Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1659

Endogenous Gender Peer Effects in the Classroom

Erik Liss
Linköping University

Maria Brandén
Linköping University

Olof Ejermo
Lund University

Martin Hällsten
Stockholm University

Karl Wennberg
Linköping University

Peer effects in schools are well documented, but it remains unclear whether they reflect con- textual effects of classmates' predetermined characteristics, or endogenous peer effects, where students respond to their peers' achievement and effort. We study gender-specific endogenous peer effects in mathematics achievement among Swedish grade-9 students, using register data with classroom identifiers. We identify endogenous peer effects using variation in peer-group size, an approach whose precision is strongest in small groups, well below typical classroom size. We instead exploit uneven gender composition to construct smaller, same-gender peer groups, which yields more identifying variation. We find substantial endogenous peer effects in mathematics achievement: approximately 0.48 standard deviations for girls and 0.20 for boys, while cross-gender effects appear comparatively modest. We also develop a simple analytical framework linking stronger within-gender endogenous peer effects among girls to the positive association between girls' classroom share and academic achievement, with implications for classroom allocation and policy design.

VERSION: November 2026

Suggested citation: Liss, Erik, Maria Brandén, Olof Ejermo, Martin Hällsten, and Karl Wennberg. (2026). Endogenous Gender Peer Effects in the Classroom. (EdWorkingPaper: 26-1659). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/vxfx-ep66

Endogenous Gender Peer Effects in the Classroom*

Erik Liss† Maria Brandén‡ Olof Ejermo§ Martin Hällsten¶
Karl Wennberg‖

2026

Abstract

Peer effects in schools are well documented, but it remains unclear whether they reflect contextual effects of classmates' predetermined characteristics, or endogenous peer effects, where students respond to their peers' achievement and effort. We study gender-specific endogenous peer effects in mathematics achievement among Swedish grade-9 students, using register data with classroom identifiers. We identify endogenous peer effects using variation in peer-group size, an approach whose precision is strongest in small groups, well below typical classroom size. We instead exploit uneven gender composition to construct smaller, same-gender peer groups, which yields more identifying variation. We find substantial endogenous peer effects in mathematics achievement: approximately 0.48 standard deviations for girls and 0.20 for boys, while cross-gender effects appear comparatively modest. We also develop a simple analytical framework linking stronger within-gender endogenous peer effects among girls to the positive association between girls' classroom share and academic achievement, with implications for classroom allocation and policy design.

Keywords: Peer effects, linear-in-means model, social interactions, gender

*We thank Vincent Boucher, Habiba Djebbari, Torsten Santavirta, Andreas Diemer and participants at the Annual Congress of the International Institute of Public Finance (IIPF), the Annual Conference of the European Society for Population Economics (ESPE), and seminar participants at Uppsala Urban Lab, the Ratio Institute, and the Swedish Institute for Social Research (SOFI) for helpful comments and suggestions. This project was generously funded by the Swedish Research Council for Health, Working Life and Welfare (Forte), project 2022-00390, and the Torsten Söderberg Foundation, project E44/22.

†Institute for Analytical Sociology at Linköping University, Swedish Institute for Social Research at Stockholm University, and Ratio Institute

‡Institute for Analytical Sociology at Linköping University

§Department of Economic History, School of Economics and Management at Lund University, and Ratio Institute

¶Swedish Institute for Social Research at Stockholm University, and Institute for Analytical Sociology at Linköping University

‖Stockholm School of Economics, and Institute for Analytical Sociology at Linköping University

1 Introduction

Research on peer effects in education examines how peers in the same learning environment shape student outcomes. Central questions concern how peer groups amplify or diminish gender gaps in educational choices and academic performance. This has direct practical implications for how students are allocated to classrooms and schools, and for policies such as whether to encourage or restrict school choice.

The literature tends to find a positive influence of the share of girls in the classroom (Hoxby 2000, Lavy & Schlosser 2011, Hu 2015). These results are typically based on studies drawing on cohort-level variation in peer composition within schools (Hoxby & Weingarth 2005) or (more rarely) experimental variation (Lu & Anderson 2015). However, the mechanisms underlying this association remain unclear. Manski (1993) distinguishes between contextual peer effects, where outcomes respond to peers' predetermined characteristics, and endogenous peer effects, where students respond to peers' outcomes. The positive association between the classroom share of girls and achievement could therefore reflect girls' characteristics improving the learning environment—through lower disruption, stronger study norms, or more cooperative interactions—or gender-specific feedback through peers' achievement and effort, such as competition, imitation, or adjustments in effort. Previous studies typically do not isolate endogenous effects and instead estimate reduced-form compositional effects, often using school- rather than classroom-level data. Two notable exceptions are Opper (2019) and Horrace et al. (2025), both of whom find that endogenous peer spillovers operate primarily within same-gender peer groups, though neither study produces separate structural estimates of the endogenous effect for boys and girls.

Recent years have seen substantial methodological development. Exploiting overlapping networks of peers-of-peers offers the hope to identify peer effects and overcome the reflection problem (Bramoullé et al. 2009, De Giorgi et al. 2010). This paper builds on these insights, more specifically following Lee (2007), Boucher et al. (2014) and Davezies et al. (2009) in using group-size variation to identify peer effects. When peer groups are defined at the full-classroom or cohort level, however, identification is disproportionately driven by the relatively few small groups, which may be atypical and unrepresentative in settings where classrooms typically contain 20 to 30 students. This combination of weaker identification in larger groups together with atypical small classes disproportionately driving identification may help explain some of the imprecision observed in prior applications of the method. For example, Boucher et al. (2014) find insignificant estimates for most school subjects, but for mathematics they report an endogenous effect of 0.82 — close to the upper bound of the stability region of a social multiplier.

In this paper, we address the identification problem posed by large group sizes by estimating the model separately by gender, exploiting the fact that most classrooms have an uneven gender composition. This provides stronger identifying variation by generating smaller effective group

sizes within each gender. This approach also aligns well with evidence that endogenous peer effects are primarily channeled within gender (Horrace et al. 2025), and enables us to address questions related to heterogeneous peer effects by gender as well as the association between the classroom share of girls and school outcomes.

Using Swedish register data on students completing compulsory schooling (grade 9), we estimate substantial endogenous peer effects: increasing grades of same-gender peers by one standard deviation raises own grades on math tests by approximately 0.48 standard deviations for girls and 0.2 for boys. The gender difference in endogenous peer effects translates into markedly different social multipliers, approximately 1.93 for girls and 1.25 for boys. A shock raising girls' initial achievement by 0.1 SD is thus amplified to roughly 0.19 SD in equilibrium, whereas the same shock among boys is amplified to about 0.125 SD. Next, randomly reassigning gender labels yields much weaker endogenous effects, corroborating the assumption that relevant peer interactions operate primarily within gender rather than through opposite-gender classmates.

Theoretically, endogenous peer effects can arise from different microfoundations, often distinguished as local-average versus local-aggregate models (Ushchev & Zenou 2020). In local-average models, individuals respond to peers' average behavior or outcomes, consistent with social norms, conformity, or peer pressure. Local-aggregate models instead capture spillovers from peers' total effort or achievement, where the combined activity of peers raises individual outcomes through channels such as knowledge sharing or complementarities (Liu et al. 2014). The stronger endogenous peer effects for girls likely reflect a greater role for local-average mechanisms—for example, social norms and conformity (Shkoza et al. 2023). Experimental research suggests that females' improvements from working in groups are often driven by feeling needed in the group, in particular that friends care about the impact of one's performance on the group outcome, pointing to a local-average mechanism (Weber et al. 2009). In an educational setting, Hahn et al. (2020) find large gains for low-ability girls from working with friends rather than non-friend peers, regardless of the group's gender composition, suggesting that friendship matters rather than gender per se. However, because friendship networks are often gender-segregated, cross-gender ties — even when they generate peer effects of comparable strength — are too rare to contribute meaningfully once peer effects are aggregated to the class level. Within-gender endogenous peer effects are therefore expected to dominate in class-level data.

In addition to estimating average peer effects, we conduct a set of supplementary analyses that shed light on the nature and heterogeneity of the estimated dynamics. First, splitting classrooms by average parental education indicates that endogenous effects are strongest in lower socioeconomic status (SES) environments and are weaker in classrooms with higher parental education. Second, when we split classrooms by the dispersion of parental education, endogenous effects are most pronounced in the lowest- and highest-dispersion terciles. Finally, endogenous effects vary with class gender composition: girls exhibit sizable effects in both

minority and majority classes, whereas boys' effects are much stronger only when they are in majority. In boys-majority classes, the endogenous effect for boys reaches 0.48—matching the girls' benchmark—and the implied social multiplier climbs from 1.25 to approximately 1.93.

Finally, we develop a simple analytical model to show how our estimated endogenous peer effects map into the reduced-form literature on classroom gender composition. Previous studies report that classes with a higher share of girls perform better on average (Hoxby 2000, Lavy & Schlosser 2011), but recent evidence suggests that this pattern is driven primarily by gains among girls, while the effects for boys are weaker and sometimes negative (Proud 2014, Lu & Anderson 2015, Lénárd 2026, Briole 2021, Oosterbeek & Van Ewijk 2014). This is also what we see descriptively in our data for mathematics: a higher share of girls is associated with higher achievement for girls, whereas the corresponding relationship for boys is slightly negative but statistically insignificant. Lénárd (2026) reports a similar pattern on Swedish data, with small positive effects for girls and small negative effects for boys.

The analytical model helps explain these patterns in three progressive simulations. Stronger within-girl endogenous peer effects alone generate a composition pattern that is positively sloped for girls and weakly negatively sloped for boys, matching most of our descriptive evidence. Intuitively, significant within-gender but no cross-gender endogenous peer effects imply that each gender benefits from a larger share of same-gender peers. When this effect is stronger for girls, the within-gender channel operates more powerfully for them, producing a steeper achievement profile for girls than for boys — precisely the pattern we observe. Next, allowing the within-boy endogenous effect to vary with composition – consistent with our finding that boys' endogenous effects are stronger in boy-majority classes – this overstates boys' outcomes in girl-scarce classes compared to the descriptive aggregated pattern we see in the data. However, adding a positive within-girl contextual effect reproduces the full empirical pattern and is consistent with our estimated peer effects. This suggests that endogenous peer effects drive the slopes with respect to gender composition, while contextual effects likely also help account for the level differences we observe across boys and girls. The importance of endogenous peer effects in explaining the gender-share relationship implies that it cannot be understood as a simple linear compositional effect, whereby reallocating girls predictably improves outcomes by creating a better classroom environment. Instead, changing the share of girls also alters the strength of the within-gender peer multiplier for both boys and girls. The effect of moving one girl therefore depends on the gender composition of both the classroom she leaves and the classroom she enters.

A further implication of the analytical model is that aggregate achievement is maximized under complete gender segregation. Because endogenous peer effects are strong within gender but weak across genders, a higher own-gender share raises individual achievement, and concentrating boys and girls into separate classes therefore raises the aggregate. Similar theoretical results have previously been shown under specific assumptions by Arnott & Rowse (1987) and Sarkisian & Yamashita (2024). This aggregate gain, however, comes with a distributional cost

under partial segregation: students who belong to a small gender minority experience a lower share of same-gender peers and therefore a weaker own-group multiplier effect, leaving them worse off than they would be in more balanced classrooms. From a distributional perspective, balanced classrooms may therefore be preferable to partial segregation, even if partial segregation raises overall achievement. Only complete segregation eliminates this minority-harm channel by leaving no minority at all.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the identification strategy, and Section 4 outlines the estimation approach. Section 5 introduces the data and measurement choices. Section 6 presents the main results and a set of robustness and heterogeneity analyses. Section 7 presents a simple analytical model to clarify how gender-specific endogenous peer effects can shape observed compositional patterns. Section 8 concludes with a discussion of implications and directions for future research.

2 Related literature

The Coleman Report, launched in 1966, showed that educational outcomes can differ substantially even when schools have similar material resources (Coleman et al. 1966). Since then, much of research and policy has focused on how sorting and peer effects interact to generate unequal educational opportunities. Early studies examined peer effects through sociological path models, notably the Wisconsin model, and through educational production functions in economics (Sewell et al. 1969, Hanushek 1972).

However, as Manski (2000) argues, much of this empirical literature lacked the methodological rigour needed to address the identification problems in the peer effects literature. Manski (1993) formalized these challenges in his seminal paper on the identification of peer effects. In Manski's framework, individual outcomes depend on endogenous peer effects (the influence of peers' outcomes), exogenous or contextual peer effects (the influence of peers' characteristics), and correlated effects arising from shared environments such as school or teacher quality (Manski 1993). This specification made explicit the reflection problem: because individuals both influence and are influenced by their peers, group outcomes emerge from a simultaneous feedback process, which makes it difficult to disentangle the causal effect of the group on the individual from the individual's contribution to the group. We discuss the technical aspects of the Manski model in more detail in the next section.

One solution to the reflection problem has been to simply refrain from estimating endogenous peer effects and instead focus on research questions relating to contextual peer effects. A large literature following Hoxby (2000) uses cohort variation in peer composition within schools to estimate how the share of female classmates, the racial mix, or the share of higher-SES classmates in a cohort affects students' outcomes (Hoxby & Weingarth 2005, Jonsson & Mood 2008, Rosenqvist 2018, Holmlund et al. 2023, Meluzzi 2024).[1] Previous studies using

[1] Technically, the cohort variation in peer composition can be seen as an IV for actual peer composition; but

versions of this approach generally show that classes with a higher share of girls perform better on average (Hoxby 2000, Lavy & Schlosser 2011). However, more recent studies using the same methodological approach suggest that this result is primarily driven by gains among girls, while the effects for boys are weaker and sometimes negative (Proud 2014, Oosterbeek & Van Ewijk 2014, Briole 2021, Lénárd 2026).

Plausibly the most common way to deal with the reflection problem is to use lagged leave-out-mean outcomes (Sund 2013, Isphording & Zölitz 2020) or, similarly, student fixed effects based on past performance (Hanushek et al. 2003, Burke & Sass 2013). Of course, this lagged design yields a reduced-form effect of pre-existing peer quality, not the full feedback loop involved in endogenous peer effects. Further, as Lavy et al. (2012) note, when classes are stable, or there is little reshuffling happening across school years, lagged peer achievement basically means moving the reflection problem back one period.

Experimental and policy variation in peer composition can also be used to assess the impact of peers without explicitly estimating structural endogenous peer effects. One such approach is to randomize students to fully segregated or fully integrated peer groups. Park et al. (2013) find that students in gender-segregated schools perform better academically. In contrast, Jackson (2012) finds no robust achievement gains on average, but shows that single-sex schooling can affect course choices, for example by girls in single-sex schools becoming less likely to take science courses. Similarly, Zölitz & Feld (2018) show that peer gender composition shapes educational and labor-market choices: girls who are randomly assigned to more female peers become less likely to choose male-dominated majors and later work fewer hours and earn less, while men become more likely to choose male-dominated majors after having more female peers, although their labor market outcomes are not affected. Their results suggest that rising female university enrollment may have paradoxically contributed to the persistent occupational segregation among university graduates.

Several studies have proposed identification strategies for estimating endogenous peer effects. Brock & Durlauf (2001) suggest exploiting differences between global (everyone interacts with everyone) and local interactions (interactions confined within reference groups). Manski (2000) proposes that instrumental variables affecting only some members of a group can help identify peer effects, and Moffitt (2001) similarly emphasize exploiting policies that affect only a subset of group members. Versions of these have allowed for estimating endogenous peer effects without necessarily separating all the parameters in the Manski framework (Opper 2019).

Overlapping peer networks have turned out to be a practical data structure for exploiting these ideas, and is now perhaps the most common approach to estimate endogenous peer effects, either with complete individual network data (Bramoullé et al. 2009) or with repeated exposure in groups (De Giorgi et al. 2010). When networks only partially overlap, students are connected to different subsets of peers, so both the leave-one-out peer mean and the higher-

this is identical to using school-fixed effects in a linear regression. Hoxby also controls for trends in the cohort variation in peer composition, which can be achieved by adding interaction terms between school fixed effects and time trends.

order averages built from peers' characteristics vary across individuals. These higher-order peer covariate matrices are then used as instruments for endogenous peer effects. This additional variation breaks the perfect collinearity that otherwise exists between the expected value of leave-out peer averages in a group and the expected value of the individuals of a group.

While social and friendship networks are a very realistic basis for social interaction, they also make it more challenging to control for social homophily in friendship formation. It is possible that a Heckman selection correction can address this (Diemer 2022). More importantly, this shifts the focus from peers not necessarily chosen by the individual (a forced social relation) to that of voluntary friendships. This changes the interpretation of the effects and their policy implications.

Another related approach, which is the one adopted in this paper, is to exploit group-size variation to identify endogenous peer effects (Lee 2007, Davezies et al. 2009, Boucher et al. 2014). Whereas the overlapping-network approach exploits variation in network structure, the group-size approach exploits variation in the mechanical relationship between individual covariates and leave-out peer averages induced by differences in group size. Because an individual covariate has a bigger weight in small groups than in large groups, the relationship between individual characteristics and their leave-out means varies with group size, breaking the perfect collinearity between the expected individual characteristic and the expected leave-out mean that exists in the Manski (1993) framework. With group fixed effects, this yields a within-group estimator in which peer effects are identified by how much the dispersion of outcomes among groups is reduced. This also means that the model controls for correlated effects such as teacher quality, which is often more difficult in the overlapping network approach mentioned above. We discuss this model in more detail in the next two sections.

A precondition for gender-specific peer effects to operate through endogenous feedback is that social interactions in school are themselves gender-segregated. A substantial body of evidence supports this. An early study by Shrum et al. (1988) documented strong gender homophily in school friendships, a pattern that has since proven remarkably robust across settings and methods. Using wearable sensor data to measure actual face-to-face contact among high school students, Fournet & Barrat (2014) and Mastrandrea et al. (2015) both find that interactions are far more frequent within gender than across gender, even after accounting for class structure. Network analyses of friendship nominations in European schools confirm the same pattern: gender is consistently the strongest predictor of tie formation, often stronger than ethnicity or socioeconomic background (Campigotto et al. 2022, Kroneberg et al. 2021, Böller & Kruse 2025). McMillan (2022) further shows that gender homophily holds for both strong and weak ties, and Block & Grund (2014) demonstrate that it is robust even in multidimensional network models that simultaneously account for other attributes. This pervasive within-gender clustering in social interaction implies that the relevant peer group for endogenous feedback is more plausibly defined within gender than across the full classroom.

3 Identification

To explain our empirical model, we start with the canonical model for peer effects found in Manski (1993):

$$y_{ri} = \alpha + \beta \sum_{j \in M_{ri}} \frac{y_{rj}}{m-1} + \gamma x_{ri} + \delta \sum_{j \in M_{ri}} \frac{x_{rj}}{m-1} + \epsilon_i. \tag{1}$$

Here, y_{ri} is the outcome of student i in class r, β is the endogenous peer effect parameter multiplying the leave–one–out mean of classmates' outcomes $\frac{1}{m-1}\sum_{j \in M_{ri}} y_{rj}$, with m being class size. Next, γ captures the effect of the individual covariates x_{ri} such as the parental education level, and δ is the contextual peer effect via the leave–one–out mean of x_{rj}. By taking the average effect of the model, and then by solving for y, we can simplify to:

$$\bar{y}_r = \frac{\alpha}{1-\beta} + \frac{\gamma + \delta}{1-\beta} \bar{x}_r \tag{2}$$

This expression illustrates the reflection problem as it shows that the parameters cannot be separated and are therefore not identified (Manski 1993). However, Lee (2007) shows that if group sizes differ, as school classes typically do, one can identify endogenous and contextual peer effects. To show how, we follow the derivation in Boucher et al. (2014). We again write a Manski-style model for peer effects, but with a few tweaks from before:

$$y_{ri} = \alpha_r + \beta \sum_{j \in M_{ri}} \frac{y_{rj}}{m_r - 1} + \gamma x_{ri} + \delta \sum_{j \in M_{ri}} \frac{x_{rj}}{m_r - 1} + \epsilon_{ri}. \tag{3}$$

First, we have added a group fixed effect α_r. However, the key feature is that we have a number of different class sizes, m_r. Because of the group fixed effect we can isolate y to the right hand side by starting with demeaning both sides:

$$y_{ri} - \bar{y}_r = \beta \left(\sum_{j \in M_{ri}} \frac{y_{rj}}{m_r - 1} - \bar{y}_r \right) + \gamma \left(x_{ri} - \bar{x}_r \right) + \delta \left(\sum_{j \in M_{ri}} \frac{x_{rj}}{m_r - 1} - \bar{x}_r \right) + (\epsilon_{ri} - \bar{\epsilon}_r). \tag{4}$$

We isolate the outcome to the left-hand side to obtain:

$$y_{ri} - \bar{y}_r = \frac{\gamma - \frac{\delta}{m_r - 1}}{1 + \frac{\beta}{m_r - 1}} (x_{ri} - \bar{x}_r) + \frac{1}{1 + \frac{\beta}{m_r - 1}} (\epsilon_{ri} - \bar{\epsilon}_r). \tag{5}$$

Because we now have different class sizes m_r, as long as at least three different class sizes exist in the data the three parameters (β, γ, δ) are generically identified. Thus, this approach allows us to separate endogenous and contextual peer effects, as well as include individual controls.

Endogenous peer effects affect outcomes through an interplay with individual characteristics and contextual peer effects, multiplying their influence on outcomes. Individual covariates and contextual peer effects are predetermined, as they consist of variables such as gender, parental education, and immigrant background, fixed prior to peer-group formation. Consequently, it is not necessary to use, for example, lagged peer achievement or individual panel data to compare outcomes before and after the peer group is formed.

As Boucher et al. (2014) note, it might seem surprising that simply endogenising differences in group sizes allows for identification. Figure 1 illustrates how differences in group sizes help distinguish parameters, and also makes clear that full school classes yield group sizes at which the identifying variation becomes too weak for precise estimation — motivating our separation of boys and girls in estimation. Recall that the reflection problem has two parts: the individual outcome is difficult to separate from the peer outcome, and the contextual peer effects are difficult to distinguish from the individual controls. Within the Manski (1993) framework, both must be solved for identification. The next section (section 4) shows how the former is resolved; here we build intuition for the latter.

Each panel of Figure 1 plots the own covariate on the horizontal axis and the peer average on the vertical axis, with each line drawn holding the overall group average constant. This isolates the purely mechanical link between an individual's value and the leave-out peer average: if one student's value increases, the average of the remaining classmates must fall to keep the group average unchanged. This is important because OLS identifies parameters from conditional expectations. Although each student's own value is excluded from their peers' mean, it still enters the peer averages of all other students — so in expectation the leave-out peer mean equals the full group mean.

Panel A shows that with just one class size, a student's own X and the class's leave-out peer average move in parallel. This entails a mechanical relationship between the average level, individual level, and peer average level. Even though there are three different classes with three different average levels, this only changes the intercepts but gives no slope variation to tease apart "own" versus "peer" effects. Next, Panel B instead shows that when class size varies, the peer line tilts differently, breaking that mirrored movement which lets us separate the role of the individual covariate from that of peer composition. The blue line for a class size of three is steeper because the ego has a larger impact on the mean, whereas the dotted brown line for a class size of five is less steep because the ego has a smaller impact on the average. Again, the logic of the linear-in-means implies that per-person influence is stronger in small groups and weaker in large ones so comparing otherwise similar students placed in groups of different sizes exploits that change in slope to identify own vs. peer composition effects.

However, Panel C shows that as we approach larger group sizes, in this case 26 and 27, the slopes are only slightly different from one another. Even when we compare to a drastically larger group size of 100 students, the slope is only slightly different. As we approach infinity, the own student has basically no effect on the difference between the peer average and the

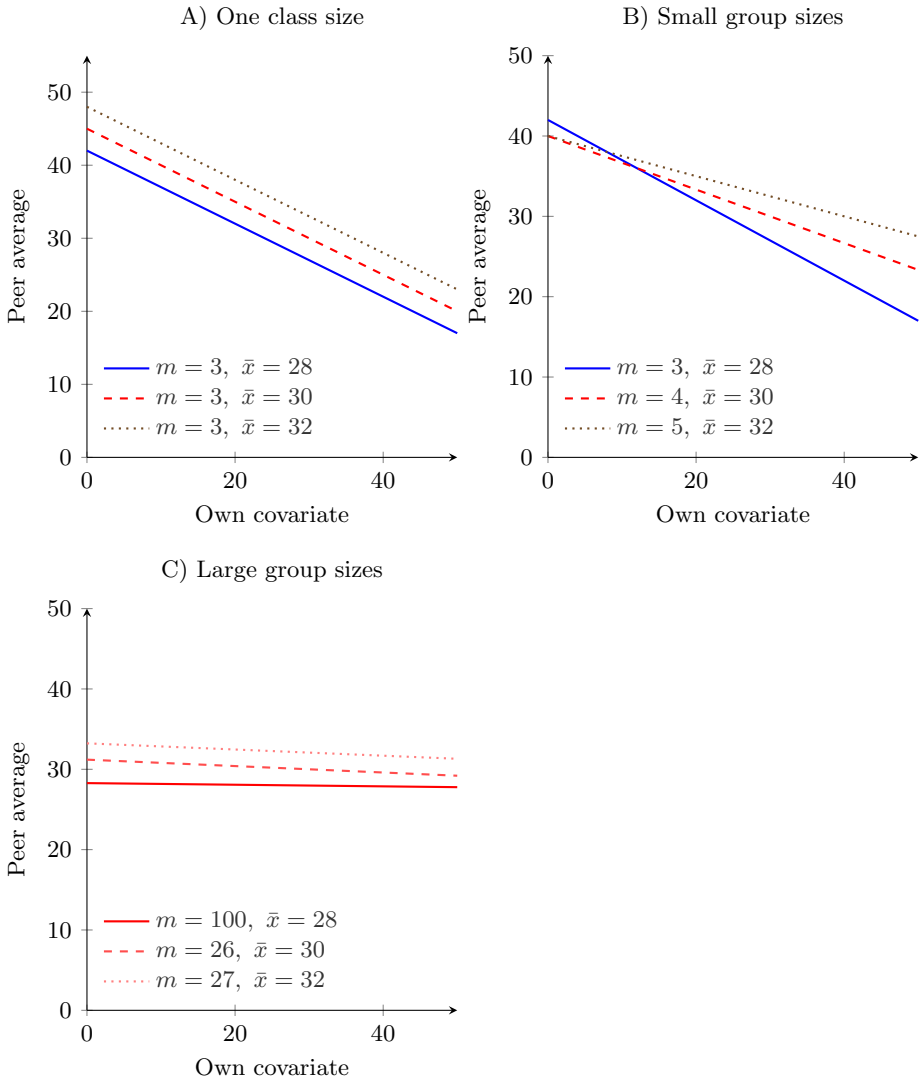
average covariate level.

This also implies that identification becomes weak for class sizes around 20–30, which are typical in real-world school settings. Our solution for empirical implementation is to estimate the model separately for boys and girls, treating same-gender classmates as the relevant peer group. This reduces the group size on which we estimate effects, yielding stronger identification. It also reduces the risk that identification is disproportionately driven by a small number of atypical outlier classes, since the gender split shrinks effective group sizes across the whole distribution of classes rather than relying on the tail of unusually small ones.[2]

As a result, our main estimates deliver gender-specific within-gender peer-effect parameters, while cross-gender endogenous spillovers are not directly identified in this design. The decision to estimate the model separately by gender is not only a statistical choice but also a substantive one. If endogenous peer interactions operate primarily within gender — as the network homophily literature strongly suggests — then pooling boys and girls into a single peer group would conflate two distinct feedback processes and attenuate the estimated peer-effect parameters toward zero.

[2]Boucher et al. (2014) estimate the model on full school-grade cohorts rather than classrooms, and argue via Monte Carlo simulations that sufficient group size dispersion can support identification even when average group sizes are large. In their data, mean group size ranges from 51 (Mathematics) to 146 (Science), with standard deviations of similar magnitude, implying that many small groups contribute to identification. Since identification is strongest for smaller groups, however, it is important that these do not differ systematically from the typical group. In their Monte Carlo simulations, Boucher et al. (2014) implicitly assume that a class of size 3 is no different from a class of size 20, which is unlikely to hold in practice. Splitting the sample by gender reduces effective group sizes to around half the class size, so that identification does not depend on outlier small groups.

Figure 1: **Peer average vs. own covariate across class sizes.**



Notes: The figure illustrates the mechanical relationship between an individual's characteristic x_{ri} (horizontal axis) and the leave-one-out peer mean $\bar{x}_{r,-i}$ (vertical axis) that arises within a group of size m_r with group mean $\bar{x}_r$. Each line corresponds to one group.

The main assumption in the model is exogeneity of outcomes conditional on the group fixed effect:

$$\mathbb{E}(\varepsilon_{ri} \mid X_r,\, m_r,\, \alpha_r) = 0.$$

Peer-group size may, however, correlate with *group-level* unobservables,

$$\mathbb{E}(\alpha_r \mid X_r,\, m_r) \neq 0,$$

but, given α_r, m_r is independent of idiosyncratic shocks ε_{ri}. For example, a teacher who is uniformly generous when grading exams does not pose a problem, but a teacher who grades different parts of the distribution differently for different class sizes may threaten identification. We therefore use mathematics test scores from the national Swedish exams taken in ninth grade as the outcome variable, since these have been shown to be graded more consistently and leave

less scope for biases in teacher grading (Vlachos 2019).

4 Estimation

While we have shown that having just three distinct class sizes is sufficient to break the mechanical link between individual characteristics and their corresponding peer averages, this condition alone does not resolve all challenges associated with the reflection problem. Characteristics such as being foreign-born, gender, and parental education are typically exogenously given from the student's point of view before the class is formed. In those cases, breaking the mechanical link through group-size variation is enough. These individual characteristics do not need to be modeled as a function of the peers' characteristics, since they are determined outside the peer group. In contrast, the outcome variable (math achievement) is endogenously affected through interaction. Therefore, individual outcome y_i appears on both sides of the linear-in-means equation, and the model thus constitutes a simultaneous system of equations (one for each individual). A standard OLS regression cannot consistently estimate such a system. However, Lee (2007) and Bramoullé et al. (2009) show that the square of the contextual peer matrix can be used as an instrument that allows for the estimation of peer effects. In linear-in-means models using network data, such a squared matrix captures the expected influence of peers' peers. Yet, in a class-level setting, where every student is a peer to every other student, this squared contextual peer matrix captures the expected back-reflection of the individual onto themselves, which means that using this as an instrument captures the endogenous peer effect (Bramoullé et al. 2009). Variation in class size implies that this back-reflection is stronger in smaller groups and weaker in larger ones, thus solving the reflection problem. As both Lee (2007) and Boucher et al. (2014) show that conditional maximum likelihood (CML) is more efficient than IV in this setting, we use CML as our estimation method. This choice is also motivated by the nature of our outcome variable: we observe test grades rather than raw test scores since the latter are not available in the data. The outcome should therefore be viewed as a discrete ordered variable rather than a fully continuous measure, implying less variation than would be available with raw scores. Thus, for our context, CML is well suited since it makes more efficient use of the variation available in the data.

The parameters are obtained by maximizing the conditional log-likelihood of the transformed outcomes, conditional on the group means and observed group sizes. As in Boucher et al. (2014), the likelihood is derived under a working assumption that the idiosyncratic error term is independently and identically normally distributed. Inference, however, is based on robust sandwich standard errors, so consistency requires correct specification of the conditional mean rather than full normality.

5 Data and measurement

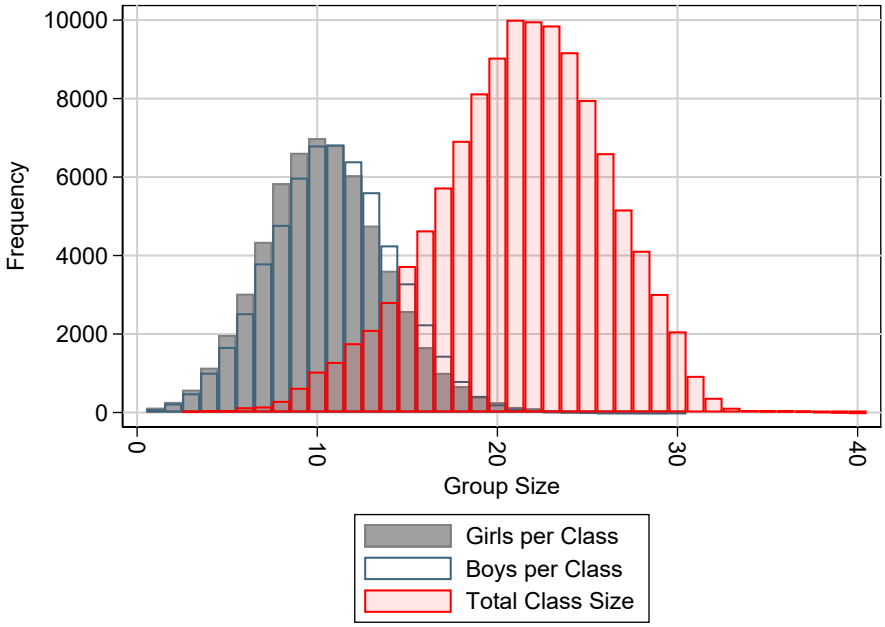
Our data consists of Swedish register data that can be linked using the Swedish social security number. We use a school-class register to obtain data on school classes from 2013 through 2019. We focus on ninth grade, the final year of compulsory school, when students are typically 15 years old. The register contains student and school IDs, and the name or label of the school class. A great majority follow the convention of using grade paired with a letter (e.g., 9A, 9B, etc), but there are also examples of using ad hoc names (animals or physical phenomena).[3]

We use a register on national test grades to construct our outcome variable. Specifically, we use the mathematics test grade. Students are linked to parents using the multigenerational register based on birth records (or immigration documents), which in turn can be linked to records on immigration status, neighborhoods, education, and occupation, among other things. For covariates used as individual controls and contextual peer variables, we first construct parental variables both at the individual level and as class-peer averages. Our socioeconomic measure is parents' years of education, computed as the average of the mother's and father's years of schooling which are then averaged over classmates to form the peer measure. Another covariate is foreign background, defined as either the student being born abroad or having two foreign-born parents. Finally, we include student age in months at the end of the year.

Figure 2 shows the distribution of girls per class, boys per class and total class sizes. The figure illustrates that estimating peer effects separately for girls and boys results in much lower group sizes, yielding stronger identification. This makes the data well suited for estimating gender-specific peer effects. The main analysis splits the sample by gender and constructs peer measures using only same-gender classmates within each class, so that identification comes from variation in the number of boys (or girls) per class.

[3]The quality of this register is overall good. There are some instances of missing information or invalid information. However, when we examine how classes connect across years (by examining pupil flows across years), it is clear that a great majority (but not all) of classes are stable, contain the same set of students across grades, and are named in a logical fashion, i.e., class 7A in 7th grade becomes class 8A in the next year and 9A in the subsequent year. We follow the precautions of Lénárd (2026), e.g., excluding unrealistic class sizes and using pupil flows between class identities to separate out flawed class information. We used the class information for the last year of a grade if a grade was repeated. We limit classes to 10 to 40 pupils. Most very small classes (10 students or fewer) are introductory classes for immigrants.

Figure 2: **Distribution of class sizes and same-gender group sizes.**



Notes: The figure plots the frequency distribution of total class size and of the number of boys and girls per class in grade-9 classes in Sweden, 2013–2019.

Table 1 shows descriptive statistics for the outcome variable and the covariates used as individual controls and contextual peer characteristics. The outcome variable, Math test grade, is standardized to have mean zero and standard deviation one in the full population, which as expected, is also reflected in the table. Girls have an average math grade 0.02 standard deviations (SD) above the population mean, while boys perform on average 0.02 SD below, implying a difference of 0.04 SD between girls and boys. However, as we later document in Figure 4, the girl–boy gap widens to roughly 0.06 SD in girl-dominated classes but is essentially zero in boy-dominated classes, underscoring that the gap is substantial in some settings and varies with classroom gender composition. Sweden's small female advantage in mathematics is unusual internationally: although girls outperform boys in reading across all OECD countries, boys outperform girls in mathematics in most countries (OECD 2019).

The covariates are very similar across gender, as expected given that they largely capture predetermined background characteristics. Mean parental education is around 12.6 years for both girls and boys. Students' age is on average 186 months, or 15.5 years. The mean share of students with a foreign background, defined as being born abroad or having two foreign-born parents, is 0.17 in the full sample.

Because identification relies on variation within gender-specific class cells, the table also reports the average within school class by gender SD. The within-cell SD of math grades is approximately 0.88, indicating substantial dispersion in achievement within gender groups in the same class. Similarly, there is meaningful within-class dispersion in parental education (around 1.80 years), age (around 3.9 months), and foreign background (around 0.25 SD).

These figures confirm that there is non-trivial within-class, within-gender variation in both the outcome and covariate variables, which is needed for identifying peer effects in the empirical analysis. Further, the similar within-class dispersion of observable characteristics among boys and girls suggests limited gender-specific sorting across classrooms along these dimensions.

Table 1: **Descriptive statistics by gender.**

	All		Girls		Boys	
	Mean	SD	Mean	SD	Mean	SD
Math score	0.00	1.00	0.02	1.00	−0.02	1.00
Parental education (years)	12.55	2.06	12.55	2.07	12.55	2.06
Age (months)	185.93	4.11	185.78	4.07	186.08	4.15
Foreign background (dummy)	0.17	0.38	0.17	0.38	0.17	0.37
Within class×gender SD: Math score		0.88		0.88		0.88
Within class×gender SD: Parental education		1.80		1.80		1.80
Within class×gender SD: Age		3.86		3.82		3.90
Within class×gender SD: Foreign background		0.25		0.25		0.25
N (students)	502,556		244,268		258,288	

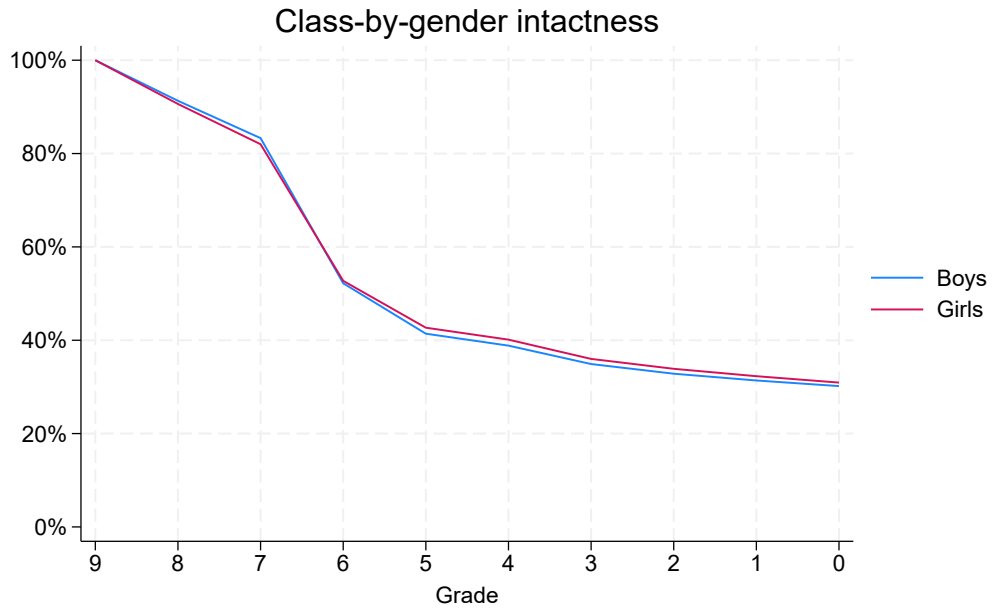
Notes: The table reports sample means and standard deviations for the outcome and covariates used in the empirical analysis, separately for girls and boys and for the full estimation sample of grade-9 students in Swedish compulsory school, 2013–2019. Math test grade is the standardized grade on the national mathematics test in grade 9, normalized to mean zero and standard deviation one in the underlying full population.

Peer exposure in Sweden is primarily classroom-based: there is no ability tracking, and most subjects are taught within the regular class, apart from limited subject-specific grouping such as modern languages. The study period also spans the 2015 refugee crisis, when about 70,000 children sought asylum in Sweden (Statistics Sweden 2023). Because newly arrived students were sometimes assigned to regular classes while receiving core instruction separately, registered class membership may not always capture actual peer exposure.

Figure 3 therefore tracks class intactness: the share of the grade-9 same-gender peer group already present in earlier grades. More than 80% remain together from grade 7, compared with about 50% from grade 6 and 30% from preschool class, suggesting that peer exposure typically spans at least three years. In classes with substantial recent turnover, however, the observed grade-9 class may be a poor measure of the relevant peer environment. We therefore exclude high-turnover classes in a robustness check, which also addresses concerns about classes affected by large inflows of newly arrived students.

The figure also highlights the difficulty of using lagged outcomes as explanatory variables. Since past outcomes are themselves shaped by at least some of the same set of peers, using lagged peer outcomes effectively shifts the reflection problem backward in time rather than resolving it (Lavy et al. 2012). Our estimates should therefore be interpreted as capturing cumulative peer exposure, typically since at least grade 7 and sometimes earlier.

Figure 3: **Class-by-gender intactness share.**



Notes: The intactness share is defined as the fraction of the grade-9 same-gender peer group that was enrolled in the same class as the focal student in each grade, plotted separately for boys and girls. Values are averaged across all students in the estimation sample. Grade 0 is the year of preschool class.

6 Results

6.1 Peer Effect Estimates

The bottom of Table 2 shows convincing statistically significant peer effect estimates, as well as pronounced differences by gender. For boys, the estimated endogenous peer effect is 0.195, implying that boys' outcomes respond positively to changes in peers' outcomes. For girls, the corresponding estimate is substantially larger at 0.481. The difference, 0.286, is statistically significant at the 0.1 percent level. The corresponding social multipliers, $1/(1-\hat{\beta})$, are thus approximately 1.93 for girls and 1.25 for boys. This implies that an intervention lifting girls' initial achievement by 0.1 SD produces an equilibrium gain of about 0.19 SD, whereas an equivalent intervention among boys produces only about 0.125 SD.

These endogenous effects are still smaller than the 0.82 that Boucher et al. (2014) report for mathematics using the same method on full school-cohorts, but are estimated with higher precision. We believe that our focus on classroom-level rather than school-cohort-level peer groups, combined with separating the analysis by gender, likely explains this improvement. Our results are more comparable to Diemer (2022), who reports endogenous peer effects between 0.20 and 0.25 using Swedish friendship data, also in ninth grade, though without breaking up the analysis by gender.

Turning to contextual peer effects, we first consider parental education. Own parents' education is strongly positively associated with achievement for both boys and girls, with

nearly identical coefficients (0.182 for boys and 0.183 for girls). The peer counterpart (peers' parents' education) is also positive and precisely estimated for both groups, but larger for boys (0.210) than for girls (0.142). This suggests that boys' performance is more sensitive to the socioeconomic composition of the peer group along this dimension. At first sight, it may seem that peers' parents' education matters more than own parents' education. However, the peer coefficients correspond to increasing the average parental education of all peers by one year, which represents a substantial shift in the socioeconomic profile of the class.

Foreign background is negatively associated with achievement for both genders. Boys with a foreign background score lower on average (-0.198), and the effect is larger for girls (-0.262). The peer analogue is likewise negative: a higher share of peers with foreign background is associated with lower outcomes for both boys and girls, with a notably stronger estimate for girls (-0.616) than for boys (-0.328). These effects may seem very large, but it is important to remember that the point estimates capture the effect of moving from 0% to 100% foreign background. The larger negative coefficient on own and peers' foreign background for girls compared to boys should not be read as implying, for example, that foreign-background girls perform worse than foreign-background boys in levels. Rather, they indicate that these characteristics are more strongly negatively associated with achievement among girls than among boys. Since girls have both a higher average mathematics performance and substantially stronger endogenous peer effects, foreign-background girls may still outperform foreign-background boys in equilibrium. As Horrace et al. (2025) note, when peer interactions operate mainly within salient groups such as gender or ethnicity, classroom spillovers can widen gaps between groups while compressing differences within groups.

Age has only a very small effect, at close to zero for boys (0.001) and is slightly negative for girls (-0.002), with the latter being marginally significant. Peer age is positive for boys (0.038) and statistically significant, while the corresponding estimate for girls is small and not statistically distinguishable from zero (0.008).

Table 2: **Conditional maximum likelihood peer effect estimates, by gender**

	Boys	Girls
Parents' education (yrs)	0.182***	0.183***
	(0.003)	(0.003)
Peers' parents' education	0.210***	0.142***
	(0.029)	(0.028)
Foreign background	-0.198***	-0.262***
	(0.017)	(0.017)
Peers' foreign background	-0.328**	-0.616***
	(0.158)	(0.152)
Age (months)	0.001	-0.002*
	(0.001)	(0.001)
Peers' age (months)	0.038***	0.008
	(0.012)	(0.012)
Endogenous peer effect (β)	0.195***	0.481***
	(0.046)	(0.047)
Observations	257,327	243,224
Log-likelihood	-79,887.4	-75,465.5

Notes: The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, $^{***}\, p < 0.01$.

The reported endogenous peer-effect coefficients in Table 2 reflect within-gender peer dynamics. The specification therefore abstracts from direct cross-gender influences. Thus, omitting cross-gender peer effects may bias the within-gender estimates. At the same time, these cross-gender interactions are themselves of interest for a full understanding of how peer effects operate within classes. In Table 3, we therefore conduct a form of placebo test in which gender labels are randomly reassigned within classes. The table reports the results from 100 repeated randomizations, with entries corresponding to the mean coefficient estimates across these repetitions.

Before turning to the estimates, we clarify the pattern expected from randomizing gender labels. Since the randomized "boys" and "girls" groups are now random mixtures of actual boys and girls, the coefficients on *individual* controls should mechanically converge toward a common value (or stay the same if they were already similar in Table 2). For *peer* effects the interpretation is sharper: if the within-gender coefficients converge toward a common value, this indicates that cross-gender peer effects are of similar magnitude to within-gender ones; if instead both coefficients attenuate toward zero, this indicates that cross-gender peer effects are weak or absent, so that mixing the peer group dilutes the estimated response.

Under this logic, the endogenous peer-effect estimates fall sharply for both groups: for boys from 0.195 to 0.09, and for girls from 0.481 to 0.281. The fact that the endogenous point estimates are approximately halved is consistent with the view that the estimated peer effects in the main specification are driven almost entirely by within-gender interactions, implying that cross-gender endogenous peer effects are close to zero. This is consistent with what Horrace et al. (2025) find for elementary schools in the US.

The contextual coefficients show both patterns. Peers' foreign background converges cleanly,

from −0.328 and −0.616 to −0.438 and −0.422, suggesting that cross-gender and within-gender contextual effects for foreign background are of similar magnitude. Peers' parental education, by contrast, attenuates: both coefficients drop below their common midpoint (from 0.210 and 0.142 to 0.142 and 0.109), consistent with cross-gender contextual effects of parental education being weaker than within-gender ones.

Table 3: **Peer-effect estimates with within-class randomized gender labels**

	Boys	Girls
Parents' education (yrs)	0.177***	0.178***
	(0.002)	(0.002)
Peers' parents' education	0.142***	0.109***
	(0.022)	(0.022)
Foreign background	-0.220***	-0.226***
	(0.011)	(0.012)
Peers' foreign background	-0.438***	-0.422***
	(0.120)	(0.108)
Age (months)	0.001	-0.002*
	(0.001)	(0.001)
Peers' age (months)	0.035***	0.008
	(0.008)	(0.008)
Endogenous peer effect (β)	0.090**	0.281***
	(0.031)	(0.030)
Observations	257,234	243,285
Log-likelihood	-81,670.9	-77,511.5

Notes: The outcome is the standardized mathematics grade on the national test. Total class composition is held fixed and only the gender variable is randomized. The reported entries are the mean of coefficient estimates across 100 independent randomizations. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

As noted, some classes may have been formed relatively late, shortly before grade 9. In such cases, a substantial share of students' peer exposure may have taken place in earlier class constellations rather than in the class observed at the end of compulsory schooling. As a result, estimates based on the grade-9 class may not fully capture the relevant peer environment for these students.

To address this concern, the analysis in Table 4 restricts the sample to classes that were at least 80% intact since grade 7. This means that at least 80% of the same-gender students in grade 9 were already part of the same peer group in grade 7. Imposing this restriction reduces the sample by 61,786 observations for boys and 61,654 observations for girls. The results in Table 4, however, remain largely unchanged to those of Table 2. For boys, the estimated endogenous peer effect changes marginally from 0.195 to 0.196, and for girls from 0.481 to 0.471. Among the contextual peer effects, the most notable difference is that the negative effect of peers with a foreign background becomes somewhat smaller in magnitude. One possible explanation is that classes with higher turnover during this period were more likely to include recently arrived immigrant students, particularly in connection with the 2015 refugee inflow. These students may have altered the learning environment to a greater extent

than a typical student with an immigrant background.

Table 4: **Peer-effect estimates restricted to classes with at least 80% grade-7 intactness**

	Boys	Girls
Parents' education (yrs)	0.186***	0.186***
	(0.004)	(0.004)
Peers' parents' education	0.215***	0.140***
	(0.035)	(0.033)
Foreign background	-0.176***	-0.261***
	(0.020)	(0.021)
Peers' foreign background	-0.035	-0.494***
	(0.190)	(0.188)
Age (months)	0.001	-0.002
	(0.002)	(0.002)
Peers' age (months)	0.028*	0.002
	(0.015)	(0.015)
Endogenous peer effect (β)	0.196***	0.471***
	(0.056)	(0.056)
Observations	195,517	181,542
Log-likelihood	-61,204.3	-56,639.6

Notes: The outcome is the standardized mathematics grade on the national test. The estimation is restricted to classes in which at least 80% of the same-gender students in grade 9 were already enrolled in the same class in grade 7. This restriction reduces the sample by 61,786 observations for boys and 61,654 observations for girls. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

6.2 Heterogeneity analysis

Because the model relies on class-size variation for identification, we cannot simply split the sample within classes to study heterogeneity. However, we can compare estimates across different types of classes in order to explore heterogeneous effects at the class level.

Splitting the sample by terciles of class-mean parental education reveals clear heterogeneity in endogenous peer effects shown in Table 5 for boys and in Table 6 for girls. In both tables, the first column repeats the pooled gendered estimates from Table 2. For boys, the endogenous peer effect is statistically significant only in the lowest-education tercile, where $\beta = 0.250$ (SE 0.082), while it is essentially zero and insignificant in the second and third terciles ($\beta = 0.005$ (SE 0.078) and $\beta = 0.009$ (SE 0.071), respectively). For girls, in contrast, the endogenous peer effect remains positive across all terciles, but is weaker in higher-education classrooms: it is estimated at $\beta = 0.639$ (SE 0.090) in the first tercile and $\beta = 0.235$ (SE 0.076) in the second tercile, before declining further in the third tercile. Thus, for both boys and girls, endogenous peer effects tend to be weaker in classrooms with higher average parental education.

Table 5: **Boys: peer-effect estimates by tercile of class-mean parental education**

	Pooled	Mean 1st tercile	Mean 2nd tercile	Mean 3rd tercile
Parents' education (yrs)	0.182***	0.166***	0.171***	0.184***
	(0.003)	(0.005)	(0.005)	(0.005)
Peers' parents' education	0.210***	0.265***	0.049	0.101*
	(0.029)	(0.047)	(0.054)	(0.052)
Foreign background	-0.198***	-0.099***	-0.151***	-0.271***
	(0.017)	(0.026)	(0.029)	(0.031)
Peers' foreign background	-0.328**	-0.233	0.302	0.169
	(0.158)	(0.232)	(0.287)	(0.307)
Age (months)	0.001	-0.001	0.002	0.001
	(0.001)	(0.002)	(0.002)	(0.003)
Peers' age (months)	0.038***	0.027	0.045*	0.040
	(0.012)	(0.017)	(0.023)	(0.025)
Endogenous peer effect (β)	0.195***	0.250***	0.005	0.009
	(0.046)	(0.082)	(0.078)	(0.071)
Observations	257,327	86,798	87,294	83,235
Log-likelihood	-79,887.4	-23,387.5	-27,425.4	-28,543.6

Notes: The sample is split by terciles of within-class average parental years of education. Column 1 reproduces the pooled estimate for boys. Columns 2–4 report estimates for classes falling in the first (lowest), second, and third (highest) tercile of class-mean parental education. The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: **Girls: peer-effect estimates by tercile of class-mean parental education**

	Pooled	Mean 1st tercile	Mean 2nd tercile	Mean 3rd tercile
Parents' education (yrs)	0.183***	0.159***	0.184***	0.179***
	(0.003)	(0.006)	(0.005)	(0.005)
Peers' parents' education	0.142***	0.126***	0.112**	0.027
	(0.028)	(0.047)	(0.050)	(0.047)
Foreign background	-0.262***	-0.134***	-0.218***	-0.325***
	(0.017)	(0.028)	(0.030)	(0.030)
Peers' foreign background	-0.616***	-0.302	0.019	-0.295
	(0.152)	(0.233)	(0.274)	(0.294)
Age (months)	-0.002*	-0.004*	-0.003	-0.003
	(0.001)	(0.002)	(0.002)	(0.002)
Peers' age (months)	0.008	0.010	0.002	-0.008
	(0.012)	(0.017)	(0.022)	(0.024)
Endogenous peer effect (β)	0.481***	0.639***	0.235***	0.146**
	(0.047)	(0.090)	(0.076)	(0.071)
Observations	243,224	76,295	85,157	81,772
Log-likelihood	-75,465.5	-20,301.7	-26,493.1	-28,212.0

Notes: The sample is split by terciles of within-class average parental years of education. Column 1 reproduces the pooled estimate for girls. Columns 2–4 report estimates for classes falling in the first (lowest), second, and third (highest) tercile of class-mean parental education. The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Tables 7 and 8 report heterogeneity in endogenous peer effects across terciles of within-class *dispersion* in parental education (in standard deviations) for boys and girls, respectively. For boys, the endogenous peer effect is statistically significant in the lowest- and highest-dispersion terciles, with $\beta = 0.185$ in the first tercile and $\beta = 0.235$ in the third tercile, while the

estimate in the middle tercile is smaller and not statistically distinguishable from zero. For girls, the endogenous peer effect is large and statistically significant in all terciles, but varies with dispersion: it is highest in the lowest-dispersion tercile, $\beta = 0.529$, declines in the middle tercile to $\beta = 0.374$, and then increases again in the highest-dispersion tercile to $\beta = 0.442$. Thus, for both boys and girls endogenous peer effects are strongest in low and high dispersion classes.

Table 7: **Boys: peer-effect estimates by tercile of within-class standard deviation of parental education**

	Pooled	SD 1st tercile	SD 2nd tercile	SD 3rd tercile
Parents' education (yrs)	0.182***	0.174***	0.177***	0.183***
	(0.003)	(0.007)	(0.005)	(0.004)
Peers' parents' education	0.210***	0.085	0.097*	0.276***
	(0.029)	(0.066)	(0.057)	(0.040)
Foreign background	-0.198***	-0.220***	-0.211***	-0.167***
	(0.017)	(0.032)	(0.030)	(0.027)
Peers' foreign background	-0.328**	-0.495*	-0.372	-0.063
	(0.158)	(0.280)	(0.299)	(0.251)
Age (months)	0.001	0.005**	0.001	-0.004
	(0.001)	(0.002)	(0.002)	(0.002)
Peers' age (months)	0.038***	0.077***	0.035	0.001
	(0.012)	(0.020)	(0.023)	(0.020)
Endogenous peer effect (β)	0.195***	0.185**	0.112	0.235***
	(0.046)	(0.076)	(0.082)	(0.080)
Observations	257,327	87,451	90,721	79,155
Log-likelihood	-79,887.4	-25,789.8	-28,818.4	-25,216.8

Notes: The sample is split by terciles of the within-class standard deviation of parental years of education. Column 1 reproduces the pooled estimate for boys. Columns 2–4 report estimates for classes falling in the first (lowest), second, and third (highest) tercile of within-class SD of parental education. The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

Table 8: **Girls: peer-effect estimates by tercile of within-class standard deviation of parental education**

	Pooled	SD 1st tercile	SD 2nd tercile	SD 3rd tercile
Parents' education (yrs)	0.183***	0.187***	0.191***	0.175***
	(0.003)	(0.007)	(0.005)	(0.004)
Peers' parents' education	0.142***	0.077	0.190***	0.130***
	(0.028)	(0.066)	(0.053)	(0.037)
Foreign background	-0.262***	-0.268***	-0.218***	-0.296***
	(0.017)	(0.034)	(0.030)	(0.026)
Peers' foreign background	-0.616***	-0.640**	-0.157	-0.945***
	(0.152)	(0.283)	(0.286)	(0.234)
Age (months)	-0.002*	-0.000	-0.003	-0.003
	(0.001)	(0.002)	(0.002)	(0.002)
Peers' age (months)	0.008	0.019	-0.005	0.007
	(0.012)	(0.020)	(0.023)	(0.019)
Endogenous peer effect (β)	0.481***	0.529***	0.374***	0.442***
	(0.047)	(0.079)	(0.084)	(0.081)
Observations	243,224	79,409	86,091	77,724
Log-likelihood	-75,465.5	-23,294.5	-27,485.0	-24,640.0

Notes: The sample is split by terciles of the within-class standard deviation of parental years of education. Column 1 reproduces the pooled estimate for girls. Columns 2–4 report estimates for classes falling in the first (lowest), second, and third (highest) tercile of within-class SD of parental education. The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.3 Gender Shares Analysis

So far, we have estimated endogenous peer effects separately for boys and girls and only indirectly shown how these estimates can explain why a higher share of girls is associated with higher achievement for the class as a whole, as well as for boys and girls separately. Endogenous peer effects may themselves vary with the gender composition of the class. We therefore estimate the model separately for classes in which each gender is in the minority versus the majority.

Because we exploit group size variation to estimate peer effects, we first have to ascertain sufficient variation for different gender shares to be studied. Classes in which boys or girls make up more than 50% may mechanically involve relatively large same-gender groups. If these groups become too large, the variation needed to estimate peer effects effectively may weaken. Appendix Figures A1 and A2 plot the distribution of same-gender group sizes separately for minority and majority classes, and show that group sizes remain mostly concentrated in the range of roughly 8 to 18 students even when a gender is in the majority. This confirms that groups are still relatively small and with sufficient variation to estimate peer effects separately for both majority and minority gender groups.

Tables 9 and 10 show that endogenous peer effects vary systematically with class gender composition, especially for boys. For boys, the endogenous peer effect is much larger in boy-majority classes: when boys constitute more than 50% of the class, $\beta = 0.480$, whereas in boy-minority classes it is substantially smaller at $\beta = 0.115$ and only marginally significant.

For girls, endogenous peer effects are large in both composition bins but are strongest in girl-majority classes: $\beta = 0.543$ when girls are in majority and $\beta = 0.443$ when girls are in minority. Thus, girls exhibit similarly large endogenous peer effects regardless of whether they are in minority or in majority, whereas for boys, these dynamics appear much stronger when boys are in majority.

Table 9: **Boys: peer-effect estimates by class gender composition**

	Pooled	Boys share $> 50\%$	Boys share $< 50\%$
Parents' education (yrs)	0.182***	0.194***	0.192***
	(0.003)	(0.005)	(0.006)
Peers' parents' education	0.210***	0.341***	0.253***
	(0.029)	(0.063)	(0.042)
Foreign background	-0.198***	-0.275***	-0.192***
	(0.017)	(0.029)	(0.031)
Peers' foreign background	-0.328**	-1.300***	-0.189
	(0.158)	(0.333)	(0.230)
Age (months)	0.001	0.004**	0.000
	(0.001)	(0.002)	(0.002)
Peers' age (months)	0.038***	0.083***	0.026
	(0.012)	(0.025)	(0.018)
Endogenous peer effect (β)	0.195***	0.480***	0.115*
	(0.046)	(0.103)	(0.066)
Observations	257,327	157,514	80,616
Log-likelihood	-79,887.4	-49,664.7	-24,244.7

Notes: The sample is split by whether boys constitute less than or more than 50% of the class. Column 1 reproduces the pooled estimate for boys. Column 2 restricts to boy-majority classes (share of boys above 50%) and column 3 restricts to boy-minority classes (share of boys below 50%). The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10: **Girls: peer-effect estimates by class gender composition**

	Pooled	Girls share $< 50\%$	Girls share $> 50\%$
Parents' education (yrs)	0.183***	0.189***	0.186***
	(0.003)	(0.005)	(0.005)
Peers' parents' education	0.142***	0.171***	0.193***
	(0.028)	(0.039)	(0.064)
Foreign background	-0.262***	-0.276***	-0.312***
	(0.017)	(0.029)	(0.032)
Peers' foreign background	-0.616***	-0.604***	-1.354***
	(0.152)	(0.210)	(0.372)
Age (months)	-0.002*	0.000	0.000
	(0.001)	(0.002)	(0.002)
Peers' age (months)	0.008	0.019	0.049*
	(0.012)	(0.017)	(0.027)
Endogenous peer effect (β)	0.481***	0.443***	0.543***
	(0.047)	(0.066)	(0.113)
Observations	243,224	103,360	120,709
Log-likelihood	-75,465.5	-30,395.1	-39,357.3

Notes: The sample is split by whether girls constitute less than or more than 50% of the class. Column 1 reproduces the pooled estimate for girls. Column 2 restricts to girl-minority classes (share of girls below 50%) and column 3 restricts to girl-majority classes (share of girls above 50%). The outcome is the standardized mathematics grade on the national test. All specifications include class-level fixed effects that absorb group-level unobservables. Robust sandwich standard errors are reported in parentheses: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

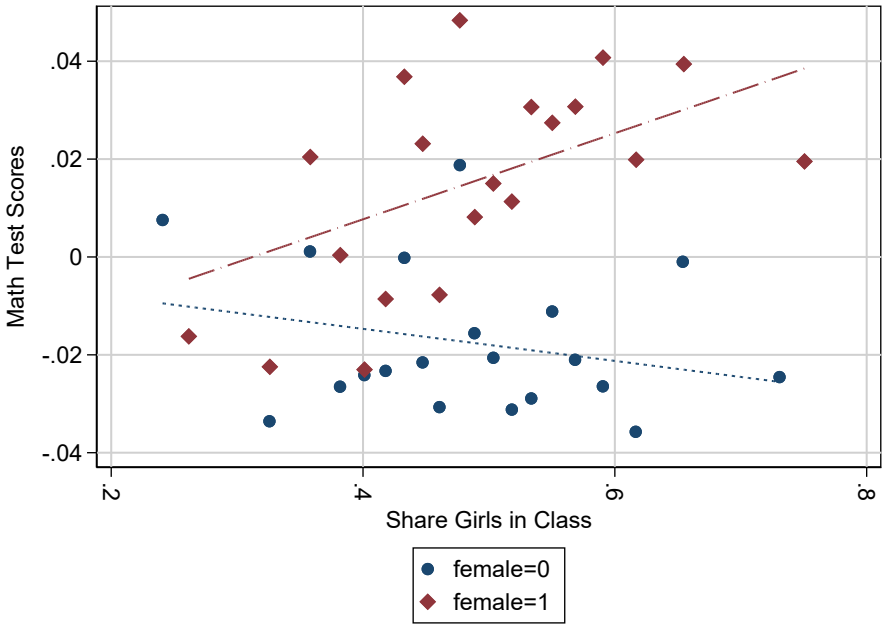
7 Analytical Model

Our empirical results point to three main gender-related regularities. First, endogenous peer effects in mathematics are stronger within gender for girls than for boys. Second, these effects become much weaker when gender labels are randomized within classes, suggesting that the relevant endogenous interactions operate primarily within gender rather than across gender. Third, while girls' endogenous peer effects are strong both when they are in minority and in majority, boys exhibit almost no endogenous peer effects when they are in minority, but strong effects when they are in majority. In this section, we develop a simple analytical framework, based on a Manski-style linear-in-means model, to illustrate how these findings map into class-level outcomes and how they relate to the reduced-form literature on classroom gender composition.

Previous literature has shown that a higher share of girls tends to improve average classroom achievement. However, this pattern appears to be driven primarily by gains among girls, while the effects for boys are more ambiguous. In the Swedish case, Lénárd (2026) even finds small negative effects for boys when the share of girls is higher. We can plot similar patterns in our data using a simple regression of mathematics test scores on the share of girls in the class, controlling for immigrant background, parental education, and class size. Figure 4 shows the results of this regression. A higher share of girls is associated with higher mathematics achievement for girls, but not for boys. Specifically, a 10 percentage-point increase in the share of girls is associated with an increase of about 0.0086 standard deviations for girls, whereas

the corresponding estimate for boys is about -0.003 standard deviations, but this negative association is not statistically significant. Thus, in our data the positive relationship is clear for girls, but much less so for boys. This difference is important for interpreting the analytical model below.

Figure 4: **Mathematics achievement by class share of girls.**



Notes: The figure plots a regression of standardized grade-9 mathematics test scores on the share of girls in the class, estimated separately for boys and girls in Swedish grade-9 classes. Controls include parental education, the share of students with foreign background, and class size.

Our starting point of the analytical model is a standard linear-in-means model in which outcomes depend on peers' outcomes (endogenous peer effects) and peers' predetermined characteristics (contextual peer effects). We extend this framework by allowing both endogenous and contextual peer effects to differ depending on whether they arise from same-gender peers or opposite-gender peers. This distinction is useful in our setting because the empirical model is estimated separately by gender and therefore identifies within-gender peer effects directly, whereas cross-gender effects are not directly estimated and must instead be calibrated in the analytical model.

For the analytical simulations, we omit individual controls because they are not varied across counterfactuals and are assumed to be equal across genders. When held fixed, such controls are thus absorbed by the intercepts. The intercepts should therefore be interpreted as reduced-form baseline terms that capture non-peer determinants of achievement that are held constant across the simulations. Consider a two-group (boys B and girls G) model in school class c, with separate equations for boys and girls, allowing both endogenous and contextual peer effects to differ by whether they arise from same-gender peers or opposite-gender peers.

For boys in class c:

$$y_{iBc} = \alpha_B + \beta_{BB}\,\bar{y}_{Bc,-i} + \beta_{BG}\,\bar{y}_{Gc} + \gamma_{BB}\,\bar{x}_{Bc,-i} + \gamma_{BG}\,\bar{x}_{Gc} + \varepsilon_{iBc}, \tag{6}$$

and for girls in class c,

$$y_{iGc} = \alpha_G + \beta_{GG}\,\bar{y}_{Gc,-i} + \beta_{GB}\,\bar{y}_{Bc} + \gamma_{GG}\,\bar{x}_{Gc,-i} + \gamma_{GB}\,\bar{x}_{Bc} + \varepsilon_{iGc}. \tag{7}$$

Here, y_{iBc} denotes the outcome for boy i in class c, and y_{iGc} the outcome for girl i in class c. The variables $\bar{y}_{Bc,-i}$ and $\bar{y}_{Gc,-i}$ denote leave-one-out mean outcomes for peers of the same gender as the focal student, for boys and girls respectively. By contrast, $\bar{y}_{Gc}$ in the boys' equation and $\bar{y}_{Bc}$ in the girls' equation denote mean outcomes for peers of the opposite gender in class c. Since the focal student does not belong to that group, no leave-one-out adjustment is required.

Similarly, $\bar{x}_{Bc,-i}$ and $\bar{x}_{Gc,-i}$ denote leave-one-out mean predetermined characteristics for peers of the same gender as the focal student, while $\bar{x}_{Gc}$ in the boys' equation and $\bar{x}_{Bc}$ in the girls' equation denote mean predetermined characteristics for peers of the opposite gender.

It should be noted that these equations do not yet make explicit how exposure is weighted when the number or shares of boys and girls differ within the class. Taking gender averages in class c, assuming mean-zero errors, and solving the resulting 2×2 system of class means yields closed-form expressions for the equilibrium gender means $\bar{y}_{Bc}$ and $\bar{y}_{Gc}$. We relegate this algebra, together with the exposure-weighted version used below, to Appendix 9.1. It is sufficient to note that each endogenous and contextual peer-effect parameter is weighted by the relevant group share, that is, by the share of girls (s_c) or the share of boys ($1 - s_c$) in the class. One implication of the solution is worth noting here: the denominator of $\bar{y}_{Bc}$ and $\bar{y}_{Gc}$ is identical for both groups, so the equilibrium social multiplier is common across boys and girls.

We examine gender-specific equilibrium means as functions of the girls' classroom share in three steps. Throughout, we set the cross-gender endogenous peer effects to zero, $\beta_{BG} = \beta_{GB} = 0$, consistent with the results that endogenous coefficients fall sharply when gender labels are randomized. Following Martinot et al. (2025), who show that boys and girls perform similarly before school entry, we normalize the baseline intercepts so that $\alpha_B = \alpha_G$. This means that the analytical exercise focuses on how gender-specific endogenous and contextual peer effects can amplify or dampen outcomes within classrooms starting from the same baseline level. Of course, curricula and teaching styles may also cater differently to boys and girls and contribute to gender differences in achievement, either independently or in combination with peer effects (Di Tommaso et al. 2024); the analytical exercise abstracts from these channels in order to isolate the role of peer effects. Specifically, we set $\alpha_B = \alpha_G = 0.1$, which calibrates well to the descriptive evidence in Figure 4. Last, we normalize the overall average to zero at gender-balanced classes ($s_c = 0.5$).

The first simulation shuts down all contextual peer effects and fixes the within-gender

endogenous effects for girls at $\beta_{GG} = 0.48$ and for boys $\beta_{BB} = 0.20$. The weighted equilibrium means for boys and girls, respectively, are given by

$$\bar{y}_{Bc} = \frac{\alpha_B}{1 - \beta_{BB}(1 - s_c)}, \qquad \bar{y}_{Gc} = \frac{\alpha_G}{1 - \beta_{GG} s_c}. \tag{8}$$

Figure 5a shows a clear positive slope for girls and a slightly negative slope for boys. As the share of girls rises, girls are increasingly exposed to the stronger within-girl multiplier, while boys are placed in smaller boy groups with a weaker within-boy multiplier. Strong within-girl endogenous peer effects are therefore sufficient to generate a positive composition gradient for girls, while weaker within-boy effects produce a negative but flatter gradient for boys as the share of girls rises. This pattern mimics the descriptive evidence in Figure 4.

Our heterogeneity results previously showed that boys' endogenous peer effects are much stronger in classes with a higher share of boys, the second simulation takes this into account by letting β_{BB} vary linearly with class composition, calibrated so that $\beta_{BB}(s_c) = 0.5$ when $s_c = 0.2$ and $\beta_{BB}(s_c) = 0.1$ when $s_c = 0.8$. The boys' equation becomes

$$\bar{y}_{Bc} = \frac{\alpha_B}{1 - \beta_{BB}(s_c)\,(1 - s_c)}. \tag{9}$$

while the girls' equation is left unchanged. Figure 5b shows that boys now perform very well in classes with a low share of girls, since they are both more numerous and subject to a stronger within-boy multiplier.[4] However, this overshoots the descriptive evidence in Figure 4, where the boys' slope is less sharp, and it also contradicts the general finding that average classroom achievement tends to be higher in classes with a larger share of girls.

A third simulation shows how contextual peer effects can reconcile the endogenous estimates with the descriptive evidence. We add a positive same-gender contextual effect for girls, $\gamma_{GG} = 0.06$, normalize $\bar{x}_{Gc} = 1$, while leaving the boys' equation unchanged. The girls' equation becomes

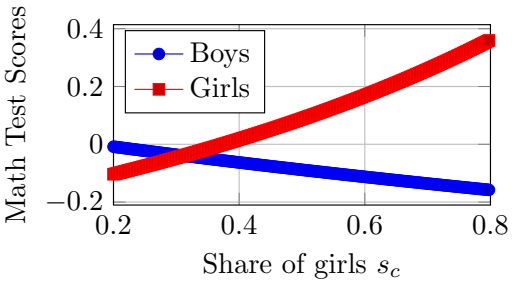
$$\bar{y}_{Gc} = \frac{\alpha_G + \gamma_{GG} s_c}{1 - \beta_{GG} s_c}. \tag{10}$$

Figure 5c shows that girls continue to benefit from a higher share of girls, while the boys' curve becomes much flatter, close to the descriptive evidence in Figure 4. Boys' absolute outcomes in boy-majority classes are unchanged from Figure 5b; girls simply gain even more from the combined contextual and endogenous effects, so the boys' relative advantage shrinks. Reconciling the endogenous estimates with the descriptive relationship therefore requires a contextual channel associated with girls: the observed pattern is consistent with a combination of stronger within-girl endogenous feedback and positive contextual effects associated with girls in the classroom environment. This calibration is illustrative, not unique — other level-shifting channels, such as a gender difference in baseline achievement unrelated to peers, likely also play

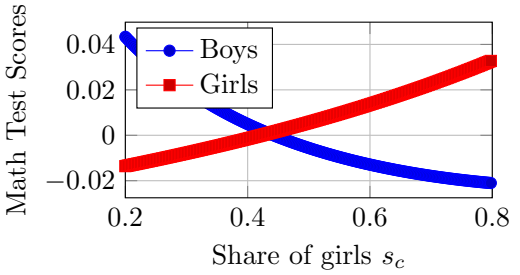
[4] What matters for group averages is the interaction between the endogenous coefficient and the size of the relevant gender group. Of course, this concerns equilibrium group averages so the stronger within-girl multiplier still implies that regardless of gender class shares, a girl gains more from a high-achieving girl than a boy does from a high-achieving boy.

a role — but reproducing the slopes still requires a sizable, composition-varying within-gender endogenous effect.

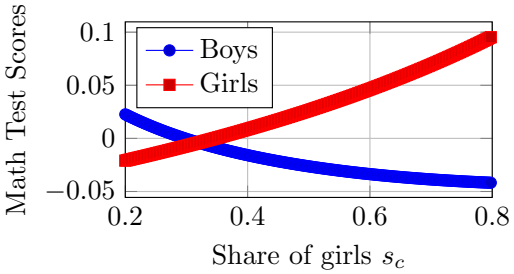
Figure 5: **Within-gender equilibrium means as functions of the girls' share s_c.**



(a) Endogenous peer effects only.



(b) Adding composition-dependent within-boy endogenous effects.



(c) Adding a positive same-gender contextual effect for girls.

Notes: Within-gender equilibrium means are measured relative to the overall class mean in a balanced class with $s_c = 0.5$. Panel (a) fixes $\beta_{GG} = 0.48$ and $\beta_{BB} = 0.20$ with all contextual effects set to zero. Panel (b) additionally lets the within-boy endogenous effect vary linearly with class composition, $\beta_{BB}(s_c)$, calibrated so that $\beta_{BB} = 0.5$ when $s_c = 0.2$ and $\beta_{BB} = 0.1$ when $s_c = 0.8$. Panel (c) additionally sets the girls' same-gender contextual peer effect to $\gamma_{GG} = 0.06$ with $\bar{x}_{Gc} = 1$; all other contextual effects remain zero. Throughout, $\beta_{BG} = \beta_{GB} = 0$ and $\alpha_B = \alpha_G = 0.1$.

It is worth noting that within-gender contextual effects could generate the same composition pattern as the endogenous-only models in Figures 5a and 5b: a positively sloped girls' curve and a negatively sloped boys' curve. Under the endogenous model, however, these curves also trace how a common shock to baseline achievement, such as improved teacher quality, is amplified differently across genders and classroom compositions. A purely contextual model lacks this feedback so that a common shock simply raises outcomes uniformly. Further, the fact that within-gender endogenous effects, rather than contextual effects alone, help explain the observed gender-share relationship also means that there is no fixed linear return to reallocating girls. Moving one girl changes the relevant peer multiplier, so its effect depends on the composition of both the sending and receiving classrooms.

Another important result of both the empirical descriptive patterns and the analytical simulations is that aggregate outcomes are maximized under complete gender segregation (see Appendix Figure A3 for the gender-specific simulations collapsed to class means). This is a natural consequence of strong within-gender endogenous effects combined with negligible cross-gender endogenous effects.

One important caveat is that only complete segregation avoids disadvantaging girls or boys who are in the minority in a class and would otherwise remain in mostly segregated classes. As long as any class contains a small minority of the opposite gender, those students face a weak own-gender multiplier and bear a disproportionate share of the cost of a segregation allocation policy. Because our empirical design does not observe classes with complete or near-complete gender segregation, this policy of complete segregation extrapolates beyond the support of our data and should be interpreted with caution.

8 Concluding Discussion

This paper studies gender-specific peer effects in Swedish ninth-grade mathematics using register data and a linear-in-means framework that distinguishes endogenous effects (peer outcomes) from contextual effects (peer characteristics). We exploit class-size variation for identification following Lee (2007) and Boucher et al. (2014), and estimate the model separately by gender. This not only strengthens identification by taking advantage of uneven gender composition across classrooms, but is also substantively motivated by evidence that the relevant peer interactions are likely to operate primarily within- rather than across-gender.

Our main finding is that endogenous peer effects are substantial and considerably stronger for girls (0.48) than for boys (0.20). The social-multiplier gap—1.93 for girls versus 1.25 for boys—implies that classroom-level interventions targeting girls will, on average, be amplified into substantially larger equilibrium effects than identical interventions targeting boys. However, we also show that in boys-majority classes the endogenous effect for boys rises to 0.48—matching the girls' benchmark magnitude—and the boys' social multiplier climbs from 1.25 to approximately 1.93. The amplification advantage girls enjoy on average thus depends critically on classroom gender composition.

We believe these estimates improve on the earlier empirical application of the model by Boucher et al. (2014), on which our paper builds. Their analysis defines peer groups at the school-grade level rather than the classroom level, resulting in substantially larger groups on average. At the same time, the standard deviation of group size in their study is roughly as large as the mean, indicating considerable variation and suggesting that their data include some very small cohorts that contribute meaningfully to identification, much as small effective groups do in our setting. A key difference is that school-grade cohorts with fewer than 10 students may be less representative of typical cohorts. By splitting classrooms by gender, our approach instead distributes the identifying variation across a broader portion of the class-size distribution and draws on classes that are likely more representative of typical school settings.

To assess whether these effects are specifically within-gender, we randomly reassign gender labels within classrooms and find much weaker endogenous effects for both groups, suggesting that the relevant peer dynamics operate primarily within gender rather than through opposite-gender classmates. Heterogeneity analyses show that endogenous effects are strongest in lower-

education classrooms, and generally decrease as average parental education rises. When splitting classes by the dispersion of parental education, peer effects tend to be strongest in the lowest- and highest-dispersion terciles.

It is important to note that the empirical results in this paper may depend on Sweden's institutional and normative environment. Academic achievement gaps across boys and girls differ across countries, making it plausible that gender peer effects also vary across contexts. Cross-country evidence generally finds that girls perform relatively better in math in more gender-equal countries (Guiso et al. 2008, Nollenberger et al. 2016, González de San Román & De La Rica 2012), although the Arab Gulf is a notable exception, with girls outperforming boys in math despite low gender-equality scores. Whether the stronger within-girl multiplier estimated here is itself culturally specific is not self-evident. Local-average mechanisms such as conformity, sensitivity to peer norms, and the motivational pull of feeling needed in the group (Shkoza et al. 2023, Weber et al. 2009) are likely shaped by gender norms and may therefore differ across contexts, but gender-segregated friendship networks are robust across countries (Shrum et al. 1988, McMillan 2022). Further, cross-country variation in gender gaps in academic achievement is also likely to operate through non-peer channels, such as curriculum content, teaching styles, grading practices, parental and teacher expectations, gendered self-concept, and math anxiety (Di Tommaso et al. 2024). These factors may shift baseline achievement gaps between boys and girls differently across countries, and endogenous peer multipliers then amplifies these baseline differences. Assuming girls' stronger endogenous peer effects generalize beyond Sweden, the same non-peer shift produces a larger equilibrium change for girls than for boys, so observed cross-country gaps reflect baseline differences filtered through gender-heterogeneous amplification.

The analytical model helps interpret these empirical findings and relate them to the reduced-form literature on classroom gender composition and academic achievement, as well as to our own descriptive results linking classroom gender composition to math test scores – where an increasing share of girls is associated with higher math scores for girls but a negative (though statistically insignificant) effect for boys. Stronger within-girl endogenous peer effects alone reproduce both the positive relationship between the share of girls and girls' achievement and the flatter profile observed for boys. Intuitively, if peer effects run within gender and are stronger for girls than for boys, a larger share of girls naturally raises girls' achievement while leaving boys' profile flatter — precisely the pattern we observe. Allowing the within-boy endogenous effect to vary with composition matches our heterogeneity result but overstates boys' outcomes in girl-scarce classes. Only when a positive within-girl contextual effect is added does the model reproduce the full empirical pattern, indicating that both an endogenous and a contextual channel are needed to reconcile the descriptive and structural estimates.

Under our calibrated analytical model, aggregate achievement rises toward the extrapolated endpoint of complete gender segregation — a qualitative pattern driven generally by strong within-gender relative to weak between-gender endogenous effects, rather than by the specific

calibration values used to generate the simulated curves. However, this aggregate gain comes at a distributional cost under partial segregation, where students of the minority gender in a class face weak own-gender multipliers. The only way to avoid this disbenefit for the minority gender is complete segregation. However, our empirical design does not specifically target classes with complete or near-complete gender segregation, so this policy conclusion extrapolates beyond the support of our data and should be interpreted with caution.

An important avenue for future research is to move beyond the linear-in-means structure within classes, which does not capture nonlinearities or richer network effects (Boucher et al. 2024). Manski (2000) suggests that individual behavior may respond to features of group behavior other than the mean. This can resolve the identification problem, provided that the functional form of group behavior is known ex ante. It is possible that this can be combined with group size variation to estimate non-linear endogenous peer effects.

Last, another natural next step for future research is to identify which specific behaviors generate the within-gender endogenous effects on academic achievement, for example through surveys. Lavy & Schlosser (2011) show that increasing the share of girls reduces classroom disruption and violence — but through a purely compositional mechanism. While this paper shows strong endogenous peer effects, the specific types of peer behavior generating them remain an open question.

References

Arnott, R. & Rowse, J. (1987), 'Peer group effects and educational attainment', *Journal of Public Economics* **32**(3), 287–305.

Block, P. & Grund, T. (2014), 'Multidimensional homophily in friendship networks', *Network Science* **2**(2), 189–212.

Böller, S. & Kruse, H. (2025), 'Friendship segregation and class composition in schools: A systematic analysis of the role of attribute consolidation', *PLOS ONE* **20**(12), e0339581.

Boucher, V., Bramoullé, Y., Djebbari, H. & Fortin, B. (2014), 'Do peers affect student achievement? evidence from canada using group size variation', *Journal of Applied Econometrics* **29**(1), 91–109.

Boucher, V., Rendall, M., Ushchev, P. & Zenou, Y. (2024), 'Toward a general theory of peer effects', *Econometrica* **92**(2), 543–565.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009), 'Identification of peer effects through social networks', *Journal of Econometrics* **150**(1), 41–55.

Briole, S. (2021), 'Are girls always good for boys? short and long term effects of school peers' gender', *Economics of Education Review* **84**, 102150.

Brock, W. A. & Durlauf, S. N. (2001), Interactions-based models, *in* J. J. Heckman & E. Leamer, eds, 'Handbook of Econometrics', Vol. 5, North-Holland, pp. 3297–3380.

Burke, M. A. & Sass, T. R. (2013), 'Classroom peer effects and student achievement', *Journal of Labor Economics* **31**(1), 51–82.

Campigotto, N., Rapallini, C. & Rustichini, A. (2022), 'School friendship networks, homophily and multiculturalism: evidence from european countries', *Journal of Population Economics* **35**(4), 1687–1722.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. & York, R. L. (1966), *Equality of Educational Opportunity*, U.S. Government Printing Office, Washington, D.C.

Davezies, L., d'Haultfoeuille, X. & Fougère, D. (2009), 'Identification of peer effects using group size variation', *The Econometrics Journal* **12**(3), 397–413.

De Giorgi, G., Pellizzari, M. & Redaelli, S. (2010), 'Identification of social interactions through partially overlapping peer groups', *American Economic Journal: Applied Economics* **2**(2), 241–275.

Di Tommaso, M. L., Contini, D., De Rosa, D., Ferrara, F., Piazzalunga, D. & Robutti, O. (2024), 'Tackling the gender gap in mathematics with active learning methodologies', *Economics of Education Review* **100**, 102538.

Diemer, A. (2022), 'Endogenous peer effects in diverse friendship networks: Evidence from swedish classrooms', *Economics of Education Review* **89**, 102269.

Fournet, J. & Barrat, A. (2014), 'Contact patterns among high school students', *PLOS ONE* **9**(9), e107878.

González de San Román, A. & De La Rica, S. (2012), 'Gender gaps in pisa test scores: The impact of social norms and the mother's transmission of role attitudes'.

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

Hahn, Y., Islam, A., Patacchini, E. & Zenou, Y. (2020), 'Friendship and female education: Evidence from a field experiment in bangladeshi primary schools', *The Economic Journal* **130**(627), 740–764.

Hanushek, E. A. (1972), 'Education and race: An analysis of the educational production process.'.

Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003), 'Does peer ability affect student achievement?', *Journal of Applied Econometrics* **18**(5), 527–544.

Holmlund, H., Lindahl, E. & Roman, S. (2023), 'Immigrant peers in the class: Effects on natives' long-run revealed preferences', *Labour Economics* **82**, 102360.

Horrace, W. C., Jung, H., Presler, J. L. & Schwartz, A. E. (2025), 'What makes a classmate a peer? examining which peers matter in nyc elementary schools', *Journal of Population Economics* **38**(3), 52.

Hoxby, C. M. (2000), 'Peer effects in the classroom: Learning from gender and race variation', *NBER Working Paper* .

Hoxby, C. M. & Weingarth, G. (2005), Taking race out of the equation: School reassignment and the structure of peer effects, Technical report, Working paper.

Hu, F. (2015), 'Do girl peers improve your academic performance?', *Economics Letters* **137**, 54–58.

Isphording, I. E. & Zölitz, U. (2020), 'The value of a peer', *University of Zurich, Department of Economics, Working Paper* (342).

Jackson, C. K. (2012), 'Single-sex schools, student achievement, and course selection: Evidence from rule-based student assignments in trinidad and tobago', *Journal of Public Economics* **96**(1-2), 173–187.

Jonsson, J. O. & Mood, C. (2008), 'Choice by Contrast in Swedish Schools: How Peers' Achievement Affects Educational Choice', *Social Forces* **87**(2), 741–765.

Kroneberg, C., Kruse, H. & Wimmer, A. (2021), 'When ethnicity and gender align: Classroom composition, friendship segregation, and collective identities in european schools', *European Sociological Review* **37**(6), 918–934.

Lavy, V., Paserman, M. D. & Schlosser, A. (2012), 'Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom', *The Economic Journal* **122**(559), 208–237.

Lavy, V. & Schlosser, A. (2011), 'Mechanisms and impacts of gender peer effects at school', *American Economic Journal: Applied Economics* **3**(2), 1–33.

Lee, L.-F. (2007), 'Identification and estimation of econometric models with group interactions, contextual factors and fixed effects', *Journal of Econometrics* **140**(2), 333–374.

Lénárd, T. (2026), 'Gender peer effects on educational achievement in swedish compulsory schools: A study of contemporaneous and cumulative effects', *Sociology of Education* **99**(2), 153–173.

Liu, X., Patacchini, E., Zenou, Y. & Lee, L.-F. (2014), 'Distinguishing between local-average and local-aggregate peer effects', *Journal of Economic Behavior & Organization* **108**, 123–141.

Lu, F. & Anderson, M. L. (2015), 'Peer effects in microenvironments: The benefits of homogeneous classroom groups', *Journal of Labor Economics* **33**(1), 91–122.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies* **60**(3), 531–542.

Manski, C. F. (2000), 'Economic analysis of social interactions', *Journal of Economic Perspectives* **14**(3), 115–136.

Martinot, P., Colnet, B., Breda, T., Sultan, J., Touitou, L., Huguet, P., Spelke, E., Dehaene-Lambertz, G., Bressoux, P. & Dehaene, S. (2025), 'Rapid emergence of a maths gender gap in first grade', *Nature* **643**(8073), 1020–1029.

Mastrandrea, R., Fournet, J. & Barrat, A. (2015), 'Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys', *PLOS ONE* **10**(9), e0136497.

McMillan, C. (2022), 'Strong and weak tie homophily in adolescent friendship networks: An analysis of same-race and same-gender ties', *Network Science* **10**(3), 283–300.

Meluzzi, F. (2024), 'The college melting pot: Peers, culture and women's job search', *arXiv preprint arXiv:2409.20225* .

Moffitt, R. A. (2001), Policy interventions, low-level equilibria, and social interactions, *in* S. N. Durlauf & H. P. Young, eds, 'Social Dynamics', MIT Press, pp. 45–82.

Nollenberger, N., Rodríguez-Planas, N. & Sevilla, A. (2016), 'The math gender gap: The role of culture', *American Economic Review* **106**(5), 257–261.

OECD (2019), 'Results (volume ii): where all students can succeed', *Organisation for Economic Co-operation and Development: Paris, France* .

Oosterbeek, H. & Van Ewijk, R. (2014), 'Gender peer effects in university: Evidence from a randomized experiment', *Economics of Education Review* **38**, 51–63.

Opper, I. M. (2019), 'Does helping john help sue? evidence of spillovers in education', *American Economic Review* **109**(3), 1080–1115.

Park, H., Behrman, J. R. & Choi, J. (2013), 'Causal effects of single-sex schools on college entrance exams and college attendance: Random assignment in seoul high schools', *Demography* **50**(2), 447–469.

Proud, S. (2014), 'Girl power? an analysis of peer effects using exogenous changes in the gender make-up of the peer group', *Regional and Sectoral Economic Studies* **14**(3), 5–18.

Rosenqvist, E. (2018), 'Two functions of peer influence on upper-secondary education application behavior', *Sociology of Education* **91**(1), 72–89.

Sarkisian, R. & Yamashita, T. (2024), 'Optimal student allocation with peer effects', *Review of Economic Design* **28**(3), 551–571.

Sewell, W. H., Haller, A. O. & Portes, A. (1969), 'The educational and early occupational attainment process', *American Sociological Review* **34**(1), 82–92.

Shkoza, L., Uysal, D. & Pohlmeier, W. (2023), Peer effects heterogeneity and social networks in education, Technical report, Discussion Paper.

Shrum, W., Cheek Jr, N. H. & MacD, S. (1988), 'Friendship in school: Gender and racial homophily', *Sociology of Education* pp. 227–239.

Statistics Sweden (2023), Ensamkommande barn 2015: Boende, utbildning och sysselsättning 2022, Report BE 57 BR2301, Statistiska centralbyrån (SCB).

Sund, K. (2013), 'Detracking swedish compulsory schools: any losers, any winners?', *Empirical Economics* **44**(2), 899–920.

Ushchev, P. & Zenou, Y. (2020), 'Mean versus aggregate peer effects in networks', *Journal of Economic Theory* **186**, 104992.

Vlachos, J. (2019), Trust-based evaluation in a market-oriented school system, *in* 'Neoliberalism and Market Forces in Education', Routledge, pp. 212–230.

Weber, B., Wittchen, M. & Hertel, G. (2009), 'Gendered ways to motivation gains in groups', *Sex Roles* **60**, 731–744.

Zölitz, U. & Feld, J. (2018), 'The effect of peer gender on major choice', *University of Zurich, Department of Economics, Working Paper* (270).

9 Appendix

9.1 Appendix for the Analytical Model

In the main text, we did not model exposure weights, in order to keep the estimated specification simple. In this appendix, we introduce an exposure-weighted version of the model in which each peer-effect term is additionally scaled by the relevant group's class share, using a large-class approximation for the scaling itself. Because this exposure weighting is not part of the specification estimated in Section 6, we treat the within gender endogenous peer effects β_{BB} and β_{GG} below as illustrative calibrations rather than as the structural estimates of Table 2 carried over unchanged. To avoid heavy notation, we use a large-class approximation in which the exact leave-one-out weights are replaced by the corresponding group shares. Let $s_c \in [0,1]$ denote the share of girls in class c, so that $1 - s_c$ is the share of boys. Under this approximation, a boy is exposed to a same-gender peer share of approximately $1 - s_c$ and an opposite-gender peer share of approximately s_c. Likewise, a girl is exposed to a same-gender peer share of approximately s_c and an opposite-gender peer share of approximately $1 - s_c$.

As in the main text, we write separate linear-in-means equations for boys and girls, where peer outcomes and peer predetermined characteristics enter with coefficients that depend on whether peers are of the same gender or the opposite gender. Since individual controls are omitted from the analytical model, the intercepts should be interpreted as reduced-form baseline terms that absorb non-peer determinants of achievement that are held fixed across simulations. For boys in class c,

$$y_{iBc} = \alpha_B + \beta_{BB}(1 - s_c)\bar{y}_{Bc,-i} + \beta_{BG}s_c\bar{y}_{Gc} + \gamma_{BB}(1 - s_c)\bar{x}_{Bc} + \gamma_{BG}s_c\bar{x}_{Gc} + \varepsilon_{iBc},$$

and for girls in class c,

$$y_{iGc} = \alpha_G + \beta_{GG}s_c\bar{y}_{Gc,-i} + \beta_{GB}(1 - s_c)\bar{y}_{Bc} + \gamma_{GG}s_c\bar{x}_{Gc} + \gamma_{GB}(1 - s_c)\bar{x}_{Bc} + \varepsilon_{iGc},$$

where $\bar{y}_{Bc,-i}$ and $\bar{y}_{Gc,-i}$ denote leave-out means, $\bar{y}_{Bc}$ and $\bar{y}_{Gc}$ denote within-class gender means, and $\bar{x}_{Bc}$ and $\bar{x}_{Gc}$ are gender-specific mean predetermined characteristics in class c. The key difference relative to the canonical linear-in-means specification is that we allow both endogenous and contextual peer effects to differ depending on whether peers are of the same gender or the opposite gender.

Taking expectations within gender in class c and assuming the mean of the error term is zero yields the gender-specific class means:

$$\bar{y}_{Bc} = \alpha_B + \beta_{BB}(1 - s_c)\bar{y}_{Bc} + \beta_{BG}s_c\bar{y}_{Gc} + \gamma_{BB}(1 - s_c)\bar{x}_{Bc} + \gamma_{BG}s_c\bar{x}_{Gc},$$

$$\bar{y}_{Gc} = \alpha_G + \beta_{GB}(1 - s_c)\bar{y}_{Bc} + \beta_{GG}s_c\bar{y}_{Gc} + \gamma_{GG}s_c\bar{x}_{Gc} + \gamma_{GB}(1 - s_c)\bar{x}_{Bc}.$$

Rearranging terms gives

$$\big(1-\beta_{BB}(1-s_c)\big)\bar{y}_{Bc}-\beta_{BG}s_c\bar{y}_{Gc}=\alpha_B+\gamma_{BB}(1-s_c)\bar{x}_{Bc}+\gamma_{BG}s_c\bar{x}_{Gc},$$

$$-\beta_{GB}(1-s_c)\bar{y}_{Bc}+\big(1-\beta_{GG}s_c\big)\bar{y}_{Gc}=\alpha_G+\gamma_{GG}s_c\bar{x}_{Gc}+\gamma_{GB}(1-s_c)\bar{x}_{Bc}.$$

We can then rewrite this as a matrix system,

$$\begin{pmatrix}1-\beta_{BB}(1-s_c) & -\beta_{BG}s_c\\ -\beta_{GB}(1-s_c) & 1-\beta_{GG}s_c\end{pmatrix}\begin{pmatrix}\bar{y}_{Bc}\\ \bar{y}_{Gc}\end{pmatrix}=\begin{pmatrix}\alpha_B+\gamma_{BB}(1-s_c)\bar{x}_{Bc}+\gamma_{BG}s_c\bar{x}_{Gc}\\ \alpha_G+\gamma_{GG}s_c\bar{x}_{Gc}+\gamma_{GB}(1-s_c)\bar{x}_{Bc}\end{pmatrix}.$$

Solving the 2×2 system yields the following equations for the boys' and girls' means:

$$\bar{y}_{Bc}(s_c)=\frac{\begin{array}{c}\big(1-\beta_{GG}s_c\big)\big[\alpha_B+\gamma_{BB}(1-s_c)\bar{x}_{Bc}+\gamma_{BG}s_c\bar{x}_{Gc}\big]\\ +\beta_{BG}s_c\big[\alpha_G+\gamma_{GG}s_c\bar{x}_{Gc}+\gamma_{GB}(1-s_c)\bar{x}_{Bc}\big]\end{array}}{\big(1-\beta_{BB}(1-s_c)\big)\big(1-\beta_{GG}s_c\big)-\beta_{BG}\beta_{GB}s_c(1-s_c)},\tag{A1}$$

$$\bar{y}_{Gc}(s_c)=\frac{\begin{array}{c}\beta_{GB}(1-s_c)\big[\alpha_B+\gamma_{BB}(1-s_c)\bar{x}_{Bc}+\gamma_{BG}s_c\bar{x}_{Gc}\big]\\ +\big(1-\beta_{BB}(1-s_c)\big)\big[\alpha_G+\gamma_{GG}s_c\bar{x}_{Gc}+\gamma_{GB}(1-s_c)\bar{x}_{Bc}\big]\end{array}}{\big(1-\beta_{BB}(1-s_c)\big)\big(1-\beta_{GG}s_c\big)-\beta_{BG}\beta_{GB}s_c(1-s_c)}.\tag{A2}$$

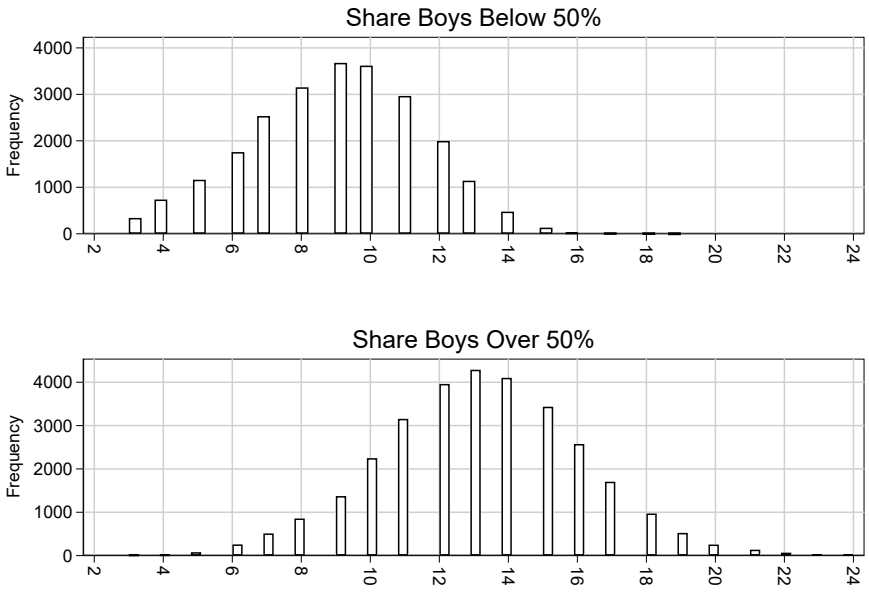
This shows that the parameters related to boys' influence are weighted by $1-s_c$, whereas those related to girls' influence are weighted by s_c.

As in the unweighted case discussed in the main text, the denominator of $\bar{y}_{Bc}(s_c)$ and $\bar{y}_{Gc}(s_c)$ is identical for both groups, including the opposite gender's within-gender feedback term. This term, however, affects the focal gender only indirectly through the simultaneous system: in the numerator of $\bar{y}_{Bc}(s_c)$, the boys' own baseline and contextual terms are scaled by $(1-\beta_{GG}s_c)$, reflecting that this term is discounted only by girls' within-gender amplifier, whereas the girls' baseline and contextual terms enter only through the cross-gender channel with weight $\beta_{BG}s_c$ (and symmetrically for $\bar{y}_{Gc}(s_c)$, with boys' terms entering through $\beta_{GB}(1-s_c)$). Thus, even when $\alpha_B=\alpha_G$, the model does not imply symmetric influence: own-group components are weighted by $(1-\beta s)$ terms, while opposite-group components are weighted by βs terms, and since $\beta<0.5$ throughout our calibrations, $(1-\beta s)>\beta s$ over the relevant range, so own-group influence dominates.

We later allow the within beta coefficient to vary with the class girl share so that we replace β_{BB} by $\beta_{BB}(s_c)$. For girls, the change in the within-girls endogenous effect is comparatively small, so we hold it constant for simplicity. For simplicity, we assume that β_{BB} depends linearly on the class girl share s_c. The coefficients in the linear specification for $\beta_{BB}(s_c)$ are chosen so that the function matches two calibration points from the empirical results shown in Table 9. Specifically, we set $\beta_{BB}(0.2)=0.5$ and $\beta_{BB}(0.8)=0.1$, and then solve for the unique linear function passing through these two points. Solving gives $\beta_{BB}(s_c)=0.6333-0.6667\,s_c$.

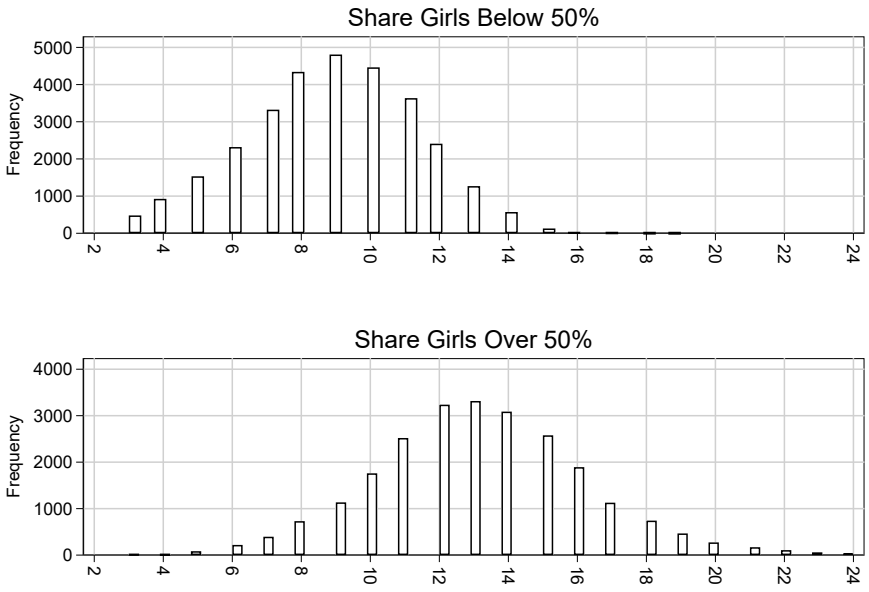
9.2 Supplementary Figures

Figure A1: **Distribution of the number of boys per class, by boy-minority vs. boy-majority classes.**



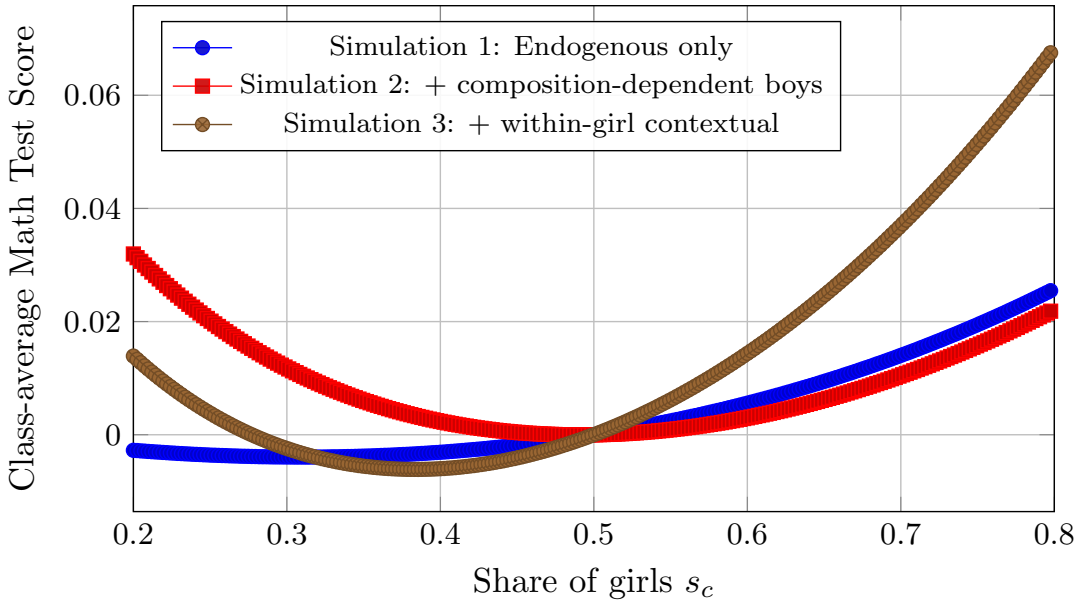
Notes: The figure plots the frequency distribution of the number of boys in a class in the estimation sample of Swedish grade-9 classes, 2013–2019. The upper panel restricts to classes in which boys constitute below 50% of the class (boy-minority classes) and the lower panel restricts to classes with boys constituting above 50% (boy-majority classes).

Figure A2: **Distribution of the number of girls per class, by girl-minority vs. girl-majority classes.**



Notes: The figure plots the frequency distribution of the number of girls in a class in the estimation sample of Swedish grade-9 classes, 2013–2019. The upper panel restricts to classes in which girls constitute below 50% of the class (girl-minority classes) and the lower panel restricts to classes with girls constituting above 50% (girl-majority classes).

Figure A3: **Class-average expected math test scores as a function of the share of girls s_c.**



Notes: Class-average expected math test scores are plotted for the three simulations of Figure 5. Each line plots the weighted mean normalized to zero at $s_c = 0.5$.