Social Forces, 2024, **102**, 1089–1110
https://doi.org/10.1093/sf/soad129
Advance access publication date 11 October 2023
Original Article

Pathways of Peer Influence on Major Choice

Brian Rubineau[1,*], Shinwon Noh[2], Michael A. Neblo[3] and David M. J. Lazer[4]

[1]McGill University
[2]University of St. Thomas
[3]The Ohio State University
[4]Northeastern University

[*]Corresponding author: Brian Rubineau, McGill University, Bronfman Building, 1001 rue Sherbrooke Ouest, Montreal, Quebec H3A 1G5 Canada. Tel.: 514-398-5663, Email: brian.rubineau@mcgill.ca

Peers influence students' academic decisions and outcomes. For example, several studies with strong claims to causality demonstrate that peers affect the choice of and persistence in majors. One remaining issue, however, has stymied efforts to translate this evidence into actionable interventions: the literature has not grappled adequately with the fact that in natural settings, students typically *select* most of their peers. The bulk of causal evidence for peer influence comes from exogenously assigned peers (e.g., roommates) because peer effects are easier to identify in such cases. However, students do not form their most important ties for the convenience of scientific inference. In order to link theory and practice, we need to understand *which* peers are influential. We employ longitudinal, multiplex network data on students' choices of and persistence in their majors from 1260 students across 14 universities to identify likely causal pathways of peer influence via self-selected peers. We introduce time-reversed analysis as a novel tool for addressing some selection concerns in network influence studies. We find that peers with whom a student reports merely spending time, rather than—e.g., close friends, study partners, esteemed peers—consistently and potently influence their college major choice.

Peers exert significant influence on individuals' academic-related decisions and outcomes such as educational aspirations, educational attainments, and academic performance in higher education (Haller and Butterworth, 1960; Duncan et al., 1968; Woelfel and Haller, 1971; Hout and Morgan, 1975; Davies and Kandel, 1981; Buchmann and Dalton, 2002; Poldin et al., 2016; Balsa et al., 2018; Min et al., 2019). Scholarship establishing conclusively causal roles for peers in influencing real-world occupational and educational outcomes has come primarily from studies that leverage natural experiments where peers are exogenously assigned (An, 2011; Sacerdote, 2014). Roommates and classmates are two types of exogenously assigned peers commonly used

The authors acknowledge with gratitude support from the following sources: Raeed Dean, Sara Edwards, Ying Qi, and Lynne Vincent for research assistance; Jason Greenberg, Drew Margolin, Katya Ognyanova, Skyler Place, and Matt Vandenberg for data assistance; Matt Brashears, Benjamin Cornwall, Ron Ehrenberg, Roberto Fernandez, Sharique Hasan, Matissa Hollister, Michael Macy, William Minozzi, Stephen Morgan, Ben Ost, and Susan Silbey for helpful and constructive comments; and Cornell University's Institute for the Social Sciences for financial support. The opinions expressed are those of the authors, and the authors are responsible for any remaining errors.

Received: December 16, 2020. **Revised:** October 12, 2022. **Accepted:** October 21, 2022

to identify causal peer effects. Unfortunately, such approaches are necessarily mute regarding the influence of endogenously assigned (i.e., self-selected) peers, and most peers are self-selected. As a result, we know very little about which types of self-selected peers are causally involved in many important outcomes. This lack of knowledge is problematic because self-selected peers, not exogenously assigned peers, are likely to be the more influential peers. Research shows that efforts to use peer influence toward policy goals can be completely negated if those efforts are based on exogenously assigned peers while neglecting the important influence of self-selected peers (Carrell et al., 2013). The current study seeks to identify a likely causal role for self-selected peers on college students' choice of major and to compare empirically the relative influence of different types of self-selected peers.

Identifying a likely causal role for peers is a distinct empirical challenge (An, 2011). In particular, three problems loom large: the selection problem (Aral et al., 2009; Shalizi and Thomas, 2011), the reflection or simultaneity problem (Manski, 1993; An, 2016), and environmental confounding (Manski, 1993; VanderWeele and An, 2013). The selection problem comes from the fact that we play a role in choosing our own social contacts. This choice creates the possibility of reverse causation from apparent peer influence: it may be the case that we do not change because of our contacts, but that our choice of contacts is a signal that is correlated with our direction of likely change. The reflection or simultaneity problem comes from the fact that if peer influence is present, then peers influence each other simultaneously. In such a system of simultaneous changes, how do we disentangle the sources of influence from their effects? Environmental confounding refers to the fact that while we are with our social contacts, we are likely to experience similar exposures along with our contacts. When these exposures themselves are influential, we and our contacts are likely to be influenced in similar ways resulting in similar outcomes between us and our contacts.

Scholars have enlisted clever research designs and analytic techniques to solve the problems of selection, reflection, and environmental confounding to identify peer effects. Both the reflection and environmental confounding problems are commonly addressed using lagged or causally prior sources for influence (VanderWeele and An, 2013). So an alter's influence on a focal college student's academic performance, for example, would not be tested using the alter's academic performance in college, but instead using a causally prior indicator of academic performance such as the alter's Scholastic Aptitude Test (SAT) score prior to starting college (e.g., Sacerdote, 2001; Zimmerman, 2003).

To address selection, researchers have used either exogenously assigned peers to side-step the selection problem or longitudinal network data approaches to try to statistically correct for the contributions of selection (Steglich et al., 2010; Lomi et al., 2011). While the former approach supports strong causal inferences, that approach obviates the ability to identify influence from self-selected social contacts. While the latter approach addresses selection at least in part, it does not support unambiguously causal inferences (VanderWeele and An, 2013). In addition, these longitudinal methods cannot yet compare in a single analysis the relative influence of different tie types. Neither of these two approaches to addressing selection can be leveraged in service of our goal to compare pathways of peer influence among self-selected peers on college major choice.

To move forward, this paper adapts an existing method—time-reversed analysis associated with Granger causality in time-series analysis (Ding et al., 2006)—toward addressing selection for self-selected ties. In settings where information flow follows a clear temporal order, but that information process involves simultaneous spurious causation effects, a time-reversed analysis is a validated method for disentangling causal information flow from noise and confounding effects (Winkler et al., 2016). As with longitudinal approaches, our time-reversed approach only partially addresses selection. Although this approach effectively disentangles time-invariant confounding effects from influence, selection processes that are temporally ordered in a manner similar to peer influence remain a source of confounding. This approach allows us to reject some cases of illusory influence while simultaneously evaluating the relative influence on college major choice from several types of self-selected peers. To our knowledge, this paper provides the first application of this time-reversal approach to identify and remove time-invariant selection effects from peer influence.

The current study compares peer influence effects on major choice across five distinct social ties (spend time, close friends, holds in high esteem, has difficulties with—a negative tie, and study partners). The data for this study come from a longitudinal panel study over 3 years of university scholarship recipients across 14 US colleges who live in a scholarship-sponsored dormitory throughout their undergraduate years. Over 1200 students were asked to complete surveys twice each fall semester. One of these annual Fall surveys, conducted near the start of each academic year (August and September), collected demographic and major choice data. The second annual Fall survey, conducted toward the end of the semester (mid-November and December), included network roster surveys for multiple tie types.

We assess peer influence effects on students' college major choice using information about students' parents' occupations (e.g., Wang, 2021). As we develop more below, parents' occupations are a causally prior basis for peer influence and thus address both the reflection (Sacerdote, 2001) and the environmental confounding (VanderWeele and An, 2013) problems. After estimating peer influence effects in the usual time-forward analyses, we examine whether these effects remain in a time-reversed analysis to test whether our influence findings can be rejected as deriving from some spurious selection effects.

Leveraging these data and methods, we pose and answer the following focal question: which types of self-selected peers exert more influence on major choice? We find that major choice is significantly influenced by exposure to peers with whom a student spends more time, rather than a more narrowly defined type of interaction or relationship. In supplemental analyses, provided in the appendix, we also ask and answer the following five questions: (1) Are there gender differences in peer influence effects on college major choice? (2) Are STEM major adoption and STEM major persistence similarly subject to peer influence effects? (3) Do the peer selection tendencies of different types of peers affect the likely success of policies seeking to leverage peer influence? (4) Does a "wrong" analysis help reveal the strengths of the analytic approach used in this paper? (5) Is there causal evidence for peer influence on major choice in these data?

Pathways of Peer Influence

Several studies in economics have examined peer effects specifically on college major choice. Leveraging the quasi-random assignment of peers via roommates and classmates to address selection and using causally prior peer outcomes such as admission test scores to address both reflection and environmental confounding, these studies yield somewhat conflicting findings regarding peer influence. While there is evidence for significant classmate influence on a student's college major choice (De Giorgi et al., 2010) and Science, Technology, Engineering, and Mathematics (STEM) major persistence (Ost, 2010), there is no evidence for such influence from roommates (Sacerdote, 2001). Both roommates and classmates are examples of exogenously assigned peers. Why might there be significant evidence for a peer effect for one but not the other? If the peer influence flows more strongly through some types of peers than through others, then restricting our empirical scrutiny to exogenously assigned peers only ensures that consequential aspects of peer influence processes remain hidden. Identifying whether and which types of self-selected peers influence major choice is needed both to resolve these prior contradictory findings and to elucidate the social pathways of peer influence. This section examines prior scholarship regarding the pathways of peer influence.

Relationship Types and Peer Influence

Peer influence processes depend importantly upon the nature of the relationships defining the ties. Positive affective links like friendship (Heider, 1958; Lazer et al., 2010), negative links like antagonism (Litzler and Young, 2012), instrumental connections like work partnerships (Hasan and Bagde, 2013; Poldin et al., 2016; Min et al., 2019), and role-model relationships like esteem

(French and Raven, 1959; Spenner and Featherman, 1978; Buchmann and Dalton, 2002), in addition to simple exposure (Leenders, 2002; De Giorgi et al., 2010; Ost, 2010; Sacerdote, 2011; Balsa et al., 2018; Min et al., 2019), have all been shown to generate peer influence effects. There is both empirical and theoretical support for significant peer influence via all five of these relationship pathways in educational contexts and decision-making.

Theory and empirical evidence broadly favor the importance of friendship ties among students (e.g., Krackhardt, 1992; Lazer et al., 2010; Torlò and Lomi, 2017). Flap and Völker (2001) hypothesize, however, that the type of contact likely to serve as a conduit for peer influence depends on the tie type matching the type of outcome under consideration. For example, in a study comparing the strength of peer influence on academic performance—arguably a more instrumental outcome than a social one—among three of our five types of ties, exposure and study-partner ties remained strong and significant sources of influence while friendship ties were not a significant source of influence (Hasan and Bagde, 2013). Similarly, Min et al. (2019) showed that a student's test grade was positively influenced by the average test grades of classmates (exposure ties) and that of a learning companion (study-partner ties), although the effects of the former were much stronger.

Other theoretical perspectives based upon Heider (1958) suggest both the positive ties of friendship and the negative ties of having difficulties getting along would have similar significant contributions to peer influence outcomes. In the case of positive ties, being close friends with someone should make the friend's choice of major more attractive to ego. Conversely, having a negative tie with an alter would be expected to make the preferences of that alter—such as their major choice—less attractive to ego. Still, other theoretical perspectives point to the likely importance of other tie types for peer influence.

Given the heterogeneity in empirical findings and theoretical guidance, college major selection and persistence decisions are likely driven by some mix of one or more of these five network types. By measuring this range of plausibly important tie types, we are in a strong position to resolve the ambiguity in the literature as to which peer types are stronger sources of influence on college major choice. We are also in an advantageous position to address the relative inattention to self-selected peers in the extant literature on major choice. Neither theory nor empirics have fully come to grips with the importance of self-selected peers in major choice. By identifying whether and which types of self-selected peers influence major choice, we aim to offer novel insights into the role of social contacts on educational and occupational outcomes.

How Are Peers Selected?

The process by which individuals select their peers is a critical consideration for understanding and leveraging peer influence toward policy goals. Strong selection processes can often give rise to the illusory appearance of peer influence (Shalizi and Thomas, 2011) and are likely to inflate greatly estimates of influence when selection is not properly taken into account (Aral et al., 2009).

Students do have a strong tendency to befriend others with similar academic achievement for various reasons including shared values and easy access to valuable resources (Flashman, 2012; Kretschmer et al., 2018). These and related tendencies in the selection of peers can confound the identification of peer influence effects. Addressing these confounding effects is necessary, but the use of exogenously assigned peers side-steps the issue of the role of selected peers in ways that can be problematic.

Problems Arising from Neglecting Self-Selected Peers in Studies of Peer Influence: An Illustration

Carrell et al. (2009) and Carrell et al. (2013) usefully illustrate this issue. In the first study, Carrell et al. (2009) found that cadets' (at the US Air Force Academy) academic performance was significantly and causally influenced by the academic ability (SAT verbal scores) of their exogenously assigned squad-mates. This identified peer effect exhibited a contingency such that cadets with lower academic ability were more affected by their peers. These findings presented an opportunity for an intervention to try to promote the academic performance of lower ability cadets.

This intervention was tested in a randomized controlled experiment at the Air Force Academy (Carrell et al., 2013). Bi-modal squads were created with many lower ability cadets, but also many more higher ability cadets. The middle ability cadets were assigned to separate squads which served as controls. The results were surprising: the lower ability cadets did significantly *worse* than expected when assigned to squads with more higher ability cadets. How could this happen? A causal peer effect that was carefully and rigorously identified in a prior study taking place in the very same setting completely vanished in a well-designed follow-up field experiment seeking to leverage that causal peer effect to help students. It is like turning on a light switch in a room, but the room remains dark. When that happens, you know something in the causal chain is broken. What was broken here? The experimenters attributed the reversal of the peer effect to the observation that lower ability cadets in the bi-modal squads often formed ties with their fellow lower ability cadets and avoided forming ties with other higher ability cadets. That is, the endogenous effects of self-selected peer interactions had stronger and disrupting effects on the small differences they had hypothesized as arising from assigned peers. In short, they were using the wrong peers. And they did not know which peers would be the right peers to use as their prior study excluded any consideration of self-selected peers by focusing solely on exogenously assigned peers. Limiting empirical scrutiny to exogenously assigned peers gives us an impoverished and incomplete understanding of peer influence that limits our ability to leverage peer influence processes toward policy goals.

To apply findings about peer influence on college major selection and persistence, peer influence from self-selected ties must also be understood. Doing so requires addressing directly the issue of selection. This paper introduces an application of a time-reversed analysis to disentangle peer influence from selection effects among self-selected ties. We do this analysis across five different tie types. This paper is unusual both in comparing across five different tie types for peer influence—most such papers examine just one or two networks—and for including negative ties among the tie types. We find consistent evidence that exposure peers—the peers students report spending the most time with—are the most influential peers for college major and persistence choices.

Addressing Selection Via A Time-Reversed Analysis

While this paper's approach to addressing the reflection and environmental confounding problems common to peer influence studies are well established in the literature, we are introducing a new approach to addressing selection. The time-reversed analysis we propose as a tool for addressing selection at least in part is taken from established methods in the sciences and engineering where the signal from a temporally ordered process needs to be distinguished from simultaneous other signals that may be spurious and noisy but are not similarly temporally ordered (Ding et al., 2006). We argue that peer influence is a temporally ordered process, while at least some spurious selection processes are not similarly temporally ordered, and so a time-reversed analysis is a useful tool to distinguish between them.

What is peer influence, and what does it mean to argue that peer influence is temporally ordered? In both basic and broad terms, peer influence is when ego experiences a counterfactually different outcome because of ego's interaction with one or more peer alters. Commonly, this outcome is one of greater conformity, similarity, or alignment with a behavior or characteristic of those peers. Academic performance; risk behaviors such as smoking, drinking, and drug use; and occupational outcomes are among the many outcomes where peer influence toward similarity or conformity has been demonstrated with strong causal certainty (Sacerdote, 2011, 2014). That is, over time, an ego's outcomes become more similar to or consistent with ego's peer alters' behaviors or characteristics as a causal result of their interactions. And perhaps obviously, in the absence of peer interactions, ego's outcomes may change or not but do so in a manner wholly independent from any peer influence. We take this growing similarity or alignment over time in a manner that is causally attributable to peer interactions as a useful definition of peer influence.

Using this definition, what happens if we were to take this simplified movie-version of peer influence, start at the end, and play the movie in reverse, going backwards in time? Watching the peer influence movie played backwards, ego appears to become more *different* from ego's peers "after" interacting with them. The people who do not interact with peers may have changed outcomes or not in this backward-playing movie, but again, their changes are wholly independent from the influence of peers.

Peer influence is a temporally ordered process with a predictable pattern when time goes forward—greater similarity with peers—and a predictable pattern when time is reversed—greater dissimilarity with peers. The temporal order of peer influence enables a time-reversed analysis to address some threats from selection. Specifically, a time-reversed analysis can be used to disentangle or exclude the directional process of peer influence from other competing spurious and simultaneous processes that do not have the same temporal directionality (Winkler et al., 2016).

One type of selection effect in the context of peer influence is when ego selects their social alters in a manner that is an indicator of ego's likely future outcome. Relevant to the current study, consider the situation where ego is uncertain about whether to be a STEM major (e.g., mathematics) or a non-STEM major (e.g., accounting). Ego is currently a non-STEM major but still has more STEM major affinity than most non-STEM majors. Because of this affinity, the peers whom ego seeks out tend to include more STEM majors than the peers of most non-STEM majors. When we ask ego again later about their major plans, the same uncertainty exists, but this time ego reports planning to pursue a STEM major. Other non-STEM major students who are certain about their non-STEM major choice neither seek out interactions with STEM majors nor change their major to a STEM major. A similar pattern could hold for someone who is uncertain about their major but is currently a STEM major. That uncertainty can give rise to seeking out more non-STEM peers that the average more certain STEM major, and that same uncertainty can increase the likelihood that asking about major plans a second time yields a different answer than the first. This pattern of interactions and outcomes can look like peer influence, but the peer influence is illusory. The selection of more STEM peers or more non-STEM peers was an indication of ego's own uncertainty or ambivalence about their major choice. The presence of this and other forms of selection effects can give a large upward bias to estimates of peer influence via normal time-forward analyses (e.g., Aral et al., 2009).

Importantly, when we take the movie-version of this uncertainty-driven process of major change and peer selection, start at the end, and play it backwards in time, the movie looks the same as it does when going forward in time. Students who are uncertain about their major choose a more diverse group of alters and are more likely to report different majors at two points in time compared to their more certain counterparts. Unlike actual peer influence, this selection-driven illusory influence effect does not have a temporal direction. It is time invariant.

Because peer influence has directionality while at least some selection processes do not, we can use time-reversed analyses to distinguish between them. For peer influence that meets our definition, a positive regression estimate indicating increasing similarity with peers resulting from interactions will be expected to reverse in sign (indicating decreasing similarity or increasing dissimilarity) when we estimate the same regression in a time-reversed way. In the context of peer influence on STEM major choice, a time-forward analysis regresses ego's Time 2 major on ego's Time 1 major and a peer influence term. The time-reversed version of the analysis regresses ego's Time 1 major on ego's Time 2 major and the same peer influence term. Comparing the estimated peer influence term coefficients between the time-forward and time-reversed regressions, a finding that the signs of these estimates are opposite is consistent with a temporally ordered process like peer influence. If we compare the estimated peer influence terms between the time-forward and time-reversed regressions and find them to be similar to each other in terms of sign and magnitude, then that finding is inconsistent with a temporally ordered process like peer influence. Instead, it is consistent with a time-invariant process like the uncertainty-based selection process described.

In this way, the time-reversed analysis can identify and reject some apparent findings of peer influence as an illusory result of time-invariant selection processes. The appendix provides a simulation-based illustration along with the R code used to generate it of how a true peer influence process exhibits a sign reversal on the peer influence term between the time-forward and time-reversed analysis, while apparent peer influence from an illusory selection process like the one described shows a time-invariant peer influence term.

One notable feature of this time-reversed test is that it is robust to the nature of peer selection for a wide variety of time-invariant selection processes. That is, it does not matter whether a student selects her peers based on their STEM major status explicitly or based on some other unobserved but correlated factor (Shalizi and Thomas, 2011). Whether the basis of selection is observed or not, the information embedded in ego's peer choices is observable, and the time-reversed approach can distinguish from temporally ordered peer influence any time-invariant selection process giving rise to illusory peer influence.

One weakness of this time-reversed analysis is that it is a negative test. It can be the basis for rejecting an apparent peer influence finding from a time-forward analysis as actually illusory (if the peer influence term does not change sign when time is reversed). However, this time-reversed analysis is not sufficient for concluding the presence of peer influence. For example, this time-reversed analysis cannot distinguish between true peer influence and some other selection process that is also temporally ordered like peer influence. As we show, this ability to reveal even some apparent peer influence as illusory represents a useful contribution to the peer influence literature. We use this time-reversed analysis as a tool to address selection effects at least in part so that we can examine the relative influence of different types of self-selected peers.

Empirical Setting and Measures

Empirical Setting

We collected multiple reports of students' college major choices and network relationships from 1260 members of a residential fellowship program located across 14 large universities predominantly from the Midwestern United States in 2010, 2011, and 2012. Similar to Newcomb (1943), students in this program live together in a campus-specific program-owned dormitory while in college. Thus, we are examining the dominant, though not exclusive, portion of the students' social environment. We collected data at two points in time during the fall semesters of each of the three academic years. Time 1 (T1) data collection solicited survey responses from all students at the very beginning of the academic year. These Time 1 surveys were administered during the first week upon students' arrival (or return) to campus at the end of August and beginning of September (depending upon each university's academic calendar). Time 2 (T2) data collection solicited survey responses again from all students approximately 10 weeks after the Time 1 survey. Together, these data reflect the experiences and exposures of 42 (14 sites over three academic years) temporally overlapping networks including 42 distinct entering cohorts of students as they join new social systems.

In 2010, 790 students participated in this scholarship program. In the subsequent years of 2011 and 2012, the participants numbered 766 and 771, respectively. The overall response rate from respondents with matched and non-missing Time 1 and Time 2 survey responses totaled 1945 out of 2337 possible responses, or 83%. Because students continue in the program from year to year, many of the 1260 unique students appear in the data more than once. In our analyses, we cluster on individual students to ensure robust standard errors that are estimated between, and not within, individuals. The mean number of students per site is 53, with a range from 34 to 120. Of our respondents, 22% self-identify as female and 80% self-identify as white. This fellowship program

Table 1. Summary Statistics for and Correlations among Regression Variables

Variable	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) T2 major (STEM = 1)	0.30	0.46								
(2) T1 major (STEM = 1)	0.31	0.46	0.86							
(3) Female	0.22	0.42	0.02	0.01						
(4) STEM parents	0.07	0.26	0.03	0.03	0.03					
STEM Peers										
(5) Negative ties	0.05	0.16	0.06	0.07	0.01	0.03				
(6) Esteem ties	0.07	0.10	0.02	0.02	0.01	0.08	0.06			
(7) Close friends	0.07	0.12	−0.002	−0.01	0.02	0.14	0.09	0.35		
(8) Spend time ties	0.08	0.10	0.02	−0.004	0.05	0.14	0.09	0.39	0.59	
(9) Study partners	0.06	0.15	0.05	0.04	0.03	0.12	0.07	0.25	0.27	0.31

Counts of categorical variable values

Year in school		**Data year**		**Dormitory sites**					
1st years	464	2010	647	Site 1	82	Site 6	156	Site 11	69
Sophomores	465	2011	633	Site 2	277	Site 7	143	Site 12	152
Juniors	456	2012	567	Site 3	122	Site 8	148	Site 13	109
Seniors	462			Site 4	156	Site 9	98	Site 14	161
				Site 5	93	Site 10	81		

Note: *N*s for table are 1847, except for those involving study partners, where $N = 1200$.

is means-tested, targeting high school graduates from relatively low-income families, many of whom are first-generation college students. Due to this relative homogeneity in students' socio-economic status, our sample is not representative of the general undergraduate population. Thus, additional research will be needed to determine whether any identified peer influence processes are affected by the nature of our sample.

Measures

Table 1 provides summary statistics for and correlations among the variables described below and involved in our analyses.

College major (STEM/non-STEM)

The Time 1 survey included open-ended questions asking students to identify their *current or intended major.* The Time 2 survey repeated the same questions. We use ego's Time 2 major plan as a dependent variable in the analyses of peer effects on major choice and persistence. Across the 14 universities, there are many common and distinct majors. Following prior research on peer influence on major choice outcomes (e.g., De Giorgi et al., 2010; Ost, 2010), we group students' major choices into a few broad categories. We dichotomized student major choices into STEM majors and non-STEM majors. Student responses to the free-response major plan questions were independently coded using a lookup table of majors from the 14 schools and their STEM designations. The criteria for coding a major as STEM were provided by an Economics and Statistics Administration report within the U.S. Department of Commerce giving definitions of both STEM majors and STEM occupations (Langdon et al., 2011). The college major categories the report defines as STEM majors are computer majors, math majors, engineering majors, and physical and life sciences majors. College majors not within those categories, including majors in the arts, humanities, social sciences, and healthcare, are defined by the report as non-STEM. Students identifying one or more majors were coded as STEM if *any* of the responses included

Table 2. The Relationship Categories and Year-Specific Network Correlations of the 5 Tie Types

		Negative	Esteem	Close friend	Spend time
Esteem	2010	−0.111***			
	2011	−0.057***			
	2012	−0.096***			
Close friend	2010	−0.057***	0.293***		
	2011	−0.023***	0.319***		
	2012	−0.039***	0.295***		
Spend time	2010	−0.039***	0.254***	0.573***	
	2011	−0.011**	0.335***	0.569***	
	2012	−0.030***	0.275***	0.586***	
Study partner	2011	−0.008†	0.232***	0.308***	0.307***
	2012	−0.007	0.263***	0.354***	0.359***

Note: The significance for each correlation was tested using QAP[a] (Krackhardt, 1987; Hubert and Arabie, 1989).
† $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$ [a]We used 1000 replications to generate the comparison distribution, permuting network matrices within-site only to get a single correlation coefficient for all 14 sites in a single year. Although each set of 14 within-year networks are independent, the combined networks across the three data years are not independent because of the year-to-year overlap of classes of students. For this reason, we report year-specific correlations. Correlations with the study partner networks come from only 2 years of data (2011 and 2012) while all other correlations include 3 years of data (also 2010).

a STEM major. Students were coded as non-STEM if their response lacked any positive plans to major in a STEM field (including answers such as "I don't know yet").

A prerequisite for our analysis is that students do indeed report changes in major plans between the T1 and T2 surveys and in sufficient numbers in light of our dichotomization of major choice into STEM and non-STEM majors. We find this to be the case. Table 2 indicates the percentages and counts of students' persistence and switching behavior regarding their STEM versus non-STEM major plans from the T1 survey to the T2 surveys for all years. It is more common for students who begin in STEM to leave for a non-STEM major (11% of T1 STEM majors are non-STEM at T2) than for students with initial non-STEM major plans to switch to a STEM major (4% of T1 non-STEM majors do this). We test whether different types of peer relationships are better than others in explaining variation in both directions of this major switching behavior.

Parents' occupation STEM status

The Time 1 survey asked students to report their parents' occupations in two text fields (for up to two parents or guardians). These text responses of parents' occupations were similarly dichotomized into STEM and non-STEM categories of occupations. Independent coders coded these responses as STEM occupations or non-STEM occupations using a lookup table based on SOC codes. The coded occupations were compared with the SOC codes for STEM occupations defined by the same Economics and Statistics Administration report (Langdon et al., 2011). The occupational categories the report defines as STEM are computer and math occupations, engineering and surveying occupations, physical and life sciences occupations, and STEM managerial occupations. If student reports of parents' occupations for either parent included a STEM occupation, the student was coded as 1, for having a parent with a STEM occupation. Otherwise, the student was coded as 0. One of the key assumptions we make to identify peer effects is that students' parents are unlikely to change their occupations in response to their children's social contact choices at college. To the extent that this assumption is valid, collecting parents' occupations' STEM status provides a causally prior indicator for STEM status. A causally prior indicator such as this is needed to address both the reflection and environmental confounding problems for identifying peer influence. Our use of peers' parents' STEM status as a causally prior source of STEM influence bears some resemblance to the instrumental variable approach sometimes used in peer influence studies (e.g., Trogdon et al., 2008; Loh and Li, 2013). In our analysis, we are also making an assumption that peers' parents are not affecting students through other pathways (similar to the exclusion assumption with instrumental variable analyses). This assumption is putatively more reasonable than, for example, the use of rainfall levels as an instrument in much of the instrumental variable literature (see Ifft and Jodlowski, 2022; Felton and Stewart, 2022 for detailed critiques and concerns regarding weak instruments in instrumental variable analyses).

Tie types

The Time 2 survey also featured a complete roster network survey (Marsden, 1990) measuring networks based on five distinct relationship types. To capture an exposure tie, we asked respondents to indicate the other students in their dormitory with whom they "spend a lot of time." To get at potential role-modeling, students identified those whom they "hold in especially high esteem," while for negative relationships those with whom they "have difficulties getting along." For an instrumental relationship, we used the same type of relationship studied in Hasan and Bagde (2013), that is, study partners, asking students to identify those whom had "assisted [the respondent] with [her/his] academics." Finally, we asked students to identify those they consider to be "a close friend."

The response format for these network items was a set of checkboxes in the online survey. Positive responses indicated the presence of a tie, and non-responses were taken as the absence of a tie. Students could report as many or as few ties with their fellow dormitory mates as they wished. Among the five ties measured, four were measured in all three years (2010, 2011, and 2012), but instrumental ties (i.e., study partners) were measured only in 2011 and 2012. In 2010 and 2011, we also asked students "Roughly what percentage of the people you spend a lot of time around are scholars?" "Scholars" is a term referring to fellow scholarship recipients who live in the same program-owned dormitory. The results indicated that within-dormitory interactions represented about 67% of students' substantive social contacts at college, providing supportive evidence that this empirical setting captures respondents' primary social milieu.

Table 3 shows the year-specific correlations among the five measured social ties. As shown in Table 3, all five networks are significantly correlated with each other. The significance levels for the network correlations were evaluated using QAP (Krackhardt, 1987). The most strongly correlated pairs of networks across the 10 possible pairings are the "Spend Time" network and "Close Friend" network, suggesting that close friends tend to spend time together, or vice versa. The correlations among the negative affective network (i.e., "having difficulties in getting along with" ties) and the other networks (esteem, close friends, spending time, and study partner) are all negative, and 9 of the 11 correlations (excepting the correlations with the study partner network) are statistically significant at the $p < .01$ level.

Peer influence terms: percent of peers (by tie type) with STEM parents

Students' responses to these network questions reveal their self-reported set of peers for each type of relationship. The responses from the peers themselves concerning their parents' occupations at Time 1 allow us to construct variables giving the percent among each student's peers with a causally prior STEM status. These influence variables are constructed via the matrix multiplication of the row-normalized version of the self-reported alters network with the vector of students' parents' STEM occupation status (in the same binary STEM/non-STEM form). This matrix multiplication is performed once for each of the five relationships measured. Both vector and matrix are specific to a dormitory and data year. This matrix product is mathematically identical to, for example, the percent of ego's close friends with STEM parents, when the social pathway of interest is close friendship. Students without any reported contacts of a particular tie type are omitted from the analysis of peer influence via that tie type. The five tie types (spending

Table 3. Major Change Summary

_T1_ \ T2	Non-STEM major	STEM major	Total
Non-STEM major	66.4% (1227)	2.5% (46)	69% (1273)
STEM major	3.6% (66)	27.5% (508)	31% (574)
Total	70% (1293)	30% (554)	100% (1847)
	From T1 STEM to T2 non-STEM	**From T1 non-STEM to T2 STEM**	**Total changers**
First-year students	7.5% (35/464)	4.7% (22/464)	12.3% (57/464)
Sophomores	3% (14/465)	3% (14/465)	6.0% (28/465)
Juniors	2.0% (9/456)	1.3% (6/456)	3.3% (15/456)
Seniors	1.7% (8/462)	0.9% (4/462)	2.6% (12/462)

Note: The apparent total of students here, 1847 differs from 1260 because in each of the 3 data years (2010, 2011, and 2012), we observed major change behaviors over the course of the first 10 weeks of the academic year for all students in each of 14 dormitories. Because of this, we have multiple years of T1 to T2 major choice observations for many students. These 1847 observations of STEM major choice change or persistence during the first 10 weeks of the academic year come from 1260 distinct individuals.

time, esteem, not getting along, close friendship, and study partners) define a set of five peer composition variables that serve as our tie-specific peer influence terms.

Using peers' parents to measure peer influence likely introduces noise that biases estimates of actual peer influence toward a null effect. This null bias obtains because the distinction between STEM-promoting peers and non-STEM-promoting peers is likely blurred when defining the distinction based on peers' parents rather than the peers themselves. This null bias also makes using peers' parents' STEM status a conservative test of peer influence.

This formulation of the peer influence term also constrains the types of influence that may be detected by making a linear-in-means assumption about the nature of peer influence (An, 2011; Sacerdote, 2014). If influence operates in a manner such that as the peer influence term (percent STEM parents among a focal student's peers of a particular tie type) increases, the probability of adopting (or leaving) STEM increases commensurately, then our formulation can potentially detect this influence. However, if influence operates differently (e.g., in a non-linear manner such as requiring uniformity or unanimity among peers before experiencing influence), then our formulation is less likely to be able to detect such influence processes. This constraint is another reason to interpret our test for peer influence as a conservative test, as peer influence operating in a manner other than linear-in-means could still be present but not be detected. We use this approach first because it has successfully revealed peer influence processes in other settings with similarly constructed datasets (e.g., Carrell et al., 2009; Lazer et al., 2010; Poldin et al., 2015, 2016), and second because peer influence studies have few useful or successful alternatives to this linear-in-means assumption (Sacerdote, 2014).

When defining our peer influence variable, we could have used students' peers' own Time 1 STEM major plans. Had we done so, our analysis would look much like the ones conducted by Poldin et al. (2015, 2016). The disadvantage of that approach is the inability to address the reflection problem and the threat of environmental confounding in the context of self-selected ties. The risk is that even Time 1 STEM major plans could be visible to students and influence their choices of peers. We illustrate the deficits of this approach in a supplemental analysis presented in the appendix and described as "the wrong analysis." Our use of peers' parents' STEM status represents a more certain solution to the reflection and environmental confounding problems in the presence of self-selected ties than other approaches such as using peers' lagged outcomes (VanderWeele and An, 2013; An, 2016).

Controls

We control for three individual-level variables including ego's self-identification as female (female = 1, not female = 0), ego's Time 1 major plans (STEM = 1, non-STEM = 0), and ego's parents' occupational STEM status (any STEM occupation among parents = 1; otherwise, 0). We also include dummy variables (fixed effects) for site (1 through 14), year of data collection (2010, 2011, and 2012), and ego's year in school (1 = first year, 2 = sophomore, 3 = junior, 4 = senior).

There is substantial scholarly interest in gender differences in both peer influence (Rubineau, 2007; Han and Li, 2009; Balsa et al., 2018; Raabe et al., 2019) and STEM major choice (Sassler et al., 2017; Mouganie and Wang, 2020; Zölitz and Feld, 2021). We test for both by adding interactions between the female-identifying variable and the peer influence term—percent of peers with STEM parents (Holland, 2008) in analyses described in the appendix.

Analytic Strategy

Our analytic goal focuses on answering the question: which types of self-selected peers exert more influence on major choice? Prior research provides evidence consistent with expecting peer influence effects on major choice via all five types of ties described above—spend time, close friends, holds in high esteem, has difficulties getting along with, and study partners. To answer the question of which types of self-selected peers are more influential on students' major choice, we use a network effects model (Leenders, 2002). A stylized version of this model is provided as equation 1:

$$\text{Ego's T2 Major} = \text{Ego's T1 Major} + \text{Peer Influence Term} + \text{Controls} \tag{1}$$

This model is based on established network influence models (Friedkin and Johnsen, 1990: 320 equation 6; Marsden and Friedkin, 1993: 135, equation 6). In this stylized estimation model, "Major" refers to the binary STEM/non-STEM indicator for major plans. Because of the binary nature of the dependent variable, we use logistic regression to estimate peer influence effects. We also replicate our regressions using linear probability models. Because our data include multiple observations of the same student (e.g., T1 to T2 major changes when a student is a first-year student in 2010, for example, and also when they are a second-year student in 2011), whenever we estimate regressions where the same student appears more than once in our data, we cluster on individual students to estimate standard errors.

The "Peer Influence Term" is the tie-specific percent of peers with STEM parents. This analysis entails five distinct peer influence terms, one for each of the five types of ties examined. In addition to using peer influence terms based on self-selected ties, this analysis also uses the following control variables: dormitory and data year fixed effects, whether the focal student identifies as female, the focal student's parents' STEM occupation status, and the focal student's year in school.

The regression analysis results for our peer influence terms provide an initial indicator for likely peer influence effects. Because these peer influence terms are based upon self-selected ties, selection concerns remain (Manski, 1993; Mouw, 2006; Hasan and Bagde, 2013). We use the time-reversed analysis described above to address some of these selection concerns. The time-reversed model simply regresses students' Time 1 major on their Time 2 major plus the same peer influence terms and controls. If a positive and significant peer influence estimate from the initial time-forward analysis remains positive and significant in the time-reversed analysis, then that effect can be rejected as coming more from some time-invariant selection process than from a peer influence process. If instead, the positive and significant peer influence estimate from the initial time-forward analysis changes sign in the time-reversed analysis, then that result remains consistent with (but not conclusive of) a temporally ordered peer influence process.

We conduct these analyses—both time forward and time reversed, using both logistic and linear regression specifications—for all five of the examined tie types. We examine each tie type separately, as well as estimating models including the tie types simultaneously. Because one of the examined tie types (study partner) is not available for one of our data years, our examination of the simultaneous effects of the peer influence terms involving the different tie types includes both the analysis of all five tie types using the available 2 years of data as well as the analysis of the other four tie types using all 3 years of data. We are looking for consistent evidence of peer influence across these different model specifications.

In the appendix, we conduct two sets of additional analyses on the operation of peer influence on major choice via self-selected ties. First, in response to the substantial scholarly interest on the topic of women in STEM fields (Beede et al., 2011; Raabe et al., 2019; National Academies of Sciences, Engineering, and Medicine, 2020), we test for gender differences in peer influence effects by adding interaction terms (female × the peer influence terms) to the analysis (Holland, 2008). Second, we test for distinct peer influence effects on STEM adoption and STEM persistence. Note: we do not use stochastic actor-oriented modeling (SAOM) because we only observe students' networks once per year. SOAM requires multiple network observations over time with only small changes between observations (Lomi et al., 2011; Flashman, 2012; Kretschmer et al., 2018).

Results

For each of the five ties in our study, we test for peer influence using the STEM status of peers' parents. This yields five models of peer influence with each peer influence term in a separate model. In addition, we include a model with the peer influence terms from all four tie types available from all three data years simultaneously (all but study partners), as well as a model with all five peer influence terms simultaneously that uses data from 2011 and 2012 only. The peer influence term estimates from the logistic regressions of these models are given in Table 4. The linear regression estimates for the same set of models are given in Table 5.

The results shown in Tables 4 and 5 reveal consistently significant positive influence from the *spend time* peers on students' major choices. The spend-time peer effects also dominate in the models involving multiple peer types simultaneously. The substantive interpretation of this exposure effect is that if the proportion of STEM peers in a student's exposure peer group increases by a standard deviation (0.10), their chance of choosing a STEM major at Time 2 increases by as much as 4.9% (per Model 7 of Table 5).

The analysis indicating spend time peers as a potential pathway for peer influence has not yet been evaluated for selection effects. To do so, we conduct a time-reversed analysis. The results from both the logistic and linear time-reversed analyses are given in Table 6. In a true peer influence process, the influence term would be significant in the time-forward model and would have the opposite sign in the time-reversed model. In a time-invariant selection-driven process giving rise to illusory peer influence, the influence term would be similar (e.g., both positive) in both the time-forward and time-reversed models. Table 6 shows the peer influence term for the spend time tie to be negative in all six models with that term, and significantly negative in all but one of the six estimates. This pattern of results provides confidence that our identified peer effect from spend time peers is more likely to be driven by actual peer influence than by a time-invariant selection process. Based on our analyses, we identify spend time peers as a likely causal source of peer influence on major choice.

Discussion

This study goes beyond prior work showing that peers play a causal role in the real-world decisions and behaviors of individuals regarding college major choice by identifying the pathway of that peer influence via self-selected peers. We find that peer influence for major choice exists and flows through a distinctive social pathway—spending time together, rather than the other four tie types that we examined. A standard deviation change in the composition of these spend

Table 4. Pathways of Peer Influence Regression Results, Logistic Regressions

	Negative (1)	Esteem (2)	Friend (3)	Time (4)	Study (5)	4 Ties (6)	5 Ties (7)
Constant	−3.138***	−2.919***	−3.041***	−3.347***	−4.899***	−3.015***	−4.822***
	(.614)	(.676)	(.740)	(.717)	(.771)	(.613)	(1.198)
T1 STEM status	5.258***	5.492***	5.485***	5.507***	5.715***	5.341***	5.868***
	(.318)	(.249)	(.254)	(.250)	(.368)	(.346)	(.599)
Female	.211	.332	.281	.246	.064	.089	−.685
	(.345)	(.257)	(.271)	(.264)	(.299)	(.356)	(.478)
STEM parents	−.156	−.264	−.280	−.399	−.298	.034	−.041
	(.454)	(.339)	(.350)	(.328)	(.402)	(.521)	(.538)
STEM peers							
Negative tie	−.479					−.132	.255
	(.585)					(.616)	(.822)
Esteem		.408				−1.299	−.849
		(.764)				(1.630)	(2.245)
Close friend			.349			−1.677	−4.773*
			(.719)			(1.438)	(1.957)
Spend time				2.173**		4.776*	10.072***
				(.818)		(1.891)	(3.032)
Study partner					1.379*		−.198
					(.622)		(.829)
Site fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Data year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	923	1584	1525	1590	808	831	419
AIC	474	714	694	728	390	426	234

Note: Estimates use clustered standard errors by student. The "STEM peers" term is based on the STEM status of peers' parents. $^{\dagger}p < .10$ $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

time peers (toward a more or less STEM-oriented set of peers) is associated with an almost five-percentage-point change in the probability of changing one's major in the same direction. Further analyses, provided in the appendix, examine both the possible gendered nature of peer influence (we found no gender differences in peer influence) and the process by which students form ties to inform the practical implications of our findings. We find that peer influence is more likely to be successful to affect STEM major persistence than STEM major adoption. This work makes theoretical, practical, and methodological contributions.

Contributions to Theory

Classical sociological theories of social influence suggest that relationship strength and affective closeness may be important relationship factors for social influence (Friedkin, 2006). Our findings join a growing body of evidence challenging this view. For major choices in our setting, it is neither our friends nor those we esteem who hold the greatest influence over our own consequential decisions, but rather those with whom we interact the most. This may reflect a social comparison process that is driven by knowledge of and experience with the alter rather than affect or respect. That is, when in Rome, do as the Romans *you spend time with* do.

Research on core discussion networks—the set of alters with whom the ego discusses important matters (Marsden, 1987; Small, 2013)—has come to a similar conclusion. Contrary to the conventional wisdom that people seek those closest to them such as spouse and close friends to speak about important issues, research shows that the core discussion network is not dominantly composed of strong ties (Pescosolido 1992; Small, 2013). Rather, the availability of alters plays a crucial role in whom ego decides to confide. For example, the study of mothers whose children

Table 5. Pathways of Peer Influence Regression Results, Linear Regressions

	Negative (1)	Esteem (2)	Friend (3)	Time (4)	Study (5)	4 Ties (6)	5 Ties (7)
Constant	.049	.064†	.056	.038	−.026	.062†	−.013
	(.041)	(.038)	(.042)	(.041)	(.044)	(.037)	(.068)
T1 STEM status	.827***	.849***	.849***	.849***	.825***	.824***	.801***
	(.021)	(.016)	(.016)	(.016)	(.022)	(.023)	(.033)
Female	.013	.016	.015	.013	−.001	.007	−.024
	(.021)	(.014)	(.015)	(.015)	(.018)	(.021)	(.028)
STEM parents	−.011	−.016	−.017	−.025	−.030	.003	−.010
	(.029)	(.020)	(.021)	(.020)	(.029)	(.032)	(.039)
STEM peers							
Negative tie	−.029					−.005	.016
	(.038)					(.035)	(.047)
Esteem		.015				−.057	−.123
		(.041)				(.079)	(.153)
Close friend			.012			−.094	−.229†
			(.042)			(.081)	(.127)
Spend time				.124*		.263*	.487*
				(.053)		(.113)	(.191)
Study partner					.082*		.034
					(.038)		(.060)
Site fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Data year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	923	1584	1525	1590	808	831	419
Adjusted R^2	.699	.738	.736	.735	.720	.697	.682

Note: Estimates use clustered standard errors by student. The "STEM Peers" term is based on the STEM status of peers' parents. $^{\dagger}p < .10$ $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

were enrolled in childcare centers found that mothers discussed critical issues with other mothers they regularly met at the center even though they did not consider those mothers their strong ties (Small, 2009). In a similar vein, Small et al. (2015) found that the core discussion networks tend to change quickly as people enter new institutional environments and develop different routines. These studies suggest that the decision to mobilize which type of tie to discuss important topics is more strongly affected by the availability of interaction opportunities in one's social routines.

Contributions to Practice and Policy

Our examination in this paper's appendix of the peer influence predictors of STEM adoption versus STEM persistence allow us to distinguish between high-potential (e.g., STEM persistence) and low-potential (e.g., STEM adoption) opportunities to promote gender equity among STEM majors via peer influence. Increasing students' exposure to STEM students could help to attract more students to STEM majors. However, explicit policies aimed at promoting such interactions may be thwarted both by students' tendency to spend time with similar others in terms of STEM major status. However, once a student has chosen a STEM major, interventions involving spend-time peers can be expected to support STEM persistence. Also, these peer influence effects do not appear to be gendered but affect the decisions of students similarly regardless of gender identification. The *absence* of significant gender homophily in forming spend time relationships suggests that policies that increase female STEM majors' exposure to other STEM majors, both women and men, could promote women's persistence in STEM majors.

Table 6. Time-Reversed Analysis of Social Influence

	Logistic regression models						
STEM peers	**Negative (1)**	**Esteem (2)**	**Friend (3)**	**Time (4)**	**Study (5)**	**4 Ties (6)**	**5 Ties (7)**
Negative tie	.708					.428	−.193
	(.555)					(.527)	(.720)
Esteem		−.213				.847	1.467
		(.730)				(1.087)	(2.246)
Close friend			−.705			.899	1.923
			(.697)			(1.014)	(1.315)
Spend time				−3.206**		−3.114†	−3.322
				(.989)		(1.706)	(3.121)
Study partner					−1.022		−.266
					(.677)		(.887)
	Linear regression models						
STEM peers	**Negative (8)**	**Esteem (9)**	**Friend (10)**	**Time (11)**	**Study (12)**	**4 Ties (13)**	**5 Ties (14)**
Negative tie	.054					.032	−.011
	(.042)					(.039)	(.056)
Esteem		−.016				.039	.138
		(.042)				(.081)	(.160)
Close friend			−.044			.055	.147
			(.041)			(.081)	(.134)
Spend time				−.170**		−.208†	−.337†
				(.052)		(.112)	(.199)
Study partner					−.044		.034
					(.039)		(.054)
Site fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Data year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls in model[a]	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N[b]	923	1584	1525	1590	808	831	419

Note: Both logistic and linear models results, regression constant, and controls are estimated but not shown. The time-reversed analysis regresses T1 major on peer influence terms controlling for T2 major. Estimates use clustered standard errors by student. The "STEM peers" term is based on the STEM status of peers' parents. Fixed effects (site and year) and controls apply to all models in the table. †$p < .10$ *$p < .05$ **$p < .01$ ***$p < .001$
[a]Controls: student's T2 (STEM) major, female, student's parents' STEM status. [b]Ns are identical in both logistic and linear model version.

Methodological Contributions

In addition to these theoretical and practical insights, this study contributes methodologically. We present the first application of a time-reversed analysis to identify and reject illusory peer influence effects arising from time-invariant selection processes. This method also highlights the primary benefit of using a causally prior indicator of peer influence (e.g., peers' parents' STEM occupation status) rather than a more proximate but selection-compromised measure of peer influence (e.g., peer's lagged STEM major status—the "wrong" analysis, provided in the appendix). By conducting parallel time-forward and time-reversed analyses using both types of peer influence terms, we show that the apparent peer influence results obtained from the time-forward peer influence analysis that used the more proximate but selection-compromised peer influence term were indeed an illusory result from time-invariant selection effects. In contrast, our finding of the influential role of spend time peers is not attributable to time-invariant selection effects. Our analytic strategy of using a causally prior indicator of peer influence with a time-reversed analysis addresses both reflection and at least some selection and thus provides improved causal certainty for our findings.

Limitations

We note three important limitations regarding the findings of our study. First and foremost, because self-selected ties are neither assigned nor the result of some exogenous manipulation, we cannot definitively rule out the possibility that other unobserved processes are giving rise to apparent peer influence effects. For example, if a Time 1 non-STEM major opts to take a popular STEM class, and this class influences both the composition of that student's social alters and that student's choice to become a STEM major, that unobserved exposure could look like a peer influence effect in our data.

Using our data, we cannot reject the possibility that unobserved non-influence processes contribute to our findings of peer influence. Still, there are some features of our study along with some of the requirements for an unobserved process to compromise our findings that make such concerns improbable albeit possible. First, whatever popular class, teacher, event, or other exposure attracts students likely to change their major (to or from STEM) and connects them with other students whose parents have the appropriate STEM status, this connection needs to operate disproportionately via the spend time type of peer connection, and not, for example, via the other types of peers examined here such as study partner or close friend. Second, this unobserved class, teacher, event, or other exposure process must operate similarly (i.e., confound influence and operate primarily via the spend time peer connection) at multiple sites across our 14 different university settings. Otherwise, this unobserved process would entail that our findings are driven primarily by a single site. We can easily test whether this is the case by re-running our influence analysis 14 times, excluding each site in turn. When we do this, we find that spend time peers remain significantly influential in all 14 results (see appendix). Finally, this unobserved process operating similarly at multiple sites must remain nonetheless variable enough (i.e., making some students more likely to leave STEM and others more likely to choose STEM majors) within individuals and across sites and data years so as not to be captured by our fixed effects for site, data year, and year in school. An unobserved process across our sites meeting all of these criteria could indeed give rise to confounded findings of influential spend time peers. For this reason, although we view causal peer influence from spend time peers as being likely, we cannot claim such causal effects with conclusive certainty.

We can and do conduct an explicitly causal analysis of peer effects in our dataset using exogenously assigned peers. This analysis is provided in the appendix and shows significant causal peer effects for first-year students from the more senior students in their dormitories.

A second limitation is that this study identifies the pathways, but not strictly the mechanisms of peer influence. We have not addressed the mechanism of peer influence because we cannot discern the mechanism from our data. The literature on peer influences on major and career plans suggests several possible mechanisms. A common view in this literature is that peers may influence choices and aspirations either through a normative "definer" role or through a "modeler" role (Spenner and Featherman, 1978; Buchmann and Dalton, 2002). This restricted dichotomy assumes ego to be a decision-maker subject to peer influence but does not encompass all the possible mechanisms of influence. For example, social capital and ego's embeddedness could generate peer effects as well. We identify the pathway of peer influence on STEM major choice as flowing through exposure ties, but we cannot completely settle whether the influence results from modeling, defining, embeddedness, or some other mechanism, except insofar as any such modeling is not based on esteem.

Finally, our results are limited in their generalizability. While we are confident that we have identified a likely causal peer effect among self-selected spend time peers on students' choice of major in our setting, as noted above, our empirical context is not representative of the undergraduate population generally. In addition, given evidence that peer influence processes may follow different pathways for different outcomes (Flap and Völker, 2001), we cannot claim that all peer influence operates through exposure or spend time peers. For these reasons, replications in other populations and for other outcomes is needed before claiming broader generalizability of the pathway of peer influence we identify—exposure or spend time peers.

This study's research design and analytic strategies address reflection, environmental confounding, and at least some selection while examining peer influence from self-selected peers. Acknowledging that the possibilities of some unobserved environmental confounding and time-ordered selection processes cannot be entirely eliminated, the likely causal pathway of peer influence on students' major choices is via the peers with whom they spend the most time.

About the Authors

Brian Rubineau is an Associate Professor of Organizational Behavior at the Desautels Faculty of Management at McGill University. Using social network analysis, mathematical modeling, and computational simulation approaches, his research investigates how informal social dynamics contribute to inequalities in organizations, occupations, and labor markets.

Shinwon Noh is an Assistant Professor of Entrepreneurship in the Opus College of Business at the University of St. Thomas. Her research explores the emergence of new occupations and organizational fields using organization theory. She is particularly interested in examining the intersection of cultural entrepreneurship and community organizations.

Michael A. Neblo is the Arts & Sciences Alumni Professor in the Department of Political Science and (by courtesy) Philosophy, Communication, and Public Affairs at The Ohio State University where he also directs the Institute for Democratic Engagement & Accountability (IDEA). His research focuses on the dynamics between institutions and individuals with special reference to deliberation in democratic politics: michaelneblo.net.

David M. J. Lazer is University Distinguished Professor of Political Science and Computer and Information Science, Northeastern University, and an elected fellow to the National Academy of Public Administration. His research focuses on computational social science, social networks, democratic deliberation, algorithm auditing, and collective intelligence. More information is available at davidlazer.com.

Supplementary Data

Supplementary material is available at *Social Forces* online.

Data Availability

The data underlying this article cannot be shared publicly due to agreements made with the focal scholarship program as well as the approved data protection procedures established with the institutional review board (IRB) for ethical research prior to data collection. Anonymized version of the data used in this study will be shared on reasonable request to the corresponding author.

References

An, Weihua 2011. "Models and methods to identify peer effects". In *The Sage Handbook of Social Network Analysis*, edited by John Scott and Peter J. Carrington, pp. 514–32. Thousand Oaks, CA: Sage Publications. https://doi.org/10.4135/9781446294413.

An, Weihua 2016. "On the Directionality Test of Peer Effects in Social Networks." *Sociological Methods & Research* **45**(4):635–50. https://doi.org/10.1177/0049124115582271.

Aral, Sinan, Lev Muchnik and Arun Sundararajan 2009. “Distinguishing Influence-Based Contagion from Homophily-Driven Diffusion in Dynamic Networks.” *Proceedings of the National Academy of Sciences* **106**(51):21544–9. https://doi.org/10.1073/pnas.0908800106.

Balsa, Ana, Néstor Gandelman and Flavia Roldán 2018. “Peer and Parental Influence in Academic Performance and Alcohol Use.” *Labour Economics* **55**:41–55. https://doi.org/10.1016/j.labeco.2018.08.010.

Beede, David, Tiffany Julian, David Langdon, George McKittrick, Beethika Khan and Mark Doms 2011. “Women in STEM: A Gender Gap to Innovation.” *Economics and Statistics Administration Issue Brief No. 04-11,* 1–11. https://doi.org/10.2139/ssrn.1964782.

Buchmann, Claudia and Ben Dalton 2002. “Interpersonal Influences and Educational Aspirations in 12 Countries: The Importance of Institutional Context.” *Sociology of Education* **75**(2):99–122. https://doi.org/10.2307/3090287.

Carrell, Scott E., Richard L. Fullerton and James E. West 2009. “Does your Cohort Matter? Measuring Peer Effects in College Achievement.” *Journal of Labor Economics* **27**(3):439–64. https://doi.org/10.1086/600143.

Carrell, Scott E., Bruce I. Sacerdote and James E. West 2013. “From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation.” *Econometrica* **81**(3):855–82. https://doi.org/10.3982/ECTA10168.

Davies, Mark and Denise B. Kandel 1981. “Parental and Peer Influences on Adolescents’ Educational Plans: Some Further Evidence.” *American Journal of Sociology* **87**(2):363–87. https://doi.org/10.1086/227462.

De Giorgi, Giacomo, Michele Pellizzari Giacomo and Silvia Redaelli 2010. “Identification of Social Interactions through Partially Overlapping Peer Groups.” *American Economic Journal: Applied Economics* **2**(April):241–75. https://doi.org/10.1257/app.2.2.241.

Ding, Mingzhou, Yonghong Chen and Steven L. Bressler 2006. “Granger causality: Basic Theory and Application to Neuroscience”. In *Handbook of time series analysis: recent theoretical developments and applications*, edited by Björn Schelter, Matthias Winterhalder, and Jens Timmer, pp. 437–60. Weinheim, Germany: Wiley-VCH.

Duncan, Otis Dudley, Archibald O. Haller and Alejandro Portes 1968. “Peer Influences on Aspirations: A Reinterpretation.” *American Journal of Sociology* **74**(2):119–37. https://doi.org/10.1086/224615.

Felton, Chris and Brandon M. Stewart 2022. “Handle with Care: A Sociologist’s Guide to Causal Inference with Instrumental Variables.” Working Paper. Pp. 1–58. https://doi.org/10.31235/osf.io/3ua7q.

Flap, Henk and Beate Völker 2001. “Goal Specific Social Capital and Job Satisfaction: Effects of Different Types of Networks on Instrumental and Social Aspects of Work.” *Social Networks* **23**(4):297–320. https://doi.org/10.1016/S0378-8733(01)00044-2.

Flashman, Jennifer 2012. “Academic Achievement and its Impact on Friend Dynamics.” *Sociological Education* **85**(1):61–80. https://doi.org/10.1177/0038040711417014.

French, John R. and Bertram Raven 1959. “The Bases of Social Power”. In *Studies in Social Power*, edited by Cartwright, D. Ann Arbor, Mich: Institute for Social Research.

Friedkin, Noah E. 2006. *A Structural Theory of Social Influence Vol. 13*. Cambridge, UK: Cambridge University Press.

Friedkin, Noah E. and Eugene C. Johnsen 1990. “Social Influence and Opinions.” *Journal of Mathematical Sociology* **15**(3–4):193–206. https://doi.org/10.1080/0022250X.1990.9990069.

Haller, A.O. and C.E. Butterworth 1960. “Peer Influences on Levels of Occupational and Educational Aspiration.” *Social Forces* **38**(4):289–95. https://doi.org/10.2307/2573036.

Han, Li and Tao Li 2009. “The Gender Difference of Peer Influence in Higher Education.” *Economics of Education Review* **28**(1):129–34. https://doi.org/10.1016/j.econedurev.2007.12.002.

Hasan, Sharique and Surendrakumar Bagde 2013. “The Mechanics of Social Capital and Academic Performance in an Indian College.” *American Sociological Review* **78**(6):1009–32. https://doi.org/10.1177/0003122413505198.

Heider, Fritz 1958. *The Psychology of Interpersonal Relations*. New York: John Wiley.

Holland, Paul W. 2008. "Causation and Race". In *White Logic, White Methods: Racism and Methodology*, edited by Zuberi, T., Bonilla-Silva, E., pp. 93–109. Lanham, MD: Rowman & Littlefield Publishers, Inc.

Hout, Michael and William R. Morgan 1975. "Race and Sex Variations in the Causes of the Expected Attainments of High School Seniors." *American Journal of Sociology* **81**(2):364–94. https://doi.org/10.1086/226078.

Hubert, Lawrence and Phipps Arabie 1989. "Combinatorial Data Analysis: Confirmatory Comparisons between Sets of Matrices." *Applied Stochastic Models and Data Analysis* **5**(3):273–325. https://doi.org/10.1002/asm.3150050309.

Ifft, Jennifer and Margaret Jodlowski 2022. "Is ICE Freezing US Agriculture? Farm-Level Adjustment to Increased Local Immigration Enforcement." *Labour Economics* **78**(102203):1–28. https://doi.org/10.1016/j.labeco.2022.102203.

Krackhardt, David 1987. "QAP Partialling as a Test of Spuriousness." *Social Networks* **9**(2):171–86. https://doi.org/10.1016/0378-8733(87)90012-8.

Krackhardt, David 1992. "The Strength of Strong Ties: The Importance of Philos in Organizations". In *Networks and Organizations: Structure, Form, and Action*, edited by Nohria, N., Eccles, R.G., pp. 216–39. Boston: Harvard University Press.

Kretschmer, David, Lars Leszczensky and Sebastian Pink 2018. "Selection and Influence Processes in Academic Achievement—More Pronounced for Girls?" *Social Networks* **52**:251–60. https://doi.org/10.1016/j.socnet.2017.09.003.

Langdon, David, George McKittrick, David Beede, Beethika Khan and Mark Doms 2011. "STEM: Good Jobs Now and for the Future". *ESA Issue Brief #03-11*. Washington, DC: US Department of Commerce. http://files.eric.ed.gov/fulltext/ED522129.pdf.

Lazer, David, Brian Rubineau, Carol Chetkovich, Nancy Katz and Michael Neblo 2010. "The Coevolution of Networks and Political Attitudes." *Political Communication* **27**(3):248–74. https://doi.org/10.1080/10584609.2010.500187.

Leenders, Roger T.A. 2002. "Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix." *Social Networks* **24**(1):21–47. https://doi.org/10.1016/S0378-8733(01)00049-1.

Litzler, Elizabeth and Jacob Young 2012. "Understanding the Risk of Attrition in Undergraduate Engineering: Results from the Project to Assess Climate in Engineering." *Journal of Engineering Education* **101**(2):319–45. https://doi.org/10.1002/j.2168-9830.2012.tb00052.x.

Loh, Chung-Ping A. and Qiang Li 2013. "Peer Effects in Adolescent Bodyweight: Evidence from Rural China." *Social Science & Medicine* **86**:35–44. https://doi.org/10.1016/j.socscimed.2013.02.042.

Lomi, Alessandro, Tom A.B. Snijders, Christian E.G. Steglich and Vanina Jasmine Torlo 2011. "Why Are some More Peer than Others? Evidence from a Longitudinal Study of Social Networks and Individual Academic Performance." *Social Science Research* **40**(6):1506–20. https://doi.org/10.1016/j.ssresearch.2011.06.010.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* **60**(3):531–42. https://doi.org/10.2307/2298123.

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* **52**(1): 122–31. https://doi.org/10.2307/2095397.

Marsden, Peter V. 1990. "Network Data and Measurement." *Annual Review of Sociology* **16**:435–63. https://doi.org/10.1146/annurev.so.16.080190.002251.

Marsden, Peter V. and Noah E. Friedkin 1993. "Network Studies of Social-Influence." *Sociological Methods & Research* **22**(1):127–51.

Min, Shi, Zhouhang Yuan, Xiaobing Wang and Lingling Hou 2019. "Do Peer Effects Influence the Academic Performance of Rural Students at Private Migrant Schools in China?" *China Economic Review* **54**:418–33. https://doi.org/10.1016/j.chieco.2019.02.004.

Mouganie, Pierre and Yaojing Wang 2020. "High-Performing Peers and Female STEM Choices in School." *Journal of Labor Economics* **38**(3):805–41. https://doi.org/10.1086/706052.

Mouw, Ted 2006. "Estimating the Causal Effect of Social Capital: A Review of Recent Research." *Annual Review of Sociology* **32**:79–102. https://doi.org/10.1146/annurev.soc.32.061604.123150.

National Academies of Sciences, Engineering, and Medicine 2020. *Promising Practices for Addressing the Underrepresentation of Women in Science, Engineering, and Medicine: Opening Doors*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25585.

Newcomb, Theodore Mead 1943. *Personality & Social Change; Attitude Formation in a Student Community*. New York: Dryden Press.

Ost, Ben 2010. “The Role of Peers and Grades in Determining Major Persistence in the Sciences.” *Economics of Education Review* **29**(6):923–34. https://doi.org/10.1016/j.econedurev.2010.06.011.

Pescosolido, Bernice A. 1992. “Beyond rational choice: The social dynamics of how people seek help.” *American Journal of Sociology* **97**(4):1096–1138. https://doi.org/10.1086/229863.

Poldin, Oleg, Diliara Valeeva and Maria Yudkevich 2015. “Choice of Specialization: Do Peers Matter?” *Applied Economics* **47**(44):4728–40. https://doi.org/10.1080/00036846.2015.1034840.

Poldin, Oleg, Diliara Valeeva and Maria Yudkevich 2016. “Which Peers Matter: How Social Ties Affect Peer-Group Effects.” *Research in Higher Education* **57**:448–68. https://doi.org/10.1007/s11162-015-9391-x.

Raabe, Isabel J., Zsófia Boda and Christoph Stadtfeld 2019. “The Social Pipeline: How Friend Influence and Peer Exposure Widen the STEM Gender Gap.” *Sociology of Education* **92**(2):105–23. https://doi.org/10.1177/0038040718824095.

Rubineau, Brian 2007. *Gendering Professions: An Analysis of Peer Effects*.” PhD Thesis. Massachusetts Institute of Technology, Cambridge, MA. http://hdl.handle.net/1721.1/40854.

Sacerdote, Bruce 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *The Quarterly Journal of Economics* **116**(2):681–704. https://doi.org/10.1162/00335530151144131.

Sacerdote, Bruce 2011. “Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?” Chapter 4 In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen J. Machin, and Ludger Woessmann pp. 249–77. Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/B978-0-444-53429-3.00004-1.

Sacerdote, Bruce 2014. “Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?” *Annual Review of Economics* **6**(1):253–72. https://doi.org/10.1146/annurev-economics-071813-104217.

Sassler, Sharon, Jennifer Glass, Yael Levitte and Katherine M. Michelmore 2017. “The Missing Women in STEM? Assessing Gender Differentials in the Factors Associated with Transition to First Jobs.” *Social Science Research* **63**:192–208. https://doi.org/10.1016/j.ssresearch.2016.09.014.

Shalizi, Cosma Rohilla and Andrew C. Thomas 2011. “Homophily and Contagion Are Generically Confounded in Observational Social Network Studies.” *Sociological Methods & Research* **40**(2):211–39. https://doi.org/10.1177/0049124111404820.

Small, Mario Luis 2009. *Unanticipated Gains*. Chicago: University of Chicago Press.

Small, Mario Luis 2013. “Weak Ties and the Core Discussion Network: Why People Regularly Discuss Important Matters with Unimportant Alters.” *Social Networks* **35**:470–83. https://doi.org/10.1016/j.socnet.2013.05.004.

Small, Mario Luis, Vontrese Deeds Pamphile and Peter McMahan 2015. “How Stable Is the Core Discussion Network?” *Social Networks* **40**:90–102. https://doi.org/10.1016/j.socnet.2014.09.001.

Spenner, Kenneth I. and David L. Featherman 1978. “Achievement Ambitions.” *Annual Review of Sociology* **4**:373–420. https://doi.org/10.1146/annurev.so.04.080178.002105.

Steglich, Christian, Tom A.B. Snijders and Michael Pearson 2010. “Dynamic Networks and Behavior: Separating Selection from Influence.” *Sociological Methodology* **40**(1):329–93. https://doi.org/10.1111/j.1467-9531.2010.01225.x.

Torlò, Vanina Jasmine and Alessandro Lomi 2017. “The Network Dynamics of Status: Assimilation and Selection.” *Social Forces* **96**(1):389–422. https://doi.org/10.1093/sf/sox040.

Trogdon, Justin G., James Nonnemaker and Joanne Pais 2008. “Peer Effects in Adolescent Overweight.” *Journal of Health Economics* **27**(5):1388–99. https://doi.org/10.1016/j.jhealeco.2008.05.003.

VanderWeele, Tyler J. and Weihua An 2013. “Social Networks and Causal Inference”. In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, pp. 353–74: Springer: Dordrecht, The Netherlands. https://link.springer.com/chapter/10.1007/978-94-007-6094-3_17.

Wang, Tianheng 2021. “Classroom Composition and Student Academic Achievement: The Impact of Peers’ Parental Education.” *The BE Journal of Economic Analysis & Policy* **21**(1):273–305. https://doi.org/10.1515/bejeap-2020-0109.

Winkler, Irene, Danny Panknin, Daniel Bartz, Klaus-Robert Müller and Stefan Haufe 2016. “Validity of Time Reversal for Testing Granger Causality.” *IEEE Transactions on Signal Processing* **64**(11):2746–60. https://doi.org/10.1109/TSP.2016.2531628.

Woelfel, Joseph and Archibald O. Haller 1971. “Significant Others, the Self-Reflexive Act and the Attitude Formation Process.” *American Sociological Review* **36**(6):74–87. https://doi.org/10.2307/2093508.

Zimmerman, David J. 2003. “Peer Effects in Academic Outcomes: Evidence from a Natural Experiment.” *Review of Economics and Statistics* **85**(1):9–23. https://doi.org/10.1162/003465303762687677.

Zölitz, Ulf and Jan Feld 2021. “The Effect of Peer Gender on Major Choice in Business School.” *Management Science* **67**(11):6963–79. https://doi.org/10.1287/mnsc.2020.3860.