally statistically significant) we are unable to reject the null hypothesis that the characteristics of the girls who remain with their exogenously-assigned GP until age 15 are the same as the characteristics of the girls who do not remain with their exogenously-assigned GP until age 15. While this speaks in favor of the generalizability of our results to the broader population, we acknowledge that these groups may differ on other dimensions that we cannot identify with our data.

²¹ However, individuals in vocational programs can take supplemental courses to attain this qualification

²² We restrict the sample to individuals with information about high school. As females are significantly more likely to complete high school, this means that our sample for females will be larger. Our results are robust to relaxing this restriction (available from the authors upon request).

GPA (average GPA across all STEMM courses). In auxiliary analyses, we extend the outcome set and examine if the potential STEMM GPA effects generalize to non-STEMM GPA as well.

The compulsory school STEMM GPA is measured in grade 10, and is used by individuals when applying to high school. The high school STEMM GPA is measured in grade 13, and is used by individuals to apply to college. Conditional on finding effects on educational choice, examining the effect of same-gender GP on academic performance in high school is particularly interesting as selection into STEMM could result either in lower academic performance (due to STEMM being a more academically challenging program) or higher academic performance (through improved program match and motivation). Panel C of Table 1 provides descriptive statistics of educational performance and educational choice variables. On average, girls and boys do equally well in school as measured by GPA, both at the compulsory level as well as at the high school level. However, girls are more likely to have attended university than boys, but much less likely to sort into STEMM. This exemplifies the STEMM gender gap in educational choice discussed in the introduction.

3.2 Method

We exploit variation in GP-patient gender match among individuals born between 1988 and 1996 who experienced an exogenous GP swap (due to patient list termination or reduction) between ages 6 and 15 (the age at which children submit their high school applications).²³ We estimate models of the following form separately for males and females:

$$Y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \varepsilon_i, \quad (1)$$

where Y_i is one of the four core educational outcomes listed above for individual i : the probability of choosing an academic high school track, the probability of graduating with a STEMM high school degree, compulsory school STEMM GPA, and high school STEMM GPA. The variable GP_Match_i is a dichotomous variable taking the value of one if the gender of the exogenously assigned GP matches that of the individual, and zero otherwise. The coefficient of interest, β_1 , thus measures the effect of being randomly assigned to a same-gender GP in childhood, compared to being randomly assigned to an opposite-gender GP in childhood, on Y_i . In the event that the individual experienced several different GPs during the analysis

²³ Data limitations do not allow us to look at swaps that took place prior to the children turning 6. See Section 3.1 for more details.

period, we focus on the gender of the first exogenous GP swap. This protects us against any potential endogeneity concerns after the initial swap, but likely attenuates our estimates. The results from estimation of equation (1) should therefore be interpreted as intent-to-treat effects.²⁴ All standard errors are clustered at the level of the exogenously-assigned GP.²⁵

Equation (1) also includes a set of birth cohort (θ_c), previous doctor (ρ_d), municipality (π_m) and year of swap (τ_t) fixed effects. The municipality fixed effects account for any systematic differences across municipalities correlated both with the probability of being assigned to a same-gender GP and the outcomes of interest. It is worth noting that the average size of a Norwegian municipality – 12,000 individuals – is small and represents a relatively fine classification of residential area. The previous doctor fixed effects subsume any fixed differences between the GPs that individuals were assigned to prior to experiencing a swap (such as the gender of the previous GP), that are correlated both with the probability of being assigned to a same-gender GP and with the outcomes of interest. Thus, if particularly able girls from specific GPs are more likely to be assigned to same-gender GPs, that will be controlled for by ρ_d . The birth cohort fixed effects control for any time-invariant differences between birth cohorts that may be correlated both with GP-gender match and the outcomes. For example, if younger cohorts are more likely to be assigned to same-gender GPs and also are more likely to perform well in school and choose STEMM programs, θ_c will account for that. Finally, the year of swap fixed effects ensures that our results are not driven by systematic differences across years that are related to GP match as well as our outcomes.

Taken together, the fixed effects included in equation (1) protect against any systematic differences across birth cohorts, previous doctors, municipalities and year of swap, that may be correlated with the probability of being assigned to a same-gender GP and with the educational outcomes we examine.²⁶ A factor that would bias our results therefore has to be correlated with the probability of being assigned to a GP of the same gender and the individual’s educational outcomes, but uncorrelated with the individual’s previous GP, municipality, birth year and year

²⁴ As shown in Appendix Table A3, 90 percent of individuals in our sample do not experience more than one exogenous swap during the analysis period. However, some of these individuals have swaps for other reasons, such as requesting a new GP or moving. We do not use these other swaps due to endogeneity concerns, but their existence means that our results should be interpreted as ITT effects.

²⁵ As a robustness check, we also estimate Equation (1) using only those GP swaps that were caused by list terminations. Results from this exercise are shown in Online Appendix Table A7. This adjustment leads to larger point estimates and smaller standard errors.

²⁶ An alternative estimation strategy would be to use family fixed effects models. However, there are very few instances in which there are two sisters under the age of 15 who are assigned to different gender GPs, and we are therefore unable to utilize this identification strategy.

of swap. As shown in Section 5, such factors would also have to be uncorrelated with a rich set of individual demographics.

Conditional on the fixed effects in equation (1), the variation we exploit comes from differences in the GP-patient gender match among those who experienced an exogenous GP swap prior to submitting their high school applications. The assumption underlying the identification of the parameter of interest, β_1 , is that the GP-patient gender match is orthogonal to other individual characteristics that influence educational choice and performance. In Section 5, we show extensive evidence that our data are consistent with this assumption. First, we conduct a balance test, which shows that the gender of the newly-assigned GP is uncorrelated with a rich set of observable characteristics. This suggests that randomization was successful. Second, we demonstrate that exogenous GP swaps that occur after high school graduation have no impact on high school educational choice or performance. Third, we conduct a series of permutation tests in which we randomly assign treatment across individuals in our sample (keeping the distribution of treatment constant). Our true estimates are meaningfully larger than the distribution of these placebo estimates. Our robustness and falsification tests are inconsistent with the presence of biases due to non-random sorting of children to GPs of a specific gender, and the permutation tests demonstrate that we are not simply picking up random noise. This provides support for a causal interpretation of our results.

4 Results

4.1 Baseline Results

Table 2 presents the effect of same-gender GP on female educational choice and performance using the four core outcomes defined above: the probability of choosing an academic high school track, the probability of graduating with a STEMM high school degree, compulsory school STEMM GPA, and high school STEMM GPA. Each column in Table 2 comes from a separate estimation of equation (1). The point estimates represent the average effect of assignment to a same-gender GP using the first exogenous GP swap experienced by the person, and should be interpreted as the intent-to-treat effect of same-gender role models.

Column 1 shows that girls who are exposed to a female GP during childhood are 5 percentage points more likely to choose an academic education program in high school (7 percent relative to the mean). As discussed in Section 2, the choice of an academic program in high school ensures immediate access to the higher education system of Norway and greatly

improves an individual's chance of obtaining a university degree.²⁷ Column 2 explores if GP gender affects the probability of graduating from high school with a STEMM degree (recall that STEMM specialization can be chosen first in the second year of high school). The point estimate closely mirrors that in Column 1, and shows that females who are assigned to female GPs in childhood are 4 percentage points more likely to obtain a STEMM degree in high school (20 percent relative to the mean). The results in Columns 1 and 2 thus suggest that exposure to same-gender GPs leads girls to sort into - and graduate from - educational programs that traditionally have been underrepresented by females and are associated with larger financial returns.

The remaining two columns of Table 2 examine if exogenous assignment to female GPs also influences the educational performance of girls. Column 3 present results for STEMM GPA at the compulsory level, and provides clear evidence of a positive effect of GP match on performance: the point estimate in Column 3 indicates that a girl who is exposed to a female GP experiences a 0.047 unit increase in STEMM GPA. This effect is 9 percent of a standard deviation (Table 1). As discussed in Section 2, the compulsory school GPA is imperative for admission to selective high school programs, and this result is thus consistent with the idea that role model exposure motivates individuals to work harder in compulsory school to get accepted to the more selective high school programs.

We also study if the performance effects identified in compulsory school persist as girls enter high school. Examining the effect of same-gender GP on academic performance in high school is particularly interesting as selection into STEMM could result either in lower academic performance (due to STEMM being a more academically challenging program) or higher academic performance (through improved program match and motivation). Column 4 of Table 2 provides clear evidence of a positive effect of GP gender match on performance in high school. The point estimate is very similar to the STEMM GPA effect in compulsory school. This suggests that the effect of same-gender GP on educational choice does not induce worse academic performance. Rather, it leads to improved educational achievement, which could be due to improved education matching and enhanced motivation.

²⁷ To better understand where these students are coming from, Appendix Table A5 shows results from estimation of equation (1) using non-STEMM academic track and non-academic track as dependent variables. The result from this exercise suggests that girls are pulled into the STEMM academic track from both of these groups rather than from one specific group, though more than 2/3 are coming from academic non-STEMM programs. Due to power issues, we are unable to look separately at each of the 12 high school programs.

To examine if the same-gender GP effects for girls identified in Table 2 extend to boys as well, Table 3 provides results obtained from estimating equation (1) on our boy sample for each of our four core outcomes. Looking across the table, the point estimates are small (all of them are less than half the size of the female estimates) and none of them are statistically significant, suggesting that boys are not affected by being exposed to same-gender GPs. The results shown in Table 3 are consistent with previous literature on this topic in a classroom setting (e.g. Carrell et al. 2010). While the lack of significant same-gender role model effects among boys could be due to gender differences in receptiveness to role model influences, it could also be because the channels through which roles models are expected to operate (eradication of stereotype threats and provision of information) are less important for boys.²⁸ Specifically, the percentage of boys in STEMM fields is substantial, and there is a disproportionate availability of public figure male STEMM role models. It is important to note that the results in Table 3 also suggest that there is no across-the-board effect of female GPs on children: if female GPs had a positive influence on male children we would expect to find a negative effect of same-gender GP assignment among boys.

Taken together, the results in Tables 2 and 3 suggest that same-gender GPs have a positive influence on females' educational performance and choice but not on males' educational performance and choice, and that these effects are not driven by an across-the-board positive effect of female GPs on children. It is worth noting that the male sample is smaller than the female sample, such that our power to detect effects among boys is smaller. However, the differences in point estimates are substantial, while the standard errors are very similar, making it unlikely that the lack of a significant effect among boys is due to sample size.

The results from this analysis provide important insights into how gender role model interactions in childhood may shape educational choices and career decisions of girls, and have important policy implications. Specifically, educational choices are likely to have significant effects on future labor market opportunities and career decisions, and may help in closing the gender wage gap (Carrell et al. 2010; Lavy and Sand 2015). In terms of policy implications, an interesting question with respect to our results is whether the effects are driven simply by increased awareness of the existence of females within the STEMM fields, or due to repeated interactions with a female in this field. To provide some suggestive evidence on this question, we estimate equation (1) separately for those that remain with their GPs until they send in their high school applications at age 15 and for those that do not remain with their GPs. Results from

²⁸ Another potential explanation could be that boys visit their GPs less often. However, the descriptive statistics in Table 1 suggest that the gender difference in visiting behavior is small, making this an unlikely explanation.

this exercise are highly speculative as the decision to not remain with their initially exogenously assigned GP may be endogenous, and should therefore be interpreted with care. The results are shown in Appendix Table A6, and suggest that the identified effects are predominantly driven by individuals who remain with their GPs. Conditional on the endogeneity issue above, this suggests that repeated interactions with same-sex GPs, rather than simply increased awareness of females pursuing STEMM careers, drive the effects.

Due to the lack of economically meaningful and statistically significant results among boys, we focus on same-gender GP matches on the educational choice and performance of girls in the remainder of the paper.

4.2 Heterogeneity Analysis

i. Mother education level

The effects of same-gender role models likely differ across individuals depending on the availability of same-gender role models in their families and surroundings. For example, a female with a highly educated mother may benefit and learn less from exposure to a female GP than a female with a low educated mother, since many of the channels through which role models may operate (providing higher aspirations, reducing stereotype threats and being a source of important information) have already been filled by the mother.

To examine this possibility, we estimate equation (1) for our core outcomes stratified by whether the female's mother has more or less than a college degree.²⁹ The results from this exercise are shown in Table 4. The results demonstrate that the role model effects identified in Section 4.1 are driven by daughters of low educated mothers. These results are consistent with our priors, and highlight that same-gender role models are not only important for closing the gender STEMM gap, but also for closing the within-gender socioeconomic STEMM gap. This suggests that same-gender role models could be an important tool for improving intergenerational occupational mobility, contributing to a long-standing debate on the intergenerational transmission of human capital and how to facilitate upward socioeconomic mobility (e.g. Black, Devereux and Salvanes 2005).

²⁹ Unfortunately, we lack sufficient power to conduct a similar exercise based on whether the mother has a STEM education or not.

ii. Distributional Effects

Equation (1) estimates the average treatment effect of assignment to a same-gender GP. However, looking only at the mean effect likely misses important heterogeneity in effect size across the ability distribution. Specifically, exposure to same-gender role models are likely to incentivize students at the right-tail of the ability distribution who satisfy – or are close to satisfying – the requirements for choosing STEMM programs, while it may not be sufficient for incentivizing students at the left-tail of the distribution. To explore this further, Table 5 show the results from estimating unconditional quantile regressions, using the methodology of Firpo et al. (2009). With respect to compulsory school STEMM GPA, the results in Panel A of Table 5 demonstrate that the effect of same-gender GP match on educational performance loads on individuals in the right tail of the ability distribution. The results for high school STEMM GPA largely mirror the results for compulsory school STEMM GPA, though quantiles 7 and 8 are slightly noisier. Taken together, the results in Table 5 are consistent with our priors and with the previous literature on the effect of same-gender teachers in the classroom (e.g. Carrell, Page and West 2010). Combined with the results in Section 4.2 (i), Table 5 suggest that female role models may not only improve intergenerational mobility, but that they may also help narrow the gifted gap for disadvantaged girls.

iii. Age-at-swap

Equation (1) estimates the average treatment effect of assignment to a same-gender GP. This may miss important treatment heterogeneity across age-at-swap. Specifically, it is unlikely that the effect of being exogenously assigned to a same-gender GP at age 6 is the same as the effect of being exogenously assigned to a same-gender GP at age 14. First, children exposed at a younger age may be differentially affected compared to children exposed at a later age since the GP-patient interaction likely is different. Second, children that experience a swap at an earlier age may be exposed to the gender of the new GP for a longer period of time.

To examine this age heterogeneity, we estimate a modified and more flexible version of equation (1) in which we replace GP_Match_i with a set of binary match variables based on the age at which the child experienced the swap. Specifically, we estimate models of the following form:

$$Y_i = \alpha + \sum_{a=1}^3 [\partial_a (ExogSwap_{ia} \times GP_Match_i)] + \tau_t + \pi_m + \theta_c + \rho_d + \varepsilon_i, \quad (2)$$

where $ExogSwap_{ia}$ takes the value of 1 if individual i experienced an exogenous swap at age a , and zero otherwise, grouping individuals into three age bins: 6-9, 10-12, and 13-15.³⁰ We pick these age groups because they represent distinct stages of the children’s educational careers. Specifically, age 6 through 9 correspond to lower primary education (*Småskoletrinnet*), age 10 through 12 correspond to upper primary education (*mellomtrinnet*), and age 13 through 15 correspond to lower secondary education (*ungdomsskolen*).

The term $(ExogSwap_{ia} \times GP_{Match_{ia}})$ represents the interaction of age-at-swap (a) and gender match. This interaction term takes the value of 1 if the person experienced an exogenous swap to a same gender GP and the swap happens at that age, and 0 if the person swapped at a different age or did not experience a GP gender match. The results from this exercise are shown in Figure 1 for each of our four core outcomes: academic high school program, STEMM high school degree, compulsory school STEMM GPA, and high school STEMM GPA. Each point in the figure corresponds to the estimated effect of same-gender GP on girls who experienced the GP swap at the age range indicated on the x-axis.

The results in Figure 1 implies that girls who are exogenously assigned a female GP at an earlier age experience larger effects than girls exposed at a later age. This suggestive heterogeneity is largest when examining STEMM GPA in compulsory school, but is also noticeable when looking at STEMM GPA in high school and STEMM high school degree. With respect to academic high school program, there is no indication of such age-at-swap heterogeneity. It is important to note the suggestive effect heterogeneity documented in Figure 1 can be driven by two different factors: (1) girls who experience an exogenous swap at an earlier age may be exposed to the GP for a longer time period, and (2) girls who are exposed at an earlier age might be more receptive to role model influences. Unfortunately, it is not possible to disentangle which of these mechanisms are driving the suggestive evidence in Figure 1.

4.3 Additional Outcomes

i. Non-STEMM GPA

The results in Table 2 demonstrate that same-gender role models positively affect the STEMM GPA of female students, both in compulsory school and high school. As discussed in Section 3, these variables represent the average GPA across all STEMM courses. It is not clear that these effects lead to an improvement in the overall GPA, because the improved STEMM GPA could be due to students spending less time on other subjects in order to perform better in

³⁰ Results using alternative groupings are available from the authors upon request.

STEMM subjects. To examine this in detail, Column 1 of Table 6 shows the result from estimating equation (1) using the non-STEMM GPA as the dependent variable.

Column 1 of Table 6 shows that the positive educational performance effect documented in Table 2 is not restricted to STEMM subjects – females exposed to same-gender role models perform better in non-STEMM subjects as well, though the effects are smaller in magnitude. This is an important finding, because admission both to high school and university is dependent not only on performance in STEMM courses, but also on overall GPA. This result is consistent with the idea that role model exposure motivates individuals to work harder—irrespective of subject - to get accepted to more selective programs.

ii. College Enrollment and College Major Choice

In Table 6, we also examine if the positive effects on educational choice in compulsory school and high school persist as individuals enter college and choose majors. Looking at these outcomes is particularly important for understanding the persistence of the role model effects, something that earlier studies have not been able to explore. We focus on two outcomes: the probability of attending university (Column 2) and the probability of choosing a STEMM major (Column 3).

Column 2 of Table 6 shows that there is no effect on enrolling in university, suggesting that same-gender role models do not have an impact on the extensive margin of females’ decision to pursue higher education. This result is consistent with the quantile regression results discussed above, which suggest that same-gender role models may be effective in motivating students on the margin, but ineffective in motivating those at the bottom of the ability distribution.

While there is no effect of same-gender role models on females’ probability of enrolling in college, column 3 of Table 6 reveals that there is a positive, economically meaningful and statistically significant effect of same-gender role models on a female’s likelihood of choosing a STEMM major at university. In terms of magnitude, same-gender role models cause a 2 percentage point increase in the probability of choosing a STEMM major for female students. Given the gender-specific enrollment rates in STEMM majors at university shown in Table 1, the 2 percentage point increase in the probability of choosing a STEMM major at university among females suggests that female role models can close the gender gap in college STEMM choice by almost 20 percent.³¹ Thus, exposure to same-gender role models do not affect an

³¹ $Gap\ Close = \left(\frac{Effect\ size\ for\ girls}{(mean\ boys - mean\ girls)} \right) * 100 = \left(\frac{0.02}{(0.185 - 0.077)} \right) * 100 = 18.5$

individual's decision to attend university, but conditional on wanting to attend university, it does affect the choice of major.³²

5 Balance, Falsification, Permutation and Alternative Pathways

5.1 Balance Test

The key assumption underlying our estimation strategy is that - conditional on birth cohort, previous doctor, municipality, and year of swap - the gender of the exogenously-assigned GP is orthogonal to other individual characteristics that influence the educational outcomes we study. To examine this assumption, we conduct a balance test in which we regress a rich set of pre-determined observable characteristics on the gender match of the exogenously-assigned GP using equation (1). The characteristics we explore are birth order, number of siblings, whether the child is born in Norway, the educational level of the mother, the age of the mother, marital status of the mother, the income of the mother, and the labor force participation status of the mother. If the GP gender is orthogonal to individual characteristics that may influence future educational decisions and performance, this exercise should return a series of economically small and not statistically significant point estimates. The results from this exercise are presented in Table 7.

The results in Table 7 provide evidence in favor of our identifying assumption, showing small and not statistically significant point estimates for each of the background variables included in this test. These results are inconsistent with non-random sorting across GP gender, and support a causal interpretation of our results.

5.2 Falsification Test

In addition to the balancing test discussed above, we perform a placebo test in which we examine the effect of exogenously generated GP-matches at ages 21 through 24 on educational outcomes at the compulsory and high school level. The idea underlying this placebo test is that the overwhelming majority of individuals above age 20 have completed their high school education and entered either college or the labor market. Thus, GP gender matches after age 20 should not affect an individual's educational performance and choice in compulsory school and high school. If this exercise returns statistically significant and economically meaningful point estimates, that suggests that individuals subject to exogenous GP swaps were

³² As can be seen in Table 5, only a small fraction of females pursue STEMM degrees at university (less than 10 percent). Unfortunately, there is therefore not enough power to study the effect of enrollment in individual STEMM programs.

non-randomly sorted across GPs of different gender, and that our results cannot be used to identify causal role model effects.

The results from this analysis are shown in Table 8. Looking across the columns in Table 8, all point estimates are much smaller, or in the wrong direction, and none of them are statistically significant. This suggests that GP matches at ages 21 through 24 do not affect educational performance and choice in compulsory school and high school. These results are inconsistent with the presence of biases due to non-random sorting of children to GPs of a specific gender, and provide support for a causal interpretation of our results.

5.3 Permutation Tests

The balance test shows that the GP-patient gender matches we use are uncorrelated with a range of predetermined individual characteristics that potentially could affect the educational outcomes of interest. The falsification test provides additional support against non-random sorting of individuals to same-gender GPs. Another concern is that we are simply picking up random noise, and that our results are independent of treatment assignment. To investigate this potential issue, we perform a series of permutation tests in which we randomly reassign treatment to individuals and re-estimate equation (1) using this re-randomization. We perform the permutations 300 times for our four core outcomes and examine where the effects identified in Table 2 fall relative to these 300 alternative placebo estimates.³³ If the results in Table 2 represent true effects of same-gender GP assignment, the estimates in Table 2 should be larger than the vast majority of these simulations. If this were not the case, the placebo test would raise concerns that our estimates are due to random noise.

Figure 2 shows histograms of the distribution of effects obtained through the 300 permutation simulations for each of our four core outcomes: academic high school track, high school STEM degree, compulsory STEM GPA, and high school STEM GPA. The vertical lines in these figures indicate where the “true” estimates fall compared to the alternative placebo estimates. As shown in Figure 2, none of the simulated estimates exceeds the actual estimates for any of the outcomes. We can therefore reject the null hypothesis that any combination of treatment would generate the same magnitude of treatment effects as that displayed in Table 2.

³³ Our choice of 300 permutations is arbitrary, but since the distribution of placebo estimates is far from the baseline estimates, we feel comfortable that additional permutations would not alter the results displayed in the figure.

5.4 Ruling Out Alternative Pathways

In addition to a direct role model mechanism between the child and the GP, some of the identified effects may operate through the children's health and parents. First, the GP may affect how an individual interacts with the health system. For example, girls may feel more comfortable with a female GP, and therefore be more likely to disclose health issues and receive treatment. It is also possible that GPs are better able to relate with same-gender patients. While it is unlikely that this would affect an individual's decision to pursue a STEMM degree, improved health could affect educational performance, and may therefore provide an additional pathway through which our effects operate. Second, it is also possible that part of the effects we identify operate through the mothers' exposure to same-gender GPs. Specifically, mothers' interactions with the GP might also expose them to female STEMM role models, which could impact their outcomes directly and their children's outcomes indirectly (through improved resources at home or through access to better information via the mother).³⁴

It is important to note that the alternative health and family mechanisms do not pose a threat to our identification strategy, and would not invalidate the findings that assignment to same-gender GPs in childhood positively affects girls educational performance in compulsory school and high school, makes them more likely to pursue academic high school programs and obtain high school STEMM degrees, and increases their likelihood of enrolling in STEMM programs at university. However, the existence of these pathways would affect the interpretation of our results. Specifically, it would suggest that some of the effects we have identified are operating through a health-based non-role model channel, or through an indirect role-model channel via the impact of the GP on the girl's mother.

To examine the existence of a potential health-based non-role model channel, we estimate equation (1) using a battery of health-related outcomes measured at age 15: GP visits, number of diagnoses, mental health diagnoses, birth control visits, fertility (admission for delivery), and probability of remaining with the assigned GP. To examine if any of our effects are operating through indirect role model influences via the mother's interaction with the GP, we estimate equation (1) using a number of education, labor market, and health outcomes of the mother

³⁴ Not only because mothers are likely to accompany their children to the GP, but also because children are generally assigned to the same GP as their mother (their legal custodian if they do not live with their mother), such that a GP swap for the child also entails a GP swap for the mother.

when the child is 15. The education and labor market outcomes we look at are years of education, total income (in logarithmic form), and labor force participation.³⁵

Panel A (health-based non-role model channel) and Panel B (indirect role-model effect through the mother) of Table 9 show results from these exercises. With respect to the potential health-based non-role model channel, the results in Panel A show that assignment to a same-gender GP does not have an impact on any of the health-related outcomes of the child. The one exception is fertility, which displays a statistically significant effect. However, the point estimate of 0.000 is not economically meaningful. With respect to the potential indirect role-model effect through the mother, the results in Panel B show that assignment to a same-gender GP does not affect the mother’s education, labor market, and health outcomes. Taken together, the results in Table 9 suggest that the effects we identify in Section 4 are unlikely to operate through a health-based non-role model channel, and they are unlikely operating through an indirect role model channel via the mother’s interaction with the GP. However, as noted in the introduction, we acknowledge that the assignment of a same-gender GP will unavoidably change the conversation at home from talking about the doctor as a “he” to talking about the doctor as a “she”. To the extent that this drives some of our results, we consider it part of the treatment of being exposed to a visible role model.

The lack of a health-based non-role model channel is a particularly important finding, as one of the main concerns with the existing research on same-gender role models in the classroom has been that it may represent differences in teaching practices rather than true role model effects. While we acknowledge that there are additional unobserved health outcomes that we are unable to examine, it is unlikely that they are driving any of the effects as they would have to be uncorrelated with all the outcomes in Table 9, not subsumed by the four one-dimensional fixed effects included in equation (1), but correlated both with the probability of being assigned a same-gender GP and the outcomes of interest.

6 Discussion and Conclusion

Females outperform males in educational attainment but remain significantly underrepresented in fields with high returns, most notably STEM. This gender imbalance in educational choice can explain a large part of the gender wage gap, and a better understanding of how to narrow this gap is imperative for attaining gender wage parity.

³⁵ We focus on the mother’s education and labor market outcomes when the child is 15 as this is when the daughter submits her high school application. If GP match has an impact on the mother’s education and labor market outcomes that indirectly affects the daughter, this is the age where one would expect to find an effect.

We use random assignment of children to general practitioners to examine if female role models can be used to eliminate some of the observed gender gap in educational choice. We test if girls who are randomly assigned and exposed to female GPs - a group of successful female STEM role models - perform better in school and are more likely to pursue STEM degrees, compared to girls assigned to male GPs. This is the first paper to study the effects of female role models in childhood on the long-run educational outcomes of girls. It is also the first paper to explore the effects of same-gender role model interactions outside the classroom.

We find that exposure to a female GP has a statistically significant and economically meaningful positive effect on the probability that girls pursue academic high school programs, graduate with STEM degrees from high school, and choose STEM majors at university. A back-of-the-envelope calculation suggests that female role models can close the gender gap in college STEM choice by up to 20 percent. These effects fall within range of what has been identified in other settings (e.g. Breda et al. 2018; Porter and Serra 2019). The persistency of the role model effect is interesting given the “leaking pipeline” phenomenon - a metaphor for the loss of women in STEM fields at every step of the career ladder.

We also document significant improvements in educational achievement among girls who are assigned to a female GP, both at the compulsory level and at the high school level. The fact that enrollment in more difficult education programs is accompanied by improved educational performance suggest that selection into STEM programs does not lead to a deterioration of educational achievement, and that same-gender role models improve education matches. We do not find any effects of same-gender assignment on male education choices or performance.

The effects we identify are larger for girls with low-educated mothers, suggesting that same-gender role models may facilitate intergenerational occupational mobility, contributing to a long-standing debate on the intergenerational transmission of human capital and how to facilitate upward socioeconomic mobility (e.g. Black, Devereux and Salvanes 2005). Performing unconditional quantile regressions reveal that it is high ability children with low-educated mothers that are driving our results. This suggests that female role models may help narrow the gifted gap for disadvantaged girls.

By tracing the individuals in our analysis sample through a range of medical registries in Norway we directly explore if our results operate through improved health, and by linking children to their parents through unique intergenerational family identifiers we explicitly examine if some of the estimate effects are operating through the mother. We find no evidence in favor of these alternative pathways: assignment to a same-gender GP has no impact on the

number of diagnoses, the likelihood of receiving a mental health diagnosis, the probability of visiting the GP for birth control reasons, or fertility. Assignment to same-gender GP also has no impact on the mother's education and labor market outcomes. This suggests that the effects we identify likely are driven by direct role model influences between the GP and the child. The lack of a health-based non-role model channel is interesting as one of the main concerns with the existing research on same-gender role models in the classroom has been that it may represent differences in teaching practices rather than true role model effects.

Our results demonstrate that female role models can be successful in not only reducing the gender STEMM gap in educational choice in both high school and college, but also in raising the educational performance of females. These effects have important policy implications. Specifically, they imply that intentionally matching girls to female role models (doctors, professors, supervisors, mentors), and scaling up existing same-gender mentorship programs, may be effective policy tools for narrowing the gender gap in educational choice and labor market outcomes.

References

- Bettinger, E., and B. Long** (2005). "Do faculty serve as role models? The impact of instructor gender on female students" *American Economic Review* 95(2): pp. 152-157
- Bhashkar, M.** (2005). "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics* 87 (2): pp. 235-255
- Black, S., P. Devereux, and K. Salvanes** (2005). "Why the apple doesn't fall far: understanding intergenerational transmission of human capital" *American Economic Review* 95(1): pp. 437-449
- Blau, F., and L. Kahn** (2017). "The gender wage gap: extent, trends, and explanations" *Journal of Economic Literature* 55(3): pp. 789-865
- Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre** (2018). "Can female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools" *PSE Working papers No. 2018-06*
- Card, D., and A. Payne** (2017). "High school choices and the gender gap in STEM" *NBER Working Paper No. 23769*
- Carneiro P., and J. Heckman** (2003). "Human Capital Policy." In Heckman J., Krueger A. (eds.), *Inequality in America: What Role for Human Capital Policy?* (MIT Press)
- Carrell, S., M. Page and J. West** (2010). "Sex and science: how professor gender perpetuates the gender gap" *The Quarterly Journal of Economics* 125 (3): pp. 1101-1144
- Ceci, S., D. Ginther, S. Kahn and W. Williams** (2014). "Women in academic science: a changing landscape" *Psychological Science in the Public Interest* 15(3): pp. 75-141
- Chetty, R., and N. Hendren** (2018). "The effects of neighborhoods on intergenerational mobility I: childhood exposure effects" *Quarterly Journal of Economics* 133(3): pp. 1107-1162
- Chetty, R., N. Hendren, and L. Katz** (2015). "The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment" Working Paper NBER 21156
- Chung, K.** (2000). "Role models and arguments for affirmative action" *American Economic Review* 90(3): pp. 640-648

- Dee, T.** (2005). “A teacher like me: does race, ethnicity, or gender matter?” *American Economic Review* 95(2): pp. 158-165
- Dee, T.** (2004). “Teachers, race, and student achievement in a randomized experiment” *The Review of Economics and Statistics* 86(1): pp. 195-210
- Eble, A., and F. Hu** (2017). “Stereotypes, role models, and the formation of beliefs” *Mimeo*.
- Elango, S., and J. Garcia, J. Heckman and A. Hojman** (2015). “Early Childhood Education.” In Mofitt R. (editor), *Economics of Means-Tested Transfer program in the United States* (University of Chicago Press)
- Firpo, S., N. Fortin, and T. Lemieux** (2009). “Unconditional quantile regressions” *Econometrica* 77 (3): pp. 953-973
- Griffith, A.** (2014). “Faculty gender in the college classroom: does it matter for achievement and major choice? *Southern Economic Journal* 49(3): pp. 730-749
- Hill, A., D. Jones and L. Woodworth** (2018). “A doctor like me: Physician-patient race-match and patient outcomes” *Mimeo*
- Hoffmann, F., and P. Oreopoulos** (2009). “A professor like me: the influence of instructor gender on college achievement” *Journal of Human Resources* 44(2): pp. 479-494
- Hyde, J.** (2005). “The gender similarities hypothesis” *American Psychologist* 60(6): pp. 581-592
- Ickovics, J., A. Carroll-Scott, S. Peters, M. Schwartz, K. Gilstad-Hayden and C. McCaslin** (2014). “Health and academic achievement: cumulative effects of health assets on standardized test scores among urban youth in the United States” *Journal of School Health* 84(1): pp. 40-48
- Kofoed, M., and E. McGovney** (2017). “The effect of same-gender and same-race role models on occupational choice: evidence from randomly assigned mentors at West Point” *Journal of Human Resources*, 0416–7838r1
- Lavy, V., and E. Sand** (2015). “On the origins of gender human capital gaps: short and long term consequences of teachers’ stereotypical biases” *NBER Working Paper No. 20909*
- Lim, J., and J. Meer** (2017). “The impact of teacher-student gender matches: random assignment evidence from South Korea” *NBER Working Paper No. 21407*
- Mansour, H., D. Rees, B. Rintala, N. Wozny** (2018). “The effects of professor gender on the post-graduation outcome of female students” *IZA DP No. 11820*

NiN (2017). *Primary Doctors*. Accessed via <http://www.nyinorge.no/en/Ny-i-Norge-velgsprak/New-in-Norway/Health/Health-services/Primary-doctor/>

WEF (2018). *The Global Gender Gap Report 2018*, accessible through <https://www.weforum.org/reports/the-global-gender-gap-report-2018>

Weinberger, C. (1999). "Mathematical college majors and the gender gap in wages"
Industrial Relations 38(3): pp. 407-413

Table 1: Descriptive Statistics of Individuals in Sample

	<i>Girls</i>		<i>Boys</i>	
	Mean	St.D	Mean	St.D
<i>Panel A: Family Composition</i>				
Birth order	1.864	0.983	1.850	0.985
Siblings	1.697	1.064	1.707	1.052
Born in Norway	0.861	0.346	0.851	0.356
Mother age (birth)	28.987	5.008	29.414	4.822
Mother marital status (birth)	0.567	0.496	0.595	0.491
Parental years of education (birth)	13.775	2.343	14.121	2.309
Both parents born abroad	0.058	0.234	0.068	0.252
<i>Panel B: GP Visiting Behavior</i>				
GP visits age 15	1.165	1.630	0.898	1.525
GP visits age 15 conditional on visiting	2.130	1.680	1.915	1.730
Still with ex. GP at age 15	0.613	0.487	0.616	0.486
<i>Panel C: Educational Performance and Choice</i>				
Compulsory school GPA	4.591	0.573	4.430	0.595
Compulsory school STEMM GPA	4.303	0.889	4.351	0.914
High school GPA	4.314	0.654	4.185	0.681
High school STEMM GPA	3.993	0.890	3.940	0.905
Academic track year 1	0.734	0.440	0.807	0.395
High school STEMM degree	0.195	0.396	0.303	0.460
Ever College	0.751	0.433	0.678	0.467
Ever College STEMM	0.077	0.266	0.185	0.388

Notes: Authors' calculations based on detailed administrative records. Sample includes all boys and girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Number of observations is approximately 8500 girls and 5500 boys.

Table 2: Effect of same-gender GP on educational choice and performance of girls

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Same-gender GP	0.052*** (0.017)	0.039** (0.018)	0.084** (0.039)	0.109*** (0.039)
Mean	0.736	0.194	4.297	3.993
Observations	8679	8424	8617	8258

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 3: Effect of same-gender GP on educational choice and performance of boys

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Same-gender GP	0.011 (0.020)	0.002 (0.025)	-0.038 (0.047)	0.010 (0.049)
Mean	0.807	0.600	4.350	3.940
Observations	5514	5338	5475	5253

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all boys born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 4: Effect of same-gender GP on females, by mother’s education

	Compulsory school STEMM GPA	High school STEMM GPA	High school academic track	High school STEMM degree
<i>Panel A: Mother college or more</i>				
Same-gender GP	0.021 (0.091)	-0.014 (0.098)	0.029 (0.033)	0.035 (0.050)
Mean	4.632	4.243	0.851	0.291
Observations	2339	2337	2341	2341
<i>Panel B: Mother less than college</i>				
Same-gender GP	0.101* (0.054)	0.137** (0.053)	0.070*** (0.026)	0.093*** (0.024)
Mean	4.212	3.874	0.675	0.152
Observations	4637	4652	4654	4643

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Panel A includes all girls who were subject to at least one exogenous GP swap prior to age 15 and has a mother with at least a college education. Panel B includes all girls who were subject to at least one exogenous GP swap before age 15 and has a mother with less than a college education. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 5: The effect of same-gender GP on STEM GPA - quantile effects

	Quantile 1	Quantile 2	Quantile 3	Quantile 4	Quantile 5	Quantile 6	Quantile 7	Quantile 8	Quantile 9
<i>Panel A: Compulsory school STEM GPA</i>									
Same-gender GP	0.006 (0.030)	-0.000 (0.044)	0.011 (0.032)	0.011 (0.032)	0.082* (0.043)	0.082* (0.043)	0.067* (0.035)	0.067* (0.035)	0.149* (0.070)
<i>Panel B: High school STEM GPA</i>									
Same-gender GP	-0.027 (0.084)	0.060 (0.066)	0.076 (0.056)	0.145** (0.062)	0.135** (0.052)	0.156** (0.063)	0.096 (0.058)	0.045 (0.067)	0.140** (0.068)

Notes: Authors' estimation of equation (1) as described in text using the unconditional quantil regression method discussed in Firpo, Fortin and Lemieux (2009). Regressions include municipality, year of swap, birth year and previous GP fixed effects. Standard errors are clustered at level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 6: Effect of same-gender GP, additional outcomes

	Compulsory school non-STEMM GPA	High school non-STEMM GPA	College enrollment	College STEMM enrollment
Same-gender GP	0.040* (0.023)	0.067** (0.028)	0.005 (0.017)	0.022* (0.013)
Mean	4.297	4.384	0.754	0.077
Observations	8617	8678	8680	8680

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Panel A includes all girls who were subject to at least one exogenous GP swap prior to age 15 and has a mother with at least a college education. Panel B includes all girls who were subject to at least one exogenous GP swap before age 15 and has a mother with less than a college education. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 7: Balance test

	Birth order	Siblings	Born in Norway	Mother years of education	Mother age	Mother married	Mother log income	Mother not in labor force
Same-gender GP	0.044 (0.042)	0.049 (0.055)	-0.025 (0.017)	0.152 (0.111)	0.086 (0.195)	0.009 (0.020)	-0.020 (0.023)	-0.003 (0.015)
Mean	1.865	1.698	0.861	13.775	28.974	0.567	12.049	0.121
Observations	8424	8424	8424	6996	8423	8034	8424	8424

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \theta_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (θ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Mother years of education, mother age and mother marital status are calculated at the year of birth. Mother earnings and employment are measured when the child is five years old as our data on earnings and employment does not date as far back. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 8: Placebo Tests; Post high school GP swaps

	Compulsory school STEMM GPA	High school STEMM GPA	High school academic track	High school STEMM degree
Same-sex GP	0.024 (0.028)	-0.048 (0.046)	0.009 (0.009)	-0.012 (0.015)
Mean	4.048	3.973	0.752	0.146

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls who subject to at least one exogenous GP swap between age 21 and 24. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

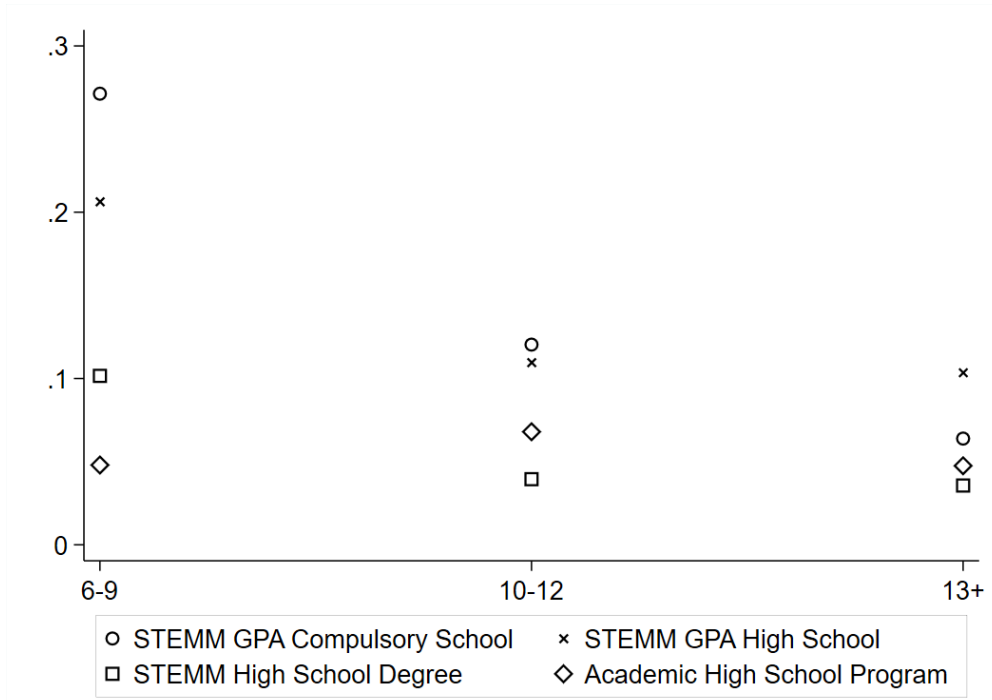
Table 9: Potential pathways and mechanisms

<i>Panel A: Potential indirect effects through health</i>							
	Number of GP visits	Still with assigned GP	Number of diagnoses	Mental health diagnosis	Bith control visits	Fertility	Fertility
Same-gender GP	-0.009 (0.070)	-0.014 (0.019)	-0.039 (0.028)	-0.000 (0.000)	0.002 (0.005)	-0.000** (0.000)	
Mean	1.166	0.613	2.241	0.078	0.069	0.000	

<i>Panel B: Potential indirect effects through mother</i>							
	Years of education	Total income	Not in labor force	Number of diagnoses	Mental health diagnosis	Bith control visits	Fertility
Same-gender GP	0.004 (0.109)	-0.004 (0.021)	0.009 (0.012)	0.176 (0.141)	-0.003 (0.018)	0.001 (0.012)	-0.001 (0.001)
Mean	14.146	12.740	0.066	3.614	0.186	0.068	0.003

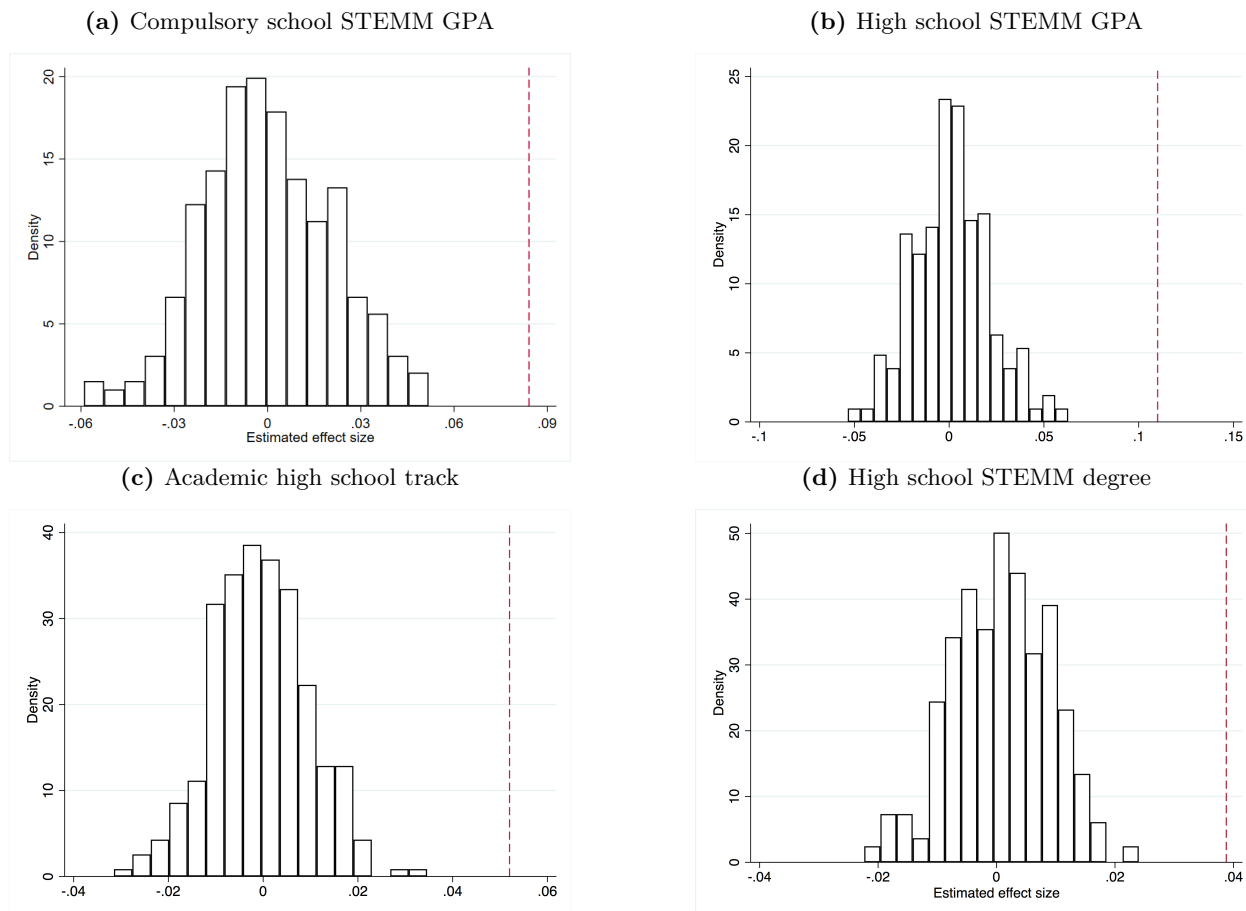
Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample in Panel (A) includes all girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Sample in Panel (B) include all mothers to girls who were born between 1988 and 1996 and who were exposed to at least one exogenous GP swap prior to age 15. All outcomes are measured when the girls are 15 years old. Total income is reported in logarithmic form. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Figure 1: Same-gender role model effects by age-at-swap, females



Notes: Authors' estimation of a equation (2) as described in text and reproduced here for the sake of clarity: $y_i = \alpha + \sum_{a=1}^3 [\partial_a(ExogSwap_{i_a} x GPMatch_i)] + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in the age range denoted on the x axis. Heterogeneity in effect size across age ranges are driven both by differences in the length of exposure and differences in how susceptible individuals are to role models influences in the different age ranges. Standard errors are clustered at level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 that were subject to at least one exogenous GP swap prior to age 15.

Figure 2: Permutation tests



Notes: Authors' estimation of equation (1) as described in text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each subfigure, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. Each subfigure plots the distribution of point estimates obtained from 300 permutations of Equation (1) in which treatment status has been randomly assigned across individuals holding the distribution of treatment constant. The dotted vertical line illustrates where in this distribution the identified effect in Table 2 falls. Regressions include municipality, year of swap, birth year and previous GP fixed effects. Standard errors are clustered at level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 that were subject to at least one exogenous GP swap prior to age 15.

Online Appendix: Not For Publication

Table A1: Overview of high school educational programs and specializations

Panel A: Academic education programs			
Year 1 (education program)	Year 2 (subfield)	Year 3 (subfield)	Competence
Art, design and architecture	Art, design and architecture	Art, design and architecture	University admissions certification
Media and communication	Media and communication	Media and communication	University admissions certification
Music, dance and drama	Dance Drama Music	Dance Drama Music	University admissions certification University admissions certification University admissions certification
Specialization in general studies	International Baccalaureate Languages, social sciences and economics Natural science and mathematics	International Baccalaureate Languages, social sciences and economics Natural science and mathematics	University admissions certification University admissions certification University admissions certification
Sports and physical education	Sports and physical education	Sports and physical education	University admissions certification
Panel B: Vocational education programs			
Year 1 (education program)	Year 2 (subfield)	Year 3 (subfield)	Competence
Agriculture, fishing and forestry	<i>Further specialization</i>	<i>Further specialization</i>	Vocational competence / craft certificate / university admissions certification
Building and construction	<i>Further specialization</i>	<i>Further specialization</i>	Craft certificate
Design, arts and crafts / media producti	<i>Further specialization</i>	<i>Further specialization</i>	Vocational competence / craft certificate
Electricity and electronics	<i>Further specialization</i>	<i>Further specialization</i>	Vocational competence / craft certificate
Healthcare, childhood and youth devel	<i>Further specialization</i>	<i>Further specialization</i>	Vocational competence / craft certificate
Restaurant and food processing	<i>Further specialization</i>	<i>Further specialization</i>	Craft certificate
Technical and industrial production	<i>Further specialization</i>	<i>Further specialization</i>	Craft certificate
Service and transport	<i>Further specialization</i>	<i>Further specialization</i>	Craft certificate

Table A2: Descriptive statistics of individuals not in sample

	<i>Girls</i>		<i>Boys</i>	
	Mean	St.D	Mean	St.D
Birth order	1.830	0.950	1.817	0.936
Siblings	1.683	1.041	1.671	0.999
Born in Norway	0.855	0.352	0.841	0.366
Mother age (birth)	28.974	4.873	29.384	4.777
Mother marital status (birth)	0.614	0.487	0.637	0.481
Parental years of education (birth)	13.937	2.353	14.300	2.301
Both parents born abroad	0.061	0.240	0.063	0.251

Notes: Authors' calculations based on detailed administrative records. Sample includes all boys and girls born between 1988 and 1996 who were not subject to an exogenous GP swap prior to age 15. Number of observations is approximately 108500 girls and 73400 boys.

Table A3: Swap frequency

Exogenous GP swap	
No Swaps	0
One Swap	12292
Two Swaps	1333
Three Swaps	131
Four Swaps	13
Five Swaps	1

Notes: Number of individuals that experienced different frequencies of exogenous swaps prior to turning 15. Sample includes all boys and girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15.

Table A4: Probability of staying with GP as a function of observable characteristics

	Birth order	Siblings	Born in Norway	Mother years of education	Mother age	Mother married	Mother log income	Mother not in labor force
Same-gender GP	-0.003 (0.028)	0.019 (0.027)	0.008 (0.009)	0.151* (0.080)	-0.022 (0.138)	-0.002 (0.014)	0.022 (0.020)	0.004 (0.008)
Mean	1.865	1.698	0.861	13.775	28.974	0.567	12.049	0.121
Observations	8424	8424	8424	6996	8423	8034	8424	8424

Notes: The table shows the β coefficients obtained through estimation of a modified version equation (1): Still with exogenously-assigned GP at age 15 $y_i = \alpha + \beta X_i + \tau_i + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_i), birth year (θ_c) and previous GP (ρ_d) fixed effects. X represents a vector of demographic characteristics and is made up of all variables presented as column titles in this table. The regression includes municipality (π_m), year of swap (τ_i), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates illustrate the difference in the observable characteristic at the top of the column between those who stayed with their exogenously-assigned GP until age 15 and those who did not stay with their exogenously-assigned GP until age 15. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Mother years of education, mother age and mother marital status are calculated at the year of birth. Mother earnings and employment are measured when the child is five years old as our data on earnings and employment does not date as far back to enable us to measure these outcomes prior to child birth. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table A5: Effect of same-gender GP on females, additional outcomes

	Non-STEMM academic tracks	Non-academic tracks
Same-gender GP	-0.028 (0.020)	-0.012 (0.010)
Mean	0.726	0.079
Observations	8449	8680

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls who were subject to at least one exogenous GP swap prior to age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table A6: Effect of same-gender GP on female educational choice and performance, stayers v. leavers

	Compulsory school STEMM GPA	High School STEMM GPA	High school academic track	High school STEMM degree
<i>Panel A: Individuals remaining with their exogenously assigned GP until age 15</i>				
Same-gender GP	0.075 (0.059)	0.143** (0.044)	0.048** (0.024)	0.070** (0.028)
Mean	4.305	3.986	0.740	0.193
Observations	5160	5040	5168	5168
<i>Panel B: Individuals not remaining with their exogenously assigned GP until age 15</i>				
Same-gender GP	0.061 (0.077)	0.021 (0.069)	0.073** (0.031)	0.011 (0.033)
Mean	4.300	4.002	0.732	0.196
Observations	3249	3169	3255	3255

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_i + \tau_i + \pi_m + \theta_c + \theta_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_i), birth year (θ_c) and previous GP (ρ_i) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Panel A includes all girls who were subject to at least one exogenous GP swap prior to age 15 and remained with that GP until age 15. Panel B includes all girls who were subject to at least one exogenous GP swap before age 15 and did not remain with that GP until age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table A7: Effect of same-gender GP on female educational choice and performance, GP list termination only

	Compulsory school STEMM GPA	High school STEMM GPA	High school academic track	High school STEMM degree
Same-gender GP	0.134*** (0.042)	0.134*** (0.045)	0.062*** (0.019)	0.063*** (0.021)
Mean	4.293	3.990	0.734	0.193
Observations	8174	7825	8236	7986

Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. The point estimates depicted in the table should be interpreted as the effect of random assignment to same-gender GP in childhood on the outcome listed at the top of the column. Standard errors are clustered at the level of the exogenously-assigned GP. Sample includes all girls who were subject to at least one exogenous GP swap caused by GP list termination prior to age 15. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Department of Economics
University of Bergen
PO BOX 7802
5020 Bergen
Visitor address: Fosswinckels gate 14
Phone: +47 5558 9200
www.uib.no/econ/