Neighbors' Effects and Early Track Choices

Manon Garrouste[*]   Camille Hémet[†]

March 2024

Preliminary version. Please do not circulate without permission.

Abstract

Choosing between vocational or academic education at the end of secondary education has important long-run effects, and is taken at an age where peers' influence might be paramount. In this paper, we investigate the effect of *neighbors*' track choices on 9th graders choices at the end of lower secondary education, in Paris. This question is central to understand the extent to which residential segregation can reinforce social segregation across vocational and academic tracks. We rely on neighbors from the preceding cohort in order to bypass the reflection problem, and use within-catchment-area variation in distance between pairs to account for residential sorting. We use a pair-wise model that enables us to study carefully the role of distance between neighbors, and to perform detailed heterogeneity analysis. Our results suggest that close neighbors do matter in track choices at the end of 9th grade, particularly for pupils pursuing a professional track. This effect is driven by neighbors living in the same building, and is larger for pairs of boys. Our results also suggest that neighbors' effects tend to accentuate social segregation across high school tracks.

Keywords: Peer effects, Education, Early track choice, Segregation, Fixed-effects models

JEL Codes: I21, C21

1 Introduction

Choosing an educational track, particularly between vocational and general training, holds significant importance in shaping future educational and labor market outcomes (Silliman and Virtanen, 2022). Importantly, track choice is socially marked, with a notable over-representation of students from lower socio-economic backgrounds in vocational training, and a higher concentration of pupils pursuing vocational tracks in poorer neighborhoods. This critical decision generally occurs during adolescence, a period where peer influence is paramount (Albert et al., 2013; Brown and Larson, 2009; Deutscher, 2020). If peers influence each other regarding orientation decisions, and in particular peers living in the same neighborhood, this type of social interaction could contribute to reinforcing social segregation across educational tracks, exacerbating existing disparities. Understanding whether individual orientation decisions are influenced by the choices of their neighbors is therefore a key question that can lead to important policy implications. Although the literature provides evidence that neighbors influence the pursuit of higher education (see, e.g. recent work by Barrios-Fernández, 2022), research on the role of close neighbors in determining early track choice is still scarce.

In this paper, we aim at understanding whether and to what extent interactions with other pupils living in the same neighborhood affect individual education decisions at the end of 9th grade. While it is established that social reproduction is an important dimension of educational track choices, especially in France (OECD, 2019), much less is known about the way in which these choices can be influenced by neighborhood peers' orientation decisions. Using administrative data on the universe of pupils in Paris in 2017, geolocalized at the adress level, we show that close neighbors do influence orientation decisions, this effect being driven by pupils enrolling in vocational training. Importantly, because we rely on a measure of distance between each pupil and any of her neighbors in a pair-wise model, we are able to precisely document the spatial scope of neighbors' effects. We find that this effect rapidly decays with distance, and is actually driven by neighbors living in the same building. To the best of our knowledge, this is the first

paper to precisely measure the spatial range of neighbors' effect, and to show that it is so local. This has important implications for the literature on neighbors' effects, where researchers might fail to identify such effect if the reference neighborhood is too large. Barrios-Fernández (2022) also provides convincing suggestive evidence that neighbors' effects are very local. However, while his estimates for various ranges of distance are obtained on different samples of students, we are able to estimate the precise effect of distance for each pupil in our sample.

Our study relies on a pair-wise model, where each ninth grader is paired with any of her neighbors, based on the Euclidian distance between each ninth grader and all other pupils from the preceding cohort (tenth graders) living in the surroundings. For computational reasons, we alternatively limit the set of neighbors to those living in the same catchment area, or to those living in a 1 kilometer radius around the reference pupil. Therefore, we measure the effect of distance between two neighbors on the likelihood that they follow the same track, and interpret a significant coefficient of distance as the existence of neihgbors' effect. This model has three main advantages. First, it enables us to precisely measure the effect of distance between peers, considering distance in a continuous way, while the literature typically looks at neighbors within a given range. Therefore, we can precisely document the spatial decay of neighborhood peers' effects, which is an important methodological contribution to the literature on neighborhood effects. Second, by relying on pairs of pupils, we can study the heterogeneity of peer effects with respect to peer characteristics. Finally, it enables us to abstract from the famous reflection problem pointed out by (Manski, 1993), as it arises in linear-in-means models. Nonetheless, for the sake of comparison with the rest of the literature, we also rely on a more standard linear-in-means model, where we estimate the effect of average neighbors' track choice on an individual choice, within various ranges of distance. In this case, we avoid the reflection problem by considering the effect of track choice from pupils of the previous cohort, whose decision making took place one year earlier and can thus not have

been influenced by the ninth graders' choices under study.

A fundamental methodological challenge for identifying peer effects comes from the fact that peers share common unobservable characteristics and environment which effects (called "correlated effects") are difficult to distinguish from social interaction effects (Manski, 1993). When looking at neighbors' effects, the crucial issue is endogenous neighborhood selection, i.e. the fact that individuals decide where to live based on unobserved preferences. In the French context, this problem is particularly salient: residential location largely determines which school will be attended, based on a catchment area system. It is therefore likely that socially and economically more advantaged families, who are more involved in their children's education and less constrained in their housing decisions, chose to reside in good quality schools' catchment areas. In turn, this means that the catchment area of residence, and consequently the set of neighbors, is not exogenous to schooling decisions and preferences. To bypass this issue, we rely on a fixed-effect strategy, whereby the identifying variation comes from comparing local neighbors *within* catchment areas, which arguably constitute a relevant spatial selection zone in the residential selection process in Paris. This method is similar in spirit to Bayer et al. (2008); Hémet and Malgouyres (2018); Solignac and Tô (2018), but is much more precise as we use individuals' exact geographic coordinates rather than the block or census tract where they reside.

Our results show that close neighbors do matter in track choices at the end of lower secondary education, after including catchment area fixed-effects accounting for residential selection across catchment areas. In particular, we find that this result is driven by *i)* pupils enrolling into vocational track after ninth grade, and by *ii)* pupils living at the exact same address. More precisely, we find that living in the same building than a one-year-older student enrolled in vocational training increases the likelihood that a ninth grader enrolls into vocational education after ninth grade by 2 percentage points. By contrast,

we do not find any significant effect of neighbors living further away. Our results thus show that neighbors' effects are extremely local in nature, and that they rapidly decay with distance. This raises the question of how to define a neighborhood, and reveals that neighborhood peer effects are likely to be diluted when considering relatively large neighborhoods. We check that these results are not driven by pupils living in social housing, which creates local clusters of low socio-economic status households, and that they hold once we exclude probable siblings from the sample. Our results are also robust to the inclusion of national test scores taken at the end of ninth grade, to account for ninth-graders ability.

Our pair-wise design crucially enables us to investigate the heterogeneity of neighbors' effects along various dimensions of the pairs' characteristics. That is, we do not only take a stand on how neigbhors' effect varies depending on neighbors' characteristics, but we are able to look at how this effect varies depending on the characteristics of the pair. In particular, we find that the close neighbor' effect is reinforced when both peers (the ninth grader and the one-year older neighbor) are from a low socio-economic background. Therefore, our results suggest that neighbors' effects may accentuate social segregation across tracks. In addition, looking at the heterogeneity of the effect with respect to gender, we find that close neighbors' effect is the strongest when the two pupils are boys.

Our paper contributes to the large and growing literature on peer effects in education. Several studies analyze the role of school peers on educational attainment (see for instance Epple and Romano, 2011; Sacerdote, 2011, for a review). Some of them use the US Add Health Survey to shed light on the role of peers on tests scores, drop out rates, or college enrollment (see Calvó-Armengol et al., 2009; Lin, 2010; Bifulco et al., 2011; Patacchini et al., 2017, among others). However, few studies look at the influence of *neighborhood* peers rather than *school* peers. In India, Helmers and Patnam (2014) show that a child's nearest neighbors' cognitive skills positively affect own cognitive achievement. Using American data, Del Bello et al. (2015) find that only peers at school influence academic

results and that peers in the neighborhood do not. Relying on French data, Goux and Maurin (2007) show that close neighbors' characteristics influence test scores and grade repetition, but their setting does not allow them to take a stand on orientation decision. Few papers directly look at track choice, and if so, they document the role of school peers' characteristics (in particular educational attainment and social background) rather than peers' choices *per se* (Battiston et al., 2020; Bifulco et al., 2011; Gibbons and Telhaj, 2016; Jonsson and Mood, 2008; Landaud et al., 2020; Ly and Riegert, 2014). Instead, we want to take a stand on how individual decision making is influenced by peers choices.

This question is addressed by De Giorgi et al. (2010) and Barrios-Fernández (2022), but only in higher education. De Giorgi et al. (2010) show that Italian students are more likely to choose a given college major when many of their peers make the same choice. Barrios-Fernández (2022) finds that the closest neighbor significantly influences individual decision to go to the university in Chile. While they focus on enrollment decision or major choice for undergraduate students or applicants, we will rather study orientation decision for students at the end of lower secondary education (at the end of ninth grade). This is a key timing in the education path, particularly in France, as students may decide, for the first time in the education process, to embark upon vocational training rather than academic track. The decision made at the end of ninth grade thus deeply influences future education and labor market outcomes. For that reason, it is crucial to understand how this decision is taken and how neighboring pupils may impact it.[1]

The remaining of this paper is organized as follows. First, we describe the French education context. Second, we present the data and some descriptive statistics. Then we explain the estimation strategy and display the estimation results. We then discuss the scope and consequences of the results. The last section concludes.

---

[1] Several closely related papers also look at this type of decision between vocational and academic training, but they focus on the role of information or aspirations rather than neighborhood peer effects *per se* (Fricke et al., 2018; Goux et al., 2017; Guyon and Huillery, 2020).

2 Institutional setting and data

2.1 Compulsory education and track choice in France

In France, education was compulsory from age 6 to age 16 up to 2019.[2] Children have a five-year primary schooling, then a four-year lower secondary education in middle schools, and a two to three-year upper secondary education in high schools. Primary and lower secondary education is the same for every one. But at the end of lower secondary education (in grade nine), pupils (usually aged 14 or 15) have to choose between differentiated school tracks. They can continue in higher secondary education, either through an academic track or through a vocational track. They can also choose to repeat ninth grade or they can leave school if they have reached the age of 16. The majority continue in an academic track, which prepares for an academic national examination ("baccalauréat") that students take at the end of high school. They can then continue in higher education. The vocational track is two-fold. A short (two-year) vocational track prepares for a professional certificate and a direct entry into the labor market. A long (three-year) vocational track prepares for a professional "baccalauréat", which gives access either to the labor market or to higher education.

The choice procedure starts at the beginning of the third quarter of ninth grade (in January). Each pupil and their family indicate a choice between an academic track, a short vocational track, or a long vocational track. At the end of the quarter, the teaching staff provides an answer in the form of a temporary proposition to each family. Before the end of the fourth quarter, pupils make a list of final track choices, as well as a choice of high school for each track. They can rank multiple choices, but they have a priority in the public high school of their catchment area. At the end of the academic year, the middle school provides a final recommendation in terms of track choice. If it matches the family's track choice, the pupil is allocated to that track. If not, the family meets the school headmaster. If they disagree on the track choice, the family can appeal against

---

[2]This is true at the time of the study. Since the 2019 school year, education has been compulsory from the age of 3.

the middle school's decision. Then an appeal board makes a final decision. In any case, pupils can always choose to repeat ninth grade.

Importantly, track choice is socially marked, with a notable over-representation of students from lower socio-economic backgrounds in vocational training, and a higher concentration of pupils pursuing vocational tracks in poorer neighborhoods. Table 1 displays the socio-economic characteristics of all 9th graders enrolled in France in 2016, and how they are distributed across track choices (observed in 2017). Strikingly, pupils from high socio-economic status background represent 26.5 % of all pupils enrolled in an academic track in 10th grade, but only 4.1 % and 6.5 % of pupils enrolled in short and long vocational tracks respectively. Conversely, pupils from a low socio-economic background represent 59.1 % of pupils embarking upon short vocational training. As a result, pupils receiving means-tested scholarships are over-represented in vocational education. We can also note that vocational training concentrates pupils from a foreign nationality.

Table 1 – Socioeconomic background of 9th grade graders by track choice in 10th grade

	Academic track	Vocational track	
		Short	Long
Socioeconomic status			
High SES	26.5	4.1	6.5
Medium SES	50.1	36.8	48.3
Low SES	23.4	59.1	45.2
Scholarship status			
No scholarship	83.4	67.8	69.7
Scholarship	16.6	32.2	30.3
Citizenship			
French	96.1	84.9	92.8
Not French	3.9	15.1	7.2

Source: MENJS DEPP, FAERE-Apprenants 2017. Sample: all 9th graders enrolled in France in 2016.

The decision to enroll in a vocational or academic education strongly influences mid to long-term education and labor market outcomes. For instance, after seven years, the employment rate of young people who left the education system in 2010 with a short vocational certificate was 73% as opposed to 84% for those who obtained a professional "baccalauréat", and 92% for those who completed a master's degree. The median monthly

wage was about 1 500 euros for young people with a vocational degree, 1 600 euros for those having an academic "baccalauréat", and 2 200 euros for those having a master's degree (DEPP, 2020).

2.2 Allocation to schools

Allocation to middle schools is based on a catchment area system. Each pupil has a priority in the public middle school of the catchment area, which depends on his or her home address. Allocation to another public school is granted by the local authority (*académie*) only if there are some remaining seats once all pupils from the catchment area are served. Allocation to another public school is granted on the basis of (in order of priority) medical reasons, scholarship, brother or sister already in the school, distance, specific option only available in the school. Furthermore, pupils can also choose to enroll at a private school, as private schools are not subject to the catchment area system. Due to this system, most of the middle school choice is actually made through location decision (Boutchenik et al., 2020; Monarrez, 2023). When deciding where to live, some families carefully choose the corresponding catchment area school. In particular, we expect socially and economically more advantaged families to locate in good quality schools' catchment areas, meaning that location in catchment area, and therefore the set of neighbors is not exogenous to schooling decisions and preferences.

Similarly, allocation to high schools is based on a catchment area system and on a centralized allocation mechanism at the local authority level. Each pupil has a priority in the public high school of the catchment area. Depending on the local authority, past academic achievement, favorable view from headmaster of middle school, or from headmaster of high school are also considered. For high schools providing vocational tracks, the catchment area is typically larger than for high schools providing academic tracks and often corresponds to the entire local authority area.

As some school choice is possible, it is important to note that not all students are enrolled at the school of their catchment area. In particular, when entering middle school

about 10% of students go to a public school different from their catchment area middle school and about 20% go to a private school (Boutchenik et al., 2020). This setting may create an additional selection process: on top of endogenous location decision, school choice is likely to create supplementary selection. As for location decisions, we expect socially and economically more advantaged families to choose good quality schools, potentially outside of their catchment areas. To account for this, we will carefully control for students' school choices in the estimations.

2.3 Data and descriptive statistics

To analyze the effect of neighbors on individual schooling decisions, we rely on geocoded administrative data on the universe of pupils enrolled in secondary education in France in 2017, provided by the statistical service of the French ministry of education. For each pupil, we observe the school and grade they are enrolled at in years t and $t-1$. We focus on two samples. First, we consider pupils enrolled in the last year of lower secondary education (in ninth grade) in 2017 and we observe the grade and track they follow in year $t+1$. Second, we consider the sample of pupils enrolled in the first year of upper secondary education (in tenth grade) in 2017 and we observe the track they follow in year t. For each pupil, we know the exact location of the home address, given by the geographic coordinates. We also observe each pupil's sex, country of birth, parents' occupations, and whether or not the pupil receives a public income-based scholarship.

In addition, we need to assign each pupil to the catchment area – and thus the public middle school – corresponding to her home address. Unfortunately, there is no nationwide GIS documenting the borders of each catchment area in France. In this study, we focus on the municipality of Paris, for which we could obtain the geocoded data of every public middle school's catchment area. We therefore restrict our sample to pupils who live in Paris and are enrolled in ninth grade in 2016 or in 2017, for whom we can thus identify the public middle school's catchment area.[3]

---

[3]The catchment area is defined as the one that prevailed at the time ninth graders entered middle school.

In total we observe 17,708 ninth graders in 2017, and 19,614 tenth graders in 2017 living in 111 middle schools' catchment areas in Paris. On average, Parisian pupils are socially more advantaged than pupils living in the rest of France. About 45% come from a high socioeconomic background in the sense than their reference parent is an executive or a highly qualified worker, 34% come from a medium socioeconomic backgrounds, and 21% from low socioeconomic backgrounds (with parents being blue-collars or unemployed). Pupils benefiting from an income-based scholarship represent about 20% of the sample. Almost 90% of children are born French citizens. About 11% of pupils live in a deprived neighborhood in the sense that it is part of the urban policy scheme.[4]

Although most pupils were enrolled at a public middle school in ninth grade (43% at their catchment-area school and 28% at another public school), a large proportion (29%) were in a private school. After ninth grade, about 82% go to an academic track and 15% to a vocational track. Among those who continue in a vocational track, the vast majority follow a long three-year track (see Tables 2 for descriptive statistics of the students' sample).

---

[4]These neighborhoods are part of the *Quartiers prioritaires de la politique de la ville*. They are defined as economically disadvantaged neighborhoods, based on the residents' median revenue.

Table 2 – Socioeconomic characteristics of pupils in the sample

	9th graders		10th graders	
Socioeconomic status				
High SES	7,987	45.1	8,344	42.5
Medium SES	6,087	34.4	6,768	34.5
Low SES	3,628	20.5	4,502	23.0
Gender				
Female	8,975	50.7	9,617	49.0
Male	8,733	49.3	9,997	51.0
Scholarship status				
No scholarship	14,153	79.9	16,429	83.8
Scholarship	3,555	20.1	3,185	16.2
Citizenship				
French	15,765	89.0	17,545	89.5
Not French	1,943	11.0	2,069	10.5
Middle school				
Public	12,658	71.5	14,198	72.4
Private	5,050	28.5	5,416	27.6
Track choice				
Repetition	494	2.8		
General track	14,645	82.7	15,517	79.1
Vocational track	2,569	14.5	4,097	20.9
Total	17,708	100.0	19,614	100.0

Source: MENJS DEPP, FAERE-Apprenants 2017

Note: High SES = children of executives or highly qualified workers, Medium SES = children of workers with intermediate occupations, or white-collar workers, Low SES = children of blue-collars or of the unemployed. Repetition refers to repeating ninth grade and is therefore not observed for the 10th graders cohort.

3 Empirical model

3.1 Conceptual framework

To bypass the issue of endogenous location, we use the within-catchment-area variation in distance between neighbors. To do that, we construct all pairs of pupils living in the same catchment area. Every ninth grader is paired with all peers of the preceding cohort living in the same catchment area.

Distance to neighbors is then defined as the Euclidean distance to all 10th graders living in the same catchment area. Figure 1 represents middle schools catchment areas in the municipality of Paris with a closer look on the 14th *arrondissement* of Paris. Each catchment area corresponds to one public middle school (black squares). The municipality of Paris includes 111 catchment areas, each of them containing about 159 ninth graders on average. Blue dots represent a (fictive) sample of ninth graders in each catchment area, grey diamonds represent a (fictive) sample of tenth graders, and Figure 2 represents the two-by-two distance between every ninth grader and every tenth grader in a given catchment area. Note that this corresponds to creating a perfectly dense network of neighbors within each catchment area.
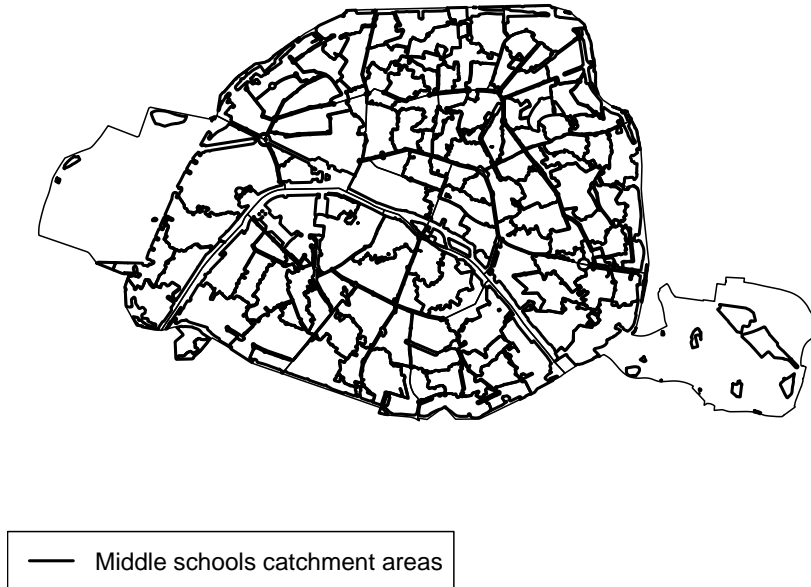
Although some families carefully choose where to live with respect to the quality of the catchment area middle school, we can make the assumption that, within a catchment area, they do not choose a particular location with respect to the location of other tenth graders. In the following, we will thus use the fact that within-catchment-area variation in two-by-two distance between neighbors is exogenous to unobserved schooling preferences.[5]

We observe more than 3,200,000 pairs of pupils living in the same catchment area (see Table 3 for descriptive statistics of the pairs' sample). Within a catchment area, the average distance between a ninth grader and a tenth grader is about 500 meters (see Figure 3). Note that we censured distance to be less than 1 kilometer for computational
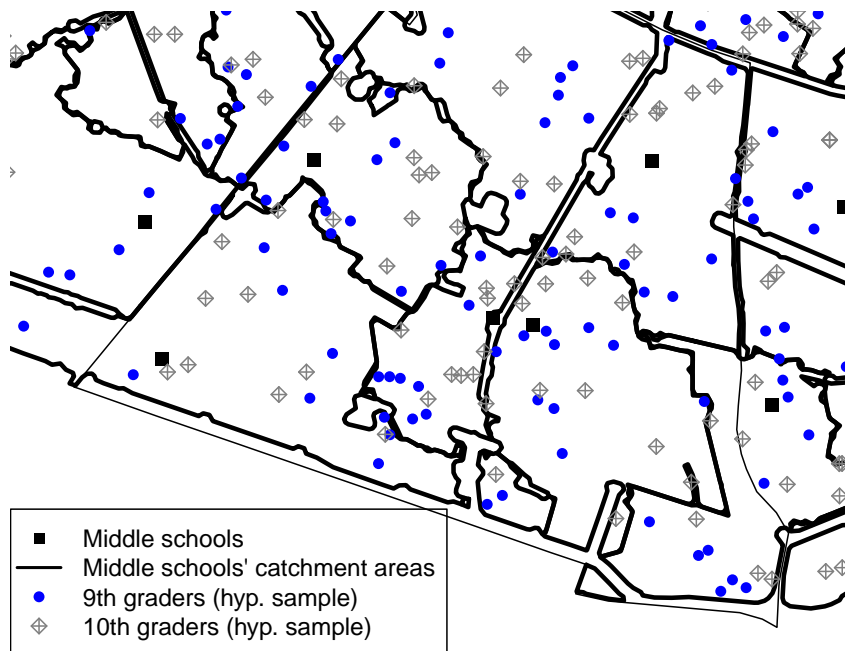
---

[5]This assumption will be discussed and tested in Section 3.2.2
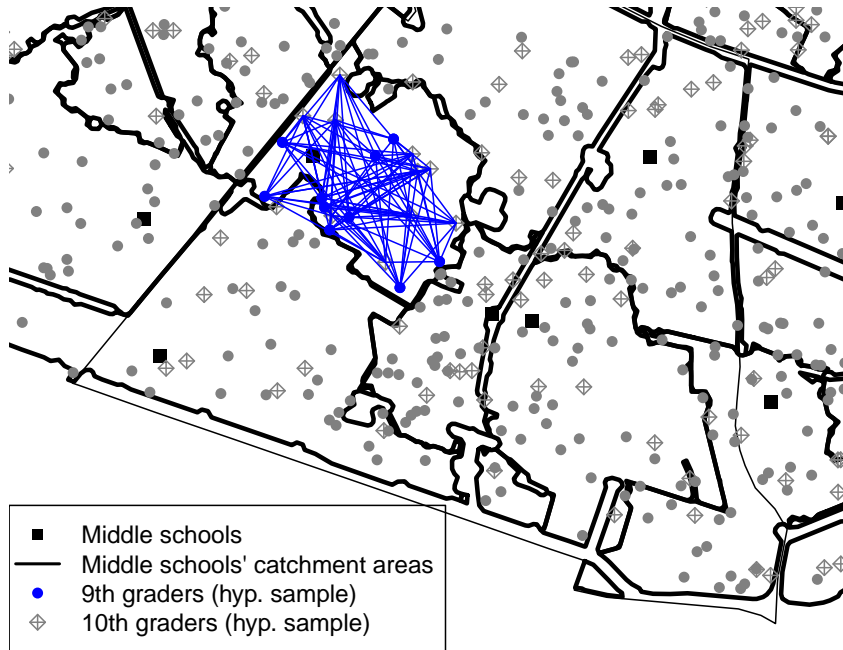
Figure 1 – Middle schools catchment areas

(a) in Paris

| — | Middle schools catchment areas |

(b) in the 14th *arrondissement* of Paris

■	Middle schools
—	Middle schools' catchment areas
●	9th graders (hyp. sample)
⊕	10th graders (hyp. sample)

Source: Insee, IGN, Direction des Affaires Scolaires - Ville de Paris

Note: Thick lines represent the boundaries of middle schools catchment areas in the municipality of Paris. Black squares represent middle schools. Blue dots represent a (fictive) sample of ninth graders. Grey diamonds represent a (fictive) sample of tenth graders.

Figure 2 – Middle schools catchment areas

Note: Thick lines represent the boundaries of middle schools catchment areas in the municipality of Paris. Black squares represent middle schools. Blue dots represent a (fictive) sample of ninth graders. Grey diamonds represent a (fictive) sample of tenth graders. Blue lines represent distances between ninth graders and tenth graders within one catchment area.

reasons and because almost 90% of pairs live less than one kilometer away from each other.

Table 3 – Track choices of pairs of ninth graders - tenth graders

	Percentage
Repetition-General	2.2
Repetition-Vocational	0.6
General-General	67.7
General-Vocational	15.7
Vocational-General	10.4
Vocational-Vocational	3.4
Observations	3272473

Note: The unit of observation is a pair of individuals living in the same catchment area.

Figure 3 – Distribution of distance between neighbors
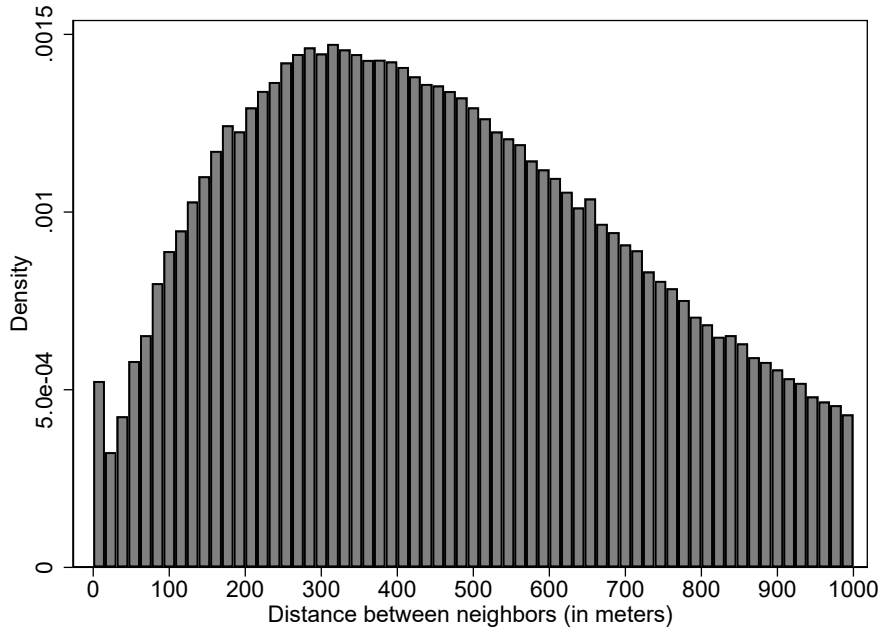
3.2 Estimation strategy

For any pair of ninth grader i and tenth grader j living in the same catchment area C, let us denote Y_{ij} the dummy variable which equals one if i follows the same track as j after ninth grade and zero otherwise. Let us consider $dist_{ij}$ the distance (in meters) between i's and j's places of residence. Let us consider the following equation:

$$Y_{ij} = \beta dist_{ij} + \zeta dist_{ij}^2 + X_i'\gamma_1 + X_j'\gamma_2 + X_{ij}'\gamma_3 + \mu_C + \varepsilon_{ij} \qquad (1)$$

with X_i, X_j, and X_{ij} a set of individual characteristics; gender and socio-economic background (parents' occupations, scholarship status, and born French or not), as well as the pair's school choice (same school or not and type of school, i.e. catchment area public school, another public school, or a private school).

3.2.1 Mitigating residence- and school-based sorting: a fixed-effect approach

As families choose the catchment area where to live, $dist_{ij}$ is supposed to be endogenous in general. For instance, if socially more advantaged families choose to live in better quality schools' catchment areas, then we expect $dist_{ij}$ to be negatively correlated with the error term (distant neighbors' unobserved preferences regarding education choices are less similar than close neighbors'). A naive estimation would then overestimate the effect of close neighbors' track choices. The estimation strategy consists of including catchment area fixed effects μ_C in the equation. Under the assumption that families do not choose exactly where to live in a particular catchment area, then the within-catchment area variation in distance can be used to recover a causal effect of neighbors' choices on own choice. In other words, we are assuming that the distribution of students within catchment areas is as good as random once we account for their choice of a catchment area. To test this assumption, we will check that there is no spatial auto-correlation within catchment areas regarding students' observed characteristics (see Section 3.2.2).

This strategy is similar to the one used in the seminal paper by Bayer et al. (2008) or more recently by Hémet and Malgouyres (2018) and Solignac and Tô (2018) for France, but uses students' exact geographic coordinates instead of individuals' census tracts. We are thus able to analyze the effect of distance to neighbors very precisely rather than using a proxy measured by living within the same census tract. As we observe the entire distribution of two-by-two distances, we can analyze the distance at which neighbor effects fade away. To test whether only very close neighbors matter, let us consider the following model where distance to neighbors is split into concentric rings of 100 meters around every student (see Figure 4). For all students i and j such that $100 \times k \leq dist_{ij} < 100 \times (k+2)$, let us consider the model:

$$Y_{ij} = \beta_k \mathbb{1}_{100 \times k \leq dist_{ij} < 100 \times (k+1)} + X_i' \gamma_1 + X_j' \gamma_2 + X_{ij}' \gamma_3 + \mu_C + \varepsilon_{ij} \qquad (2)$$

$\forall k = 0, \ldots 5.$

17

The parameters of interest, β_k, measure the additional probability of choosing the same track as a close neighbor rather than someone living further away, controlling for the selection of the catchment area. For instance, β_0 corresponds to the effect on a pair's track choice of living less than 100 meters away rather than between 100 and 200 meters away, within a given catchment area.

The magnitude of these parameters therefore provides a test for the importance of residence-based peer influence in determining the type of track that students chose. Obviously, there might be several channels explaining such neighborhood peers' effect. The most direct one would be that pupils are (more) influenced by their neighbors (than by their non-neighbors) because they interact more frequently, for instance going home together, or hanging out together more frequently during evenings and weekends (remember that whether or not they were enrolled at the same middle school is included in the model). Another more indirect channel would be through parents: it could be that parents who are close neighbors interact more frequently with each other than parents who live further away, and have therefore more opportunities to exchange information on track choices and discuss the plans they have in mind for their children, thus influencing each other, and eventually affecting student's actual choice.[6] Given the data at hand, disentangling the precise mechanisms through which neighbors affect students track choice is out of the scope of this paper. However, the heterogeneity analysis can give us some insights. For instance, if same-sex peers interact more frequently or easily, then one could expect larger neighbor effects for pairs of girls or for pairs of boys than for opposite-sex pairs.

3.2.2 Testing for the no-spatial-autocorrelation assumption in the *within* estimations

In order to test the identifying assumption of equations (1) and (2), we conduct a formal statistical test consisting in verifying that neighbors are not distributed within catchment areas in a way that would create spatial auto-correlation with respect to certain observed

---

[6]Note that these two channels are not mutually exclusive, and can even be complementary.

Figure 4 – Effect of distance to neighbors (Equation 2)

Middle schools
Middle schools' catchment areas
9th graders (hyp. sample)
10th graders (hyp. sample)
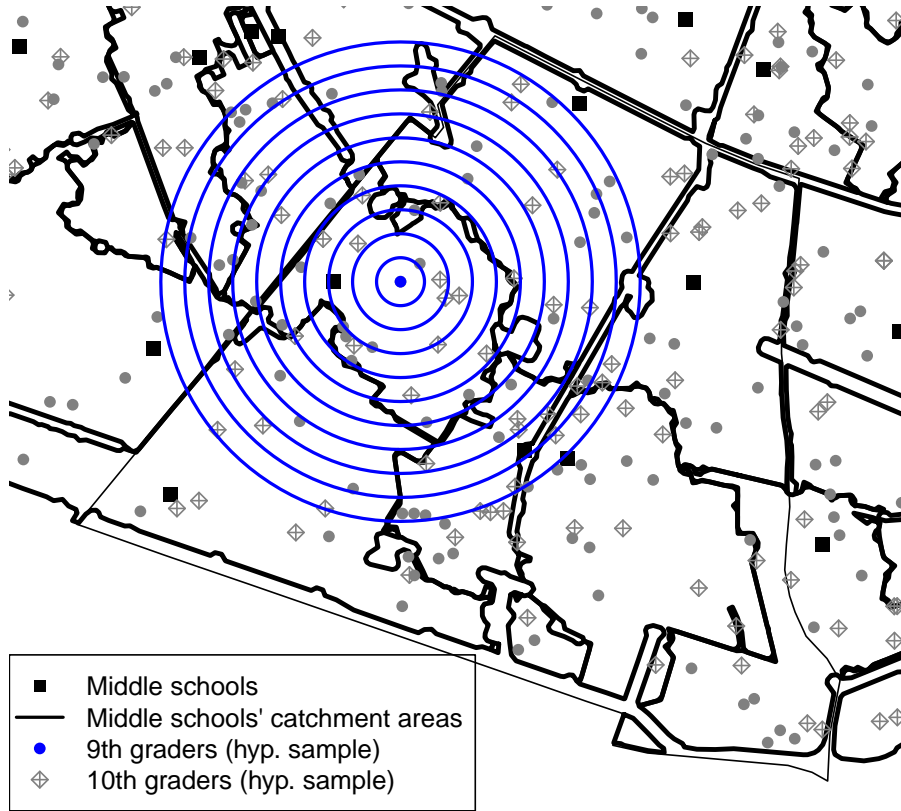
Notes: For each student i at the center of the rings, Equation (2) considers sequential sub-samples of all neighbors living at less than 100 meters , more than 100 meters but less than 200 meters, etc. Coefficient β_0 corresponds to the effect of living at distance zero rather than less than 100 meters away. Coefficient β_1 corresponds to the effect of living less than 100 meters away rather than between 100 and 200 meters away, etc.

characteristics. In particular, even within catchment areas, we could expect individuals to be clustered on the basis of their social background. For instance, this would be the case if residential segregation (particularly high in the Paris municipality) makes that wealthy and poor families live in separate neighborhoods, or if social housing makes that some low socioeconomic status families live close to each other. To test for this, we run join count tests, which consist in comparing the theoretical probability that two students share the same characteristic, under the assumption of absence of spatial auto-correlation, with the observed probability. This test corresponds to assessing whether pairs with same characteristic (for instance both coming from the same social background) are closer to

each other than what would be expected by chance, taking into account the proportion of students from each social background in the catchment area.[7]

The results show that the null hypothesis of absence of spatial auto-correlation related to the pairs' social background is rejected in 30 catchment areas out of 111 (this represents about 28% of the sample of students' pairs) with a 5% significance level. Note that some spatial auto-correlation due to randomness is of course likely to happen (in about 5% of cases). For sake of comparison, testing for spatial auto-correlation with respect to the pairs' sex composition (dimension in which there cannot be spatial sorting) leads to reject the null of absence of spatial auto-correlation in 5 catchment areas out of 111.

In the following, we will first estimate the model on the entire sample, then we will exclude these 30 catchment areas to make sure that the estimation results are not driven by the potential clustering of families within catchment areas based on socioeconomic characteristics.

4   Results

4.1   Estimation results

Table 4 presents the estimation results corresponding to estimating equation (1) by ordinary least squares with catchment area fixed effects including all Parisian catchment areas. The first column corresponds to the effect of the distance between neighbors (measured in 100 meters) on the probability of choosing the same track after ninth grade, once controlled for the choice of catchment area. The remaining columns give the results of estimating the equation separately for each type of track (repetition, academic, or vocational).

Interestingly, the only significant effect of distance is on the probability that i and

---

[7]Social backgrounds are defined as parents' occupation status, defined by the statistical service of the French Ministry of education as a four-category variable. Very high socio-economic status students are children of executives, High SES students are children of parents with intermediate professions, Medium SES students are children of white-collars, and Low SES students are children of blue collars or the unemployed.

j both choose the vocational track. The results suggest that a one hundred meter increase in distance between neighbors corresponds to a 0.2 percentage point increase in the probability that i and j choose the vocational track. Compared to estimating the same model without catchment-area fixed effects, the coefficient is a bit smaller (see Table A1 in Appendix A), suggesting that the effect of distance between neighbors would be over-estimated without controlling for sorting across catchment areas.

Our estimation strategy relies on the assumption that the distribution of students within catchment areas is as good as random once we account for their choice of a catchment area. This assumption may not hold if households are spatially clustered on the basis of characteristics that are correlated to their unobserved preferences regarding schooling trajectories. In this case, one would wrongly attribute to neighbors' influence something that is actually due to correlated effects.

Residential segregation, poverty clusters, or social housing may be reasons why this assumption would be violated, especially in the Paris municipality. To make sure that the estimation results are not driven by the residential sorting of families within catchment areas, we exclude from the analysis the 30 catchment areas for which the no-spatial-sorting test does not hold.

Table 5 shows that the results are not due to spatial sorting on the basis of the social composition of the pairs as the results are robust to the exclusion of the catchment areas that do not pass the spatial-auto-correlation test. The estimate of the effect of distance on the probability that i and j choose the vocational track is smaller (0.16 percentage point) but still very significant. The coefficient corresponding to the square of distance suggests that neighbor effects are likely non linear in distance between students.[8]

Order of magnitude of the results. Although statistically significant, the estimated effects of neighbors on track choices may seem economically small. Moreover, the results are not easily comparable to other studies on neighbor effects on education outcomes

_____

[8]The effect of distance between students fades away or changes sign after $\frac{0.0016}{2\times -0.0001} = 8$ hundred meters.

(Goux and Maurin, 2007; Del Bello et al., 2015), as they use measures of schooling performance and not schooling decisions and as they use linear-in-means models at the individual level, whereas we use models at the pair level. We can nevertheless compare neighbor effects with the size of coefficients of other covariates in the model, for instance students' parents' occupation. Of course, we do expect neighborhood peer effects to be much smaller than the effect of own social background. For the sake of comparison, the 2 percentage point increase in the probability to follow a vocational track after ninth grade when the same-building older neighbor follows a vocational track is about ten times smaller than the increase in the same probability of coming from a high socioeconomic background.

Table 4 – Neighborhood peer effects on track choices at the end of ninth grade - Equation (1) - All catchment areas

	i's track-j's track						
	Same track	Rep^ion-Gen^al	Rep^ion-Voc^al	Gen^al-Gen^al	Gen^al-Voc^al	Voc^al-Gen^al	Voc^al-Voc^al
Dist. btw neighbors (100 meters)	-0.0005	0.0006*	-0.0001	0.0015	-0.0002	0.0001	-0.0020***
	(0.0010)	(0.0003)	(0.0001)	(0.0012)	(0.0010)	(0.0007)	(0.0005)
Dist.^{2}	0.0000	-0.0000	0.0000	-0.0001	-0.0000	0.0000	0.0001***
	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Intercept	0.7422***	0.0144***	0.0031***	0.7069***	0.1258***	0.1146***	0.0353***
	(0.0062)	(0.0022)	(0.0008)	(0.0082)	(0.0053)	(0.0053)	(0.0028)
Nbr obs	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014
Nbr clusters	111	111	111	111	111	111	111

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a pair of individuals living in the same catchment area. Standard errors in parenthesis are clustered at the catchment area level. All estimations include catchment area fixed effects, individual characteristics, and characteristics of the pair (gender, parents' occupations, scholarship status, born French or not, and middle school choice).

Table 5 – Neighborhood peer effects on track choices at the end of ninth grade - Equation (1) - Excluding the catchment areas that do not pass the no spatial auto-correlation test

	i's track-j's track						
	Same track	Rep^{ion}-Gen^{al}	Rep^{ion}-Voc^{al}	Gen^{al}-Gen^{al}	Gen^{al}-Voc^{al}	Voc^{al}-Gen^{al}	Voc^{al}-Voc^{al}
Dist. btw neighbors (100 meters)	-0.0003	0.0004	-0.0001	0.0013	-0.0006	0.0007	-0.0016***
	(0.0011)	(0.0004)	(0.0002)	(0.0014)	(0.0012)	(0.0008)	(0.0005)
Dist.^{2}	0.0000	-0.0000	0.0000	-0.0001	0.0000	-0.0000	0.0001*
	(0.0001)	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Intercept	0.7465***	0.0175***	0.0041***	0.7139***	0.1252***	0.1069***	0.0326***
	(0.0068)	(0.0024)	(0.0008)	(0.0090)	(0.0064)	(0.0064)	(0.0034)
Nbr obs	2,318,373	2,318,373	2,318,373	2,318,373	2,318,373	2,318,373	2,318,373
Nbr clusters	81	81	81	81	81	81	81

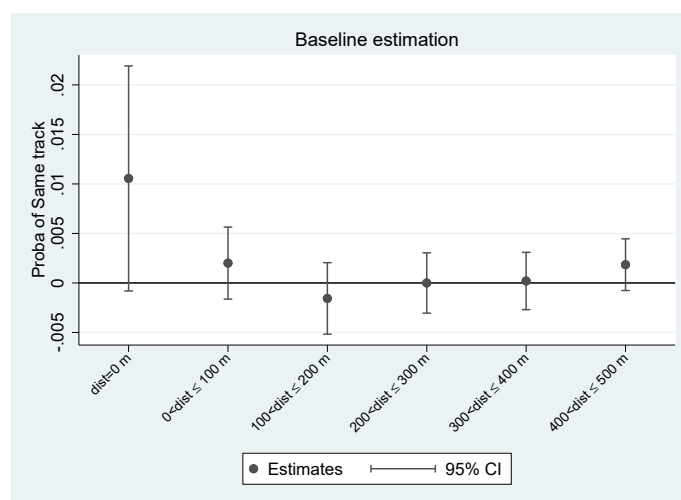*Source:* MENJS DEPP, FAERE-Apprenants 2017

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a pair of individuals living in the same catchment area. Standard errors in parenthesis are clustered at the catchment area level. All estimations include catchment area fixed effects, individual characteristics, and characteristics of the pair (gender, parents' occupations, scholarship status, born French or not, and middle school choice).

Considering linear or quadratic effects of distance between neighbors may be inaccurate as neighborhood peer effects are presumably not linear and mixing very close and very distant neighbors is likely misleading.

To further understand neighborhood peer effects on track choices, let us present the estimation results corresponding to estimating Equation 2. Figure 5 plots the coefficients corresponding to estimating the β_k parameters. When looking at more comparable neighbors, the results suggest that older neighbors have no effect on track choice when we consider all tracks. However, very close neighbors seem to influence younger neighbors when they follow a vocational track. Figure 5 (b) illustrates that, compared to pairs of neighbors living between 0 and 100 meters away, pairs living at distance zero, i.e. in the same building or block of buildings, have a 2 percentage point lower probability that i follows an academic track after ninth grade when j is following a vocational track. Symmetrically, compared to pairs of neighbors living between 0 and 100 meters away, pairs living at distance zero have a 2 percentage point higher probability that i follows a vocational track after ninth grade when j is following a vocational track. Older neighbors living further away do not influence their younger neighbors.

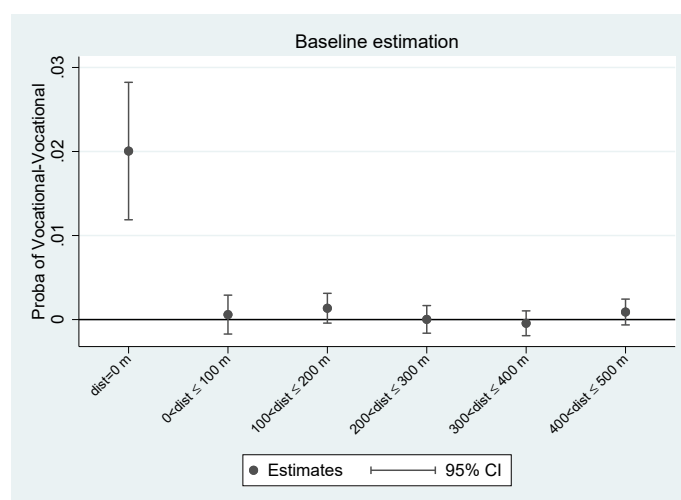Figure 5 – Neighborhood peer effects on track choices at the end of ninth grade - Equation (2)

(a) All tracks

(b) Academic track
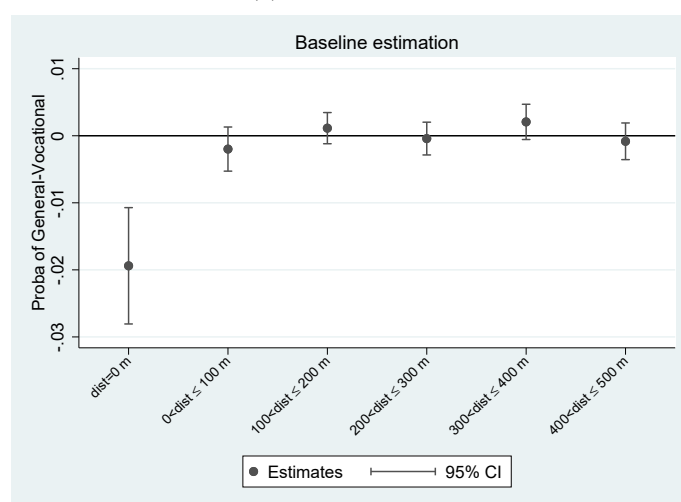
(c) Vocational track

Source: MENJS DEPP, FAERE-Apprenants 2017

4.2 Heterogeneity analysis

The results presented so far measure average effects. However, we may expect neighbor effects to be different according to the pairs' characteristics. For instance, the influence of neighbors may be different between pairs of boys and pairs of girls. In particular, we expect the effect of distance to be larger for same-sex pairs. We may also hypothesize that neighbor effects vary with the pairs' social backgrounds and expect that the effect of distance is higher for same-SES pairs. To test these assumptions, let us pre-multiply the model with the different characteristics of the pairs (gender, parents' socio-economic status, scholarship status, country of birth).

Table 6 presents the results with differentiated effects according to the pairs' gender composition, with pairs of girls as the reference group. It shows that the effect of distance is not significantly different from zero for pairs of girls (first row), nor for opposite-sex pairs (second and third row). Interestingly, the effect of distance is entirely driven by pairs of boys who follow a vocational track. Every one hundred meter decrease in distance between neighbors increases the probability that i and j follow a vocational track by 0.38 percentage point more for pairs of boys than for pairs of girls (the reference). Importantly, this heterogeneity with respect to the sex composition of pairs suggests that the effect of distance captures genuine interactions between students and is not only due to unobserved characteristics or family ties for instance.

Table 7 gives the results with differentiated effects according to the pairs' socioeconomic backgrounds.[9] Compared with pairs of high-SES students (the reference), pairs with a low-SES younger student i and a high-SES older student j have a 0.35 percentage point larger probability to both go to a vocational track after ninth grade with a one hundred meter decrease in distance. Remarkably, the effect of distance is larger when only the older student j comes from a low socioeconomic background (third row), as compared to same-high-SES pairs. In that case, every one hundred decrease in distance increases the

---

[9]To simplify, let us consider two categories: low-SES pupils are defined as children of blue-collar workers or of the unemployed; high-SES pupils are defined as children of executives or intermediate occupations.

probability that i follows the same track as j by 1.1 percentage point more for high-SES-low-SES students than for same-(high)-SES pairs. This is due to a larger effect of distance on the probability that they both choose the academic track. At the same time, there is a smaller effect of distance on the probability that they choose different tracks (academic and vocational, with a 0.85 percentage point smaller effect with every one hundred meter decrease, or vocational and academic, with a 0.38 percentage point smaller effect with every one hundred meter decrease). Compared with pairs of high-SES pupils, pairs of two low-SES pupils have a 0.75 percentage point larger probability to both go to a vocational track after ninth grade with a one hundred meter decrease in distance. They also have a 0.19 percentage point smaller probability that i repeats ninth grade when j follows a vocational track, and a marginally significant 0.81 percentage point smaller probability to both go to an academic track when distance decreases by one hundred meters.

There is no significantly different neighbor effects either with respect to the pair's scholarship status or country of birth.

To summarize the results, one-year-older neighbors matter in individual track choices, but only when they live in the same building or block of buildings, and only when they follow a vocational track. Furthermore, the effect is driven by pairs of boys, and is larger for low-SES students.

Table 6 – Neighborhood peer effects on track choices at the end of ninth grade - Equation (1) - By gender

	Same track	Rep$^{\text{ion}}$-Gen$^{\text{al}}$	Rep$^{\text{ion}}$-Voc$^{\text{al}}$	Gen$^{\text{al}}$-Gen$^{\text{al}}$	Gen$^{\text{al}}$-Voc$^{\text{al}}$	Voc$^{\text{al}}$-Gen$^{\text{al}}$	Voc$^{\text{al}}$-Voc$^{\text{al}}$
				i's track-j's track			
Dist. (100 meters)	0.0020	0.0009	-0.0001	0.0022	-0.0032*	0.0005	-0.0003
	(0.0021)	(0.0007)	(0.0002)	(0.0021)	(0.0019)	(0.0014)	(0.0006)
i Male $\times$ Dist.	-0.0011	-0.0009	0.0000	-0.0025	0.0027	-0.0007	0.0014
	(0.0027)	(0.0011)	(0.0003)	(0.0028)	(0.0018)	(0.0024)	(0.0009)
j Male $\times$ Dist.	-0.0042	-0.0003	-0.0001	-0.0045	0.0040	0.0006	0.0003
	(0.0031)	(0.0004)	(0.0002)	(0.0029)	(0.0032)	(0.0014)	(0.0010)
Male $\times$ Male $\times$ Dist.	-0.0007	0.0009	0.0000	0.0031	-0.0036	0.0034	-0.0038**
	(0.0036)	(0.0007)	(0.0004)	(0.0028)	(0.0023)	(0.0023)	(0.0018)
Intercept	0.1721***	0.1626***	0.0356***	-0.0564***	-0.1524***	0.7821***	0.2285***
	(0.0244)	(0.0155)	(0.0035)	(0.0202)	(0.0237)	(0.0257)	(0.0196)
Nbr obs	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653
Nbr clusters	81	81	81	81	81	81	81

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a pair of individuals living in the same catchment area. Standard errors in parenthesis are clustered at the catchment area level. All estimations include catchment area fixed effects and characteristics of the pair (dummies for each combination of genders, parents' occupations, scholarship status, and born French or not).

Table 7 – Neighborhood peer effects on track choices at the end of ninth grade - Equation (1) - By socioeconomic status (Low SES)

	i's track-j's track						
	Same track	Rep^{ion}-Gen^{al}	Rep^{ion}-Voc^{al}	Gen^{al}-Gen^{al}	Gen^{al}-Voc^{al}	Voc^{al}-Gen^{al}	Voc^{al}-Voc^{al}
Dist. (100 meters)	0.0025	0.0007	-0.0001	0.0018	-0.0033*	0.0002	0.0007*
	(0.0019)	(0.0006)	(0.0001)	(0.0020)	(0.0019)	(0.0009)	(0.0004)
i Low SES $\times$ Dist.	-0.0019	-0.0009	0.0001	0.0016	-0.0006	0.0033	-0.0035**
	(0.0029)	(0.0013)	(0.0004)	(0.0028)	(0.0019)	(0.0032)	(0.0015)
j Low SES $\times$ Dist.	-0.0110***	-0.0003	-0.0009**	-0.0119***	0.0085**	0.0038***	0.0009
	(0.0036)	(0.0005)	(0.0004)	(0.0039)	(0.0039)	(0.0011)	(0.0010)
Low SES $\times$ Low SES $\times$ Dist.	0.0006	0.0002	0.0019**	0.0081*	0.0026	-0.0052	-0.0075**
	(0.0044)	(0.0013)	(0.0009)	(0.0041)	(0.0044)	(0.0035)	(0.0030)
Intercept	0.2368***	0.1643***	0.0338***	0.0194	-0.2294***	0.7945***	0.2173***
	(0.0241)	(0.0161)	(0.0033)	(0.0191)	(0.0237)	(0.0266)	(0.0191)
Nbr obs	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653	1,033,653
Nbr clusters	81	81	81	81	81	81	81

Source: MENJS DEPP, FAERE-Apprenants 2017

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a pair of individuals living in the same catchment area. Standard errors in parenthesis are clustered at the catchment area level. All estimations include catchment area fixed effects and characteristics of the pair (dummies for each combination of genders, parents' occupations, scholarship status, and born French or not).

Consequences in terms of social segregation across tracks The choice between an academic track and a vocational track at the end of ninth grade is an important stage, as they lead to very different educational paths and labor market opportunities. In France, the academic track is perceived as the most academically prestigious, whereas the vocational track is often considered a low-status track or a default option for pupils who do not perform well enough to follow the academic track. If neighbor effects tend to increase the probability to follow the vocational track for pairs of low-SES pupils and, on the contrary, to increase the probability to follow the academic track for pairs of high-SES pupils, then it suggests that neighbor effects tend to accentuate social segregation across high school tracks.

5 Robustness checks

5.1 Controlling for academic achievement

At the end of middle school, students take a standardized national exam called *Brevet*, which consists of a written exam in four subjects (French, Mathematics, History-Geography, and Sciences), and an oral presentation. The total *Brevet* score also includes students' overall assessment in all subjects as evaluated by their teachers during the academic year.

Passing the *Brevet* exam is not a prerequisite for continuing in higher secondary education. However, almost all students take it and their score is positively correlated to their probability to follow an academic track. In order to check that our neighbor effect estimates are not driven by students academic achievement, we include ninth graders' *Brevet* exam scores as a covariate in the estimation.[10] Figure 6 compares the preceding estimation results (baseline estimates) with the estimation including ninth graders' *Brevet* exam scores as an additional covariate (second estimate in the graph). The estimated coefficient corresponding to the effect of very close neighbors (living in the same building) is smaller but still significantly different from zero and not significantly different from the preceding

---

[10]The *Brevet* exam score is unfortunately not available for the 10th graders cohort.

estimate.

5.2 Social housing

Some very close neighbors, living in the same building or block of buildings, may very well be living in large social housing estates. These students are likely to share unobserved characteristics that could create the estimated correlation between tracks. To exclude any effects driven by social housing, we matched our data with the national geocoded dataset on social housing. Every student's place of residence is matched with the national registry of social housing (*Répertoire des logements locatifs des bailleurs sociaux*) to determine the ones who live in such type of housing. Whether i, j, or both of them live in social housing is then included as additional covariates in the estimation (i.e. on top of the *Brevet* exam score).

Figure 6 shows that our results are robust to controlling for social housing as the estimate (third estimate in the graph), although smaller, is still significant and not significantly different from the preceding estimate.

5.3 Siblings

One-year-apart students living in the same building may very well be siblings. As there is no identifier for the family in the data, we cannot observe real siblings. However, we define probable siblings as any two students living in the same place and sharing the same parental characteristics (parents' occupations and scholarship status). As a robustness checks, we then exclude all probable siblings from the estimation sample. This is more conservative than excluding actual siblings, as two students living in the same place and sharing the same parental characteristics may not be siblings, and leads to remove 835 pairs.

Figure 6 confirms that our results are not driven by siblings and are robust to excluding probable siblings as the estimate is unchanged (fourth estimate in the graph) and, although

32

Figure 6 – Robustness checks

Excelling probable siblings
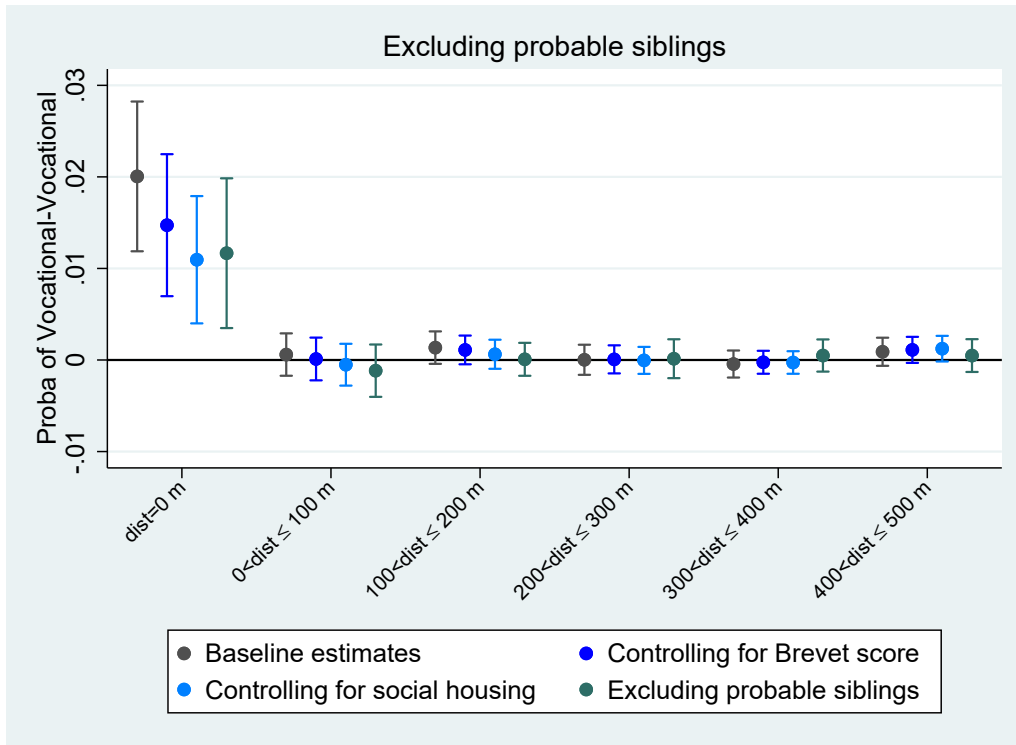
a bit less precise, still significant and not significantly different from the preceding estimate.

5.4 Linear-in-means model

The peer effect literature often considers linear-in-means models, in which the individual outcome depends on the mean outcome of their peers. The pairwise model is our preferred specification because it is the only way to measure the effect of the exact distance between neighbors and to measure heterogeneous effect with respect to the social composition of pairs. However, to check that our results are not due to the pairwise specification as well as to make our results more comparable to the existing literature, let us estimate a linear-in-means model in which a ninth grader's probability to follow the academic (respectively vocational) track after middle school is a function of the proportion of her neighbors from the preceding cohort in each bin of distance following an academic (respectively vocational) track.

33

Figure 7 confirms that only very close neighbors (at distance zero) matter in the probability to follow and academic or a vocational track. The magnitude of the estimates is 0.035, meaning that a 1% increase in the proportion of older neighbors following the vocational (respectively academic) track corresponds to a 3.5 percentage point increase in the individual probability to follow vocational (respectively academic) track.

Figure 7 – Linear-in-means model estimates

(a) Academic track
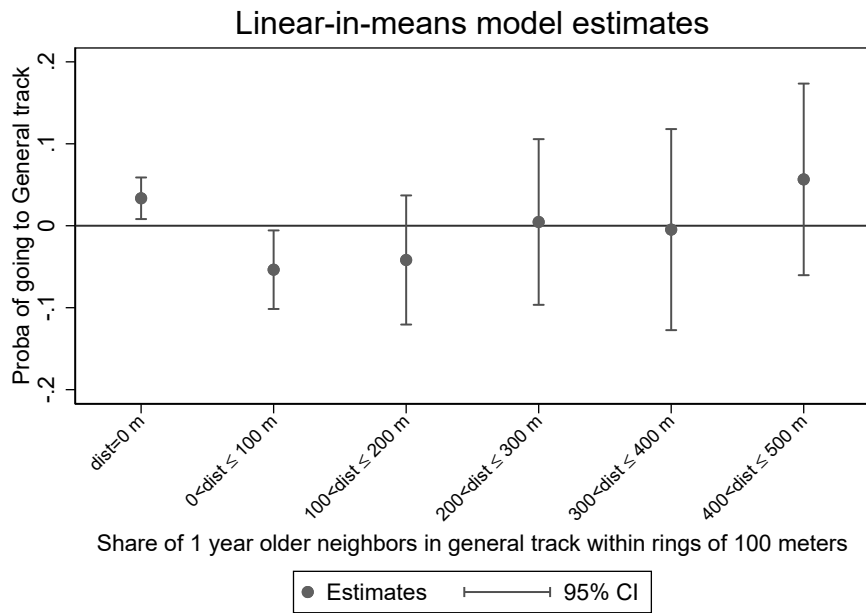
(b) Vocational track
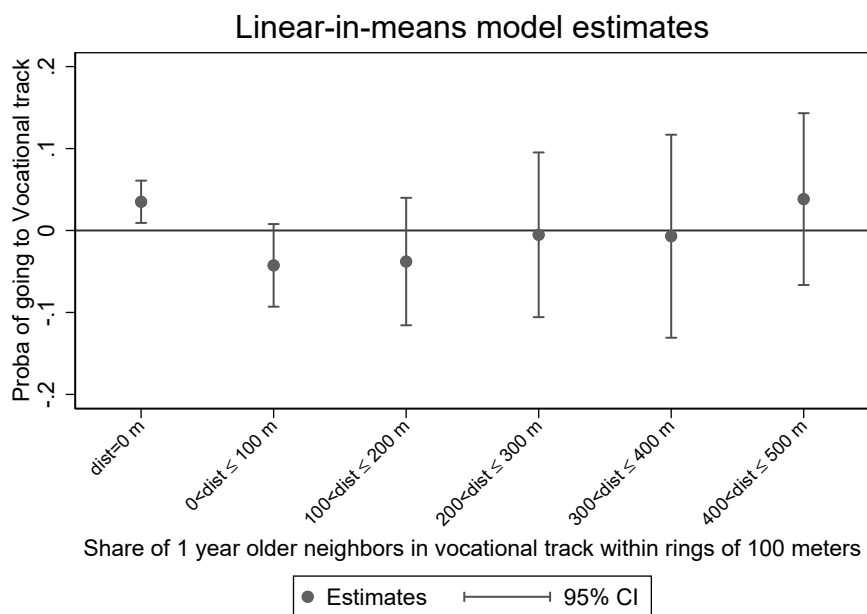
Source: MENJ, DEPP, FAERE-Apprenants 2017

5.5 Non linear model

Thus far and out of simplicity, we have estimated independent linear probability models to analyze pairs' track choices. However, choosing the same track as one's neighbor could rather be modelled as a multinomial model where:

$$
\begin{aligned}
Y_i &= k \ \text{ if } \ U_i^k > U_i^l, \ \ \forall k \neq l \\
U_i^k &= \beta^k \bar{Y}_i^k + X_i' \gamma^k + \lambda^k \bar{X}_i^k + \mu_C^k + \varepsilon_i^k
\end{aligned}
\tag{3}
$$

with $k = \{1, 2, 3\}$ each possible track : Repetition, Academic, or Vocational track, and $\bar{Y}_i^k$ the share of i's older neighbors following track k.

Estimating Equation (3) using a multinomial Logit give results that are perfectly equivalent both qualitatively and quantitatively to estimating linear probability models by ordinary least squares. Results are displayed in Appendix B in Figure B1. The corresponding odd-ratio is 1.5, meaning that a 1% increase in the proportion of older neighbors following the vocational track multiplies by 1.5 the individual probability to follow vocational track rather than the academic track (the reference).

6 Conclusion

The aim of the paper is to analyze the effect of neighbors' choices on individual track choices at the end of lower secondary education, which is a key turning point in the French school system. To take care of endogenous location decisions, we use the within-catchment-area variation in distance between neighbors.

Our results suggest that one-year-older close neighbors do matter in track choices at the end of lower secondary education, especially for pupils going to a vocational track. The effect fades away very rapidly and only same-building neighbors seem to matter.

These results are not driven by spatial sorting across neighborhoods as they hold even when excluding catchment areas where spatial auto-correlation based on socio-demographic characteristics is likely. They are not driven by previous academic achieve-

ment, social housing, nor siblings.

Neighbor effects are higher for pairs of boys and tend to increase the probability to follow an vocational track for pairs of low-SES peers, suggesting that neighbor effects are likely to accentuate social segregation across high school tracks.

These results suggest that education place-based policies relying on people's place of residence may matter, especially for the most socially and academically disadvantaged students. However, they need to be designed carefully, in a manner that would take into consideration potential effects on social segregation across education tracks.

References

Albert, D., J. Chein, and L. Steinberg (2013). Peer Influences on Adolescent Decision Making. *Current Directions in Psychological Science 22*(2), 114–120.

Barrios-Fernández, A. (2022, July). Neighbors' Effects on University Enrollment. *American Economic Journal: Applied Economics 14*(3), 30–60.

Battiston, A., S. Hedges, T. Lazarowicz, and S. Speckesser (2020, January). Peer Effects and Social Influence in Post-16 Educational Choice. Research Discussion Paper 025 025, Centre for Vocational Education Research.

Bayer, P., S. L. Ross, and G. Topa (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy 116*(6), 1150–1196.

Bifulco, R., J. M. Fletcher, and S. L. Ross (2011). The effect of classmate characteristics on post-secondary outcomes: Evidence from the add health. *American Economic Journal: Economic Policy 3*(1), 25–53.

Boutchenik, B., P. Givord, and O. Monso (2020, February). How do restrictive zoning and parental choices impact social diversity in schools?: A methodological contribution to the decomposition of segregation indices applied to France. Sciences Po publications 105, Sciences Po.

Brown, B. B. and J. Larson (2009). *Peer Relationships in Adolescence*, Chapter 3. John Wiley Sons, Ltd.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009, 10). Peer Effects and Social Networks in Education. *The Review of Economic Studies 76*(4), 1239–1267.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010). Identification of Social Interactions through Partially Overlapping Peer Groups. *American Economic Journal: Applied Economics 2*(2), 241–275.

Del Bello, C. L., E. Patacchini, and Y. Zenou (2015). Neighborhood effects in education.

DEPP (2020). L'état de l'école. Technical report, Direction de l'évaluation, de la prospective et de la performance - Ministère de l'Éducation nationale, de la Jeunesse et des Sports.

Deutscher, N. (2020, April). Place, peers, and the teenage years: Long-run neighborhood effects in australia. *American Economic Journal: Applied Economics 12*(2), 220–49.

Epple, D. and R. E. Romano (2011). Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics*, Volume 1, pp. 1053–1163. Elsevier.

Fricke, H., J. Grogger, and A. Steinmayr (2018). Exposure to academic fields and college major choice. *Economics of Education Review 64*(C), 199–213.

Gibbons, S. and S. Telhaj (2016). Peer effects: Evidence from secondary school transition in england. *Oxford Bulletin of Economics and Statistics 78*(4), 548–575.

Goux, D., M. Gurgand, and E. Maurin (2017). Adjusting your dreams? high school plans and dropout behaviour. *The Economic Journal 127*(602), 1025–1046.

Goux, D. and E. Maurin (2007). Close neighbours matter: Neighbourhood effects on early performance at school*. *The Economic Journal 117*(523), 1193–1215.

Guyon, N. and E. Huillery (2020, 06). Biased Aspirations and Social Inequality at School: Evidence from French Teenagers. *The Economic Journal 131*(634), 745–796.

Helmers, C. and M. Patnam (2014). Does the rotten child spoil his companion? Spatial peer effects among children in rural India. *Quantitative Economics 5*(1), 67–121.

Hémet, C. and C. Malgouyres (2018). Diversity and Employment Prospects Neighbors Matter! *Journal of Human Resources 53*(3), 825–858.

Jonsson, J. O. and C. Mood (2008). Choice by contrast in swedish schools: How peers' achievement affects educational choice. *Social Forces 87*(2), 741–765.

Landaud, F., S. T. Ly, and É. Maurin (2020). Competitive schools and the gender gap in the choice of field of study. *Journal of Human Resources 55*(1), 278–308.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics 28*(4), 825–860.

Ly, S. T. and A. Riegert (2014, September). Persistent Classmates: How Familiarity with Peers Protects from Disruptive School Transitions. Working Paper 2013 − 21, Paris School of Economics. working paper or preprint.

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

Monarrez, T. E. (2023, July). School attendance boundaries and the segregation of public-schools in the us. *American Economic Journal: Applied Economics 15*(3), 210–237.

OECD (2019). PISA 2018 Results - Volume I-III. Technical report, OECD.

Patacchini, E., E. Rainone, and Y. Zenou (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior and Organization 134*, 190 − 227.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, Volume 3, pp. 249–277. Elsevier.

Silliman, M. and H. Virtanen (2022, January). Labor market returns to vocational secondary education. *American Economic Journal: Applied Economics 14*(1), 197–224.

Solignac, M. and M. Tô (2018). Do workers make good neighbours? The impact of local employment on young male and female entrants to the labour market. *Annals of Economics and Statistics* (130), 167–198.

A No-fixed effect estimates

Table A1 – **No fixed-effect** estimates of neighborhs' effects on track choices - All catchment areas

	i's track-j's track						
	Same track	Rep$^{\text{ion}}$-Gen$^{\text{al}}$	Rep$^{\text{ion}}$-Voc$^{\text{al}}$	Gen$^{\text{al}}$-Gen$^{\text{al}}$	Gen$^{\text{al}}$-Voc$^{\text{al}}$	Voc$^{\text{al}}$-Gen$^{\text{al}}$	Voc$^{\text{al}}$-Voc$^{\text{al}}$
Dist. btw neighbors (100 meters)	-0.0011	0.0009**	0.0000	0.0010	0.0002	0.0000	-0.0022***
	(0.0012)	(0.0004)	(0.0002)	(0.0016)	(0.0011)	(0.0008)	(0.0006)
Dist.^{2}	0.0004**	-0.0001**	-0.0000	0.0003	-0.0002	-0.0001	0.0001*
	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Intercept	0.7316***	0.0143***	0.0033***	0.6943***	0.1307***	0.1201***	0.0373***
	(0.0087)	(0.0026)	(0.0008)	(0.0110)	(0.0064)	(0.0069)	(0.0030)
Nbr obs	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014	3,271,014
Nbr clusters	111	111	111	111	111	111	111

Source: MENJ, DEPP, FAERE-Apprenants 2017

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a pair of individuals living in the same catchment area. Controls at the pair level = gender, SES, scholarship, born French, and school choice. Standard errors in parenthesis are clustered at the catchment area level.

B   Estimation results - Non-linear model

Figure B1 – Multinomial logit model
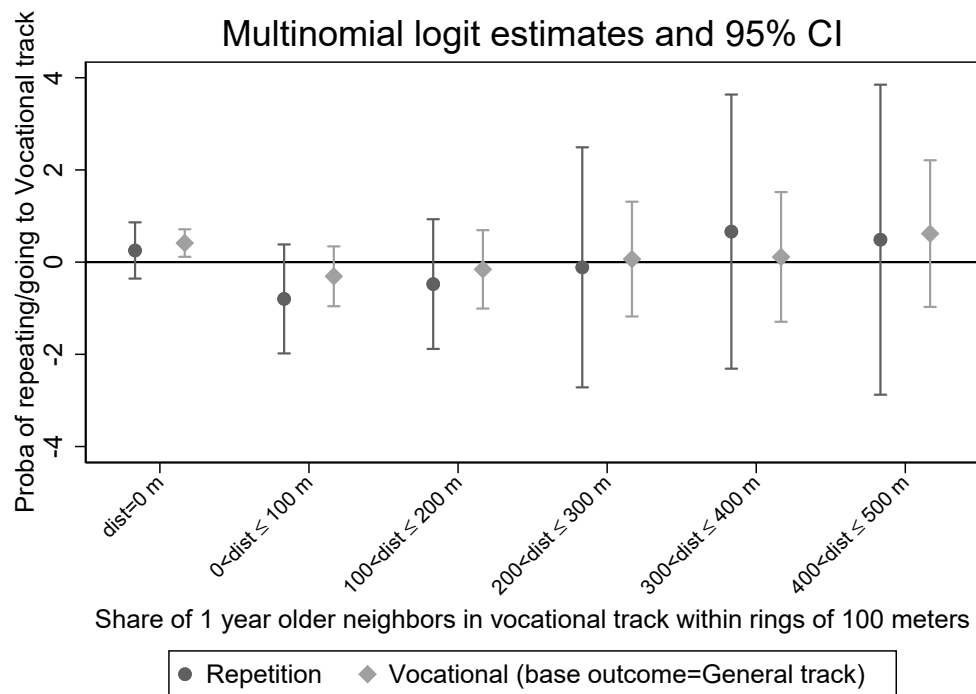
Source: MENJS DEPP, FAERE-Apprenants 2017, RPLS 2017