Supplemental Appendix

Friendship Networks and Political Opinions

By Yann Algan[1], Nicolò Dalvit[2], Quoc-Anh Do[3], Alexis Le Chapelain[4], Yves Zenou[5]

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.

[1]HEC and CEPR. Email: algan@hec.fr.
[2]World Bank. Email: ndalvit@worldbank.org.
[3]Monash University, CEPR, and CESifo. Email: quoc-anh.do@monash.edu.
[4]Email: alexis.lechapelain@gmail.com.
[5]Monash University, University of Southampton, and CEPR. Email: yves.zenou@monash.edu.

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))

Table A3: Additional Descriptive Statistics of Covariates

Panel A: Monadic Independent Variables

Variable	(1) Full Sample			(2) Benchmark Sample		
	Mean	Standard deviation	Obs.	Mean	Standard deviation	Obs.
Gender (1= Female)	0.608	(0.489)	543	0.579	(0.494)	323
Honors Graduation	0.786	(0.410)	543	0.839	(0.368)	323
Tuition Fees ('000 EUR)	3,675	(3,469)	496	3,870	(3,308)	323

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	0.027	(0.162)	145,530	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 EUR)	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

B Appendix Robustness Check: Precision of retrospective question on opinion

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,[6] the bias on opinions may also relate to the rationalization 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,

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

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.[7]

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 opinion suffer a serious recall error, the resulting bias on our results would probably be small. Second, we first start in Table 4 with the measures of directional changes of opinions, which are much more robust to recall errors. Third, if the recalled opinion is biased towards actual opinion of March 2024, it would create an attenuation bias of the effect on the changes in opinion gaps. 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.

C Appendix Robustness Check: Potential violation of exclusion restriction

C.1 Overview of the control function approach

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. This subsection summarizes the main result of our approach, while the following subsection gets into the details.

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.

[7] 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.

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 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).[8]

The two plots in Figure A1 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.

C.2 Detail of the control function approach

In this subsection, we further detail 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.

C.2.1 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}.

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

[8] On γ, see precise definitions and more thorough discussion in Appendix C.

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).[9]

[9]Because of the estimated nature of $\hat{\nu}$, the calculation of the standard error of $\hat{\beta}_{CF}$ involves more steps. As we are

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 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.

C.2.2 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 $\frac{\beta_0 \pi_0}{\beta \pi}$.

mostly concerned about the potential bias when the exclusion restriction is violated, and not statistical inferences, we bypass those steps.

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.

C.2.3 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 C.2.2, 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 -1.1989, -0.0948, and 1.1241, and $\pi^* = 3.6584$. 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^*:[10]

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

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.

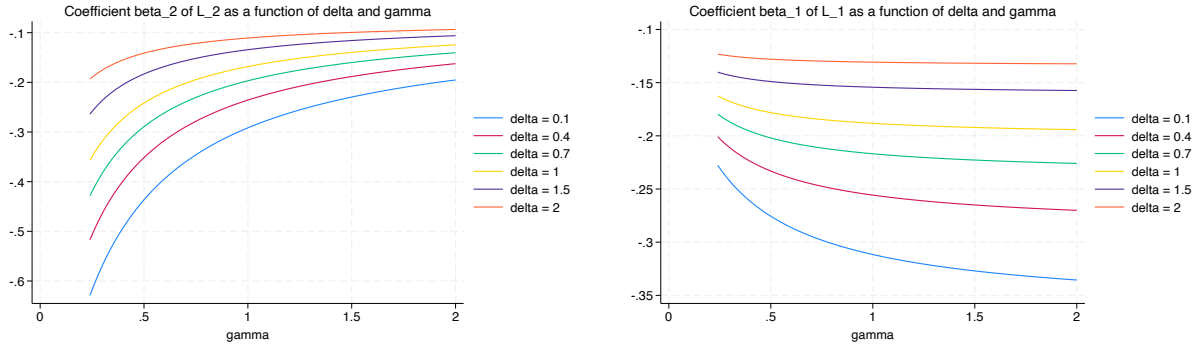
Second, $\gamma = \frac{\beta_0\rho_{01}+\rho_{u1}}{\beta_0\rho_{02}+\rho_{u2}}$ measures the relative endogeneity biases of L^1 versus L^2. Those biases involve different channels, including the unobservable component U of the outcome Y that may correlate with all measures of relationships (in the parameters 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 A1 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

[10] 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.

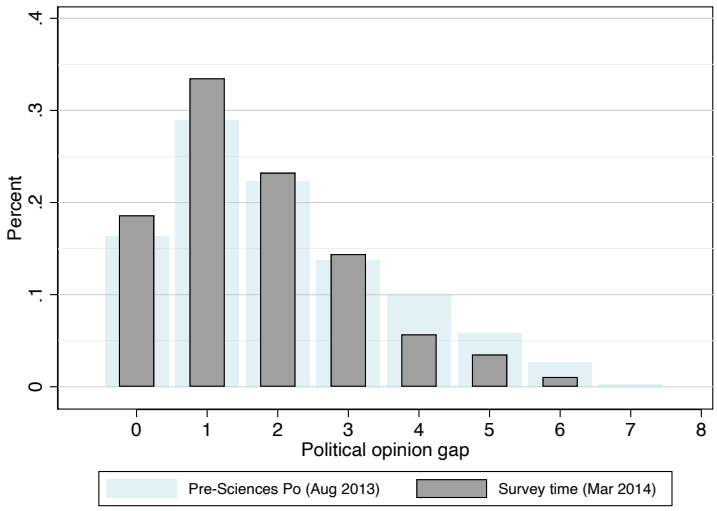
Figure A1: TRUE EFFECTS OF L^2 AND L^1 ON POLITICAL OPINIONS



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.

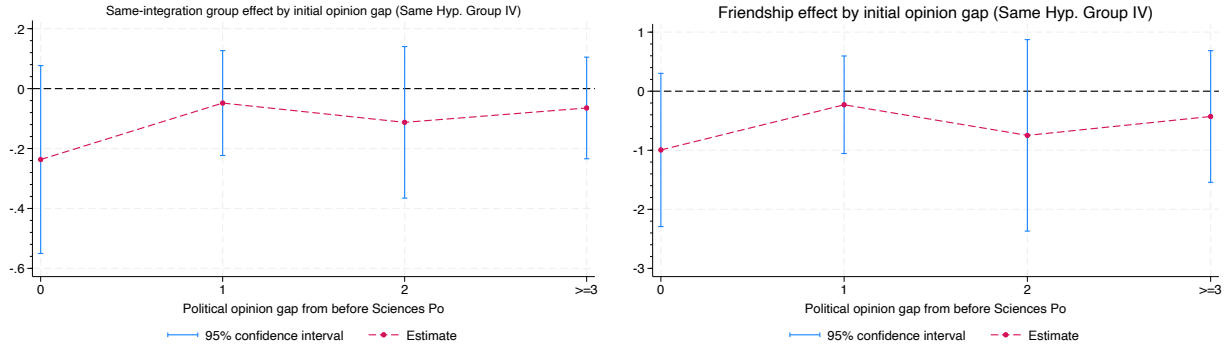
D Appendix: Additional Figures and Tables

Figure A2: 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.

Figure A3: 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 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	0.457
Tuition Fees	Within-/Between- Standard Deviation Ratio	1.576	0.190
Gender	Within-/Between- Standard Deviation Ratio	1.891	0.660
Affirmative-Action Admission	Within-/Between- Standard Deviation Ratio	1.434	0.227
Second Nationality	Within-/Between- Standard Deviation Ratio per Category	2.299	0.650
Admission Type	Within-/Between- Standard Deviation Ratio per Category	4.335	0.163
Program	Within-/Between- Standard Deviation Ratio per Category	3.081	0.650
Parents' Profession	Within-/Between- Standard Deviation Ratio per Category	4.431	0.593
High School Major	Within-/Between- Standard Deviation Ratio per Category	2.797	0.143
Département of High School	Within-/Between- Standard Deviation Ratio per Category	5.872	0.850
Region of High School	Within-/Between- Standard Deviation Ratio per Category	4.622	0.817
ZIP Code	Within-/Between- Standard Deviation Ratio per Category	6.800	0.297

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).

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: IV specification with IG

	(1)	(2)	(3)	(4)
Dependent Variable:	**Change in Political Opinion Gap**			
Specification:	IV			
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 deeper IVs & quasi-RDD specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Change in Political Opinion Gap**					
Specification:	IV					
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)	
Same Integration Group						0.1701***
						(0.0185)
Kleibergen-Paap Weak IV F-stat	12.745	21.770	26.291	35.085	58.268	84.572

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: IV specification with IG

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **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 with Same Hypothetical Group

Dependent Variable:	(1) **Weak Convergence**	(2) **Strong Convergence**	(3) **Weak Divergence**	(4) **Strong Divergence**	(5) **Co-movement**	(6) **Change in Opinion Gap**
Friendship	0.2325	0.0898	-0.1892	0.1455	0.1648	-1.0474
	(0.2902)	(0.1492)	(0.2111)	(0.1487)	(0.2331)	(0.8860)
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	24.03	24.03	24.03	24.03	24.03	24.03

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:						
Instrumental Variable:	0.0400***	0.0559***	0.0855***	0.0496***		
Same Hypothetical Group	(0.0141)	(0.0181)	(0.0187)	(0.0102)		
					-0.0022***	
Alphabetical Distance					(0.000376)	
						0.0716***
Integration Group						(0.0114)
Kleibergen-Paap Weak IV F-stat	8.01	9.54	21.00	23.87	35.07	39.80

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	0.0350**	0.0099	-0.0278*	-0.0161***	0.0103	-0.1125**
	(0.0151)	(0.0112)	(0.0152)	(0.00583)	(0.0112)	(0.0483)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dyadic Group Clustering	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.0108	0.0017	0.0051	0.0056	0.0035	0.0056
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.)	0.520	0.0953	0.226	0.0368	0.180	-0.266
Std. Dev. (Dep. Var.)	0.500	0.294	0.418	0.188	0.384	1.408

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

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Weak Convergence**	**Strong Convergence**	**Weak Divergence**	**Strong Divergence**	**Co-movement**	**Change in Opinion Gap**
Friendship	0.2204**	0.0626	-0.1754*	-0.1012***	0.0647	-0.7091**
	(0.0920)	(0.0714)	(0.0957)	(0.0363)	(0.0716)	(0.3070)
R-Squared	0.0081	0.0009	0.0029	0.0021	0.0029	0.0025
First Stage:						
Instrumental Variable:	0.1586***	0.1586***	0.1586***	0.1586***	0.1586***	0.1586***
Same Integration Group	(0.0177)	(0.0177)	(0.0177)	(0.0177)	(0.0177)	(0.0177)
Kleibergen-Paap Weak IV F-stat	80.65	80.65	80.65	80.65	80.65	80.65
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.)	0.520	0.0953	0.226	0.0368	0.180	-0.266
Std. Dev. (Dep. Var.)	0.500	0.294	0.418	0.188	0.384	1.408

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.

E 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.

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