Cultural Norms and Neighbourhood Exposure: Impacts on the Gender Gap in Math*

Sanna Ericsson[1]

June 8, 2021

Abstract

This paper investigates the interaction between cultural norms and neighbourhood characteristics, in the context of the gender gap in math. Using high-quality Swedish administrative data, I estimate the effect of mothers' country-of-origin gender norms on the gender gap in math among second-generation immigrant siblings, and exploit a refugee placement policy to obtain random variation in regional characteristics. I find that the sibling gender gap in math increases with cultural gender norms, such that girls with mothers from more gender-traditional cultures perform worse compared to their brothers, but that local labour market gender equality can completely mitigate this cultural norm effect.

Keywords: cultural gender norms, math gender gap, epidemiological approach, refugee placement policy, family fixed effects

JEL codes: I21, I24, J15, J16, Z13

---

*I would like to thank my advisers, Petter Lundborg and Therese Nilsson, for their support and helpful insight during the process of writing this paper. I also thank Andreas Kotsadam, Nùria Rodrìguez-Planas, David Figlio, two anonymous referees, and the conference participants at Lund University, Copenhagen University, IFN and the Linneaus University for their useful comments. The data used in this paper comes from the Swedish Interdisciplinary Panel (SIP), administered by the Centre for Economic Demography, Lund University, Sweden. Corresponding author: Sanna Ericsson (sanna.ericsson@nek.lu.se).

[1]Lund University, Department of Economics and Centre for Economic Demography.

1. Introduction

Girls and boys differ in terms of their educational achievement. In most cases, a gender gap in education implies one that favours girls, as girls outperform boys along most educational dimensions (DiPrete and Buchmann, 2013; Lundberg, 2020). However, one exception is math. Girls systematically perform worse than boys on math tests, particularly at the top of the performance distribution (Bedard and Cho, 2010; Pope and Sydnor, 2010). The gender gap in math has been shown to correlate with gender equality and norms regarding women's role in society, which suggests that social forces may, at least in part, be driving the differential performance of boys and girls (Guiso et al., 2008; Pope and Sydnor, 2010; Nollenberger et al., 2016; Rodríguez-Planas and Nollenberger, 2018).[1]

But from where do individuals perceive the norms that may impact their educational outcomes? Several studies find that culture and historical traditions have a significant impact on both the attitudes and the economic outcomes of individuals today (see e.g. Fernández, 2011; Alesina et al., 2013; Nollenberger et al., 2016; Finseraas and Kotsadam, 2017; Rodríguez-Planas and Nollenberger, 2018; Figlio et al., 2019; Dahl et al., 2020), which demonstrates that cultural norms are important determinants of our behaviour and outcomes. However, another strand of the literature documents significant behavioural impacts of neighbourhood exposure and peer effects (see e.g. Dahl et al., 2014; Chetty et al., 2016; Chetty and Hendren, 2018; Nicoletti et al., 2018; Olivetti et al., 2020), which indicates that our on-going exposure to institutions, peers, and other surrounding factors may also be an important determinant of our behaviour and outcomes. Taken together, a large literature shows that both family culture and neighbourhood characteristics affect our economic, including educational, outcomes.[2] It is likely that these two channels do

[1] Geographical variation in the math gender gap indicates that it is not driven solely by innate ability differences between boys and girls (Bedard and Cho, 2010). In addition, the gender gap in math does not exist at the point of school entry, but rather emerges over time when children are socialized into school (Fryer Jr and Levitt, 2010).

[2] The distinction between cultural values and neighbourhood exposure is similar to the framework developed by Bisin and Verdier (2011), who contrast vertical and horizontal transmission of norms. Vertically transmission of norms occurs within the family, from parents to children, and happens if parents believe that their children will benefit from certain cultural traits. Horizontal transmission denotes

not operate independently of each other, however, there is limited empirical evidence combining the two channels.

This paper investigates the interaction between cultural norms and neighbourhood characteristics, in the context of their impact on the gender gap in math. I ask two research questions: first, is there an effect of cultural gender norms on the gender gap in math, and second, to what extent can this effect be mitigated by surrounding neighbourhood gender equality? This way, I explore the effect of cultural norms on individuals' behaviour (measured by their educational outcomes), and I investigate how this effect interacts with exposure to neighbourhood characteristics and peers.

The main empirical challenges associated with estimating the effect of cultural gender norms is the possible correlation of cultural norms with institutional settings and with unobserved individual characteristics. The the correlation between cultural norms and institutional settings is problematic as institutional settings likely affect educational outcomes (Alesina and Giuliano, 2015). To account for this and isolate cultural gender norms from formal institutions, I estimate the effect of gender norms in mothers' countries of origin on the gender gap in math among second-generation immigrants. Second-generation immigrants, in this context, are all born and raised in Sweden, and encounter the same formal institutions, but potentially differ in their cultural heritage. Assuming that parents transmit norms to their children (through intergenerational transmission and through exposure to parents' peers from the same culture), and that these norms differ systematically depending on the parents' source country, second-generation immigrants provide the ideal experiment to isolate the effect of cultural norms from the effect of formal institutions. This method is commonly referred to as the epidemiological approach (see e.g. Fernández, 2011).

Cultural norms are also likely correlated with other parental characteristics that could affect the educational outcomes of the children. To account for this correlation, I follow

the socialisation of norms that takes place within a community context, where norms are transmitted from peers and surroundings. However, the culture/neighbourhood distinction noted above is broader compared to that of Bisin and Verdier (2011), as the cultural (i.e. "vertical") channel includes also parents' peers and networks (sharing the same cultural beliefs), and the neighbourhood (i.e. "horizontal") channel includes not only the effect of norms, but also that of more formal institutions.

Finseraas and Kotsadam (2017) and compare the gender gap in math only between opposite-sex siblings in a family fixed effects model. The family fixed effects control for everything that affects siblings equally, including everything that correlates with source-country norms but that is unrelated to gender. By construction, the variation that remains is the *gender-specific* component of the cultural norms that affects opposite-sex siblings differently, i.e. gender norms. By comparing my within-family estimates to those found in the previous literature, I am able to investigate whether the standard epidemiological approach that do not rely on within-family comparisons creates biased estimated of cultural norm impacts.

To answer the second research question, I investigate the extent to which the cultural gender norm effect can be mitigated by gender equality of the family's neighbourhood, in this setting defined as the local labour market area. The main empirical challenge of estimating the effect of neighbourhood characteristics is that there is selection in where people choose to live. Families will choose to reside in places that have certain desirable characteristics, and, in doing so, they themselves contribute to these characteristics. To account for this selection, and obtain exogenous variation in local labour market characteristics, I exploit a refugee placement policy. Under this policy, government officials assigned asylum-seeking immigrants their initial location of residence. As these immigrants were not free to chose where they would be placed, their initial location of residence is independent of unobserved individual characteristics.

I rely on high-quality Swedish administrative data on the universe of ninth-grade students who took the national standardised math test between 2004–2012. I proxy cultural gender norms with the Gender Gap Index from the World Economic Forum's Gender Gap Report, which reflects women's economic and political opportunities, education, and health in the mothers' source countries (Schwab et al., 2018). I proxy neighbourhood gender equality with female-over-male labour force participation ratios of the mothers' assigned local labour market areas.

I show that increasingly traditional cultural gender norms increase the sibling gender gap in math, such that girls whose mothers come from more gender-traditional countries perform worse relative to their brothers.[3] A one-standard-deviation increase in cultural gender norms (i.e. towards more traditional norms) increases the sibling gender gap in math by 14% of the cross-country standard deviation, in favour of boys. This result shows that while I can replicate the results of Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018) in terms of the sign of the effect, my effect size is smaller in magnitude. However, comparing estimates with and without family fixed effects reveals that the smaller effect size is likely not due to bias in the standard epidemiological approach.

Having established that traditional gender norms have a detrimental effect on girls' math performance relative to their brothers', I extend the analysis to investigate the possible mitigation by neighbourhood exposure. Importantly, I show that gender equality of the local labour market area can completely mitigate the negative cultural gender norm effect. This result suggests that even though the sibling gender gap in math increases with culture of traditional gender norms, this increase is smaller for siblings whose mothers were placed in more gender-equal labour market contexts. Comparing estimates with and without family fixed effects demonstrates that specifications that compare the math performance of boys and girls with mothers who were placed in different local labour market areas may be biased, thus, within-family comparisons are crucial to capture the gender-specific effect of regional characteristics. Taken together, my results show that while cultural gender norms play an important role for the gender gap in math, they are not immune to the influence of neighbourhood exposure.

The first research question of my study relates to the growing body of literature on gender norms and educational outcomes.[4] Several studies show that girls' relative educational performance correlates positively with gender equality, both across countries and US states (Guiso et al., 2008; Fryer Jr and Levitt, 2010; Pope and Sydnor, 2010). One important issue here is the risk of reverse causality, as these studies cannot

[3]Gender-traditional countries are defined as those with a lower Gender Gap Index score, i.e. worse economic, political, education and health outcomes for women.

[4]In addition, my paper contributes to the literature on the effects of source country culture. Using the standard epidemiological approach, studies show that source country culture affects gender roles, women's work and fertility, social trust, political regulation, domestic violence, corruption, migration etc. See Fernández (2011) for a literature review.

determine whether girls perform better because of increased gender equality or whether high-performing girls grow up to themselves contribute to increased gender equality. Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018) account for this reverse causality, as they estimate the effect of cultural gender norms (proxied by the World Economic Forum's Gender Gap Index) on the gender gaps in math, reading and science, using PISA data on the test scores of second-generation immigrants. The authors find that the educational gender gaps between boys and girls vary with cultural gender norms, such that girls with more gender-traditional mothers perform worse in all three subjects relative to boys with mothers from the same source country.

Dossi et al. (2021a) show that girls who grow up in families that exhibit a preference for boys perform worse on standardised math tests, when compared to girls growing up in other families. Furthermore, the authors use survey data to show that mothers' attitudes regarding women's role in society are transmitted to the next generation, and that they correlate with girls' performance in mathematics but not with boys'. Dossi et al. (2021b) build on these results, and demonstrate that gender role norms only can explain the lower performance of girls in mathematics for White and relatively affluent families.

Dahl et al. (2020) show that birthright citizenship for immigrant girls lowers their life satisfaction and self-esteem, a result they argue is due to the conflicting identities of German citizenship and parents' traditional cultural norms. They do not find the same effect for boys, which indicates that girls are pushed comparatively harder by their parents to conform with traditional gender norms, whereas boys are allowed to take advantage of the citizenship opportunities. Finally, Aldén and Neuman (2019) show that cultural gender norms affect the probability of girls choosing STEM, or other male-dominated fields, as their major in high school or university.[5]

The second research question of my study relates to the literature on the effects of neighbourhood characteristics on children's educational outcomes. Two well-identified

[5]Unrelated to gender norms, but related to the gender gap in educational outcomes, Figlio et al. (2019) show that family disadvantage is disproportionally detrimental to the educational outcomes of boys relative to girls, and that this result is robust to specifications within neighbourhoods, schools and families.

papers by Chetty et al. (2016) and Chetty and Hendren (2018) show that neighbourhoods shape children's earnings, college attendance rates, marriage and fertility patterns, and that the effect increases linearly with time of exposure. In addition, they show that boys and girls are affected differently in neighbourhoods that are particularly beneficial for either gender.[6] Damm and Dustmann (2014) exploit random variation in neighbourhood exposure, caused by a refugee placement policy in Denmark, to show that children who were placed in high-crime neighbourhoods are more likely to themselves commit crimes. They find that social interaction is the key channel through which this neighbourhood exposure effect operates. Bratsberg et al. (2020) exploit a refugee resettlement policy in Norway to show that initial neighbourhood exposure influences subsequent political participation of immigrants.

My study also relates to the literature on peer effects, where several papers show that the behaviour of peers affect individuals' economic outcomes, including labour market decisions (Maurin and Moschion, 2009; Dahl et al., 2014; Mota et al., 2016; Nicoletti et al., 2018; Olivetti et al., 2020). Finally, my study contributes to the literature that use the Swedish refugee placement policy to evaluate neighbourhood exposure effects, such as, for example, Edin et al. (2003) who show that living in ethnic enclaves improves immigrants' labour market outcomes, and Åslund et al. (2011) who show that immigrants' school performance is increasing with the number of highly educated individuals of shared ethnicity residing in the same neighbourhood.

To the best of my knowledge, my study is the first to estimate the interaction between cultural norms and neighbourhood characteristics. Thus, the novel and important contribution of my paper is that I merge the literatures on cultural norms, neighbourhood exposure and educational outcomes. A secondary contribution of my paper is that I replicate and extend the results of previous studies on cultural gender norms and the gender gap in math, using an extended version of the epidemiological approach that contrasts

[6] Chetty et al. (2016) estimate the effect of neighbourhood exposure using variation caused by the Moving To Opportunity experiment. They find that only children who moved when young experience positive effects of moving to a low-poverty neighbourhood. Chetty and Hendren (2018) exploit variation in age of children when families move, finding that the outcomes of children who move converge towards the outcomes of the children in the new neighbourhood at a rate of 4% per year of exposure.

the outcomes of opposite-sex siblings. This way, I am able to investigate any potential bias caused by the standard epidemiological approach.

The remainder of the paper is organised as follows. Section 2 provides institutional background on the Swedish marking system and the refugee placement policy. Section 3 describes the data, after which Section 4 outlines the empirical strategy. Section 5 presents the results for both research questions, and further investigates potential mechanisms and heterogeneous effects. Section 6 ensures that my results are robust. Finally, Section 7 concludes this paper.

2. Institutional background

2.1. The Swedish marking system and the national standardised tests

Ninth grade is the last year of mandatory schooling in Sweden, and students at this level are between 15 and 16 years old.[7] Students take 16 different subjects and receive a final mark for each subject. Marks are goal-oriented and not relative. The marks during this time period are IG (fail), G (pass), VG (pass with distinction) and MVG (pass with special distinction), which correspond to 0, 10, 15 and 20 points, all of which count towards the final total mark. The final total mark is the sum of all subject marks, with a maximum of 320 points.

The national standardised tests are issued in math, Swedish, English, social science and natural science. These tests, which students take during the spring semester of the ninth grade, are nationally standardised and mandatory for all students. The tests are developed to give all students an equal opportunity to demonstrate their knowledge, and act as means of supporting the teacher in making marking more fair. Students receive a mark on each subject test, which weigh heavily on the final mark for that subject.

The national standardised test in math consists of about 40 "pass-level" questions and about 35 "pass with distinction-level" questions. Students receive a test result, which is the sum of all correctly answered questions on each level, and an overall mark for the test.

[7] Mandatory school starting age in Sweden is age 6.

2.2. The Swedish refugee placement policy

The refugee placement policy was introduced in 1985 and lasted until 1994. Its aim was to relieve pressure on the larger cities in times of large refugee inflows. By placing the asylum seekers in those municipalities with the most suitable reception characteristics, the government hoped to speed up the integration process.[8]

During the policy period the Immigration Board assigned asylum seekers to their initial municipality of residence. Initially 60 municipalities participated in the policy, but due to the increased inflow of asylum seekers during the late 1980's the number of municipalities involved increased until 277 of the total 284 municipalities were participating. The strictest application of the policy was implemented during 1987 to 1991, when the assignment rate was almost 90% and the asylum seekers had very little ability to influence where they were assigned. For this reason, I focus my analysis on the refugees who arrived during the period 1987–1991.

The asylum seekers were placed in refugee centres while they waited for a decision from the immigration authorities. Mean duration before receiving a residence permit was between three and twelve months. While the assignment process did take the asylum seekers' preferences into account, most applied for residence in Stockholm, Gothenburg or Malmö. The Swedish housing market was booming at the time, and housing opportunities in these locations were very scarce; in practice this meant that the immigrants had very little influence over their placement.

When the number of applicants exceeded the number of available slots, municipal officers had the opportunity to choose which asylum seekers would be offered a residence permit. However, all selection was based on observable characteristics, as there was no interaction between the asylum seekers and the municipal officers. Priority was given to individuals who had attained higher education and to those who spoke the language of some of the residing immigrants. Furthermore, housing availability was dependent on family size. Edin et al. (2003) argue that municipality assignment can be viewed as

[8]The information in this section is obtained from Edin et al. (2003) and Åslund et al. (2011), who in turn base their information regarding the practical implementation of the policy on interviews with Immigration Board placement officials.

random, conditional on observable immigrant characteristics.

However, Nekby and Pettersson-Lidbom (2017) identify some important empirical challenges regarding the use of the refugee placement policy as a proxy for random placement. They argue that the municipalities had more to say in the placement process than what had been previously understood. Refugees could not be placed without consent from the municipality, and the placement within a municipality was carried out by local municipality immigration agencies. For example, the municipality of Bollnäs did not place immigrants in the first available apartment, but rather, waited until housing opened up in areas with few social problems. These types of decisions were not nationally standardised and could differ sharply between municipalities. As a result, there is a risk that municipality characteristics may correlate with placement and subsequent treatment, and therefore also with the outcomes of the asylum seekers. I address these empirical challenges in Section 4.2.

Following initial assignment, there were no restrictions on moving to a different location, provided that the immigrants themselves could find housing. Leaving the assigned municipality did not affect the welfare payments; the main cost of moving was delayed enrolment in Swedish courses.

3. Data

I rely on data from The Swedish Interdisciplinary Panel, which is administered by the Centre for Economic Demography. This is a two-generational dataset consisting of merged administrative registers. Using unique parental identifiers, I identify the parents and siblings of each individual in the data. The data contain information on the national standardised test results from years 2004–2012. Students take the tests when they are 16 years old; thus, my study contains cohorts born between 1988 and 1996. I restrict my sample to second-generation immigrants, which I define as individuals born in Sweden whose mothers were born in another country.[9]

[9] As a robustness check I use a sample of second-generation immigrants defined by having a foreign-born father; the results can be found in Online Appendix Tables OA2–OA4.

Table 1 presents the summary statistics for five sample specifications: column (1) contains the total population of students taking the national standardised math test, which will act only as a reference point to the following samples. Column (2) contains the sample used for the first research question (henceforth the RQ1 sample). In regressions that rely on within-family comparisons only families with at least two siblings will be included in the analysis, and column (3) present the summary statistics for this sub-sample of the RQ1 sample.[10] Column (4) contains the sample used for the second research question (henceforth the RQ2 sample), and column (5) present the summary statistics for the sub-sample of families in the RQ2 sample that has more than one child in the family.

3.1. Sample restrictions for RQ1

My sample consists of second-generation immigrant students, who all took the national standardized math test. To be a part of my sample, the foreign-born mother must have been born in a country I can identify in the data. For reasons of anonymity, Statistics Sweden only report the source continent for any immigrant who comes from a country from which Swedish receive few immigrants. Thus, I can only identify the source countries from which Sweden received large immigrant flows, and if a mother immigrated from a country Sweden received very few immigrants from she is excluded from this study. Nevertheless, I can identify the source country for mothers born in 20 different countries, Sweden excluded, and 75% of all immigrating mothers in my data come from one of these identifiable countries. Thus, the immigrants I can identify are highly representative of immigrants to Sweden.

A second restriction is that the source country need to be included in the World Economic Forum's Gender Gap Report in order to have a value for the Gender Gap Index. Of the identifiable countries only Somalia is missing from the Gender Gap Report and therefore excluded from the analysis. These restrictions leaves me with a total of 19 different source countries.

[10]Singleton observations are dropped in regression analysis using the Stata command -reghdfe-, so all single-child families are automatically excluded in models with family fixed effects.

Table 2 lists the source countries I can identify, and column (6) lists the frequency of children with mothers born in the various source countries of my sample. Because I define a child as a second-generation immigrant based on their mother's country of birth, the children of my sample all have mothers born in one of the 19 foreign countries I can observe, but their fathers can be missing from the data or born anywhere, including Sweden or one of the unobserved foreign countries. Siblings are defined as children who share both biological parents, or who share mother and all have missing information for the father.[11]

Column (2) of Table 1 presents the summary statistics for this sample of second-generation immigrants. There are 75,780 children in this sample, of which 39,007 children have at least one sibling and will therefore be included in the analysis with family fixed effects. Compared to the full population of ninth grade students in column (1), the second-generation immigrants receive lower final marks, score lower on the math tests and have a larger average gender gap in math. In addition, they are more likely to live in households with below-median income, and their mothers are less likely to have some university education. Apart from the mechanic difference in family size, the sub-sample of children with at least one sibling in column (3) do not differ from the sample in column (2).

3.2. Sample restrictions for RQ2

To answer the second research question, I impose additional sample restrictions, which creates a subset of the sample for the first research question. I require mothers to have immigrated to Sweden sometime between 1987–1991, the years in which the refugee placement policy was active. Unfortunately I cannot observe refugee status directly, but, following Edin et al. (2003) and Åslund et al. (2011), I assign refugee status using country of birth. The Swedish Migration Board lists the countries from which Sweden received asylum seekers between 1984–1999 (Migrationsverket, 1999). Anyone immigrating from

[11]In total, 52% of the children have parents born in the same foreign country, and 77% of those with parents from different countries have a father born in Sweden. As a robustness check I exclude all children with a Swedish father, and this does not change my results.

one of the countries on this list is defined as an asylum seeker in my sample. I exclude mothers whose spouses immigrated before them, as family immigrants were not subjected to the placement reform. This definition of refugee status leaves me with 9 different source countries. Column (6) of Table 2 shows the frequency of children with mothers coming from these countries as asylum seekers.[12]

Column (4) of Table 1 presents the summary statistics for the children whose mothers were subjected to the refugee placement policy. The total number of observations is 11,773, of which 6,102 children have at least one sibling and will therefore be included in the analysis with family fixed effects. Compared to the full population of ninth grade students in column (1), and to the second-generation sample in column (2) and (3), the children with mothers affected by the refugee placement policy receive lower final marks and math scores. The households are less likely to have a household income above median, and the mothers are less likely to have a university education, and they immigrated from source countries with more traditional gender norms, as measured by the GGI score. Apart from the mechanic difference in family size and a small difference in mother's education level, the sub-sample of children with at least one sibling in column (5) do not differ from the sample in column (4).

3.3. Dependent variable: standardised math test score

The main outcome variable is students' test scores on the national standardised math test. The raw test scores range between 0 to 75; about 40 of these are pass-level points and about 35 are pass with distinction-level points. A significant benefit of using the math test score as the outcome variable is that it contains more, and continuous, variation than the final marks, which can only take one of four values. As the tests differ slightly over the years, I standardise the scores by year to obtain a mean of zero and a standard deviation of one.

Column (1) of Table 2 demonstrates the average gender gap in math for students with

[12]In total, 80% of the children in this sub-sample have parents born in the same foreign country, and 45% of those with parents from different countries have a father born in Sweden. As a robustness check I exclude all children with a Swedish father, and this does not change my results.

mothers from each source country of my study.[13] The gender gap in math is negative (i.e. boys outperform girls) for all source countries except Czechoslovakia, the United States and Vietnam. For the full sample, girls' test scores are on average 6–7% of a standard deviation lower than boys' test scores.

3.4. Independent variable: cultural gender norms

To proxy gender equality in the mothers' source country I follow Guiso et al. (2008), Fryer Jr and Levitt (2010), Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018), and use the Gender Gap Index score (henceforth GGI) from the World Economic Forum's Gender Gap Report (Schwab et al., 2018). The GGI measures the relative position of women in society, and is an unweighted average of four subindex scores. The subindexes are 1) the economic participation and opportunity subindex, which is based on gender gaps in labour force participation, earnings, and holding certain job positions, 2) the educational attainment subindex, which is based on the gender gaps in literacy and educational enrolment, 3) the health and survival subindex, which is based on sex ratios at birth and the gender gap in healthy life expectancy, and 4) the political empowerment subindex, which is based on women in parliament, in ministerial positions or as head of state. All indexes, including the combined GGI score, range from 0 to 1, with a higher value indicating a better position for women in society.

The information on the GGI is available from 2006 and onwards. In 2006 a total of 115 countries participated and this number have been rising since then. 2018 is the first year that includes information on all source countries available in my sample, therefore, my main analysis is based on the GGI score of 2018. To use contemporaneous measures of gender norms rather than those observed when parents migrate is common practice in the literature (see e.g. Nollenberger et al., 2016; Rodríguez-Planas and Nollenberger, 2018), and it is reasonable to expect that country-level norms about women's role in society evolve slowly over time. Using the GGI of 2018 goes even further and measures cultural norms with *future* values of gender equality. The reason to use GGI from 2018

[13] The gender gap in math is obtained from estimating linear regressions of standardised math score on a female indicator for each source country of the study.

is to keep as many source countries as possible in the analysis, moreover, as the GGI of 2018 is over 90 % correlated with all previous GGI scores, the impact of using GGI of 2018 instead of GGI from previous years is small. Nevertheless, I ensure that my results are robust to using the source-country GGI score from 2010 instead of 2018.

I define the source-country gender norms as $1-GGI$, such that the measure increases as the source-country GGI score decreases; thus, the measure increases as gender norms become more traditional ("traditional" will be used in the text to denote cultures that are less favourable towards women). Column (2) of Table 2 present the source-country level of $1-GGI$. Column (3) and (4) of Table 2 display source-country GDP/capita (measured in current USD) and gross enrolment in secondary school, which I use as country-level controls in all estimations.[14] The bottom panels of Table 2 present the total number of observations, mean and cross-country standard deviations for each country-level variable, for both the RQ1 and the RQ2 sample.

As robustness checks I test the sensitivity of my results to the measure used as a proxy for cultural gender norms. Appendix Table A1 present the source-country level, mean and standard deviations of all these alternative norm measures.[15] Furthermore, Appendix Table A2 present the correlation of all country-level variables. The purpose of the $1-GGI$ proxy is to capture variation that reflects source-country norms regarding the relative importance of boys' and girls' educational achievements. Reassuringly, Appendix Table A2 shows that the correlation between the $1-GGI$ proxy and the source-country level of agreement with the statement *"A university education is more important for a boy than for a girl."*, derived from the World Values Survey, is positive and strong. Furthermore, the correlation between $1-GGI$ and the gender gap in math from the 2015 PISA evaluations is negative and strong, indicating that girls math performance improves

[14]GDP/capita and secondary education enrolment is measured in contemporaneous levels, i.e. as the average between 2007–2010. The data is collected from the World Bank Indicators (e.g. World Bank, 2021).

[15]The measures are 1-GGI from the 2010 Gender Gap Report, country-level of agreement with the World Values Survey statement *"A university education is more important for a boy than for a girl."*, country-level $1-FLFP$ rates (both defined as a contemporaneous measure from 2007–2010 and measured in the decade of mothers' emigration), average 1-GGI using both parents' source-countries, and average country-level math gender gap in the 2015 PISA evaluations.

compared to boys in more gender-equal countries. Taken together, these correlations indicate that $1 - GGI$ is a relevant proxy for source-country norms regarding gender and educational outcomes.

3.5. Independent variable: gender equality of the local labour market

To investigate the mitigation of cultural gender norms by neighbourhood characteristics, I derive a measure of gender equality at the mothers' local labour market (LLM) area. Even though asylum-seekers were assigned to municipalities and not LLM areas, the LLM areas better capture the available labour market opportunities for women.[16] I proxy LLM area gender equality with female-to-male labour force participation ratios:

$$FLFP_{lt} = \frac{FemaleLFP_{lt}}{MaleLFP_{lt}} \tag{1}$$

The measure varies at the level of assigned LLM area l and year of immigration t. At the time of mothers' immigration Sweden consisted of 120 different LLM areas, which are defined by Statistics Sweden based on commuting patterns. Asylum-seekers are assigned to municipalities, and I observe the female-to-male labour force participation ratio of each municipality at the time of mothers' immigration in the Statistics Sweden Labour Statistics Database (SCB, 2021). Some LLM areas consists of only one municipality, and some consists of several. To construct a measure of gender equality at the LLM area for those that contain multiple municipalities, I average labour force participation ratios of all municipalities belonging to the LLM area in question.

To make the mitigation estimates of my study easier to interpret, I define the LLM area measure to work in the opposite direction of the cultural norm measure. For this reason, I define the LLM gender equality proxy as $FLFP$, such that it increases as the female-to-male labour force participation ratio increases. I measure gender equality at the time of immigration and in the assigned LLM, as this is the only time it can be

[16] My results are robust to measuring gender equality at the municipality level, see Table 10. My data does not contain neighbourhood information at a more granular level than municipality. However, as the mitigation effect of labour force participation is likely to work through mothers' exposure to labour markets and peers, local labour market area is an appropriate definition of neighbourhoods for this context.

assumed to be exogenously given.

As an alternative measure of LLM area gender equality, to ensure that my results are robust to the measure of gender equality, I use female-to-male wage ratio. Furthermore, all regressions control for LLM area population, average income and cultural density (defined as the share of LLM area residents with the same nationality as the immigrating mother). Appendix Table A3 present the mean, standard deviation and cross-correlation of all these LLM area characteristics.

4. Empirical strategy

4.1. Do cultural gender norms affect the math gender gap?

To investigate the effect of cultural norms, it is necessary to disentangle the effects of culture from the effect of formal institutions. To do so, I rely on an extended version of the epidemiological approach. The epidemiological approach aims to identify the effect of culture by examining variation in outcomes among individuals who share formal and economic institutions, but who potentially differ in their social beliefs (Fernandez, 2007; Fernandez and Fogli, 2009; Fernández, 2011). Second-generation immigrants provide a useful experimental cohort for such a study. The assumptions behind this strategy are threefold: 1) parents (and parents' peers from the same source countries) transmit cultural beliefs to their children, 2) these cultural beliefs vary in a systematic way that reflects the culture of the parent's source country, and 3) individuals growing up in the same country encounter similar economic and formal institutions.

In an ideal experimental setup cultural norms would be randomly assigned, such that the gender norms are orthogonal to everything that relates to the childhood environment. Unfortunately, such ideal setups are hard to come by, and even if one were able to account for many observed differences across countries and individuals (e.g. income, parental education, etc.), one would still not be able to account for unobserved differences. Thus, a significant drawback of the epidemiological approach is that it cannot account for unobserved factors that varies in a systematic way by parents' source country. For example, different immigrant groups may have different reasons for migrating, they may be more or less likely to live in ethnic enclaves, and they may face different levels and types of discrimination in the migrant country. In addition, family structure may be endogenous to the sex of children, as suggested by Dahl and Moretti (2008), and this endogeneity could correlate with source-country culture.

To account for unobserved factors that may correlate with parents' cultural gender norms, I include family fixed effects in the epidemiological model and thus estimate the effect of cultural gender norms within families only. The family fixed effect absorbs any variation that is constant across siblings; hence, it controls for everything that affects all siblings equally, such as childhood environment, unobserved parental characteristics and endogenous family structure. The effect of cultural gender norms will be identified by families with within-variation in child gender, thus, families with at least one child of each gender. Thus, by construction, the model identifies only the effect of any *gender-specific* components of culture, i.e. gender norms, as anything that is not gender-specific will not vary across siblings and will therefore be absorbed by the family fixed effect.

This approach follows Finseraas and Kotsadam (2017), with the exception that instead of using sibling-pair fixed effects and a sample of matched opposite-sex siblings, I use the full sample of second-generation immigrants and my main specification includes family fixed effects. Thus, I compare every girl to the average of her brother's test scores (if she has more than one brother). By not restricting my sample to only families with at least one child of each gender I can better compare my results to previous research and discuss the possible bias in the standard epidemiological approach. The families with more than one child but all of the same gender will help identify the control variables and increase the sample size. However, I also ensure that my model is robust to fully following Finseraas and Kotsadam (2017) and include only a matched sample of opposite-sex siblings and sibling-pair fixed effects.

I estimate the following model:

$$MathScore_{ij} = \alpha Girl_{ij} + \beta Girl_{ij} \times Norms_j + \gamma Girl_{ij} \times X_j + \delta Birthorder_{ij} + Cohort_{ij} + \eta_j + \epsilon_{ij} \quad (2)$$

where i refers to individuals and j to families. α captures the baseline math score difference between opposite-sex siblings. The baseline cultural gender norm measure does not vary within families and will be absorbed by the family fixed effect η_j. The coefficient of interest is β, which captures how the sibling gender differences in math performance vary depending on the gender norms of the mother's source-country culture. X_j contains source-country controls for GDP/capita and gross enrolment in secondary school. These controls are included to ensure that my model captures the effect of source-country gender equality, and not just some other source-country characteristic that correlates with gender equality. I control for birth order and cohort fixed effects, and I cluster standard errors at source-country level.[17] β identifies the effect of cultural gender norms under the identifying assumption that any latent gap in childhood outcomes between brothers and sisters is as good as randomly assigned between families from different source countries.[18]

4.2. Is the effect mitigated by LLM area gender equality?

When investigating a mitigation effect of neighbourhood characteristics, the ideal setup would be to randomise gender norms among children's mothers and then randomise the types of neighbourhoods in which the children grow up. While, again, the ideal setup is not available, the refugee placement policy described in Section 2.2 offers quasi-random

[17] I cluster standard errors at the level of treatment, which is source country level. However, I only observe 19 different source countries, and there is a risk that having too few clusters may decrease the standard errors (Angrist and Pischke, 2009) For this reason, I report the p-value corresponding to standard errors obtained using a wild cluster bootstrap procedure (see e.g. Roodman et al. (2019)). Furthermore, Online Appendix Table OA1 shows that my result are robust to clustering standard errors at source country times LLM area, source country and LLM area, source country times cohort and family levels.

[18] This assumption would be violated if, for example, source-country culture was correlated with in-utero factors that affect opposite-sex foetuses differently. Following Figlio et al. (2019), I address this assumption by assessing whether cultural gender norms can predict neonatal characteristics, such as birth weight. Reassuringly, Appendix Table A4 shows that there is only a weak and negligible relationship between cultural gender norms and the gender gap in birth weight.

variation in LLM area characteristics, as asylum seekers were not free to choose their assigned location of residence.

An important issue regarding the refugee placement policy, which was raised by Nekby and Pettersson-Lidbom (2017), is that placement into municipalities may correlate with how the refugees were treated, and therefore also with the outcomes of the second generation. For example, if some municipalities systematically placed the refugees in areas with fewer social problems, this would most likely imply that the children of those refugees would attend schools of higher quality, which may affect their educational outcomes. If the likelihood of placing refugees in neighbourhoods of "higher quality" correlates with LLM area FLFP this would bias my results, as I would be comparing boys and girls attending different types of schools and living in different types of neighbourhoods. By relying on within-family comparisons, I increase the likelihood that my model identifies only the mitigation effect of neighbourhood gender equality, rather than the effect of some unobserved area characteristic that correlates with placement, as siblings are placed in the same neighbourhood, school and house, and therefore experience the same special treatment (if such treatment exists). Again, the family fixed effect absorbs everything that affects siblings equally, and I estimate the effect of only the *gender-specific* component of LLM area characteristics. I estimate the following equation:

$$\begin{aligned} MathScore_{ij} = {} & \alpha Girl_{ij} + \beta Girl_{ij} \times Norms_j + \gamma Girl_{ij} \times X_j \\ & + \delta Girl_{ij} \times LLMareaFLFP_{ij} + \lambda Girl_{ij} \times Norms_j \times LLMareaFLFP_j \\ & + \zeta Girl_{ij} \times Z_j + \theta Birthorder_{ij} + Cohort_{ij} + \eta_j + \epsilon_{ij} \end{aligned} \quad (3)$$

where i refers to individuals and j to families. α captures the baseline math gender gap between brothers and sisters, and β captures how this gender gap varies with cultural gender norms. δ captures the gender-specific effect of LLM area gender equality. λ captures how the gender-specific effect of cultural gender norms varies with the gender equality of the mothers' assigned LLM area. The mitigation effect is measured by the combination of δ and λ. The baseline cultural norm and LLM area characteristic effects will be absorbed by the family fixed effect η_j. X_j contains source-country controls for GDP/capita and gross enrolment in secondary school, and Z_j contains controls for population size, average income and cultural density of the assigned LLM area at time of immigration. The inclusion of LLM area controls ensures that my model captures the effect of LLM area gender equality, and not just some other LLM area characteristic that correlates with gender equality. I cluster standard errors at source country times assigned LLM area times year of immigration.[19]

A threat to identification is the possibility that asylum seekers were systematically placed in gender-equal areas based on the country they migrated from, which would cause correlation between LLM area FLFP and cultural gender norms and bias the results. To investigate this, I test whether cultural gender norms can predict LLM area gender equality. Table 3 presents the results of this balance test. Column (1) presents the raw correlation between cultural gender norms and LLM area gender equality. When municipalities could choose which asylum seekers were allocated to them, priority was given to the more highly educated and those who spoke the language of some of the resident immigrant stock. In addition, family size determined housing availability. Column (2) controls for these relevant placement characteristics, and column (3) adds immigration year fixed effects, which accounts for the fact that the number of participating municipalities increased over time. Column (4) adds the cohort of the mother.

The coefficient for Norms $(1 - GGI)$ shows the correlation between asylum-seeking mothers' cultural gender norms and the FLFP of the assigned LLM area. Reassuringly, the correlation is very small and statistically insignificant in all four specifications, which indicates that the allocation of asylum-seeking future mothers were random with respect to cultural gender norms. The share of LLM area residents from the same source country as the mother (cultural density) and family size predicts LLM area FLFP; this is expected as these were observed characteristics that influenced placement decisions. However, this

[19]Furthermore, Online Appendix Table OA1 shows that my results are robust to clustering the standard errors at only source country times LLM area, source country and LLM area, LLM area times immigration year, and family level.

correlation is not problematic, as it will be controlled for by the family fixed effects.

Appendix Table A5 illustrates the results of a placebo balance test when I allow selection, i.e. it investigates the correlation between asylum-seeking mothers' cultural gender norms and the FLFP of the LLM area they live in when their child is in ninth grade. The correlations shown in this placebo balance test are positive and statistically significant, which indicates that families select into LLM areas partly based on gender norms and gender equality, or based on variables that correlate with gender norms and gender equality. If this selection is related to the relative educational outcomes of boys and girls, it would bias my results. Thus, this result highlights the importance of using the refugee placement policy in order to obtain exogenous variation in LLM area characteristics.[20]

5. Results

5.1. Do cultural gender norms affect the math gender gap?

Table 4 shows the results for the first research question: the effect of cultural gender norms on the gender gap in math. All columns control for birth order and cohort fixed effects. Column (2) adds source-country fixed effects and controls for any gender-specific effect of source-country GDP/capita and gross enrolment in secondary schooling. Column (3) adds controls for family size, household income, education level and cohort of the mother. Although these are potentially endogenous to cultural gender norms and therefore bad controls, it is still useful to observe whether they change the estimate when included. Column (4) contains the preferred model specification, which adds family fixed effects. Column (5) follows Finseraas and Kotsadam (2017) and restricts the sample to a matched

[20]Refugees were not required to stay in their assigned municipality. Following initial assignment, there were no restrictions on moving to a different location, and refugees could move to another area if they found housing on their own. For this reason, I am only able to estimate the intention-to-treat effect of the assigned LLM area characteristics. However, 56% of the asylum-seeking mothers still lived in their assigned municipality (and LLM area) at the time when their children graduated from ninth grade. As expected, younger and more highly educated individuals are more likely to change municipality. Some measurement error may exist if individuals move within the first year of assignment, as I observe municipality of residence in the end of the year. This issue is investigated thoroughly by Edin et al. (2003), who use a weighting scheme based on the aggregate data on municipality reception of refugees. The weighting does not change their estimates significantly, which suggests that measurement error is not a substantial concern.

sample of siblings of opposite genders and controls for sibling-pair fixed effects — thus comparing each girl to every one of her brothers separately.

The girl indicator captures the gender gap in math for students whose mothers' cultural gender norm proxy takes the value zero (i.e. complete gender equality). This baseline math score gender gap is about 7% of a standard deviation in all specifications, which demonstrates that girls score lower on the math tests compared to boys even in the absence of the effect of cultural gender norms. The estimates in column (1) show that both boys' and girls' math performance decrease as the source-country culture becomes more traditional. The interaction effect of the girl indicator and the cultural gender norm measure shows that the gender gap in math increases as mothers' source-country culture becomes more traditional. The effect is calculated for a one-standard-deviation increase in the cultural gender norm proxy (about 0.09), which corresponds to the difference between Norway and Yugoslavia or Italy and Turkey. The interaction estimate of the preferred specification in column (4) demonstrates that girls whose mothers come from more gender-traditional countries fall behind their brothers by an additional 6.5% of a standard deviation. The p-value obtained using bootstrapped standard errors shows that clustering on source country may indeed decrease the standard errors slightly, however, using a wild cluster bootstrap demonstrates that the estimates are still statistically significant from zero on conventional levels.

Figure 1 graphically illustrates the interaction effect from the preferred specification, with the cultural gender norm proxy on the x-axis and the sibling math gender gap on the y-axis. The distribution of the gender norm proxy is plotted in the background. The interaction effect is negative; going from the most gender-equal source country to the least gender-equal source country corresponds to an increase in the math gender gap of about 20% of a standard deviation.

Comparing columns (2) and (3) with columns (4) and (5) in Table 4 allows me to evaluate the possible bias caused by the standard epidemiological approach, compared to when I estimate the effect within families. The results show that even though including family or sibling-pair fixed effects increase the explanatory power of my model significantly, it does not alter the estimates much. Comparing my estimates to those found in previous literature, my estimates are of the same sign but smaller in magnitude. Nollenberger et al. (2016) and Rodríguez-Planas and Nollenberger (2018) study PISA data on children from 35 different source-countries, and find that a one-standard-deviation increase in the source-country GGI is associated with a reduction in the math gender gap of 29% of the standard deviation of the average cross-country gender gap. After being recalculated to match their definition, my effect sizes correspond to an effect size of about 14% of the standard deviation of the average cross-country math score gender gap.[21]

A smaller effect size is in line with the results of Finseraas and Kotsadam (2017), who find that the effect of cultural gender norms on the gender gap in employment among second-generation immigrants in Norway is about 50% of the size of corresponding US estimates. The authors argue that the smaller effect is partly due to less bias as they include sibling-pair fixed effects, and partly due to the institutional context of Norway. I contribute to this discussion as I show that within-family comparisons do not change the estimates much compared to the standard epidemiological approach even when the data is not restricted to only contain opposite-sex siblings. Furthermore, I show that including family fixed effects (i.e. comparing every girl to the average of her brothers math scores) yields very similar estimates to including sibling-pair fixed effects (i.e. comparing every girl to each one of her brothers separately).

These results indicate that the smaller effect sizes found in my study are primarily due to the institutional contexts of Sweden rather than bias in the standard epidemiological approach. Finseraas and Kotsadam (2017) discuss several reasons why a smaller effect size might be expected in Norway that may also be applicable to Sweden. For example, both Sweden and Norway are social democratic welfare states with high female labour force participation and a relatively small gender gap in earnings, and the average gender gap in math is smaller in Sweden compared to many other countries.[22] For both Sweden

[21] Following Rodríguez-Planas and Nollenberger (2018), I calculate effect size as $\frac{NormSD \times \beta}{MathSD} = \frac{0.09 \times 0.065}{0.043} = 0.136$.

[22] For example, the average gender gap in math from the 2015 PISA evaluations show that the gender gap in math among Swedish students is positive (2.2), meaning that girls outperform boys. The only OECD countries with a higher gender gap in math in favour of girls are Norway and Finland (PISA,

and Norway humanitarian reasons are an important determinant of immigration, thus, the immigrants that come to Sweden and Norway may differ from the immigrants that come to other countries, for example like the U.S.

Rodríguez-Planas and Nollenberger (2018) show that parents' cultural gender norms affect not only the gender gap in math, but also the gender gaps in reading and science. This result implies that the cultural norms affecting girls' educational outcomes are not only math-specific, but rather reflect general stereotypes about gender and educational outcomes. Appendix Table A6 replicates Rodríguez-Planas and Nollenberger (2018) and estimates the effect of cultural norms on the gender gaps in final marks in Swedish, English, math and the total ninth grade mark. The table reports estimates from both the specification of Rodríguez-Planas and Nollenberger (2018) and my preferred specification with family fixed effects. The results of this exercise show that girls outperform boys in all subjects, and they get a higher total ninth grade mark. The effect of cultural gender norms is negative for all outcomes, however, it is smaller in magnitude and not statistically significant when the model relies on within-family comparisons. Even though these results indicate that the traditional epidemiological approach may be biased towards finding an effect in this setting, they should be interpreted with caution. Final marks in my data only can take one of four values (0, 10, 15 and 20), and may not contain enough within-family variation to credibly estimate the effect of cultural gender norms between siblings.

5.2. Is the effect mitigated by LLM area gender equality?

Table 5 presents the results for the second research question: whether LLM area gender equality can mitigate the negative effect of cultural gender norms. To transition from the first research question, column (1) replicates the preferred specification of Table 4 and estimates the effect of source-country gender norms on the sibling gender gap in math for the subset of children whose mothers were asylum seekers under the placement policy. The estimates demonstrate that the results for the first research question holds also on

2015).

the sub-sample created to evaluate the second research question.

Column (2) presents the mitigation effect, controlling only for birth order, cohort and source country fixed effects, and LLM area times immigration year linear trends. Column (3) adds controls for predetermined mother characteristics that may have influenced placement. Column (4) controls for any gender-specific effect of source-country GDP/capita and gross enrolment in secondary schooling, and any gender-specific effect by LLM area population size, average income level or cultural density. Column (5) contains the preferred model specification, which includes family fixed effects. Finally, column (6) follows Finseraas and Kotsadam (2017) and restricts the sample to a matched sample of opposite-sex sibling, and controls for sibling-pair fixed effects.

The baseline gender gap in math is about 6% of a standard deviation, which is similar to the gender gap in the sample used for the first research question. Cultural gender norms have a negative effect, such that the gender gap in math increases in favour of boys as the source-country culture becomes more traditional.

The estimate of the baseline gender-specific effect of LLM area gender equality shows that, if anything, boys seem to benefit more than girls from being in a more gender equal LLM area. However, when controls for potentially confounding LLM area characteristics are included this estimate becomes smaller and statistically insignificant — and when the model relies on within-family comparisons the estimate becomes even smaller and changes in sign. These estimates demonstrate that there may be some selective treatment happening at the LLM area that correlates with the gender gap in math and with some family time-invariant unobserved heterogeneity. Thus, even though Table 4 showed that the standard epidemiological approach do not introduce bias compared to a model that relies on within-family variation, the results in Table 5 demonstrate that relying on within-family comparisons is crucial to capture the effect of LLM area characteristics. This result is in line with the arguments of Nekby and Pettersson-Lidbom (2017), and indicates that there may be some selective treatment at the LLM area level that correlates with placement of refugees, and therefore need to be taken into account if we use the refugee placement policy to evaluate the impact of regional characteristics.

The three-way interaction estimate, showing how the gender-specific effect of cultural norms varies with LLM area gender equality, is both positive and statistically significant. Furthermore, the estimate increases in magnitude as the model controls for family or sibling-pair fixed effects, which again indicates that models which allow between-family comparisons may be biased towards finding a more favourable treatment of boys.

The combined mitigation effect of LLM area gender equality (i.e. the difference in the gender gap in math between siblings with mothers who were assigned to less gender-equal LLM areas and siblings with mothers who were assigned to more gender-equal LLM areas) is shown at the bottom of the table. Increasing the assigned LLM area FLFP by one standard deviation (0.04) leads to a mitigation effect of about 9% of a standard deviation of the math score in the within-family comparisons, which demonstrates that LLM area gender equality can completely mitigate and even reverse the negative effect of gender-traditional cultural norms. This result suggests that while girls who have mothers from traditional cultures do relatively worse in math compared to their brothers, this negative effect is mitigated for those girls whose mothers were assigned to more gender-equal LLM areas. Furthermore, the estimates suggest that the effect is not driven by a baseline gender-specific effect of LLM area characteristics, but rather through the impact the LLM area have on mothers from traditional cultures.

Figure 2 illustrates the graphical representation of the three-way interaction effect, with LLM area gender equality ($FLFP$) on the x-axis and the effect of cultural gender norms on the sibling gender gap in math on the y-axis. The distribution of assigned LLM area $FLFP$ is plotted in the background. Again, the results show that going from the least gender-equal LLM area to the most gender-equal LLM area would completely mitigate the negative effect of traditional gender norms.

Although there is less comparable estimates in the previous literature for my results on the second research question, my estimates can be evaluated in relation to those of Chetty and Hendren (2018). They estimate the effects of neighbourhood exposure on children's outcomes, and find that the outcomes of children who move to a new neighbourhood converge to the outcomes of the residents of the new neighbourhood at a rate of 4% per year of exposure. Extrapolating this result linearly implies that children who move to a new neighbourhood at birth and stay there until they are 20 years old would pick up about 80% of the difference in residents' outcomes between their origin and destination neighbourhood. My estimates could be viewed as the assimilation of the cultural gender norms to the neighbourhood setting, from before the birth of the child to the age of 16. My results show that neighbourhood gender equality completely mitigates the negative effect of cultural norms for girls whose mothers come from gender-traditional cultures, thus my results are in line with the extrapolated estimates of Chetty and Hendren (2018).

5.3. Possible mechanisms and heterogeneity

One possible mechanism behind my findings is an influence of nurture in the formation of girls gender identities and preferences over educational investments.[23] In line with previous findings of Kleven et al. (2019), who show that child penalties transmit through generations, or Fernández et al. (2004), who focus on intergenerational transmission of labour supply, it is possible that observing the labour market behaviour of ones mother may influence the gender identity and educational investment of the second-generation girls of my study.

Columns (1) and (2) in the top panel of Table 6 demonstrate that culture of traditional gender norms is negatively associated with the probability that a mother has positive labour earnings. Likewise, columns (1) and (2) in the bottom panel shows that being assigned to a gender-equal LLM area is positively associated with the probability that a mother has positive labour earnings. However, the third column shows that controlling for a gender-specific effect of working mothers do not alter any of my results, which indicates that even though the effect that cultural norms and neighbourhood exposure has on mothers' labour supply may be one possible mechanism of my findings, it is not

[23] Akerlof and Kranton (2000) develop a theoretical framework in which individuals choose to identify with different social categories, and derive utility from complying with the behaviour prescribed by these chosen categories. One salient category is gender, where everyone is assigned to either being a "man" or a "woman", and where there are prescribed attributes and behaviours that are considered "manly" or "womanly". This way, gender identity changes the pay-off of different actions, and the choice of identity may impact our economic outcomes, including educational performance. In this framework, as well as in empirical studies testing its implications, identity concerns have a significant impact on educational outcomes (Akerlof and Kranton, 2002; Schüller, 2015).

the only one. In line with this reasoning, the top left graphs of Figure 5 and Figure 6 show that the estimated effects of cultural norms and neighbourhood exposure do not differ when they are interacted with the mother's work status.

Furthermore, both the effect of cultural gender norms and the mitigation effect may differ depending on other characteristics of the mother. For example, a recent study in the U.S. by Dossi et al. (2021b) shows that growing up in a family that exhibits a preference for boys only has a negative impact on the math performance of girls from White families with higher income and maternal education. Moreover, education level could influence the assimilation of a mother's cultural values to the neighbourhood setting, as learning Swedish and finding employment may be easier for the more highly educated. The intergenerational transmission of cultural gender norms, and the mitigation by LLM area gender equality, could differ in strength depending on how long the mother has been living in Sweden, or how strong the cultural influence of the neighbourhood is, that is, whether the family is surrounded by peers with similar cultural background or not.

The relative math performance of a girl may also differ by the number of siblings she has, if, for example, girls in larger families have increased responsibilities at home. Furthermore, Figlio et al. (2020) show that the gender gaps in educational outcomes are larger at the tails of the school performance distribution, and that the gender-specific effect of socio-economic status is more pronounced in the bottom quantile. In line with these results, the effect of cultural gender norms could differ by the baseline school performance of both siblings.

Figure 3 presents the heterogeneous effects of cultural gender norms. I find no heterogeneous effects depending on mothers employment, whether or not the mother has some university education, by the cultural density of the LLM area, or by time since the mother's immigration. The effects are driven by families with two or three children, but this is not surprising as families with four or more children are much less common. The bottom right graph shows that the effect of cultural gender norms is strongest for families where siblings on average are in the middle quintiles of the final grade distribution. It is reasonable to believe that high-achieving students are not affected in the same way as struggling students are. However, as both math performance and final grades are outcomes of gender norms, and because the sibling gender gap mechanically depends on the siblings' absolute performance, this is a difficult issue to investigate and the results should be interpreted with caution.

Figure 4 presents the heterogeneous mitigation effects.[24] The figure shows that the mitigation effect appears stronger for siblings whose mothers are not working, did not have any university education upon arrival to Sweden, and were assigned to LLM areas with a below-median share of same-ethnicity residents. However, none of the estimates are statistically different from each other, and I find no heterogeneous effects by time since the mother immigrated or by family size. The bottom right graph mirrors the results in the corresponding graph of Figure 5, and shows that the mitigation effect is stronger for siblings whose grades are in the middle quintiles of the grade distribution.

6. Robustness checks and sensitivity analysis

6.1. Robustness checks

Tables 7 demonstrates the results from a series of robustness checks. Column (1) shows that my results are robust to dropping the country which contributes to the largest immigrant flows (Finland). Even though the estimate loses some precision, it is still negative, of about the same magnitude as the main results and close to statistically significant at conventional levels (p-value is 0.13). Column (2) restricts the sample to only families where both parents are born abroad, to ensure that the results are not driven by a confounding relationship between a future mothers' likelihood to have children with a native man and the intergenerational transmission of cultural norms. To increase the likelihood of the siblings being exposed to the same family environment as they grow up, Column (3) restricts the sample to siblings who are at most five years apart in age. In both columns the results remain robust.

Appendix Table A4 demonstrates a negative relationship between cultural gender

[24]The figure present the mitigation effect, i.e. the combination of the LLM area effect and the three-way interaction effect.

norms and the gender gap in birth weight. Although this relationship is weak and the estimate small in magnitude, it could indicate that there is some relationship between correlates of cultural gender norms and in-utero environment which is not picked up by the other source-country controls. However, the results in column (4) of Table 7 demonstrates that controlling for a gender gap in birth weight does not alter my results.

Figlio et al. (2019) find that boys from families with lower socio-economic status perform worse relative to their sisters. If family disadvantage correlates with cultural gender norms, there is a risk that my study may capture a relative effect of family disadvantage for boys rather than a relative effect of gender norms for girls. Reassuringly, column (5) shows that controlling for a gender-specific effect of mothers' education level does not affect the estimates. Finally, I address the concern of selective migration by controlling for a gender-specific effect of geographical distance between mothers' birth country and Stockholm, which, as column (6) shows, also does not alter my estimates.[25]

Table 8 show the results for the same robustness checks when estimating the mitigation effect, which the exception that column (1) drops the largest and most urban local labour market area (Stockholm) instead of the most common source country. The results show that the three-way interaction estimate, showing how the gender-specific effect of cultural norms varies with LLM area gender equality, is positive, statistically significant, and large in magnitude in all specifications. Thus, the main results for the mitigation effect remain robust.

6.2. Alternative measures of gender norms and gender equality

In order to test the sensitivity of my results to the chosen proxy for gender norms and gender equality, I replicate my main results using alternative measures of both cultural gender norms and LLM area gender equality. Table 9 demonstrates that I can replicate my results when measuring cultural gender norms with the $1 - GGI$ score from 2010

[25]Furthermore, Appendix Figure A1 addresses the concern that migrants might leave their source country precisely because they do not agree with the gender norms there. Such selection would lead me to understate the effect of cultural gender norms and is therefore not a threat to my study. Still, the figure shows that the correlation between source country GGI and the female-over-male labour force participation of the first generation in Sweden is positive, which indicates that, at least in relative terms, immigrants behave in accordance to their cultural gender norms when they reside in Sweden.

(column (1)), with source-country level of agreement with the statement *"A university education is more important for a boy than for a girl."* derived from the World Values Survey (column (2)), with contemporaneous $1 - FLFP$ of the mother's source country (column (3)), with $1 - FLFP$ of the mother's source country measured in the decade she emigrated to Sweden (column (4)), with the average $1 - GGI$ of both parents' birth countries (column (5)), and finally, using the average gender gap in math from the 2015 PISA evaluations (column (6)).[26]

Comparing the results in column (3) and (4) demonstrates that the effect is not sensitive to when in time the cultural norms are measured, and column (5) shows that the results are not affected by assortative matching on cultural gender norms and its potential correlation with the gender gap in math. Recalculating the effect sizes of these estimations to be comparable to the results of Rodríguez-Planas and Nollenberger (2018) shows that the effect sizes range between 13–24% of the standard deviation of the average cross-country gender gap in math. Thus, my results are robust to using alternative measures of cultural gender norms, and all alternative measures produce effect sizes that are smaller than what has been found in previous literature.

Table 10 demonstrates the mitigation effect when LLM area gender equality is measured using the female-over-male wage ratio (column (1)), or the municipality-level FLFP (column (2)). Column (3) and (4) estimate the mitigation effect of LLM area FLFP when cultural gender norms are measured using the $1-GGI$ of 2010 or average $1-GGI$ of both parents' source countries, respectively. In the interest of preserving space the mitigation of cultural norms measured using the WVS measure, contemporaneous source-country FLFP, or the source-country FLFP when mothers emigrated are presented in Appendix Table A8.[27] In all specifications the results remain robust to using alternative measures

[26] In order for all norm measures to increase as gender norms become less favourable towards women, I define the PISA gap as $-(PISAgap)$ in the regressions. The sample size fluctuates somewhat between specifications as some countries/decades/fathers are missing for different variables.

[27] Information on the PISA gender gap is missing for a third of the source-countries of asylum-seeking mothers, therefore I cannot credibly estimate the mitigation effect using the PISA gender gap as the cultural norm proxy. Finally, Online Appendix Tables OA2–OA4 show that the gender norm effect is similar for measuring cultural gender norms using fathers' source-country $1 - GGI$, but that LLM area gender equality does not mitigate the effect of fathers' cultural gender norms. This result is not surprising, as LLM area gender equality likely has a stronger relative impact on women then on men.

of both cultural norms and neighbourhood gender equality.[28]

7. Conclusion

This paper estimates the effect of cultural gender norms on the gender gap in math, and explores whether this effect is mitigated by local labour market (LLM) area gender equality. I employ an extended version of the epidemiological approach and estimate the effect of source-country culture of traditional gender norms on the gender gap in math among second-generation immigrant siblings of opposite genders. This way, I replicate and extend previous research on the impact of cultural norms on the gender gap in math, and I am able to investigate whether the standard epidemiological approach causes biased results by not accounting for family-specific correlates of cultural gender norms.

I show that cultural gender norms have a negative impact on girls' relative math performance, such that the sibling gender gap in math increases with the gender-traditional culture of the mother's country of origin. Compared to previous literature, my effects are of the same sign but smaller in magnitude. However, comparing the results of the standard epidemiological approach to when the model is extended with within-family comparisons reveals that the difference in effect sizes is likely not due to bias in the standard epidemiological approach, but rather, to institutional factors specific to Sweden.

To investigate the mitigation effect of LLM area gender equality, I exploit a refugee placement policy to obtain random variation in LLM characteristics. Again, in the preferred specification I compare the outcomes of opposite-sex siblings, allowing me to control for any differential treatment by LLM areas that is unrelated to the gender of the child. I show that LLM area gender equality can completely mitigate the negative effect of cultural gender norms. This result means that even though the sibling gender gap in math increases with culture of traditional gender norms, this increase is smaller or even reversed for siblings whose mothers were placed in more gender-equal LLM areas.

Previous research have shown that both culture and neighbourhood exposure have an

[28]The mitigation estimates using municipality gender equality is similar to the main results, albeit slightly weaker, which again suggests that the effect of neighbourhood exposure on mothers from gender-traditional cultures works primarily through labour market exposure.

impact in children's educational outcomes, but to my knowledge there is yet no evidence on how these two channels cooperate. Thus, the main contribution of my study is to merge the literature on educational outcomes, cultural norms and neighbourhood exposure — and investigate how the effect of cultural gender norms interacts with neighbourhood gender equality. A secondary contribution of my study is that I replicate previous evidence on the link between cultural gender norms and the gender gap in math, and investigate whether the standard epidemiological approach causes biased results.

Regarding validity, my study is based on variation among the most common source countries of immigrants to Sweden and the results are internally valid for the Swedish context. Even though my number of observable source countries is lower than in previous studies such as Rodríguez-Planas and Nollenberger (2018), the included source countries are those most common for Swedish immigration flows and, therefore, the immigrants I observe are representative of Swedish immigrants. However, as my study is based on findings for Sweden and source-countries common for immigrants to Sweden, the external validity of the results to an international context may be limited to countries similar to Sweden, such as other high-income welfare states for which humanitarian reasons are an important determinant for immigration. Future research could extend my results by investigating if the results on the mitigation effect can be extrapolated to an international context.

Finally, my results show that while cultural gender norms play an important role for the gender gap in math, they are not immune to the influence of surrounding characteristics. This result is important from a policy perspective. It would most likely be difficult to influence the norms transmitted to children by their parents; however, the understanding that these norms are affected by surrounding characteristics, which can be influenced, provides policy-makers with opportunities to affect change.

References

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Akerlof, G. A. and R. E. Kranton (2002). Identity and schooling: Some lessons for the economics of education. *Journal of economic literature 40*(4), 1167–1201.

Aldén, L. and E. Neuman (2019). Culture and the gender gap in gajor choice: an analysis using sibling comparisons. Working Paper 2019:3, Linnaeus University Centre for Discrimination and Integration Studies.

Alesina, A. and P. Giuliano (2015). Culture and institutions. *Journal of Economic Literature 53*(4), 898–944.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Angrist, J. and J.-S. Pischke (2009). Mostly harmless econometrics: an empiricists guide.

Åslund, O., P.-A. Edin, P. Fredriksson, and H. Grönqvist (2011). Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *American Economic Journal: Applied Economics 3*(2), 67–95.

Bedard, K. and I. Cho (2010). Early gender test score gaps across oecd countries. *Economics of Education Review 29*(3), 348–363.

Bisin, A. and T. Verdier (2011). The economics of cultural transmission and socialization. In *Handbook of social economics*, Volume 1, pp. 339–416. Elsevier.

Bratsberg, B., J. Ferwerda, H. Finseraas, and A. Kotsadam (2020). How settlement locations and local networks influence immigrant political integration. *American Journal of Political Science*.

Chetty, R. and N. Hendren (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

Dahl, G. B., C. Felfe, P. Frijters, and H. Rainer (2020). Caught between cultures: Unintended consequences of improving opportunity for immigrant girls. Working Paper 26674, National Bureau of Economic Research.

Dahl, G. B., K. V. Løken, and M. Mogstad (2014). Peer effects in program participation. *American Economic Review 104*(7), 2049–74.

Dahl, G. B. and E. Moretti (2008). The demand for sons. *The Review of Economic Studies 75*(4), 1085–1120.

Damm, A. P. and C. Dustmann (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review 104*(6), 1806–32.

DiPrete, T. A. and C. Buchmann (2013). *The rise of women: The growing gender gap in education and what it means for American schools.* Russell Sage Foundation.

Dossi, G., D. Figlio, P. Giuliano, and P. Sapienza (2021a). Born in the family: preferences for boys and the gender gap in math. *Journal of Economic Behavior & Organization 183*, 175–188.

Dossi, G., D. N. Figlio, P. Giuliano, and P. Sapienza (2021b). The family origin of the math gender gap is a white affluent phenomenon. Technical report, National Bureau of Economic Research.

Edin, P.-A., P. Fredriksson, and O. Åslund (2003). Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The quarterly journal of economics 118*(1), 329–357.

Fernandez, R. (2007). Women, work, and culture. *Journal of the European Economic Association 5*(2-3), 305–332.

Fernández, R. (2011). Does culture matter? In *Handbook of social economics*, Volume 1, pp. 481–510. Elsevier.

Fernandez, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics 1*(1), 146–77.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Figlio, D., P. Giuliano, U. Özek, and P. Sapienza (2019). Long-term orientation and educational performance. *American Economic Journal: Economic Policy 11*(4), 272–309.

Figlio, D., K. Karbownik, J. Roth, M. Wasserman, et al. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics 11*(3), 338–81.

Figlio, D. N., K. Karbownik, J. Roth, M. Wasserman, et al. (2020). Males at the tails: How socioeconomic status shapes the gender gap. Working Paper 27196, National Bureau of Economic Research.

Finseraas, H. and A. Kotsadam (2017). Ancestry culture and female employment—an analysis using second-generation siblings. *European sociological review 33*(3), 382–392.

Fryer Jr, R. G. and S. D. Levitt (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics 2*(2), 210–40.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008). Culture, gender, and math. *Science 320*(5880), 1164–1165.

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

Lundberg, S. (2020). Educational gender gaps. *Southern Economic Journal 87*(2), 416–439.

Maurin, E. and J. Moschion (2009). The social multiplier and labor market participation of mothers. *American Economic Journal: Applied Economics 1*(1), 251–72.

Migrationsverket (1999). Asylsökande till sverige 1984–1999. Data retrieved from The Swedish Migrations Sgency's Statistics Database, https://www.migrationsverket.se/Om-Migrationsverket/Statistik/Asyl.html Accessed: 2019-11-01.

Mota, N., E. Patacchini, and S. S. Rosenthal (2016). Neighborhood effects, peer classification, and the decision of women to work.

Nekby, L. and P. Pettersson-Lidbom (2017). Revisiting the relationship between ethnic diversity and preferences for redistribution: Comment. *The Scandinavian Journal of Economics 119*(2), 268–287.

Nicoletti, C., K. G. Salvanes, and E. Tominey (2018). The family peer effect on mothers' labor supply. *American Economic Journal: Applied Economics 10*(3), 206–34.

Nollenberger, N., N. Rodríguez-Planas, and A. Sevilla (2016). The math gender gap: The role of culture. *American Economic Review 106*(5), 257–61.

Olivetti, C., E. Patacchini, and Y. Zenou (2020). Mothers, peers, and gender-role identity. *Journal of the European Economic Association 18*(1), 266–301.

Program for International Student Assessment (2015). Oecd pisa scores in reading and mathematics, by gender. Data retrieved from The PISA website, https://pisadataexplorer.oecd.org/ide/idepisa/dataset.aspx Accessed: 2021-05-20.

Pope, D. G. and J. R. Sydnor (2010). Geographic variation in the gender differences in test scores. *Journal of Economic Perspectives 24*(2), 95–108.

Rodríguez-Planas, N. and N. Nollenberger (2018). Let the girls learn! it is not only about math... it's about gender social norms. *Economics of Education Review 62*, 230–253.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal 19*(1), 4–60.

Statistics Sweden (2021). Population and gainfully employed by year, sex and region of residence. Data retrieved from Statistics Sweden Labour Statistics Database (RAMS), https://www.scb.se/en/finding-statistics/statistics-by-subject-area/labour-market/employment-and-working-hours/labour-statistics-based-on-administrative-sources/ Accessed: 2021-01-11.

Schüller, S. (2015). Parental ethnic identity and educational attainment of second-generation immigrants. *Journal of Population Economics 28*(4), 965–1004.

Schwab, K., R. Crotti, T. Geiger, and S. Zahidi (2018). The global gender gap report 2018. World Economic Forum.

The World Bank (2021). School enrollment, secondary (Data retrieved from World Development Indicators, https://data.worldbank.org/indicator/SE.SEC.ENRR Accessed: 2021-01-11.

Table 1: Summary statistics

	(1)	(2)	(3)	(4)	(5)
	All	RQ1 sample	RQ1: > 1 siblings	RQ2 sample	RQ2: > 1 siblings
2nd generation	0.22				
	(0.42)				
Girl	0.49	0.49	0.49	0.49	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Cohort	1991.74	1992.02	1992.09	1992.45	1992.56
	(2.42)	(2.42)	(2.39)	(2.12)	(2.18)
9th grade total mark	212.90	205.72	207.12	204.02	203.18
	(58.93)	(62.14)	(60.97)	(61.72)	(61.04)
Math test score	36.40	33.02	33.44	31.05	30.67
	(15.64)	(15.56)	(15.62)	(15.24)	(15.26)
Math gender gap (std)	-0.01	-0.07	-0.07	-0.07	-0.07
	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)
Family size	1.73	1.64	2.25	1.65	2.25
	(0.71)	(0.73)	(0.53)	(0.73)	(0.51)
HH income > median	0.50	0.33	0.35	0.21	0.21
	(0.50)	(0.47)	(0.48)	(0.41)	(0.41)
Mothers: university ed.	0.35	0.26	0.26	0.21	0.18
	(0.48)	(0.44)	(0.44)	(0.41)	(0.38)
Mother's age	44.85	45.14	44.41	44.57	43.75
	(5.02)	(5.42)	(5.03)	(5.32)	(4.99)
Mother's norms (1-GGI)		0.28	0.28	0.35	0.36
		(0.09)	(0.09)	(0.06)	(0.06)
Assigned LLM FLFP				0.89	0.89
				(0.04)	(0.04)
Observations	840,296	75,780	39,007	11,773	6,102

Notes: The table reports means and standard deviations of the main variables. Column (1) includes all ninth grade students taking the national standardised tests between 2004–2012. Column (2) includes children whose mothers are immigrants, and column (3) restricts this sample to only children with at least one sibling (the sample which identifies effects in models with family fixed effects). Column (4) includes children whose mothers were subjected to the refugee placement policy, and column (5) restricts this sample to only children with at least one sibling (the sample which identifies effects in models with family fixed effects). The gender gap in math is calculated from estimating linear regressions of standardised math score on a female indicator.

Table 2: Source-country variables and number of children

	(1)	(2)	(3)	(4)	(5)	(6)
	Math gap (std)	1 - GGI (2018)	GDP/capita (USD)	Secondary edu.	N RQ1	N RQ2
Chile	-0.073	0.283	11,067	0.818	3,759	1,434
Czechoslovakia	0.064	0.307	18,642	0.912	554	
Denmark	-0.119	0.222	59,753	1.149	2,772	
Ethiopia	-0.038	0.344	409	0.268	2,420	1,034
Finland	-0.036	0.179	48,930	1.125	19,753	
Germany	-0.046	0.224	42,508	1.024	1,526	
Greece	-0.087	0.304	29,363	0.909	600	
Iran, Islamic Rep.	-0.130	0.411	5,732	0.699	4,562	2,217
Iraq	-0.144	0.449	4,082	0.465	3,105	829
Italy	-0.115	0.294	37,920	0.871	280	
Lebanon	-0.119	0.405	6,608	0.651	4,960	1,750
Norway	-0.089	0.165	87,438	1.077	3,764	
Poland	-0.055	0.272	12,347	0.925	5,146	1,427
Russian Federation	-0.143	0.299	9,993	0.918	586	
Thailand	-0.072	0.298	4,410	0.489	1,690	
Turkey	-0.098	0.372	10,144	0.64	6,714	1,533
Unites States	0.034	0.28	47,981	0.945	1,006	
Vietnam	0.059	0.302	1,147	0.455	1,170	449
Yugoslavia	-0.045	0.267	10,265	0.892	11,413	1,100
N RQ1	75,780	75,780	75,780	75,780	75,780	
Mean RQ1	-0.067	0.277	26,269	0.871		
SD RQ1	0.043	0.086	23,962	0.238		
N RQ2	11,773	11,773	11,773	11,773		11,773
Mean RQ2	-0.064	0.352	7,554	0.68		
SD RQ2	0.047	0.063	3,675	0.185		

Notes: The table reports country-specific values for the math gender gap, and each of the source-country variables. Column (1) present country-level gap in math score, obtained from country-specific linear regressions of standardised math score on a female indicator. Column (2) contains source-country 1-GGI (2018) score. Column (3) contains average GDP per capita (in USD) between 2007-2010, and column (4) contains the average gross enrolment ratio in secondary school between 2007–2010. Columns (5) and (6) report frequencies of children with mothers from each source country for both sample definitions. The bottom panels reports mean and cross-country standard deviation of each of the variables, for both sample definitions.

Table 3: Balance test for refugee placement policy

	(1)	(2)	(3)	(4)
Norms (1-GGI)	-0.023	0.035	0.031	0.031
	(0.058)	(0.042)	(0.043)	(0.043)
Cultural density		-0.566***	-0.553***	-0.553***
		(0.049)	(0.051)	(0.051)
Mother's edu (im.)		-0.013	-0.006	-0.006
		(0.008)	(0.009)	(0.008)
Family size		-0.059***	-0.063***	-0.063***
		(0.014)	(0.015)	(0.014)
Mother's cohort				-0.000
				(0.003)
R-squared	0.00	0.33	0.34	0.34
Observations	11,773	11,773	11,773	11,773
	Indicators			
Immigration year FE	No	No	Yes	Yes

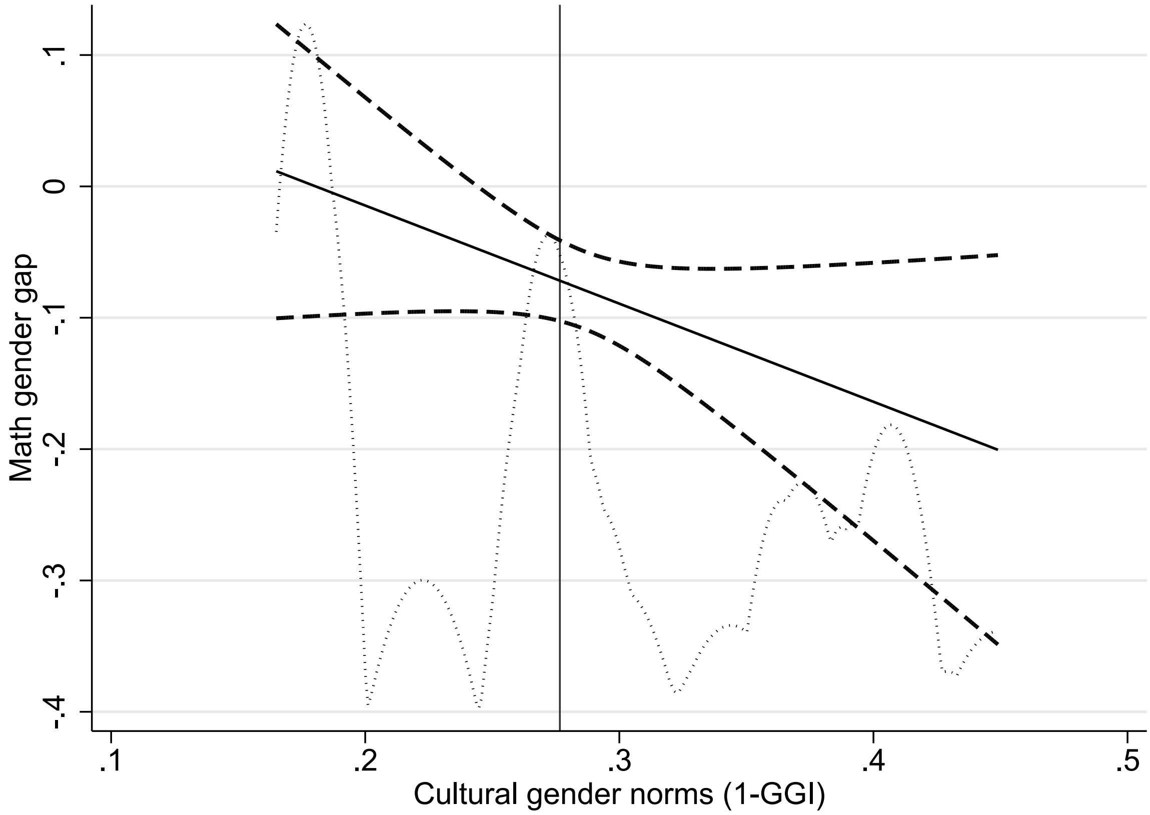
Notes: The table reports the correlation between local labour market area FLFP and mothers' source-country gender norms and individual characteristics. The dependent variable is the FLFP of the assigned local labour market. All maternal characteristics are measured at the time of immigration. Cultural density is the share of local labour market residents from the same source country as the mother who is being placed there. Standard errors are clustered at the source country × assigned local labour market × immigration year level. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Effect of culture of traditional gender norms on the math gender gap

	(1)	(2)	(3)	(4)	(5)
Girl	-0.070***	-0.066***	-0.063***	-0.072***	-0.074***
	(0.006)	(0.004)	(0.004)	(0.012)	(0.012)
Norms	-0.097**				
	(0.041)				
Girl × norms	-0.029***	-0.064***	-0.064***	-0.065**	-0.063*
	(0.005)	(0.010)	(0.010)	(0.030)	(0.030)
p(girl × norms = 0)	0.066	0.004	0.005	0.077	0.093
R-squared	0.01	0.05	0.14	0.70	0.71
Observations	75,780	75,780	75,780	39,007	23,650
	Indicators				
Birth order	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	No	No
Mother controls	No	No	Yes	No	No
Country controls	No	Yes	Yes	Yes	Yes
Family FE	No	No	No	Yes	No
Sibling-pair FE	No	No	No	No	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.09). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. Mothers' controls are cohort, education level, household income and family size. Standard errors in parentheses are clustered at the source country. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * p<0.1, ** p<0.05, *** p<0.01.

Figure 1: Effect of culture of traditional gender norms on the math gender gap



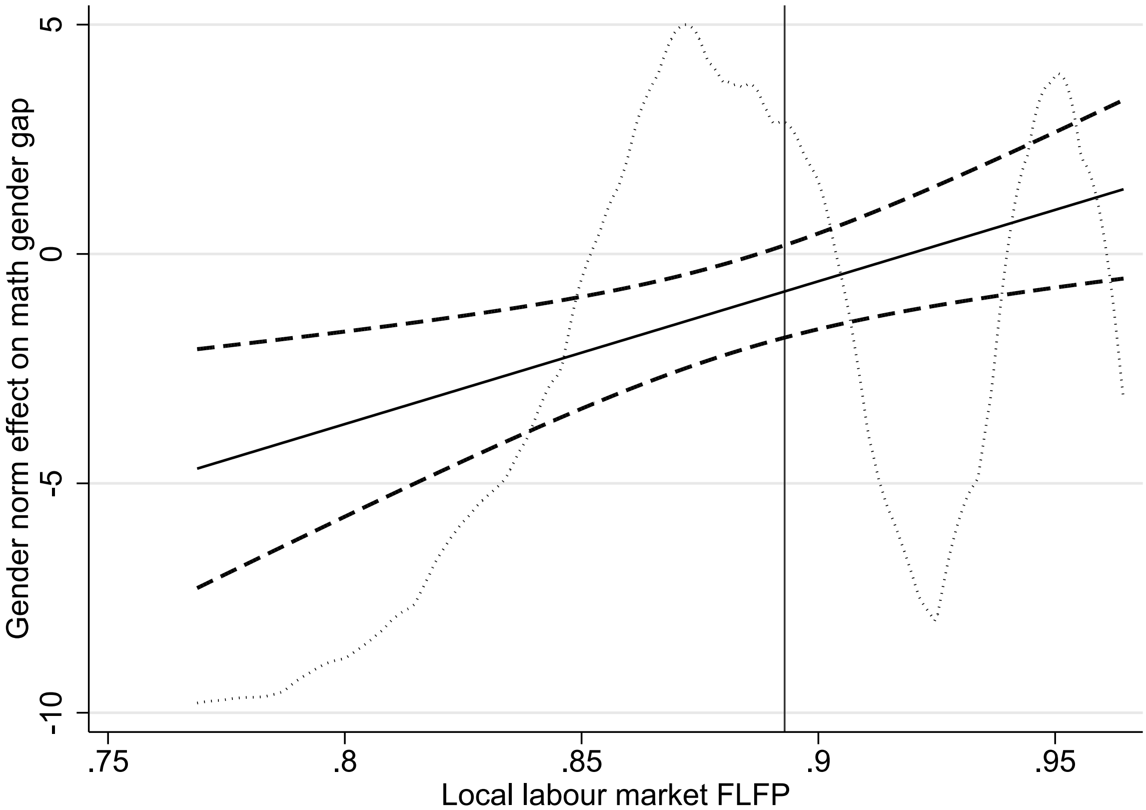
Notes: The figure plots the effect of cultural gender norms (girl × source-country 1-GGI) on the gender gap in math. The dashed lines depict the 95% confidence interval. The kernel density distribution of the source-country 1-GGI is plotted with the dotted line. The vertical line indicates the mean of 1-GGI.

Table 5: Local labour market mitigation effect

	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.066**	-0.057***	-0.056***	-0.056***	-0.064**	-0.065**
	(0.028)	(0.018)	(0.017)	(0.017)	(0.027)	(0.028)
Girl × norms	-0.062*	-0.033*	-0.033*	-0.051**	-0.054*	-0.055*
	(0.032)	(0.017)	(0.017)	(0.020)	(0.032)	(0.032)
Girl × LLM FLFP		-0.047***	-0.045***	-0.033	0.008	0.011
		(0.016)	(0.016)	(0.035)	(0.052)	(0.054)
Girl × norms × LLM FLFP		0.033**	0.037**	0.033**	0.085***	0.084***
		(0.016)	(0.016)	(0.016)	(0.029)	(0.029)
Mitigation		-0.014	-0.008	-0.000	0.093	0.095
p(mitigation = 0)		0.565	0.752	0.992	0.126	0.130
R-squared	0.70	0.12	0.15	0.15	0.70	0.71
Observations	6,102	11,725	11,725	11,725	6,102	3,684
			Indicators			
Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes	No	No
LLM × Im.Year FE	No	Yes	Yes	Yes	No	No
Country controls	Yes	No	No	Yes	Yes	Yes
LLM controls	Yes	No	No	Yes	Yes	Yes
Mother controls	No	No	Yes	Yes	No	No
Family FE	Yes	No	No	No	Yes	No
Sibling-pair FE	No	No	No	No	No	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Local labour market FLFP is measured at the assigned local labour market, in the year in which the mother immigrated to Sweden. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.06), and a one standard deviation increase in local labour market FLFP (0.04). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration, all interacted with a gender indicator. Mothers' controls are cohort, education level and family size, all of which are measured at time of immigration. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl × LLM FLFP and girl × norms × LLM FLFP, the corresponding p-value is based on the standard errors of this combined estimate. Standard errors are clustered at the source country × assigned local labour market × immigration year level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Figure 2: Local labour market mitigation effect



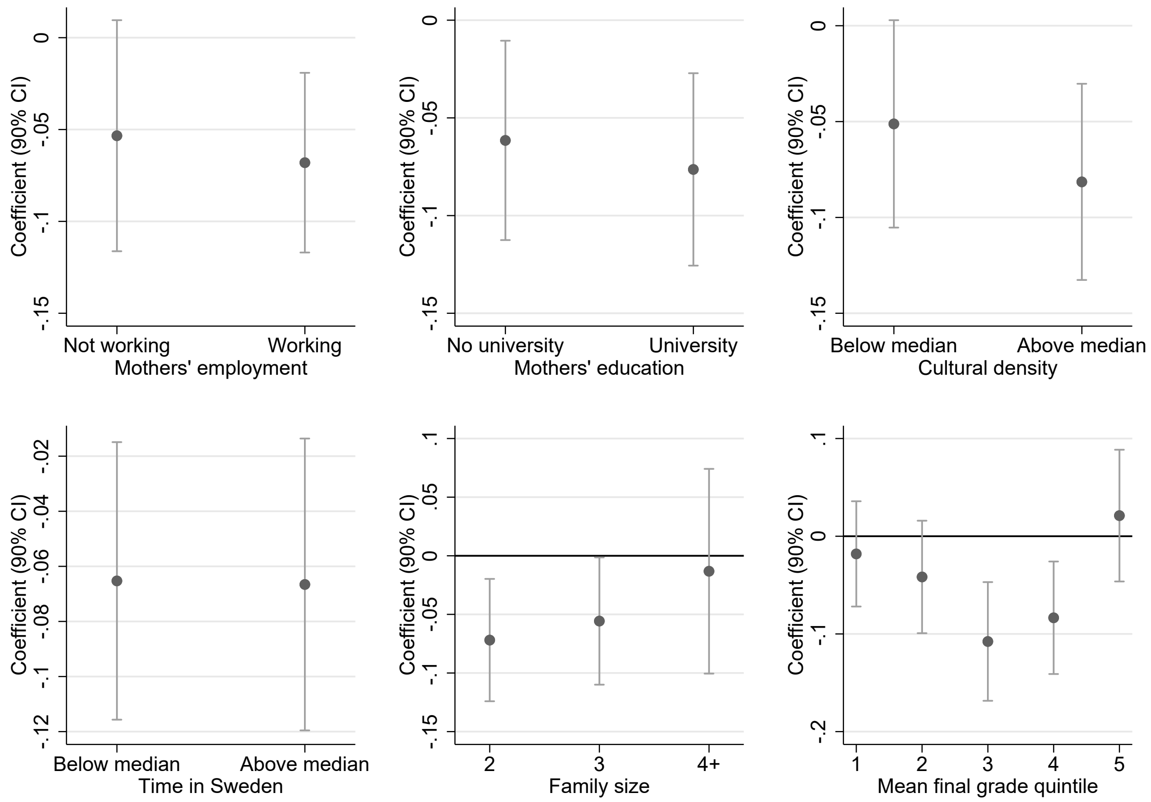
Notes: The figure plots the three-way interaction effect (girl × source-country 1-GGI × local labour market FLFP) on the gender gap in math. The dashed lines depict the 95% confidence interval. The kernel density distribution of the local labour market level FLFP is plotted with the dotted line, and the vertical line indicates the mean local labour market FLFP. FLFP is measured at the assigned local labour market, in the year in which the mother immigrated to Sweden.

Table 6: Possible mechanism: working mother

	(1)	(2)	(3)
	Outcome: working mother		Outcome: math score
Norms	-0.078***	-0.081***	
	(0.017)	(0.018)	
Mother's cohort		0.007***	
		(0.001)	
Family size		0.029**	
		(0.011)	
Mother: uni		0.154***	
		(0.020)	
Girl			-0.081***
			(0.023)
Girl × norms			-0.064*
			(0.031)
Girl × working m.			0.011
			(0.017)
p(norms = 0)	0.014	0.016	
p(girl × norms = 0)			0.099
Mean	0.79	0.79	0.82
R-squared	0.04	0.08	0.70
Observations	75,780	75,780	39,007
	Outcome: working mother		Outcome: math score
LLM FLFP	0.018***	0.007	
	(0.007)	(0.006)	
Norms		-0.052***	
		(0.006)	
Mother's cohort		0.002*	
		(0.001)	
Family size		-0.074***	
		(0.005)	
Mother: edu (im.)		0.090***	
		(0.011)	
Girl			-0.014
			(0.052)
Girl × norms			-0.061*
			(0.032)
Girl × LLM FLFP			0.008
			(0.052)
Girl × norms × LLM FLFP			0.085***
			(0.029)
Girl × working m.			-0.067
			(0.059)
Mean	0.73	0.73	0.75
R-squared	0.00	0.07	0.70
Observations	11,773	11,773	6,102
		Indicators	
Year FE	No	Yes	No
Birth order	No	No	Yes
Cohort FE	No	No	Yes
Country controls	No	No	Yes
LLM controls	No	No	Yes
Family FE	No	No	Yes

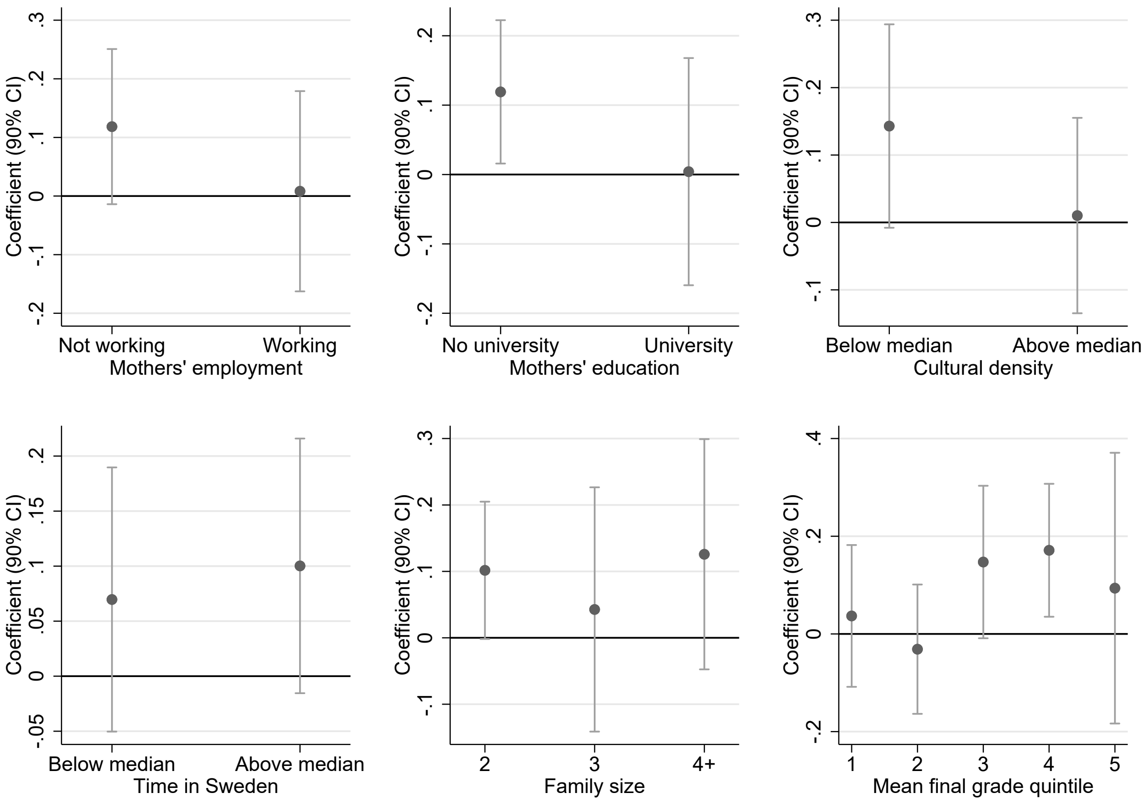
Notes: The dependent variable in column (1)–(2) is an indicator that takes the value one if the mother is earning positive labour earnings in the current year (mean reported in the bottom of the table). The dependent variable in column (3) is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Mother characteristics in top panel are measured at the time when her child is in 9th grade, and the top panel do not contain LLM controls in column (3). Mother characteristics in the bottom panel are measured at time of mother's immigration. Standard errors are clustered at the source country level in the top panel (the p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom), and at the source country × assigned local labour market × immigration year level in the bottom panel. * p<0.1, ** p<0.05, *** p<0.01.

Figure 3: Heterogeneity in cultural gender norm effect



Notes: The figure plots the effect of culture of traditional gender norms on the gender gap in math, when the effect is allowed to differ by mothers' education level (no, or at least some university education), cultural density of the LLM area (measured by share of same-ethnicity residents, below or above the median share), and how long it has been since the mother immigrated to Sweden (below or above the median number of years (21)), family size, and the mean final grade mark quintile of the siblings. The figure plots 90% confidence intervals. The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Regressions are based on the specification in column (4) of Table 4. Standard errors are clustered at the source country.

Figure 4: Heterogeneity in mitigation effect



Notes: The figure plots the mitigation effect of local labour market FLFP, when the effect is allowed to differ by mothers' education level (no, or at least some university education), cultural density of the LLM area (measured by share of same-ethnicity residents, below or above the median share), and how long it has been since the mother immigrated to Sweden (below or above the median number of years (19)), family size, and the mean final grade mark quintile of the siblings. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl $\times$ LLM FLFP and girl $\times$ norms $\times$ LLM FLFP. The figure plots 90% confidence intervals. The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Regressions are based on the specification in column (5) in Table 5. Standard errors are clustered at the source country $\times$ assigned local labour market $\times$ immigration year level.

Table 7: Robustness checks: cultural gender norm effect

	(1)	(2)	(3)	(4)	(5)	(6)
	Drop FI	Both immigrants	Max 5 years	Birth weight	SES advantage	Selective migration
Girl	-0.093***	-0.090***	-0.076***	-0.066***	-0.073***	-0.073***
	(0.011)	(0.016)	(0.011)	(0.013)	(0.011)	(0.011)
Girl × norms	-0.048	-0.054*	-0.073**	-0.069**	-0.066**	-0.073**
	(0.031)	(0.031)	(0.030)	(0.030)	(0.031)	(0.029)
Girl × birth weight				-0.012*		
				(0.006)		
Girl × mother uni					0.005	
					(0.017)	
Girl × distance						-0.021**
						(0.010)
p(girl × norms = 0)	0.179	0.085	0.060	0.081	0.085	0.047
R-squared	0.70	0.69	0.71	0.70	0.70	0.70
Observations	28,243	23,575	36,508	37,667	39,007	39,007
	Indicators					
Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.09). Column (1) drops Finland, the country contributing to the largest immigrant inflow. Column (2) includes only children whose parents are both immigrants. Column (3) includes only children born a maximum of five years apart. Column (4) controls for the potentially confounding gender gap in birth weight. Column (5) controls for a potential gender gap effect of mothers' socio-economic status (measured by a university indicator), while column (6) controls for a potential gender gap effect of geographical distance between source countries and Stockholm, Sweden. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. Standard errors in parentheses are clustered at the source-country level. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 8: Robustness checks: mitigation effect

	(1)	(2)	(3)	(4)	(5)	(6)
	Drop Sthlm	Both immigrants	Max 5 years	Birth weight	SES advantage	Selective migration
Girl	-0.058**	-0.066**	-0.068**	-0.051*	-0.067**	-0.066**
	(0.029)	(0.029)	(0.028)	(0.029)	(0.030)	(0.027)
Girl × norms	-0.036	-0.049	-0.054*	-0.049	-0.055*	-0.068*
	(0.035)	(0.033)	(0.033)	(0.033)	(0.032)	(0.035)
Girl × LLM FLFP	0.009	0.010	-0.013	0.012	0.014	0.006
	(0.053)	(0.054)	(0.053)	(0.054)	(0.054)	(0.052)
Girl × norms × LLM FLFP	0.103***	0.083***	0.090***	0.077**	0.084***	0.078**
	(0.033)	(0.031)	(0.030)	(0.031)	(0.029)	(0.031)
Girl × mother uni					0.016	
					(0.074)	
Girl × mother uni × LLM FLFP					-0.035	
					(0.069)	
Girl × birth weight				-0.012		
				(0.027)		
Girl × birth weight × LLM FLFP				0.014		
				(0.025)		
Girl × distance						-0.032
						(0.029)
Girl × distance × LLM FLFP						-0.013
						(0.026)
Mitigation	0.112	0.093	0.076	0.090	0.099	0.085
p(mitigation = 0)	0.076	0.130	0.216	0.158	0.105	0.170
R-squared	0.69	0.69	0.70	0.70	0.70	0.70
Observations	5,421	5,619	5,746	5,909	6,102	6,102
	Indicators					
Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes
LLM controls	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.06), and a one-standard-deviation increase in local labour market FLFP (0.04). Column (1) drops Stockholm, the largest and most urban LLM. Column (2) includes only children whose parents are both immigrants. Column (3) includes only children born a maximum of five years apart. Column (4) controls for a potentially confounding gender gap in birth weight. Column (5) controls for a potential gender gap effect of mothers' socio-economic status (measured by a university indicator), while column (6) controls for a potential gender gap effect of geographical distance between source countries and Stockholm, Sweden. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration, all interacted with a gender indicator. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl × LLM FLFP and girl × norms × LLM FLFP, the corresponding p-value is based on the standard errors of this combined estimate. Standard errors in parentheses are clustered the source country × assigned local labour market × immigration year level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 9: Gender norm effect: alternative norm measures

	(1)	(2)	(3)	(4)	(5)	(6)
	1-GGI 2010	WVS	1-FLFP	1-FLFPdec	Avg 1-GGI	-PISA
Girl	-0.065***	-0.075***	-0.072***	-0.086***	-0.067***	-0.070***
	(0.016)	(0.012)	(0.013)	(0.016)	(0.012)	(0.012)
Girl × 1-GGI2010	-0.094*					
	(0.046)					
Girl × WVS		-0.050**				
		(0.019)				
Girl × 1-FLFP			-0.043**			
			(0.019)			
Girl × 1-FLFPdec				-0.043**		
				(0.018)		
Girl × avg 1-GGI					-0.062***	
					(0.023)	
Girl × -PISA						-0.037**
						(0.017)
p(girl × norms = 0)	0.120	0.085	0.088	0.035	0.004	0.062
R-squared	0.70	0.70	0.70	0.70	0.70	0.70
Observations	32,446	37,189	39,007	25,544	37,063	30,820
	Indicators					
Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Country cntrl	Yes	Yes	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The measures of culture of traditional gender norms are the source-country GGI of the Gender Gap Report from 2010 (column (1)), source-country-level of agreement with the statement *"A university education is more important for a boy than for a girl."* from the World Values Survey in column (2), source-country-level 1-FLFP rates measured in 2007-2010 in column (3), source-country-level 1-FLFP rates from the decade of mothers' immigration in column (4), average source-country 1-GGI scores of both mothers' and fathers' source countries in column (5), and the average gender gap (girls' scores - boys' scores) from the 2015 PISA evaluations in column (6). In order to have the norm measures work in the same direction (i.e. increase with gender-traditionality), the PISA gender gap is defined as -PISAgap in the regressions. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling (at decade of immigration in column (4)), both interacted with a gender indicator. Standard errors in parentheses are clustered at the source-country level. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 10: Mitigation effect: alternative measures

	(1)	(2)	(3)	(4)
	Wage ratio	Mun FLFP	1-GGI 2010	Avg 1-GGI
Girl	-0.065**	-0.134	-0.073**	-0.071***
	(0.028)	(0.188)	(0.029)	(0.025)
Girl × norms (1-GGI)	-0.056*	-0.060*		
	(0.032)	(0.032)		
Girl × LLM wage ratio	0.017			
	(0.028)			
Girl × norms (1-GGI) × LLM wage ratio	0.076***			
	(0.029)			
Girl × mun FLFP		-0.058		
		(0.039)		
Girl × norms (1-GGI) × mun FLFP		0.049*		
		(0.026)		
Girl × 1-GGI2010			-0.061*	
			(0.036)	
Girl × avg 1-GGI				-0.039
				(0.033)
Girl × LLM FLFP			0.025	0.011
			(0.056)	(0.071)
Girl × 1-GGI 2010 × LLM FLFP			0.080***	
			(0.029)	
Girl × avg 1-GGI × LLM FLFP				0.079***
				(0.029)
Mitigation	0.094	-0.009	0.105	0.090
p(mitigation = 0)	0.007	0.861	0.093	0.283
R-squared	0.70	0.70	0.70	0.69
Observations	6,102	6,102	5,095	5,704
	Indicators			
Birth order	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes
LLM controls	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Column (1) uses 1-GGI score as the norm measure and local labour market area female-over-male wage ratio (at time of immigration) as the measure of gender quality. Column (2) uses 1-GGI score as the norm measure and municipality FLFP rates (at time of immigration) as the measure of gender quality. Column (3) and (4) investigate the mitigation effect of LLM FLFP when using 1-GGI from 2010 or average 1-GGI of both parents' source countries as the norm measures, respectively. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration (assigned municipality in column (2)), all interacted with a gender indicator. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl × LLM FLFP and girl × norms × LLM FLFP, the corresponding p-value is based on the standard errors of this combined estimate. Standard errors in parentheses are clustered the source country × assigned local labour market × immigration year level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Appendix

Table A1: Alternative source-country measures of gender norms

	(1)	(2)	(3)	(4)	(5)	(6)
	1-GGI (2010)	WVS	1-FLFP	1-FLFPdec	Avg 1-GGI	PISA gap
Chile	0.299	0.262	0.391	0.616	0.273	-18.3
Czechoslovakia		0.358	0.264	0.244	0.269	-6.4
Denmark	0.223		0.139	0.284	0.204	-9.4
Ethiopia	0.398	0.08	0.153	0.391	0.339	
Finland	0.174	0.105	0.12	0.263	0.183	7.5
Germany	0.247	0.136	0.214	0.377	0.210	-16.6
Greece	0.309		0.321	0.560	0.283	-0.1
Iran, Islamic Rep.	0.407	0.476	0.771	0.882		
Iraq		0.446	0.831	0.878	0.447	
Italy	0.324	0.074	0.364	0.528	0.248	-19.9
Lebanon	0.392	0.312	0.683	0.789	0.402	-22.1
Norway	0.160	0.075	0.116	0.336	0.173	2.3
Poland	0.296	0.218	0.255	0.224	0.255	-11.4
Russian Federation	0.296	0.288	0.198	0.198	0.277	-6.0
Thailand	0.309	0.283	0.203	0.164	0.246	2.9
Turkey	0.412	0.258	0.639	0.499	0.369	-5.9
Unites States	0.259	0.098	0.183	0.320	0.233	-8.5
Vietnam	0.322	0.221	0.116	0.092	0.297	3.0
Yugoslavia		0.183	0.294	0.280	0.263	-5.6
N RQ1	60,708	72,408	75,780	60,163	71,919	65,693
Mean RQ1	0.277	0.207	0.330	0.400	0.271	-3.8
SD RQ1	0.098	0.115	0.239	0.233	0.088	9.6
N RQ2	9,844	11,773	11,773	11,773	11,101	7,693
Mean RQ2	0.368	0.292	0.512	0.563	0.348	-12.4
SD RQ2	0.049	0.120	0.241	0.260	0.069	7.6

Notes: The table reports country-specific values for the math gender gap, and each of the source-country variables. Column (1) presents 1-GGI from the 2010 Gender Gap Report. Column (2) contains the country level of agreement with the statement *"A university education is more important for a boy than for a girl."* from the World Values Survey, and column (3) contains country-level 1-FLFP rates where FLFP is defined as female-over-male labour force participation rates in 2007–2010. Column (4) contains country-level 1-FLFP rates where FLFP is measured in the decade of mothers' emigration, and column (5) present the average 1-GGI, using both parents GGI (2018) score. Columns (6) present the country-level average gender gap in math from the 2015 PISA evaluations (defined as girls' scores' - boys scores). The bottom panels reports mean and cross-country standard deviation of each of the variables, for both sample definitions.

Table A2: Cross-correlation of source-country variables

	1-GGI	GDP	Edu	1-GGI 2010	WVS	1-FLFP	1-FLFPdec	Avg 1-GGI	PISA gap
RQ1 sample									
1-GGI	1								
GDP	-0.834	1							
Edu	-0.905	0.789	1						
1-GGI 2010	0.996	-0.841	-0.878	1					
WVS	0.901	-0.927	-0.878	0.900	1				
1-FLFP	0.912	-0.661	-0.738	0.920	0.768	1			
1-FLFP dec	0.448	-0.237	-0.311	0.443	0.384	0.635	1		
Avg 1-GGI	0.931	-0.748	-0.816	0.933	0.814	0.903	0.462	1	
PISA gap	-0.653	0.613	0.504	-0.660	-0.698	-0.599	-0.513	-0.574	1
RQ2 sample									
1-GGI	1								
GDP	-0.532	1							
Edu	-0.722	0.840	1						
1-GGI 2010	0.934	-0.388	-0.734	1					
WVS	0.811	-0.450	-0.471	0.602	1				
1-FLFP	0.912	-0.209	-0.508	0.888	0.832	1			
1-FLFP dec	0.754	-0.185	-0.380	0.638	0.929	0.903	1		
Avg 1-GGI	0.917	-0.500	-0.679	0.860	0.751	0.844	0.701	1	
PISA gap	-0.213	-0.043	-0.266	0.078	-0.694	-0.351	-0.622	-0.192	1

Notes: The table reports the cross-correlation between the source-country variables used in this study. The top panel reports the correlation for the sample of children whose mothers are immigrants, and the bottom panel reports the correlation for children whose mothers were subjected to the refugee placement policy.

Table A3: Cross-correlation of LLM area variables

	RQ2 sample				
	FLFP	Wage ratio	Population	Income	Cultural density
FLFP	1				
Wage ratio	0.572	1			
Population	0.833	0.507	1		
Income	0.475	0.179	0.515	1	
Cultural density	-0.571	-0.294	-0.554	-0.361	1
Mean	0.891	0.353	335,346	101,550	0.04
SD	0.043	0.021	360,268	15,377	0.006

Notes: The table reports the cross-correlation, as well as the mean and standard deviation, of the local labour market area variables used in this study. The sample consist of children whose mothers were subjected to the refugee placement policy.

Table A4: Effect of culture of traditional gender norms on the gender gap in birth weight

	(1)	(2)	(3)
Girl	-134.202***	-134.005***	-133.667***
	(6.827)	(5.622)	(2.957)
Norms	-52.112***		
	(2.318)		
Girl × norms	-4.837	-12.992	-15.801*
	(5.801)	(12.618)	(8.091)
p(girl × norms = 0)	0.451	0.341	0.210
R-squared	0.03	0.03	0.76
Observations	74,106	74,106	37,667
	Indicators		
Birth order	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Country FE	No	Yes	No
Country controls	No	Yes	Yes
Family FE	No	No	Yes

Notes: The dependent variable is birth weight, measured in grams. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.09). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. Standard errors in parentheses are clustered at the source country. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A5: Placebo balance test for refugee placement policy

	(1)	(2)	(3)	(4)
Norms (1-GGI)	0.086**	0.122***	0.124***	0.124***
	(0.042)	(0.033)	(0.031)	(0.031)
Cultural density		-0.297***	-0.315***	-0.316***
		(0.036)	(0.037)	(0.037)
Mother's edu (im.)		0.010	0.004	-0.000
		(0.007)	(0.007)	(0.007)
Family size		-0.060***	-0.051***	-0.062***
		(0.013)	(0.014)	(0.014)
Mother's cohort				-0.005*
				(0.003)
R-squared	0.01	0.10	0.12	0.12
Observations	11,773	11,773	11,773	11,773
	Indicators			
Immigration year FE	No	No	Yes	Yes

Notes: The table reports the correlation between local labour market area FLFP and mothers' source-country gender norms and individual characteristics. The dependent variable is the FLFP of the local labour market the family lives in when the child is attending ninth grade. All maternal characteristics are measured at the time of immigration. Cultural density is the share of local labour market residents from the same source country as the mother who is being placed there. Standard errors are clustered at the source country × assigned local labour market × immigration year level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A6: Effect of culture of traditional gender norms on gender gap in final marks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math		English		Swedish		Final	
Girl	0.023***	0.024**	0.147***	0.169***	0.515***	0.524***	0.324***	0.325***
	(0.007)	(0.011)	(0.014)	(0.014)	(0.010)	(0.013)	(0.008)	(0.011)
Girl × norms	-0.058***	-0.021	-0.061**	-0.014	-0.072***	-0.033	-0.064***	-0.016
	(0.014)	(0.022)	(0.028)	(0.032)	(0.019)	(0.029)	(0.015)	(0.025)
p(girl × norms = 0)	0.010	0.407	0.138	0.703	0.024	0.324	0.009	0.569
R-squared	0.04	0.67	0.06	0.69	0.09	0.67	0.05	0.73
Observations	75,163	38,583	75,163	38,583	62,686	29,749	75,780	39,007
	Indicators							
Birth order	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	No	Yes	No	Yes	No	Yes	No
Family FE	No	Yes	No	Yes	No	Yes	No	Yes
Country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

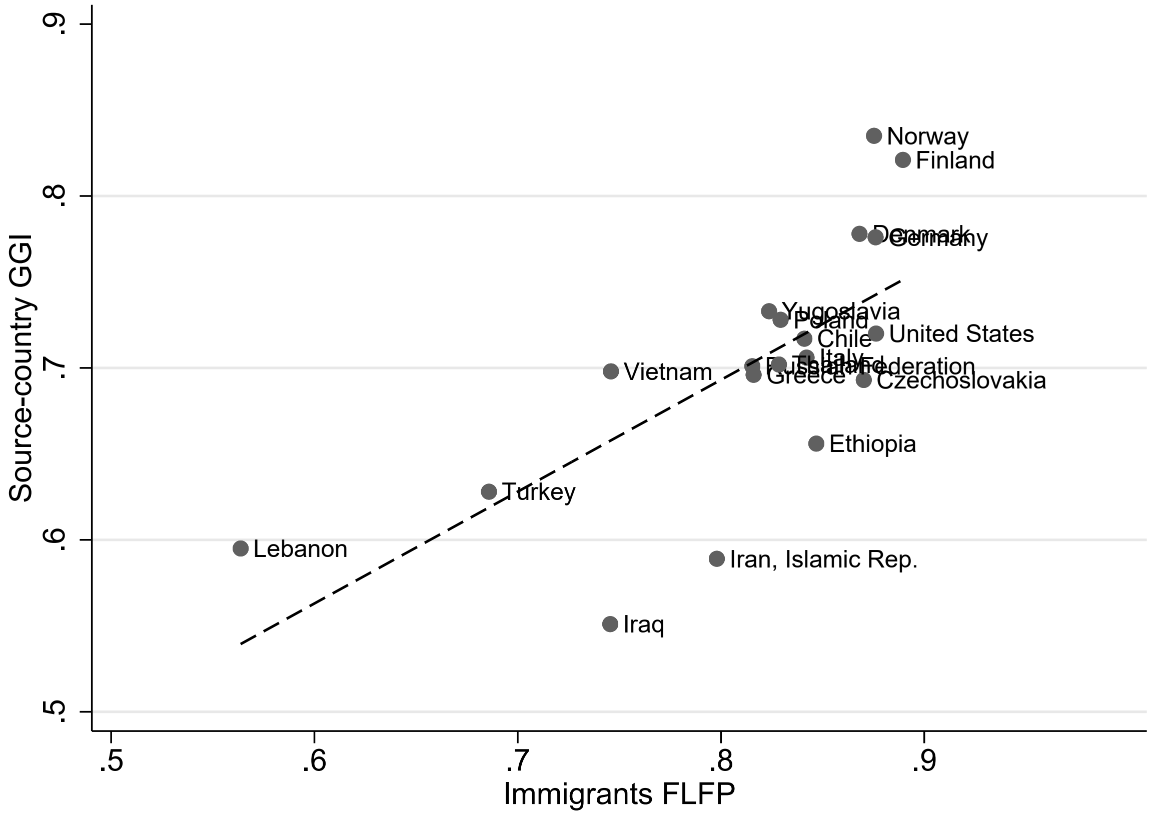
Notes: The dependent variables are the student's final mark in math (1)–(2), English (3)–(4), Swedish (5)–(6) and the total final ninth grade mark (7)–(8). All outcomes are standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in culture of traditional gender norms (0.09). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. Standard errors in parentheses are clustered at the source-country level. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A7: Mitigation effect: alternative measures continued

	(1)	(2)	(3)
	WVS	FLFP	FLFP Dec.
Girl	-0.064**	-0.065**	-0.065**
	(0.028)	(0.027)	(0.028)
Girl × WVS boy uni	-0.055*		
	(0.032)		
Girl × FLFP		-0.051*	
		(0.027)	
Girl × FLFP decade			-0.063**
			(0.029)
Girl × LLM FLFP	0.013	0.015	0.013
	(0.052)	(0.052)	(0.052)
Girl × WVS boy uni × LLM FLFP	0.072**		
	(0.029)		
Girl × FLFP × LLM FLFP		0.090***	
		(0.028)	
Girl × FLFP decade × LLM FLFP			0.076***
			(0.029)
Mitigation	0.085	0.105	0.088
p(mitigation = 0)	0.170	0.083	0.145
R-squared	0.70	0.70	0.70
Observations	6,102	6,102	6,102
		Indicators	
Birth order	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Country controls	Yes	Yes	Yes
LLM controls	Yes	Yes	Yes
Family FE	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. All columns use LLM area FLFP as the measure of neighbourhood gender equality. Column (1) source-country-level of agreement with the statement *"A university education is more important for a boy than for a girl."* from the World Values Survey as the norm measure. Column (2) and (3) uses source-country-level 1-FLFP rates measured between 2007–2010 or in the decade of mothers' immigration (around 1990), respectively. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration, all interacted with a gender indicator. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl × LLM FLFP and girl × norms × LLM FLFP, the corresponding p-value is based on the standard errors of this combined estimate. Standard errors in parentheses are clustered the source country × assigned local labour market × immigration year level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Figure A1: Correlation: GGI and immigrant FLFP rates



Notes: The figure plots the correlation between source-country GGI 2018 score and the average current FLFP ratio of the first generation. The correlation coefficient is 0.7.

Online Appendix

Table OA1: Norm and mitigation effect: different cluster strategies

	(1)	(2)	(3)	(4)
	Norm effect (RQ1 sample)			
Girl	-0.072***	-0.072***	-0.072***	-0.072***
	(0.017)	(0.012)	(0.011)	(0.011)
Girl × norms	-0.065**	-0.065**	-0.065***	-0.065***
	(0.031)	(0.028)	(0.023)	(0.022)
R-squared	0.70	0.70	0.70	0.70
Observations	39,007	39,007	39,007	39,007
Cluster SE	Country × LLM	Country & LLM	Country × cohort	Family
Birth order	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes
	Mitigation effect (RQ2 sample)			
Girl	-0.064**	-0.064**	-0.064**	-0.064**
	(0.028)	(0.025)	(0.026)	(0.027)
Girl × norms	-0.054*	-0.054	-0.054	-0.054*
	(0.032)	(0.037)	(0.035)	(0.031)
Girl × LLM FLFP	0.008	0.008	0.008	0.008
	(0.052)	(0.071)	(0.052)	(0.051)
Girl × norms × LLM FLFP	0.085***	0.085**	0.085***	0.085***
	(0.028)	(0.026)	(0.032)	(0.027)
Mitigation	0.093	0.093	0.093	0.093
p(mitigation = 0)	0.119	0.271	0.130	0.113
R-squared	0.70	0.70	0.70	0.70
Observations	6,102	6,102	6,102	6,102
Cluster SE	Country × LLM	Country & LLM	LLM × im. year	Family
Birth order	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Country controls	Yes	Yes	Yes	Yes
LLM controls	Yes	Yes	Yes	Yes
Family FE	Yes	Yes	Yes	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in cultural gender norms. Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration, all interacted with a gender indicator. Standard errors are reported in parentheses, and the cluster strategy is reported for each column. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table OA2: Effect of fathers' gender norms on the math gender gap

	(1)	(2)	(3)	(4)	(5)
Girl	-0.064***	-0.062***	-0.059***	-0.064***	-0.064***
	(0.007)	(0.005)	(0.006)	(0.008)	(0.008)
Norms	-0.034				
	(0.042)				
Girl × norms	-0.048***	-0.067***	-0.070***	-0.049**	-0.048**
	(0.006)	(0.010)	(0.009)	(0.021)	(0.022)
p(girl × norms = 0)	0.010	0.012	0.006	0.173	0.204
R-squared	0.01	0.04	0.12	0.69	0.71
Observations	77,903	77,903	77,903	41,023	25,116
			Indicators		
Birth order	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	No	No
Father controls	No	No	Yes	No	No
Country controls	No	Yes	Yes	Yes	Yes
Family FE	No	No	No	Yes	No
Sibling-pair FE	No	No	No	No	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. The effects are estimated for a one-standard-deviation increase in father's cultural gender norms (0.09). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. Fathers' controls are cohort, education level, household income and family size. Standard errors in parentheses are clustered at the source country. The p-value from standard errors obtained using a wild cluster bootstrap are shown at the bottom. * p<0.1, ** p<0.05, *** p<0.01.

Table OA3: Balance test for refugee placement policy: fathers

	(1)	(2)	(3)	(4)
Norms	-0.128**	-0.032	-0.037	-0.035
	(0.058)	(0.044)	(0.046)	(0.046)
Cultural density		-0.516***	-0.498***	-0.498***
		(0.045)	(0.047)	(0.047)
Father's edu (im.)		-0.007	0.001	0.003
		(0.007)	(0.007)	(0.006)
Family size		-0.021*	-0.033***	-0.027**
		(0.012)	(0.012)	(0.013)
Father's cohort				0.003
				(0.004)
R-squared	0.02	0.28	0.29	0.29
Observations	11,726	11,726	11,726	11,726
		Indicators		
Immigration year FE	No	No	Yes	Yes

Notes: The table reports the correlation between local labour market area FLFP and fathers' source-country gender norms and individual characteristics. The dependent variable is the FLFP of the assigned local labour market. All father characteristics are measured at the time of immigration. Cultural density is the share of local labour market residents from the same source country as the mother who is being placed there. Standard errors are clustered at the source country × assigned local labour market × immigration year level. * p<0.1, ** p<0.05, *** p<0.01.

Table OA4: Local labour market and mitigation effect: fathers

	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.069**	-0.067***	-0.066***	-0.067***	-0.070**	-0.076***
	(0.027)	(0.020)	(0.019)	(0.019)	(0.029)	(0.028)
Girl × norms	-0.004	-0.049**	-0.048**	-0.040*	-0.003	-0.016
	(0.033)	(0.021)	(0.020)	(0.023)	(0.033)	(0.033)
Girl × LLM FLFP		-0.006	-0.003	0.002	-0.013	-0.008
		(0.019)	(0.019)	(0.031)	(0.045)	(0.048)
Girl × norms × LLM FLFP		0.028	0.033	0.028	-0.005	-0.010
		(0.021)	(0.020)	(0.021)	(0.034)	(0.034)
Mitigation		0.022	0.030	0.030	-0.018	-0.019
p(mitigation = 0)	0.430	0.455	0.299	0.400	0.734	0.746
R-squared	0.70	0.11	0.13	0.13	0.70	0.71
Observations	6,009	11,687	11,687	11,687	6,009	3,564
	Indicators					
Birth order	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	Yes	Yes	No	No
LLM × Im.Year FE	No	Yes	Yes	Yes	No	No
Country controls	Yes	No	No	Yes	Yes	Yes
LLM controls	Yes	No	No	Yes	Yes	Yes
Father controls	No	No	Yes	Yes	No	No
Family FE	Yes	No	No	No	Yes	No
Sibling-pair FE	No	No	No	No	No	Yes

Notes: The dependent variable is the student's test score on the national standardised test in math, standardised to have a mean of 0 and standard deviation of 1. Local labour market FLFP is measured at the assigned local labour market, in the year in which the father immigrated to Sweden. The effects are estimated for a one-standard-deviation increase in cultural gender norms (0.06), and a one standard deviation increase in local labour market FLFP (0.04). Country controls are standardised source-country gdp per capita and gross enrolment in secondary schooling, both interacted with a gender indicator. LLM controls are standardised population, average income and cultural density of the assigned local labour market area at the time of immigration, all interacted with a gender indicator. Fathers' controls are cohort, education level and family size, all of which are measured at time of immigration. The estimate of the mitigation effect is obtained by linearly adding the coefficients for girl × LLM FLFP and girl × norms × LLM FLFP, the corresponding p-value is based on the standard errors of this combined estimate. Standard errors are clustered at the source country × assigned local labour market × immigration year level. * p<0.1, ** p<0.05, *** p<0.01.