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 2018; 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 body of work within the field of education economics supports the role model hypothesis, showing that females who are exposed to female role models in high school and college perform better in school, and are more likely to select into traditionally male dominated fields (Bettinger and Long 2005; Breda et al. 2018; Carrell et al. 2010; Dee 2004; Dee 2005; Eble and Hu 2017; Griffith 2014; Hoffman and Oreopoulos 2009; Kofoed and McGovney 2017; Lim and Meer 2017; Mansour et al. 2018; Porter and Serra 2019). However, several questions remain: are these same-sex role model effects temporary or permanent? Do they extend to more general settings outside of educational institutions? Can exposure in childhood – before any substantial education investments have been made – produce similar effects?

The goal of this paper is to move beyond the existing role model literature by addressing each of these 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 referrals to hospitals and specialists. 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.[2] 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 childhood exposure to

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

female GPs – a group of successful female STEMM role models – has an effect on the educational choice and performance of girls.[3]

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 persistence as the children age. Third, the role model interactions we study take place outside the classroom. This allows us to better understand to what extent the same-sex role model effects that have been identified in the education literature extend to more general settings outside of educational institutions. This is interesting as the majority of role model studies in the classroom have focused on teacher-student interactions. Such 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. Additionally, while it is very difficult to disentangle same-gender role model effects from other potential mechanisms 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 10^{th} grade – the earliest age at which students have options regarding subject specialization – into college. To ensure that our results are not identified off of a systematic correlation between the characteristics of the previous GP (such as gender) and the gender of the new exogenously-assigned GP, all results are obtained through regressions that incorporate a full set of previous doctor fixed effects. In addition, we empirically test that there are no correlations between GP gender and observable patient characteristics.

[3] 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.

Exposure to a female GP has a statistically significant and economically meaningful 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 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.[4] For example, Breda et al. (2018) find that a one-hour classroom intervention by an external female role model raises the probability that high school girls enroll in a selective STEM track in higher education the following year by 30 percent, and Porter and Serra (2019) find that a 15-minute classroom intervention in which a female economics alumni discussed her experiences as an economics major and her career path increased the likelihood that a female student would major in economics by 8 percentage points, over a baseline of 9 percent (a 90 percent increase in the share of female economics majors).

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 suggest that it is mainly high ability children with low-educated mothers who are driving our results. This demonstrates that female role models may help narrow the gifted gap for disadvantaged girls.[5] This is consistent with evidence suggesting that investments in early childhood, in particular among children from disadvantaged backgrounds, can reduce inequality and help

[4] Our decision to focus on STEMM (STEM plus Medicine) rather than STEM is driven by the fact that the role models in our setting belong to the former rather than the latter group. While this distinction is irrelevant for the high school outcomes we examine (as there are no specific high school medicine tracks or courses), it does affect the construction of certain college outcomes we study (major choice). In the result section we will demonstrate that our results on college major choice likely are driven both by an increased propensity to pursue traditional STEM degrees and an increased propensity to pursue medicine. This suggests that the effects we identify represent an effect of female GPs on STEMM interest more generally, rather than just an effect on the decision to pursue medicine.

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

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 educational settings such as classrooms 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 2018). Further, female students with same-gender role models are not only selecting into traditionally male-dominated education programs, but they are also 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 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.[6] 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

[6] 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.

(mothers' interactions with the GP might also expose the mother to a female STEMM role model, which could impact the mothers' outcomes directly and their children's outcomes indirectly).[7] 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. In addition, our findings do not indicate 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.

Our paper contributes to the existing literature in several ways. First, this is the first paper to study the effect of one-on-one 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 educational settings such as classrooms 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. In a seminal paper, Beaman et al. (2012) show that female representation among politicians can affect the gender gap in aspirations and education among adolescents. Although that study does not examine the implications of one-on-one interactions between individuals and potential role models, it does provide suggestive evidence that role model effects may exist in 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

[7] 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).

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. Most importantly, we check for balance on characteristics between girls who get assigned women vs men GP and we provide detailed evidence on the exogeneity of individual assignment to GPs of different genders. In Section 6, we discuss our findings and the implications of these findings for future public policy discussions.

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.[8]

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, 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. The average time that the GP spends with the patient during an appointment is 20 minutes (Mjølstad and Stund 2019), and most children therefore spend several hours interacting with their GP before making their first decision on education specialization at age 15.

[8] There are currently 422 municipalities and 4 health regions in Norway.

Beginning in 2001, the government assigns every individual residing in Norway to a local GP.[9] In most cases, patients interact with their assigned GP every time they engage with the healthcare system.[10] 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, and the initial assignment in 2001 was primarily based on patient preferences. As of 2015, there were 4,500 GPs in Norway, and each GP had an average of 1,200 patients on their list.[11] 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.[12] 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 affects the educational choice and performance of females.[13] We do not use the initial assignment, nor any swaps initiated by the patients, due to endogeneity concerns. 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.

---

[9] Specifically, The Norwegian Health Economics Administration (part of the Norwegian Directorate of Health), assigns individuals to local GPs on behalf of the government.
[10] There are a few exceptions to this, for example if the patient is brought into the ER. In this case, there is not necessarily any interaction between the patient and the GP.
[11] Patient list refers to the roaster of individual patients that the doctor is responsible for.
[12] Individuals can 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 3 for more details.
[13] The exogenous GP swaps we use in the main analysis are thus driven both by GPs reducing, and by GPs terminating, their patient lists. List terminations are driven primarily by GP retirement, and could be perceived as a cleaner source of exogenous variation than list reductions. We therefore also perform a robustness check in which we use only those GP swaps that were caused by list terminations. See Section 4 for more details.

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 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 most 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 STEMM degree with university admission certification, a student has to select the "Specialization in General

[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] Individuals in vocational programs can take supplemental courses to attain this qualification.

Studies" education program in the first year of high school, and then specialize in "Natural Science and Mathematics" in the second and third years.

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 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

[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.

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 education data is 2014. Panel A of Table 1 shows demographic information for these individuals. Note that only individuals who have attended high school are included in the data. As there is a substantial difference in the probability of attending high school between females and males in Norway, this means that our female sample will be noticeably larger than our male sample.[17]

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).[18]

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.[19] To this end, we focus exclusively on GP swaps that occurred due to the doctor deciding to terminate, or significantly reduce, the patient list.[20] 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.[21] The data

[17] This is based on authors' calculations using publicly-available data from Statistics Norway. See "Upper Secondary Education Advanced Course II/Certificate" in Table 5 at https://www.ssb.no/en/vgu/.
[18] 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.
[19] 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.
[20] Note that our GP data starts in 2006, and we therefore do not have information on – and cannot use – the initial GP assignments in 2001. However, even if we had this information, as stated in the Background Section, the initial assignment was influenced by patient preferences. Due to endogenous selection, we would therefore not want to use this information for the purpose of our analysis.
[21] Frequency of GP swaps are shown in Appendix Table A3.

show that approximately 40 percent of girls, and 60 percent of boys, are matched to a same-sex GP when experiencing an exogenous swap. This indicates that there are more male than female GPs in our sample. The data further 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. Given the average appointment time of 20 minutes, this means that each child in our sample, conditional on going to the GP, spends approximately 10 hours with their GP before deciding which high school education program to pursue.[22] 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).

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.[23]

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 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),

[22] This is the upper bound on exposure to GPs as role models. For instance, GP visits as an infant likely do not have role model effects. Additionally, as shown in Panel B of Table 1, many children switch GPs. The average number of years that the individuals in our sample remain with their exogenously-assigned GP is 3, such that they were exposed to the GPs through one-on-one consultations for an average of 120 minutes. However, (1) the total time spent with GPs is much longer than in previous studies and (2) these are one-on-one interactions, which contrasts with much of the previous role model research (e.g. Breda et al. 2018; Porter and Serra 2019).

[23] 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.

while most students in vocational programs do not.[24] Next, we focus on the primary research question of this paper: Does a same-gender GP role model encourage girls to enroll in STEMM fields? We examine this question by estimating the effect of GP gender match on the probability of graduating with a high school STEMM degree. As discussed in Section 2, this choice is important for STEMM-eligibility at the university level (the correlation between high school STEMM and college STEMM 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.

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 STEMM gender gap, we focus specifically on STEMM 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)

[24] Individuals in vocational programs can take supplemental courses to attain this qualification.

between ages 6 and 15 (the age at which children submit their high school applications).[25] We estimate models of the following form separately for males and females:

$$Y_i = \alpha + \beta_1 GP\_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \varepsilon_i, \tag{1}$$

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. As individuals' decision to swap GPs or comply with the assigned GP is endogenous, we focus exclusively on the first exogenous swap of the individuals in our sample. 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.[26] All standard errors are clustered at the level of the exogenously-assigned GP.[27] Except for restricting the sample to children who experienced an exogenous GP switch between the ages of 6 and 15, we do not impose any restrictions on the sample.[28] However, as the exogenous swaps we study do not affect all individuals in the country, the final analysis sample is significantly smaller than the total population.

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

[25] 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.
[26] 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. To the extent that girls who are exogenously assigned to male GPs are more likely to endogenously swap to female GPs, that would attenuate our results and push us towards finding no effects.
[27] 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.
[28] However, as noted in Section 3, only individuals who have attended high school are included in the registry. As there is a substantial difference in the probability of attending high school between females and males, this means that our female sample will be noticeably larger than our male sample.

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.[29] 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 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

[29] 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.

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), which includes a full set of birth cohort, previous doctor, municipality, and year of swap fixed effects. As individuals' decision to swap GPs or comply with the assigned GP is endogenous, we focus exclusively on the first exogenous swap of the individuals in our sample. This protects us against any potential endogeneity concerns after the initial swap, but likely attenuates our estimates. The point estimates thus 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. Evidence on the validity of our identification assumptions, most importantly the random assignment of GP gender, is provided in Section 5.[30]

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.[31] Column 2 explores if GP

[30] Note that missing data for a small number of individuals for some of the outcomes means that there are minor differences in sample sizes between the different columns (most of which is driven by individuals dropping out of high school prior to graduation, or not taking STEMM classes).

[31] 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.

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 0.084 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.

As discussed in Section 2, the exogenous GP swaps are driven by GPs reducing, or completely terminating, their patient list. List terminations are driven primarily by GP retirement, and may be perceived as a cleaner source of exogenous variation than list reductions. We therefore perform a robustness check in which we 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.

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.[32] Specifically, the percentage of boys in STEMM fields is substantial, and there is a disproportionate availability of public figure male STEMM role models.

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. This is an interesting result, especially in light of a few cross-sectional studies suggesting that female GPs are associated with slightly better outcomes for patients than male GPs (e.g. Tsugawa et al. 2017). However, it is also worth noting that the individuals in our sample are very young, and on average have much fewer health problems than the general population.

Taken together, the results in Tables 2 and 3 suggest that same-gender GPs have a significant effect 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 2018). In terms of policy implications, an interesting question with respect to our results is whether the effects are driven simply by

[32] 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.

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 this exercise are highly speculative as the decision to not remain with their initially exogenously assigned GP likely are endogenous, and should therefore be interpreted with great 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 a college degree or not.[33] 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

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

transmission of human capital and how to facilitate upward socioeconomic mobility (e.g. Black, Devereux and Salvanes 2005).

To ensure that this effect heterogeneity is not simply driven by female GPs being better at communicating with low educated families than male GPs, we also estimate Equation (1) for boys to low-educated mothers. The results are provided in Appendix Table A8. Looking across the table, none of the coefficients are statistically significant, supporting the conclusion reached in our main text.

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 suggests that the effect of same-gender GP match on educational performance loads on individuals in the right tail of the ability distribution. While the results for high school STEMM GPA are a bit noisier with respect to quantiles 7 and 8, and therefore have to be interpreted a bit more cautiously, the general pattern of results is similar to that for compulsory school STEMM GPA: the effect of same-gender GP match on educational performance loads on individuals in the higher end of the distribution, with small and not significant effects in the bottom third.

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:

$$Y_i = \alpha + \sum_{a=1}^{3}[\partial_a(ExogSwap_{ia}\ x\ GP\_Match_i)] + \tau_t + \pi_m + \theta_c + \rho_d + \varepsilon_i, \qquad (2)$$

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.[34] 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 ($ExogSwap_{ia}\ x\ GP_{Match_{ia}}$) 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.[35]

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

[34] Results using alternative groupings are available from the authors upon request.
[35] It is important to note that any 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.

noticeable when looking at STEMM GPA in high school and STEMM high school degree. However, as shown in Appendix Table A9, which provide the full set of point estimates and associated standard errors of the results shown in Figure 1, these effects are relatively noisy and not statistically significantly different from each other. Thus, this is only suggestive evidence regarding heterogeneity by age of swap, and the results need to be interpreted with care.

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 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 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 3) and the probability of choosing a STEMM major (Column 4). While it would also be interesting to study college graduation effects, this is unfortunately not possible due to data limitations. Specifically, the children in our sample are too young to have graduated from college by the time our analysis period ends.

Column 3 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 4 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.[36] Thus, exposure to same-gender role models do not affect an individual's decision to attend university, but conditional on wanting to attend university, it does affect the choice of major.[37] It should be noted that the effect we identify is economically large, but fall within the range of the effects that have been identified from shorter information interventions in the classroom. For example, Breda et al. (2018) find that a one-hour classroom intervention by an external female role model raises the probability that high school girls enroll in a selective STEM track in higher education the following year by 30 percent, and Porter and Serra (2019) find that a 15-minute classroom intervention in which a female economics alumni discussed her experiences as an economics major and her career path increased the likelihood that a female student would major in economics by 8 percentage points, over a baseline of 9 percent (a 90 percent increase in the share of female economics majors).[38]

---

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

[37] As can be seen in Table 6, 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.

[38] The STEMM versus STEM distinction is irrelevant for the high school outcomes we examine as there are no specific high school medicine tracks or courses. However, it is possible that college STEMM major effect is driven exclusively by treated individuals being more inclined to pursue medicine. To examine this in detail, we have reestimated equation (1) using the probability of pursuing medicine, and the probability of pursuing STEM, as the dependent variables. We find that about half of the effect comes from enrolling in medicine. This suggests that the identified effect in Table 6 likely is driven both by an increase in the probability of pursuing STEMM, and in the probability of pursuing medicine.

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.[39] A joint test of significance for the covariates in Table 7 further supports the randomization assumption, failing to reject the null hypothesis that the covariates are jointly unable to predict the gender of the exogenously-assigned GP (p-value of 0.557). These results are inconsistent with non-random sorting across GP gender, and support a causal interpretation of our results.[40]

It is worth noting that the sample used for the balance test is slightly smaller than the sample used for the main analysis. This is due to missing information on detailed parental characteristics for a small number of individuals in our analysis. To ensure that this is not biasing our results, we have also restricted the sample to only those for which we have a full set of balance variables, and reestimated our main results. The results from this exercise on our core results are provided in Appendix Table A12. The results from these exercises show that our findings are robust to discarding those observations for which we lack some of the characteristics used in the balance test.

[39] The education balance is robust to using a dummy variable for whether the mother has a college degree or not (coefficient estimate of -0.006 with a standard error of 0.020).

[40] The balance test for males is provided in Appendix Table A10, and the balance test for females without previous GP fixed effects are provided in Appendix Table A11.

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 20 through 25 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 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 20 through 25 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.[41] If the results in Table 2 represent true effects of same-gender GP assignment, the estimates in Table 2 should be larger

[41] 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.

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 STEMM degree, compulsory STEMM GPA, and high school STEMM 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.

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).[42]

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

[42] 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.

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 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.[43]

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.[44] 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

[43] 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.
[44] The results on the mother's education and labor market outcomes are already known from the balance test in Table 7.

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 STEMM. 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.

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 STEMM role models - perform better in school and are more likely to pursue STEMM 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 STEMM degrees from high school, and choose STEMM majors at university. A back-of-the-envelope calculation suggests that female role models can close the gender gap in college STEMM choice by up to 20 percent. This effect is economically large, but falls 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). 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.

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 STEMM 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

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India" *Science* 335(6068): pp. 582-586

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 (2016). “The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity experiment” *American Economic Review* 106 (4): pp. 855-902

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 (2018). "On the origins of gender gaps in human capital: short and long term consequences of teachers' biases" *Journal of Public Economics* 167(C): pp. 263-279

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*

Mjølstad, B., and M. Sund (2019). "Er 20 minutter hos fastlegen nok?" *Dagens Medisin*, September 30. Accessed via https://www.dagensmedisin.no/artikler/2019/09/30/er-20-minutter-hos-fastlegen-nok/

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

Porter, C., and D. Serra (2019). "Gender Differences in the Choice of Major: The Importance of Female Role Models" *American Economic Journal: Applied Economics* (forthcoming)

Tsugawa, Y., A. Jena., J. Figueroa, E. Orav, D. Blumenthal, and A. Jha (2017). "Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs female physicians" *JAMA Intern Med. 177(2)*: pp. 206-213

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

	Girls		*Boys*	
	Mean	St.D	Mean	St.D
Panel A: Family Composition				
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	28.987	5.008	29.414	4.822
Mother marital status	0.567	0.496	0.595	0.491
Mother years of education	14.146	2.396	14.445	2.356
Mother log earnings	12.740	0.560	12.782	0.587
Mother not in labor force	0.067	0.250	0.070	0.255
Panel B: GP Visiting Behavior				
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
GP gender match	0.405	0.491	0.627	0.483
Panel C: Educational Performance and Choice				
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. Mother age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of birth).

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***	0.039**	0.084**	0.109***
	(0.017)	(0.018)	(0.039)	(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	0.002	-0.038	0.010
	(0.020)	(0.025)	(0.047)	(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

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

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 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 STEMM GPA - quantile effects

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

Notes: Authors' estimation of equation (1) as described in text using the unconditional quantile 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*	0.067**	0.005	0.022*
	(0.023)	(0.028)	(0.017)	(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. 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 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	0.049	-0.025	0.004	0.086	0.009	-0.004	0.009
	(0.042)	(0.055)	(0.017)	(0.109)	(0.195)	(0.020)	(0.021)	(0.012)
Mean	1.865	1.698	0.861	14.146	28.974	0.567	12.740	0.066
Observations	8424	8424	8424	8034	8423	8034	8357	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: $y_i = \alpha + \beta_1 GP\_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. 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. Mother age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of birth). * 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

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Same-sex GP	0.008	-0.005	-0.030	-0.022
	(0.017)	(0.018)	(0.039)	(0.036)
Mean	0.759	0.199	4.120	3.986
Observations	6730	6730	6730	6730

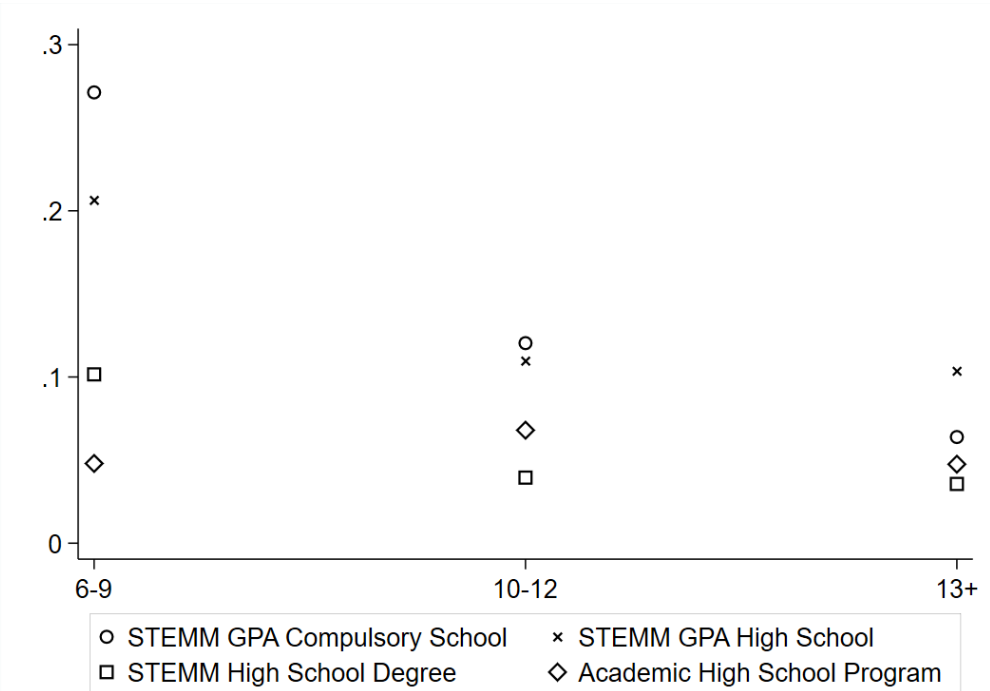
Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text and reproduced here for the sake of clarity: $y_i = \alpha + \beta_1 GP\_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top 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 their first exogenous GP swap between age 20 and 25. * 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

Panel A: Potential indirect effects through health							
	Number of GP visits	Still with assigned GP	Number of diagnoses	Mental health diagnosis	Bith control visits	Fertility	
Same-gender GP	-0.009	-0.014	-0.039	-0.000	0.002	-0.000**	
	(0.070)	(0.019)	(0.028)	(0.000)	(0.005)	(0.000)	
Mean	1.166	0.613	2.241	0.078	0.069	0.000	
Panel B: Potential indirect effects through mother							
	Years of education	Total income	Not in labor force	Number of diagnoses	Mental health diagnosis	Bith control visits	Fertility
Same-gender GP	0.004	-0.004	0.009	0.176	-0.003	0.001	-0.001
	(0.109)	(0.021)	(0.012)	(0.141)	(0.018)	(0.012)	(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.

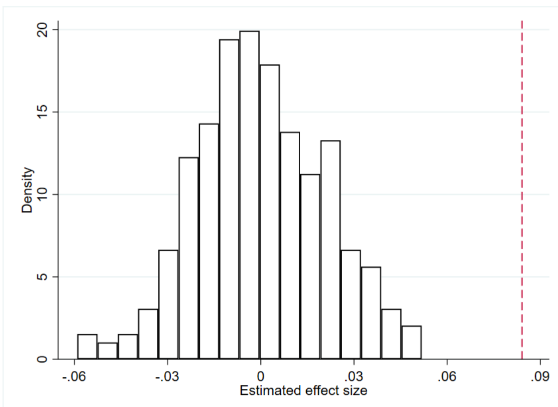
Figure 1: Same-gender role model effects by age-at-swap, females



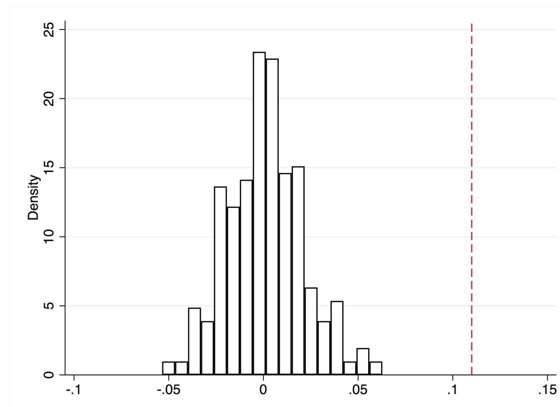
Notes: Authors' estimation of a equation (2) as described in text and reproduced here for the sake of clarity: $y_i = \alpha + \sum_{a=1}^{3}[\partial_a(ExogSwap_{ia} \times GPMatch_i)] + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each 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 susceptive 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

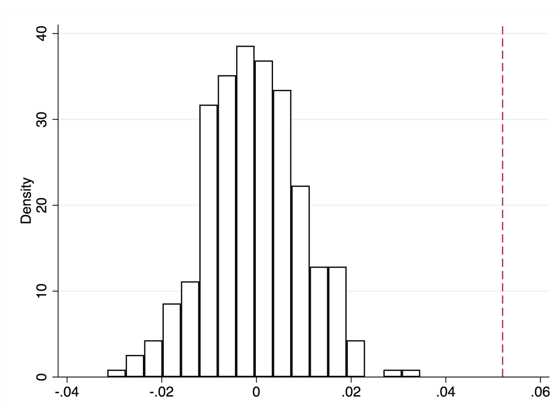
(a) Compulsory school STEMM GPA



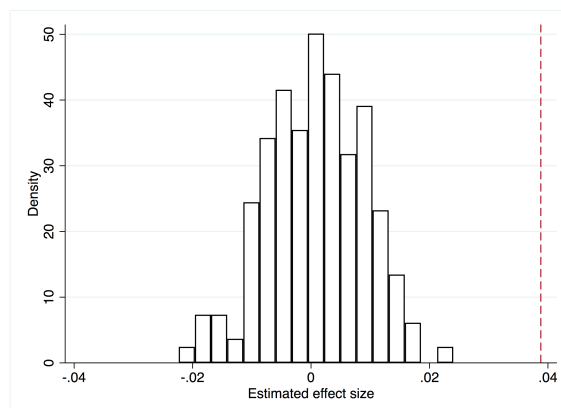
(b) High school STEMM GPA



(c) Academic high school track



(d) High school STEMM degree



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	Dance	University admissions certification
	Drama	Drama	University admissions certification
	Music	Music	University admissions certification
Specialization in general studies	International Baccalaureate	International Baccalaureate	University admissions certification
	Languages, social sciences and economics	Languages, social sciences and economics	University admissions certification
	Natural science and mathematics	Natural science and mathematics	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	*Further specialization*	*Further specialization*	Vocational competence / craft certificate / university admissions certification
Building and construction	*Further specialization*	*Further specialization*	Craft certificate
Design, arts and crafts / media producti	*Further specialization*	*Further specialization*	Vocational competence / craft certificate
Electricity and electronics	*Further specialization*	*Further specialization*	Vocational competence / craft certificate
Healthcare, childhood and youth devel	*Further specialization*	*Further specialization*	Vocational competence / craft certificate
Restaurant and food processing	*Further specialization*	*Further specialization*	Craft certificate
Technical and industrial production	*Further specialization*	*Further specialization*	Craft certificate
Service and transport	*Further specialization*	*Further specialization*	Craft certificate

Table A2: Descriptive statistics of individuals not in sample

	Girls		*Boys*	
	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	28.974	4.873	29.384	4.777
Mother marital status	0.614	0.487	0.637	0.481
Mother years of education	14.225	2.447	14.575	2.408
Mother log earnings	12.685	0.627	12.717	0.654
Mother not in labor force	0.075	0.264	0.072	0.259
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. Mother age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of birth).

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	0.019	0.008	0.076	-0.022	-0.002	0.011	-0.007
	(0.028)	(0.027)	(0.009)	(0.062)	(0.138)	(0.014)	(0.016)	(0.007)
Mean	1.865	1.698	0.861	13.146	28.974	0.567	12.740	0.066
Observations	8424	8424	8424	8367	8423	8034	8357	8424

Notes: The table shows the β coefficients obtained through estimation of a modified version equation (1): $y_i = \alpha + \beta_i S + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i is a general term denoting the outcome listed on top of each column. S is an indicator for whether the female is still with the exogenously-assigned GP at age 15. The regression includes municipality (π_m), year of swap (τ_t), 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 age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of 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	-0.012
	(0.020)	(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

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

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 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

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

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.

Table A8: Effect of same-gender GP on male educational choice and performance, low-educated mothers

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Same-gender GP	0.031	0.023	-0.010	-0.026
	(0.040)	(0.036)	(0.084)	(0.083)
Mean	0.756	0.254	4.223	3.827
Observations	2604	2594	2604	2594

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 with low educated mothers 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 A9: Effect of same-gender GP on female educational choice and performance, age-at-swap

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Age 6-9	0.048	0.102**	0.271**	0.206
	(0.051)	(0.049)	(0.116)	(0.126)
Age 10-12	0.068***	0.039	0.120**	0.110*
	(0.025)	(0.025)	(0.060)	(0.056)
Agre 13-15	0.048***	0.036*	0.064	0.104**
	(0.018)	(0.019)	(0.041)	(0.043)

Notes: Authors' estimation of a equation (2) as described in text and reproduced here for the sake of clarity: $y_i = \alpha + \sum_{a=1}^{3}[\partial_a(ExogSwap_{ia} \times GPMatch_i)] + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. 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 in the different rows. Heterogeneity in effect size across age ranges are driven both by differences in the length of exposure and differences in how susceptive 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.

Table A10: Balance test, boys

	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	0.031	0.009	-0.044*	-0.186	-0.155	0.006	-0.027	0.017
	(0.049)	(0.060)	(0.023)	(0.132)	(0.280)	(0.030)	(0.036)	(0.016)
Mean	1.850	1.707	0.851	14.449	29.416	0.595	12.781	0.070
Observations	5342	5340	5340	5310	5340	5103	5285	5285

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 males born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Mother age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of birth). * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table A11: Balance test without previous doctor fixed effects

	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	-0.014	0.007	-0.023*	-0.029	0.042	-0.004	-0.014	0.002
	(0.0.029)	(0.034)	(0.012)	(0.084)	(0.140)	(0.013)	(0.016)	(0.008)
Mean	1.865	1.598	0.861	14.146	28.974	0.567	12.740	0.066
Observations	8424	8424	8424	8367	8423	8034	8357	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: $y_i = \alpha + \beta_1 GP\_Match_i + \tau_t + \pi_m + \theta_c + \epsilon_i$. 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) and birth year (θ_c) 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 females born between 1988 and 1996 who were subject to at least one exogenous GP swap prior to age 15. Mother age and marital status are calculated at the year of birth. Mother earnings, education and employment are measured when the child is fifteen years old (when the children in our sample make their high school choices; our data does not extend far back enough for us to get this information at the year of birth). * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table A12: Effect of same-gender GP on educational choice and performance of girls, restricted sample

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Same-gender GP	0.048***	0.033*	0.089**	0.122***
	(0.019)	(0.019)	(0.041)	(0.041)
Mean	0.733	0.191	4.312	4.000
Observations	7953	7954	7942	7749

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. Sample restricted to girls with no missing data on the balance test variables. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.