Let’s Stick Together: Peer Effects in Secondary School Choice and Variations by Student Socio-Economic Background

Zwier, D.; Geven, S.; Bol, T.; van de Werfhorst, H.G.

DOI
10.1093/esr/jcac033

Publication date
2023

Document Version
Final published version

Published in
European Sociological Review

License
CC BY-NC-ND

Citation for published version (APA):
https://doi.org/10.1093/esr/jcac033
Let’s Stick Together: Peer Effects in Secondary School Choice and Variations by Student Socio-Economic Background

Dieuwke Zwier ⋆, Sara Geven, Thijs Bol and Herman G. van de Werfhorst

Department of Sociology, University of Amsterdam, 1018 WV Amsterdam, The Netherlands

*Corresponding author. Email: d.zwier@uva.nl

Submitted February 2021; revised May 2022; accepted June 2022

Abstract

Despite the vast body of research focusing on peer effects in education, the role of the immediate peer environment in school choice has been understudied to date. We study the extent to which students from the same primary school cluster in the same secondary school, and how this effect varies by a student’s socio-economic background. We use register data from the Netherlands, covering six cohorts of students (2013–2019), that enable us to account for selection into primary schools and other endogeneity issues when identifying peer effects. The results indicate that students are more likely to choose a secondary school when students from their primary school cohort also choose this school, even after accounting for school popularity trends. We find evidence that students from socially disadvantaged backgrounds are more likely to cluster in the same secondary school as their primary school peers, yet these differences are small.

Introduction

Secondary schools tend to be highly segregated in terms of socio-economic status (SES) (Boterman, 2019; Oosterbeek, Sóvágó and Van der Klaauw, 2019). This segregation may drive inequalities in learning opportunities and educational outcomes, and impede the establishment of everyday contact with dissimilar others, which is considered crucial for fostering mutual understanding (Burgess and Wilson, 2005).

The dominant explanations for school segregation are (i) admission rules and ability tracking (Jenkins, Micklewright and Schnepf, 2008; Burgess, Greaves and Vignoles, 2019), (ii) social class-based parental preferences (e.g. Reay and Ball, 1998; Denessen, Driessena and Sleegers, 2005), and (iii) residential segregation (e.g. Boterman, 2019). These explanations have in common that they tend to reduce school choice to a choice made within the context of the family. The importance of peers for other educational outcomes is well-established: (the composition of) the peer group at school is found to affect student performance, aspirations, and school-related behaviour (Sacerdote, 2011; Perry, 2012). While a large body of research has focused on peer effects in education, studies on school choice have paid surprisingly little attention to the immediate peer environment (Lauen, 2007).

The main aim of our article is to study the extent to which students from the same primary school cluster in the same secondary school, and whether this effect
varies by a student’s socio-economic background. Our contribution to the literature on school choice is two-fold. First, we focus on primary school peers as a largely overlooked, potential source of influence in secondary school choice. We define peers as grade mates, as we expect that the influential peer group for school choice is not restricted to (close) friends (cf. Jonsson and Mood, 2008; Rosenqvist, 2018). While a few studies show that peers play a role in educational decision-making (e.g. Jonsson and Mood, 2008; Fletcher, 2012; Rosenqvist, 2018; Jansson, Birkelund and Lillehagen, 2020), they all focus on track choice or continuation decisions after compulsory education. Such outcomes comprise only part of the educational decisions students have to make. A notable exception is a recent study by Ruijs and Oosterbeek (2019). They show that students tend to choose secondary schools that many of their former grade mates in primary school choose as well. However, this study is mostly focused on school quality effects, and relies on data from just one Dutch city.

Our second contribution is that we examine the social gradient in peer effects in secondary school choice. Previous studies, including the study by Ruijs and Oosterbeek (2019), have not assessed such heterogeneities in peer effects. However, theoretically, there are reasons to believe that parents from socio-economically advantaged backgrounds play a more decisive role in the school choice processes of their offspring, thereby reducing the impact of peers. Higher SES parents generally have more knowledge about the educational system (Lareau, 2015), and tend to actively guide children into choosing the ‘best’ school (Reay and Ball, 1998), making peers redundant as a source of information to ‘navigate’ the educational system.

We study these questions in the context of the Netherlands by using comprehensive longitudinal register data from the Netherlands Cohort Study on Education (NCO, 2013–2019). The Netherlands is characterized by free school choice and relatively high levels of school segregation (Alegre and Ferrer, 2010). Given the introduction of free school choice policies in various countries (Burgess, Greaves and Vignoles, 2019; Ramos Lobato and Groos, 2019), the Netherlands provides an interesting case to assess how these policies may work out (Ladd, Fiske and Ruijs, 2009).

The register data offer important advantages compared to data that has been the main source for previous studies on school choice. First, our data cover six full cohorts of students in the transition from primary to secondary education, thereby avoiding common problems associated with selective non-response in survey data. Second, the longitudinal character and rich set of student and school characteristics included in the data allow us to control for several alternative explanations that may drive the relationship between a student’s school choice and the school choice of peers.

Using discrete choice models, we are able to account for characteristics of the secondary schools that students can choose, such as their distance from a student’s home and the ability tracks offered. Though we cannot fully account for all potential sources of endogeneity, we (partly) capture self-selection effects and systematic feeder patterns by using enrolment information from adjacent cohorts. Intuitively, we estimate the extent to which peers cluster together in secondary school above and beyond the general popularity of secondary schools within primary schools.

Theory

School Choice

School choice is often conceptualized as the outcome of a social class-dependent decision-making process. Together with their parents, students are assumed to evaluate the costs and benefits of the schools available to them. Existing research focuses on three main determinants that influence this process: (i) school admission rules and ability tracking, (ii) parental preferences, and (iii) residential segregation.

First, in countries characterized by testing, tracking, or other restrictions in school choice based on prior performance, students from higher-SES backgrounds generally attain better grades and score higher on standardized achievement tests (Jenkins, Micklewright and Schnepf, 2008), and consequently gain access to more selective schools.

Second, students from higher-SES backgrounds typically make different educational choices, net of performance. Research shows that higher-SES parents typically have more information about the educational system (Forster and van de Werfhorst, 2019), and value other criteria when making educational decisions (Ball, Bowe and Gewirtz, 1996; Croll, 2004). How parents can exercise school choice depends on the educational system. Studies in the Netherlands find that higher SES parents tend to pay more attention to school quality (Borghans, Golsteyn and Zöllitz, 2015) and the school’s pedagogical approach (Denessen, Driessena and Sleegers, 2005). They also tend to avoid schools with a high concentration of students from socially disadvantaged or ethnic minority backgrounds, due to quality or reputation concerns as well as homophilic tendencies (Karsten et al., 2006; Boterman, 2019). In contrast, lower SES parents
tend to be more pragmatic, and value locality and proximity of the school more (Karsten et al., 2006).

Third, students from different social backgrounds live in different neighbourhoods. If catchment areas determine access to local schools, residential and school segregation are inherently linked. In recent years, many Western countries have experienced the introduction of policies that grant families more autonomy in school choice (Burgess, Greaves and Vignoles, 2019; Ramos Lobato and Groos, 2019). Nonetheless, even in Dutch urban contexts (without catchment areas), residential segregation explains the largest share of primary school segregation (Boterman, 2019), though its importance is reduced in secondary school (Oosterbeek, Sóvágó and Van der Klaauw, 2019).

Peer Effects in School Choice

Next to these determinants, we expect primary school peers to be important in shaping secondary school choice. In this study, we define peers as school cohorts—i.e. all other students enrolled in the same primary school and grade. This definition has been applied in prior research examining how peers affect educational decisions (e.g. Jonsson and Mood, 2008; Rosenqvist, 2018). Another common definition of peers in the literature is (close) friends, yet studies suggest that the influential peer group for educational outcomes is not restricted to (close) friends (Frank et al., 2008; Müller et al., 2016). Empirical evidence indicates that a too narrow definition of the peer group may lead to an underestimation of peer effects (Carrell, Fullerton and West, 2009).

Three theoretical mechanisms may underlie the influence of this larger peer group in secondary school choice. First, preserving social relationships over educational transitions may be beneficial for students’ social and academic adjustment in secondary school. Jæger (2007) argues that the ‘utility’ a student gains by making educational decisions does not only have an economic but also a social component. During educational transitions, students are exposed to new peers. This could drive students who fail to integrate in the new peer environment into social isolation (Weller, 2007; Andrew and Flashman, 2017; Felmlee et al., 2018).

Existing social relationships can ease this transition, and function as ‘transitional support’. Students who enter a new school with peers they know from primary school are suggested to have more confidence, and to experience less difficulties with expanding their social network (Weller, 2007). As students with a wider network are more attractive potential friends, knowing others from primary school—even if they are not (close) friends—can reduce the risk of social isolation (Aikins, Bierman and Parker, 2005). Hence, choosing a school that maximizes the likelihood of preserving existing social relationships (Jæger, 2007) is considered a source of social and academic support.

Second, the peer group in primary school has a normative function. According to social influence theories, students are motivated to receive social approval and avoid social punishment. This is not only considered to be a goal in itself but also crucial to foster a positive self-concept. To achieve this, students tend to align their behaviour with the social norms of relevant peer groups (Cialdini and Goldstein, 2004). Conformist tendencies have been documented for various educational outcomes, including aspirations (Raabe and Wölfer, 2019), school-related behaviour (Geven, Weesie and van Tubergen, 2013), and performance (Gremmen et al., 2017). While most studies focus on adolescents, research also demonstrates the influential power of peers in primary education (Gremmen et al., 2018).

Conformity to peer norms in primary school may also shape secondary school choice. Secondary schools may differ in terms of reputation (Ball and Vincent, 1998; Reay and Lucey, 2000; Kosunen, 2014). For instance, schools could be known in the local community for their (lack of) strictness, academic orientation, or for attracting a certain type of students (e.g. nerds, snobs/ elites). In this way, peers could install a norm of ‘appropriate’ choices, and choosing a particular school can be a way to attain social approval, or avoid social punishment (Fletcher, 2012).

Third, primary school peers can be conceptualized as a source of information or social capital that students use when making school-enrolment decisions (Crosnoe, Cavanagh and Elder, 2003; Crosnoe and Muller, 2014). When school choices are discussed among classmates, students may gain access, intentionally or not, to information (Crosnoe, Cavanagh and Elder, 2003). This information may, for instance, include school quality, social composition, and selectivity in terms of ability tracks. Students may also become generally acquainted with different schools via their peers, thereby learning which schools to consider in the first place.

Based on aforementioned mechanisms via which peers may influence secondary school choice (i.e. providing transitional support, acting as a normative reference group, or providing information), we hypothesize:

H1: Students from the same primary school are likely to cluster in the same secondary school.
Social Gradient in Peer Effects

Are all students equally affected by peers in primary school? We argue that there are reasons to expect that especially students from lower SES backgrounds will be influenced by peers in choosing a secondary school.

Parents of different SES backgrounds tend to be involved in children’s education in distinct ways. According to theories on cultural capital or ‘concerted cultivation’, higher SES parents generally have more knowledge about the complexities and (un)written rules of the educational system, and how to make the system work in securing educational opportunities (Lareau, 2015). Consequently, information on the educational system that socially advantaged students may get through their school peers is likely to be redundant (Ball and Vincent, 1998). In contrast, lower SES parents are typically less equipped with knowledge relevant for making educational decisions (Forster and van de Werfhorst, 2019), and students may compensate for this at school (Crosnoe and Muller, 2014). More specifically, lower SES students may receive more new information about prospective schools or choice criteria via primary school peers.

Additionally, prior research suggests that parents from higher SES families tend to play a more decisive role in the school choice process. For instance, higher SES parents often actively guide their children into choosing what is considered by them as the ‘best’ school (in terms of, e.g. ideology or quality) (Reay and Ball, 1998; Reay and Lucey, 2000). Consequently, if social norms in the local peer context do not align with parental preferences, parents—rather than peers—are more likely to have a final say in the decision. In other words, family constraints on making certain choices are arguably larger among higher SES students, and limit the potential influence of school peers. Conversely, in less advantaged families, children are more often considered to be equipped to make educational decisions themselves (Reay and Ball, 1998; Reay and Lucey, 2000). For this reason, criteria valued by the child, such as wishes to stay close to familiar peers, may become more decisive. This links to related research showing that low-SES children are taught to solve problems in school on their own from an early age, while high-SES parents more actively intervene in these matters (Calarco, 2014).

Taken together, we expect that lower SES students are more likely to acquire new information related to secondary school choice through primary school peers. Furthermore, the typically higher levels of children autonomy in lower SES families leave more space for following choices of peers for reasons related to (i) peer group continuation (i.e. transitional support) and/or (ii) choosing a school that is valued in the local peer context (i.e. behavioural conformity). Hence, we expect:

H2: Peer effects in secondary school choice (H1) are stronger for low-SES students than for high-SES students.

The Dutch Context

Figure 1 gives an overview of the Dutch educational system. In the Netherlands, the transition from primary to secondary education (around age 12) marks the start of between-school ability tracking and is crucial for understanding educational inequalities.

At the end of primary school, students receive a quasi-binding, formally registered track recommendation by their teacher, after which they participate in a centralized exit test. Based on students’ test performance, their final track recommendation may be adjusted upwards (not downwards). Few secondary schools offer all academic tracks: some schools only offer one track, while other schools offer several adjacent tracks. In school, students are generally separated by track, though some schools offer (a) mixed ‘bridge year(s)’ to postpone the final selection moment. While students are formally allowed to apply to secondary schools that offer tracks above or below their track recommendation, this is uncommon, as schools place students based on their track recommendation. Hence, the track recommendation strongly determines the pool of schools a student can choose from.

Within these constraints, students can freely choose the school they would like to enter in almost all of the Netherlands.1 There are no catchment areas, and virtually all students are enrolled in publicly funded schools. Therefore, geographical and financial factors are less important for school choice than in other countries (Boterman, 2019; Ruijs and Oosterbeek, 2019). Due to the high population density, the vast majority of students has multiple secondary schools to choose from, and the type of transition where virtually all students from a primary school feed into the same secondary school (see ‘uniform’ pathway in Langenkamp, 2009) is very rare (i.e. 2.8 per cent in our sample). Secondary schools differ in the tracks they offer, religious denomination or learning ideology, as well as student composition, extracurricular activities, drop-out rates, etc. This makes the Netherlands a particularly interesting case to study (the social gradient in) peer effects in secondary school choice.

Data and Methods

Data

We used register data covering six full cohorts of Dutch students who made the transition from primary to
secondary school (2013/2014–2018/2019). The NCO data include both primary and secondary school enrolment, student background characteristics, and school information (see for more information Haelermans et al., 2020).

To reduce computational complexity, we selected a random 10 per cent sample in each cohort, resulting in a starting sample of 110,998 students. We restricted this sample in multiple ways (see Appendix A, Table A1 for details). Particularly, we excluded students in special needs education, and students recommended to enter a vocational programme tailored to students who would otherwise not be able to obtain a school leaving qualification (Praktijkonderwijs, see Figure 1), as their school choice processes often deviate from those of students in regular education. Additionally, we removed students choosing a school greater than 20 kilometres from home. Such travel distances are rare in the Netherlands (Statistics Netherlands, 2019), and likely include measurement errors (e.g. students living at an address where they were not registered).

We also excluded students with missing values on key school or student characteristics, such as school quality, home/school address, track recommendation, or parental education. Overall, we lost few cases for these reasons: the register data are very complete. The one exception is parental education, which is not registered for every resident and is for some imputed using the Dutch Labour Force Survey. Parental education was missing for about 15 per cent of the sample. We conducted several robustness checks to examine if our sample restrictions affect our main findings, including analyses where students with missing information on parental education are included.

After further selections when creating choice sets (see section ‘Defining Choice Sets’), the final sample consisted of 62,620 students, clustered in 6,159 primary schools and 1,110 secondary schools. Table 1 describes the student sample.

Defining Choice Sets
To analyse the influence of primary school peers on secondary school choice, we used discrete choice models. This framework allowed us to model the choice among alternatives (here, secondary schools) as a function of the alternatives, by comparing the characteristics of the chosen alternative with those of the rejected ones, within individuals. For all students, we had information on the secondary school they enrolled (i.e. revealed preference). As we could not observe the set of schools considered by
the family, a key challenge was to identify the schools each student realistically could have chosen, but did not.

We define individual choice sets based on (i) proximity to home and urbanization degree, and (ii) the tracks offered. The intuition behind our approach is visualized in Figure 2. Our results are robust to another specification of choice sets including all schools in a 10-kilometre radius (see Supplementary Appendix B, Table B.3).

First, the pool of schools was constrained by school distance. Due to the high population density in the Netherlands, the nearest secondary school is on average located 2.4 kilometres from home, with a mean number of 18.1 schools in a 10-km radius (Statistics Netherlands, 2019). Nonetheless, regional differences in school supply are large, and students in rural areas are probably willing and/or forced to travel further to their preferred school.

To take these differences into account, we first calculated the 90th percentile of the distance to all chosen schools by urbanization degree. These distances were calculated as straight-line distances (cf. Burgess et al., 2015; Ruijs and Oosterbeek, 2019). The 90th percentile of distance to chosen schools ranges from ~5 kilometres in the most densely populated areas (≥ 2,500 addresses per square kilometre) to ~12 kilometre in the most rural areas (< 500 addresses per square kilometre). Subsequently, we used these percentile scores to select schools in a certain proximity to home. We limited the sample to students choosing a school in this distance radius, excluding a small subgroup of students choosing a school relatively far away from their registered address (see Appendix A, Table A1).

Second, school choice was constrained by the track recommendation students received at the end of primary education. Students can receive a recommendation for one or two adjacent tracks. We selected all schools that offer the track(s) that match(es) the student’s track recommendation. We excluded a small share of students choosing a school (i) for which we cannot determine the tracks offered, or (ii) that does not match their track recommendation (see Appendix A, Table A1).

Taken together, the choice sets included all schools in the 90th percentile distance radius (based on urbanization degree) that fitted the student’s track recommendation. The mean number of schools to choose from is 8.0 (SD = 6.0). Figure 3 visualizes the number of schools in each choice set by urbanization degree. This illustrates that, accounting for urbanization degree, students from densely populated areas generally have a larger choice set (M = 11.5, SD = 6.9) than students from rural areas (M = 7.8, SD = 6.4).

### Method
For the main analysis, we estimated conditional logit models (CLMs). The CLM follows from an additive random utility framework, in which we assumed that students—together with their parents—faced a set of schools, and preferred the school that maximized their utility (McFadden, 1973). Every school in the individual choice set was associated with expected utility $U_{ist}$, a student $i$ derived from a given school $s$ in year $t$:

$$U_{ist} = W_{it} \beta + X_{ist} \beta + Z_{ist} \beta + \epsilon_{ist}$$

(1)

where $W_{it}$ was a vector of student-specific characteristics (e.g. parental education), $X_{it}$ included school-specific characteristics (e.g. quality), $Z_{ist}$ represented student-school specific characteristics (e.g. home-school distance), and $\epsilon_{ist}$ was a non-systematic random error term. Note that the model estimated within-person preferences for schools and that student characteristics ($W_{it}$) did not vary between alternatives. To illustrate, a student’s social background or primary school attended is the same across all alternatives, and cannot explain in itself the outcome of the choice process. Hence, student-
specific characteristics cannot be included as main effects, but can be interacted with school characteristics to study preference heterogeneity across subgroups of students.

Though utility levels $U_{ist}$ were not observed, the outcome of the choice process was. Given that a student would only choose school $k$ if its utility was higher than that of any other school in the choice set ($U_{ist} > U_{ist}$, $\forall s \neq k$), we could estimate how differences in school or student-school characteristics relate to choice probabilities. The probability $P_{ist}$ that a student chose school $s$ was represented as:

$$P_{ist} = \frac{e^{W_{ist} \beta + X_{ist} \beta + Z_{ist} \beta}}{\sum_{k=1}^{n} e^{W_{ik} \beta + X_{ik} \beta + Z_{ik} \beta}}.$$  (2)

The CLM makes the Independence of Irrelevant Alternatives (IIA) assumption. This means that outcome categories should be distinct, and adding or excluding alternatives should not affect the relative odds of choosing certain outcomes. This required choice probabilities for any two schools in the choice set to be equally affected, and preferences to remain stable, if a school would suddenly be opened or closed. This is a strong assumption: some schools may be close substitutes, and the inclusion or exclusion of a school in a choice set may change the relative importance of preferences.

For this reason, we compared our CLM estimates with those from mixed logit models (see section ‘Robustness Checks’), which do not impose the IIA assumption (Train, 2009). Due to computational limitations, we used CLMs for our main analysis. In all models, standard errors were clustered at the primary school level.

**Measurement**

We focused on parental education to measure SES, as the role of knowledge on educational institutions is emphasized in our theoretical argumentation. *Parental education* is measured as the highest education level of both parents and recoded into three categories: (i) lower secondary education or less (ISCED 0–2), (ii) upper secondary or post-secondary non-tertiary education (ISCED 3–4), and (3) higher education (ISCED 5–8).

The identification of (endogenous) peer effects in school choice is complicated for various reasons. Most importantly, primary school peers may choose the same secondary school because they share (i) certain individual characteristics (which may have also led to self-selection into the same primary school); and (ii) the same primary school environment (i.e. correlated peer effects). To illustrate the latter, teachers at a student’s primary school may recommend certain secondary
schools that previous student cohorts have good experiences at, or primary schools may have ties to or cooperate with secondary schools in the neighbourhood, which may lead students from the same primary school to flow into the same secondary school. Estimates may also be biased by the reflection problem (Manski, 1993): since students’ peers in primary school may affect their school choice, and vice versa, the causation runs in both directions. Unfortunately, we cannot account for the reflection problem as we use cross-sectional data on a behavioural outcome. However, we did try to account for self-selection and unobserved school characteristics that could lead to feeder patterns between primary and secondary schools.

More specifically, we studied peer effects in school choice with two indicators. First, we measured the actual percentage of primary school peers enrolled in the same secondary school (cf. Ruijs and Oosterbeek, 2019). We defined the peer group as school cohorts—i.e. all other students who are enrolled in the same primary school and grade. Since Dutch primary schools are generally not very large, it is likely that students also interact with peers from other classes in the same grade (e.g. during breaks, school trips, or other activities). Note that, to construct this variable, we did not only use information on peers who are included in the analytical sample themselves. Instead, we used information on the school choices of all of the sampled students’ grade mates. While we define peers as grade mates in our main analysis, we conducted supplementary analyses with more specific peer group conceptualizations, since students might be stronger influenced by choices of similar peers (i.e. same-gender and similar-track peers, see section ‘Robustness Checks’).

The actual percentage of peers is considered a rather naive estimator of the true peer effect, as its estimated effect may also include self-selection and/or feeder effects. Hence, as a second measure, we included the extent to which the actual percentage of peers deviates from the percentage of peers we predicted to enrol in the same secondary school. We calculated this predicted share of peers by using the predicted values of a regression in which the actual share of primary school peers enrolled in a specific secondary school was regressed on year (cf. Ruijs and Oosterbeek, 2019). Intuitively, these predicted values captured the fact that some secondary schools are generally expected to be more popular among students from a certain primary school, irrespective of peer effects. Several factors may underlie this popularity, such as distance, denomination, or teachers advising particular schools to their students.

By subtracting the predicted share from the actual share of peers, we accounted for these feeder effects. That is, the deviation between the actual and predicted percentage of peers captures the extent to which school choices of a given cohort deviate from general enrolment patterns. While this indicator enabled us to estimate the influence of peers over and above the general (linear) popularity trend of the secondary school in a given primary school, estimates may be too conservative. Little variation in school choice patterns between consecutive student cohorts may also be a consequence of consistently strong peer effects over time. In such a case, we will not find support for peer effects, even if they did play a role. Conversely, accounting for popularity trends does not fully solve the correlated error problem, as there may still exist (time-varying) correlated unobservables that could explain estimated peer effects, such as active student recruitment by secondary schools in some years (Ruijs and Oosterbeek, 2019). Also note that we assumed popularity trends to be linear while this does not necessarily have to be the case empirically.

We controlled for several school characteristics. We included straight-line home-school distance to control for travel distance. Given the importance of proximity among low-SES students (Reay and Ball, 1998), home-school distance was also interacted with parental education. Families prefer schools that match their religious
beliefs (Karsten et al., 2006; Borghans, Golsteyn and Zölitz, 2015) and primary schools may be more likely to have ties to secondary schools with a similar religious denomination or learning ideology. Lacking data on individual religiosity, we accounted for denomination match between primary and secondary school. We distinguished between secondary schools without denomination (i), and schools with a denomination or learning ideology that either did not match (ii), or matched (iii) that of the primary school.

Furthermore, we controlled for secondary school composition. We measured the percentage of high-SES students as the percentage of students with at least one higher educated parent enrolled in the first grade of secondary school. Similarly, we calculated the percentage of students with a non-western migration background. Both composition measures were lagged with one school year. We interacted the composition measures with parental education, given empirical research showing that higher-SES parents are on average more likely to avoid schools with large shares of socially disadvantaged or ethnic minority students (Karsten et al., 2006; Bjerre-Nielsen and Gandil, 2020).

Finally, we controlled for secondary school quality by using graduation data that are made publicly available on a website that families can consult when choosing a secondary school. More specifically, we used the (i) yearly graduation rate, (ii) central exam grade point average (CE GPA), and (iii) discrepancy between CE and school exam (SE) GPAs. We standardized all measures, and combined them into one factor capturing general quality. The items loaded on one factor, explaining 70.1 per cent of the variance, with loadings of 0.78 (graduation rate), 0.90 (CE GPA), and 0.83 (CE-SE discrepancy). Our quality index was lagged with 2 years to ensure that this information was publicly available at the moment students made their school choice. School quality was also interacted with parental SES, as this is demonstrated to be a stronger driver of school choice in socio-economic advantaged families (Borghans, Golsteyn and Zölitz, 2015; Burgess et al., 2015).

**Results**

**Descriptive Results**

Table 2 summarizes school characteristics of (i) all choice alternatives in our sample, and (ii) the schools chosen by the students. This table shows that the mean actual and predicted shares of primary school peers are substantially higher among the chosen schools (26.4 and 24.0 per cent), compared to all choice alternatives (6.6 and 6.3 per cent), providing descriptive evidence that students from the same primary schools tend to cluster in the same secondary schools. As mentioned before, this difference does not have to be explained by peer effects only but also by other factors, such as conventional student flows or (mis)matches in denomination.

If we separate the sample by parental education (not shown in Table 2), no clear socio-economic differences in peer-related school preferences arise. The mean difference in the share of actual peers between the chosen schools and all alternatives is roughly similar among students with lower/medium educated parents (26.7–6.7 per cent = 20.0%p), and students with at least one college-educated parent (26.0–6.5 per cent = 19.5%p). Note, however, that these descriptive patterns cannot illuminate (the social gradient in) peer effects in individual choice sets.

Table 2 also highlights other school preferences. For instance, distances to chosen schools are on average lower than to all choice alternatives. Furthermore, when choosing for a school with a religious denomination or learning ideology, students display a relative preference for a school that matches the denomination of the primary school they went to. Students seem generally less likely to choose for a school with a higher percentage of students with a non-western migration background. We find no clear preference for the school’s socio-economic composition. If we distinguish between students with lower and higher educated parents (not shown in Table 2), we see that schools chosen by students from disadvantaged SES backgrounds tend to have a lower share of students from high-SES backgrounds (42.0 per cent) than schools chosen by socially advantaged students (59.4 per cent). In schools available to students with lower educated parents, a lower share of high-SES students is enrolled (42.2 per cent) than in schools available to students from advantaged backgrounds (55.6 per cent). This demonstrates that these compositional patterns are partly due to social inequalities in track recommendations. Finally, we observe no clear descriptive preferences for school quality.

**Regression Results**

The main CLM results are summarized in Table 3. We ran similar models for both peer group indicators. We mean-centred and standardized all continuous variables to ease interpretation (except for distance, in kilometres). Results are presented in log odds—i.e. positive coefficients denote school attributes that increase the probability of choosing a school, whereas negative coefficients denote attributes that make schools less
attractive. Note that it is not possible to present results as (differences in) predicted probabilities: though each student faces a limited choice set, the total number of outcome categories exceeds 1,000.

Model 1a includes the actual percentage of peers and the control variables. In line with hypothesis 1, the percentage of primary school peers is positively associated with the probability of choosing a secondary school. More specifically, with every SD increase in the percentage of peers enrolled in a secondary school, the odds of choosing this school increases with a factor of \( e^{0.827} \approx 2.29 \). In other words, a student is 2.29 times more likely to enroll in a secondary school chosen by a substantial share of grade mates from primary school, compared to another school in the choice set that is less popular among grade mates—net of proximity, denomination match, school composition, and school quality. This is a substantial effect: according to this model, students are willing to travel 3.3 km further to attend a school chosen by a larger share (+1SD) of primary school peers.\(^{15}\)

In Model 2a, an interaction term between parental education and the actual percentage of primary school peers is added to study preference heterogeneity (hypothesis 2). As previously mentioned, we estimated within-student preferences, so parental education could not be included as main effect. This model shows that the positive effect of the actual percentage of peers enrolled in a secondary school on the probability of choosing that school is significantly less strong among students with medium \( e^{0.952-0.133} \approx 2.27 \) and higher \( e^{0.952-0.146} \approx 2.24 \) educated parents, compared to students with lower educated parents \( e^{0.952} \approx 2.59 \). As predicted by hypothesis 2, this suggests that peer effects are slightly stronger for children with lower educated parents. Compared to its main effect, the interaction term is small in size.

### Table 2. Descriptive statistics secondary schools

<table>
<thead>
<tr>
<th></th>
<th>(a) All alternatives</th>
<th>(b) Chosen schools</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary school peer characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual % primary school peers</td>
<td>6.62 (13.66)</td>
<td>26.39 (22.41)</td>
</tr>
<tr>
<td>Predicted % primary school peers</td>
<td>6.30 (12.83)</td>
<td>23.96 (21.34)</td>
</tr>
<tr>
<td>Deviation actual-predicted peers</td>
<td>0.31 (4.18)</td>
<td>2.44 (7.66)</td>
</tr>
<tr>
<td><strong>Other school characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance (kilometres)</td>
<td>4.75 (2.80)</td>
<td>3.24 (2.38)</td>
</tr>
<tr>
<td>Denomination match</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public school</td>
<td>0.26 (0.27)</td>
<td></td>
</tr>
<tr>
<td>No denomination match</td>
<td>0.47 (0.37)</td>
<td></td>
</tr>
<tr>
<td>Denomination match</td>
<td>0.28 (0.36)</td>
<td></td>
</tr>
<tr>
<td>% high-SES students</td>
<td>49.82 (23.96)</td>
<td>51.39 (21.22)</td>
</tr>
<tr>
<td>% non-western students</td>
<td>23.98 (20.42)</td>
<td>18.06 (15.50)</td>
</tr>
<tr>
<td>School quality (index)</td>
<td>-0.13 (1.00)</td>
<td>-0.05 (0.87)</td>
</tr>
<tr>
<td>Graduation rate</td>
<td>91.37 (5.47)</td>
<td>91.77 (4.73)</td>
</tr>
<tr>
<td>Mean GPA central exam</td>
<td>6.42 (0.20)</td>
<td>6.43 (0.17)</td>
</tr>
<tr>
<td>Deviation GPA CE-SE</td>
<td>-0.05 (0.15)</td>
<td>-0.04 (0.14)</td>
</tr>
<tr>
<td><strong>School type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vwo</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Vwo havo</td>
<td>0.19</td>
<td>0.21</td>
</tr>
<tr>
<td>Vwo havo vmbo-gt</td>
<td>0.39</td>
<td>0.36</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>0.11</td>
<td>0.14</td>
</tr>
<tr>
<td>Vmbo</td>
<td>0.27</td>
<td>0.24</td>
</tr>
<tr>
<td>Other</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>500,044</td>
<td>62,620</td>
</tr>
</tbody>
</table>

Source: Netherlands Cohort Study on Education (NCO).

\(^{a}\)School type is simplified into six categories. Vwo havo vmbo-gt refers to schools offering either vwo, havo, and vmbo-gt, or vwo and vmbo-gt, or havo and vmbo-gt. The category ‘Comprehensive’ refers to schools offering all, or all but one, tracks. Vmbo includes all four pre-vocational tracks (in different combinations). See Figure 1 for an overview of the Dutch educational system.
Robustness Checks

We performed several robustness checks. First, our main models do not account for unobserved differences between secondary schools. However, it is plausible that unobserved school characteristics—such as teacher quality, extracurricular activities, and reputation—affect school choice. Hence, we also estimate models where we include secondary school-fixed effects as dummies (cf. Ruijs and Oosterbeek, 2019). This implies that we estimate how school choices relate to changes in quality, composition, etc., instead of absolute scores on these variables. These results confirm the main findings (see Supplementary Appendix B, Table B.5).

Second, we compare our CLM estimates to mixed logit estimates (see Supplementary Appendix B, Table B.6). The variance components illustrate that students vary in their preference for attending the same secondary school as their primary school peers. Compared to Table 3, the main effects of both peer measures increase in size. In the model including the actual-predicted deviation indicator, estimated peer effects are significantly stronger for students with lower educated parents than for students with higher educated parents (but not versus students with medium educated parents).

Third, given the stratified nature of the Dutch educational system with not all secondary schools offering all academic tracks, students may be influenced more by peers who are recommended the same/a similar track. To explore these dynamics, we repeated the main analysis with a more specific conceptualization of the peer group, where we only included peers who were recommended the same/a similar track. The results (see Supplementary Appendix B, Table B.7) suggest that effects for both peer measures are stronger when only considering similar-track peers (rather than all grade mates). Results for the interaction between the peer group measures and parental education are consistent with the main findings.

Finally, it could be argued that boys and girls respond differently to peers, or that students are more affected by same-sex peers in educational decision-making. Supplementary analyses on subsamples by gender suggest that there are no significant gender differences in peer effects on school choice (see Supplementary Appendix B, Table B.8). Additionally, models including same-gender peer group variables are consistent with the main findings (see Supplementary Appendix B, Table B.9).

Conclusion and Discussion

While theories of social capital and norm formation and empirical studies stress the importance of the peer environment in education, the role of peers in school choice processes has received surprisingly little attention so far. In the current paper, we studied the extent to which students from the same primary school cohort cluster in the same secondary school, as well as the social gradient in these patterns. For this purpose, we used register data from the Netherlands for six student cohorts transitioning from primary to secondary education (2013−2019).

In line with our expectation, we found that students were more likely to choose a secondary school when students from their primary school cohort also enrolled in this school. This is the case when we estimated peer effects by taking the share of other students from the same primary school cohort enrolling in a given secondary school, but also when we used a more conservative approach that accounted for the general feeder patterns of primary school students in a given secondary school (cf. Ruijs and Oosterbeek, 2019). We believe that the actual magnitude of the peer effects lies somewhere between these two estimates: whereas our first estimator may give an upper bound estimate as it captures conventional student flows, the conservative estimator provides a lower-bound estimate as it underestimates peer effects that are stable over time.

The finding that peers influence students’ secondary school choice resonates with prior research pointing to the importance of the peer environment for educational decision-making, and, more specifically, school choice. While past studies have mostly highlighted the importance of ability tracking, parental preferences, and residential segregation in school choice processes, we demonstrated that peers in primary school are important actors in shaping these choices.
Table 3. Conditional logit models, secondary school choice

<table>
<thead>
<tr>
<th>Main effects</th>
<th>Actual percentage of peers</th>
<th>Dev. actual-predicted peers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1a)</td>
<td>(2a)</td>
</tr>
<tr>
<td>% actual peers</td>
<td>0.827***</td>
<td>0.952***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Deviation actual-predicted peers</td>
<td>-0.331***</td>
<td>-0.309***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Distance (kilometres)</td>
<td>-0.331***</td>
<td>-0.309***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Denomination match (ref. = public)</td>
<td>No match</td>
<td>Match</td>
</tr>
<tr>
<td></td>
<td>-0.078***</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
</tr>
<tr>
<td></td>
<td>-0.076***</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>0.305***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.014)</td>
</tr>
<tr>
<td></td>
<td>-0.249***</td>
<td>-0.248***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Interactions parental education (ref. = low)</td>
<td>Medium × % actual peers</td>
<td>High × % actual peers</td>
</tr>
<tr>
<td></td>
<td>-0.133***</td>
<td>-0.146***</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.025)</td>
</tr>
<tr>
<td></td>
<td>0.058***</td>
<td>0.058***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.010)</td>
</tr>
<tr>
<td></td>
<td>0.010***</td>
<td>0.082***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.011)</td>
</tr>
<tr>
<td></td>
<td>-0.011</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.019)</td>
</tr>
<tr>
<td></td>
<td>0.067***</td>
<td>0.065***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
</tr>
<tr>
<td></td>
<td>0.085**</td>
<td>0.085**</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.030)</td>
</tr>
<tr>
<td></td>
<td>0.597***</td>
<td>0.614***</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.031)</td>
</tr>
<tr>
<td></td>
<td>-0.194***</td>
<td>-0.183***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.028)</td>
</tr>
<tr>
<td></td>
<td>-0.036</td>
<td>-0.027</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(0.030)</td>
</tr>
</tbody>
</table>

Loglikelihood

<table>
<thead>
<tr>
<th>(1a)</th>
<th>(2a)</th>
<th>(1b)</th>
<th>(2b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-79,074.98</td>
<td>-79,040.02</td>
<td>-90,778.39</td>
<td>-90,766.36</td>
</tr>
</tbody>
</table>

Source: Netherlands Cohort Study on Education (NCO).

Notes: N = 500,044, n = 62,620. All continuous variables are mean-centred and standardized (SD = 1), except for distance (in kilometres). Cluster-robust standard errors in parentheses.

*P < 0.05, **P < 0.01, ***P < 0.001.
We also found evidence for a social gradient in peer effects in school choice, yet these differences are relatively small. Compared to students with at least one parent with a post-secondary or higher educational degree, students with parents without post-secondary or higher educational degrees were more likely to choose a secondary school when a larger share of primary school peers also enrolled in this particular school. These heterogeneities in the strength of peer effects remain after correction for school popularity effects.

Future research may benefit from examining why students tend to choose the same secondary schools as primary school peers, and variations herein by students’ socio-economic background. We proposed three related theoretical mechanisms: students may prefer to enter a new school environment with familiar peers, students may conform to social norms produced in local peer contexts, and peers may provide students with information relevant for making school choices. It is possible that students from disadvantaged backgrounds are more likely to rely on peers for information purposes, but that there is no social gradient in norm conformity or the desire to transfer to a new school with familiar peers. Another direction for future research would be to explore the role of student performance in explaining or moderating the social gradient in peer effects in school choice.

Our rich longitudinal register data enabled us to systematically examine the extent to which students from the same primary school cluster in the same secondary school for a large sample of students and schools, while accounting for various alternative explanations. However, our analyses do not provide a direct test of the proposed mechanisms. Our findings could still be consistent with alternative explanations, such as active recruitment by secondary schools in specific years, or non-linear effects of previous primary school cohorts’ choices on secondary school choice. Relatedly, though we try to address self-selection in schools by correcting for school popularity effects, it cannot be ruled out that our results partly reflect self-selection patterns. Furthermore, we cannot capture friendships dynamics, and the avoidance of negative ties, in secondary school choice. Though peer influence may reach further than (close) friendships, (longitudinal) network data are needed to answer remaining questions, such as the influence of befriended versus non-befriended peers on school choice.

By using discrete choice models, we were able to model the role of various factors in school choice processes and heterogeneities herein across socio-economic groups. Nonetheless, these methods have some caveats that may be addressed in future work. While we accounted for geography and track recommendation, and school choice is relatively free in the Netherlands, estimates may still be affected by unobserved differences in access constraints (Burgess et al., 2015). To illustrate, the social gradient in peer effects could also be interpreted as lower SES students having a narrower choice set in reality due to unobserved resources or preferences. Furthermore, a developing literature defines school choice as a two-stage process: i.e. families are assumed to screen available options and reduce this set to alternatives they are willing to consider, before evaluating options within this smaller set (Burdick-Will et al., 2020). We did not distinguish between peer effects in choice set formation and evaluation.

We studied the case of the Netherlands, but the findings may also be relevant for other national contexts. Over the past decades, many Western countries such as the United States and the United Kingdom have introduced policies in a more choice-driven direction (Burgess, Greaves and Vignoles, 2019; Ramos Lobato and Groos, 2019). Prior research raised concerns that this autonomy may enhance social class-based criteria of choice, thereby potentially increasing segregation (Wilson and Bridge, 2019).

With this article, we showed that primary school peers matter in shaping secondary school choice. To understand (inequality in) school choice, the peer context needs to be taken into account. Peer effects in secondary school choice may reproduce and even amplify segregation by social background in primary school, on top of segregation induced by ability tracking in educational transition. In other words, segregation in primary school feeds into further educational trajectories, potentially causing inequalities in attainment and future life chances. When developing policies to combat school segregation, one should thus not merely focus on isolated choice processes (within families) but also consider the peer context in which school preferences are formed.

Notes
1 Due to oversubscriptions and capacity problems, some secondary schools (especially in cities) organize lotteries to allocate students. The goal is to match students to their preferred school, but the exact procedure differs between regions and schools. Schools are allowed to apply priority rules, such as prioritizing prospective students with siblings at the same school.
2 Students from the same primary school cohort spread on average over 6.96 ($SD = 4.08$) secondary
suggested that if parental education is missing, a student’s parents likely did not follow higher education. Therefore, we repeated the main analysis on a sample where we recoded parental education into two categories (1 = high, 0 = low/medium) and replaced missing values with 0. These results can be interpreted as conservative estimates (some higher educated parents with missing information are falsely coded to low/medium), and confirm the main findings (see Supplementary Appendix B, Table B.2).

6 The students selected into the sample do not substantially differ from those excluded in terms of age and gender. Students from non-western migration backgrounds are slightly overrepresented (18.7 per cent among sampled members, versus 15.4 per cent among excluded students), as well as students with higher educated parents (53.9 versus 45.3 per cent).

7 We use mean GPA coordinates at the neighbourhood level for the home and school address. Neighbourhoods are demarcated parts of a municipality that are homogenous in terms of building structure or socio-economic function (e.g. housing, industry, recreation). The median neighbourhood is 0.8 square kilometres.

8 We repeated the main analyses with a sample including students who pick a school outside the distance radius. For them, we included all schools with a distance up to the chosen school in the choice set. This does not alter our main findings (see Supplementary Appendix B, Table B.4).

9 Registered track enrolment is unreliable in grades 7–8. For this reason, we follow the Dutch Inspectorate of Education in using track enrolment of grade 9 students in a school at the time our sample members enter secondary education (grade 7) to determine the tracks offered in a particular school.

10 The median cohort size in our sample is 29 students; 52.9 per cent of all primary schools have 30 students or less in grade 6, and 7.3 per cent of the schools have more than 60 students.

11 We only calculate these trends if the primary–secondary school combination can be observed for at least 3 years in the study’s time span (some schools opened after 2013/2014, or do not exist anymore in 2018/2019).

12 To illustrate, for a hypothetical primary–secondary school combination with actual shares of peers of 18, 23, 21, 37, and 22 per cent in consecutive cohorts from 2013/2014 to 2017/2018, the predicted share of peers equals 26.4 per cent in 2016/2017. Hence, the deviation for the 2016/2017 cohort is 37–26.4 = 10.6%.

13 In the Netherlands, graduation from secondary school is determined for 50 per cent by school exams and for 50 per cent by standardized central exams. A large negative CE-SE GPA discrepancy suggests that school exams are too easy. Until recently, the Inspectorate of Education used this indicator to assess the performance of secondary schools.

14 There may be reasons not to control for school composition and quality and their interaction with parental education. The fact that higher-SES parents actively guide their child into choosing what they consider the ‘best’ school may partly cause a social gradient in peer effects in school choice. Therefore, we repeated the analyses without controlling for school composition and quality. The results (not shown) are consistent with the presented results.

15 To illustrate the magnitude of peer effects, we calculated this willingness to travel by dividing the coefficient for the percentage of peers ($b = 0.827$) by the distance coefficient (cf., Borghans, Golsteyn and Zöllitz, 2015). As distance is interacted with parental education, we used the net effect for students with medium educated parents ($-0.331 + 0.081 = -0.250$).

16 In this model, the effects of some control variables substantially increase in size. This may be due to the fact that feeder effects are partly captured by school characteristics instead of being conflated with the effect of the actual percentage of peers.

Supplementary Data

Supplementary data are available at ESR online.

Acknowledgements

This study uses non-public microdata from Statistics Netherlands (CBS) and the Netherlands Cohort Study on Education (NCO). Under certain conditions, these data are accessible for statistical and scientific research. For further
information, and to request access to the data, please visit the website of CBS (https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research) and NCO (https://www.nationaalcohortonderzoek.nl/onderzoek). Earlier versions of this paper were presented at the European Consortium for Sociological Research (ECSR) 2020 (online) and the Interuniversity Center for Social Science Theory and Methodology (ICS) Forum Day 2020 (online). The authors thank the participants at those meetings, and the anonymous reviewers for their helpful comments and suggestions on earlier drafts of this paper.

Funding
This study was made possible through a grant awarded to René Veenstra, Jan Kornelis Dijkstra, Herman G. van de Werfhorst, Thijs Bol, and Sara Geven for the project ‘Peer Relations in the Transition from Primary to Secondary school: Social, Behavioral and Academic Aspects of Social Integration’ by the Netherlands Initiative for Education Research (NRO), grant number 40.5.18325.001, a Veni grant awarded to Sara Geven by the Netherlands’ Organisation for Scientific Research (NWO), grant number 016.Veni.195.125, and a grant awarded to Dieuwke Zwier for usage of the Netherlands Cohort Study on Education (NCO) data by NRO, grant number 40.5.20326.005.

References


**Diewuize Zwier** is a PhD student in the Department of Sociology at the University of Amsterdam. Her research interests lie in social inequalities and education. Her PhD research focuses on how peer relations in primary education affect educational decisions and outcomes in secondary education. This project is embedded in the Transition from PRIMary to Secondary school (PRIMS) data collection.

**Sara Geven** is an assistant professor in the Sociology Department of the University of Amsterdam. Her research focuses on educational inequality, particularly between students from different socioeconomic and ethnic backgrounds. In her past work she has examined the role of social networks in educational inequality. Currently she studies how institutional arrangements structure inequality in teacher expectations and track recommendations, funded by an NWO-Veni grant. Her work appeared in Sociology of Education, Social Networks, Journal of Youth and Adolescence, and Journal of Ethnic and Migration Studies.

**Thjis Bol** is an associate professor in the Department of Sociology at the University of Amsterdam. His research interests are in social stratification and inequality in education, the labor market, and science. Recent publications include “School-to-Work Linkages, Educational Mismatches, and Labor Market Outcomes” (American Sociological Review, 2019). For the next five years he...
will work on CAREER, a project that studies the career effects of vocational and general education, funded by an ERC Starting Grant.

Herman van de Werfhorst started as chair of sociology at the European University Institute, Florence, in the summer of 2022, while on leave from the University of Amsterdam. His research focuses on inequalities in and through education, focusing on educational outcomes, labour market returns, and civic engagement. He combines comparative research with longitudinal data. Van de Werfhorst is co-principal investigator of the Transition from PRImary to Secondary school (PRIMS) study and the Adolescent Panel Study of Democratic Values and School Careers (Adolescentenpanel Democratische Kernwaarden en Schoolloopbanen, ADKS in Dutch).
Appendix A. Sample restrictions

### Table A1. Sample restrictions

<table>
<thead>
<tr>
<th>Primary education exit cohort</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
<th>N</th>
<th>%Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting sample</td>
<td>19,407</td>
<td>19,243</td>
<td>18,733</td>
<td>18,239</td>
<td>17,779</td>
<td>17,597</td>
<td>110,998</td>
<td></td>
</tr>
<tr>
<td>1. Drop if special education or praktijkonderwijs</td>
<td>18,501</td>
<td>18,330</td>
<td>17,863</td>
<td>17,412</td>
<td>17,018</td>
<td>16,850</td>
<td>105,974</td>
<td>-4.2%</td>
</tr>
<tr>
<td>2. Drop if missing or unclear track advice</td>
<td>18,215</td>
<td>18,186</td>
<td>17,789</td>
<td>17,278</td>
<td>16,942</td>
<td>16,729</td>
<td>105,139</td>
<td>-0.7%</td>
</tr>
<tr>
<td>3. Drop if missing school type</td>
<td>17,759</td>
<td>17,704</td>
<td>17,314</td>
<td>16,874</td>
<td>16,523</td>
<td>16,269</td>
<td>102,443</td>
<td>-2.7%</td>
</tr>
<tr>
<td>4. Drop if home/school address not registered</td>
<td>17,644</td>
<td>17,607</td>
<td>17,220</td>
<td>16,793</td>
<td>16,481</td>
<td>16,117</td>
<td>101,862</td>
<td>-0.9%</td>
</tr>
<tr>
<td>5. Drop if chosen school greater than 20 kilometres</td>
<td>17,404</td>
<td>17,366</td>
<td>16,961</td>
<td>16,564</td>
<td>16,224</td>
<td>15,853</td>
<td>100,372</td>
<td>-1.6%</td>
</tr>
<tr>
<td>6. Drop if chosen school does not fit track advice</td>
<td>16,180</td>
<td>16,111</td>
<td>15,851</td>
<td>15,690</td>
<td>15,376</td>
<td>15,013</td>
<td>94,221</td>
<td>-5.3%</td>
</tr>
<tr>
<td>7. Drop if chosen school outside distance range</td>
<td>14,719</td>
<td>14,658</td>
<td>14,453</td>
<td>14,289</td>
<td>14,042</td>
<td>13,713</td>
<td>85,874</td>
<td>-8.7%</td>
</tr>
<tr>
<td>8. Drop if school-level data are missing</td>
<td>13,608</td>
<td>13,883</td>
<td>13,847</td>
<td>13,623</td>
<td>13,275</td>
<td>13,122</td>
<td>81,358</td>
<td>-4.3%</td>
</tr>
<tr>
<td>9. Drop if parental education is missing</td>
<td>10,355</td>
<td>10,825</td>
<td>11,055</td>
<td>11,241</td>
<td>11,159</td>
<td>11,178</td>
<td>65,813</td>
<td>-14.8%</td>
</tr>
<tr>
<td>10. Drop if one choice alternative left</td>
<td>9,788</td>
<td>10,301</td>
<td>10,567</td>
<td>10,635</td>
<td>10,647</td>
<td>10,682</td>
<td>62,620</td>
<td>-4.4%</td>
</tr>
<tr>
<td>Final sample</td>
<td>9,788</td>
<td>10,301</td>
<td>10,567</td>
<td>10,635</td>
<td>10,647</td>
<td>10,682</td>
<td>62,620</td>
<td></td>
</tr>
</tbody>
</table>

*Source: Netherlands Cohort Study on Education (NCO).*