Friendship Networks and Political Opinions

Yann Algan* Nicolò Dalvit† Quoc-Anh Do‡ Alexis Le Chapelain§
Yves Zenou¶

March 2, 2025

Abstract

We examine how social interactions and friendships shape students' political opinions in a natural experiment at Sciences Po, a leading French university specializing in social and political sciences. The quasi-random assignment of students into short-term integration groups before their academic curriculum reduces political opinion gaps and fosters friendship formation. Using same-group membership as an instrumental variable for friendship, we find that friendship reduces opinion differences by 40% of a standard deviation in the opinion gap. Our evidence supports a homophily-enforced mechanism: friendships form among initially politically similar students, leading them to join political associations together, reinforcing their similarity. However, friendship does not significantly influence politically dissimilar pairs. Instead, it reduces opinion divergence without enforcing ideological convergence.

Keywords: Political opinion, social networks, friendship effect, polarization, homophily, natural experiment.

JEL classification codes: C93, D72, Z13.

*HEC and CEPR. Email: algan@hec.fr.

†World Bank. Email: ndalvit@worldbank.org.

‡Monash University, CEPR, and CESifo. Email: quoc-anh.do@monash.edu.

§Email: alexis.lechapelain@gmail.com.

¶Monash University, University of Southampton, and CEPR. Email: yves.zenou@monash.edu.

[1]We thank valuable comments from Devin Pope, Katia Zhuravskaya, three anonymous referees, Alberto Alesina, Ghazala Azmat, Roland Bénabou, Yann Bramoullé, Emily Breza, Jing Cai, Arun Chandrasekhar, Aureo De Paula, Ruben Durante, Georgy Egorov, Ruben Enikolopov, Marcel Fafchamps, Gigi Foster, Jeff Frieden, Matt Gentzkow, Ben Golub, Sanjeev Goyal, Nicolaj Harmon, Emeric Henry, Yannis Ioannides, Emir Kamenica, Ethan Kaplan, Rachel Kranton, Eliana La Ferrara, Horacio Larreguy, Alexandre Mas, Friederike Mengel, Markus M. Möbius, Nancy Qian, Gautam Rao, Imran Rasul, Andrei Shleifer, Erik Snowberg, Enrico Spolaore, Andreas Steinmayr, Uwe Sunde, Adam Szeidl, Fabian Waldinger, Noam Yuchtman, and from seminar and conference participants at ANU, U Melbourne, Kellogg SoM, UIUC, Adansonia conference at U Bocconi, HSE Moscow, NES, UBC, CREST Political Economy workshop, UNSW, Monash U, NUS, Queen Mary U, U of Copenhagen, Royal Holloway, Brazilian Econometric Society meeting, LMU Munich, Harvard U, U of Chicago, U Autònoma de Barcelona, Monash U Symposium on Social and Economic Networks, Barcelona Summer Forum, the Social Impacts of Economic Policies and Change conference in Heidelberg, the AEA Meeting, the Information Transmission in Networks conference at Harvard, and Aix-Marseille School of Economics. We are thankful for Sri Srikandan's assistance with the survey website. Algan acknowledges support from the European Research Council Grant 647870. Do is grateful for the financial support of the Ford Center for Global Citizenship, Kellogg School of Management, Northwestern University, the hospitality of Harvard University's Department of Economics and Weatherhead Center For International Affairs, and financial support from the French National Research Agency's (ANR) "Investissements d'Avenir" grants ANR-11-LABX-0091 (LIEPP) and ANR-11-IDEX-0005-02 (USPC), and the Banque de France - Sciences Po partnership. This paper was previously

1 Introduction

The recent rise of populism and political polarization has attracted a burgeoning research area on the related role of social interactions in social networks. Several authors attribute political polarization to the rise of social media (e.g., Sunstein, 2009, 2018; Pariser, 2011; Allcott et al., 2020; Levy, 2021; Allcott et al., 2022; DellaLena et al., 2023), which incubates echo chambers that facilitate more interactions between like-minded individuals, thus strengthen polarization of views. Others debate the quantitative importance of such mechanism (e.g., Boxell et al., 2018; Allcott and Gentzkow, 2017; Guess et al., 2018). A key yet understudied factor in this debate is the causal impact of social interactions on political opinions, as echo chambers alone do not necessarily lead to opinion polarization.

This paper provides causal estimates of this impact by addressing the endogeneity bias arising from individuals' choices to interact. We exploit a natural experiment at Sciences Po, a leading French university specializing in social and political sciences, where students are quasi-randomly assigned to integration groups during an introductory week before their first year. Being assigned to the same integration group significantly reduces the gap in political opinion between two students. We then consider friendship as the main mechanism behind this result, in the tradition of Lazarsfeld et al.'s (1944) seminal study of friends' influence on US voters. Using the group allocation as instrument for friendship, we find that friendship strongly reduces students' political opinion gap. We show how friendship influences students' pairwise shifts in opinions and shared activities, highlighting a mechanism where homophily in political views reinforces the effect of friendship.

Students at Sciences Po, the setting of this study, are known for their enthusiasm and active engagement in political movements and associations. According to the latest student survey (Foucault and Muxel, 2022), 90% of Sciences Po students express an interest in politics and 79% consider politics a central part of their lives. The curriculum orientation of incoming Sciences Po students provides a natural experiment to examine the complex issue of homophily—namely, the tendency to befriend similar individuals[1]—and its influence on the magnitude of political peer effects.

In the presence of homophily, students' social interactions become endogenous choices. To address this concern, we exploit the quasi-random assignment of students to integration groups

circulated with the title "Friendship Networks and Political Opinions: A Natural Experiment among Future French Politicians".

[1]The concept of homophily was first introduced in seminal sociological studies such as Lazarsfeld and Merton (1954) and McPherson et al. (2001), and later explored in economics by, e.g., Currarini et al. (2009) and Golub and Jackson (2012).

during the introductory week before their first year begins. During this week, students are arbitrarily assigned by alphabetical order into small groups of approximately 16, where they engage in social activities designed to facilitate integration into their new environment. To analyze the impact of these interactions, we estimate dyadic regressions of pairwise opinion differences on the same-integration group indicator across all possible student pairs. Additionally, we conduct non-parametric permutation tests to assess the effect of same-integration group interactions.

We obtain data on integration groups participation as well as other administrative data for the entry cohort of 2013. We survey their current and pre-Sciences Po political opinions and association activities in March 2014, and use Leider et al.'s (2009) incentive-compatible method to elicit their social networks of friendship.

We find a negative and statistically significant effect of the same-integration group indicator on students' pairwise political opinion gap. On average, being randomly assigned to the same-integration group reduces the political opinion gap by 0.16, which corresponds to 8% of the mean (1.93) and 11% of the standard deviation (1.47) of the political opinion gap,[2] and increases a pair's probability of convergence and reduces their probability of divergence. To allay concerns on group compliance and confounders of students' last names, we closely follow Harmon et al.'s (2019) approach of using hypothetical membership in fictive groups formed by the 16-person rule as instruments of being in the same integration group.

We then investigate the mechanisms driving these results, particularly how friendship influences political opinion convergence. The integration week is explicitly designed to foster friendship formation, and we estimate that being in the same-integration group increases the likelihood of a lasting friendship by 17 percentage points—substantially more than the effect of any observable similarity. Since integration groups are dissolved before the academic year begins, we argue that same-integration group membership affects surveyed political opinions only through friendships that persist beyond the integration period. Under this exclusion assumption, we use the same-integration group indicator as instrument for pairwise friendship to estimate its effect on political opinion gap. We further provide lower bounds of this effect in plausible cases when this exclusion restriction does not fully hold.

Our method yields powerful effects of social interactions and friendship. A friendship link between two students reduces their differences in political opinions by almost a point after 6 months, equivalent to 50% of the mean and 65% of the standard deviation of opinion differences.

[2]Political opinions are measured on a scale from 1 (extreme left) to 10 (extreme right).

We further find that the effect of friendship is strongest among students with similar pre-Sciences Po political views. This evidence supports what we term the "homophily-enforced mechanism," in which homophily along a particular dimension—such as political views—reinforces the friendship effect on that dimension. Specifically, when two individuals initially share similar views in politics, friendship increases their interactions on politics, further strengthening their similarity over time. In contrast, among initially dissimilar pairs, friendship may have little impact on that dimension.

In particular, among initially politically similar pairs, friendship strongly increases the likelihood of joining the same political associations, thereby encouraging more political interactions. Conversely, among pairs with divergent pre-Sciences Po political views, friendship does not lead them to join the same political associations and, consequently, does not significantly affect their subsequent political opinion gap. In short, similarity fosters friendship, which in turn reinforces similarity along the same dimension.[3] This mechanism highlights the complementarity between social and political proximities in producing political similarity as previously shown in Harmon et al. (2019).

Our paper contributes to different strands of literature. A long-standing question in political economics and related fields is whether social peers influence each other's political attitudes. Answering this question presents two main empirical challenges. First, obtaining high-quality data that accurately identifies a person's social peers while also capturing relevant measures of political attitudes is difficult. Second, and even more challenging, is the estimation of causal peer effects, which is complicated by identification issues stemming from the non-random assignment of peers.

One strand of the literature has sought to address this second issue through careful econometric modeling, as seen in studies by Battaglini and Patacchini (2018), Canen and Trebbi (2016), and Canen et al. (2023). However, these findings critically rely on strong structural assumptions embedded in their models. In terms of quasi-experimental evidence, political peer effects have been identified in certain contexts, though primarily among legislators, using crude proxies for peer relationships, such as seating arrangements (e.g., Fowler, 2006; Gabel and Scheve, 2007; Cohen and Malloy, 2014; DellaVigna and Gentzkow, 2010; Rogowski and Sinclair, 2012; Saia, 2018; Harmon et al., 2019; Carlsson et al., 2021; Lowe and Jo, 2025; Darmofal et al., 2025). While these findings on peer effects among politicians are insightful, they may not necessarily generalize to the broader public. Peer effects among legislators could be driven by strategic considerations, such as vote trading, which

[3]This mechanism helps reconcile this paper's strong friendship effect with the small and sometimes insignificant peer effects on academic performance found in the literature (e.g., Angrist and Lang, 2004). When friendship is voluntarily formed and strengthened around a specific dimension, it significantly influences that dimension. In contrast, in peer effect studies, peer groups may not necessarily be formed around the dimension being measured as the outcome, meaning peer effects are not always observed.

are unlikely to operate outside the realm of professional politics.

Two major unanswered questions remain: do political peer effects extend beyond professional politics to other contexts, such as among voters, and do they occur among more traditional peer groups, such as friends, partners, and relatives? This paper makes an important contribution to these questions by leveraging students' quasi-random peer assignments along with detailed data on friendships and political attitudes.

This paper also contributes to an active literature on the impacts of exogenous variations in social networks and social media on political polarization (e.g., Gentzkow, 2006; DellaVigna and Kaplan, 2007; Gerber et al., 2009; Gentzkow et al., 2011; Kendall et al., 2015; Allcott et al., 2020; Levy, 2021; Allcott et al., 2022; DellaLena et al., 2023).[4] In comparison with this literature, we study the impact of social interactions on the convergence or divergence of pairwise political opinions and provide a homophily-enforced mechanism, by which friendship causes initially politically-similar students to join political associations together, which reinforces their political similarity.

Our focus on the role of friendship as a mechanism, in the tradition of Lazarsfeld et al.'s (1944), extends beyond the extensive literature on peer effects under randomized assignment (surveyed by, e.g., Sacerdote, 2014). Indeed, our findings align with Carrell et al.'s (2013)'s discovery that individuals' peer choices, when interacting with their characteristics, can significantly impact educational outcomes.

Our findings also relate to the literature on the contact hypothesis. Since Allport's (1954) seminal argument that intergroup contact can reduce prejudice between ethnic groups under certain conditions, this theory has inspired a vast body of empirical research across various contexts (Pettigrew and Tropp, 2006; Paluck et al., 2019). Notably, many studies examine how interactions with ethnically diverse groups influence the political attitudes of white individuals (e.g., Billings et al., 2021; Kaplan et al., 2022; Polipciuc et al., 2023). According to contact theory, intergroup contact fosters improved relations, particularly when both groups share equal status or common goals (Boisjoly et al., 2006; Cai and Szeidl, 2018; Merlino et al., 2019; Rao, 2019; Lowe, 2021; Corno et al., 2022; List et al., 2023; Siddique et al., 2024). Our results are in line with the contact hypothesis in that students' exposure to each other strengthen their interactions and reduce their divergence in political opinion. We further deepen the underlying mechanism by showing the intricate role of friendship in shaping opinions.

[4]There is also significant literature in political science demonstrating how political interest develops in young people, with childhood experiences and college socialization playing crucial roles in shaping political attitudes (e.g., Prior, 2010; Healy and Malhotra, 2013; Mendelberg et al., 2017; Prior, 2018).

The rest of the paper unfolds as follows. Section 2 describes the study's context. Section 3 shows the impact of the quasi-random allocation to the integration group on political opinion gap. Section 4 presents the main mechanism of friendship on opinions and behaviors. Section 5 then investigates the main drivers at work and explores the homophily-enforced mechanism behind friendship. Section 6 concludes.

2 Background on Sciences Po and relationship among its students

This section provides an overview of the natural experiment's context. Sciences Po is a leading French university specializing in social and political sciences. Comparable in size to the London School of Economics in the UK or Yale in the US, it enrolls approximately 15,000 students across its undergraduate and graduate programs each year, with an annual intake of 1,500 first-year bachelor's students. Founded in 1870 by Émile Boutmy to train new political leaders following France's defeat by Bismarck's Germany, Sciences Po has distinguished itself for over 150 years through the political engagement of its students. According to the latest Sciences Po student survey (Foucault and Muxel, 2022), 90% of students express an interest in politics, with 54% stating they are very interested—compared to just 11% of 18-25-year-old in France (Foucault and Teinturier, 2022). This engagement translates into active participation in debates and student associations: 79% of students report that politics plays a very important role in their lives, and 74% frequently discuss politics with their friends at Sciences Po, whereas only 45% engage in political discussions with their parents.

Moreover, most students actively participate in associations, allowing them to rally around shared aspirations and engage with the broader Sciences Po community. These associations span various fields, including politics, humanitarian work, environmental activism, identity-based movements (such as #MeToo, LGBTQ+ rights, and anti-discrimination initiatives), as well as arts and sports. According to the latest official report on student life,[5] 231 student associations are officially registered and receive institutional support to develop their projects. Political associations attract the most members, with 40% of students reporting participation in an association focused on policy debates. These associations primarily aim to invite distinguished external figures—senior civil servants, politicians, and business leaders—to discuss national and international political issues. Notably, the vast majority remain non-partisan. For instance, the most representative association, "Parlement des Étudiants", organizes debates featuring political figures from across the ideological spectrum. This strong culture of political engagement provides a particularly compelling setting for studying

[5]Student Life Report 2023/2024, Sciences Po.

how political opinions form and evolve among students and their peers.[6]

Beyond political associations, student involvement spans various causes. According to official records, 36% of students participate in a humanitarian association, 16% in an environmental association, 14% in a women's rights association, 11% in a political party, 9% in a religious association, 8% in an LGBTQ+ rights association, and 6% in an anti-racism association.

Before the first year. Sciences Po students generally do not know each other before their first year. Approximately 75% come from high schools across France and gain admission through the standard selection process, which involves two highly competitive rounds of written and oral exams. Another 5% are international students from a variety of schools abroad. The remaining 20% are admitted through an affirmative action program known as Convention Éducation Prioritaire (CEP). This program selects the top students from designated high schools in disadvantaged areas across France through a separate evaluation process based on academic records and oral assessments. Across all admission channels, the likelihood of students coming from the same high school is extremely low, making pre-existing friendships rare. In our sample, we identify only two pre-Sciences Po friendship pairs within the same integration group—accounting for just 0.2% of friendship pairs or 0.02% of all observed friendships.

Integration Groups. During the integration week, held just before the start of the academic year, incoming undergraduate students are formally introduced to Sciences Po and assigned to integration groups of approximately 16 students each, based on alphabetical order. In our sample, this integration week took place in the last week of August 2013. Throughout the week, students participate in various extracurricular activities—such as games and guided tours of Paris—exclusively within their assigned groups. These activities are specifically designed to foster and strengthen social connections among students. Notably, no activities during this week are related to academic or political topics, nor do they involve discussions on students' political opinions. Conversations with students indicate that they primarily remember the integration week for the friendships they formed rather than for

[6]Anecdotes abound about how Sciences Po students influence one another's political views. A notable example is former President Jacques Chirac, who underwent a drastic shift from communist sympathizer to Gaullist and later became a leader of the French political right. Upon entering Sciences Po in 1951, the young communist met and formed a close friendship with his classmate Bernadette Chodron de Courcel. They got engaged in 1953 and married in 1956. Bernadette's aristocratic and bourgeois background, along with her family's influence, played a significant role in Chirac's political transformation and career (Chirac, 2012). Conversely, former President François Mitterrand underwent a political shift in the opposite direction, evolving from a nationalist militant to the leader of the Socialist Party during his time at Sciences Po.

any specific content.[7] After the week concludes, integration groups are dissolved and are not used in any subsequent activities.

After the integration week. At the start of their first year at Sciences Po, students immerse themselves in a wide range of academic and extracurricular activities. No longer restricted to their integration groups, they interact with the entire cohort. Academically, all first-year students follow the same core curriculum, with the only variation being their choice of language courses. For extracurricular activities, students can join multiple associations from a selection of approximately one hundred. These include politically affiliated groups linked to parties and movements, as well as associations focused on sports, the arts, and cultural or ethnic identities. Many of these associations meet regularly for practices, events, and social gatherings, fostering deeper engagement and community building. Our primary focus is on students' association choices as a key behavioral outcome during their first year.

The choice of associations primarily takes place during the first two weeks of September, coinciding with the start of academic classes. During this period, student associations set up stands at *La Péniche*, the central campus venue, allowing students to explore and join groups that align with their interests. Although we do not have precise data on when friendships are formed, this sequence strongly suggests that being in the same integration group facilitates both the selection of the same associations and the development of friendships. Friendships continue to form throughout the academic year, likely influenced by students' exposure to one another and their shared interests.

3 Effect of integration group assignment on political opinions

This section examines the quasi-random allocation of students into integration groups and its impact on their political opinions.

3.1 Data sources, survey design, and measurement

Administrative data. First, data on the integration group organization and student characteristics are obtained from Sciences Po's official administrative source. They include gender, nationality, academic program (e.g., dual-degree programs joint with another institution), admission type (such as regular exam admission, international admission, or priority admission through the affirmative ac-

[7] While the effectiveness of integration week in fostering lasting friendships has faced some skepticism, our results in Table 7 provide strong empirical support for its impact. Anecdotal evidence from students further corroborates these findings.

tion channel), département (the French administrative district) and region of high school, high school major, profession of parents, permanent address' postal code, and tuition fee that proxies for household income.[8] Based on those variables, we construct dyadic variables, including the same-integration group indicator $IG_{ij} = \mathbf{1}_{\{I_i = I_j\}}$, and the dyadic differences over individual characteristics $\mathbf{X}_{ij}$.

Surveyed measures of political opinions. We ran an internet-based survey in March 2014 on the cohort of Sciences Po first-year students who start in September 2013 to measure students' political opinions and friendship.[9] Our survey took place during a vacation week in the middle of the second semester. We ask students' current political opinions Y_i as well as their pre-Sciences Po political opinions Y_i^0 from before their arrival at Sciences Po in August 2013. The answers are given on common scale from 1 to 10, 1 being extreme left and 10 extreme right. Given Sciences Po's emphasis on political science and politics, students are highly familiar with this scale in the French context.

We define the current political opinion gap between two students i and j as $DY_{ij} = |Y_i - Y_j|$, and the pre-Sciences Po opinion gap as DY_{ij}^0. The surveyed change in the opinion gap is then given by $ChgD_{ij} = DY_{ij} - DY_{ij}^0$. To further characterize the evolution of political opinions, we define the direction of opinion change as $DIR_i = \mathbf{sign}(Y_i - Y_i^0)$. Apart from the trivial case where both $DIR_i = DIR_j = 0$, we classify a pair (i, j) as a (weakly) *converging* pair if both DIR_i and DIR_j move toward each other's opinions:

$$\begin{cases} DIR_i \geq 0 \text{ and } DIR_j \leq 0 & \text{if} \quad Y_i \leq Y_j, \\ DIR_i \leq 0 \text{ and } DIR_j \geq 0 & \text{if} \quad Y_i \geq Y_j. \end{cases}$$

Among converging pairs, we define *strong convergence* if $DIR_i \times DIR_j \neq 0$, meaning both students actively shift their positions toward each other. Similarly, apart from the trivial case where both $DIR_i = DIR_j = 0$, we classify a pair (i, j) as a *diverging* pair if both DIR_i and DIR_j move away from each other's opinions:

$$\begin{cases} DIR_i \geq 0 \text{ and } DIR_j \leq 0 & \text{if} \quad Y_i \geq Y_j, \\ DIR_i \leq 0 \text{ and } DIR_j \geq 0 & \text{if} \quad Y_i \leq Y_j. \end{cases}$$

[8] At Sciences Po, a student's tuition fee is determined by parents' tax bracket from their tax declaration(s), ranging between 0 and 10,000 euros. There is no better information on the precise household income, as the administrative data are mostly missing when it comes to very rich parents' declared incomes.

[9] The translated survey can be found in Appendix D. We also surveyed the cohort that entered in 2009 for a different purpose (the integration groups were not available to that cohort).

Among diverging pairs, we define *strong divergence* if $DIR_i \times DIR_j \neq 0$, indicating both individuals shift their opinions further apart. Finally, we classify *co-movement* pairs as cases where $DIR_i \times DIR_j \geq 0$, except for the trivial case where both $DIR_i = DIR_j = 0$.

While the political opinion gap $ChgD_{ij}$ is our most quantitatively relevant outcome, measures of pairwise opinion movements are qualitatively important for understanding the directions in which opinions evolve. Additionally, surveyed directional changes in opinions are far less susceptible to recall bias on the retrospective assessment of pre-Sciences Po opinions (Fischhoff and Beyth, 1975; Bradburn et al., 1987). Even if recalled opinions are biased toward aligning with respondents' current views at the time of the survey, this bias generally does not affect the reported direction of opinion change.[10]

3.2 Main empirical designs

We employ two approaches to test the randomness of integration group assignments and estimate the causal effect of the intervention on students' political opinion changes: (i) randomized permutation tests, which require no parametric or functional assumptions about the relationship between group assignment, observable characteristics, and outcomes,[11] (ii) dyadic regression specifications, which link the pairwise same-integration group indicator with other pairwise characteristics and outcomes to assess the impact of group assignment on political opinion changes.

Permutation tests. This nonparametric approach tests the null hypothesis of randomized group assignment against the alternative hypothesis of selection of similar individuals into groups by (i) computing a test statistic that measures within-group similarity in the original sample as well as in a large number of its randomized permutations, and (ii) comparing the original sample's statistic with its distribution from the permutations to obtain the test's p-value. In our context, we choose the test statistic as the ratio between the within-group and the between-group standard deviations of a predetermined variable, such as gender or initial opinion. Selection into groups by this variable implies a small value of the test statistic (zero in case of perfect selection into group), hence the test's p-value is calculated as the left-tail quantile of the test statistic's distribution from 300 randomly drawn permutations. This test statistic also has the advantage of being invariant to any affine

[10] This bias may influence the variable DIR_i by causing some respondents to report no change ($DIR_i = 0$) instead of acknowledging a minor shift. To account for this potential issue, we consider both weak and strong definitions of convergence and divergence.

[11] See, e.g., Kennedy's (1995) for a summary of the advantages of permutation tests.

transformation of the variable, such as a change of unit or a multiplication of the scale.[12]

Once we have established integration group assignment's randomness, we will use the same approach to test the null hypothesis that there is no effect of group assignment on political opinion against the alternative hypothesis that group assignment makes same-integration group members' political opinions more similar. The implementation simply replaces predetermined variables in the previous tests by individuals' present political opinion, so the test statistic is now the ratio between the within-group and the between-group standard deviations of political opinion.

Dyadic specifications. This parametric approach considers the sample of unordered pairs of students (i, j) and dyadic variables over those observations,[13] including the same-integration group indicator IG_{ij} and pairwise covariates X_{ij} calculated as pairwise similarity/difference over all predetermined variables obtained from administrative data (see details in section 3.1). First, in order to test the exogeneity of IG_{ij}, we regress it on the dyadic covariates X_{ij}'s. Exogeneity is rejected if there are more such statistically significant coefficients than in the randomized case.

Next, we use this dyadic approach to estimate the effect of being in the same integration group on pairwise outcomes, e.g., the pair's change in opinion gap $ChgD_{ij}$. The regression of $ChgD_{ij}$ on IG_{ij} estimates the Average Treatment Effect $\beta_{IG} = \mathbb{E}[ChgD_{ij}|IG_{ij} = 1, \mathbf{X_{ij}}] - \mathbb{E}[ChgD_{ij}|IG_{ij} = 0, \mathbf{X_{ij}}]$.

Statistical inference in dyadic specifications. Our statistical inference framework follows closely Harmon et al. (2019) in addressing potential correlated error terms due to correlated shocks that may have occurred during the integration week within each integration group, e.g., accidental events that may have shifted the group's opinions.[14] In presence of such correlated shocks, the estimation of the estimators' variance-covariance matrix allows for arbitrary correlations between the error terms of two pairs (i, j) and (i', j') if and only if either (i, i'), (i, j'), (j, i'), or (j, j') belongs to the same integration group.[15]

[12] The statistic is directly computed for continuous and binary variables. For category variables, e.g., each student's high school major, we first break it down to binary variables (indicators) representing each category (e.g., an indicator whether a student's high school major is "scientific" or not), then compute the within-group/between-group ratio statistic, and average it over all categories.

[13] We consider (i, j) and (j, i) as the same pair, since the main intervention variable IG_{ij} is symmetric by nature.

[14] Those shocks are uncorrelated to the intervention variable IG_{ij}, so they cannot bias our OLS or IV estimates.

[15] This form of correction for clustering has been discussed in Fafchamps and Gubert (2007), Cameron and Miller (2014), and systematically used in Harmon et al. (2019). In the balance tests using dyadic specifications, as we aim to detect significant correlations between any covariate and the same-integration group indicator, we impose a stronger restriction on the clustering structure of the error terms: clustering by the interaction of i's group and j's group.

Both approaches are useful. Permutation tests are particularly helpful to examine the exogeneity of group assignment, because they are nonparametric by nature, and needs no distributional assumption. This advantage addresses parametric specifications' problem of incorrect inference due to misspecification. In our context, parametric specifications usually over-simplify the covariance structure across individuals and pairs, and it is not clear whether standard error clustering can best address the issue. This complexity arises because the data generating process in group assignment is equivalent to each individual sequentially drawing a group variable without replacement, so an individual's drawn group can be dependent on earlier drawn groups. On the other hand, permutation tests only yield statistical inference, not quantifiable estimates.[16] In addition, the IV design to estimate the effect of friendship on political opinion can only be adapted to the dyadic specification. The two methods are thus complementary and both help support the results' persuasiveness.

3.3 Identification concerns and robustness tests

Confounded imperfect compliance to group assignment. We first address the concern that compliance with the alphabetical assignment to integration groups may be imperfect, as students could opt out of their assigned group. In principle, noncompliance could pose a problem if it correlates with political opinions—for instance, if noncompliers deliberately avoid groups with opposing views and instead switch to groups with more similar views.

Can this be a major issue for our empirical strategy? First, only 4% among participants in integration groups have last names that are distant from the rest of the group (i.e., the likely noncompliers).[17] Dropping them does not affect the empirical results. Second, it seems unlikely that new students could select into integration groups by peers' political views, given that almost all of them had not known each other before. Furthermore, they would have no valid reason to convince the organizers to switch to a specific group, as all groups took place at the same time. Based on our understanding, most noncompliance was due to idiosyncratic, arbitrary reasons.

To make sure that the main results do not depend on those issues, we first follow closely Harmon et al. (2019) in designing an instrumental variable of the same-integration group indicator IG_{ij}. We rank the cohort's 800 students in alphabetical order of their last names and assign the rank $AlphRank_i$ to each student i. We then go through this list and pick sequentially each group of exactly 16 students, to end up with exactly 50 hypothetical groups. The dyadic indicator whether

[16] Furthermore, while it is possible to condition on values of covariates in permutation tests if they are discrete, a large number of covariates will quickly reduce the range of possible permutations and limit the tests' validity. Without covariates, we cannot test for random assignments conditional on covariates.

[17] Recorded participation rate in the integration groups is 92% among the 800 students of the 2013 cohort.

two students belong to the same hypothetical group $SameHypGroup_{ij}$ serves as instrument for IG_{ij}.[18] It is by construction orthogonal to compliance issues, and turns out to be highly correlated with the same-integration group indicator.

Based on similar intuitions, we further introduce a more flexible instrument, the alphabetical distance between two students (i, j).[19] While this instrumental variable is less stringent than the main instrumental variable $SameHypGroup_{ij}$ because it is not based on a strict partition into groups, it is more flexible to fit certain features in the data, such as groups with irregular sizes.

Alphabetical order is correlated with omitted factors. There remains a potential concern that last names' order can correlate with certain confounders based on family, ethnicity, as well as social and cultural heritages that may also influence social connections and social and political outcomes. For example, students with hard-to-pronounce names may share certain characteristics such as race and ethnicity, as well as similar educational and labor market outcomes (Ge and Wu, 2024).[20] To address this issue, we further follow Harmon et al. (2019) in controlling for the alphabetical distance between each pair of students' names within a much larger list of last names that are still drawn from the population in a similar way. For this purpose, we obtain the last names of all students from cohorts that entered Sciences Po in any year from 2009 to 2014 (note that our sample comes from the 2013 cohort), then compute the alphabetical distance within this extended list of names.

Next, we can further strengthen our approach by restricting the sample to only pairs of students whose alphabetical distance is close. Intuitively, we consider same-group and different-group pairs of students within a bandwidth of the cutoff between two consecutive groups. Analogous to the logic of a Regression Discontinuity Design, around the threshold between two groups, same-group and different-group pairs are almost identical in both observable and unobservable characteristics (Lee and Lemieux, 2010), which reinforces the identification assumption of exogeneity of integration

[18] Mathematically, $SameHypGroup_{ij} = \mathbf{1}\left\{\left\lceil \frac{AlphaRank_i - 1}{16} \right\rceil = \left\lceil \frac{AlphaRank_j - 1}{16} \right\rceil\right\}$.

[19] Mathematically, $AlphDist_{ij} = \min(|AlphRank_i - AlphRank_j|, 24)$. We winsorize the alphabetical distance at 24, 1.5 times the average length of a group to exclude excessive variation above 24 that unlikely matters to same-integration group membership.

[20] In particular, it would be a concern to find certain integration groups overpopulated by individuals with very common family names from the same ethnic origin, such as Nguyen among Vietnamese, Kim among Koreans, and last names starting with W, X, Y, and Z among Chinese. Our manual check of all last names reveals no such concern, as only one last name appears more than twice and carries ethnicity information, while the others are mostly common French names that should not correlate with a student's background or political view. Excluding repeated last names does not alter the empirical results. Furthermore, we perform checks in which we drop (i) all names starting with each specific letter, or (ii) all students with each specific non-French nationality, or all French family names starting with "de" and similar prefixes that can correlate with an aristocratic background (shown in Appendix Tables A29 and A31).

Table 1: Descriptive Statistics of Dependent Variables

Panel A: Monadic Dependent Variables

Variable	(1) Full Sample			(2) Benchmark Sample		
	Mean	Standard deviation	Obs.	Mean	Standard deviation	Obs.
Political Opinion in March 2014 (1-10)	5.044	(1.755)	472	5.091	(1.712)	331
Initial (Pre-Sciences Po) Political Opinion (August 2013) (1-10)	5.108	(1.958)	463	5.148	(1.934)	331
Political Opinion in 2015	4.853	(1.807)	285	4.818	(1.746)	331
Membership in an Association in 2014	0.597	(0.491)	499	0.642	(0.480)	330

Panel B: Dyadic Dependent Variables

Variable	(1) Full Sample			(2) Benchmark Sample		
	Mean	Standard deviation	Obs.	Mean	Standard deviation	Obs.
Difference in Political Opinions in March 2014	1.932	(1.467)	105,111	1.927	(1.469)	52,326
Initial (Pre-Sciences Po) difference in Political Opinions (August 2013)	2.211	(1.631)	101,025	2.194	(1.621)	52,326
Difference in Political Opinions in 2015	2.014	(1.538)	27,027	1.940	(1.496)	15,920
Participants in the Same Sports Activities	0.600	(0.490)	52,003	0.586	(0.493)	23,436
Membership in the Same non-Sports Association	0.085	(0.279)	52,003	0.097	(0.296)	23,436
Membership in the Same Political Association	0.018	(0.131)	52,003	0.023	(0.149)	23,436
Membership in the Same Activism Association	0.007	(0.084)	52,003	0.008	(0.088)	23,436
Membership in the Same Identity-related Association	0.005	(0.072)	52,003	0.004	(0.061)	23,436

Notes: Statistics in (1) are computed on the full sample of data available for each variable, while statistics in (2) are computed on the benchmark sample, which is detailed in Table A1.

group assignment.[21]

3.4 Data description

In Table 1, Panel A lists the descriptive statistics of students' political opinion and behavior. The average of political opinion slowly shifts to the left over time (i.e., to lower value, as 5.5 represents the center). Its variance decreases by about 10 percent from before Sciences Po until the survey in March 2014. Panel B shows a similar pattern from the analogous statistics for the corresponding dyadic outcome variables. The average dyadic opinion gap experiences a reduction of about 13 percent from 2.21 before Sciences Po to 1.93. Appendix Table A1 (Appenidx A) describes in detail all variable definitions, and Appendix Table A3 completes the descriptive statistics of other variables used in the empirical analysis.

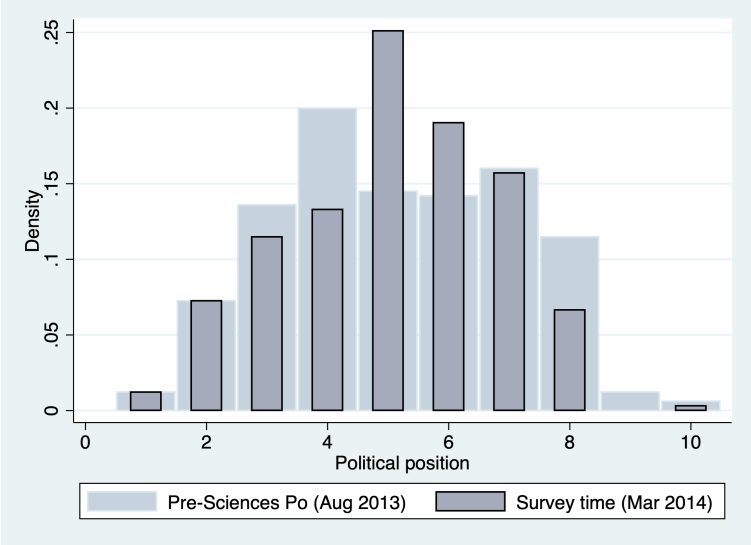
Figure 1 shows the distributions of political opinions in March 2014 and in August 2013. The net decrease in variance of political opinion is illustrated by two major changes. First, the bimodal distribution in 2013, with two modes at 4 and 7 corresponding to mainstream left-right politics, becomes unimodal in 2014 with a strongly dominant center in 5-6. Second, the right to extreme right positions (8-9-10) experience a strong reduction over the period.

[21]While similar, this is not a proper Regression Discontinuity Design, since there is no exact cutoff. It is thus not possible to implement standard RDD methods, or choose an optimal bandwidth. We pick the bandwidth $AlphDist_{ij} \leq 24$, noting that the results remain similar for a broad range of bandwidths.

3.5 Exogeneity tests of assignment into integration groups

This subsection evaluates the claim that the assignment into integration groups by alphabetical order of the students' family name is exogenous, by checking that alphabetically close family names do not carry other information that could stack up students with similar backgrounds in the same group. We first start with a permutation test, as described in subsection 3.2. Table 2 shows the p-values of the left-sided permutation tests corresponding to all predetermined covariates, calculated based on the empirical distribution of the test statistics (the within-/between-standard deviation ratio) drawn from 300 random permutations of the original sample. None of the tests can reject the null hypothesis of randomized assignment into integration groups at 5%.

Figure 1: Distributions of Political Opinions



Notes: Distributions of Individual Political Opinions just before joining Sciences Po (August 2013) and at the time of survey (March 2014).

Second, in Table 3 we report dyadic linear regressions of IG_{ij} on observable pairwise covariates, either altogether (Panel A) or one by one (Panel B), as explained in subsection 3.2. In the pooled regression, the joint hypothesis that they are all equal to zero has a very small F-stat. All coefficients are not statistically different from zero at 95% confidence, except one. Even this statistically significant coefficient shows the opposite of homophily, namely that individuals with similar high school major are less, not more likely in the same integration groups. It is quite natural that among 15 estimates, one coincidentally has a p-value below 0.05. Furthermore, given their very small mag-

Table 2: Permutation Tests of Integration Group Assignment's Randomness

Variable	Within-Group Statistics	Actual value	p-value
Initial Political Opinion	Within-/Between- Standard Deviation Ratio	2.282	0.810
Tuition Fees	Within-/Between- Standard Deviation Ratio	1.842	0.140
Gender	Within-/Between- Standard Deviation Ratio	1.954	0.290
Affirmative-Action Admission	Within-/Between- Standard Deviation Ratio	1.753	0.167
Second Nationality	Within-/Between- Standard Deviation Ratio per Category	1.261	0.813
Admission Type	Within-/Between- Standard Deviation Ratio per Category	2.496	0.473
Program	Within-/Between- Standard Deviation Ratio per Category	2.244	0.560
Parents' Profession	Within-/Between- Standard Deviation Ratio per Category	2.356	0.233
High School Major	Within-/Between- Standard Deviation Ratio per Category	2.310	0.363
Département of High School	Within-/Between- Standard Deviation Ratio per Category	2.744	0.993
Region of High School	Within-/Between- Standard Deviation Ratio per Category	2.614	0.977
ZIP code	Within-/Between- Standard Deviation Ratio per Category	2.806	0.867

Notes: Permutation tests of integration group assignment's exogeneity by 300 Monte Carlo permutations of the full sample. For continuous and binary variables, the test is performed on the distribution of the ratio of within-group and between-group standard deviations. For category variables, the test is performed on the distribution of the average of this ratio across all indicators representing each category. p-values are computed with respect to the left tail (rejection of low within-group variation with respect to between-group variation). See Appendix A and Appendix Table A1 for description of variables and sample.

nitude, their corresponding 95% confidence intervals remain small,[22] suggesting that their inclusion in the main analysis does not matter much to the coefficient of IG_{ij}, a fact that we can later verify. To remain cautious, we do control for all of those covariates throughout the empirical exercises.

3.6 Effects of integration group membership: Permutation tests

Based on the claim of random assignment of students into integration groups, we move on to establish its causal effects on political opinions. We first implement permutation tests (as described in subsection 3.2 and also performed in Table 2) confronting the alternative hypothesis of lowered within-group variation due to integration group assignment against the null hypothesis of no integration group assignment effect. We apply this procedure to (i) changes in political opinions from before Sciences Po until March 2014 and (ii) individual political opinions surveyed in March 2014, and plot the distributions of simulated test statistics in Figures 2A and 2B, respectively. In both cases, we can reject at 5% the null hypothesis that there is no effect from group assignment on individual opinions.[23]

3.7 Effects of integration group membership: Dyadic regressions

To better quantify those nonparametric results, in this subsection we proceed with the parametric dyadic specification (subsection 3.2). First, Panel A of Table 4 shows the regressions of the indicators of different types of pairwise opinion evolution on the same-integration group indicator, controlling for the full set of observable covariates. Being in the same integration group increases a pair's chance

[22]The largest 95% confidence interval, corresponding to the variable whether the pair are both admitted via Sciences Po's affirmative action program, is still contained within [-0.02,0.04].

[23]The result remains robust to permutations stratified by each covariate, as shown in Appendix Table A30.

Table 3: Balance Test of Same Integration Group Indicator

Panel A: Balance Test by Pooled Dyadic Regression

Dependent Variable	Same integration group	Dependent Variable	Same integration group
7pt Same Gender	0.000604	Same High School Major	-0.00235**
	(0.002079)		(0.00110)
Both Female	-0.00184	Diff. in Tuition Fees ('000 Euros)	-0.000329
	(0.00390)		(0.000454)
Same Nationality	-0.00663	Both Free Tuition	-0.00214
	(0.00764)		(0.00231)
Same Admission Type	-0.000897	Same Parents' Profession	0.00123
	(0.003050)		(0.00212)
Both Affirmative Action	0.00847	Same ZIP Code	-0.000499
	(0.01475)		(0.003941)
Same Département of High School	0.00251	Both living in Paris	0.0000606
	(0.00713)		(0.0013025)
Same Region of High School	0.00548	Both High School in Ile de France	-0.00336
	(0.00631)		(0.00776)
		Same Program	0.00477
			(0.00481)
Observations		52,326	
R-squared		0.0004	
F-stat		0.61	

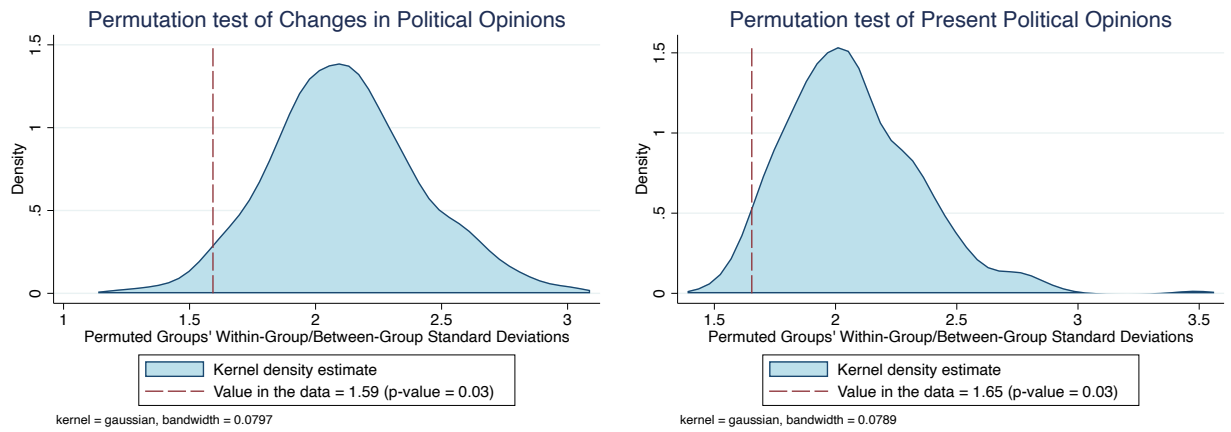
Notes: Balance test by OLS regression of *Same integration group* on all covariates altogether. F-stats are for the joint significance of the covariates. Standard errors in brackets are clustered by individual 1's group × individual 2's group. See Appendix A and Appendix Table A1 for description of variables and sample.

Panel B: Balance Test by Separate Regressions

Dependent Variable	Same integration group	Dependent Variable	Same integration group
Same Gender	0.000665	Same High School Major	-0.00221**
	(0.001134)		(0.00112)
Both Female	-0.00145	Diff. in Tuition Fees ('000 Euros)	-0.000243
	(0.00263)		(0.000391)
Same Nationality	-0.00826	Both Free Tuition	-0.00110
	(0.00749)		(0.00208)
Same Admission Type	0.000300	Same Parents' Profession	0.00138
	(0.003004)		(0.00217)
Both Affirmative Action	0.00983	Same ZIP Code	-0.000645
	(0.01500)		(0.004066)
Same Département of High School	0.00452	Both living in Paris	0.000247
	(0.00722)		(0.001343)
Same Region of High School	0.00305	Both High School in Ile de France	-0.00263
	(0.00361)		(0.00401)
		Same Program	0.00460
			(0.00483)
Observations		52,326	

Notes: Balance test by OLS regression of *Same integration group* on each covariate separately. Standard errors in brackets are clustered by individual 1's group × individual 2's group. See Appendix A and Appendix Table A1 for description of variables and sample.

Figure 2: Permutation Tests of Integration Group Effects on Political Opinions



(A) On Changes in Political Opinions

(B) On Present Political Opinions

Notes: Permutation tests of the effects of integration group assignment. In each case, we permute individuals' group assignment across the sample 300 times, and compute the distribution of the test statistic of the ratio between within-group and between-group standard deviations of the outcome. We then calculate the p-value of a left-sided test as the quantile of the observed sample's test statistic with respect to this distribution over permutations. The outcome in Figure 2A is the change of political opinions from before Sciences Po (August 2013) to March 2014, and that in Figure 2B is individual political opinion surveyed in March 2014.

of weakly converging (column 1) by 4.5%, and reduces the chance of diverging, both weakly (by 4%) and strongly (by 2%) (columns 3 and 4). In proportion to the incident of each type, the effect is particularly strong for strong divergence, as the effect's magnitude is 57% of the share of strongly diverging pairs.

Those results are further strengthened in Panel B, where we use the indicator whether a pair (i, j) belong to the same hypothetical group created by alphabetical order $SameHypGroup_{ij}$ as instrument for the same-integration group indicator IG_{ij}. The corresponding estimates across all columns are rather similar (with the exception of the small, insignificant effect on strong convergence), highlighting the reliability of the OLS estimates in Panel A. Overall, the exposure to the same integration group makes pairs of students more likely to converge and less likely to diverge.

Beyond the direction of change, Table 5 examines the impact of the indicator of being the same integration group IG_{ij} on the change in the opinion gap between each pair $ChgD_{ij}$, as defined in subsection 3.1. Panel A presents the OLS estimate in the full sample (column 4) as well as its components across subsamples of pairs that converge, diverge, or move in the same direction (columns 1 to 3).[24] The overall effect is strong and statistically significant, amounting to 11% of the dependent variable's standard deviation. The component effect among diverging pairs is the

[24]Since the sample is conditioned on outcome variables in Table 5's columns 1 to 3, the corresponding estimates should be taken with caution. They are mostly useful to show the different parts of column 4's full-sample estimate.

Table 4: Effects of Integration Group on Movement of Opinion Pairs

Panel A: OLS specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0447***	0.0095	-0.0396***	-0.0216***	0.0131
	(0.0161)	(0.0114)	(0.0154)	(0.0063)	(0.0128)
R-Squared	0.0142	0.0018	0.0048	0.0071	0.0048

Panel B: IV specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0371*	-0.0079	-0.0388*	-0.0292***	0.0109
	(0.0215)	(0.0112)	(0.0199)	(0.0109)	(0.0241)
R-Squared	0.0142	0.0018	0.0048	0.0071	0.0048
First Stage:					
Instrumental Variable:	0.5757***	0.5757***	0.5757***	0.5757***	0.5757***
Same Hypothetical Group	(0.0401)	(0.0401)	(0.0401)	(0.0401)	(0.0401)
Kleibergen-Paap Weak IV F-stat	205.9	205.9	205.9	205.9	205.9
Panels A & B's common features					
Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.517	0.0968	0.228	0.038	0.182
Std. Dev. (Dep. Var.)	0.500	0.296	0.419	0.191	0.386

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the same integration group indicator. In Panel B, the same-integration group indicator is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

strongest: Not only does exposure to the same integration group reduce a pair's probability of diverging (Table 4), but it also reduces the magnitude of divergence if they do diverge (the effect's magnitude of 0.17 is 11% of the average increase in the opinion gap in this case). The effect is similarly large among pairs that co-move, but small and insignificant among pairs that converge.

Those results are further strengthened in Table 5's Panel B, where the same integration group indicator IG_{ij} is instrumented by the same hypothetical group indicator $SameHypGroup_{ij}$ (columns 1 to 4), with a slightly smaller overall estimate (column 4). In addition, in column 5 we use the alphabetical distance between a pair's last names in the entire cohort of 2013 as another instrument for IG_{ij}. To further address potential confounders of alphabetical ranking, we also control for pairwise difference in alphabetical ranking based on a larger sample of Sciences Po's cohorts of 2009 to 2014. Last, in column 6 we limit the sample to pairs within a short band of alphabetical distance of at most 24. Those robustness checks all produce estimates similar to the baseline result Panel A's column 4, within a range from 0.12 to 0.21 (45% to 78% of the mean dependent variable, and 8% to 15% of its standard deviation).

Table 5: Same Integration Group Membership and Changes in Political Opinion Gaps

Panel A: OLS specification

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	OLS			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Same Integration Group	0.0264	-0.1678***	-0.1523**	-0.1584***
	(0.0526)	(0.0536)	(0.0731)	(0.0509)
R-squared	0.0251	0.0226	0.0056	0.0053

Panel B: Robustness with IV & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Full	Close Alphabetical Ranks
Same Integration Group	0.0969	-0.2240**	-0.0936	-0.1192**	-0.2091***	-0.1498**
	(0.0627)	(0.0946)	(0.110)	(0.0528)	(0.0693)	(0.0593)
R-squared	0.0250	0.0225	0.0055	0.0053	0.0053	0.0058
First Stage:						
Instrumental Variable:						
Same Hypothetical Group	0.5746***	0.5703***	0.6298***	0.5757***		
	(0.0432)	(0.0560)	(0.0627)	(0.0401)		
Alphabetical Distance					-0.0294***	
					(0.00153)	
Kleibergen-Paap Weak IV F-stat	177.3	103.6	100.8	205.9	367.44	

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,075	11,918	9,519	52,326	52,326	4,268
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	-1.151	1.474	0.000210	-0.267	-0.267	-0.281
Std. Dev. (Dep. Var.)	1.034	0.730	0.834	1.415	1.415	1.393

Notes: This table shows dyadic specifications of the effect of being in the same integration group on Changes in Political Opinion Gaps, estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). Panel B's columns 1 to 4 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 5 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and control for the pairwise alphabetical distance within an extended sample of last names of all students that entered Sciences Po from 2009 to 2014. Column 6 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Persuasion rate. We take DellaVigna and Gentzkow's (2010) approach to calculate the persuasion rate, as previously done in Harmon et al. (2019), to compare our results with others estimated in similar contexts. First, we recode the dyadic outcome into a binary variable that simply signifies a pair's agreement when their opinion gap is 0 or 1 versus disagreement when their opinion gap is 2 or greater.[25] Based on this definition, the baseline rate of disagreement among the untreated group ($IG = 0$) is close to 60%.

Regressing the binary variable $Agreement_{ij}$ on IG_{ij}, instrumented by the indicator of being in the same hypothetical group, and the same set of covariates, we obtain a treatment effect of 0.034. This implies a persuasion rate of 5.7%. This number is very close to the "reduced form" estimated persuasion rate in Harmon et al.'s (2019) (their Table 4) of 5.4%-5.5%, and slightly smaller than their 2SLS-estimated persuasion rate of 7%. It is remarkable to find such similarity based on rather similar treatments conducted on two completely different contexts and pools of individuals.

4 How the integration group's effect works through friendship

4.1 Survey of friendship

To understand the role of friendship in shaping the same-integration group effect, in the same survey in March 2014, we measure the undirected friendship link $Link_{ij}$ as the indicator whether either of the two students names the other as a friend.[26]

We offer strong material incentives in the survey in the form of a lottery for 50 mini iPads at approximately 300 Euros each, with an average probability of about 9% to win one. We seek a high rate of participation to avoid the problem of complex biases in network measures due to missing information on network structure (Chandrasekhar and Lewis, 2011). Eventually, 68.4% (547 out of 800) of the students answer to at least some question in the survey, and 65.6% (526 out of 800) complete the whole survey. This participation rate is similar to some of the most participated studies of social networks of students, such as Leider et al. (2009, 2010) or Goeree et al. (2010). It is well above the standard participation rate of around 20% found in studies using online surveys (Cantoni et al., 2017).

A follow-up survey conducted in June 2015 on the same cohort is unfortunately much less well-funded, and attracted a much smaller matched sample. Since the long-run dyadic sample size

[25] Note that the agreement relation defined as such is not transitive.

[26] This definition specifies the OR network of undirected friendship, similar to Leider et al. (2009, 2010) and other papers that consider surveyed friendship. The results remain robust to using the AND network, which counts a link between i and j if both list the other as friend, as shown in Appendix Tables A22 to A24 that replicate Tables 7, 8, and 9 on the AND network.

is reduced 5 times compared with the first survey's sample, we only use it in robustness checks.

In order to incentivize truthful answers, we design the elicitation of friendships as a coordination game, similarly to Leider et al. (2009, 2010). We ask students to name a list of up to 10 friends in the same cohort, and also details on how they meet each of them, how much time they spend together, and in which activities, as well as how strong they evaluate their relationships. We announce in the survey that their answers would be cross-validated with those of their named friends, and that if the two sets of answers match sufficiently, each would gain points, later converted into an additional probability of winning the iPad. When a respondent starts typing some characters of a friend's name, the survey displays a dropdown list of names in the same cohort that match those characters, so as to facilitate the input of long, unfamiliar, and hard-spelling names. Those details are designed to (i) encourage respondents to list all their friends, including those whose names can be long, uncommon, unfamiliar, and hard to spell, such as students from immigrant origin, (ii) encourage them to list their strongest friendships first, as those friends are most likely to list them back,(iii) discourage them from listing non-friends, as it is highly unlikely that non-friends reciprocate with cross-validated answers, and (iii) discourage respondents from overlooking the friendship questions just because they are more time-consuming.

In this design, the cross-validated incentive may raise the concern of collusion among respondents to maximize their gains. From our interaction with students, we believe this possibility is rather rare. First, student pairs who succeed in coordinating their answers are likely already friends in some way, in which case their coordination cements the validity of the friendship measure. Second, the survey is carried out during a vacation week, which limits the possibility that two students interact in person and complete the survey together. Third, since students only know the content of the questions once they open the survey website, and those who coordinate must spend much time to call each other and agree on a strategy, we further avoid potential colluders by censoring the top 5% of the sample by the amount of time spent on friendship questions.[27]

The survey incentive also pushes students to exhaust all 10 name slots. This may be problematic in two directions: students list many "mere acquaintances", namely relationships that do not matter as much as friendship (9% in our sample), and they may not be able to list all of the real friendships if there are more than 10 of them. In our subsequent analysis, we consider this issue as one of omitted

[27]This means dropping individuals who spend more than 81.625 seconds per friend on that question. Right tail truncation looks necessary, given that at the top of the distribution certain students spend up to half an hour per friend. Appendix Tables A25 to A28 show similar results of the main specifications in case of no truncation and two-sided truncation (2.5% each side).

relationships that can matter to outcomes but that are likely less important than the reported ones.

4.2 Empirical design to study friendship and opinions

4.2.1 Friendship and integration groups

We explore the role of pairwise friendship $Link_{ij}$ in two empirical exercises. First, in subsection 4.4, we explore the effect of participating in the same integration group on the chance that two students i and j become friends. The regression of $Link_{ij}$ on IG_{ij}, controlling for $\mathbf{X}_{ij}$, estimates $\beta^L_{IG} = \mathbb{E}[Link_{ij}|IG_{ij} = 1, \mathbf{X}_{ij}] - \mathbb{E}[Link_{ij}|IGij = 0, \mathbf{X}_{ij}]$.

4.2.2 Political opinion and friendship

Second, we consider the impact of friendship on a pair of students' change in political opinion gap. We consider the following linear dyadic specification:

$$ChgD_{ij} = \alpha + \zeta IG_{ij} + \beta_L Link_{ij} + \psi \mathbf{X}_{ij} + \eta_{ij}, \quad \mathbb{E}[\eta_{ij}|IG_{ij}, \mathbf{X}_{ij}] = 0. \tag{1}$$

The effect of friendship on the changes of students' political opinion gap β_L is expected to be negative, i.e., friendship causes more similarity between friends' opinions.[28] The coefficient ζ denotes the effect of being assigned to the same integration group beyond what works through friendship—its absence would mean the excludability of IG_{ij} in equation (1).

Homophily bias. The most important barrier to the identification of β_L in equation (1) is the potential homophily bias due to a certain unobserved factor U_{ij} such that (i) individuals' similarity U_{ij} correlates with the formation of friendship links $Link_{ij}$ (homophily), and (ii) it directly influences the outcome DY_{ij} through η_{ij} (outcome-relevance). In the standard case of homophily by political opinion, as politically similar students are more likely to become friends (Lazarsfeld and Merton, 1954; McPherson et al., 2001), one should expect a negative correlation between $Link_{ij}$ and the error term η_{ij}, leading to a bias of β_L away from zero.[29]

Same integration group membership as instrument for friendship. We can address this homophily bias and identify β_L by using IG_{ij} as instrument for $Link_{ij}$ under the standard assumptions of relevance and exclusion stated below:

Assumption 1 (Relevance) $\mathbb{E}[L_{ij}|IG_{ij} = 1, \mathbf{X}_{ij}] \neq \mathbb{E}[L_{ij}|IG_{ij} = 0, \mathbf{X}_{ij}]$, *so* IG_{ij} *predicts* L_{ij}.

[28] For simplicity, we take the IV approach in a linear model with a homogenous effect β_L.

[29] The size of the bias is proportionate to how much the omitted variable U_{ij} matters to the outcome, and its correlation with $Link_{ij}$.

Assumption 2 (Exclusion) *$\zeta = 0$, so IG_{ij} does not directly affect DY_{ij}.*

The relevance of Assumption 1 will be established in subsection 4.4. The key exclusion of Assumption 2 requires that IG_{ij} does not affect the differences in political opinions through any other channel beyond friendship, in which case equation (1) is simplified as:

$$ChgD_{ij} = \alpha + \beta_L Link_{ij} + \psi \mathbf{X}_{ij} + \eta_{ij}. \tag{2}$$

4.2.3 Assessment of the exclusion restriction

The structure of the integration groups provides some support for the exclusion assumption. The integration week was designed solely to help students acclimate to their new peers and environment in Paris, without any academic or political activities. After the week concludes, integration groups are dissolved and play no role in any subsequent academic or extracurricular activities. Additionally, no other large-scale events at Sciences Po are organized based on alphabetical order.

A major remaining concern is that individuals who do not declare each other as friends may still have influenced one another. In other words, there may be weaker relationships—undocumented in the survey—that are nonetheless significant in shaping political opinions. This suggests the presence of an unobserved variable, L^0_{ij} which correlates with IG_{ij} and affects $ChgD_{ij}$, but is omitted from the right-hand side of equation (2).[30]

Non-excludability bias. To further examine the implications of this possibility, we elaborate specification (2) to include L^0_{ij}:

$$ChgD_{ij} = \alpha + \beta_L Link_{ij} + \beta_0 L^0_{ij} + U_{ij} + \varepsilon_{ij}. \tag{3}$$

We introduce three unobservable terms. First, U_{ij} captures the endogeneity of friendship, as it may correlate with $Link_{ij}$. Second, L^0_{ij} accounts for potential violations of the exclusion assumption (2), representing unobserved dyadic relationships influenced by the instrument IG_{ij} that may directly affect the outcome $ChgD_{ij}$. Third, the centered idiosyncratic error ε is assumed to be uncorrelated with $Link_{ij}$ (i.e., $\mathbb{E}[\varepsilon Link_{ij}] = 0$). For simplicity, we omit control variables $\mathbf{X}_{ij}$.[31] When the exclusion assumption is violated, $\beta_0 \neq 0$, resulting in a non-excludability bias in the estimation of β_L.

[30] In addition, there may be unobserved animosity between certain pairs, which may also correlate with their pairwise outcomes. If we expect enemies to likely move in opposite directions, this kind of omitted unobservable would contribute negatively to the main estimate of the friendship effect, thus making the estimate weaker than it truly is.

[31] In this subsection's analysis, we can partial out $\mathbf{X}_{ij}$ from all other variables (i.e., subtract each variable's linear projection on $\mathbf{X}_{ij}$), preserving all results. In practice, we include the full set of controls $\mathbf{X}_{ij}$ in all regressions.

In the spirit of Altonji et al. (2005) and Oster (2019), in Appendix B we evaluate the size of the potential non-excludability bias using interpretable statistics based on relations between unobservables and observables in this context. Our analysis relies on the equivalence between the IV 2SLS estimator and a control function (CF) estimator to overcome the nonlinear nature of the 2SLS estimator. We obtain the following IV-CF estimator of β_L:

$$\widehat{\beta}_L^{IV} = \beta_L \left(1 + \frac{\beta_0 \pi_0}{\beta_L \pi_L}\right), \tag{4}$$

with a bias to true value ratio of $\frac{\beta_0 \pi_0}{\beta_L \pi_L}$, where π_L and π_0 are respectively first-stage coefficients when $Link_{ij}$ and L_{ij}^0 are regressed on IG_{ij}. This bias ratio represents the relative importance of the omitted channel through unsurveyed links L_{ij}^0 versus the modeled channel through declared friendship links $Link_{ij}$ in terms of the influence from IG_{ij} on the outcome DY_{ij}. It is null if the exclusion restriction is valid ($\beta_0 = 0$). When the exclusion restriction is invalid, it is likely positive, leading to an overestimation of the coefficient β_L.

Following Altonji et al.'s (2005) heuristic argument, one may consider a similarly simple argument that the channel through declared friendship likely has as much influence on opinions as the omitted channel.[32] Under this assumption, the bias ratio $\frac{\beta_0 \pi_0}{\beta_L \pi_L} \leq 1$, hence we can obtain a lower bound and an upper bound of β_L: $\beta_L \in \left[\frac{1}{2}\widehat{\beta}_L^{IV}, \widehat{\beta}_L^{IV}\right]$.

More conservatively, we may assume that the direct influence of declared friendships on the outcome $ChgD_{ij}$ is stronger than that of omitted relationships, namely $\frac{\beta_0}{\beta_L} \leq 1$. To shed light on $\frac{\pi_0}{\pi_L}$, we can estimate the various impacts π_k ($k \in [1, 4]$) of IG_{ij} on different declared types of friendship by intensity k (Appendix Table A8). We then assume that $\frac{\pi_0}{\pi_L}$ falls within the range of the highest ratio between those estimates π_k, denoted as $r = \max \frac{|\pi_k|}{|\pi_l|}$ ($k \neq l \in [1, 4]$). Under those heuristic assumptions, the bounds become $\beta_L \in \left[\frac{1}{r+1}\widehat{\beta}_L^{IV}, \widehat{\beta}_L^{IV}\right]$.[33]

Appendix B further presents a more refined analysis of the non-excludability bias using the surveyed intensity of friendship from 1 to 4. With such additional information, we can better assess the bias as a function of two statistics that relate to the relative importance of different levels of friendship intensity.

[32] In similar veins, Altonji et al. (2005) argues that observables that have been selected through the process of survey design are likely a more reliable source of predictors of outcomes than unobservables. Oster (2019) later quantifies this argument in terms of the comparison between those two sets of variables.

[33] Appendix Table A8 presents the estimates of those β_k from our survey. Among OLS estimates of β_k's the maximal ratio is $r = 2.70$, and $r = 1.97$ among estimates using $SameHypGroup_{ij}$ as instrument for IG_{ij}.

4.3 Description of friendship data

We consider the (symmetric) OR network in which two students are linked if at least one nominates the other. Table 6 Panel A describes the quality of the network survey. About half of the nominated friends reciprocate, a considerably higher rate than typically reported in the literature (e.g., Calvó-Armengol et al., 2009, Leider et al., 2009). The probabilities of a well-matched response between two friends regarding the context of their first meeting (76%), the amount of time spent together each week (52%), the types of activities they engage in most frequently (46%), and the self-evaluated strength of their friendship (52%) are all significantly higher than those reported in Leider et al. (2009). If responses were entirely fabricated or randomized, the probability of matching on any of these dimensions would be considerably lower, given the wide range of possible answer choices—particularly for the question about the context of their first meeting. Collectively, these statistics suggest that the survey responses are highly reliable, especially for accurately identifying friendships.

Panel B reports the major statistics on the number of friends and the social network structure. The average and maximum number of nominated friends per student is 8.8 and 21, respectively, with a very high variance.[34] Moreover, there seems to be some small world properties with a very small average path length (3.7) and a relatively small diameter (9). The clustering coefficient is also relatively high, which means that roughly 25 percent of students have friends of friends who are friends. In terms of network position, the mean eigenvector centrality is relatively low (0.0361).

Panel C shows the descriptive statistics of the friendship dyadic measures. We distinguish between the full sample (column 1) of all students who have participated and the benchmark sample (column 2) that corresponds to the benchmark regression (the difference is due to certain missing values). By nature, the share of measured friendship links is relatively small at 1.6%, and that of second and third order indirect links are larger at 9.3% and 38%, respectively. The dyadic same group variables are of similar magnitudes, at an average of 1.6% for same-integration group membership, and 2.3% for same tutorial groups. The friendships are partitioned rather evenly across different levels of friendship strength, especially from 2 (ordinary friends) to 4 (very close friends). We also observe that there is little difference between the full sample and the benchmark sample.

[34] Even if the maximum number of friends that someone can nominate is 10, a student can have 21 friends since we use an undirected network approach so that a friend is assigned to a person if either her or her friend has nominated the other.

Table 6: Descriptive Statistics

Panel A: Quality of the Survey

	(1) Full Sample	(2) Benchmark Sample
Number of reported friends	8.234	8.613
	(2.522)	(1.984)
Probability of reciprocal friend	0.461	0.479
	(0.499)	(0.500)
Correct answer: meeting	0.800	0.815
	(0.400)	(0.389)
Correct answer : time spent	0.483	0.497
	(0.500)	(0.501)
Correct answer : activity	0.568	0.587
	(0.496)	(0.493)
Correct answer : strength of the relationship	0.532	0.532
	(0.499)	(0.500)

Panel B: "OR" Network Statistics

Mean of degree per individual	8.8625
Variance of degree per individual	18.4842
Median of degree per individual	10
Maximum of degree per individual	21
Minimum of degree per individual	0
Diameter of the network	9
Average path length	3.7008
Overall clustering coefficient	0.241
Average clustering coefficient	0.271
Mean eigenvector centrality	0.0361
Standard deviation of eigenvector centrality	0.0200

Notes: Summary statistics are computed on the full sample.

Panel C: Dyadic Links and Groups

Variable	(1) Full Sample			(2) Benchmark Sample		
	Mean	Standard deviation	Obs.	Mean	Standard deviation	Obs.
Friendship	0.0160	(0.1240)	147,153	0.0178	(0.1324)	52,326
2nd Order Links	0.0930	(0.2900)	147,153	0.1014	(0.3019)	52,326
3rd Order Links	0.3800	(0.4850)	147,153	0.4081	(0.4914)	52,326
Mere relationship (strength 1)	0.0014	(0.0382)	147,153	0.0018	(0.0428)	52,326
Friendship link (strength 2)	0.0063	(0.0791)	147,153	0.0070	(0.0832)	52,326
Close friendship (strength 3)	0.0041	(0.0642)	147,153	0.0047	(0.0681)	52,326
Very close friendship (strength 4)	0.0035	(0.0593)	147,153	0.0041	(0.0641)	52,326
Same Integration Group	0.0160	(0.1280)	147,153	0.0188	(0.1359)	52,326

Notes: Statistics in (1) are computed on the full sample of data available for each variable, while statistics in (2) are computed on the benchmark sample, which is detailed in Table A1.

4.4 Same-integration group exposure and friendship formation

We now proceed to estimate the causal effect of participating in the same integration group on forming and maintaining a lasting friendship 6 months later, which would confirm the relevance of the instrumental variable IG_{ij} in the strategy described in subsection 4.2. Columns 1 and 2 of Table 7 present the regression of $Link_{ij}$ on IG_{ij}, with and without observable covariates $\mathbf{X_{ij}}$, yielding an estimate of $\beta_{IG} = \mathbb{E}[Link_{ij}|IG_{ij} = 1, \mathbf{X_{ij}}] - \mathbb{E}[Link_{ij}|IGij = 0, \mathbf{X_{ij}}]$ of around 17%. Columns 3 to 5 report robustness checks, as described in subsection 3.3 and performed in subsection 3.7. The estimates remain largely stable across those columns, using either the same-hypothetical group instrument, the alphabetical distance instrument, or the quasi-RD design that focuses on pairs of students within a close alphabetical distance of each other.

This estimate is remarkably large in comparison with other coefficients of the included control variables, as shown in detail in Appendix Table A7. It is indeed more than 10 times larger than any coefficient on observable predetermined characteristics (the next largest coefficients are on students' ZIP code).[35] It shows that "exposure by chance" to other students during the first week

[35] All covariates are binary variables, thus their coefficients are easily comparable, with the exception of the difference in tuition fees, our best proxy for the difference in parents' income brackets. Each student's tuition fees range from zero (i.e., full scholarship) to 10,000 euros (full tuition), with a mean of 3,900 and standard deviation of 3,000 euros.

of a student's college life has an effect on friendship formation several orders of magnitude larger than that of most typical predetermined characteristics obtained from administrative records.

Following Table 7, we can further explore the effect of the same integration group membership on specific features of a friendship link. Those effects are shown respectively by the four degrees of friendship intensity as provided in the survey (from mere acquaintance to very close friends) in Appendix Table A8, by the type of the most-frequent activity the pair spend together (such as academic or leisure) in Appendix Table A9, and by the range of time they typically spend together per week in Appendix Table A10. Those results demonstrate the broad basis of friendship links caused by the same integration group membership.

The result can be interpreted as evidence of the first week's special role as a "window of opportunity" for friendship formation. It hints that friendships tend to form at the beginning of college, in activities meant to facilitate socialization with same-cohort peers, and familiarization with a completely new environment, when almost all students still have no friends there yet. It is all the more striking that those friendships can last much longer beyond the window of opportunity, even when the special exposure ends right after this window, and all students become fully exposed to the whole cohort.

While we also find statistically significant evidence of homophily based on some predetermined characteristics, its role is rather limited in comparison with the effect of integration group exposure. The inclusion of those covariates hardly alters the coefficient of IG_{ij}.

4.5 Friendship and students' political opinions

We move on to examine friendship as the channel through which belonging to the same integration group affects students' political opinions, using the method described in subsection 4.2. We carry over the settings of Tables 4 and 5 in subsection 3.7 to Tables 8 and 9 below, in which the main treatment variable, friendship link, is now instrumented by the same-integration group indicator. Appendix Tables A13 and A14 show all specifications that follow Tables 4 and 5, including those in which friendship is instrumented by the same hypothetical group indicator, by alphabetical distance, and in the quasi-RD setting by alphabetical order.

Similar to the effect of belonging to the same integration group, as shown in Table 8, friendship increases the likelihood of convergence and decreases the likelihood of divergence between students. In particular, friendship has a strong effect in reducing strong divergence—where both students move further apart—lowering its occurrence by a factor of 3.5 relative to the average probability of strong divergence. The friendship effect is considerably higher than that of the same integration

Table 7: Same Group Membership and Friendship Formation

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Friendship**				
Specification:	OLS		IV		Quasi RD
Instrumental Variable:			Same Hyp. Group	Alpha. Distance	
Sample:	Full		Full		Close Alpha. Ranks
Same Integration Group	0.1660***	0.1647***	0.1784***	0.1684***	0.1701***
	(0.0186)	(0.0186)	(0.0245)	(0.0197)	(0.0247)
Controls	No	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	4,268
Number of Integration Groups	52	52	52	52	52
R-Squared	0.029	0.064	0.064	0.064	0.129
Kleibergen-Paap Weak IV F-stat	79.645	78.380	52.893	73.017	50.139
Mean (Dep. Var.)	0.018	0.018	0.018	0.018	0.052
Std. Dev. (Dep. Var.)	0.132	0.132	0.132	0.132	0.222

Notes: This table shows dyadic specifications of the effect of being in the same integration group on friendship formation. Column 3 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 4 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and control for the pairwise alphabetical distance within an extended sample of last names of all students that entered Sciences Po from 2009 to 2014. Column 5 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the standard set of controls.

Table 8: Friendship and Movement of Opinion Pairs

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Friendship	0.2714***	0.0574	-0.2405**	-0.1314***	0.0793
	(0.0988)	(0.0698)	(0.0949)	(0.0377)	(0.0779)
Bounds	[0.0914, 0.2714]	[0.0193, 0.0574]	[-0.2405, -0.0810]	[-0.1314, -0.0442]	[0.0267, 0.0793]
R-Squared	0.0096	0.0011	0.0001	0.0007	0.0038
First Stage:					
Instrumental Variable:	0.1646***	0.1646***	0.1646***	0.1646***	0.1646***
Same Integration Group	(0.0186)	(0.0186)	(0.0186)	(0.0186)	(0.0186)
Kleibergen-Paap Weak IV F-stat	78.381	78.381	78.381	78.381	78.381
Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.517	0.0968	0.228	0.038	0.182
Std. Dev. (Dep. Var.)	0.500	0.296	0.419	0.191	0.386

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the friendship indicator. Friendship is instrumented by the same-integration group indicator. The bounds are calculated as $\beta_L \in \left[\frac{1}{1+r}\widehat{\beta_L}, \widehat{\beta_L}\right]$, with $r = 1.97$, as discussed in subsection 4.2.3. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

group indicator, even after we take into account the bounds of $\beta_L \in \left[\frac{1}{2.97}\widehat{\beta_L}, \widehat{\beta_L}\right]$ as discussed in subsection 4.2.3.

Overall, Tables 8 and 9 indicate the particularly strong role of friendship as a barrier against divergence of opinions. Not only does it reduce a pair's probability of diverging, but it also reduces the magnitude of divergence of they do diverge.

In further robustness checks that echo concerns about integration group assignments previously discussed in subsection 3.2, Appendix Tables A13 and A14 show more specifications that follow Tables 4 and 5 by instrumenting friendship by the same hypothetical group indicator and by alphabetical distance, and in the quasi-RD setting by alphabetical order. All results remain quantitatively close to those shown in Tables 8 and 9.

How much could friendships have contributed to the reduction of the average pairwise opinion difference in the sample, from 2.211 before Sciences Po to 1.932 at the survey? Per dyad, there is on average 0.0178 friendships, so an effect of -0.962 can explain $\frac{0.962 \times 0.0178}{2.194 - 1.927} \sim 6.4\%$ of the change in total pairwise differences (Table 6 Panel C and Appendix Table A3 Panel B). This modest proportion is mostly due to the very low frequency of direct friendships in the dyadic sample.

Finally, we examine whether our results vary based on the initial political views of student pairs. Figure 3(A) presents the heterogeneous effects of friendship on political opinions among pairs with different starting positions. Figure 3(B) illustrates the impact of friendship on the *sum* of the pair's individual political opinions, categorized by their initial stance. This analysis captures the direction of political convergence. A positive estimate indicates that a pair of friends with a given initial

Table 9: Friendship and Changes in Political Opinion Gaps

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Friendship	0.1640	-1.0531***	-0.8592**	-0.9618***
	(0.324)	(0.311)	(0.397)	(0.314)
Bounds	[0.0552, 0.1640]	[-1.0531, -0.3546]	[-0.8592, -0.2893]	[-0.9618, -0.3238]
R-squared	0.0248	-0.0012	-0.0094	-0.0011
First Stage:				
Instrumental Variable:				
Same Integration Group	0.1610***	0.1594***	0.1772***	0.1647***
	(0.0266)	(0.0292)	(0.0321)	(0.0186)
Kleibergen-Paap Weak IV F-stat	36.68	29.89	30.56	78.38
Controls	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes
Observations	27,075	11,918	9,519	52,326
Number of integration groups	52	52	52	52
Mean (Dep. Var.)	-1.151	1.474	0.000210	-0.267
Std. Dev. (Dep. Var.)	1.034	0.730	0.834	1.415

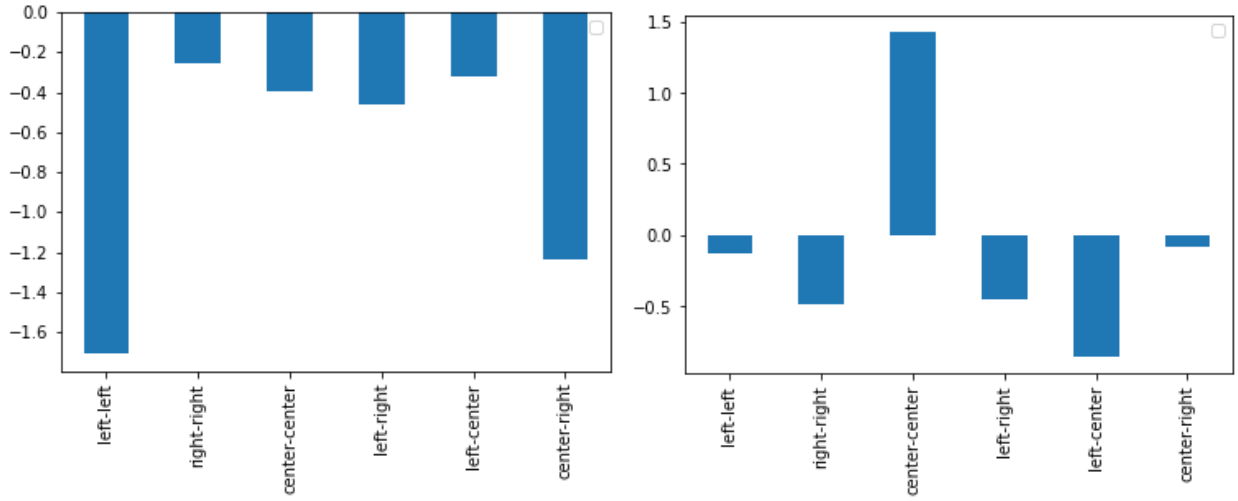
Notes: This table shows dyadic specifications of the effect of friendship on Changes in Political Opinion Gaps, estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). Friendship is instrumented by the same-integration group indicator. The bounds are calculated as $\beta_L \in \left[\frac{1}{1+r}\widehat{\beta_L}, \widehat{\beta_L}\right]$, with $r = 1.97$, as discussed in subsection 4.2.3. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

position, on average, shifts toward the right of the political spectrum, while a negative estimate suggests a shift to the left. Our findings suggest that convergence is strongest among left-left and center-right pairs, with both groups gravitating toward the middle of their initial positions. This is further supported by the small and statistically insignificant effect of friendship on the sum of their political opinions. Center-center pairs tend to shift collectively to the right, while all other types of pairs converge toward the left.[36] Overall, we find no evidence that friendship, on average, drives individuals toward more extreme political positions.

Longer term effects. We rerun the survey in June 2015 on a reduced budget, and obtain a much reduced dyadic sample. We estimate the effects of the same-integration group indicator and of friendship on pairwise opinion movements and the change in pairwise opinion gaps, and report them in Appendix Tables A15 and A16. While the small sample size has much reduced the estimates' precision, the magnitude of the effects of being in the same integration group and being friends on the change in pairwise opinion gaps remains quite comparable with the effects found in Tables 5 and 9. Thus, one year later, there is no evidence that the main effects have faded away.

[36]We also explore other dimensions of heterogeneity, particularly whether the impact of friendship on political opinions differs between same-gender and mixed-gender pairs. Interestingly, the effect is stronger for mixed-gender pairs, where the reduction in the political opinion gap reaches 1.4 points, compared to 0.62 points for same-gender pairs.

Figure 3: Heterogeneous effects of friendship by pairs of initial political opinions



(A) Effect on the pair's political opinion gap

(B) Effect on the pair's sum of political opinions

Notes: This figure shows the heterogeneous effects of friendship on students' political opinions in terms of student pair's initial political opinions.

Persuasion rate. Similar to subsection 3.7, we can use DellaVigna and Gentzkow's (2010) and Harmon et al.'s (2019) approach to calculate the persuasion rate based on the treatment effect of friendship on the binary outcome of pairwise agreement. The regression of $Agreement_{ij}$ on friendship, instrumented by the indicator of being in the same hypothetical group, yields a treatment effect estimate of 0.188, corresponding to a friendship persuasion rate of 31%, and with a lower bound of 10.4% as calculated using the method in subsection 4.2.3.

4.6 Discussion

We draw three major conclusions from these results. First, friendship appears to have a strong effect in reducing opinion gaps between students, particularly by reducing the probability and magnitude of divergence. While estimating the effect of friendship in subsection 4.5 requires considerably stronger assumptions than estimating the effect of belonging to the same integration group in subsection 3.7, the evidence suggests that friendship effects are robust enough to remain meaningful even if these assumptions are violated within certain bounds.

Second, it is crucial to consider the nonlinearity of friendship's effect on political opinions, as its impact may depend on the direction of opinion change. This finding challenges the common assumption of homogeneous, linear effects of direct social links on belief formation, as typically modeled and estimated in the theoretical and empirical literature on non-Bayesian learning in networks (Möbius and Rosenblat, 2014). Examples include theories based on average-based belief

Table 10: Friendship, Integration Group, and Association Activities

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Both Are Members of the Same Association in						
	Any Area		Politics		Humanitarian	Identity	Sports Clubs
Specification:	OLS	IV	OLS	IV	OLS	OLS	OLS
Same Integration Group	0.0306*		0.0229*		-0.0020	0.0008	0.0071
	(0.0157)		(0.0127)		(0.00420)	(0.00320)	(0.0186)
Friendship		0.1761**		0.1314*			
		(0.0826)		(0.0743)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weak IV Test Statistic		38.562		38.562			
Observations	23,436	23,436	23,436	23,436	23,436	23,436	23,436
Number of IGs	52	52	52	52	52	52	52
R-Squared	0.0071	0.0179	0.0060	0.0186	0.0111	0.0022	0.0138
Mean (Dep. Var.)	0.097	0.097	0.023	0.023	0.008	0.004	0.586
Std. Dev. (Dep. Var.)	0.296	0.296	0.149	0.149	0.088	0.061	0.493

Notes: This table shows dyadic specifications relating the indicator of being members of the same association with the same-integration group indicator (columns 1, 3, 5-7) and friendship (columns 2 and 4), where friendship is instrumented by the same-integration group indicator. Columns 3 and 4 focus on associations related to politics and policies, column 5 on associations with humanitarian and social purposes, column 6 on identity-related associations, and column 5 on sports clubs. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, including association categorization, and the set of controls.

updating processes, such as Golub and Jackson's (2012)'s generalized framework of DeGroot's (1974) belief updating, as well as alternative updating mechanisms (Molavi et al. 2018, Campbell et al. 2025).

4.7 Effects on students' participation in associations

We proceed to study whether the uncovered effects on students' political opinion also manifest in their behaviors. We focus on their participation in students' political associations, the most important type of extra-curricular activities at Sciences Po.

In our survey, students reported being involved in a total of 90 different associations, with 59.5% declaring participation in at least one. Among those involved, 37.2% are members of at least one association related to policy debates (including political parties and student unions), a figure closely aligned with the most recent official report from Sciences Po Student Life (40%). Additionally, 17.3% participate in humanitarian associations focused on social and environmental causes, 22.3% are involved in identity-based associations (including women's rights, LGBTQ+, anti-racism, national, and religious groups), and 27.6% take part in sports associations.

Table 10 presents the effects of the same-integration group indicator and friendship on the likelihood that a pair of students enrolls in the same organization, using both OLS and IV estimation strategies (as described in subsections 3.2 and 4.2).[37]

[37]We do not observe the intensity of participation and can only analyze the extensive margin. Another limitation is that formal political party enrollment remains rare among first-year students. Therefore, we use participation in politically related associations as a proxy instead of direct political party membership or voting behavior. Additionally, most students in our sample had not yet reached voting age in previous elections.

Being in the same integration group (column 1) and being friends (column 2) increase the chance of a pair of students joining at least one common (non-sports) association by respectively 3% and 18%. Those effects are sizable in comparison with the average of the dyadic outcome of 8.5% (Table 6), and different from zero at 5% statistical significance. The effect is primarily driven by common membership in political associations, with the corresponding estimates of 2.3% (columns 3) and 13.1% (column 4).

For other categories of associations, including associations with humanitarian and social agendas (column 5) and associations defined based on students' origins and identity (e.g., those centered around a certain religion or an ethnic origin) (column 6), the OLS estimate of the same-integration group indicator is indistinguishable from zero. The estimate for common participation in a sports activity is slightly higher (yet a much smaller proportion of the mean outcome variable of 60%), but still statistically insignificant.[38]

In summary, these findings highlight the significant impact of exposure within integration groups and the influence of subsequent friendships on students' actual choices, extending beyond their self-reported beliefs. This effect is particularly pronounced in associations focused on policy debates, likely due to both the prevalence and nature of these organizations.

As noted earlier, the number of student associations dedicated to political debate is significantly higher than that of other types of associations. Consequently, being part of the same integration group may have a stronger influence on students' selection of a particular political debate association, given the wide range of options available.

Moreover, these associations tend to be more diverse and inclusive in their membership. Their primary objective is to facilitate discussions on national and international issues by inviting prominent figures such as senior civil servants, politicians, and business leaders. For example, the most prominent association, "Parlement des Étudiants", organizes debates featuring representatives from across the ideological spectrum.

In contrast, humanitarian associations serve a fundamentally different purpose, focusing primarily on civic action, such as outreach programs for the homeless or aid initiatives in African countries. Similarly, identity-based associations (e.g., religious groups or LGBTQ+ organizations) and sports associations typically require a certain level of personal commitment or interest prior to students' arrival at Sciences Po—for instance, a preference for Catholicism over Islam or soccer over rock

[38]Since the reduced form estimates in columns 5 to 8 are close to zero and statistically insignificant, the unreported corresponding IV estimates are even less precise.

climbing. As a result, the impact of integration groups on these specific choices is expected to be much smaller.

5 The homophily-enforced mechanism among similar students

In this section, we further investigate the mechanism behind the friendship effect by distinguishing between pairs of students according to their pre-Sciences Po opinion gap. We conjecture the "homophily-enforced mechanism" that friendship matters more to political opinions among pairs with similar initial opinions, because those pairs' interactions are more relevant to politics. This conjecture is equivalent to Harmon et al.'s (2019) finding of the complementarity between social and political proximities in terms of producing similar political outcomes.

To illustrate this mechanism, consider two pairs of students. The first pair, François (F) and Ségolène (S), enter Sciences Po with similar political opinions, while the second pair, Michel (M) and Dominique (D), hold very different political views.[39] Despite these differences, both pairs form friendships: F and S bond over their shared political interests, while M and D connect through other commonalities, such as a mutual love of literature.[40] Over time, these friendships shape their interactions at Sciences Po. As F and S engage more deeply in political discussions, their views become further aligned. Similarly, M and D's bond strengthens around their shared appreciation for the fine arts, reinforcing that common interest.

Consequently, friendship matters in binding the political views of pairs of friends whose initial views are already similar, while it has little effect on the political gap between those whose initial views are dissimilar.

5.1 Effects on political opinions among similar students

First, we document how friendships form based on initial political views—that is, how the gap in political opinions affects the likelihood of forming a friendship. Initially, 16.1% of pairs exhibited no political differences, while 66.8% had a difference of 2 or less. This indicates that the vast majority of friends initially held similar, though not identical, political views (e.g., left vs. center-left or center-right vs. right). Furthermore, 90.8% of friend pairs had an initial political difference of 4 or less. Notably, an initial gap of 3 or 4—observed in 24% of pairs—can be considered significant, as it typically reflects a contrast between far-left and left, far-right and right, or even left and right.

[39] More broadly, this mechanism can operate along any dimension that fosters friendship, not just political opinions. Moreover, there is no inherent relationship between the initial opinion gap and the estimated effects.

[40] Given that M and D become friends despite differing political views, homophily suggests they are more likely than F and S to share another non-political interest.

While there was substantial heterogeneity in political opinions among friends, extreme divergence was rare.

Additionally, the likelihood of forming friendships within the same integration group decreases as the initial political difference increases. The probability of becoming friends is 20.5% for pairs with no initial political difference, 18.2% for those with a difference of 1, 15.1% for a difference of 2, and 12.5% for a difference of 3. Beyond an initial gap of 4, the likelihood of friendship drops below 5% and continues to decline.

In line with the main findings of this study, the distribution of political differences shifts toward greater alignment by March 2014 among the pairs of friends. By then, 18.1% of pairs exhibited no political differences (against 16.1% initially), while 74.6% had differences of 2 or less (against 66.8% initially) and 95.2% had differences of 4 or less (against 90.8% initially).

Table 11 tests the homophily-enforced mechanism's implication on the effect of same-integration group exposure and friendship on changes in a pair's political opinion gap. We partition the sample into (i) pairs with a below-average (less than 1.9) initial opinion gap (columns 1 and 2) and (ii) pairs with an above-average initial opinion gap (columns 3 and 4). Columns 1 and 3 estimate the effect of the same-integration group indicator in an OLS specification as previously used in Table 5's Panel A. Columns 2 and 4 estimate the effect of friendship, instrumented by the same-integration group indicator, as previously used in Table 9.[41]

Comparing the two subsamples, the same-integration group effect in column 1 is more than 2 times larger than that in column 3. Similarly, the friendship effect in column 2 is 1.6 times larger than that in column 4. The effects among more politically similar pairs are thus both much more statistically and economically significant than those among more politically dissimilar pairs.

This pattern is further illustrated graphically in Figure 4, in which Panel A shows the effects of the same-integration group exposure by the value of a pair's initial opinion gap, and Panel B shows the corresponding effects of friendship from the regressions using the same-integration group indicator as instrument.[42]

The contrast between similar and dissimilar pairs in terms of initial political opinions is further demonstrated in the effects of being in the same integration group on the different types of movements

[41] It is important to take the estimates in this section with caution, because the subsamples are conditioned on recalled political opinions, which may suffer from a recall bias (discussed in subsection 3.1). Provided that the recall bias is not large and would only affect marginal pairs in the conditioned subsamples (i.e., those with initial opinion gaps around 2), this section's results can still be informative.

[42] Appendix Figure A4 further shows the corresponding estimates of the same-integration group indicator and of friendship in specifications using the same hypothetical group indicator as instrument for those treatment variables.

Table 11: Effects of Integration Group and Friendship by Initial Political Opinion Gaps

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion**			
Sample:	Initial Opinion Gap<2		Initial Opinion Gap≥2	
Specification:	OLS	IV	OLS	IV
Same Integration Group	-0.1728***		-0.0813	
	(0.0594)		(0.0606)	
Friendship		-0.8984***		-0.5471
		(0.3037)		(0.4216)
R-squared	0.0132	0.0038	0.0099	0.0089
First Stage:				
Instrumental Variable:		0.1924***		0.1487***
Same Integration Group		(0.0249)		(0.0203)
Kleibergen-Paap Weak IV F-stat		59.885		53.867
IV	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes
Observations	21,054	21,054	31,272	31,272
Mean (Dep. Var.)	0.548	0.548	-0.816	-0.816
Std. Dev. (Dep. Var.)	1.097	1.097	1.338	1.338

Notes: This table shows dyadic specifications of the effects of being in the same integration group and of friendship on Changes in Political Opinion Gaps. Columns 1 and 2 use the subsample of pairs that started out similarly before Sciences Po (with pre-Sciences Po opinion gaps of 0 or 1). Columns 3 and 4 use the subsample of pairs that started out differently before Sciences Po (with pre-Sciences Po opinion gaps of 2 or more). Friendship is instrumented by the same-integration group indicator. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

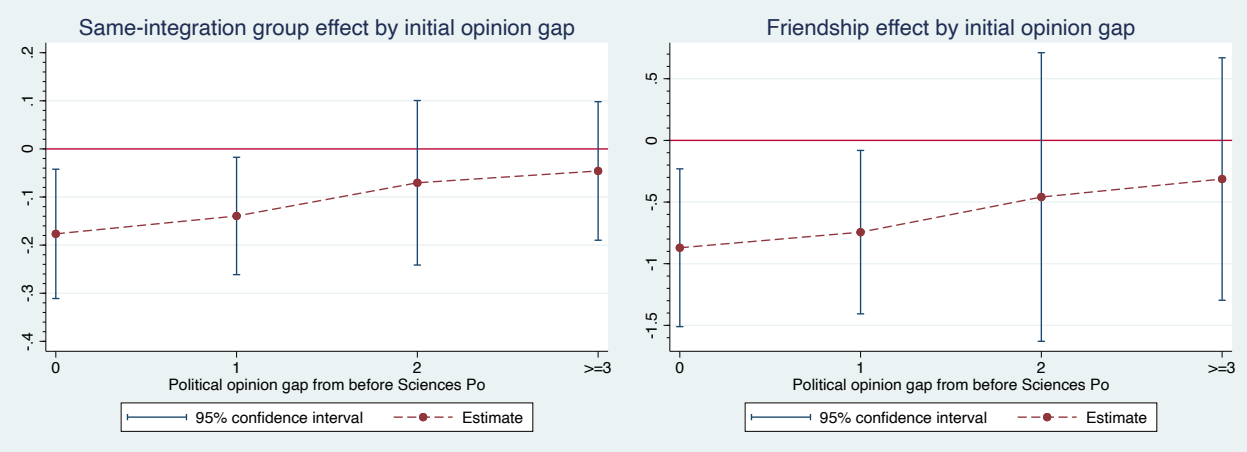
of opinions, as shown in Appendix Tables A17's and A18's replications of different strategies used in Table 4 for the subsamples of initially similar and dissimilar pairs, respectively. This contrast is also demonstrated in Appendix Tables A19's and A20's replications of Table 5 for the subsamples of initially similar and dissimilar pairs, respectively.

In sum, the evidence suggests that initial political proximity complements social proximity, such as shared integration groups and friendships, consistent with Harmon et al. (2019).

5.2 Similar students' participation in associations

As the homophily-enforced mechanism works through interactions on the dimensions that friends are more similar, it is useful to investigate whether pairs of students with initially more similar views tend to interact more in students' associations. In this direction, Table 12 estimates the effects of being in the same integration group (columns 1, 3, 5, 7) and being friends (columns 2, 4, 6, 8) on a pair's changes in political opinion gap. Among pairs that hold similar opinions from before Sciences Po (opinion gap of at most 1), column 1 shows that being in the same integration group increases a pair's chance of joining the same association by 5.2%, and column 2 shows that the effect of friendship is 23.4%. Those numbers are markedly larger than their statistically insignificant counterparts shown on columns 3 and 4, where the sample is limited to pairs of dissimilar views

Figure 4: Effects of Integration Group and Friendship by Initial Political Opinion Gap



(A) Effect of the same-integration group indicator (B) Effect of friendship

Notes: Panel A shows the estimates of the same-integration group indicator on changes in political opinion gap as in Table 5's Panel A, based on subsamples of pairs with different initial pre-Sciences Po opinion gaps. Panel B shows the corresponding estimates of friendship, instrumented by the same-integration group indicator, on changes in political opinion gap as in Table 9, based on subsamples of pairs with different initial pre-Sciences Po opinion gaps. All bands show 95% confidence intervals based on dyadic clustered standard errors, which allow for error correlations between dyads sharing a common integration group.

before Sciences Po (opinion gap of at least 2).

Columns 5 to 8 show a similar pattern when we focus on only associations related to politics. Most notably, the same-integration group effect in column 5 (friendship effect in column 6) is 7.2 (4.9) times larger among student pairs with initially similar political views than those with dissimilar views. Column 5's effect (column 6's effect) is 2.4 (1.9) times larger than its corresponding effect in the full sample shown in Table 10's column 3 (column 4). Overall, the evidence thus suggests that most of the effects on common association participation found in Table 10 come from pairs of students with initially similar political views.

Interactions and common interest in politics. The effects of same-integration group exposure and friendship on common political association membership may work through (i) a general political interest channel, in that friends affect each other's general interest in politics, or through (ii) a friend-specific channel, such that friends' influences on each other's particular political views, or their choices to interact more in political associations. To test for the first channel, we estimate friends' participation in different political associations, as two friends who reinforce each other's general interest in politics may choose to join some political associations, but not necessarily the same. We examine this possibility in Appendix Table A21, and find no significant evidence of any

Table 12: Friendship, Integration Group, and Association Activities

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Both Are Members of the Same Association in							
	Any Area				Politics			
Sample	Initial Opinion Gap < 2		Initial Opinion Gap ≥ 2		Initial Opinion Gap < 2		Initial Opinion Gap ≥ 2	
Specification:	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Same Integration Group	0.0517**		0.0221		0.0556***		0.0077	
	(0.0223)		(0.0187)		(0.0194)		(0.0133)	
Friendship		0.2338**		0.1469		0.2516**		0.0513
		(0.1027)		(0.1147)		(0.1029)		(0.0872)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weak IV Test Statistic		35.044		27.850		35.044		27.850
Observations	9,393	9,393	14,043	14,043	9,393	9,393	14,043	14,043
Number of IGs	52	52	52	52	52	52	52	52
R-Squared	0.0071	0.0179	0.0069	0.0145	0.0060	0.0186	0.0031	0.0069
Mean (Dep. Var.)	0.110	0.110	0.0882	0.0882	0.0327	0.0327	0.0163	0.0163
Std. Dev. (Dep. Var.)	0.313	0.313	0.284	0.284	0.178	0.178	0.127	0.127

Notes: This table shows dyadic specifications relating the indicator of being members of the same association with the same-integration group indicator (columns 1, 3, 5, 7) and friendship (columns 2, 4, 6, 8), where friendship is instrumented by the same-integration group indicator. Columns 5 to 8 focus on associations related to politics and policies. Columns 1, 2, 5, 6 use the subsample of pairs that started out similarly before Sciences Po (with pre-Sciences Po opinion gaps of 0 or 1). Columns 3, 4, 7, 8 use the subsample of pairs that started out differently before Sciences Po (with pre-Sciences Po opinion gaps of 2 or more). Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, including association categorization, and the set of controls.

effect of being in the same integration group or being friends on a pairs' joining different associations related to politics. Those results thus highlight that the discovered effect is targeted towards friends' interactions and their specific political views, and does not work simply through heightened general interest in politics.

6 Concluding remarks

This paper empirically examines how incoming Sciences Po students' exposure to one another in integration groups and their newly formed friendships shape their political views over six months. We find that both shared exposure and friendship reduce the likelihood of political divergence and significantly narrow opinion gaps—by 8% for students in the same integration group and by 50% for friends, relative to the average opinion gap. Additionally, treated pairs are more likely to join the same political association.

The empirical results support what we term the "homophily-enforced" mechanism. When students with similar initial political views become friends, e.g., following chance encounters in an integration group, they continue to engage in related discussions, as reflected in their shared participation in political associations. Those sustained interactions, reinforced by homophily, drive the strong influence of friendship among politically like-minded students. As a result, such pairs are strongly discouraged from diverging and tend to have closer opinion gaps than between non-friend pairs. In contrast, friendships between students with initially dissimilar political opinions do not

follow this pattern, as they may bond over other shared affinities. Their friendship is hence largely ineffective in shaping their political views.

This mechanism aligns with Golub and Jackson's (2012) analysis on homophily and the speed of belief convergence, while introducing an endogenous selection of interaction dimensions based on homophilous preferences. However, our empirical findings reveal a nonlinear diffusion of beliefs, particularly evident in the asymmetry between convergence and divergence—an aspect not explored in Golub and Jackson (2012).

References

Allcott, Hunt and Matthew Gentzkow, "Social Media and Fake News in the 2016 Election," *The Journal of Economic Perspectives*, 2017, *31* (2), 211–235. 2

_ , Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, "The Welfare Effects of Social Media," *American Economic Review*, 2020, *110* (3), 629–676. 2, 5

_ , Matthew Gentzkow, and Lena Song, "Digital Addiction," *American Economic Review*, 2022, *112* (2), 2424–2463. 2, 5

Allport, Gordon Willard, *The Nature of Prejudice*, Cambridge, MA: Addison-Wesley, 1954. 5

Altonji, Joseph, Todd Elder, and Christopher Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184. 25, A12

Angrist, Joshua D. and Kevin Lang, "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *The American Economic Review*, 2004, *94* (5), 1613–1634. 4

Battaglini, Marco and Eleonora Patacchini, "Influencing Connected Legislators," *Journal of Political Economy*, 2018, *126* (6), 2277–2322. 4

Biais, Bruno and Martin Weber, "Hindsight Bias, Risk Perception, and Investment Performance," *Management Science*, 2009, *55* (6), 1018–1029. A6

Billings, Stephen B., Eric Chyn, and Kareem Haggag, "The Long-Run Effects of School Racial Diversity on Political Identity," *American Economic Review: Insights*, 2021, *3* (3), 267–284. 5

Boisjoly, Johanne, Greg J. Duncan, Michael Kremer, Dan M. Levy, and Jacque Eccles, "Empathy or Antipathy? The Impact of Diversity," *American Economic Review*, 2006, *96* (5), 1890–1905. 5

Boxell, Levi, Matthew Gentzkow, and Jesse Shapiro, "Greater Internet Use is Not Associated

with Faster Growth in Political Polarization among US Demographic Groups," *Proceedings of the National Academy of Sciences*, 2018, *115* (3), 10612–7. 2

Bradburn, Norman M., Lance J. Rips, and Steven K. Shevell, "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys," *Science*, 1987, *236* (4798), 157–161. 10, A5

Cai, Jing and Adam Szeidl, "Interfirm Relationships and Business Performance," *The Quarterly Journal of Economics*, 2018, *133* (3), 1229–1282. 5

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou, "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 2009, *76* (4), 1239–1267. 26

Camerer, Colin, George Loewenstein, and Martin Weber, "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 1989, *97* (5), 1232–1254. A6

Cameron, Colin and Douglas Miller, "Robust Inference for Dyadic Data," *Unpublished, University of California-Davis*, 2014. 11

Campbell, Arthur, C. Matthew Leister, Philip Ushchev, and Yves Zenou, "Network Structure and Polarization," Unpublished manuscript, Monash University, 2025. 33

Canen, Nathan and Francesco Trebbi, "Endogenous Network Formation in Congress," NBER Working Paper 22756, 2016. 4

_ , **Matt O. Jackson, and Francesco Trebbi**, "Social Interactions and Legislative Activity," *Journal of the European Economic Association*, 2023, *21* (3), 1072–1118. 4

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang, "Curriculum and Ideology," *Journal of Political Economy*, 2017, *125* (2), 338–392. 21

Carlsson, Magnus, Gordon B. Dahl, and Dan-Olof Rooth, "Backlash in Attitudes After the Election of Extreme Political Parties," *Journal of Public Economics*, 2021, *204*, 104533. 4

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*, 2013, *81* (3), 855–882. 5

Chandrasekhar, Arun and Randall Lewis, "Econometrics of sampled networks," *Unpublished, MIT*, 2011. 21

Chirac, Jacques, *My Life in Politics*, St. Martin's Press, 2012. 7

Cohen, Lauren and Christopher J. Malloy, "Friends in High Places," *American Economic Journal: Economic Policy*, 2014, *6* (3), 63–91. 4

Corno, Lucia, Eliana La Ferrara, and Justine Burns, "Interaction, Stereotypes and Performance: Evidence from South Africa," *American Economic Review*, 2022, *112* (12), 3848–3875. 5

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin, "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, 2009, *77* (4), 1003–1045. 2

Darmofal, David, Charles J. Finocchiaro, and Indridi H. Indridason, "Roll-call voting under random seating assignment," *Political Science Research and Methods*, 2025, *13* (1), 76–95. 4

DeGroot, Morris H., "Reaching a Consensus," *Journal of the American Statistical Association*, 1974, *69* (345), 118–121. 33

DellaLena, Sebastiano, "The Spread of Misinformation in Networks with Individual and Social Learning," *European Economic Review*, 2024, *168*, 104804.

_ **, Luca Paolo Merlino, and Yves Zenou**, "Affective Polarization, Media Outlets, and Opinion Dynamics," CEPR Discussion Paper 18508, 2023. 2, 5

DellaVigna, Stefano and Ethan Kaplan, "The Fox News effect: Media Bias and Voting," *The Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234. 5

_ **and Matthew Gentzkow**, "Persuasion: Empirical Evidence," *Annual Review of Economics*, 2010, *2* (1), 643–669. 4, 21, 32

Fafchamps, Marcel and Flore Gubert, "Risk Sharing and Network Formation," *American Economic Review*, 2007, *97* (2), 75–79. 11

Fischhoff, Baruch and Ruth Beyth, "I Knew it Would Happen ? Remembered Probabilities of Once-Future Things," *Organizational Behavior and Human Performance*, 1975, *13*, 1–16. 10, A6

Foucault, Martial and Anne Muxel, *Une jeunesse engagée. Enquête sur les étudiants de Sciences Po, 2002-2022*, Paris: Presses de Sciences Po, 2022. 2, 6

_ **and Brice Teinturier**, *Baromètre politique des Français*, ENEF IPSOS, 2022. 6

Fowler, James H., "Connecting the Congress: A Study of Cosponsorship Networks," *Political Analysis*, 2006, *14* (4), 456–487. 4

Gabel, Matthew and Kenneth Scheve, "Estimating the Effect of Elite Communications on Public Opinion Using Instrumental Variables," *American Journal of Political Science*, 2007, *51* (4), 1013–1028. 4

Ge, Qi and Stephen Wu, "How do you Say your Name? Difficult-to-Pronounce Names and Labor Market Outcomes," *American Economic Journal: Economic Policy*, 2024, *16* (4), 254–279. 13

Gentzkow, Matthew, "Television and Voter Turnout," *The Quarterly Journal of Economics*, 2006, *121* (3), 931–972. 5

_ , **Jesse Shapiro, and Michael Sinkinson**, "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, December 2011, *101* (7), 2980–3018. 5

Gerber, Alan, Dean Karlan, and Daniel Bergan, "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions," *American Economic Journal: Applied Economics*, April 2009, *1* (2), 35–52. 5

Goeree, Jacob, Margaret McConnell, Tiffany Mitchell, Tracey Tromp, and Leeat Yariv, "The 1/d Law of Giving," *American Economic Journal: Microeconomics*, 2010, *2* (1), 183–203. 21

Golub, Benjamin and Matthew O. Jackson, "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *The Quarterly Journal of Economics*, 2012, *127* (3), 1287–1338. 2, 33, 40

Guess, Andrew, Brendan Nyhan, Benjamin Lyons, and Jason Reifler, *Avoiding the Echo Chamber about Echo Chambers*, Knight Foundation, 2018. 2

Harmon, Nikolaj, Raymond Fisman, and Emir Kamenica, "Peer Effects in Legislative Voting," *American Economic Journal: Applied Economics*, 2019, *11* (4), 156–180. 3, 4, 11, 12, 13, 21, 32, 35, 37

Healy, Andrew and Neil Malhotra, "Childhood Socialization and Political Attitudes: Evidence from a Natural Experiment," *The Journal of Politics*, 2013, *75* (4), 1023–1037. 5

Kaplan, Ethan, Jörg L. Spenkuch, and Cody Tuttle, "School Desegregation and Political Preferences: Long-run Evidence from Kentucky," Unpublished manuscript, University of Maryland, 2022. 5

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi, "How Do Voters Respond to Information? Evidence from a Randomized Campaign," *American Economic Review*, 2015, *105* (1), 322–53. 5

Kennedy, Peter E., "Randomization Tests in Econometrics," *Journal of Business & Economic Statistics*, 1995, *13* (1), 85–94. 10

Lazarsfeld, Paul and Robert Merton, "Friendship as a Social Process: A Substantive and Methodological Analysis," in Morroe Berger, Theodore Abel, and Charles H. Page, eds., *Freedom and Control in Modern Society*, New York: Van Nostrand, 1954, pp. 18–66. 2, 23

_ , **Hazel Gaudet, and Bernard Berelson**, *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign.*, New York, NY: Columbia University Press, 1944. 2, 5

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355. 13

Leider, Stephen, Markus Möbius, Tanya Rosenblat, and Quoc-Anh Do, "Directed Altruism and Enforced Reciprocity in Social Networks," *The Quarterly Journal of Economics*, 2009, *124* (4), 1815–1851. 3, 21, 22, 26

_ , _ , _ , and _ , "What do we Expect from our Friends?," *Journal of the European Economic Association*, 2010, *8* (1), 120–138. 21, 22

Levy, Ro'ee, "Social Media, News Consumption, and Polarization: Evidence from a Field Experiment," *American Economic Review*, 2021, *111* (3), 831–870. 2, 5

List, John, Fatemeh Momeni, Michael Vlassopoulos, and Yves Zenou, "Neighborhood Spillover Effects of Early Childhood Interventions," CEPR Discussion Paper No. 18134, 2023. 5

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 2021, *111* (6), 1807–1844. 5

_ and Donghee Jo, "Legislature Integration and Bipartisanship: A Natural Experiment in Iceland," *Journal of Politics,* forthcoming, 2025. 4

McPherson, Miller, Lynn Smith-Lovin, and James M Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 2001, *27* (1), 415–444. 2, 23

Mendelberg, Tali, Katherine T. McCabe, and Adam Thal, "College Socialization and the Economic Views of Affluent Americans," *American Journal of Political Science*, 2017, *61* (3), 606–623. 5

Merlino, Luca P., Max F. Steinhardt, and Liam Wren-Lewis, "More than Just Friends? School Peers and Adult Interracial Relationships," *Journal of Labor Economics*, 2019, *37*, 663–713. 5

Möbius, Markus and Tanya Rosenblat, "Social Learning in Economics," *Annual Review of Economics*, 2014, *6* (1), 827–847. 32

Molavi, Pooya, Alireza Tahbaz-Salehi, and Ali Jadbabaie, "A Theory of Non-Bayesian Social Learning," *Econometrica*, 2018, *86* (2), 445–490. 33

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204. 25, A12

Paluck, Elizabeth Levy, Seth A Green, and Donald P Green, "The Contact Hypothesis Re-Evaluated," *Behavioural Public Policy*, 2019, *3* (2), 129–158. 5

Pariser, Eli, *The Filter Bubble: What the Internet is Hiding from You*, Penguin UK, 2011. 2

Pettigrew, Thomas F and Linda R Tropp, "A Meta-Analytic Test of Intergroup Contact Theory," *Journal of Personality and Social Psychology*, 2006, *90* (5), 751. 5

Polipciuc, Maria, Frank Corvers, and Raymond Montizaan, "Peers' Race in Adolescence and Voting Behavior," *Economics of Education Review*, 2023, *97*, 102486. 5

Prior, Markus, "You've Either Got It or You Don't? The Stability of Political Interest over the Life Cycle," *The Journal of Politics*, 2010, *72* (3), 747–766. 5

_ , *Hooked: How Politics Captures People's Interest*, Cambridge, UK: Cambridge: Cambridge University Press, 2018. 5

Rao, Gautam, "Familiarity Does Not Breed Contempt: Generosity, Discrimination and Diversity in Delhi Schools," *American Economic Review*, 2019, *109* (3), 774–809. 5

Rogowski, Jon C. and Betsy Sinclair, "Estimating the Causal Effects of Social Interaction with Endogenous Networks," *Political Analysis*, 2012, *20* (3), 316–328. 4

Sacerdote, Bruce, "Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?," *Annual Review of Economics*, 2014, *6* (1), 253–272. 5

Saia, Alessandro, "Random Interactions in the Chamber: Legislators' Behavior and Political Distance," *Journal of Public Economics*, 2018, *164*, 225–240. 4

Siddique, Abu, Michael Vlassopoulos, and Yves Zenou, "Leveraging Edutainment and Social Networks to Foster Interethnic Harmony," Unpublished manuscript, Monash University, 2024. 5

Sunstein, Cass, *Republic.com 2.0*, Princeton, NJ: Princeton University Press, 2009. 2

_ , *# Republic: Divided Democracy in the Age of Social Media*, Princeton, NJ: Princeton University Press, 2018. 2

Wagenaar, Willem, "My Memory: A Study of Autobiographical Memory over Six Years," *Cognitive Psychology*, 1986, *18* (2), 225–252. A5

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press, 2010. A8

(Not-for-publication) ONLINE APPENDIX

A Appendix: Description of data

Sample construction: The sample excludes observations (pairs of students) in which any of the above-mentioned variables is missing, when at least one of the two individuals in the couple did not answer to the related question in the survey. We also drop pairs that contain at least one individual in the top 5 percent of the distribution of time taken to name each friend (about 82 seconds per friend or 13.5 minutes for individuals with 10 friends).

Controls: The standard set of controls in dyadic specifications throughout the paper include the following variables: Same Gender, Both Female, Same Second Nationality, Same Admission Type, Both Affirmative Action, Same Département of High School, Same Region of High School, Same High School Major, Difference in Tuition Fees, Both Free Tuition, Same Parents Profession, Same ZIP Code, Both from Paris, Both from Ile de France (Greater Paris) Region of High School, Same Special Program.

Table A1: Description of Variables in Dyadic Data

Variable	Description
Friendship	1 if at least one of the two individual has named the other as one of her friends (the 'OR' network of undirected friendship), zero otherwise.
Same Integration Group (IG)	1 if the two individuals have attended the same integration group before starting the first school year at Sciences Po, 0 otherwise.
Difference in political opinion (March 2014)	Absolute difference in political opinions of the two individuals, as declared on a 1-10 scale in the main survey (March. 2014).
Difference in initial (pre-Sciences Po) political opinion (August 2013)	Absolute difference in political opinions of the two individuals from before entering Sciences Po (August 2013), as declared on a 1-10 scale in the main survey (March. 2014).
Difference in political opinion in 2015	Absolute difference in political opinions of the two individuals, as declared on a 1-10 scale in the 2015 survey.
Difference in political opinion in 2014 (Recalled)	Absolute difference in political opinions of the two individuals in 2014, as declared on a 1-10 scale in the 2015 survey.
Both members of some association	1 if the two individuals are members of some student association. Missing if at least one of them did not answer this question. 0 otherwise.
Both members of some association of type T	1 if the two individuals are members of some student association of type T (see classification of association types below). Missing if at least one of them did not answer this question. 0 otherwise.
Both members of different association of type T	1 if each of the two individuals is member of some student association of type T, without both being members of the same association of type T (see classification of association types below). Missing if at least one of them did not answer this question. 0 otherwise.

Both members of the same association	1 if the two individuals are members of the same student association. Missing if at least one of them did not answer this question. 0 otherwise.
Both members of the same association of type T	1 if the two individuals are members of the same student association of type T (see classification of association types below). Missing if at least one of them did not answer this question. 0 otherwise.
Association Types	Survey participants are members of 107 student associations at Sciences Po. We classify them into four types: • Political associations, including those directly affiliated to political parties, those that focus their actions and debates on political issues, and student unions (usually committed to political struggles), • Sports associations, • Humanitarian associations with a clear humanitarian/charity agenda that is not politically controversial, such as human right issues or environmental protection, • "Identity" associations that gather individuals based on common personal characteristics, such as the LGBTQ+ group, religious groups, and associations based on geographical origins (such as province or country of origin).
Movement in Same Direction	1 if both individuals have changed their political opinion between August 2013 and March 2014 and their new political opinion have moved in the same direction relative to their initial one ($\Delta Y_i \Delta Y_j \geq 0$), 0 otherwise.
Strong Convergence	1 if the two individuals have different initial political positions, none of them have moved away from and at least one of them has moved towards the other initial political opinion relative to her own initial position ($\Delta Y_i(Y_{j0} - Y_{i0}) > 0$ & $\Delta Y_j(Y_{i0} - Y_{j0}) > 0$). Missing if the two individuals have the same initial political opinion. 0 otherwise.
Weak Convergence	1 if the two individuals have different initial political positions, none of them have moved away from the other initial political opinion relative to her own initial position. Missing if the two individuals have the same initial political opinion ($\Delta Y_i(Y_{j0} - Y_{i0}) \geq 0$ & $\Delta Y_j(Y_{i0} - Y_{j0}) \geq 0$). 0 otherwise.
Strong Divergence	1 if the two individuals have both moved away from each others initial political position relative to their own initial position ($\Delta Y_i(Y_{j0} - Y_{i0}) < 0$ & $\Delta Y_j(Y_{i0} - Y_{j0}) < 0$), 0 otherwise.
Weak Divergence	1 if the two individuals have not moved towards each others political position relative to their own initial position ($\Delta Y_i(Y_{j0} - Y_{i0}) \leq 0$ & $\Delta Y_j(Y_{i0} - Y_{j0}) \leq 0$), 0 otherwise.
Friendship Strength 1	1 if at least one of the two individual has named the other as one of her friends and has stated that their friendship is at least as intense as a "mere relationship", 0 otherwise.
Friendship Strength 2	1 if at least one of the two individual has named the other as one of her friends and has stated that their friendship is at least as intense as a "friendship link", 0 otherwise.
Friendship Strength 3	1 if at least one of the two individual has named the other as one of her friends and has stated that their friendship is at least as intense as a "close friendship", 0 otherwise.
Friendship Strength 4	1 if at least one of the two individual has named the other as one of her friends and has stated that their friendship is at least as intense as a "very close friendship", 0 otherwise.

Alphabetical distance between last names	The entire cohort's last names are ordered alphabetically, and placed on a circle (so that after last names starting with 'Z' we return to last names starting with 'A'). Any pair of last names on this circle are connected through two different arcs. Their alphabetical distance refers to the number of last names between them in the shorter arc, plus one. Put differently, denoting their ranks on the alphabetically ordered list of the cohort's last names as $r_1, r_2 \in [1, N]$, $r_1 < r_2$, the alphabetical distance is $\min(r_2 - r_1, N + r_1 - r_2)$, N being the total number of last names (exactly 800 for the cohort in consideration).
Difference in Differences in Political Opinion	Difference in Political Opinion in March 2014 minus Difference in Political Opinion from before entering Sciences Po.
Same Gender	1 if the two individuals are of the same gender, 0 otherwise.
Both Female	1 if the two individuals are both female, 0 otherwise.
Same Second Nationality	1 if the two individuals share a common second (i.e., non-French) nationality, 0 otherwise.
Same Admission Type	1 if the two individuals have been admitted through the same admission procedure, 0 otherwise. The three main procedures include the standard admission procedure (consideration of dossier, written tests, and oral tests), the international procedure (consideration of dossier and oral tests), and the priority admission (consideration of dossier and oral interview among students from schools in disadvantaged areas).
Both Affirmative Action	1 if the two individuals have both been admitted through the priority admission procedure, 0 otherwise. This is Sciences Po's affirmative action channel that targets high schools in disadvantaged areas of France (the ZEP, prioritized educational zones) under its Prioritized Education Convention (CEP). This admission procedure includes examination of dossier and of an oral interview, but not the standard written test.
Same Département of High School	1 if the two individuals have completed their high school diploma in the same French département, 0 otherwise. Metropolitan France is composed of 96 départments.
Same Region of High School	1 if the two individuals have completed their high school diploma in the same French region, 0 otherwise. Metropolitan France is composed of 22 regions.
Both from High Schools in Ile de France	1 if the two individuals have completed their high school diploma in the same Greater Paris region of 'Ile de France', 0 otherwise.
Same High School Major	1 if the two individuals have a high school diploma with the same major classification, 0 otherwise. The categories include ES (Economic and Social), L (Literary/Language-Mathematics), S (Sciences), and Foreign Diplomas (grouped into one category).
Difference in Tuition Fees	Absolute difference in tuition fees among the couple (proxy for family income). At Sciences Po, the amount of tuition is a function of the parents' official income tax quotient, which is calculated based on total household income and household size.
Both Free Tuition	1 if both individuals do not pay tuition fees, 0 otherwise. Students pay no tuition when their parents' income tax quotient is below a threshold.
Same Parents' Profession	1 if at least one of an individual's parents has a common profession with at least one of the other individual's parents, 0 otherwise. The information on parents' profession is based on the French government's official socio-professional categories.
Same ZIP code	1 if the two individuals live in the same ZIP code area, 0 otherwise. The Greater Paris region of 'Ile de France' contains more than 528 areas with separate ZIP codes, mostly corresponding to arrondissements (districts) inside Paris and cantons outside Paris.

Both from Paris	1 if the two individuals' ZIP codes are both inside Paris, 0 otherwise.
Same Program	1 if the two individuals are enrolled in the same study program, 0 otherwise. In our sample, apart from the common undergraduate program that all students undertake, some students are enrolled in double-degree programs joint between Sciences Po and other, French or non-French educational institutions. In some cases they are subject to additional constraints in terms of course timing.

Table A2: Description of Data Cleaning Steps

Number of observations	Data cleaning stage
Individual dataset	
800	Full population of students
543	After dropping individuals that did not respond to any question on friendship
Dyadic dataset	
$294,849 = 543^2$	After creating the dyadic dataset by crossing the individual dataset for individual 1 and individual 2
$294,306 = 543^2 - 543$	After dropping dyads with same individual (diagonal of the adjacency matrix)
$121,452 = 349^2 - 349$	After keeping only observations with non-missing information on political opinions (contemporaneous and pre-Sciences Po) and controls
$109,230 = 331^2 - 331$	After dropping observations at least one individual was in the top 5 percent of the distribution of time taken to respond to the friendship questions
$104,652 = 324^2 - 324$	After dropping observations where the integration group is missing
$52,326 = 104,652 \div 2$	After keeping only the upper triangular adjacency matrix (i.e., keeping only one of the two pairs (i, j) and (j, i))

A.1 Appendix: Robustness check: Potential violation of exclusion restriction

As discussed in subsection 4.2, the major concern about the IV strategy's exclusion restriction is the existence of individuals in the same IG whose relationships are not declared as friendship, but who may still have influenced each other's political opinions. To deal with this concern, we apply the approach sketched in subsection 4.2.3 to evaluate how this concern may affect the main estimates. Recall that we consider two observable levels of intensity of relationships, denoted as L^1 (acquaintances) and L^2 (friends, close friends, and very close friends). Two parameters are sufficient in determining the biases, hence the true effect β_2 (β_1) of L^2 (L^1) on outcome.

The first parameter δ measures the relative importance of the two channels of L^0 and L^1 in terms of the influence of the IV IG_{ij} on the outcome DY_{ij}. In our context, it is most likely that declared acquaintances (L^1) are at least as important as omitted acquaintances (L^0), so it is likely that $\delta < 1$. We will explore a broad range of δ from as small as 0.1 to 2.

The second parameter, γ, measures the relative endogeneity biases of L^1 versus L^2. Intuitively, homophily, the main force behind those biases, is likely stronger for higher friendship intensity, so

Table A3: Additional Descriptive Statistics of Covariates

Panel A: Monadic Independent Variables

Variable	(1) Full Sample		(2) Benchmark Sample			
	Mean	Standard deviation	Mean	Obs.	Standard deviation	Obs.
Gender (1= Female)	0.592	(0.492)	796	0.583	(0.494)	331
Honors Graduation	0.754	(0.431)	796	0.831	(0.375)	331
Tuition Fees	3602	(3495)	713	3826	(3328)	331

Panel B: Dyadic Independent Variables

Variable	(1) Full Sample			(2) Benchmark Sample		
	Mean	Standard deviation	Observations	Mean	Standard deviation	Observations
Same Gender	0.522	(0.500)	147,153	0.511	(0.500)	52,326
Both Female	0.369	(0.483)	147,153	0.336	(0.472)	52,326
Same Second Nationality	0.036	(0.187)	147,153	0.009	(0.094)	52,326
Same Admission Type	0.565	(0.496)	147,153	0.697	(0.459)	52,326
Both Affirmative Action	0.0291	(0.168)	147,153	0.0127	(0.112)	52,326
Same Département of High School	0.0517	(0.221)	132,870	0.0613	(0.240)	52,326
Same Region of High School	0.253	(0.435)	132,355	0.250	(0.433)	52,326
Both from Ile de France High Schools	0.212	(0.408)	147,153	0.231	(0.422)	52,326
Same High School Major	0.363	(0.481)	147,153	0.382	(0.486)	52,326
Difference in Tuition Fees ('000)	3.879	(3.005)	122,760	3.744	(2.811)	52,326
Both Free Tuition	0.476	(0.499)	147,153	0.624	(0.484)	52,326
Same Parents' Profession	0.422	(0.494)	119,316	0.445	(0.497)	52,326
Same ZIP code	0.0264	(0.160)	146,611	0.0252	(0.157)	52,326
Both from Paris	0.506	(0.500)	147,153	0.505	(0.500)	52,326
Same Program	0.028	(0.165)	147,153	0.031	(0.174)	52,326

Notes: Statistics in (1) are computed on the full sample of data available for each variable, while statistics in (2) are computed on the benchmark sample, which is detailed in Table A1

we expect $\gamma < 1$. We will consider a broad range of γ from 0.25 (L_2's endogeneity bias is four times that of L_1) to 2 (L_2's endogeneity bias is half of that of L_1).[1]

The two plots in Figure A2 show that for very broad ranges of δ and γ, both coefficients β_1 and β_2 are clearly negative. If we are mostly concerned with the effect of friendship beyond simple acquaintance, namely β_2, we can see that its magnitude is very strong and barely goes below 0.6 for the range of δ and γ below 1. So we can safely claim that our result is very much robust to the concern of possible violation of the exclusion restriction.

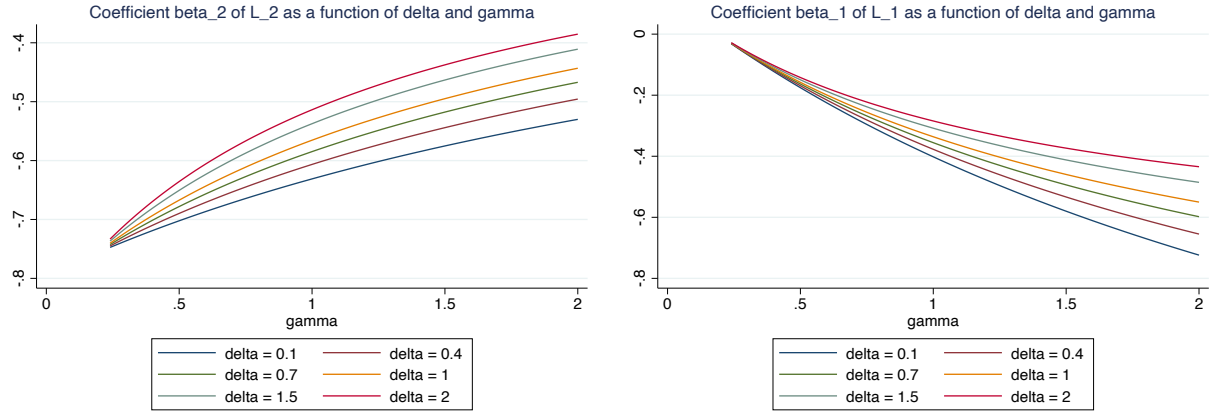
A.2 Robustness check: Precision of retrospective question on opinion

Second, we use a retrospective question in the survey in March 2014 on students' political opinions just before they join Sciences Po (see description in subsection 4.1), which raises a potential concern that retrospective answers may incorporate a bias in the direction of the respondent's opinion today. While such a measurement error regarding retrospective survey questions on events and answers may be rather small after only 6 months,[2] the bias on opinions may also relate to the rationalization

[1] On γ, see precise definitions and more thorough discussion in Appendix B.

[2] Wagenaar (1986) finds that 20% of subjects forget key personal events after one year. See review by Bradburn et al. (1987).

Figure A1: True effects of L^2 and L^1 on political opinions



Notes: The subgraphs show respectively the effect β_2 of L^2 and β_1 of L^1 on political opinions as functions of the values of δ and γ. $\delta = \frac{\beta_0 \pi_0}{\beta_1 \pi_1}$ measures the relative importance of the two channels of L_0 and L_1 in terms of the influence of the IV IG on outcome Y. $\gamma = \frac{\beta_0 \rho_{01} + \rho_{u1}}{\beta_0 \rho_{02} + \rho_{u2}}$ measures the relative endogeneity biases of L^1 versus L^2.

of new information that results in a hindsight bias, according to which individuals reconstruct their past opinion in light of their newly updated opinion (Fischhoff and Beyth, 1975). It is thus useful to investigate our method's robustness to this issue.

To evaluate the magnitude of the retrospective answer measurement error, we use the second survey in June 2015 to compare the answers to its retrospective question on recalled opinion back in March 2014 with the actual answers in 2014. First, Appendix Table A11 shows the joint distribution of both surveyed and recalled opinions for 2014. The mass is clearly concentrated on the diagonal, with 90% of the observations not differing more than 1 point between the two measures, implying a very strong correspondence between recalled and actual answers. This lends confidence to the accuracy of the recalled opinion expressed in March 2014 over the political opinion in August 2013.[3]

Appendix Table A12 presents further results on students' recall error, measured as recalled opinion for 2014 minus actual opinion surveyed in 2014. The absolute magnitude of the recall error has practically zero partial correlations with past and present actual political opinions, as shown in column 1. However, in column 2 we do find evidence that the signed recall error is strongly correlated with the change in opinions from 2014 to 2015, signifying that recalled opinions are biased towards present opinions (as surveyed in 2015) by the same magnitude as estimated, e.g., by Fischhoff and Beyth (1975); Biais and Weber (2009); Camerer et al. (1989).

Can the recall error strongly affect our results? First, as less than 10% of answers of the recalled

[3]Unfortunately, due to reduced budget in 2015, the participation rate in 2015 is much lower than in 2014, resulting in a small sample that overlaps between the two waves that we cannot use as a panel to study friendship effect.

opinion suffer a serious recall error, the resulting bias on our results would probably be small. Second, when we control for the recalled opinion gap in the OLS specification, if this variable is biased towards actual opinion gap of March 2014, it would create an attenuation bias of our coefficient of interest towards zero. The effect of friendship is attenuated because the biased variable tends to absorb more variation in the outcome variable than does the latent true opinion. Third, the control variables are not needed for the IV strategy's validity, and are only included to improve estimates' precision. Indeed, the results remain very similar, albeit less precise, if we do not control for pre-Sciences Po political opinions.

B Appendix: Control function approach and the friendship effect

In this section, we propose a framework to evaluate the potential bias on the estimator of the effect of friendship on political opinion gap as implemented with the Instrumental Variable strategy in section 4. To do so, we first rely on the equivalence between the IV 2SLS estimator and a control function estimator. We then use the latter specification to evaluate the potential biases due to the exclusion of different measures of friendship at different levels of intensity. Thanks to the surveyed intensity of friendship, we can then assess those biases and provide bounds for the main coefficient of interest. The calculation implies a rather tight interval of the potentially biased estimate, thus high robustness of the results reported in section 4.

For simplicity, we will use the notation $\mathbb{E}^*[U|X_1, \ldots, X_n]$ for the linear projection of a random variable U on a constant and $X_1, \ldots, X_n$. With a slight abuse of notations, we will reuse the letter α to indicate a constant term.

B.1 IV and control function estimators

Let us first consider the following simple linear model in which the outcome of pairwise opinion gap Y_{ij} is influenced by friendship L_{ij}.

$$Y_{ij} = \alpha + \beta L_{ij} + U_{ij} + \varepsilon_{ij}. \tag{5}$$

The unobserved errors include a centered idiosyncratic error ε uncorrelated with L (so $\mathbb{E}[\varepsilon L] = 0$), and an unobserved centered term U that captures the issue of the endogeneity of friendship, in that it may correlate with L_{ij}. We expect a negative coefficient β, i.e., friendship causes a reduction in pairwise opinion gap. The OLS estimate of β contains a homophily bias when people of similar political views are more likely to become friends. To understand this bias, we can write the linear

projection of U on L as $\mathbb{E}^*[U|L] = \alpha + \kappa L$, so $U = \alpha + \kappa L + \eta$ for an error term η uncorrelated with L. Replacing the last expression in equation (5), we deduce that, when U is unobserved, the OLS estimate's bias is κ.

We first recall the control function (CF) approach that can solve this endogeneity problem with the use of an exogenous instrumental variable IG (see, e.g., Wooldridge, 2010, c.6). We make the following standard IV assumptions, based on the exogeneity and importance of IG membership as discussed in section 4.2.

Assumption A1 (Relevance) *IG helps linearly predicts L, so that* $\mathbb{E}^*[L|IG] = \alpha + \pi IG, \quad \pi \neq 0$.

Assumption A2 (Excludability) *The unobserved terms U and ε are mean independent of IG, i.e.,* $\mathbb{E}[U|IG] = \mathbb{E}[\varepsilon|IG] = 0$.

Given those assumptions, the CF approach first estimates the residual $\hat{\nu}$ from the linear regression

$$L_{ij} = \alpha + \pi IG_{ij} + \nu_{ij}, \tag{6}$$

then use it as an additional control in equation (5) to obtain a consistent estimator of β.

To see how the CF approach works, we can write $\mathbb{E}^*[U|IG, \nu] = \rho\nu$ and define a new idiosyncratic noise $\xi = U - \mathbb{E}^*[U|IG, \nu]$. IG does not appear in $\mathbb{E}^*[U|IG, \nu]$ since both U and ν are mean independent of IG. So ξ is uncorrelated with both ν and IG, hence it is also uncorrelated with $L = \alpha + \pi IG + \nu$. We thus obtain the following expression of Y based on L and ν, which produces an OLS regression that yields a consistent estimator of β:

$$Y_{ij} = \alpha + \beta L_{ij} + \rho\nu_{ij} + \xi_{ij} + \varepsilon_{ij}. \tag{7}$$

In practice, we do not observe ν, but the use of the estimated residuals $\hat{\nu}$ instead of ν still produces a consistent estimator of β. In fact, this procedure produces an estimator $\hat{\beta}_{CF}$ that is identical to the 2SLS estimator (Wooldridge, 2010, c.6).[4]

The identification of β relies on both assumptions A1 and A2. The IV's excludability guarantees that the new error term ξ is uncorrelated with the regressors L and $\hat{\nu}$, and the IV's relevance establishes that those two regressors are not perfectly collinear, so that β can be identified in the

[4]Because of the estimated nature of $\hat{\nu}$, the calculation of the standard error of $\hat{\beta}_{CF}$ involves more steps. As we are mostly concerned about the potential bias when the exclusion restriction is violated, and not statistical inferences, we bypass those steps.

regression equation (7). The strategy is no longer valid if the excludability is violated. In what follows, we will establish the magnitude of the bias in this case.

B.2 Invalidated exclusion restriction

The empirical strategy discussed in subsection 4.2 is subject to the concern that the IV is not excludable, in that it may affect outcome through a channel other than friendship L. We model this concern as a form of unobserved dyadic relationship L^0 that is not captured by L, that is made more likely between two individuals in the same IG, and that has a direct effect on opinion gap Y. We now attempt to assess the bias in the following modified regression due to the unobserved nature of L^0:

$$Y_{ij} = \alpha + \beta L_{ij} + \beta_0 L^0_{ij} + U_{ij} + \varepsilon_{ij}. \tag{8}$$

Building on the CF approach in equation (6), we further write the linear projection $\mathbb{E}^*[L^0|IG, \nu] = \alpha + \pi_0 IG + \rho_0 \nu$ and denote the residual $\nu^0 = L^0 - \mathbb{E}^*[L^0|IG, \nu]$, so that:

$$L^0_{ij} = \alpha + \pi_0 IG_{ij} + \rho_0 \nu_{ij} + \nu^0_{ij}. \tag{9}$$

Since $IG = \frac{1}{\pi}(L - \alpha - \nu)$, it follows that $L^0 = \alpha + \frac{\pi_0}{\pi}(L - \nu) + \nu^0$, and that ν^0 is also uncorrelated with L. We further write the linear projection $\mathbb{E}^*[U|IG, \nu, \nu_0] = \rho_u \nu + \rho_{u0} \nu^0$, and denote the error $\xi = U - \mathbb{E}^*[U|IG, \nu, \nu^0]$, which is also uncorrelated with L. Plugging the expansions of U and L^0 into (8), we obtain:

$$\begin{aligned} Y_{ij} &= \alpha + \beta L_{ij} + \beta_0 \left(\frac{\pi_0}{\pi}(L_{ij} - \nu_{ij}) + \nu^0_{ij} \right) + (\rho_u \nu_{ij} + \rho_{u0} \nu^0_{ij} + \xi_{ij}) + \varepsilon_{ij} \\ &= \alpha + \beta \left(1 + \frac{\beta_0 \pi_0}{\beta \pi} \right) L_{ij} + \left(\rho_u - \frac{\beta_0 \pi_0}{\pi} \right) \nu_{ij} + (\beta_0 + \rho_{u0}) \nu^0_{ij} + \xi_{ij} + \varepsilon_{ij} \end{aligned} \tag{10}$$

As the unobserved term $(\beta_0 + \rho_{u0})\nu^0_{ij} + \xi_{ij} + \varepsilon_{ij}$ is orthogonal to the regressors L_{ij} and ν (which can be approximated by $\hat{\nu}$), the OLS estimator of β produces $\beta \left(1 + \frac{\beta_0 \pi_0}{\beta \pi}\right)$, with a bias to estimate ratio of $\frac{\beta_0 \pi_0}{\beta \pi}$.

Interpretation of the bias. Considering the specification in (8), the IV IG matters to the outcome Y through two separate channels, either via the main measure of friendship L or the omitted relationship L^0. Its "reduced form" impact on Y through each channel is the product of the "first stage" coefficient π (π_0) of L (L^0) on IG and its main effect β (β_0). So the bias ratio $\frac{\beta_0 \pi_0}{\beta \pi}$ represents the relative importance of the omitted channel versus the main channel through which

the IV works.

If the surveyed relationship is rather exhaustive, and the omitted channel unimportant, then the bias will likely be small. However, it is difficult to assess the size of the bias, and impossible to rule it out completely. In the next subsection, we will rely on the detailed intensity of friendship to gauge more precisely the magnitude of the bias.

B.3 Model with two levels of friendship intensity

Our survey elicits the intensity of each declared relationship by values of 1 (acquaintance), 2 (friendship), 3 (close friendship), and 4 (very close friendship). In what follows, we will define two new variables from the data, L^1 as the indicator of level-1 relationships and L^2 as the indicator of relationships of level 2 or higher. By construction, $L^1 + L^2 = L$. The regression equation, including the unobserved relationship L^0, is now written:

$$Y_{ij} = \alpha + \beta_2 L^2_{ij} + \beta_1 L^1_{ij} + \beta_0 L^0_{ij} + U_{ij} + \varepsilon_{ij}. \tag{11}$$

We are now interested mostly in estimating β_2. We will examine the size of the bias when L^0 is omitted, and provide a useful benchmark to gauge the size of the bias and bound the true parameter.

Keeping both L^2 and L^1. Following the deduction in subsection B.2, we write the linear projections of L^k on IG:

$$\begin{aligned} L^k_{ij} &= \alpha + \pi_k IG_{ij} + \nu^k_{ij}, \quad \mathbb{E}^*[\nu^k | IG] = 0, \quad k \in \{1, 2\}, \\ L^0_{ij} &= \alpha + \pi_0 IG_{ij} + \rho_{01}\nu^1_{ij} + \rho_{02}\nu^2_{ij} + \nu^0_{ij}, \quad \mathbb{E}^*[\nu^0 | IG, \nu^1, \nu^2] = 0. \end{aligned} \tag{12}$$

Because we only have one IV for two endogenous regressors (L^2, L^1), the variables (L^2, L^1, ν^2, ν^1) are collinear. As we are interested in β_2, we use the control function approach with only ν^2. To do so, we further project U on the residuals as $\mathbb{E}^*[U|IG, \nu^2, \nu^1, \nu^0] = \rho_{u2}\nu^2 + \rho_{u1}\nu^1 + \rho_{u0}\nu^0$, and denote the error $\xi = U - \mathbb{E}^*[U|IG, \nu^2, \nu^1, \nu^0]$. We then express $IG = \frac{1}{\pi_2}(L^2 - \nu^2)$, and $\nu^1 = L^1 - \pi_1 IG =$

$L^1 - \frac{\pi_1}{\pi_2}(L^2 - \nu^2)$, in order to rewrite equation (11) in terms of a projection on (L^2, L^1, ν^2):

$$\begin{aligned} Y_{ij} &= \alpha + \beta_2 L^2_{ij} + \beta_1 L^1_{ij} + \beta_0 \left[\frac{\pi_0}{\pi_2}(L^2_{ij} - \nu^2_{ij}) + \rho_{01}\left(L^1_{ij} - \frac{\pi_1}{\pi_2}(L^2_{ij} - \nu^2_{ij})\right) + \rho_{02}\nu^2_{ij} + \nu^0_{ij}\right] + \\ &+ \left[\rho_{u2}\nu^2_{ij} + \rho_{u1}\left(L^1_{ij} - \frac{\pi_1}{\pi_2}(L^2_{ij} - \nu^2_{ij})\right) + \rho_{u0}\nu^0_{ij} + \xi_{ij}\right] + \varepsilon_{ij} \\ &= \alpha + \left[\beta_2 + \frac{\beta_0\pi_0}{\pi_2} - \frac{\beta_0\rho_{01}\pi_1}{\pi_2} - \frac{\rho_{u1}\pi_1}{\pi_2}\right] L^2_{ij} + \left(\beta_1 + \beta_0\rho_{01} + \rho_{u1}\right) L^1_{ij} + \\ &+ \left[-\frac{\beta_0\pi_0}{\pi_2} + \frac{\beta_0\rho_{01}\pi_1}{\pi_2} + \beta_0\rho_{02} + \rho_{u2} + \frac{\rho_{u1}\pi_1}{\pi_2}\right] \nu^2_{ij} + \left(\beta_0 + \rho_{u0}\right)\nu^0_{ij} + \xi_{ij} + \varepsilon_{ij}. \end{aligned} \tag{13}$$

Equation (13) shows the feasible regression of Y on (L^2, L^1, ν^2), in which the coefficient of L^2 estimates $\beta_2\left(1 + \frac{\beta_0\pi_0}{\beta_2\pi_2} - \frac{(\beta_0\rho_{01}+\rho_{u1})\pi_1}{\beta_2\pi_2}\right)$. In addition to the bias expressed in equation (10), the new term $-\frac{(\beta_0\rho_{01}+\rho_{u1})\pi_1}{\beta_2\pi_2}$ comes from the endogeneity of L^1 that is not addressed with an IV. This term is proportionate to the bias of the coefficient of L^1, denoted $B_1 = \beta_0\rho_{01} + \rho_{u1}$.

Intuitively, it represents the bias due to the unaddressed endogeneity of L^1 in equation (13). It works through two channels, including L^1's correlation with the omitted L^0 ($\beta_0\rho_{01}$) and the unobservable U (ρ_{u1}). It is equivalent to the coefficient when we regress the omitted part $\beta_0 L^0 + U$ on ν^1, the residual of L^1.

Similar to B_1, denote $B_2 = \beta_0\rho_{02} + \rho_{u2}$, namely the coefficient when we regress the omitted part $\beta_0 L^0 + U$ on ν^2, the residual of L^2. Analogously to the case of B_1, B_2 represents the degree of endogeneity of L_2 through its correlations with with the omitted L^0 ($\beta_0\rho_{02}$) and the unobservable U (ρ_{u2}). Hence we denote the ratio of those two measures of the endogeneity of L_2 and L_1 as $\gamma = \frac{B_1}{B_2}$.

To fully utilize the recovered coefficients of specification (13), we further introduce $\delta = \frac{\beta_0\pi_0}{\beta_1\pi_1}$. As $\beta_i\pi_i$ represents the effect of the IV IG on outcome Y through the channel of L_i, the quantity δ measures the relative importance of the two channels of L_0 and L_1.

Analysis of biases. We can now write all estimated coefficients in specification (13), corresponding to (L^2, L^1, ν^2), in terms of the estimands β_i, the two measures of endogeneity biases B_1 and B_2, the parameter of relative importance δ, and the ratio of the estimated first-stage coefficients $\pi^* = \frac{\pi_1}{\pi_2}$:

$$\begin{aligned} \tilde{\beta}_{L^2} &= \beta_2 + (\delta\beta_1 - B_1)\,\pi^* \\ \tilde{\beta}_{L^1} &= \beta_1 + B_1 \\ \tilde{\beta}_{\nu^2} &= -(\delta\beta_1 - B_1)\pi^* + B_2. \end{aligned} \tag{14}$$

In the data, the three estimates are respectively -0.7922, 0.1222, and 0.6852, and $\pi^* = 3.5031$. Replacing $B_2 = B_1/\gamma$ in (14), we can generically solve this linear system for (β_1, β_2, B_1) in terms of δ, γ, π^*:[5]

$$\begin{aligned} B_1 &= \frac{\delta\pi^*\tilde{\beta}_{L^1} + \tilde{\beta}_{\nu^2}}{(\delta+1)\pi^* + \frac{1}{\gamma}} \\ \beta_1 = \tilde{\beta}_{L^1} - B_1 &= \frac{(\pi^* + \frac{1}{\gamma})\tilde{\beta}_{L^1} - \tilde{\beta}_{\nu^2}}{(\delta+1)\pi^* + \frac{1}{\gamma}} \\ \beta_2 = \tilde{\beta}_{L^2} - (\delta\beta_1 - B_1)\,\pi^* &= \tilde{\beta}_{L^2} - \frac{\frac{\delta\pi^*}{\gamma}\tilde{\beta}_{L^1} - (\delta+1)\pi^*\tilde{\beta}_{\nu^2}}{(\delta+1)\pi^* + \frac{1}{\gamma}} \end{aligned} \tag{15}$$

Robustness to the parameters γ and δ. The robustness of the estimates of β_2 and β_1 in (15) depends on the values of the parameters of relative importance γ and δ, which we cannot know from the data. We follow Altonji et al.'s (2005) and Oster's (2019) approach in trying to explore the range of variation of those two parameters, and infer the implications on the estimates of interest.

First, $\delta = \frac{\beta_0\pi_0}{\beta_1\pi_1}$ measures the relative importance of the two channels of L_0 and L_1 in terms of the influence of the IV IG on outcome Y. In our context, it is most likely that declared acquaintances (L^1) are more important than omitted acquaintances (L^0), so it is quite likely that $\delta < 1$). In the next numerical analysis, we will explore a broad range of δ from as small as 0.1 to 2.

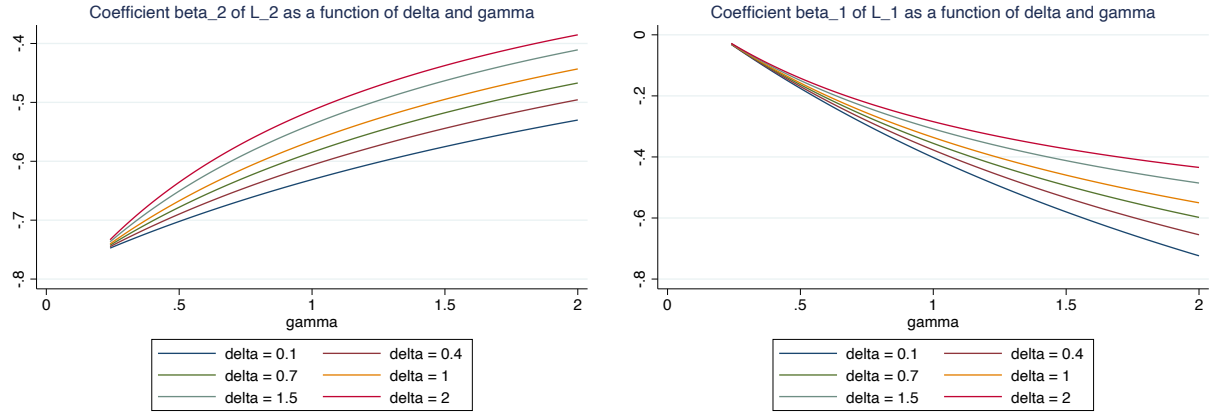
Second, $\gamma = \frac{\beta_0\rho_{01}+\rho_{u1}}{\beta_0\rho_{02}+\rho_{u2}}$ measures the relative endogeneity biases of L^1 versus L^2. Those biases involve different channels, including the unobservable component U of the outcome Y that may correlate with all measures of relationships (in the parameters ρ_{u2} and ρ_{u1}) , and the omitted variable of acquaintances L^0 (in the composite parameters $\beta_0\rho_{02}$ and $\beta_0\rho_{01}$). One intuition on those components is that the biases through L^0 are likely small, because of the substitutability nature of L^0 with L^1 and L^2 (they are mutually exclusive), and that β_0 are also likely small. Another intuition is that homophily is likely stronger for higher friendship intensity, so that $|\rho_{u2}| > |\rho_{u1}|$. Those intuitions imply an informed guess that the endogeneity bias of L^2 is likely higher than that of L^1, hence γ is likely below 1. In the numerical analysis, we consider a broad range of γ from 0.25 (L_2's endogeneity bias is four times that of L_1) to 2 (L_2's endogeneity bias is half of that of L_1).

The two plots in Figure A2 show that for very broad ranges of δ and γ, both coefficients β_1 and β_2 are clearly negative. If we are mostly concerned with the effect of friendship beyond simple

[5] β_2 cannot be solved only in the improbable case when $(\delta+1)\pi^* + \frac{1}{\gamma} = 0$ (impossible if both δ and γ are positive).

acquaintance, namely β_2, we can see that its magnitude is very strong and barely goes below 0.6 for the range of δ and γ below 1. So we can safely claim that our result is very much robust to the concern of possible violation of the exclusion restriction.

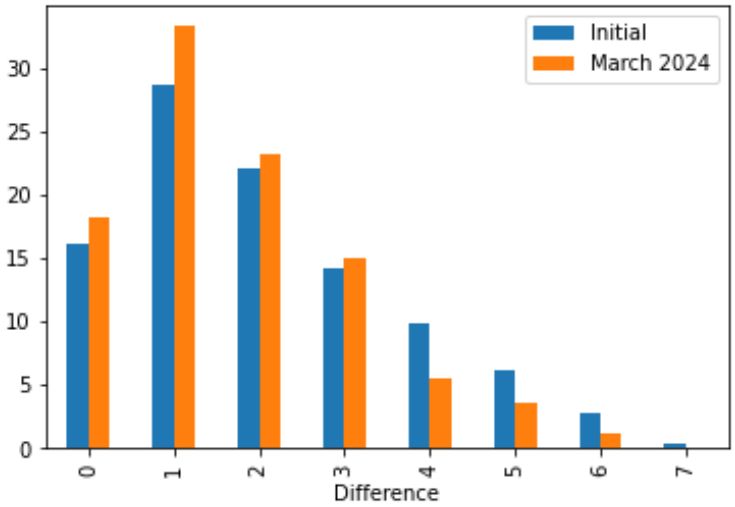
Figure A2: TRUE EFFECTS OF L^2 AND L^1 ON POLITICAL OPINIONS



Notes: The subgraphs show respectively the effect β_2 of L^2 and β_1 of L^1 on political opinions as functions of the values of δ and γ. $\delta = \frac{\beta_0 \pi_0}{\beta_1 \pi_1}$ measures the relative importance of the two channels of L_0 and L_1 in terms of the influence of the IV IG on outcome Y. $\gamma = \frac{\beta_0 \rho_{01} + \rho_{u1}}{\beta_0 \rho_{02} + \rho_{u2}}$ measures the relative endogeneity biases of L^1 versus L^2.

C Appendix: Additional Figures and Tables

Figure A3: Friendship formation by pairs of initial political opinions



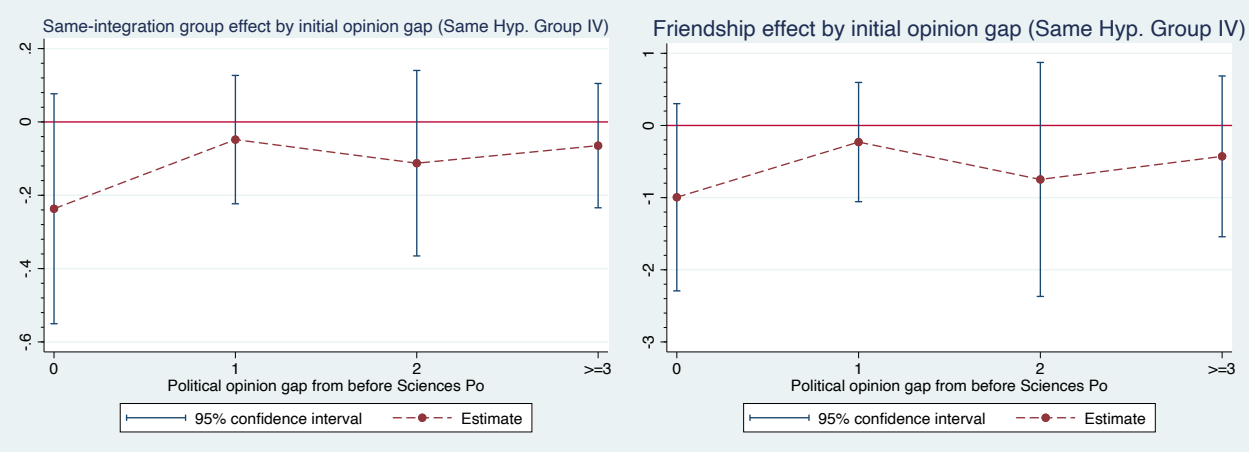
Notes: This figure shows the shares of friendship by student pair's political opinion gap, both before joining Sciences Po and at the time of the survey.

Table A4: Permutation tests of randomness of last name's first letter

Variable	Within-Group Statistics	Actual value	p-value
Initial Political Opinion (August 2013)	Within-/Between- Standard Deviation Ratio	1.806	0.410
Tuition Fees	Within-/Between- Standard Deviation Ratio	1.576	0.147
Gender	Within-/Between- Standard Deviation Ratio	1.891	0.670
Affirmative-Action Admission	Within-/Between- Standard Deviation Ratio	1.434	0.253
Second Nationality	Within-/Between- Standard Deviation Ratio per Category	2.299	0.667
Admission Type	Within-/Between- Standard Deviation Ratio per Category	4.335	0.197
Program	Within-/Between- Standard Deviation Ratio per Category	3.081	0.670
Parents' Profession	Within-/Between- Standard Deviation Ratio per Category	4.431	0.543
High School Major	Within-/Between- Standard Deviation Ratio per Category	2.797	0.110
Département of High School	Within-/Between- Standard Deviation Ratio per Category	5.872	0.810
Region of High School	Within-/Between- Standard Deviation Ratio per Category	4.622	0.810

Notes: Permutation tests over the full sample are performed over 300 Monte Carlo draws. For continuous and binary variables, the test is performed on the distribution of the ratio of within-group and between-group standard deviations. For category variables, the test is performed on the distribution of the average of this ratio across all binary (dummy) variables representing each category. p-values are computed with respect to the left tail (rejection of low within-group variation with respect to between-group variation).

Figure A4: Effects of Integration Group and Friendship by Initial Political Opinion Gap with Same Hypothetical Group IV



(A) Effect of the same-integration group indicator

(B) Effect of friendship

Notes: Panel A shows the estimates of the same-integration group indicator, instrumented by the same hypothetical group indicator, on changes in political opinion gap, based on subsamples of pairs with different initial pre-Sciences Po opinion gaps. Panel B shows the corresponding estimates of friendship, instrumented by the the same hypothetical group indicator, on changes in political opinion gap in specification (2), based on subsamples of pairs with different initial pre-Sciences Po opinion gaps. All bands show 95% confidence intervals based on dyadic clustered standard errors, which allow for error correlations between dyads sharing a common integration group.

Table A5: Effects of Integration Group on Movement of Opinion Pairs Using Alphabetical Distance

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0667**	0.0060	-0.0480**	-0.0418***	0.0008
	(0.0331)	(0.0187)	(0.0186)	(0.0122)	(0.0230)
R-Squared	0.0146	0.0018	0.0050	0.0075	0.0048
First Stage:					
Instrumental Variable:	-0.0294***	-0.0294***	-0.0294***	-0.0294***	-0.0294***
Alphabetical Distance	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Kleibergen-Paap Weak IV F-stat	367.44	367.44	367.44	367.44	367.44
Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.517	0.0968	0.228	0.038	0.182
Std. Dev. (Dep. Var.)	0.500	0.296	0.419	0.191	0.386

Notes: This table shows dyadic specifications of the effect of being in the same integration group on relating indicators of convergence, divergence, and co-movements of a pair's political opinions. The same-integration group indicator is instrumented by the alphabetical distance between two students' last names, calculated as the difference between their ranks on the alphabetical order of last names in the same cohort and further winsorized at 24. In addition, all regressions control for the alphabetical distance between their last names in the list of students who entered in 2009 to 2014. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A6: Same Integration Group Membership and Changes in Political Opinion Gaps Using Alphabetical Distance

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Change in Political Opinion Gap**				
Specification:	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Close Alphabetical Ranks
Same Integration Group	0.0816	-0.2933***	-0.1610	-0.2091**	-0.1445**
	(0.137)	(0.108)	(0.115)	(0.0693)	(0.0598)
R-squared	0.0253	0.0235	0.0056	0.0054	0.0060
First Stage:					
Instrumental Variable:	-0.0303***	-0.0269***	-0.0317***	-0.0294***	
Alphabetical Distance	(0.00154)	(0.00240)	(0.00346)	(0.00153)	
Kleibergen-Paap Weak IV F-stat	384.49	125.45	102.99	367.44	
Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	27,075	11,918	9,519	52,326	4,268
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	-1.151	1.474	0.000210	-0.267	-0.281
Std. Dev. (Dep. Var.)	1.034	0.730	0.834	1.415	1.393

Notes: This table shows dyadic specifications of the effect of being in the same integration group on Changes in Political Opinion Gaps, estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (column 4). In columns 1 to 4, the same-integration group indicator is instrumented by the alphabetical distance between two students' last names, calculated as the difference between their ranks on the alphabetical order of last names in the same cohort and further winsorized at 24. Column 5 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. In addition, all regressions control for the alphabetical distance between their last names in the list of students who entered in 2009 to 2014. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A7: Same Group Membership and Friendship Formation
Showing All Coefficients

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Friendship**				
Specification:	OLS		IV		Quasi RD
Instrumental Variable:			Same Hyp. Group	Alpha. Distance	
Sample:	Full		Full		Close Alpha. Ranks
Same Integration Group	0.1660***	0.1647***	0.1784***	0.1684***	0.1701***
	(0.0186)	(0.0186)	(0.0245)	(0.0197)	(0.0247)
Same Gender		0.0121***	0.0121***	0.0121***	0.0305***
		(0.0019)	(0.0019)	(0.0019)	(0.0106)
Both Female		-0.0082***	-0.0081***	-0.0081***	-0.0153
		(0.0027)	(0.0027)	(0.0027)	(0.0116)
Same Nationality		-0.0002	-0.0001	-0.0004	0.0417**
		(0.0020)	(0.0020)	(0.0019)	(0.0208)
Same Admission Type		0.0049***	0.0049***	0.0049***	0.0071
		(0.0013)	(0.0013)	(0.0013)	(0.0080)
Both Affirmative Action		0.0133**	0.0132**	0.0131**	0.0160
		(0.0061)	(0.0060)	(0.0061)	(0.0239)
Same Département of High School		0.0105**	0.0105**	0.0104**	0.0162
		(0.0043)	(0.0043)	(0.0043)	(0.0125)
Same Region of High School		0.0096	0.0096**	0.0096**	-0.0149**
		(0.0061)	(0.0061)	(0.0061)	(0.0244)
Same High School Major		0.0052***	0.0053***	0.0052***	0.0304***
		(0.0010)	(0.0010)	(0.0010)	(0.0087)
Diff. in Tuition Fees		-0.0005	-0.0004	-0.0004	-0.0038***
(in thousand euros)		(0.0003)	(0.0003)	(0.0003)	(0.0014)
Both Free Tuition		0.0014	0.0014	0.0014	-0.0060
		(0.0020)	(0.0020)	(0.0020)	(0.0081)
Same Parents' Profession		0.0003	0.0002	0.0002	0.0028
		(0.0013)	(0.0013)	(0.0013)	(0.0080)
Same ZIP Code		0.0148***	0.0149***	0.0149***	0.0187
		(0.0035)	(0.0035)	(0.0035)	(0.0216)
Both living in Paris		-0.0006	-0.0006	-0.0006	0.0048
		(0.0012)	(0.0012)	(0.0012)	(0.0064)
Both High School in Ile de France		-0.0058	-0.0058	-0.0057	0.0303
		(0.0060)	(0.0061)	(0.0060)	(0.0259)
Same Program		0.1351***	0.1350***	0.1352***	0.1601***
		(0.0155)	(0.0155)	(0.0155)	(0.0378)
Controls	No	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	4,268
Number of Integration Groups	52	52	52	52	52
R-Squared	0.029	0.064	0.064	0.064	0.129
Kleibergen-Paap Weak IV F-stat	79.645	78.380	52.893	74.919	50.139
Mean (Dep. Var)	0.018	0.018	0.018	0.018	0.052
Std. Dev. (Dep. Var)	0.132	0.132	0.132	0.132	0.222

Notes: This table shows the full set of coefficients for Table 7. It contains dyadic specifications of the effect of being in the same integration group (IG) on friendship formation. Column 3 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 4 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and control for the pairwise alphabetical distance within an extended sample of last names of all students that entered Sciences Po from 2009 to 2014. Column 5 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the standard set of controls.

Table A8: Same Group Membership and Friendship Formation by Friendship Intensity

Dependent Variable	Friendship with: Low friendship intensity (acquaintances)		Medium friendship intensity (friends)		Higher friendship intensity (close friends)		Highest friendship intensity (very close friends)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same integration group	0.0431***	0.0520***	0.0621***	0.0572***	0.0361***	0.0405***	0.0234***	0.0288***
	(0.0083)	(0.0130)	(0.0105)	(0.0129)	(0.0085)	(0.0081)	(0.0065)	(0.0088)
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap Weak IV F-stat.		205.879		205.879		205.879		205.879
Observations	52,326	52,326	52,326	52,326	52,326	52,326	52,326	52,326
Number of IGs	52	52	52	52	52	52	52	52
R-Squared	0.016	0.015	0.024	0.024	0.021	0.021	0.011	0.011
Mean (Dep. Var.)	0.00231	0.00231	0.00763	0.00763	0.00472	0.00472	0.00319	0.00319
Std. Dev. (Dep. Var.)	0.0480	0.0480	0.0870	0.0870	0.0685	0.0685	0.0564	0.0564

Notes: This table shows dyadic specifications relating indicators of having a friendship link with different intensity (from acquaintances in column 1 to very close friends in column 4) to the same-integration group indicator. The instrumental variable is the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A9: Same Group Membership and Friendship Formation by Frequent Activity

Dependent Variable	Friendship with most interaction in: Academic activities		Association activities		Politics-related activities		Leisurely activities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same integration group	0.0118***	0.0121**	0.0063***	0.0056*	0.0033*	0.0023	0.1118***	0.1196***
	(0.0036)	(0.0051)	(0.0024)	(0.0033)	(0.0017)	(0.0022)	(0.0186)	(0.0222)
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap Weak IV F-stat		205.879		205.879		205.879		205.879
Observations	52,326	52,326	52,326	52,326	52,326	52,326	52,326	52,326
Number of IGs	52	52	52	52	52	52	52	52
R-Squared	0.014	0.014	0.001	0.001	0.001	0.001	0.045	0.045
Mean (Dep. Var.)	0.00336	0.00336	0.000936	0.000936	0.000745	0.000745	0.0112	0.0112
Std. Dev. (Dep. Var.)	0.0579	0.0579	0.0306	0.0306	0.0273	0.0273	0.105	0.105

Notes: This table shows dyadic specifications relating indicators of having a friendship link with different most-frequent activities to the same-integration group indicator. The most-frequent activities range include academic activities (column 1), association activities (column 2), politics-related activities (column 3), and leisurely activities (column 4). The instrumental variable is the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A10: Same Group Membership and Friendship Formation by Time Spent Together

Dependent Variable	Friendship with interaction time							
	$\leq 30min$ weekly		$[30min, 1h]$ weekly		$[1h, 2h]$ weekly		$> 2h$ weekly	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same integration group	0.0900***	0.0966***	0.0309***	0.0313***	0.0321***	0.0353***	0.0415***	0.0520***
	(0.0127)	(0.0178)	(0.0057)	(0.0079)	(0.0068)	(0.0091)	(0.0076)	(0.0103)
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap Weak IV F-stat		205.879		205.879		205.879		205.879
Observations	52,326	52,326	52,326	52,326	52,326	52,326	52,326	52,326
Number of IGs	52	52	52	52	52	52	52	52
R-Squared	0.032	0.032	0.011	0.011	0.019	0.019	0.026	0.026
Mean (Dep. Var.)	0.00583	0.00583	0.00474	0.00474	0.00466	0.00466	0.00648	0.00648
Std. Dev. (Dep. Var.)	0.0761	0.0761	0.0687	0.0687	0.0681	0.0681	0.0802	0.0802

Notes: This table shows dyadic specifications relating indicators of having a friendship link with different amount of weekly interaction time to the same-integration group indicator. The interaction time groups include cases of less than 30 minutes (column 1), from 30 minutes to 1 hour (column 2), from 1 hour to 2 hours (column 3), and above 2 hours (column 4). The instrumental variable is the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A11: Descriptive Statistics on Recall Bias

		Actual (Individual) Political Opinion in 2014										
		1	2	3	4	5	6	7	8	9	10	Total
Recalled Political Opinion in 2014	1	0	1	0	0	0	0	0	0	0	0	1
	2	0	5	1	2	0	0	0	0	0	0	8
	3	1	6	19	7	3	1	0	0	0	0	37
	4	0	0	7	16	21	4	1	0	0	0	49
	5	0	0	2	7	25	6	1	0	0	0	41
	6	1	0	0	1	6	21	8	3	0	0	40
	7	0	0	0	1	0	6	12	5	0	0	24
	8	0	0	0	0	0	1	6	6	1	0	14
	9	0	0	0	0	0	0	2	1	0	1	4
	10	0	0	0	0	0	0	0	0	0	0	0
	Total	2	12	29	34	55	39	30	15	1	1	218

Notes: The joint empirical distribution of actual (horizontal axes) and recalled (vertical axes) individual political opinion in 2014, based on the main survey in 2014 and the additional survey in 2015. Individuals with a missing observation in either year are excluded.

Table A12: Recall Bias Regression on Individual Data

Dependent Variable:	Absolute Recall Bias	Recall Bias
	(1)	(2)
Actual Political Opinion in 2015	0.00426	-
	(0.116)	
Actual Political Opinion in 2014	0.00609	-
	(0.137)	
Diff. in Actual Political Opinion	-	0.574***
Between 2015 and 2014		(0.0437)
Observations	216	216
Double Group Clust.	Yes	Yes

Notes: OLS predictions of recall bias based on actual opinions, on the individual linked 2014-2015 sample, including individuals present in both surveys for which the variables "political opinion in 2015", "actual political opinion in 2014" and "recalled political opinion in 2014" are not missing. The outcome variable "Recall Bias" is calculated as recalled political opinion of 2014, as answered in the 2015 survey, minus actual political opinion in 2014, as answered in the 2014 survey. "Absolute Recall Bias" is the absolute value of Recall Bias. Standard errors are clustered at the group level.

Table A13: Friendship and Movement of Opinion Pairs

Panel A: Same Integration Group as instrumental variable

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Friendship	0.2714***	0.0574	-0.2405**	-0.1314***	0.0793
	(0.0988)	(0.0698)	(0.0949)	(0.0377)	(0.0779)
R-Squared	0.0096	0.0011	0.0001	0.0007	0.0038
First Stage:					
Instrumental Variable:	0.1646***	0.1646***	0.1646***	0.1646***	0.1646***
Same Integration Group	(0.0186)	(0.0186)	(0.0186)	(0.0186)	(0.0186)
Kleibergen-Paap Weak IV F-stat	78.381	78.381	78.381	78.381	78.381

Panel B: Same Hypothetical Group as instrumental variable

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Friendship	0.2081**	-0.0445	-0.2175**	-0.1637***	0.0613
	(0.1060)	(0.0641)	(0.1043)	(0.0561)	(0.1350)
R-Squared	0.0115	0.0015	0.0010	-0.0033	0.0042
First Stage:					
Instrumental Variable:	0.1027***	0.1027***	0.1027***	0.1027***	0.1027***
Same Hypothetical Group	(0.0173)	(0.0173)	(0.0173)	(0.0173)	(0.0173)
Kleibergen-Paap Weak IV F-stat	35.085	35.085	35.085	35.085	35.085

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.517	0.0968	0.228	0.038	0.182
Std. Dev. (Dep. Var.)	0.500	0.296	0.419	0.191	0.386

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the friendship indicator. Panel A reproduces Table 8, where friendship is instrumented by the same-integration group indicator. In Panel B, friendship is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A14: Friendship and Changes in Political Opinion Gaps

Panel A: OLS specification

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	OLS			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Friendship	0.1640	-1.0531***	-0.8592**	-0.9618***
	(0.324)	(0.311)	(0.397)	(0.314)
R-squared	0.0248	-0.0012	-0.0094	-0.0011
First Stage:				
Instrumental Variable:				
Same Integration Group	0.1610***	0.1594***	0.1772***	0.1647***
	(0.0266)	(0.0292)	(0.0321)	(0.0186)
Kleibergen-Paap Weak IV F-stat	36.68	29.89	30.56	78.38

Panel B: Robustness with IV & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Full	Close Alphabetical Ranks
Friendship	0.6119	-1.3867**	-0.3960	-0.6679***	-1.2419***	-0.8803***
	(0.4177)	(0.5508)	(0.4381)	(0.2471)	(0.3732)	(0.3339)
R-squared	0.0196	0.0226	0.0025	0.0024	0.0003	0.0043
First Stage:						
Instrumental Variable:						
Same Hypothetical Group	0.0910***	0.0921***	0.1489***	0.1027***		
	(0.0255)	(0.0197)	(0.0290)	(0.0173)		
Alphabetical Distance					-0.0049***	
					(0.0006)	
Kleibergen-Paap Weak IV F-stat	12.745	21.770	26.291	35.085	58.268	

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,075	11,918	9,519	52,326	52,326	4,268
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	-1.151	1.474	0.000210	-0.267	-0.267	-0.281
Std. Dev. (Dep. Var.)	1.034	0.730	0.834	1.415	1.415	1.393

Notes: This table shows dyadic specifications of the effect of friendship on Changes in Political Opinion Gaps, estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). Panel A reproduces Table 9 where friendship is instrumented by the same-integration group indicator. Panel B's columns 1 to 4 uses the indicator of being in the same hypothetical group as instrument for friendship, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 5 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and control for the pairwise alphabetical distance within an extended sample of last names of all students that entered Sciences Po from 2009 to 2014. Column 6 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A15: Long-Term Effects of Integration Group on Pairwise Opinions

Panel A: OLS specification

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Same Integration Group	0.0357	0.0274	-0.0053	0.0082	-0.0098	-0.0998
	(0.0267)	(0.0174)	(0.0355)	(0.0134)	(0.0201)	(0.0744)
R-Squared	0.0285	0.0063	0.0120	0.0120	0.0121	0.0101

Panel B: IV specification

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Same Integration Group	0.0503	0.0194	-0.0409	0.0315	0.0356	-0.2266
	(0.0602)	(0.0322)	(0.0467)	(0.0320)	(0.0506)	(0.1877)
R-Squared	0.0285	0.0063	0.0119	0.0118	0.0119	0.0100
First Stage:						
Instrumental Variable:	0.5460***	0.5460***	0.5460***	0.5460***	0.5460***	0.5460***
Same Hypothetical Group	(0.0666)	(0.0666)	(0.0666)	(0.0666)	(0.0666)	(0.0666)
Kleibergen-Paap Weak IV F-stat	67.20	67.20	67.20	67.20	67.20	67.20

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,628	11,628	11,628	11,628	11,628	11,628
Number of integration groups	48	48	48	48	48	48
Mean (Dep. Var.)	0.414	0.0835	0.367	0.0581	0.226	-0.0991
Std. Dev. (Dep. Var.)	0.492	0.277	0.482	0.234	0.418	1.577

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in a pair's political opinion gap (column 6), to the same-integration group indicator. In Panel B, the same-integration group indicator is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A16: Long-Term Effects of Friendship on Pairwise Opinions

Panel A: OLS specification

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**	**Change in Opinion Gap**
Friendship	0.2113	0.1622	-0.0316	0.0483	-0.0577	-0.5905
	(0.1692)	(0.1082)	(0.2089)	(0.0789)	(0.1207)	(0.4491)
R-Squared	0.0262	0.0017	0.0122	0.0109	0.0118	0.0089
First Stage:						
Instrumental Variable:	0.1690***	0.1690***	0.1690***	0.1690***	0.1690***	0.1690***
Same Integration Group	(0.0288)	(0.0288)	(0.0288)	(0.0288)	(0.0288)	(0.0288)
Kleibergen-Paap Weak IV F-stat	34.51	34.51	34.51	34.51	34.51	34.51

Panel B: IV specification

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**	**Change in Opinion Gap**
Friendship	0.2325	0.0898	-0.1892	0.1455	0.1648	-1.0474
	(0.0602)	(0.0322)	(0.0467)	(0.0320)	(0.0506)	(0.1877)
R-Squared	0.0257	0.0051	0.0111	0.0043	0.0094	0.0049
First Stage:						
Instrumental Variable:	0.1181***	0.1181***	0.1181***	0.1181***	0.1181***	0.1181***
Same Hypothetical Group	(0.0241)	(0.0241)	(0.0241)	(0.0241)	(0.0241)	(0.0241)
Kleibergen-Paap Weak IV F-stat	67.20	67.20	67.20	67.20	67.20	67.20

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,628	11,628	11,628	11,628	11,628	11,628
Number of integration groups	48	48	48	48	48	48
Mean (Dep. Var.)	0.414	0.0835	0.367	0.0581	0.226	-0.0991
Std. Dev. (Dep. Var.)	0.492	0.277	0.482	0.234	0.418	1.577

Notes: This table shows dyadic specifications that estimate the effect of friendship on indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in a pair's political opinion gap (column 6). In Panel A, friendship is instrumented by the same-integration group indicator. In Panel B, the friendship is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A17: Effects of Integration Group on Movement of Opinion Pairs among Initially Politically Similar Pairs

Panel A: OLS Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0301	.	-0.0428	-0.0404***	0.0208
	(0.0280)	.	(0.0284)	(0.0124)	(0.0157)
R-Squared	0.0179		0.0041	0.0130	0.0053

Panel B: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0631*	.	-0.0285	-0.0524**	-0.0168
	(0.0363)	.	(0.0394)	(0.0231)	(0.0435)
R-Squared	0.0178		0.0040	0.0130	0.0051
First Stage:					
Instrumental Variable:	0.5863***	0.5863***	0.5863***	0.5863***	0.5863***
Same Hypothetical Group	(0.0489)	(0.0489)	(0.0489)	(0.0489)	(0.0489)
Kleibergen-Paap Weak IV F-stat	143.8	143.8	143.8	143.8	143.8

Panels A & B's Common Features

Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	21,054	21,054	21,054	21,054	21,054
Number of Integration Groups	52	52	52	52	52
Mean (Dep. Var.)	0.289	0	0.373	0.0715	0.224
Std. Dev. (Dep. Var.)	0.453	0	0.483	0.258	0.417

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the same integration group indicator, restricted to the sample of pairs with pre-Sciences Po political opinion gap of 0 or 1. In Panel B, the same-integration group indicator is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A18: Effects of Integration Group on Movement of Opinion Pairs among Initially Politically Dissimilar Pairs

Panel A: OLS Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0330*	0.0064	-0.0248*	-0.0068	0.0115
	(0.0172)	(0.0173)	(0.0143)	(0.00441)	(0.0163)
R-Squared	0.0162	0.0022	0.0064	0.0033	0.0074

Panel B: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Same Integration Group	0.0128	-0.0158	-0.0424*	-0.0110	0.0312
	(0.0287)	(0.0159)	(0.0227)	(0.00776)	(0.0254)
R-Squared	0.0162	0.0021	0.0064	0.0033	0.0074
First Stage:					
Instrumental Variable	0.5686***	0.5686***	0.5686***	0.5686***	0.5686***
Same Hypothetical Group	(0.0389)	(0.0389)	(0.0389)	(0.0389)	(0.0389)
Kleibergen-Paap Weak IV F-stat	213.2	213.2	213.2	213.2	213.2

Panels A & B's Common Features

Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	31,272	31,272	31,272	31,272	31,272
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.671	0.162	0.130	0.0155	0.154
Std. Dev. (Dep. Var.)	0.470	0.368	0.337	0.124	0.361

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the same integration group indicator, restricted to the sample of pairs with pre-Sciences Po political opinion gap of 2 or higher. In Panel B, the same-integration group indicator is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A19: Effects of Integration Group on Changes in Political Opinion Gaps among Initially Politically Similar Pairs

Panel A: OLS specification

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	OLS			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Same Integration Group	0.0155	-0.2217***	-0.0264	-0.1728***
	(0.0413)	(0.0684)	(0.0742)	(0.0594)
R-squared	0.0085	0.0325	0.0136	0.0132

Panel B: Robustness with IV & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Full	Close Alphabetical Ranks
Same Integration Group	0.0546	-0.2960**	0.1585	-0.1540*	-0.1634**	-0.1782**
	(0.0795)	(0.1331)	(0.1791)	(0.0829)	(0.0698)	(0.0756)
R-squared	0.0250	0.0225	0.0055	0.0053	0.0053	0.0058
First Stage:						
Instrumental Variable:						
Same Hypothetical Group	0.5966***	0.5668***	0.6548***	0.5863***		
	(0.0831)	(0.0598)	(0.0786)	(0.0489)		
Alphabetical Distance					-0.0284***	
					(0.00156)	
Kleibergen-Paap Weak IV F-stat	51.58	89.80	69.47	143.8	331	

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,087	7,844	4,706	21,054	21,054	1,682
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	-0.437	1.555	0.203	0.548	0.548	0.527
Std. Dev. (Dep. Var.)	0.496	0.786	0.811	1.097	1.097	1.092

Notes: This table shows dyadic specifications of the effect of being in the same integration group on Changes in Political Opinion Gaps, restricted to the sample of pairs with pre-Sciences Po political opinion gap of 0 or 1, and estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). Panel B's columns 1 to 4 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 5 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and column 6 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A20: Effects of Integration Group on Changes in Political Opinion Gaps among Initially Politically Dissimilar Pairs

Panel A: OLS specification

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	OLS			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Same Integration Group	0.0436	-0.0641	-0.2590**	-0.0813
	(0.0586)	(0.0615)	(0.102)	(0.0606)
R-squared	0.0252	0.0144	0.0142	0.0099

Panel B: Robustness with IV & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Full	Close Alphabetical Ranks
Same Integration Group	0.0734	-0.0946	-0.2960**	-0.0863	-0.1014	-0.0855
	(0.0787)	(0.0956)	(0.1238)	(0.0699)	(0.0636)	(0.0718)
R-squared	0.0252	0.0144	0.0142	0.0099	0.0099	0.0148
First Stage:						
Instrumental Variable:						
Same Hypothetical Group	0.5678***	0.5777***	0.6081***	0.5686***		
	(0.0383)	(0.0805)	(0.0773)	(0.0389)		
Alphabetical Distance					-0.0296***	
					(0.00154)	
Kleibergen-Paap Weak IV F-stat	219.5	51.54	61.95	213.2	367.1	

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,988	4,074	4,813	31,272	31,272	2,586
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	-1.357	1.318	-0.198	-0.816	-0.816	-0.807
Std. Dev. (Dep. Var.)	1.057	0.577	0.808	1.338	1.338	1.314

Notes: This table shows dyadic specifications of the effect of being in the same integration group on Changes in Political Opinion Gaps, restricted to the sample of pairs with pre-Sciences Po political opinion gap of 2 or higher, and estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). Panel B's columns 1 to 4 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 5 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and column 6 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A21: Friendship, Integration Group, and Same Interests in Association Types

Panel A: Pairs with Similar Initial Political Opinions

	(1)	(2)	(3)	(4)
Dependent Variable:	**Both Are Members of Some Political Associations**			
	Any		Not the Same Association	
Sample	Initial Opinion Gap < 2			
Specification:	OLS	IV	OLS	IV
Same Integration Group	0.0606*		0.0044	
	(0.0347)		(0.0332)	
Friendship		0.2741		0.0197
		(0.1738)		(0.1507)
Controls	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes
Kleibergen-Paap Weak IV F-stat		35.044		35.044
Observations	9,393	9,393	9,393	9,393
Number of IGs	52	52	52	52
R-Squared	0.0407	0.0423	0.0307	0.0304
Mean (Dep. Var.)	0.176	0.176	0.144	0.144
Std. Dev. (Dep. Var.)	0.381	0.381	0.351	0.351

Panel B: Pairs with Dissimilar Initial Political Opinions

	(1)	(2)	(3)	(4)
Dependent Variable:	**Both Are Members of Some Political Associations**			
	Any		Not the Same Association	
Sample	Initial Opinion Gap ≥ 2			
Specification:	OLS	IV	OLS	IV
Same Integration Group	-0.0112		-0.0190	
	(0.0203)		(0.0156)	
Friendship		-0.0744		-0.1262
		(0.136)		(0.104)
Controls	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes
Kleibergen-Paap Weak IV F-stat		27.850		27.850
Observations	14,043	14,043	14,043	14,043
Number of IGs	52	52	52	52
R-Squared	0.0378	0.0363	0.0345	0.0330
Mean (Dep. Var.)	0.174	0.174	0.158	0.158
Std. Dev. (Dep. Var.)	0.379	0.379	0.364	0.364

Notes: This table shows dyadic specifications relating the indicator of being members of the same association with the same-integration group indicator (columns 1 and 3) and friendship (columns 2 and 4), where friendship is instrumented by the same-integration group indicator. Panel A restricts the sample to pairs of students with similar political opinions from before Sciences Po, i.e., an initial difference in opinions of less than 2. Panel B restricts the sample to pairs of students with dissimilar political opinions from before Sciences Po, i.e., an initial difference in opinions of 2 or above. Columns 1 and 2 consider all associations, while columns 3 and 4 focus on associations related to politics and policies. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, including association categorization, and the set of controls.

Table A22: Same Group Membership and Friendship Formation on "AND" Network

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Friendship**				
Specification:	OLS		IV		Quasi RD
Instrumental Variable:			Same Hyp. Group	Alpha. Distance	
Sample:	Full		Full		Close Alpha. Ranks
Same Integration Group	0.0689***	0.0681***	0.0862***	0.0765***	0.0716***
	(0.0105)	(0.0106)	(0.0149)	(0.0122)	(0.0141)
Same Gender		0.0057***	0.0057***	0.0057***	0.0064
		(0.0012)	(0.0012)	(0.0012)	(0.0065)
Both Female		-0.0027	-0.0026	-0.0026	0.0026
		(0.0018)	(0.0018)	(0.0018)	(0.0061)
Same Nationality		-0.0012	-0.0011	-0.0012	0.0238
		(0.0013)	(0.0012)	(0.0013)	(0.0221)
Same Admission Type		0.0031***	0.0031***	0.0031***	0.0044
		(0.0010)	(0.0010)	(0.0010)	(0.0044)
Both Affirmative Action		0.0073	0.0071	0.0072	-0.0035
		(0.0056)	(0.0055)	(0.0055)	(0.0154)
Same Département of High School		0.0083***	0.0083***	0.0083***	0.0175
		(0.0025)	(0.0025)	(0.0025)	(0.0119)
Same Region of High School		-0.0013	-0.0014	-0.0014	-0.0083
		(0.0035)	(0.0035)	(0.0035)	(0.0152)
Same High School Major		0.0020**	0.0020***	0.0020**	0.0132**
		(0.0008)	(0.0008)	(0.0007)	(0.0057)
Diff. in Tuition Fees		-0.0000**	-0.0000**	-0.0000**	-0.0000**
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
Both Free Tuition		-0.0011	-0.0010	-0.0011	-0.0091
		(0.0014)	(0.0014)	(0.0014)	(0.0058)
Same Parents Profession		0.0010	0.0010	0.0010	-0.0011
		(0.0006)	(0.0006)	(0.0006)	(0.0043)
Same ZIP Code		0.0118***	0.0118***	0.0118***	0.0121
		(0.0040)	(0.0040)	(0.0040)	(0.0146)
Both High School in Paris		-0.0006	-0.0006	-0.0006	0.0016
		(0.0008)	(0.0008)	(0.0008)	(0.0049)
Both High School in Ile de France		0.0029	0.0030	0.0029	0.0145
		(0.0034)	(0.0034)	(0.0034)	(0.0166)
Same Program		0.0753***	0.0752***	0.0753***	0.1105***
		(0.0111)	(0.0111)	(0.0111)	(0.0231)
Observations	52,326	52,326	52,326	52,326	4,268
Number of Integration Groups	52	52	52	52	52
R-Squared	0.0104	0.0332	0.0325	0.0330	0.0634
Kleibergen-Paap Weak IV F-stat	42.969	41.361	33.292	39.066	34.144
Mean (Dep. Var.)	0.00852	0.00852	0.00852	0.00852	0.0225
Std. Dev. (Dep. Var.)	0.0919	0.0919	0.0919	0.0919	0.148

Notes: This table replicates Table 7 on the AND network, in which a link is defined between two students when they both report each other as friend. It shows dyadic specifications of the effect of being in the same integration group on friendship formation. Column 3 uses the indicator of being in the same hypothetical group as instrument for the same-integration group indicator, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 4 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and column 5 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the standard set of controls.

Table A23: Friendship and Movement of Opinion Pairs on "AND" Network

Panel A: Same Integration Group as instrumental variable

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Friendship	0.6559***	0.1387	-0.5813***	-0.3176***	0.1916
	(0.2263)	(0.1705)	(0.2132)	(0.0973)	(0.1846)
Bounds	[0.2208,0.6559]	[0.0467, 0.1387]	[-0.1957, -0.5813]	[-0.1069, -0.3176]	[0.0645, 0.1916]
R-Squared	0.0010	0.0004	-0.0106	-0.0129	0.0026
First Stage:					
Instrumental Variable:	0.0681***	0.0681***	0.0681***	0.0681***	0.0681***
Same Integration Group	(0.0106)	(0.0106)	(0.0106)	(0.0106)	(0.0106)
Kleibergen-Paap Weak IV F-stat	41.36	41.36	41.36	41.36	41.36

Panel B: Same Hypothetical Group as instrumental variable

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**
Friendship	0.4306**	-0.0921	-0.4500**	-0.3388***	0.1269
	(0.2099)	(0.1295)	(0.1928)	(0.1202)	(0.2748)
Bounds	[0.1450, 0.4306]	[-0.0921,-0.0310]	[-0.4500, -0.1515]	[-0.3388, -0.1141]	[0.0427, 0.1269]
R-Squared	0.0087	0.0008	-0.0043	-0.0158	0.0038
First Stage:					
Instrumental Variable:	0.0496***	0.0496***	0.0496***	0.0496***	0.0496***
Same Hypothetical Group	(0.0102)	(0.0102)	(0.0102)	(0.0102)	(0.0102)
Kleibergen-Paap Weak IV F-stat	23.87	23.87	23.87	23.87	23.87

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52
Mean (Dep. Var.)	0.517	0.0968	0.228	0.038	0.182
Std. Dev. (Dep. Var.)	0.500	0.296	0.419	0.191	0.386

Notes: This table replicates Table 8 and Appendix Table A13 on the AND network, in which a link is defined between two students when they both report each other as friend. It shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions to the friendship indicator on the AND network. In Panel A, friendship is instrumented by the same-integration group indicator. In Panel B, friendship is instrumented by the indicator of being in the same hypothetical group, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A24: Friendship and Changes in Political Opinion Gaps on "AND" Network

Panel A: Same Integration Group as instrumental variable

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	OLS			
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full
Friendship	0.4448	-2.0431***	-1.6749**	-2.3246***
	(0.8554)	(0.7195)	(0.8530)	(0.7530)
Bounds	[0.1498, 0.4448]	[-2.0431, -0.6879]	[-1.6749, -0.5639]	[-2.3246, -0.7825]
R-squared	0.0231	-0.0274	-0.0232	-0.0142
First Stage:				
Same Integration Group	0.0594***	0.0822***	0.0909***	0.0681***
	(0.0158)	(0.0202)	(0.0202)	(0.0106)
Kleibergen-Paap Weak IV F-stat	14.08	16.57	20.36	41.36

Panel B: Robustness with IV & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV	IV	IV	IV	IV	OLS
Sample:	Weak Convergence	Weak Divergence	Co-movement	Full	Full	Close Alphabetical Ranks
Friendship	1.3933	-2.2843**	-0.6900	-1.3819***	-2.0356***	-2.0907***
	(0.8666)	(1.0717)	(0.7917)	(0.4779)	(0.4229)	(0.7509)
R-squared	0.0086	-0.0418	0.0010	-0.0010	-0.0094	-0.0431
First Stage:						
Same Hypothetical Group	0.0400***	0.0559***	0.0855***	0.0496***		
	(0.0141)	(0.0181)	(0.0187)	(0.0102)		
Alphabetical Distance					-0.0022***	
					(0.000376)	
Kleibergen-Paap Weak IV F-stat	8.01	9.54	21.00	23.87	35.07	

Panels A & B's common features

Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,075	11,918	9,519	52,326	52,326	4,268
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	-1.151	1.474	0.000210	-0.267	-0.267	-0.281
Std. Dev. (Dep. Var.)	1.034	0.730	0.834	1.415	1.415	1.393

Notes: This table replicates Table 9 and Appendix Table A14 on the AND network, in which a link is defined between two students when they both report each other as friend. It shows dyadic specifications of the effect of friendship on Changes in Political Opinion Gaps, estimated in subsamples of pairs that have converged (column 1), diverged (column 2), or co-moved in the same direction (column 3), as well as in the full sample (columns 4-6). In Panel A, friendship is instrumented by the same-integration group indicator. Panel B's columns 1 to 4 uses the indicator of being in the same hypothetical group as instrument for friendship, where hypothetical groups are created as consecutive 16-member groups based on the alphabetical order of last names of the entire cohort. Column 5 uses the pairwise alphabetical distance (winsorized at 1.5 times the average group size) as instrument for being in the same integration group, and column 6 focuses on the subsample of pairs within an alphabetical distance below 1.5 times the average group size. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A25: Effects of Integration Group on Pairwise Opinions No Data Restriction

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Same Integration Group	0.0294**	0.0092	-0.0217*	-0.0143**	0.0076	-0.1176**
	(0.0144)	(0.0101)	(0.0127)	(0.00643)	(0.0116)	(0.0459)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.0104	0.0019	0.0044	0.0051	0.0030	0.0049
Observations	58,311	58,311	58,311	58,311	58,311	58,311
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	0.519	0.0958	0.228	0.0380	0.180	-0.275
Std. Dev. (Dep. Var.)	0.500	0.294	0.420	0.191	0.384	1.415

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in their political opinion gap (column 6), to the same integration group indicator. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The sample used in this table does not truncate observations in which respondents take too long to answer questions on friendship. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A26: Effects of Friendship on Pairwise Opinions No Data Restriction

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Friendship	0.1867**	0.0584	-0.1379*	-0.0905**	0.0480	-0.7460**
	(0.0923)	(0.0643)	(0.0813)	(0.0394)	(0.0735)	(0.2914)
R-Squared	0.0102	0.0010	0.0038	0.0036	0.0031	0.0013
First Stage:						
Instrumental Variable:	0.1577***	0.1577***	0.1577***	0.1577***	0.1577***	0.1577***
Same Integration Group	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)
Kleibergen-Paap Weak IV F-stat	86.19	86.19	86.19	86.19	86.19	86.19
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	0.519	0.0958	0.228	0.0380	0.180	-0.275
Std. Dev. (Dep. Var.)	0.500	0.294	0.420	0.191	0.384	1.415

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in their political opinion gap (column 6), to a pair's friendship. Friendship is instrumented by the indicator whether the pair participated in the same integration group. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. The sample used in this table does not truncate observations in which respondents take too long to answer questions on friendship. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A27: Effects of Integration Group on Pairwise Opinions Two-Sided Data Restriction

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Same Integration Group	0.0414***	0.0105	-0.0313**	-0.0159**	0.0074	-0.1337***
	(0.0150)	(0.0110)	(0.0145)	(0.00708)	(0.0130)	(0.0517)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.0158	0.0014	0.0052	0.0071	0.0050	0.0050
Observations	52,326	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	0.509	0.0930	0.236	0.0408	0.182	-0.237
Std. Dev. (Dep. Var.)	0.500	0.290	0.425	0.198	0.386	1.413

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in their political opinion gap (column 6), to the same integration group indicator. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The sample used in this table truncates observations in which respondents take too long (top 2.5%) or too short (bottom 2.5%) to answer questions on friendship. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A28: Effects of Friendship on Pairwise Opinions Two-Sided Data Restriction

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Friendship	0.2551***	0.0649	-0.1929**	-0.0979**	0.0455	-0.8247***
	(0.0946)	(0.0685)	(0.0910)	(0.0424)	(0.0805)	(0.3195)
R-Squared	0.0120	0.0006	0.0025	0.0043	0.0047	0.0006
First Stage:						
Instrumental Variable:	0.1621***	0.1621***	0.1621***	0.1621***	0.1621***	0.1621***
Same Integration Group	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)
Kleibergen-Paap Weak IV F-stat	80.60	80.60	80.60	80.60	80.60	80.60
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	52,326	52,326	52,326	52,326	52,326	52,326
Number of integration groups	52	52	52	52	52	52
Mean (Dep. Var.)	0.509	0.0930	0.236	0.0408	0.182	-0.237
Std. Dev. (Dep. Var.)	0.500	0.290	0.425	0.198	0.386	1.413

Notes: This table shows dyadic specifications relating indicators of convergence, divergence, and co-movements of a pair's political opinions (columns 1 to 5), as well as the change in their political opinion gap (column 6), to a pair's friendship. Friendship is instrumented by the indicator whether the pair participated in the same integration group. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. The sample used in this table truncates observations in which respondents take too long (top 2.5%) or too short (bottom 2.5%) to answer questions on friendship. See Appendix A and Appendix Table A1 for variable and sample definitions, and the set of controls.

Table A29: Friendship Effect after Excluding Each Nationality

Dependent Variable:	Difference in Political Opinion						
Excluding:	Algeria	Germany	Belgium	Spain	Italy	Madagascar	Morocco
Friendship	-0.8391***	-0.9682***	-0.9419***	-0.9293***	-0.9673***	-0.9601***	-0.9834***
	(0.324)	(0.318)	(0.314)	(0.312)	(0.315)	(0.315)	(0.317)
First Stage:							
Instrumental Variable:	0.1608***	0.1665***	0.1651***	0.1654***	0.1668***	0.1642***	0.1648***
Same Integration Group	(0.0179)	(0.0189)	(0.0187)	(0.0186)	(0.0187)	(0.0187)	(0.0189)
Kleibergen-Paap Weak IV F-stat	80.43	77.33	77.78	78.78	79.60	77.50	76.31
Observations	51,040	51,040	51,681	52,003	51,360	52,003	51,360
R-squared	-0.0010	-0.0016	-0.0015	-0.0006	-0.0009	-0.0006	-0.0014
Mean (Dep. Var.)	-0.288	-0.259	-0.263	-0.272	-0.270	-0.267	-0.262
Std. Dev. (Dep. Var.)	1.399	1.414	1.412	1.411	1.421	1.411	1.412

Notes: This table shows dyadic specifications relating the change in a pair's political opinion gap to their friendship, instrumented by the same-integration group indicator. Each column excludes all individuals of a nationality present in the sample. Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the standard set of controls.

Table A30: Stratified Permutation Tests of Integration Group Membership

Stratified by	Number of strata	p-value
Gender	2	0.033
Second Nationality	23	0.017
Admission Type	8	0.040
Parents' Profession	26	0.007
High School Major	6	0.010
Département of High School	65	0.047
Region of High School	23	0.047
Tuition Fees	12	0.027
ZIP code	128	0.067

Notes: Permutation tests of the effect of integration group membership on the variation surveyed political opinion exogeneity by 300 Monte Carlo permutations of the full sample. The test is based on the distribution of the ratio of within-group and between-group standard deviations. The actual value of this ratio on the sample is 1.654. p-values are computed with respect to the left tail (rejection of low within-group variation with respect to between-group variation). See Appendix A and Appendix Table A1 for description of variables and sample.

Table A31: Friendship Effect after Excluding Names Starting with a Given Letter

Dependent Variable:	Political Opinion Gap								
Excluded First Letter	A	B	C	D	E	F	G	H	I
Friendship	-1.0204***	-0.9492**	-0.7060**	-1.2801***	-1.0056***	-0.9614***	-0.8779***	-0.9474***	-1.0228***
	(0.319)	(0.413)	(0.336)	(0.324)	(0.304)	(0.317)	(0.326)	(0.327)	(0.321)
First Stage:									
Instrumental Variable:	0.1675***	0.1494***	0.1682***	0.1467***	0.1667***	0.1664***	0.1696***	0.1689***	0.1668***
Same Integration Group	(0.0192)	(0.0193)	(0.0206)	(0.0158)	(0.0199)	(0.0190)	(0.0190)	(0.0198)	(0.0187)
Kleibergen-Paap Weak IV F-stat	76.16	59.87	66.39	86.12	70.28	77.05	79.41	72.85	79.82
Observations	48,205	38,781	43,956	43,071	51,360	48,516	46,360	47,895	51,040
R-squared	-0.0007	-0.0023	0.0034	-0.0032	-0.0021	-0.0009	0.0002	-0.0001	-0.0022
Mean (Dep. Var.)	-0.262	-0.274	-0.241	-0.267	-0.270	-0.259	-0.303	-0.278	-0.257
Std. Dev. (Dep. Var.)	1.413	1.430	1.420	1.401	1.415	1.423	1.429	1.428	1.416

Excluded First Letter	J	K	L	M	N	O	P	Q	R
Friendship	-0.9592***	-0.8954***	-0.7704**	-0.8287***	-0.9496***	-1.0356***	-1.0693***	-0.9895***	-0.9819***
	(0.318)	(0.322)	(0.307)	(0.311)	(0.318)	(0.326)	(0.315)	(0.310)	(0.327)
First Stage:									
Instrumental Variable:	0.1687***	0.1645***	0.1570***	0.1672***	0.1644***	0.1636***	0.1666***	0.1656***	0.1672***
Same Integration Group	(0.0189)	(0.0188)	(0.0191)	(0.0197)	(0.0187)	(0.0188)	(0.0192)	(0.0187)	(0.0197)
Kleibergen-Paap Weak IV F-stat	80.06	76.76	67.75	71.86	76.90	75.89	75.09	78.63	72.08
Observations	49,455	51,360	41,328	44,253	51,681	51,040	47,895	52,003	46,971
R-squared	-0.0011	-0.0006	0.0047	-0.0005	-0.0010	-0.0017	-0.0023	-0.0020	-0.0010
Mean (Dep. Var.)	-0.274	-0.268	-0.246	-0.283	-0.278	-0.257	-0.274	-0.264	-0.276
Std. Dev. (Dep. Var.)	1.423	1.402	1.401	1.401	1.403	1.395	1.420	1.412	1.417

Excluded First Letter	S	T	U	V	W	X	Y	Z	De/Du/D'
Friendship	-0.9865***	-0.9367***	-0.9767***	-0.8960***	-0.9682***	-0.9618***	-0.9629***	-0.9607***	-1.0709***
	(0.323)	(0.315)	(0.318)	(0.320)	(0.318)	(0.314)	(0.319)	(0.313)	(0.332)
First Stage:									
Instrumental Variable:	0.1695***	0.1699***	0.1646***	0.1657***	0.1634***	0.1647***	0.1642***	0.1648***	0.1560***
Same Integration Group	(0.0194)	(0.0187)	(0.0187)	(0.0192)	(0.0187)	(0.0186)	(0.0188)	(0.0186)	(0.0166)
Kleibergen-Paap Weak IV F-stat	76.15	82.23	77.54	74.68	76.04	78.38	76.48	78.23	88.41
Observations	46,360	48,205	52,003	49,770	52,003	52,326	51,681	52,003	49,455
R-squared	-0.0023	-0.0012	-0.0014	-0.0016	-0.0011	-0.0011	-0.0009	-0.0009	-0.0015
Mean (Dep. Var.)	-0.247	-0.265	-0.266	-0.274	-0.263	-0.267	-0.266	-0.269	-0.261
Std. Dev. (Dep. Var.)	1.417	1.422	1.415	1.421	1.416	1.415	1.417	1.417	1.427

Notes: This table shows dyadic specifications relating the change in a pair's political opinion gap to friendship, instrumented by the same-integration group indicator. Each column excludes all individuals whose family name starts with the corresponding letter, or with "De", "Du", or "D' ". Standard errors are corrected for clustering to allow for error correlations between dyads that share a common integration group. The Kleibergen-Paap Weak IV test's F statistic takes into account clustered standard errors. See Appendix A and Appendix Table A1 for variable and sample definitions, and the standard set of controls.

D Appendix: Translated survey

Consent Form

We invite you to participate in a study on the attitudes and opinions of Sciences Po students.

This information is collected solely for scientific research purposes. It will not be used for administrative or commercial purposes. Your responses will remain strictly anonymous. Once the data has been collected, your first and last names will be replaced with a code that does not allow for identification. Therefore, the individuals using this data will never have access to your real identity.

This questionnaire is conducted as part of a scientific research project funded by the European Commission (ERC Starting Grant "Trust No. 240923" by Yann Algan) and approved by the European Commission's ethics committee on data protection and anonymity. By answering the questions honestly and thoroughly, you will contribute to advancing scientific knowledge.

Participation in this questionnaire will take approximately twenty minutes.

Participating in this study entitles you to a reward and offers the possibility of taking part in future studies, which will also include a reward.

You are under no obligation to participate in this study, and if you change your mind, you may withdraw at any time. In this case, you will not receive any payment.

☐ **I have read and understood the terms and conditions of the study and accept them.**

Labeds - SciencesPo. 2014

Introduction: Your Social Network

Attention! Some questions may appear similar but require different responses. It is important to read each question carefully!

We will now ask you about your friends during your time at Sciences Po.

By participating in this study, you have a chance to receive a reward: 50 mini iPads are up for grabs!

Important! Answer this questionnaire truthfully: your chances of winning will increase based on the accuracy of your responses, which will be cross-checked with those of your peers. This verification is conducted by a computer, and the individuals you mention will never have access to your responses. Similarly, you will never know who has mentioned you.

There is no way to deduce other participants' responses at the end of this survey.

You will be asked some identical questions multiple times, but they will refer to different time periods. Please make sure to check which period each question corresponds to.

We ask that you complete this questionnaire individually and without discussing your answers with your friends. This survey is strictly personal. Additionally, we request that you complete the questionnaire in one sitting and as spontaneously as possible.

Labeds - SciencesPo. 2014

Q1.a. List your friends among the Sciences Po students from your entering class.

Example: I met Z in **September 2013** at Sciences Po, and we became friends. I list their name.

☐ I have no friends.

On the next page, we will ask you to describe your friendship (from "just an acquaintance" to "very close"), the origin of the friendly bond, and the time spent together for each friend. Your responses will be automatically cross-checked by computer with those of your friends; matching answers will increase your chances of winning a mini iPad.

Labeds - SciencesPo. 2014

If you do not wish to declare any friends, please check the box **"I have no friends among my classmates."**
Please uncheck the box **"I have no friends among my classmates,"** if you wish to mention friends.
Please list at least 3 friends.

Q1.b. Please complete the table below for each of your friends:

If you and your friend mention each other, you will both receive an additional token. Furthermore, if your answers match regarding how you met and the time spent together, you will each receive an extra token.

Example: I met my friend Z in September 2013 in a Sciences Po association, and during the 2013-2014 academic year, we spent an average of less than 30 minutes together per week. I list their name, and they also list mine in this questionnaire. We each earn one token. We select the answers "In a Sciences Po association" and "Less than 30 minutes", earning one additional token each.

Friend	**How did you meet this friend?**	**Indicate the average time spent each week during the 2013-2014 school year with this person (outside of class hours).**	**What activity did you mainly do with this person outside of class?**	**How would you evaluate your friendship with this person in 2013-2014?**
LAST NAME First Name				
LAST NAME First Name				

Please answer all the questions to proceed to the next step.

Labeds - SciencesPo. 2014

Q1.c. During the 2013-2014 school year, how many very close friends did you have outside of Sciences Po?

Q1.d. During the 2003-2014 school year, how many close friends did you have outside of Sciences Po?

Labeds - SciencesPo. 2014

Please answer all the questions to proceed to the next step.

Q3: Have you been an active member of any Sciences Po organizations?

These organizations may include officially recognized Sciences Po associations, active political parties at Sciences Po, student unions, the BDE, the BDA, the AS, and others.

(a) List up to 5 organizations in which you have been most active since you entered Sciences Po in September 2013.

Labeds - SciencesPo. 2014

If you do not wish to declare any organizations, please check the box **"I have no activity in an organization."**
Please uncheck the box **"I have no activity in an organization,"** if you wish to mention organizations. Please list at least one organization.

Q4.a: Today, how would you position yourself politically on a scale from 1 to 10, where 1 represents the far-left, 5-6 corresponds to the center-left or center-right, and 10 represents the far-right?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ I prefer not to answer

Q4.b: In July 2013, how would you have positioned yourself politically on a scale from 1 to 10, where 1 represents the far-left, 5-6 corresponds to the center-left or center-right, and 10 represents the far-right?

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ I prefer not to answer

---

Q5.a. Have you been a member of a political party in the past?

☐ Yes ☐ No

Which ones:

Party 1:

Period of membership:

Party 2:

Period of membership:

Q5.b: Are you currently a member of a political party?

☐ Yes ☐ No

Which ones:

Party:

Since when?

Labeds - SciencesPo. 2014

Q6.a. Generally speaking, on a scale from 1 to 10, would you say that most people try to take advantage of you whenever they can, or that they do their best to behave properly? (1 if everyone always tries to take advantage of you, and 10 if everyone always behaves properly. Intermediate ratings allow you to refine your judgment.)

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ I prefer not to answer

Q6.b1. Let's now talk about people coming from less developed countries to work here. In your opinion, what should the government do?

- Let anyone who wishes come.
- Let people come as long as there is work available.
- Set strict limits on the number of foreigners who can come here.
- Prohibit people from other countries from coming here.
- I don't know.

Q6.b2. Generally speaking, on a scale from 1 to 10, do you agree or disagree with the following statement: When jobs are scarce, employers should give priority to hiring French citizens? (1 if you completely disagree, 10 if you completely agree. Intermediate ratings allow you to refine your judgment.)

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐ I prefer not to answer

Q6.c. Do you think the following factors will influence your professional success in the future?

- your degree
- the knowledge and skills acquired during your studies
- your personal efforts
- your network formed at Sciences Po
- your family network

Q6.f. What rating out of 10 would you give to the organization of courses around the "triplettes" in the first year? What are the advantages?

- Facilitates integration at Sciences Po
- Makes it easier to make friends
- Allows for a more diverse group of friends
- Facilitates collaborative work
- Improves the quality of collaborative work
- Other: Please specify

End of questionnaire

Thank you for your participation.

Labeds - SciencesPo. 2014