Can Female Doctors Cure the Gender STEMM Gap?

Evidence from Exogenously-assigned General Practitioners

By JULIE RIISE, BARTON WILLAGE AND ALEXANDER WILLÉN*

Abstract

We use exogenously-assigned general practitioners to study the effects of female role models on educational outcomes of girls. Girls who are exposed to female GPs are more likely to sort into male-dominated education programs in high school, most notably STEMM. These effects persist as females enter college and select majors. The effects are larger for high-ability girls with low educated mothers, suggesting that female role models improve intergenerational mobility and narrow the gifted gap. This demonstrates that role model effects in education need not involve individuals in the classroom, but can arise due to everyday interactions with medical professionals.

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, Lise Vesterlund, Aparna Soni, and Analisa Packham, as well as several other colleagues and seminar participants for valuable comments and suggestions.

1 Introduction

Females outperform males in educational attainment but remain underrepresented in fields with high financial returns, most notably STEMM (Science, Technology, Engineering, Mathematics, and Medicine). This gender imbalance in education 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 same-gender role models, influencers, and mentors.

Exposure to same-gender role models may affect educational decisions through several channels. Role models may fuel higher aspirations, reduce "stereotype threats", and convey important information (Breda et al. 2018).[1] A growing body of work within 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 male dominated fields.[2] However, several questions remain: are these effects temporary or permanent? Do they extend to more general settings outside the classroom? Can exposure in childhood – before important education investments have been made – produce similar effects?

This paper aims to move beyond the existing role model literature by addressing 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,

---

[1] We define role models as "a person whose behavior in a particular role is imitated by others" (Merriam-Webster 2020). However, we acknowledge that several definitions and types of role models exist. For a concise discussion on this topic, see Chung (2000).

[2] Bettinger and Long 2005; Breda et al. 2018; Carrell et al. 2010; Dee 2004, 2005; Eble and Hu 2017; Griffith 2014; Hoffman and Oreopoulos 2009; Kofoed and McGovney 2017; Lim and Meer 2017; Mansour et al. 2018; Gershenson et al. 2018; Porter and Serra 2019; Canaan and Mougaine 2019.

and are responsible for diagnosing and treating patients and referring them to hospitals and specialists. When individuals are no longer able to consult with their existing GPs (e.g. due to GP retirement), the Norwegian Health Economics Administration (HELFO) randomly reassigns them to new GPs conditional on municipality and availability. We use GP reassignments due to GP retirement or other causes outside the patient's control as a source of exogenous variation to test if childhood exposure to female GPs – a group of successful female STEMM role models – has an effect on the educational choice and performance of girls.

Using doctor-patient interactions to test for same-gender role model effects has several benefits. First, the interactions take place in childhood before any educational decisions have been made. Second, by tracing the role model effects throughout a child's education career – from compulsory school to college – we can examine effect persistence as children age. Third, the role model interactions take place outside the classroom. This allows us to better understand to what extent same-sex role model effects, as identified in the education literature, extend to more general settings. This is interesting as the majority of role model studies in the classroom have focused on teacher-student interactions. Such effects may be driven by gender differences in teaching practices, the student-teacher interaction may be different from other types of social interactions, and the purpose of a classroom (to create an environment conducive to academic development) may make children more receptive to role model influences. Additionally, while disentangling same-gender role model effects from other potential mechanisms in the classroom is difficult (e.g. differences in teaching practices), our setting allows us to test for these cofounders (gender differences in health practices) directly. Finally, while females remain underrepresented among GPs (37 percent of GPs were females during our analysis period), this is a less male-dominated field than for example math, physics, and computer science. Understanding if same-gender role models can impact girls' educational performance and career choices even in relatively more balanced fields has important policy implications.

For our analysis, we leverage rich matched doctor-patient administrative data on all children in Norway who were subject to an exogenous GP reassignment between 2002 and 2011. We link these data to detailed information on educational performance and choices made throughout individuals' academic careers, from 10th grade – the earliest age at which students have subject specialization options – into college. To account for the potential systematic correlation between previous GP characteristics (such as gender) and the gender of the new exogenously-assigned GP, all analyses incorporate a full set of previous GP fixed effects.

Exposure to a female GP has a statistically significant and meaningful effect on both STEMM choice and educational performance among girls. Specifically, assignment to a female GP during childhood increases the probability of choosing a STEMM program in high school by 4 percentage points (20 percent relative to the mean), and increases high school STEMM GPA by 0.09 standard deviations. These effects persist as girls enter college: assignment to a female GP increases 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 identified from shorter information interventions in the classroom (Breda et al. 2018; Porter and Serra 2019).[3] Our paper shows that role model effects in education need not involve the classroom, but can arise due to everyday interactions with medical professionals.

We find significantly larger effects for girls with low-educated mothers – a group that may be 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

[3] We focus on STEMM (STEM plus Medicine) rather than STEM since the role models in our setting belong to the former rather than the latter group.

facilitate socioeconomic mobility (e.g. Black et al. 2005). Quantile regressions suggest high-ability children with low-educated mothers drive the results. This demonstrates that female role models may help narrow the gap between high-ability students from more and less advantaged socioeconomic backgrounds. This is consistent with evidence suggesting that investments in early childhood, in particular among children from disadvantaged backgrounds, can reduce inequality (e.g. Carneiro and Heckman 2003; Elango et al. 2015). Consistent with papers on classroom role models (e.g. Carell et al. 2010), we find no effect of GP gender match on male educational outcomes.

This paper provides novel insights into how same-gender role models in childhood outside educational settings may shape educational choices and career decisions. These results have important policy implications. Specifically, educational choices in high school and college likely have significant effects on future labor market opportunities and career decisions, and may help close the gender wage gap (Carrell et al. 2010; Lavy and Sand 2018). Further, female students with same-gender role models not only select into traditionally male-dominated education programs but also perform better in school, suggesting that same-gender role models improve education matches. Intentionally matching girls to female role models may be an effective tool for narrowing the gender gap in educational choice and labor market outcomes.

In addition to a direct role model effect, some of our results could operate through health (same-gender GPs may impact interactions with the health system) and family (mothers' interactions with the GP might also expose her to a female role model, which could impact the mothers' outcomes directly and their children's outcomes indirectly). While these alternative pathways do not threaten our identification strategy, they would affect the interpretation of our results, because some of the effects could operate through a health-based non-role model channel, or through an indirect role-model channel via the impact on the girl's mother.

We find no evidence of these alternative pathways: assignment to a same-gender GP has no impact on the number of diagnoses, the likelihood of a mental health diagnosis, the probability of visiting the GP for birth control reasons, or fertility. In addition, our findings do not indicate that the effects operate through the mother as measured by her education, labor market, and health outcomes. This suggests the effects we identify likely are driven by direct role model influences between GPs and children.

Our paper contributes to the existing literature in several ways. First, this is the first paper to study one-on-one role model interactions in childhood. This is an important contribution as these children have not made any educational decisions prior to exposure, such that the potential impact of role models is greater. Second, no other paper has examined effects of same-gender role models outside educational settings on educational choices. Our setting is important for understanding the extent to which the earlier research on teacher-student interactions generalizes to non-classroom settings.[4] Third, by examining the effect of GP gender match on health and parents, we can rule out alternative pathways and better isolate the role model effect. Finally, Norwegian registry data allow us to trace children throughout their educational careers and explore long-term effects.

2 Background

2.1 Health Care System and GP Assignment in Norway

The Norwegian public health care system is based on universal access, and enrollment is automatic. As in much of Europe, the health system is a two-part system, with primary care provided by the local municipalities and specialist care provided by larger health regions.[5]

---

[4] 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 suggests that role model effects may exist in non-classroom settings.

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

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 care, and is responsible for diagnosis and treatment. When GPs deem it necessary, they refer patients to specialists. In other words, the Norwegian GPs act as gatekeepers to country's health system. The average time GPs spend with a patient during an appointment is 20 minutes (Mjølstad and Stund 2019), and most children therefore spend several hours interacting with their GP before deciding on education specialization at age 15.

Since 2001, the government assigns every resident to a local GP.[6] In most cases, patients interact with their assigned GP every time they use the healthcare system.[7] Prior to this system, individuals were not tied to a specific GP, and had to find a GP every time they needed care. The system was meant to improve doctor-patient relationships and ensure appropriate use of health care, and the initial assignment in 2001 was primarily based on patient preferences. As of 2015, there were 4,500 GPs, and each GP had an average of 1,200 patients. The average GP was 47 years old, and 60 percent were male.

When GPs retire, move, or for some other reason decide to terminate/reduce their patient list, patients on that GP's list are reassigned to a new GP in the municipality.[8] Within the system's legal framework, there are two important things to note (FOR 2018). First, in the event of list reductions, GPs must randomly select which patients to remove from the list. Second, in the event of reassignment, patients should be randomly assigned to new GPs in the municipality

---

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

[7] There are a few exceptions to this, for example if the patient is brought into the ER.

[8] Individuals can independently change the GP they have been assigned twice per year. Using information on the exact cause of the GP swap, we ignore such endogenous swaps.

conditional on availability.[9] However, the regulatory framework does not specify a randomization device. In Section 5, we perform balance tests and falsification checks to show that the data are consistent with the new GPs being exogenously-assigned to children.

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. 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 effects of same-gender GP on girls' education are operating through their mothers.

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 (around 30 percent of their income), fee-for-service from the Health Economics Administration (almost 70 percent of their income) and out-of-pocket payments from patients. GP financing is determined nationally through collective bargaining.

With respect to the gender balance of physicians in Norway, the number of female physicians has steadily increased over the past several decades, from 10 percent in 1930 to 40 percent in 2010. This number will increase over time, as the current share of female medical students exceed 50 percent. Consistent with other OECD countries, female share across different specializations varies widely, ranging from 6 percent in thoracic surgery to 100 percent in medical genetics (Legeforeningen 2019a). In general, specializations that guarantee

[9] In certain cases, entering GPs can take over the entire list from a retiring GP. However, our identification strategy consists of comparing individuals who had the same GP that then got allocated to different GPs, such that these do not contribute to the identifying variation.

a fixed monthly pay (non self-employed) and that do not require on-call duty tend to have greater female representation. As the majority of GPs are self-employed, and as all GPs are required by law to be on-call at the ER a certain number of weeks per year, females remain underrepresented as GPs.[10] In 2010, 37 percent of GPs were females (Legeforeningen 2019b).

2.2 The Norwegian Education System

The Norwegian education system consists of 10 years of tuition-free compulsory education starting at age 6, with the curriculum set by the central government. During the first seven years, 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 receive grades. Following successful completion of compulsory education, each child has a right to three years of tuition-free high school.

High school in Norway is very different from the US, and includes 13 distinct education programs: 5 academic and 8 vocational. 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 programs. Following the first year, students choose a specialization within their broad education program. Around 60 percent of students pursue an academic track.

High school education provides students with university admission certification, vocational competence, or basic (craft) competence.[11] Students in academic programs receive a university admission certification and can apply to college, while most students in vocational programs do not.[12] Table A1 provides an overview of the education programs and tracks available at the high school level. To obtain a high school STEMM degree, a student has to

---

[10] There are three exceptions to the on-call requirement of GPs: Individuals over the age of 60, females in the last three months of pregnancy, and females with children under the age of one.

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

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

Students apply to high school through a centralized system based on grades from their final year of compulsory education. The application consists of ranking three programs in the county of residence. If the number of applications exceeds program capacity, students are assigned based on compulsory school grades. However, even though admission to specific programs are based on compulsory school grades, national law ensures that all students gain admission to one of the three programs on their list.

A range of universities and colleges offer higher education in Norway, and most are tuition-free public institutions. The Norwegian Universities and Colleges Admission Service coordinates the admissions process. Students apply to specific programs at the different universities, and if the number of applications exceeds the number of seats, students are assigned based on high school grades. Admission is conditional on having graduated from high school with a university admission certification. Additionally, some programs – most notably STEMM – impose specific high school course pre-requisites. This makes it difficult for students with non-STEMM high school diplomas to enroll in STEMM programs.

3 Data and Method

3.1 Data

Our data come from rich administrative records on the universe of Norwegian residents. We restrict our sample to individuals who graduated from high school between 2006 and 2014 (born between 1988 and 1996). We start with 2006, because the current high school structure was introduced in 2006. We stop with 2014, because that is the final year for which we have education data. The GP system was introduced in 2001, so we have information on all exogenous swaps starting in 2002, when the youngest individuals in our sample (born in 1996) were 6 years old. As high school applications are submitted when individuals are aged 15, swaps after age 15 should not impact the outcomes we examine. As such, our main analysis exploits variation in GP-patient gender match among individuals born between 1988 and 1996 who experienced an exogenous GP swap in 2002 or later and are between ages 6 and 15. Panel A of Table 1 shows demographic information for these individuals. Only individuals who have attended high school are included in the data. Since the probability of attending high school differs between females and males (24500 males completed high school in 2016, compared to 31600 females), our female sample is larger than our male sample.[13]

The main strength of our data is that we can link individuals across different longitudinal data registries through unique identifiers. This allows us to combine the demographic information in Panel A of Table 1 with detailed information on GP interactions (through GP and health care registries), educational choice and performance (through education registries) and family characteristics (through intergenerational identifiers and the population registry).[14]

The GP registry provides 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 health care registry information, which includes 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 our study, we are interested in GP swaps that are outside the patient's control, which generates plausibly exogenous variation in the gender of the patients' new GPs. To this end, we focus on GP swaps that are due to the doctor deciding to terminate, or reduce, the patient list. In Section

[13] See "Upper Secondary Education Advanced Course II/Certificate" at https://ssb.no/341521/completed-educational-programmes-in-upper-secondary-education-by-gender-and-results.

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

5, we show that these swaps are unlikely to be correlated with the gender of the assigned GP.

With respect to availability of male and female GPs within each municipality (the level of assignment), there are an average of 9 GPs per municipality, with around two-thirds being male. Approximately 90 percent of all girls live in municipalities that had at least one GP of each gender. These girls are different from the small subset of girls that live in municipalities without both a male and a female GP.[15] Even though these differences are small, we need to be careful to extrapolate our results to the municipalities in which there were not both a male and a female GP available. However, seeing as this represents less than 10 percent of the girls in the country, we do not believe that this constitutes a significant limitation of our paper.

Panel B of Table 1 provides summary statistics for GP visits of the individuals in our sample, as well as information on GP swaps. Approximately 40 percent of girls, and 60 percent of boys, are matched to a same-gender GP after an exogenous swap. This indicates that there are more male than female GPs in our sample. The individuals in our sample met with their GP an average of 2 times per year, conditional on going to the GP, amounting to 30 times before deciding on a high school education program. Given the average appointment time of 20 minutes, this means that each child in our sample spends approximately 10 hours with their GP before deciding which high school education program to pursue. This is an upper bound on exposure to GPs as role models, since the average number of years that the individuals in our sample remain with their exogenously-assigned GP is 3 (indicating 120 minutes of mean exposure). The GP intervention we examine is less intense than the classroom experiments

[15] Girls who lived in municipalities that had at least one male and one female GP were slightly less likely to have siblings (0.180) and be Norwegian born (0.078), and were more likely to have mothers who were more educated (0.17 years of education), slightly older (0.508), more likely to be married (0.028) and had higher earnings (0.070 log points).

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 females come to classrooms to talk (e.g. Breda et al. 2018; Porter and Serra 2019). Importantly, these are one-on-one interactions, which contrasts with much of the previous literature.

Panel B of Table 1 shows 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 study period, very few individuals swapped GPs more than twice (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.[16]

Our education data include detailed information on educational choice and academic performance. In terms of educational choice, we begin by using high school registry data to examine if GP gender match has an effect on the probability of choosing one of the five academic tracks discussed in Section 2. Next, we focus on the primary research question of this paper: Does a same-gender GP role model encourage girls to enroll in STEMM? We examine this question by estimating the effect of GP gender match on the probability of graduating with

[16] Table A4 compares the characteristics of girls who remain with their exogenously-assigned GPs until age 15 with the characteristics of girls who do not remain with their exogenously assigned GP. We fail to reject the null hypothesis that the characteristics of the girls who remain with their exogenously-assigned GP are the same as the characteristics of the girls who do not remain with their exogenously-assigned GP. 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.

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 in college.

After exploring the impact of GP gender match on educational choice, we examine if it had an impact on educational performance. The outcomes we examine include compulsory and high school GPA. Since we are interested in the impact of same-gender role models on the STEMM gender gap, we focus on STEMM GPA. In auxiliary analyses, we extend the outcome set and examine if potential STEMM GPA effects extend to non-STEMM GPA.

Compulsory school STEMM GPA is measured in grade 10, and is used to apply to high school. High school STEMM GPA is measured in grade 13, and is used 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 shows descriptive statistics of the educational performance and choice variables. On average, girls and boys do equally well in school as measured by GPA. However, girls are more likely to attend university, but much less likely to sort into STEMM.

3.2 Method

We estimate the following model separately for males and females:

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

where Y_i is one of the educational outcomes listed above for individual i. The variable GP_Match_i is a dichotomous variable equal to 1 if the gender of the exogenously-assigned GP matches that of the individual, and zero otherwise. The coefficient β_1 measures the effect of being exogenously-assigned to a same-gender GP in childhood, compared to being exogenously-assigned to an opposite-gender GP.[17] As the decision to swap GPs or comply with the assigned GP is endogenous, we focus on the first exogenous swap of the individuals in our sample.[18] This protects us against potential endogeneity concerns after the initial swap, but attenuates our estimates. The results should therefore be interpreted as intent-to-treat effects.

Equation (1) includes birth cohort (θ_c), previous doctor (ρ_d), municipality (π_m) and year of swap (τ_t) fixed effects. Municipality fixed effects account for systematic differences across municipalities correlated both with being assigned to a same-gender GP and the outcomes. The average size of a municipality – 12,000 individuals – is small.[19]

---

[17] There are two types of papers in the literature: papers comparing exposure to female role models to exposure to nothing (e.g. Porter and Serra 2019 and Breda et al. 2018), and papers comparing exposure to female role models to exposure to male role models (e.g. Carell et al. 2010 and Canaan and Mouganie 2019). Our paper is part of the second group, and a limitation with this approach is that it is not possible to know if female GPs positively affect girls' likelihood to study STEMM or if male GPs negatively affect it. However, all individuals require a GP and will be exposed to either a male or a female GP, and this paper provides evidence of the implications associated with assigning girls to female GPs.

[18] To understand what the results are when we ignore the endogeneity of treatment (recognizing these are only correlational), we also estimate a version of equation (1) in which the match variable is based on the gender of the first GP that the girl was observed with. These results are consistent with a story of positive selection of girls to male GPs; failing to account for the endogeneity in GP selection would bias us towards finding no effects (Table A5). Such bias may occur because parents have a belief of male GPs being more competent than female GPs.

[19] While the legal framework contains no mention of geographic matching within municipalities, there could potentially be systematic differences across areas within

The previous doctor fixed effects absorb systematic differences between the GPs that individuals were assigned to prior to a swap, such that our identifying variation is driven by girls who had the same previous GP, but that got exogenously moved to differently-gendered new GPs (one girl assigned to a male GP, and one to a female GP). This eliminates the risk of our results being confounded by characteristics of the girls' previous GPs, and increases the similarity of the girls we compare. The cohort fixed effects control for any time-invariant differences between cohorts that may be correlated both with GP-gender match and outcomes. Finally, the year of swap fixed effects account for systematic differences across years.

4 Results

4.1 Baseline Results

Panel A of Table 2 presents the effect of same-gender GP on female educational choice and performance using the four outcomes defined above: academic high school track, STEMM high school degree, compulsory school STEMM GPA, and high school STEMM GPA.[20] Column 1 shows that girls who are exposed to a female GP during childhood are 5 percentage points more likely to choose an academic program in high school (7 percent relative to the mean). The choice of an academic program in high school ensures access to higher education and greatly improves an individual's chance of obtaining a university degree.[21] Column 2 explores if GP

municipalities that are not absorbed by the municipality fixed effects. However, this concern is alleviated by the previous GP fixed effects; our results are identified off of girls who went to the same GP in the same location but then got moved to differently-gendered new GPs.

[20] 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.

[21] Table A6 shows results from equation (1) using non-STEMM academic track and non-academic track as outcomes. The result show that girls are pulled into the STEMM track from both of these groups, though more than 2/3 are coming from academic non-STEMM 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). Columns 1 and 2 show that female GPs lead girls to sort into - and graduate from - educational programs that traditionally have been underrepresented by females and are associated with larger financial returns.

The last two columns of Panel A in Table 2 examine if female GPs affect the educational performance of girls. Column 3 presents results for STEMM GPA at the compulsory level, and provides clear evidence of a positive effect of GP match on performance: the point estimate indicates that a girl who is exposed to a female GP experiences a 0.084 unit increase in STEMM GPA. This effect is 0.09 standard deviations (Table 1). As discussed in Section 2, compulsory school GPA is imperative for admission to selective high school programs, and this result is consistent with role model exposure motivating individuals to work harder in compulsory school to get accepted to more selective high school programs.

We also study if the performance effects identified in compulsory school persist as girls enter high school. Examining academic performance in high school is interesting as selection into STEMM could result either in lower academic performance (due to STEMM being a more 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 similar to the compulsory school STEMM GPA effect. This suggests that the effect of same-gender GP on educational choice does not induce worse academic performance.[22] Rather, it leads to improved educational achievement, which could be due to improved education matching and enhanced motivation.[23]

The interpretation of our estimates are complicated by the fact that they are reduced form, and that they may be attenuated by the fact that non-matched individuals may also be exposed to female GPs at some point during their childhood. We have also estimated a modified version of equation (1) in which we instrument whether or not the child had a female GP at any point between age 6 and 15 using the gender of the exogenously assigned GP as the instrumental variable. The results from this exercise return first-stage F statistics in the 800-range, and second-stage results that are approximately fifty percent larger than our baseline results (Table A8). However, we encourage caution when interpreting these results, because the first stage of this analysis is exclusively driven by girls whose only female GP is the exogenously assigned

[22] 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 Table A7. This adjustment leads to larger point estimates and smaller standard errors.

[23] Our baseline equation includes a rich set of fixed effects to limit the risk of biases arising from geography, time, and previous GP characteristics. To examine the sensitivity of our results to the choice of model specification, Figure 2 show estimates and confidence intervals from estimating Equation (1) on our main outcomes using nine different combinations of fixed effects. The figure shows that the estimates are relatively insensitive to the choice of specification, though the results are attenuated when we do not account for the previous GP and when we replace the conservative previous GP fixed effects with previous GP characteristics.

GP (compliers). Girls who would have a female GP regardless of the exogenously assigned female GP (always-takers) do not contribute to identification of the first stage. Yet, exogenous assignment to a female GP could impact education outcomes even for girls who have other female GPs, and the always-takers might thus contribute to the reduce form effect.

To examine if the same-gender GP effects for girls identified in Panel A of Table 2 extend to boys, Panel B provides results from estimating equation (1) on our boy. The point estimates are small and none are statistically significant, suggesting that boys are not affected by same-gender GPs. The results are consistent with previous literature on this topic in a classroom setting (e.g. Carrell et al. 2010). While the lack of significant effects among boys could be due to gender differences in receptiveness to role models, it could also be because the channels through which roles models operate (eradication of stereotype threats and provision of information) are less important for boys. Specifically, the percentage of boys in STEMM is substantial, and there is an abundance of public figure male STEMM role models.

Panel B of Table 2 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 find a negative effect of same-gender GP on 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 very few health problems.

Taken together, the results in Table 2 indicate same-gender GPs have a significant effect on females', but not males', educational performance and choice.[24] The education effects we

[24] An interesting and policy-relevant question with respect to our results is whether the effects are driven by increased awareness of females within the STEMM fields, or due to repeated interactions with a female in STEMM. To provide suggestive evidence, we estimate equation

identify are large, but are considerably smaller than existing differences in the outcome variables by socioeconomic factors such as mother's income, education, and marital status (Table A10), and fall within the range of the effects identified from shorter information interventions in the classroom (e.g. Breda et al. 2018; Porter and Serra 2019).[25] Due to the lack of 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 effect of same-gender role models likely differ across girls depending on the availability of same-gender role models in their families and surroundings. For example, a girl with a highly educated mother may benefit less from a female GP than a girl with a less educated mother, as many of the channels through which role models operate (providing higher aspirations, reducing stereotype threats and conveying information) are 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. These results are in Table 3, and show that the role model effects identified in Section 4.1 are driven by daughters of less educated mothers. The results highlight that same-gender role models are not only important for closing

(1) separately for those that remain with their GPs until age 15 and for those that do not remain with their GPs. These results are speculative as the decision to remain with an exogenously assigned GP is likely endogenous. The results are in Table A9, and suggest that the effects are driven by girls who remain with their GPs. Conditional on the endogeneity issue, this suggests that repeated interactions, rather than simply increased awareness, drive the effects.

[25] The male sample is smaller than the female sample, such that our power to detect effects among boys is smaller. However, the differences in estimates are substantial, while the standard errors are similar, making it unlikely that the lack of effects among boys is due to sample size.

the gender STEMM gap, but also for closing the within-gender socioeconomic STEMM gap. This suggests same-gender role models may be an important tool for improving intergenerational occupational mobility.[26] To ensure that this heterogeneity is not 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 Table A11. None of the coefficients are statistically significant, consistent with our main findings.

ii. Distributional Effects

Equation (1) estimates the average treatment effect. 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 is 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. To explore this, Table 4 show 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 4 suggest 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 slightly noisier with respect to quantiles 7 and 8 (Panel B of Table 4), the general pattern of results is similar to that for compulsory school STEMM GPA.

iii. Age-at-swap

Average treatment effects from equation (1) may also miss treatment heterogeneity across age-at-swap. Specifically, it is likely that the effect of being assigned to a same-gender GP at age 6 is different from being assigned to a same-gender GP at age 15: young children may be

[26] The fact that we observe no statistically significantly different effect by the presence of high-educated fathers suggests that the effects are not due to a general information flow, but due to information provision unique to the mother or other same-gender role models (Table A12).

differentially affected since the GP-patient interaction likely is different, and children that experience a swap at an earlier age may be exposed to the new GP for a longer period. However, this does not mean that we expect the effect to be zero among girls exposed at a later age. First, the average person spends around 3 years with the assigned GP, and the average exposure time for young girls is not much different from that of older girls. Second, individuals exposed at the end of compulsory school are preparing their high school applications, and it is possible that exposure during this critical time period is important.

To examine age heterogeneity, we estimate models of the following form:

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

where $ExogSwap_{ia}$ takes the value of 1 if individual i experienced an exogenous swap at age a, and zero otherwise, grouping individuals into three age bins: 6-9, 10-12, and 13-15. These age groups represent distinct stages of the children's educational careers: lower primary education, upper primary education, and lower secondary education. The term ($ExogSwap_{ia} \times GP\_Match_i$) is 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. These results are shown in Figure 1 for our four core outcomes.

Figure 1 implies that girls who are exogenously-assigned a female GP at an earlier age experience larger effects than girls exposed at a later age. This suggestive heterogeneity is largest when examining STEMM GPA in compulsory school, but is also noticeable with STEMM GPA in high school and STEMM high school degree. With respect to academic high school track, the effect appears constant across the different age groups. However, as shown in Table A13, which provides the full set of point estimates and standard errors of the results in Figure 1, the effects are noisy and not statistically significantly different from each other. Thus, this is only suggestive evidence regarding heterogeneity by age of swap.

It is also interesting that exposure at ages 13 through 15, when children are preparing their high school applications, is associated with statistically significant and economically meaningful effects. This is consistent with existing literature on role models in the classroom (e.g. Breda et al. 2018 and Porter and Serra 2019) that finds that information interventions have an immediate impact on female's educational choice. Having said that, the point estimate is monotonically declining with age for all but the academic high school track outcome.

4.3 Additional Outcomes

i. Non-STEMM GPA

Table 2 shows that female GPs improve STEMM GPA of female students. It is not clear that these effects lead to an improvement in overall GPA, because the improved STEMM GPA could be due to students spending less time on other subjects. To examine this, Column 1 of Table 5 shows the result from estimating equation (1) using non-STEMM GPA as the dependent variable. This result shows that the STEMM GPA effect in Table 2 is not restricted to STEMM subjects – females exposed to same-gender GPs perform better in non-STEMM subjects as well, though the effects are smaller. It is important to note that admission to high school and university depends not only on subject-specific GPA, but also on overall GPA. Thus, a high STEMM GPA is necessary, but not sufficient, for being admitted to STEMM programs.[27] The effect on non-STEMM GPA is consistent with the idea that STEMM role models motivate individuals to work harder to get accepted into the more selective programs.

ii. College Enrollment and College Major Choice

In Table 5, we examine if the effects on educational choice in compulsory school and high school persist as individuals enter college. Looking at these outcomes is important for understanding the persistence of the role model effects. While Gershenson et al. (2018) have documented effect persistence with respect to same-race teachers, we are aware of no studies

[27] This is a different setting from Carrell et al. (2010) for example, where performance in non-STEM courses don't affect the students' ability to major in STEM (conditional on passing).

that have examined this with respect to same-sex role models. We focus on two outcomes: the probability of attending university and the probability of choosing a STEMM major.

Column 3 of Table 5 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 not in motivating those at the bottom of the ability distribution.

Column 4 of Table 5 shows that female GPs increase girls' likelihood of choosing a STEMM major at university by 2 percentage point. This suggests that female role models can close the gender gap in college STEMM choice by almost 20 percent. The effect is economically large, but falls 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 a female role model raises the probability that girls enroll in selective STEM tracks in higher education by 30 percent, and Porter and Serra (2019) find that a 15-minute classroom intervention by a female economics alumni increases the likelihood that female students major in economics by 8 percentage points (a 90 percent increase).

It is possible that the college STEMM major effect is driven exclusively by treated individuals being more inclined to pursue medicine. To examine this, we estimate equation (1) using the probability of pursuing medicine, and the probability of pursuing STEM, as outcomes. The results are shown in the last two Columns of Table 5. While we lack statistical power to identify precise effects at this disaggregated level, the standard errors are smaller than the point estimates, and the results show that about half of the effect comes from enrolling in medicine. However, as the share of girls pursuing medicine is considerably smaller than the share of females pursuing STEM, the effect of enrolling in medicine is much larger as a percentage of the mean. This provides evidence on the channel through which the STEMM effect operates.

5 Balance, Falsification, Permutation and Alternative Pathways

5.1 Balance Test

The key assumption underlying our estimation strategy is that the gender of the exogenously-assigned GP is orthogonal to other characteristics that influence the educational outcomes we study. To examine this assumption, we conduct a balance test in which we regress a set of pre-determined characteristics on the gender match of the assigned GP using equation (1). If the GP gender is orthogonal to individual characteristics that may influence future educational decisions, this exercise should produce small and not statistically significant point estimates.

The results from this exercise are presented in Table 6 and support our identifying assumption, showing small and not statistically significant point estimates for each characteristic. A joint test of significance for the covariates in Table 6 further supports the identifying 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).[28]

It is worth noting that the sample used for the balance test is slightly smaller than the sample used for the main analysis due to missing information on parental characteristics for a small number of girls in our sample. To ensure that this is not biasing our results, we have also restricted the sample to only those girls for which we have all balance variables, and re-estimated our main results. Our results are robust to this sample restriction (Table A17).

5.2 GP Characteristics

---

[28] Balance test for males is in Table A14, and balance test for females without previous GP fixed effects is in Table A15. An alternative balance test is to use the out-of-sample data (girls with no GP swaps) to predict the four main outcomes with the independent covariates. We can then compare the predicted outcomes for girls with a female GP and a male GP (after a swap). These results are provided in Table A16, and show that there are no statistically significant differences between the predicted outcomes for girls with a female GP and a male GP.

If GPs are as good as randomly assigned to patients, then GP characteristics should be orthogonal to patient characteristics. The balancing test in Section 5.1 does not reject this assumption. However, a remaining issue is that GP gender may correlate with other GP characteristics such as age, confounding our estimates. To examine if non-gender GP characteristics are orthogonal to GP gender, we have examined the distribution of GP characteristics among newly assigned GPs by estimating the effect of being randomly assigned to a same-gender GP on non-gender GP characteristics. We have also estimated our main equation including controls for the non-gender characteristics of the newly assigned GP.

The non-gender GP characteristics we look at are age and nationality. These are the only non-gender GP demographic characteristics available in our data. Fortunately, we believe that they represent two of the most important non-gender characteristics, since characteristics such as education, certification and income are similar across GPs (GPs have the same level of base training, the quality of medical education is constant across the four public medical schools in the country, and the GPs are paid based on national pay scales). The results from the first exercise are in Table A18, and the results from the second exercise are in Table A19. These results support the claim that GP gender is orthogonal to other GP characteristics such as age.

5.3 Falsification Test

In addition to balance tests, we perform placebo tests in which we examine the effect of exogenous GP-matches at ages 20-25 on educational outcomes at the compulsory and high school level, and the effect of GP-matches at ages 17-20 on educational outcomes at the compulsory school level. The vast majority of people aged 20-25 have completed high school education, and the vast majority of people aged 17-20 have completed compulsory school. Thus, GP matches after age 19 should not affect compulsory and high school outcomes, and GP matches after age 16 should not affect compulsory school outcomes. Looking at the results in Table 7, all estimates are much smaller or in the wrong direction, and none are statistically significant. These results are inconsistent with the presence of biases due to non-random sorting of children to GPs of a specific gender, and support a causal interpretation of our results.[29]

5.4 Permutation Tests

Another concern is that we are simply picking up random noise, and that our results are independent of treatment assignment. To investigate this issue, we perform a series of permutation tests in which we randomly reassign treatment to GPs 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 placebo estimates. If the results in Table 2 represent true effects of same-gender GP assignment, the estimates in Table 2 should be larger than the vast majority of these simulations.

Table 8 displays our baseline estimates and the p-values obtained from the permutation exercises. The stars accompanying the estimates correspond to the level of statistical significance of our baseline estimates, and are included to facilitate the interpretation of the results. The p-values produced by the permutation exercises greatly resemble those obtained in the baseline estimation. 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.

5.5 Ruling Out Alternative Pathways

In addition to a direct role model mechanism between girls and GPs, some of the identified

[29] Another placebo test is to examine the effect of GP gender match on girls who never visited their GP. The idea behind this test is that these girls did not interact with the GP, and should therefore not be affected by her. None of the estimates obtained from equation (1) using this sample are statistically significant, and with the exception of the coefficient on academic high school track, the magnitude of these effects is very close to zero. However, it is difficult to interpret these results due to endogenous selection into the zero-visit sample, and due to the fact that assignment to same-sex GP may affect the number of times a girl visits the GP.

effects may operate through children's health and parents. First, GPs may affect how girls interacts with the health system. For example, girls may be more comfortable with female GPs, and be more likely to disclose health issues and receive treatment. It is also possible GPs are better able to relate with same-gender patients. While it is unlikely that this would affect education choices, improved health could affect educational performance, and may provide an additional pathway through which our effects operate. Second, it is also possible that part of the effects we identify operate through mothers' exposure to same-gender GPs. Specifically, mothers' interactions with the GP might expose them to female STEMM role models, which could impact their outcomes directly and children's outcomes indirectly (through improved resources at home or through access to better information via the mother).[30]

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 operate through indirect role model influences via the mother's interaction with the GP, we estimate equation (1) using a set of education, labor, and health outcomes of the mother when the child is 15.

Panel A (health-based channel) and Panel B (indirect effect through the mother) of Table 9 show these results. The results in Panel A show that GP gender match does not impact the health-related outcomes of the child. The one exception is fertility, which displays a statistically significant effect. However, the point estimate is not economically meaningful. The results in Panel B show that assignment to a same-gender GP does not affect the mother's education, labor market, and health outcomes. Taken together, the results in Table 9 suggest that the effects

[30] 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.

we identify in Section 4 are unlikely to operate through a health-based non-role model channel or an indirect role model channel via the mother's interaction with the GP.

The lack of a health-based non-role model channel is an important finding, as one of the main concerns with 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 cannot examine, it is unlikely that they are driving our findings as they would have to be uncorrelated with all the outcomes in Table 9, not subsumed by the fixed effects in equation (1), but correlated with the probability of being assigned a same-gender GP and the outcomes of interest.

6 Discussion and Conclusion

We use exogenous assignment of children to GPs to examine if female role models can reduce the gender gap in educational choice. 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 female GPs 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 20 percent. This effect is large, but within the range of effects identified from shorter information interventions in the classroom. The persistence of role model effects is interesting given the "leaking pipeline" phenomenon discussed in the literature.

In terms of mechanisms, our results show that high-ability girls with limited access to female role models at home (as measured by mother's education) drive the results. This suggests that a main channel through which the effects operate is information and the elimination of stereotype threats, enabling individuals with high underlying ability to realize their full potential. That more than two-thirds of the girls who switch into STEMM come from other relatively demanding non-STEMM academic tracks reinforces this argument. We observe no differential effect by the presence of high-educated fathers, which suggests that the effects are not due to a general information flow, but due to information provision unique to the mother or other same-gender role models. By showing that the effects are confined to those who remain with their GPs over time, we also provide suggestive evidence that repeated interactions, rather than just general awareness, matter for the realization of general role model effects.

While previous literature in economics of education has documented the value of same-gender role models, we are the first to show that role model effects in education need not involve classrooms but can arise due to everyday interactions with other potential non-teacher role models. A particularly novel feature of our study is ruling out the possibility that the effects operate through gender differences in health practices. This is interesting as the majority of role model studies in the classroom have focused on teacher-student interactions. A concern with these studies has been that the effects may simply be driven by gender differences in teaching practices, something that appears unlikely in light of the results presented here.

Another novel aspect of our study is the ability to examine the effects of role models before any educational decisions have been made and how these effects differ by age at swap. In general, the results provide suggestive evidence that earlier exposure generates larger effects (consistent with research on early childhood investments). However, the results show that exposure during the final phase of compulsory school, when students prepare their high school applications and decide which type of track to choose, also matter.

Finally, by following individuals over time and examining the persistence of the effects as individuals enter college and choose majors, we demonstrate that role model effects are long-lasting and perhaps deserving of even more policy attention. Specifically, the findings 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.

In contextualizing our results, it is important to note that Norway ranks as one of the most gender equal countries in the world (WEF 2018). However, it is difficult to speculate whether we would expect larger or smaller effects in countries with greater gender inequality. On the one hand, female role model effects may be larger in countries where female role models are rare. On the other hand, same-gender role models may interact positively with a country that is more progressive on gender, such that the effects are smaller in countries where female role models are rare.[31] Nevertheless, we believe that the general pattern of our results is relevant to a large number of countries and settings. Norway is still far from reaching perfect gender equality, and similar to other OECD countries, the gender-divided labor market represents the largest challenge to reaching this goal. For example, females are more likely to work part-time (36.8 percent compared to 12.5 percent), more likely to work in the public sector (70.1 percent of public employees are females), less likely to hold leadership positions (35.3 percent of individuals in leadership positions are females), and the annual income is only 70 percent of the income of males (SSB 2017). In terms of education, females dominate nursing, welfare and teaching degrees at the university and are less likely to do natural science, leading to gender-segregated occupational patterns (SSB 2018). Thus, we cannot speak to how the effects would differ across countries, but we believe that our results generalize to the majority of OECD countries that are actively working towards gender parity and face similar labor market barriers.

References

[31] This ambiguity is further exemplified by the gender-equality paradox, in which countries with greater equality have been observed to have a more unequal gender distribution in STEM fields. For example, the share of female tertiary graduates in science in Norway is lower than that in countries that are commonly considered less gender equal, see UNESCO (2015).

Beaman, Lori, Esther Duflo, Rohini Pande, & Petia Topalova (2012). "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India" *Science* 335(6068): 582-586

Bettinger, Eric, & Bridget Long (2005). "Do faculty serve as role models? The impact of instructor gender on female students" *American Economic Review* 95(2): 152-157

Black, Sandra, Paul Devereux, & Kjell Salvanes (2005). "Why the apple doesn't fall far: understanding intergenerational transmission of human capital" *American Economic Review* 95(1): 437-449

Breda, Thomas, Julien Grenet, Marion Monnet, & Clémentine 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*

Canaan, Serena, & Pierre Mougaine (2019). "Female Science Advisors and the STEM Gender Gap" *IZA Working Paper No. 12415*

Card, David, & Abigail Payne (2017). "High school choices and the gender gap in STEM" *NBER Working Paper No. 23769*

Carneiro, Pedro, & James Heckman (2003). "Human Capital Policy." In Heckman J., Krueger A., *Inequality in America: What Role for Human Capital Policy?* (MIT Press)

Carrell, Scott, Marianne Page, & James West (2010). "Sex and science: how professor gender perpetuates the gender gap" *Quarterly Journal of Economics* 125 (3): 1101-1144

Ceci, Stephen, Donna Ginther, Shulamit Kahn & Wendy Williams (2014). "Women in academic science: a changing landscape" *Psychological Science in the Public Interest* 15(3): 75-141

Chung, Kim-Sau. (2000). "Role models and arguments for affirmative action" *American Economic Review* 90(3): 640-648

Dee, Thomas. (2005). "A teacher like me: does race, ethnicity, or gender matter?" *American*

Economic Review 95(2): 158-165

Dee, Thomas. (2004). "Teachers, race, and student achievement in a randomized experiment" *The Review of Economics and Statistics* 86(1): 195-210

Eble, Alex, & Feng Hu (2017). "Stereotypes, role models, and the formation of beliefs" *Mimeo.*

Elango, Sneha, Jorge Garcia, James Heckman & Andrés Hojman (2015). "Early Childhood Education." In Mofitt R. (editor), *Economics of Means-Tested Transfer program in the United States* (University of Chicago Press)

Firpo, Sergio, Nicole Fortin, & Thomas Lemieux (2009). "Unconditional quantile regressions" *Econometrica* 77 (3): 953-973

FOR (2018). *Forskrift om fastlegeordning I kommunene* (Helse- og omsorgsdepartementet, Norge). Accessed via https://lovdata.no/dokument/SF/forskrift/2012-08-29-842

Gershenson, Seth, Cassandra Hart, Joshua Hyman, Constance Lindsay & Nicholas Papageorge (2020). "The Long-run Impacts of Same-Race Teachers" *NBER Working paper No. 25254*

Griffith, Amanda (2014). "Faculty gender in the college classroom: does it matter for achievement and major choice? *Southern Economic Journal* 49(3): 730-749

Hoffmann, Florian, & Philip Oreopoulos (2009). "A professor like me: the influence of instructor gender on college achievement" *Journal of Human Resources* 44(2): 479-494

Hyde, Janet (2005). "The gender similarities hypothesis" *American Psychologist* 60: 581-92

Kofoed, Michael, & Elizabeth 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, Victor, & Edith Sand (2018). "On the origins of gender gaps in human capital: short and long term consequences of teachers' biases" *Journal of Public Economics* 167: 263-79

Table 1: Descriptive Statistics of Individuals in Sample

	Girls		*Boys*	
	Mean	St.Dev	Mean	St.Dev
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: 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

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Panel A: Girls				
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
Panel B: Boys				
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 girls (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 3: 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 4: 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 5: Effect of same-gender GP, additional outcomes

	Compulsory school non-STEMM GPA	High school non-STEMM GPA	College enrollment	College STEMM enrollment	College Medicine enrollment	College STEM enrollment
Same-gender GP	0.040*	0.067**	0.005	0.022*	0.011	0.009
	(0.023)	(0.028)	(0.017)	(0.013)	(0.008)	(0.011)
Mean	4.297	4.384	0.754	0.077	0.019	0.059
Observations	8617	8678	8680	8680	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 6: 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 7: Placebo Tests; Post high school GP swaps

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Panel A: All girls who were subject to their first exogenous GP swap between age 20 and 25				
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
Panel B: All girls who were subject to their first exogenous GP swap between age 17 and 20				
Same-sex GP	-0.007		-0.028	
	(0.010)		(0.023)	
Mean	0.741		4.199	
Observations	16868		14862	

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 who were subject to their first exogenous GP swap between age 20 and 25. Sample in Panel B includes all girls who were subject to their first exogenous GP swap between age 17 and 20. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level and *** denotes significance at the 1 percent level.

Table 8: P-values of permutation test

	High school academic track	High school STEMM degree	Compulsory school STEMM GPA	High school STEMM GPA
Baseline estimate	0.052***	0.039**	0.084**	0.109***
% less than baseline	0.010	0.053	0.053	0.020
Number of replications	300	300	300	300

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. The table shows the proportion of times the estimates from the permutation tests are smaller than the baseline estimate. We run 300 simulations in which we randomly assign GPs to children. 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. The stars accompanying the estimates correspond to the level of statistical significance of our baseline estimates, and are included to facilitate the interpretation of the results. The stars accompanying the estimates correspond to the level of statistical significance of our baseline estimates, and are included to facilitate the interpretation of the results. * 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

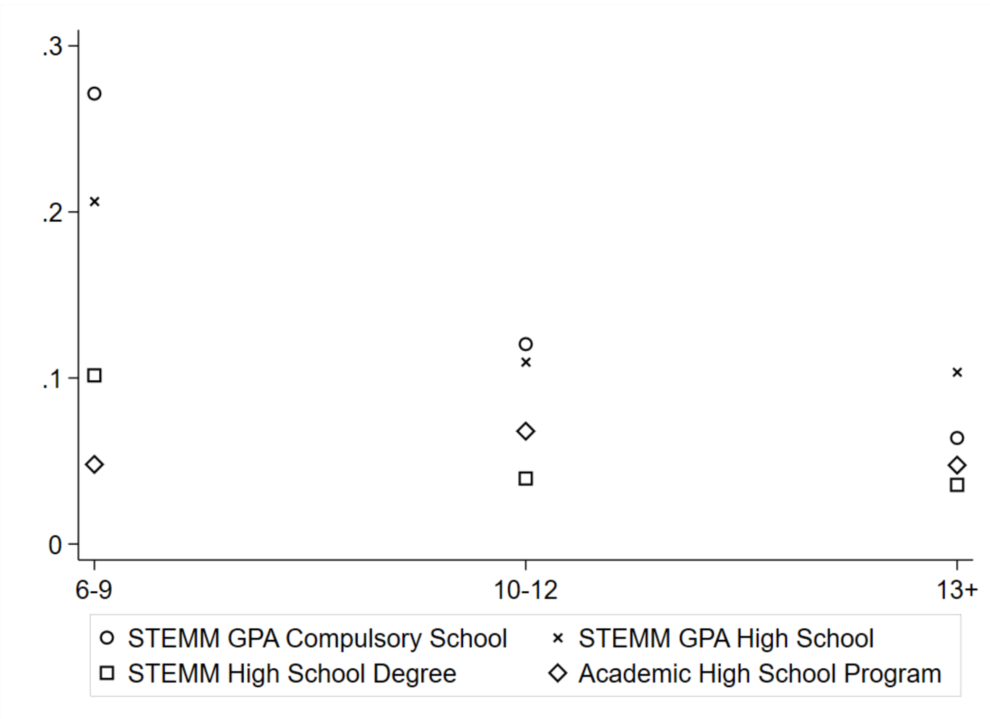
	Number of GP visits	Still with assigned GP	Number of diagnoses	Mental health diagnosis	Birth control visits	Fertility
Same-gender GP	-0.009	-0.014	-0.039	-0.000	0.002	-0.000**
	(0.070)	(0.019)	(0.028)	(0.000)	(0.005)	(0.000)
Mean	1.166	0.613	2.241	0.078	0.069	0.000

Panel B: Potential indirect effects through mother

	Years of education	Total income	Not in labor force	Number of diagnoses	Mental health diagnosis	Birth 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

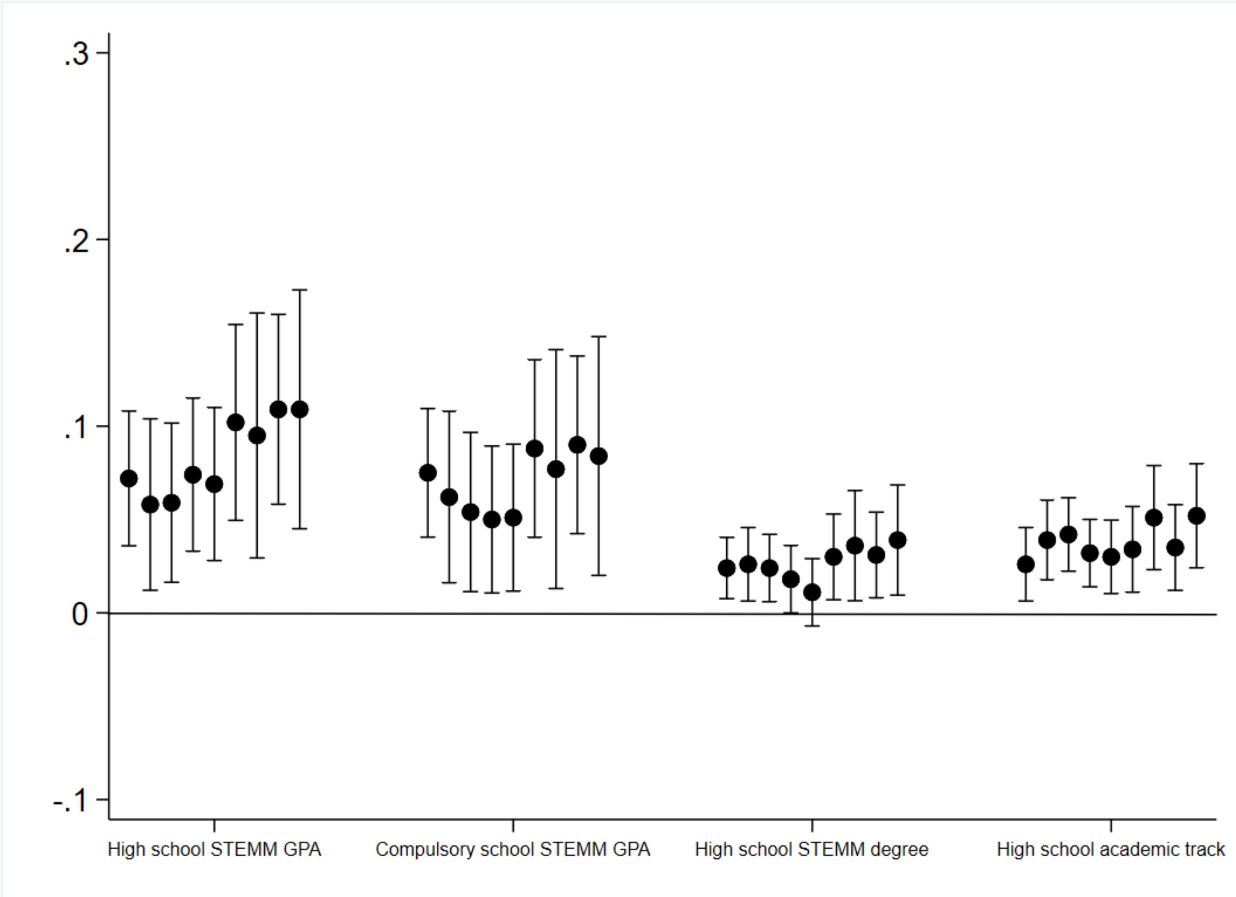
Notes: The table shows the β_1 coefficients obtained through estimation of equation (1) as described in the text: $y_i = \alpha + \beta_1 GP\_Match_i + \tau_t + \pi_m + \theta_c + \rho_d + \epsilon_i$. y_i denotes the outcome, and each estimation includes municipality (π_m), year of swap (τ_t), birth year (θ_c) and previous GP (ρ_d) fixed effects. Standard errors are clustered at the GP. Sample in Panel (A) includes all girls born between 1988 and 1996 who were subject to an exogenous GP swap prior to age 15. Sample in Panel (B) include all mothers to girls who were born between 1988 and 1996 and exposed to an exogenous GP swap prior to age 15. Outcomes are measured when the girls are 15 years old. Income is reported in log form. * significance at the 10 percent level, ** significance at the 5 percent level and *** significance at the 1 percent level.

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



Notes: Authors' estimation of a modified version 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 the age range denoted on the x axis. As Equation (2) contains both year-of-swap and birth cohort fixed effects, the equation does indirectly control for the main effects for age-at-assignment as well (as this is a linear combination of year-of-swap and bith cohort). 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: Estimated effects with varying fixed effects



Notes: Authors' estimation of a modified version 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. 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. The specification which replaces the previous GP fixed effects with previous GP characteristics include GP age, gender, and Norwegian-born status.

For each of the four outcomes, the following specification legend applies:

Specification 1: No fixed effects; Specification 2: Cohort-by-municipality fixed effects; Specification 3: Cohort fixed effects and municipality-specific trends; Specification 4: Cohort, municipality, and year fixed effects; Specification 5: Cohort, municipality, and year fixed effects, and controls for previous GP characteristics; Specification 6: Cohort and previous GP fixed effects; Specification 7: Cohort, municipality, and previous GP fixed effects; Specification 8: Cohort, year, and previous GP fixed effects; Specification 9: Cohort, municipality, year, and previous GP fixed effects