



HAL
open science

Disentangling peer effects in transportation mode choice: The example of active commuting

Mathieu Lambotte, Sandrine Mathy, Anna Risch, Carole Treibich

► To cite this version:

Mathieu Lambotte, Sandrine Mathy, Anna Risch, Carole Treibich. Disentangling peer effects in transportation mode choice: The example of active commuting. *Journal of Environmental Economics and Management*, 2023, 121, pp.102868. 10.1016/j.jeem.2023.102868 . hal-04194873

HAL Id: hal-04194873

<https://hal.science/hal-04194873v1>

Submitted on 12 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Journal Pre-proof

Disentangling peer effects in transportation mode choice: The example of active commuting

Mathieu Lambotte, Sandrine Mathy, Anna Risch, Carole Treibich



PII: S0095-0696(23)00086-4
DOI: <https://doi.org/10.1016/j.jeem.2023.102868>
Reference: YJEEM 102868

To appear in: *Journal of Environmental Economics and Management*

Received date: 13 July 2022

Please cite this article as: M. Lambotte, S. Mathy, A. Risch et al., Disentangling peer effects in transportation mode choice: The example of active commuting. *Journal of Environmental Economics and Management* (2023), doi: <https://doi.org/10.1016/j.jeem.2023.102868>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 Published by Elsevier Inc.

Disentangling peer effects in transportation mode choice: the example of active commuting

Mathieu Lambotte^{1,*}, Sandrine Mathy¹, Anna Risch¹, and Carole Treibich¹

¹Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, 38000 Grenoble, France

*Corresponding author: Mathieu Lambotte, mathieu.lambotte@univ-grenobles-alpes.fr, 04 76 74 29 27, bureau 201b, BATEG, 1241 rue des Résidences, 38400 Saint-Martin-d'Hères

July 26, 2023

Abstract

We investigate the role of peer effects in the workplace on individual active transportation mode choices. We collect original data through an online survey on networks and sustainable behaviors among 334 individuals working in ten research laboratories at the University of Grenoble Alps in February 2020. We apply linear and nonlinear models of peer effects on active modal choice, untangling the role of conformism and strategic complementarity in social influence. We show that given our data, a linear local-average specification is the preferred empirical model of peer effects, and we estimate strong and significant endogenous peer effects.

Keywords— peer effects, social network, workplace, transportation choice, conformism

JEL Classification— Q53, D91, R41, C31, C35

Acknowledgments

We would like to thank Paolo Crosetto, Habiba Djebbari, Aristide Houndetoungan and Yann Bramoullé for their comments at different stages of this article. We are also deeply grateful for the detailed and inspiring comments of two anonymous reviewers. We also thank the participants of the International Workshop of Spatial Econometrics and Statistics, of the JMA, the Congrès of the AFSE and the Journées Louis-André Gérard-Varet. This work was partially funded by IRS IDEX UGA RESMOB and by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) through the projects CDP IDEX UGA MOBILAIR and CDTool IDEX UGA ACME. All authors declare that they have no conflicts of interest.

Disentangling peer effects in transportation mode choice: the example of active commuting

Abstract

We investigate the role of peer effects in the workplace on individual active transportation mode choices. We collect original data through an online survey on networks and sustainable behaviors among 334 individuals working in ten research laboratories at the University of Grenoble Alps in February 2020. We apply linear and nonlinear models of peer effects on active modal choice, untangling the role of conformism and strategic complementarity in social influence. We show that given our data, a linear local-average specification is the preferred empirical model of peer effects, and we estimate strong and significant endogenous peer effects.

Keywords— peer effects, social network, workplace, transportation choice, conformism

JEL Classification— Q53, D91, R41, C31, C35

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 Introduction

1
2 Reducing the use of private cars is a public policy priority due to both health and environmen-
3 tal concerns. Indeed, road traffic is one of the main contributors to greenhouse gas emissions,
4 reinforcing climate change. Emissions of atmospheric pollutants by cars also have serious health
5 consequences in terms of mortality and morbidity. Last, the important use of private cars also
6 leads to noise pollution and congestion issues as well as sedentary lifestyles, which have become
7 an important public health issue, contributing to obesity and a lack of physical activity. Thus,
8 inducing modal shifts to active transportation modes (i.e., walk or bicycle) would simultane-
9 ously reduce the negative environmental and economic externalities of road traffic (pollution,
10 road wear and accidents (Bouscasse et al., 2022)) and the prevalence of several diseases linked to
11 inactivity (cardiovascular diseases, type-2 diabetes, and colo-rectal and breast cancers (Rojas-
12 Rueda et al., 2013)). Furthermore, due to both the high investment costs necessary to deploy
13 additional public transportation services and the important health benefits induced by the use
14 of active modes, it seems more interesting from a cost-benefit perspective to favor shifts from
15 cars to active modes rather than to public transportation (Bouscasse et al., 2022). In addition,
16 individuals generally highly value the private gains from health improvements when shifting
17 from car use to active modes (Bouscasse et al., 2023). Moreover, when shifting from car use
18 to public transport, members of the public are more reluctant to change their habits and to
19 experience a loss of utility for pure environmental motivations. Indeed, while several surveys
20 show that most people are aware of the benefits of reducing car use for environmental or health
21 reasons (IPSOS and BCG, 2017), in many cases, intentions do not translate into actual behav-
22 ioral changes (Sheeran and Webb, 2016). In 2017, 74% of French people who commute to work
23 used their car, 16% took public transportation and only 8% used an active mode (Brutel and
24 Pages, 2021). Even for distances of less than five kilometers, which might easily be traveled by
25 walk or bicycle, car travel still accounts for 60% of home-work commuting trips.
26
27

28
29 Public authorities are actively attempting to reduce the use of cars by introducing various
30 measures to encourage the uptake of alternative modes. For example, in France since 2010,
31 employers must reimburse 50% of annual public transportation passes. Since 2021, a so-called
32 *sustainable mobility package* (“forfait mobilité durable”) enables companies to voluntarily and
33 monetarily compensate their employees for traveling to work by means other than cars. These
34 two schemes are cumulative and exempt of taxes up to 500 euros per year per employee. Despite
35 the existence of these legal schemes and the development of bicycle and public transportation in-
36 frastructures, resulting changes in transportation behavior are rare. Indeed, behavioral changes
37 in transportation modes are particularly difficult to initiate. Like most other pro-environmental
38 behaviors, transportation mode choice depends on several intertwined economic, social and psy-
39 chological determinants and may be constrained by the institutions and infrastructure in which
40 these choices are embedded (De Witte et al., 2013; Sherwin et al., 2014). Such determinants of
41 modal choice include individual characteristics (age, gender, physical capabilities and cycling
42 experience, for instance) as well as specificities of the trip (speed, cost, infrastructure, slope,
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 safety, availability of alternatives, congestion, etc). While accounting for these classical deter-
2 minants of modal choice, in this paper, we propose to estimate to what extent peer behavior
3 influences individual transportation choice.
4

5
6 Indeed, the literature on transportation mode choices has recently begun a shift from an in-
7 dividual focus to a social focus, and behavioral models have started to account for the social
8 context in which choices are made (Maness et al., 2015). A study in Canada has shown for ex-
9 ample that individuals widely shift from car to bicycle when commuting with a bicycle becomes
10 the social norm (Kormos et al., 2015). Similarly, a case study in England provides qualitative
11 evidence that most of the interviewees who were new bicyclists had been influenced by their
12 peers to change their transportation mode (Sherwin et al., 2014). However, to the best of our
13 knowledge, the quantitative literature on social influence on transportation mode choice has
14 only considered arbitrarily defined sources of social influence. The most common proxy for
15 explicit peer networks is spatial neighborhoods, which assumes that all individuals living in the
16 same area are sources of social influence (Dugundji and Gulyás, 2008; Goetzke and Andrade,
17 2010; Goetzke and Rave, 2011; Kim et al., 2017; Phithakkitnukoon et al., 2017; Pike and Lubell,
18 2016, 2018; Walker et al., 2005; Wang et al., 2015). Another approximation of peer networks are
19 implicit social networks, where all individuals sharing a specific characteristics are assumed to
20 be connected, as in a study of transportation behavior on US army bases by Morrison and Lawell
21 (2016), in which all employees who recently immigrated to the US or who were born in Latin
22 America are considered peers, although they might have never met. The fact that the literature
23 on social influence on modal choice has focused on analyses of spatial neighborhoods or implicit
24 social network is an important limitation, as social influence is more likely to operate and to
25 be identified through peer networks and explicit social relationships (Pike, 2014). Moreover,
26 studying the role of peer networks on individual modal choices broadens the scope of analy-
27 sis to the large body of literature on network econometrics and causal estimation of so-called
28 *peer effects*, i.e., accounting for the fact that individual choices are embedded in social networks.
29
30

31
32
33
34
35
36
37
38
39
40 Indeed, we feel that the recent developments in network econometrics are underused in the
41 modeling strategies of social influence in pro-environmental behaviors in general and in trans-
42 portation modal choice in particular (Wolske et al., 2020). Social environments play a significant
43 role in explaining individual behavior and in modeling individual preferences (Clark and Os-
44 wald, 1998). We are especially interested here in the estimation of *endogenous peer effects*, the
45 influence of one's peers' behavior on one's own behavior. Peer effects can either be positive
46 or negative, revealing a complementarity or a substitutability, respectively, between individual
47 choices in a social setting. Negative endogenous peer effects relate, for example, to the infamous
48 free rider problem in public good provision, where one's private contribution might be nega-
49 tively correlated with those of one's peers. Positive endogenous peer effects are more diverse
50 and entail two main social phenomena depending on the type of pro-environmental behaviors
51 considered.
52

53 First, *conformism to the social norm* prevailing in one's social group is an important source
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

of positive endogenous peer effects, as individuals may account for their peers' choices "to develop and preserve meaningful social relationships, and to maintain a favorable self-concept" (Cialdini and Goldstein, 2004, p.591). Individuals may conform to either or both descriptive and injunctive norms, with the former referring to the individual expectation of peers' average behavior and the latter concerning how one's peers might expect one to behave. As information on injunctive norms is difficult to measure and collect, their impact is mostly studied in experimental economics (see, for example, the work of De Groot and Schuitema (2012) and Smith et al. (2012)). However, injunctive and descriptive norms are found to be strongly correlated in pro-environmental behaviors and both can be approximated in empirical studies by the observed behavior of peers (Farrow et al., 2017). In the context of transportation modal choice, a desire for conformity might be observed if individuals select their transportation modes to match the norm in their groups of peers.

Second, another source of positive endogenous peer effects is *strategic complementarity*, whereby the utility an individual derives from a given behavior is enhanced if her peers also choose the same behavior. Strategic complementarity is particularly relevant to transportation modal choice or physical activities (Liu et al., 2014), as commuting to work with a given mode or practicing a given activity might yield a higher utility when these actions are performed with peers.

These endogenous peer effects are crucial to analyze because they can promote the dissemination of "good practices" through a change in a social norm or because they generate strategic complementarities between individual choices. Endogenous peer effects are thus able to amplify the impact of public policies beyond their direct effects, acting as a social multiplier of policies' impacts. More precisely, if endogenous peer effects are driven by conformism, targeting the peers of individuals with few social connections might be more effective, as the social norm will be easier to change in relatively small social groups. On the other hand, targeting the peers with the highest number of social connections – through which policies' effects could spread – would be more effective if endogenous peer effects are driven by strategic complementarity.

Social influence can also stem from peers' characteristics rather than their behavior, which is generally referred to as *exogenous peer effects*. Indeed, specific peers' characteristics may influence individual behavior, such as peers' income, position on climate change, political orientation or age (Manski, 1993). For example, an individual sensitive to exogenous peer effects could refrain from using her car if most of her peers are environmentalists, regardless of their own transportation modes.

In this paper, we study peer effects on transportation modal choices in the workplace using explicit data on colleagues network. Indeed, although most people spend a large part of their daily lives at work, the diffusion of sustainable behaviors in this key arena of social interaction has not been extensively studied. As social links with coworkers may not be perfect substitutes for links with family or neighbors (Videras et al., 2011), one might also want to explore the impacts of coworkers on individual behavior. To the best of our knowledge, studies of peer effects in the workplace have almost exclusively sought to evaluate the impact of colleagues' productivity

1 on individual productivity (Beugnot et al., 2019; Falk and Ichino, 2006; Lindquist et al., 2022).
2 However, regular and friendly exchanges with colleagues during coffee breaks and lunches taken
3 together may lead to the adoption of new practices in non-work-related areas. Private and pub-
4 lic firms thus appear to be potentially favorable places in which to focus public policy efforts so
5 that more environmentally friendly mobility behaviors emerge and then spread to the private
6 sphere. Employers themselves can also favor the development of pro-environmental behaviors
7 in the workplace by informing employees that such behaviors are supported by the company
8 (Robertson and Barling, 2013; Wesselink et al., 2017).
9

10
11
12
13 The analysis of peer effects in the workplace on active modal choice proposed here is designed
14 to assess the extent to which interactions with colleagues can lead to the adoption of an active
15 transportation mode. To answer this question, we deployed a survey among ten research labo-
16 ratories (henceforth labs) on the main campus of the University of Grenoble-Alps in France and
17 constructed an original database gathering information on individual behaviors and their peer
18 networks in the workplace. To estimate causal peer effects, we adapt the generalized framework
19 of peer effects à la Liu et al. (2014) to a binary outcome, active transportation choice, following
20 recent developments in network econometrics (Boucher and Bramoullé, 2020). This generalized
21 model distinguishes between two sources of endogenous peer effects, conformism and strategic
22 complementarity, and disentangles endogenous and exogenous peer effects as well as correlated
23 effects due to lab specificities. Confirming the intuition in the economic literature on social
24 influence in pro-environmental behaviors, we show that conformism drives peer effects in active
25 transportation mode choice and we estimate that an individual whose peers all use an active
26 transportation mode has a 50 to 70 percentage point (pp) higher probability of using an ac-
27 tive mode than an individual with no peers using an active transportation mode. In addition
28 to various alternative specifications of the econometric models and robustness analyses of the
29 network data, we propose a nonlinear model of peer effects as an extension of our main model,
30 building on the classical papers in the network econometrics of binary outcomes by Brock and
31 Durlauf (2001) and Lee et al. (2014). Last, we extend our nonlinear framework to a multino-
32 mial outcome model, adapting Guerra and Mohnen (2022) to the transportation modal choice
33 among active modes, public transportation and cars. We believe that the empirical strategy
34 we develop in this paper demonstrates how peer effects can be estimated for a wide range of
35 pro-environmental behaviors.
36
37
38
39
40
41
42
43
44
45
46

47 The rest of the paper is organized as follows. Section 2 presents the study protocol, the data
48 collection process and descriptive statistics. Section 3 displays the empirical and identification
49 strategies of the generalized linear model as well as its microeconomic foundations. Section 4
50 reports the results of the estimation of peer effects, the extensions to nonlinear models as well
51 as alternative specifications and robustness analyses. Finally, section 5 discusses the results,
52 the methodological and empirical challenges of the study, its limitations and future research
53 opportunities.
54
55
56
57
58
59
60
61
62
63
64
65

2 Data collection and descriptive statistics

2.1 Census and social network survey

We seek to study the impacts of the work environment on transportation choice. More precisely, we wonder whether colleagues from the workplace influence behaviors that are not directly related to work but to personal choice. To carry out this study, we set up an original survey on the university campus of Grenoble Alps. The data collection process is described below.

One of the challenges we face is the construction of the networks per se. While Grenoble University hosts more than 90 research labs at two distinct geographical sites, we drastically restricted the number of research labs in which we wanted to deploy the survey to maximize the share of each lab's members participating in the study. Obtaining a high response rate per lab is a crucial objective of the data collection phase to be able to trace the social interactions within each research unit, and properly represent the peer network of each respondent. The higher the response rate, the better able we will be to minimize bias in the analysis.

Therefore, the selection of labs was based on two criteria: (i) the lab must be located on the main university campus and have only one geographical site and (ii) very large labs are excluded. This resulted in a selection of 14 labs. In a second step, the director of each selected research lab was contacted by email and/or phone to present him/her the study, confirm that the lab fulfilled the two inclusion criteria and ask whether he/she would agree to us coming to present the study in the lab to all the personnel. This second phase resulted in a final selection of ten labs varying in size from 29 to approximately 300 persons (including permanent and non-permanent staff and research and research support staff, excluding non-doctoral students). Ultimately, this represents 1,335 workers. At informational meeting about the study in each lab, flyers were distributed presenting the study¹. This visit was followed by an email¹ sent to each member of the selected research labs, recalling the purpose of the study and providing the consent form to be completed to receive the baseline questionnaire. The description of the purpose of the study was intentionally vague and referred only to social influence at work and not to transportation mode choice to avoid self-selection bias. The collection of consent forms started on January 7, 2020 and ended on January 27, 2020. Following the *General Data Protection Regulation* guidelines, only staff who provided their consent and thus agreed to participate in the survey received an email containing a link to the "baseline" questionnaire¹. This survey was conducted using the Sphinx software, and the questionnaire was therefore completed online by each respondent. This survey was intended to collect information on the following elements:

Network. We asked participants to name the coworkers with whom they have the most frequent and regular interactions. A list of colleagues appeared in the questionnaire. Only individuals who consented to participate to the study appear on this list.

Behavior. We asked respondents what was the main transportation mode that they used to commute to work in the last three months.

¹Available at <https://osf.io/nqat4/>.

Sociodemographic characteristics. We collected additional variables to account for individual heterogeneity and other determinants of modal choice.

The baseline questionnaire was sent on February 6, 2020, to the 465 individuals who agreed to participate in the study. A first reminder was sent on February 11, 2020, and a second was sent on February 18, 2020. In total, 407 individuals actually completed the baseline questionnaire. Of these 407 individuals, some did not report social network information or were isolated (i.e., did not have any peers among the other participants), yielding 334 exploitable observations for estimation. Table 1 reports the consent and participation rates for each research lab. The protocol was approved by the multidisciplinary ethics committee of the University of Grenoble Alps and complies with the *General Data Protection Regulation*. All documents related to the survey were presented in both French and English.

Table 1: Survey

Labs	1	2	3	4	5	6	7	8	9	10	Total
Nb of obs. in the lab †	110	77	300	227	100	121	77	182	29	112	1,335
Nb of obs. consenting to participate	54	38	41	105	33	46	38	44	13	53	465
Nb of obs. filling the questionnaire	51	28	34	94	30	39	34	40	12	45	407
Nb of obs. fully exploitable	47	20	22	86	27	28	22	39	10	33	334
Potential participation rate ‡	46%	36%	11%	41%	33%	32%	44%	22%	41%	40%	35%
Actual participation rate ¶	43%	26%	7%	38%	27%	23%	29%	21%	35%	30%	25%

Notes: † The number of individuals in the research labs is an estimation at the time of the survey. ‡ The potential participation rate corresponds to the share of individuals consenting to participate who actually fulfilled the questionnaire. ¶ The actual participation rate corresponds to the share of individuals who correctly reported network information and are not isolated, thus considered in the analysis.

Our database contains ten different professional networks composed of 10 to 86 nodes (individuals, Table 2). While individuals could cite up to six colleagues, the labs' average number of peers varies from 2.1 to 4, with a global average of 3.14 peers. We face a problem of mismeasurement of the networks, as only lab members who agreed to participate to the study appeared on the list of colleagues that could be cited (Table 1). Therefore, many respondents reported that they wished to cite a colleague who did not participate to the study.

Table 2: Network characteristics

Labs	1	2	3	4	5	6	7	8	9	10	Average
Nodes¶	47	20	22	86	27	28	22	39	10	33	33.4
Directed edges±	146	47	76	337	104	72	67	141	31	75	109.6
Average degree†	3.43	2.14	2.94	3.97	4.00	2.69	3.44	3.55	2.58	2.67	3.14
Share of nodes with missing links‡	82%	59%	88%	70%	97%	64%	74%	93%	75%	82%	78%

Notes: ¶ number of observations, ± number of directed links between observations, † average number of peers cited by an individual, ‡ people who wished to name someone who did not participate in the survey.

Given that some respondents did not correctly complete the sociodemographic entries or network information, the final data used for the empirical analysis encompass 334 individuals.

2.2 Descriptive statistics

In the survey, we collected data on the main transportation mode used by individuals to commute to work (Table 3). In our sample, 43% and 33% commute with an active mode or public transportation, respectively, while only 24% use a car. Although not representative of the distribution of modal shares in France, the data we collect are representative of the situation in the Grenoble urban area. Indeed, the modal shares from a representative sample² of 548 individuals in 2020 are similar: 35% for active modes, 23% for public transportation and 38% for car use. The differences between the two samples might be because we study a specific population (members of research labs) working at the same campus which is particularly well served by bicycle paths and public transportation relative to the rest of the Grenoble urban area.

We also collected sociodemographic characteristics: gender, age, number of dependent children (less than 12 years old) in the household, household income range and the urban neighborhood or rural city in which they live. Using the *OpenTripPlanner* tool (Morgan et al., 2019), we estimate the duration and distance of the commuting trips (between the lab in which they most often work and the centroid of their neighborhood at 8:30am) for each mode (car, walk, bicycle, public transportation). In addition to the distance from home to work (measured as the commuting distance by car), we created a variable measuring the relative opportunity cost (in time) of the fastest active mode (walking or bicycle) to commute versus the fastest non-active mode (public transportation or car) for a given trip. This variable, named *relative opportunity cost of active mode*, averages at 27.6%, which means that commuting with the fastest active mode (usually bicycle) is on average 27.6% longer than the fastest non-active mode (usually car). This relatively low opportunity cost of active transportation can be explained by congestion and traffic jams on roads and by the fact that some bike paths are less constrained by traffic lights. It is thus possible to commute faster with the fastest active mode than with the fastest non-active mode, which implies a negative relative opportunity cost of using an active mode. We also gather information on health issues that may prevent the use of an active mode (“Do you suffer from a health issue preventing you from using a bicycle or walking?”). We compute an index of three questions about recycling and waste-sorting behaviors to proxy for pro-environmental behaviors.³ There are 41% of women in the sample. The average age is 40 years; 63% of the sample holds a PhD. The individuals in our sample have an average of 0.52 dependent children and have worked in their research lab for 9.6 years (see Table 3). Note that there are no significant differences between the share of individuals that have a driver’s license or own a car between the individuals that use an active transportation mode and the others.

² *Enquête Mobilité Certifié Cerema (EMC2)* (obtained from the French Ministry of Ecological Transition and Regional Cohesion: <https://www.cerema.fr/en/cerema>)

³ The three questions are the following: “How often do you recycle items that you do not need anymore?”, “How often do you sort your household waste?”, “How often do you purchase items with low packaging?”. Answers were provided on the following four-level scale: Never/ Rarely/ Often/ Always.

Table 3: Descriptive statistics

	Obs.	Mean	Std. Dev.	Min	Max
Main commuting mode in the past 3 months					
Car	334	0.24			
Active mode †	334	0.43			
Public transportation	334	0.33			
Sociodemographic characteristics					
Male	334	0.59	0.49	0	1
Age in years	334	39.8	11.3	23	75
Years working in the research lab	334	9.5	9.2	0	50
Number of dependent children	334	0.52	0.88	0	3
Holds a PhD	334	0.63	0.48	0	1
Health issue	334	0.05	0.21	0	1
Pro-environmental behaviors ³	334	0.78	0.17	0	1
Household monthly income < 2 000 €	334	0.22			
Household monthly income 2 000 - 4 000 €	334	0.32			
Household monthly income 4 000 - 6 000 €	334	0.32			
Household monthly income > 6 000 €	334	0.14			
Commuting trips' characteristics					
Commuting distance via car (km)	334	14.9	22.6	0.9	233
Commuting time via car (min)	334	26.2	15.3	5.2	141
Relative opportunity cost of active mode	334	27.6%	92.2%	-56%	429%

Notes: † the share of people whose transportation mode is bicycle or walk.

3 Empirical strategy

To measure the influence of peer effects on active transportation choice, we adopt a sequential strategy in which we estimate different specifications of linear models of peer effects. We first propose a generalized model of peer effects in active transportation choice, where we distinguish the two main sources of endogenous peer effects, strategic complementarity and conformism. We then test for simpler specifications, which include only one of the two mechanisms, yielding the local-average model with conformism and the local-aggregate model with strategic complementarity. All three models are able to disentangle endogenous peer effects, exogenous peer effects and correlated effects and are identified when we use instrumental variables for peers' modal choices (Bramoullé et al., 2009; Liu et al., 2014). We also account for network endogeneity by estimating in a first step an unobserved individual variable that captures the unexplained individual heterogeneity in the formation of the network (Houndetoungan, 2020; Hsieh et al., 2020; Johnsson and Moon, 2021). We assume that active transportation choice depends on individual spatial variables (distance from home to work), sociodemographic characteristics of the individuals (age, gender, number of young children, highest diploma, household's monthly income), the transportation mode of their peers (endogenous peer effects), peers' characteristics (exogenous peer effects), the common environment faced by individuals from a given lab (correlated effects) and gregariousness (which accounts for network endogeneity). Furthermore, the generalized linear model of peer effects and the two local restrictions can be microfounded

on individual utility maximization as shown below.

3.1 Microfoundations

As explained in the introduction, endogenous peer effects can be distinguished with respect to two mechanisms: a desire for conformity to the norm prevailing in one's peer group and strategic complementarity, where an individual's utility depends on the sum of peers' *behavior or effort*, not the norm. Following the literature on network econometrics (Akerlof, 1997; Liu et al., 2014), we propose the following microfoundations for our econometric model of peer effects in transport modal choice, where individuals simultaneously choose y_i to maximize the following utility function in a network game:

$$u_i(y_i) = \underbrace{\left(\pi_i^* + \psi_1 \sum_{j \neq i} a_{ij} y_j \right)}_{\text{payoff}} y_i - \underbrace{\frac{1}{2} \left[y_i^2 + \psi_2 \left(y_i - \sum_{j \neq i} g_{ij} y_j \right)^2 \right]}_{\text{cost}}, \quad i, j = 1, \dots, n. \quad (1)$$

where a_{ij} is a cell of the adjacency matrix A that represents the network and $a_{ij} = 1$ if individuals i and j are peers, 0 otherwise. Denoting by n the number of nodes in the network, i.e., the sample size, A is thus an $n \times n$ matrix and $a_{ii} = 0$, i.e., self-influence is not allowed in the model. Moreover, a_{ij} may be different from a_{ji} , as the network is directed: individual i may consider individual j a peer, whereas j may not. The sum of each i^{th} row in A gives the number of peers of individual i , i.e., its out-degree. We obtain the social interaction matrix G by normalizing each row of A so that its sum equals unity: $\sum_{j \neq i} g_{ij} = 1$. $\sum_{j \neq i} g_{ij} y_j$ is thus the share of individual i 's peers who use an active transportation mode, i.e., a measure of the social norm in individual i 's social group.

The utility gain for individual i is modeled as a linear function of the individual modal choice y_i , with π_i^* capturing the direct benefit from choice y_i and $\sum_{j \neq i} a_{ij} y_j$, a social benefit emerging from the complementarity of individual i 's choice with the mobility behavior of her peers. Individuals also suffer costs from their modal choice, modeled as a quadratic function of y_i (private cost of choice y_i) and the deviation from the average choice of the peers, $(y_i - \sum_{j \neq i} g_{ij} y_j)$. The average choice of one's peers, $\sum_{j \neq i} g_{ij} y_j$, is commonly used as a measure of the norm prevailing in individual i 's social group. Following Akerlof (1997), failing to conform to the social norm is costly, and the cost is increasing in the social distance between one's choice and the norm.

Social influence thus affects the utility function in two ways: as an increasing function of the number of peers choosing a similar transportation mode and as a decreasing function of the distance between one's choice and the social norm (i.e., the average choice among one's peers). The first-order condition for the maximization of the utility function (Equation 1) yields the following best-response function:

$$y_i = \pi_i + \beta_1 \sum_{j \neq i} a_{ij} y_j + \beta_2 \sum_{j \neq i} g_{ij} y_j, \quad i, j = 1, \dots, n. \quad (2)$$

with $\beta_1 = \frac{\psi_1}{1+\psi_2}$, $\beta_2 = \frac{\psi_2}{1+\psi_2}$ and $\pi_i = \frac{\pi_i^*}{1+\psi_2}$. The consistency between the utility function and the best-response function is ensured because the relative difference of β_1 and β_2 is proportional to the relative difference of ψ_1 and ψ_2 ($\frac{\beta_1}{\beta_2} = \frac{\psi_1}{\psi_2}$). β_1 thus captures the endogenous peer effects stemming from strategic complementarity, while β_2 estimates the endogenous peer effects due to the desire to conform.

Moreover, as Equation 2 is the first-order condition for the maximization of the utility function 1, the global endogenous peer effects are given by $\frac{\partial^2 u_i(y_i)}{\partial y_i \partial y_j} = \frac{\partial y_i}{\partial y_j} = \beta_1 \sum_{j \neq i} a_{ij} + \beta_2 \sum_{j \neq i} g_{ij}$. If $\beta_1 > 0$, we detect strategic complementarity, and we identify strategic substitution if $\beta_1 < 0$. If $\beta_2 > 0$, we uncover conformism in individual behaviors and observe deviance if $\beta_2 < 0$. Let $\beta_1 > 0$ and $\beta_2 > 0$, as previous empirical evidences on social influence on active transportation choice tend to illustrate. Then, β_1 represents the increase in the probability of choosing an active transportation mode when, ceteris paribus, an additional peer chooses an active transportation mode. β_2 captures the increase in the probability of choosing an active transportation mode when, ceteris paribus, the share of peers using an active transportation mode increases from 0% to 100%.

Finally, under the assumption that $\beta_1 > 0$ and $\beta_2 > 0$, the Nash equilibrium derived from by the best-response function 2 is both unique and interior if the condition $g^{max} \beta_1 + \beta_2 < 1$ is met, where g^{max} is the maximum number of peers observed in the network (Liu et al., 2014). This condition imposes a restriction on the intensity of endogenous peer effects, ensuring the existence of a single and stable interior Nash equilibrium of the network game. The satisfaction of this condition can be assessed with the estimated parameters.

3.2 Econometric models of private utility

The best-response function (Equation 2) can be estimated by letting π_i , the individual payoff from choosing y_i , be expressed as a function of individual characteristics and unobserved heterogeneity (*preference shocks*):

$$\pi_i = \alpha + X_i \gamma + \epsilon_i, \quad i, j = 1, \dots, n. \quad (3)$$

where X_i is the i^{th} row of the $n \times k$ matrix of sociodemographic and spatial variables X , γ is a $k \times 1$ vector of parameters and ϵ_i is the i^{th} element of an $n \times 1$ vector of preference shocks ϵ .

Then, a simple and temporary expression for the econometric model stemming from the network game is:

$$y_i = \alpha + X_i \gamma + \beta_1 \sum_{j \neq i} a_{ij} y_j + \beta_2 \sum_{j \neq i} g_{ij} y_j + \epsilon_i, \quad i, j = 1, \dots, n. \quad (4)$$

Adapting the strategy of Boucher and Bramoullé (2020), we estimate Equation 4 and further specifications as a linear model. As y_i is a binary variable in our analysis, i.e., $y_i = 1$ if individual i commutes with an active transportation mode and 0 otherwise, Equation 4 is a linear probability model of active transportation mode choice with generalized social interactions.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Boucher and Bramoullé (2020) show that classical linear models of social interactions provide a sound alternative to more complex nonlinear models when the outcome of interest is a binary variable. The main advantages of such linear models of social interactions are the inclusion of fixed effects without yielding the incidental parameters problem, as nonlinear models would, the simplicity of their implementation and the fact that the estimated regression coefficients are directly interpretable as marginal effects on the probability of choosing an active transportation mode.

The simple econometric model of Equation 4 does not allow us to disentangle endogenous and exogenous peer effects. In that case, β_1 and β_2 estimate a mixture of endogenous peer effects and of the potential influence of peers' sociodemographic characteristics on one's behavior. To separate these two effects, we introduce the mean sociodemographic and spatial characteristics of peers, $\sum_{j \neq i} g_{ij} X_j$, in the individual payoff of choosing an active transportation mode, π_i (Equation 3):

$$\pi_i = \alpha + X_i \gamma + \sum_{j \neq i} g_{ij} X_j \theta + \epsilon_i, \quad i, j = 1, \dots, n. \quad (5)$$

where θ is a $k \times 1$ vector of parameters, which captures the exogenous (contextual) peer effects. A positive sign on θ_q indicates that peers' average of the q^{th} sociodemographic variable, household income for example, increases one's private payoff from choosing an active transportation mode. Indeed, if biking is well perceived among high-income peers, the income of one's peers would influence one's transportation mode.

In addition to endogenous and exogenous peer effects, Manski (1993) distinguishes correlated effects as another source of social influence. Correlated effects can be understood as the propensity of peers to behave similarly because they either share common unobserved individual characteristics or face a common environment. In the context of transportation modal choice in research labs, such variables might include unobserved variables at the lab level, as coworkers share a common environment and thus face the same set of incentives or external shocks (the presence or not of bicycle sheds, subsidies when using active transportation modes, ...). This information is unobserved by the econometrician, resulting in classical omitted variable bias for peer effects. Indeed, if these unobserved variables are not accounted for, some of their effects might be confounded with the peer effects we are interested in. Considering this common environment, we can expect that individuals in the same network will act in a similar way and therefore are likely to have positively correlated outcomes even in the absence of endogenous and exogenous peer effects.

We thus introduce lab fixed effects in the expression for π_i to control for potential correlated effects, i.e., unobserved variables at the lab level:

$$\pi_i = L_i \alpha^* + X_i \gamma + \sum_{j \neq i} g_{ij} X_j \theta + \epsilon_i, \quad i, j = 1, \dots, n. \quad (6)$$

where α^* is an $r \times 1$ vector of lab fixed effects for lab $l = 1, \dots, r$ and L is the i^{th} of an $n \times r$

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

selection matrix where $L_{il} = 1$ if individual i belongs to lab l . Each individual belongs to only one lab and can only have social interactions with other members of this same lab.

Last, correlated effects might also emerge from unobserved individual characteristics. Indeed, in most cases, we can assume that individuals select their peers and therefore share similar characteristics or interests, as homophily behaviors are dominant in peer selection and network formation (Graham, 2017). This specific issue is referred in the literature as *network endogeneity* as some omitted variables may influence both the likelihood of forming a link with a colleague (thus the social network) and the choice of transportation mode. More precisely, the unobserved characteristics are captured in the error term ϵ_i but are also correlated with the network.

Building on Houndetoungan (2020) and Graham (2017), we use a dyadic link formation model that estimates the probability that two individuals in a given lab are linked, i.e., are peers, based on their social distance and their respective gregariousness. In doing so, we estimate the unobserved individual gregariousness as the individual fixed effects in this model of link formation and we can include them as an additional sociodemographic variable, thus removing an important source of network endogeneity (Johnsson and Moon, 2021).

Let g_{ij}^* be a latent variable such that:

$$g_{ij}^* = \Delta X_{ij} \bar{\gamma} + \mu_i + \mu_j + \xi_{ij} \quad (7)$$

where $\Delta X_{ij} = |X_i - X_j|$, i.e., a $1 \times k$ vector of absolute difference of the sociodemographic space between individuals i and j , $\bar{\gamma}$ is a $k \times 1$ vector of parameters, μ_i and μ_j are individual effects that capture the degree heterogeneity, i.e., the gregariousness heterogeneity among individuals, and ξ_{ij} are the associated errors terms that follow logistic distribution (Houndetoungan, 2020). Assuming that $a_{ij} = 1$ if $g_{ij}^* > 0$, g_{ij}^* thus represents the utility obtained from link formation. We can write the probability P_{ij} of link formation between individuals i and j using the logistic distribution:

$$P_{ij} = \frac{\exp(\Delta X_{ij} \bar{\gamma} + \mu_i + \mu_j)}{1 + \exp(\Delta X_{ij} \bar{\gamma} + \mu_i + \mu_j)} \quad (8)$$

Note that $P_{ij} = P_{ji}$, and thus the link formation model is symmetric. However, the simulated links are still directed, as ξ_{ij} and ξ_{ji} may differ. Fixed effects for each network, i.e., lab, are included in the estimation. Let $\eta_i = (\xi_i, \mu_i)$ be bivariate and normally distributed with

$\Sigma_{\mu, \xi} = \begin{pmatrix} \sigma_\xi^2 & \rho \sigma_\xi \sigma_{\mu_i} \\ \rho \sigma_{\mu_i} \sigma_\xi & \sigma_{\mu_i}^2 \end{pmatrix}$ where ρ is the partial correlation between μ_i and ξ_i . The errors

ξ_i can be rewritten as $\xi_i = \rho \sigma_\xi \frac{\mu_i - m_{\mu_i}}{\sigma_{\mu_i}} + v_i$, where m_{μ_i} is the lab's average gregariousness and $v_i \sim \mathcal{N}(0, (1 - \rho^2) \sigma_\xi^2)$ and $Cov(\mu_i, v_i) = 0$. The fixed effects μ_i , which capture individual gregariousness heterogeneity, are estimated using 5000 iterations of a Markov chain Monte Carlo (MCMC) estimator. The estimate of μ_i with the highest posterior density is selected and the lab-wise standardized $\tilde{\mu}_i$ are computed as $\tilde{\mu}_i = \frac{\mu_i - m_{\mu_i}}{\sigma_{\mu_i}}$. The standardized individual fixed effects

$\tilde{\mu}_i$ are included in the previous model of peer effects as additional variables:

$$\pi_i = L_i\alpha^* + X_i^*\gamma + \sum_{j \neq i} g_{ij}X_j^*\theta + \epsilon_i, \quad i, j = 1, \dots, n. \quad (9)$$

where $X_i^* = [X_i, \tilde{\mu}_i]$ is the i^{th} row of an $n \times (k + 1)$ matrix of sociodemographic characteristics X^* while γ and θ are now $(k + 1) \times 1$ vectors of parameters. Then, if the $k + 1$ element of θ , θ_{k+1} , is different from 0, we can conclude that gregariousness influences individual active mode choice and confirm the presence of (now corrected) network endogeneity. Note that the standard errors are slightly underestimated, as this specification assumes that μ_{il} are observed instead of simulated via MCMC. Houndetoungan (2020) suggests to correct the variance via a Monte Carlo procedure, and the results presented here are obtained via 1000 simulations. Moreover, we infer the results in the next section assuming that using lab fixed effects and accounting for network endogeneity via the inclusion of the estimated gregariousness lead to exogenous errors terms, i.e., the correlated effects issue is solved and $E(\epsilon_i | L\alpha, X^*, G) = 0$. However, the errors might still be heteroskedastic and correlated between workers in the same lab; thus, the standard errors of the estimated coefficients are cluster-robust at the lab level (Pustejovsky and Tipton, 2018).

Hereinafter, we first detail the econometric estimation of the generalized linear model of peer effects. We then describe a specification test that allows us to test whether a generalized model of social interaction is informative or if a simpler specification, including social interactions only as strategic complementarity or as conformism, is preferable.

3.3 Generalized model of peer effects

The generalized econometric model of peer effects is then:

$$y_i = L_i\alpha^* + X_i^*\gamma + \beta_1 \sum_{j \neq i} a_{ij}y_j + \beta_2 \sum_{j \neq i} g_{ij}y_j + \sum_{j \neq i} g_{ij}X_j^*\theta + \epsilon_i, \quad i, j = 1, \dots, n \quad (10)$$

The generalized model thus accounts for two mechanisms of endogenous peer effects, strategic complementarity (β_1) and desire for conformity (β_2), for exogenous peer effects – captured by θ – for correlated effects – captured by the lab fixed effects α^* – and finally for network endogeneity through the estimation of the unobserved gregariousness (included in X^*). We can express the model given by Equation 10 using matrix notation such that:

$$Y = L\alpha^* + X^*\gamma + \beta_1 AY + \beta_2 GY + GX^*\theta + \epsilon \quad (11)$$

3.4 J test for the local-aggregate or local-average models

Setting $\beta_1 = 0$ or $\beta_2 = 0$ in Equation 10, we obtain the *local-average* or the *local-aggregate* econometric model, respectively. Liu et al. (2014) extended Kelejian (2008)'s *J test* of spatial

model specification to network models with fixed effects. Let M_1 and M_2 be the local-aggregate and local-average models, respectively, such that:

$$M_1 : Y = L\alpha^* + X^*\gamma + \beta_1 AY + GX^*\theta + \epsilon \quad (12)$$

$$M_2 : Y = L\alpha^* + X^*\gamma + \beta_2 GY + GX^*\theta + \epsilon \quad (13)$$

To test the null hypothesis that M_1 is the true specification of our model of peer effects of modal choice versus the alternative model of M_2 , we proceed in two steps. First, we estimate model M_2 to obtain the fitted values \hat{Y}_{M_2} . Second, we estimate an augmented version of model M_1 , including the fitted values \hat{Y}_{M_2} as an additional predictor:

$$Y = \omega_1 \hat{Y}_{M_2} + L\alpha^* + X^*\gamma + \beta_1 AY + GX^*\theta + \epsilon \quad (14)$$

This model is estimated using two-stage least squares (2SLS) with the instrumental variables $Z = [\alpha^*, X^*, GX^*, AX^*, G^2X^*]$. Then, the J test consists of the null hypothesis $H_0 : \omega_1 = 0$ and the alternative hypothesis $H_1 : \omega_1 \neq 0$; model M_1 is thus not a correct specification if $\omega_1 \neq 0$. We can also test the reciprocal case, where the null hypothesis is that M_2 is the correct specification and \hat{Y}_{M_1} are estimated in a first step then included in Equation 13. Specification M_2 is then not correct if $\omega_2 \neq 0$. If we reject H_0 in both tests, we can thus conclude that the generalized model is preferred to the local alternatives.

3.5 Identification via instrumental variables

Following Liu et al. (2014), the reduced form of the generalized model defined in Equation 11 can be expressed as:

$$Y = (1 - \beta_1 A - \beta_2 G)^{-1} (L\alpha^* + X^*\gamma + GX^*\theta + \epsilon) \quad (15)$$

Identification of peer effects has been shown to be problematic when the outcomes of peers, AY or GY , and the characteristics of peers, GX , are linearly dependent (Manski, 1993). In Manski's model, individuals interact in groups, such that every individual in a given group has ties with everyone else in the group. In that case, the group's average outcome is linearly dependent with the group's average sociodemographic variables. A similar argument holds for the group's aggregate outcome. Using a network structure with intransitive triads guarantees that individual i 's peers' average outcome is not perfectly collinear with individual i 's peer average sociodemographic variables, as peers' average outcome also depends on the sociodemographic variables of individuals that are not directly individual i 's peers. The network structure thus naturally provides exclusion restrictions allowing the identification of peer effects using instrumental variables for peers' outcomes (Bramoullé et al., 2009).

From Equation 15, the expectation of peers' aggregate or average outcomes can be written as:

$$E(AY) = (1 - \beta_1 A - \beta_2 G)^{-1}(AX^*\gamma + AGX^*\theta) + A(1 - \beta_1 A - \beta_2 G)^{-1}L\alpha^* \quad (16)$$

$$E(GY) = (1 - \beta_1 A - \beta_2 G)^{-1}(GX^*\gamma + G^2X^*\theta) + G(1 - \beta_1 A - \beta_2 G)^{-1}L\alpha^* \quad (17)$$

Using Equations 16 and 17, Liu et al. (2014) show that G^2X^* can be used as instruments of the generalized model of peer effects, as in the local-average model of Bramoullé et al. (2009). In the local-average model of peer effects, the average outcome of one's peers, GY , can be instrumented by the characteristics of one's peers. Bramoullé et al. (2009) show that conditional on the existence of *intransitive triads*, i.e., peers of peers that are not directly one's peers, one can estimate a local-average model via 2SLS and instruments $Z = [\alpha^*, X^*, GX^*, G^2X^*]$. The local-aggregate model is also estimated using 2SLS with $Z = [\alpha^*, X^*, GX^*, G^2X^*, AI]$, where I is the identity matrix of size n (Liu et al., 2014). The additional instrumental variable AI is the number of peers of each individual.

The generalized model of peer effects can similarly be estimated using 2SLS and instrumental variables consisting of $Z = [\alpha^*, X^*, AX, GX^*, G^2X^*, AI]$. Due to the linear structure of the models, they can all be estimated using a deviation from the lab's mean projector, similar to the within transformation in fixed effects estimators in panel data. For example, in the generalized model, the instruments expressed in deviation from the lab's mean are given by $Z = J[X^*, AX^*, GX^*, G^2X^*, AI]$, where J is the matrix that obtains the deviation from the lab's mean, $J = \text{diag}(J_l)$, $J_l = I_l - \frac{1}{n_l}1_l1_l'$, where I_l is an $n_l \times n_l$ identity matrix, n_l is the size of lab l and 1_l is a $n_l \times 1$ vector with all entries equal to 1.

Interestingly, the linear models of peer effects may be presented as systems of linear simultaneous equations, where instrumenting peers' outcomes also breaks down the simultaneity between the endogenous modal choices (Manski (1993)'s famous reflection problem). Moreover, instrumenting peers' modal choices GY by the instruments Z makes these linear models inherently robust to the assumption that peers' outcomes are unobservable by the individuals (but observed by the econometrician) in a given lab. Indeed, if peers' modal choices are unobserved, individuals form expectations on peers' choices based on observable information, i.e., their own characteristics, peers' characteristics and peers of peers' characteristics (Blume et al., 2011). In both cases, peers' observed or unobserved outcomes enter the econometric models through the same matrix of instruments and are thus identical.

4 Results

4.1 Estimation of peer effects – local-average, local-aggregate and generalized models

The main results concerning the role of peer effects in individual mode choices are presented in Table 4, grouping the local-average model (Equation 13), the local-aggregate model (Equation 12) and the generalized model (Equation 11). Given our data, the preferred empirical specification of the linear model of peer effects on transportation mode choice is the linear local-average model.

Our results show that while both local-average and local-aggregate endogenous peer effects are significant in the generalized model, the influence of peers' behavior in transportation mode choice is mostly captured by the local-average peer effect. Indeed, in the generalized model, the local-average peer effect is estimated at $\hat{\beta}_2 = 0.509$, which implies that the marginal effect of having an additional 33pp in the fraction of peers (one out of 3 peers, the average number of elicited peers in our sample) who uses an active transportation mode results in a 17pp increase in the own probability of choosing that transportation mode. Conversely, in the generalized model, the local-aggregate peer effect is only estimated at $\hat{\beta}_1 = 0.066$, which indicates that the estimated marginal effect of having an additional peer using an active transportation mode increases by 6.7pp the own probability of choosing an active mode. In our case, the estimated local-average peer effects in the generalized model are thus almost 10pp higher than the local-aggregate peer effects. This shows that social influence in transportation mode choice in the workplace is mostly driven by conformism.

As expected, the estimated endogenous peer effect stemming from conformism in the local-average model is larger ($\hat{\beta}_2 = 0.717$) than in the generalized model, as both conformism and strategic complementarity are confounded in the local specification. This indicates that the local-average model of peer effects tends to overestimate the strength of endogenous peer effects stemming from conformism when strategic complementarity is ignored. In addition, the J tests for both the local-average and local-aggregate models are significant, which indicates that the generalized model improves the fit. The generalized model of peer effects should thus be preferred to provide unbiased estimates of endogenous peer effects stemming from both conformism and strategic complementarity.

In addition, the three linear models have a satisfying share of predicted probabilities inside the interval $[0, 1]$. Indeed, 86%, 90%, 87% of the predicted probability of the local-average, the local-aggregate and the generalized models, respectively, are consistent with a probabilistic interpretation, bounded between 0 and 1. Thus, the linear modeling of peer effects appears to be adequate for a binary outcome such as active transportation choice. Furthermore, the local-average, local-aggregate, and generalized models satisfy the conditions for the existence and uniqueness of an interior Nash equilibrium. Specifically, the estimated parameters satisfy the conditions $0 \leq \hat{\beta}_2 < 1$, $0 \leq g^{max} \hat{\beta}_1 < 1$ and $g^{max} \hat{\beta}_1 + \hat{\beta}_2 < 1$, respectively (Liu et al., 2014).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

These conditions intuitively indicate that network games, like the one proposed here (Equation 1), have a unique interior equilibrium only when social interactions are at a moderate level (Brock and Durlauf, 2001; Lee et al., 2014).

However, given our sample size, the generalized model of peer effects is not our preferred empirical specification. Indeed, while all models pass the weak instrument test (the null hypothesis is that the instruments are weak and p -value < 0.05 in the three models), indicating that the instruments are not weak, i.e., they are strongly correlated with peers' active transportation mode choices, we reject the null hypothesis of the overidentification test (*Wooldridge's robust score test*, an extension of *Sargan test* for cluster-robust standard errors) in the generalized model (p -value = 0.027). The null hypothesis of this test is that the instruments are exogenous and uncorrelated with the error terms. Thus, the instruments are endogenous in the generalized model, and the identification strategy is not valid. Conversely, for the local-aggregate and local-average models, the null hypothesis of the overidentification tests cannot be rejected (p -value = 0.130 and p -value = 0.438, respectively), indicating that the instruments are exogenous and the adjacency (A) or social interaction (G) matrices can be considered exogenous conditional on the sociodemographic variables and lab fixed effects (Liu et al., 2014). Moreover, the null hypothesis of the Wu-Hausman test is rejected in the local-average model (p -value = 0.028), indicating that the 2SLS estimator is more consistent than simple OLS in that specification. Our preferred empirical specification is thus the local-average model, as it is the only specification that satisfies these three important tests for 2SLS (Boucher and Bramoullé, 2020). The interpretation of the results, robustness checks and nonlinear extensions are thus only based on the local-average model, although we do not exclude that the generalized model would be appropriate with a larger sample.

We thus estimate an endogenous peer effect of 72pp in the linear local-average model (Equation 13), which is similar to the endogenous peer effects found by Boucher and Bramoullé (2020) on smoking behavior or by Fletcher (2012) on students' drinking behavior, both using a similar linear local-average model of peer effects for binary outcomes. This result implies that a peer shifting to an active transportation mode, thus raising the share of peers using that mode by 33pp on average (individuals cited 3 peers on average in our sample), increases the individual probability of using an active mode by 24pp.

Table 4: Estimation of peer effects on active transportation choice

	Local-average model		Local-aggregate model		Generalized model	
	Individual effects	Contextual effects	Individual effects	Contextual effects	Individual effects	Contextual effects
Local-aggregate peer effect (β_1)	-	-	0.071 (0.029)	-	0.066* (0.023)	-
Local-average peer effect (β_2)	0.717* (0.262)	-	-	-	0.509* (0.159)	-
Gregariousness	0.042 (0.037)	-0.016 (0.065)	0.063 (0.035)	0.031 (0.064)	0.047 (0.038)	-0.016 (0.063)
Age	-0.005 (0.005)	0.015* (0.006)	-0.004 (0.004)	0.008* (0.004)	-0.004 (0.005)	0.014* (0.006)
Male	0.094 (0.060)	-0.011 (0.153)	0.107 (0.049)	0.119 (0.093)	0.097 (0.054)	-0.008 (0.123)
Years in lab	0.004 (0.005)	-0.012 (0.006)	0.002 (0.004)	-0.005 (0.007)	0.003 (0.005)	-0.011 (0.007)
Distance	0.001 (0.002)	-0.003 (0.006)	0.000 (0.002)	-0.005 (0.004)	0.000 (0.002)	-0.004 (0.006)
Relative opportunity cost of active mode	-0.002** (0.000)	0.001 (0.001)	-0.002** (0.000)	0.000 (0.001)	-0.002** (0.000)	0.001 (0.001)
Number of children under 12 years	0.025 (0.027)	0.033 (0.069)	0.028 (0.028)	0.081 (0.079)	0.024 (0.031)	0.033 (0.066)
PhD	0.024 (0.059)	-0.065 (0.144)	0.044 (0.050)	0.009 (0.111)	0.024 (0.066)	-0.066 (0.143)
Income lower than 2000€	0.038 (0.096)	0.103 (0.163)	0.065 (0.099)	0.185 (0.172)	0.040 (0.102)	0.070 (0.150)
Income between 4000€ and 6000€	0.017 (0.057)	-0.048 (0.098)	-0.002 (0.059)	-0.030 (0.058)	0.009 (0.056)	-0.059 (0.086)
Income higher than 6000€	0.211 (0.119)	-0.010 (0.130)	0.169 (0.091)	0.052 (0.120)	0.188 (0.116)	-0.009 (0.120)
Pro-environmental behavior	0.515* (0.181)	-0.098 (0.272)	0.564* (0.170)	0.135 (0.171)	0.532* (0.177)	-0.083 (0.240)
Health issue	-0.266 (0.116)	0.600* (0.186)	-0.159 (0.095)	0.512** (0.127)	-0.245 (0.109)	0.591* (0.176)
Lab fixed effects		Yes		Yes		Yes
Nb. Observations		334		334		334
R ²		0.548		0.304		0.223
Weak IV test p-value		0.040		0.000		0.040 & 0.000†
Wu-Hausman test p-value		0.028		0.632		0.031
Overidentification test p-value		0.438		0.130		0.027

Notes: † There are two p-values as the generalized model has two endogenous variables that are instrumented. The dependent variable is the choice of an active transportation mode and the parameters are estimated using a 2SLS estimator. The numbers between brackets are the cluster-robust standard errors.

J test p-value with the local-aggregate model as null: 0.019.

J test p-value with the local-average model as null: 0.032.

*p-value < 0.05; **p-value < 0.01; ***p-value < 0.001.

Most of the sociodemographic variables are not significant, with the exception of the index of pro-environmental behaviors and the relative opportunity cost (in time spent on the trip) of an active mode compared to other alternatives. As expected, individuals that have strong environmental awareness and for which using an active transportation mode is not too costly are more likely to use an active transportation mode.

Similarly, contextual peer effects are mostly nonsignificant, indicating that the main source of social influence is endogenous peer effects, especially conformism. However, having older peers or peers with health issues preventing the use of an active transportation mode has a positive and significant impact on an individual's probability of making an active transportation mode choice. Intuitively, one may expect that being confronted with peers with health concerns (either due to specific diseases or aging) could give an additional motivation to use an active transportation mode to reduce one's own exposure to future health issues.

4.2 Extension – nonlinear models of peer effects in transportation mode choice

Nonlinear model of active transportation choice. We assess the robustness of the local-average model of linear social interactions by estimating a related nonlinear model (Boucher and Bramoullé, 2020). We also estimate a multinomial logit model of peer effects on transportation modal choice, building on Guerra and Mohnen (2022). Indeed, individual modal choices are generally estimated with multinomial logit models where one may uncover different and even divergent peer effects on active mode, public transportation and car use.

Although we provide in the next sections a variety of alternative specifications and robustness checks to the linear local-average model, confirming that such models are applicable to binary outcomes such as active transportation choice, we present here an alternative nonlinear model. Indeed, in both transport economics (Pike and Lubell (2018) for example) and network econometrics (Lee et al., 2014), binary dependent variables such as transportation mode choice are mostly estimated using nonlinear models, i.e., logit models in the random utility maximization framework (McFadden, 1984). In comparison with linear models, logit models have the advantage of bounding the predicted probabilities of an active transportation modal choice in $[0, 1]$.

To derive the microfoundations of the nonlinear model of peer effects, recall that the linear local-average model is given by the best-response function:

$$y_i = L_i\alpha^* + X_i^*\gamma + \beta_2 \sum_{j \neq i} g_{ij}y_j + \sum_{j \neq i} g_{ij}X_j^*\theta + \epsilon_i, \quad i, j = 1, \dots, n \quad (18)$$

Boucher and Bramoullé (2020) shows that if ϵ_i are independent and continuously distributed, such a linear local-average model of peer effects is consistent with a random utility framework and the classical logit specification of individual modal choices.

Assuming that peers' outcomes are observed by individuals and that individual relative preference shocks follow a logistic distribution, we can easily derive such a logit model of peer effects:

$$P_i = P(y_i = 1) = \frac{\exp(L_i\alpha^* + X_i^*\gamma + \beta_2 \sum_{j \neq i} g_{ij}y_j)}{1 + \exp(L_i\alpha^* + X_i^*\gamma + \beta_2 \sum_{j \neq i} g_{ij}y_j)} \quad (19)$$

Note that due to the limited sample size in our analysis and to be as close as possible as to the "one in ten rule" (logit models are highly biased if the sample provides fewer than ten observations per independent variable per outcome category (Peduzzi et al., 1996)), we no longer account for the contextual effects in this extension⁴. Indeed, lab fixed effects capture most of the variation in peers' characteristics because the networks in our data are small. Furthermore, given the small networks in our data and the fact that transportation mode choices are highly salient in the workplace, we assume throughout this paper that individuals know their peers' transportation modes and thus maximize their utility in a situation with perfect information. This differs for the large networks studied by Brock and Durlauf (2001) or Lee et al. (2014), where they assume that individuals form expectations about peers' behavior and where the econometric models are solved using fixed-point interactions. By contrast, the binary logit model proposed here can be straightforwardly estimated by assuming that a single equilibrium exists in the sample, as is usually done in the literature (Bajari et al., 2010; Lin et al., 2021), although multiple equilibria could exist in the case of strong peer effects ($|\beta_2| > 2$) (Guerra and Mohnen, 2022; Lee et al., 2014).

For completeness, we also estimate Lee et al. (2014)'s nonlinear model of peer effects based on expected peers' outcomes. This model is obtained by replacing y_j , the observed peer's modal choice in Equation 19, with $P_j^e = P^e(y_j = 1)$, the expectation by individual i of the probability that individual j chooses an active transportation mode.

$$P_i = P(y_i = 1) = \frac{\exp(L_i\alpha^* + X_i^*\gamma + \beta_2 \sum_{j \neq i} g_{ij}P_j^e)}{1 + \exp(L_i\alpha^* + X_i^*\gamma + \beta_2 \sum_{j \neq i} g_{ij}P_j^e)} \quad (20)$$

where $P^e = (P_1^e, \dots, P_n^e)$ is the column vector of modal choice. The rational expectations equilibrium⁵ is the vector $P^* = (P_1^*, \dots, P_n^*)$ corresponding to Equation (2) in Lee et al. (2014):

$$P^* = \begin{pmatrix} P_1^* \\ \vdots \\ P_n^* \end{pmatrix} = \begin{pmatrix} \frac{\exp(L_1\alpha^* + X_1^*\gamma + \beta_2 \sum_{j \neq 1} g_{1j}P_j^*)}{1 + \exp(L_1\alpha^* + X_1^*\gamma + \beta_2 \sum_{j \neq 1} g_{1j}P_j^*)} \\ \vdots \\ \frac{\exp(L_n\alpha^* + X_n^*\gamma + \beta_2 \sum_{j \neq n} g_{nj}P_j^*)}{1 + \exp(L_n\alpha^* + X_n^*\gamma + \beta_2 \sum_{j \neq n} g_{nj}P_j^*)} \end{pmatrix} \quad (21)$$

⁴Although this specification does not allow us to disentangle endogenous and exogenous peer effects, the estimated coefficient in the linear local-average model without contextual effects (first column of Table 5) is very close to the endogenous peer effects in the complete linear local-average model (Table 4)

⁵In this context, the assumption of rational expectations implies that P^e does not systematically differ from P in equilibrium and can thus be consistently estimated by a fixed-point iteration

Simplifying Equation 21 yields:

$$P^* = F(L\alpha^*, X^*\gamma, \beta_2 GP^*) \quad (22)$$

where $F(\cdot)$ is defined as a logistic function, as the probability of selecting an active transportation mode is based on the logistic distribution. P^* is estimated iteratively as the fixed-point solution of Equation 22 until convergence.

Nonlinear model of multinomial transportation choices. Another extension of our analysis is to consider multinomial modal choices instead of binary choices, which is common practice in discrete choice analyses. Indeed, even if our primary interest lies in identifying peer effects for active transportation, evidence on the existence of peer effects in other transportation modes (public transportation and car) is also important for public policy and future research⁶. The extension of the binary logit model of peer effects (Equation 19) to a multinomial outcome with M alternatives is straightforward (Brock and Durlauf, 2003; Guerra and Mohnen, 2022):

$$P_{i,m} = P(y_i = m) = \frac{\exp(L_i\alpha_m^* + X_i^*\gamma_m + \beta_m \sum_{j \neq i} g_{ij} y_{j,m})}{\sum_{m'=1}^M \exp(L_i\alpha_{m'}^* + X_i^*\gamma_{m'} + \beta_{m'} \sum_{j \neq i} g_{ij} y_{j,m'})}, \quad m = 1, \dots, M. \quad (23)$$

where

$$y_{j,m} = \begin{cases} 1, & \text{if individual } j \text{ chooses the transportation mode } m \\ 0, & \text{otherwise} \end{cases}$$

As for the binary logit model, we assume in Equation 23 that individuals know their peers' modal choices and thus the multinomial logit model can be estimated straightforwardly.

Alternatively, we can extend the nonlinear model for unobserved binary outcomes of Equation 22 to a multinomial setting:

$$P_{i,m} = P(y_i = m) = \frac{\exp(L_i\alpha_m^* + X_i^*\gamma_k + \beta_m \sum_{j \neq i} g_{ij} P_{j,m}^e)}{\sum_{m'=1}^M \exp(L_i\alpha_{m'}^* + X_i^*\gamma_{m'} + \beta_{m'} \sum_{j \neq i} g_{ij} P_{j,m'}^e)}, \quad m = 1, \dots, M. \quad (24)$$

Letting $P_m^* = (P_{1,m}^*, \dots, P_{n,m}^*)$ be the column vector of the rational expectation equilibrium:

$$P_m^* = \begin{pmatrix} P_{1,m}^* \\ \vdots \\ P_{n,m}^* \end{pmatrix} = \begin{pmatrix} \frac{\exp(L_1\alpha_m^* + X_1^*\gamma_m + \beta_m \sum_{j \neq 1} g_{1j} P_{j,m}^*)}{\sum_{m'=1}^M \exp(L_1\alpha_{m'}^* + X_1^*\gamma_{m'} + \beta_{m'} \sum_{j \neq 1} g_{1j} P_{j,m'}^*)} \\ \vdots \\ \frac{\exp(L_n\alpha_m^* + X_n^*\gamma_m + \beta_m \sum_{j \neq n} g_{nj} P_{j,m}^*)}{\sum_{m'=1}^M \exp(L_n\alpha_{m'}^* + X_n^*\gamma_{m'} + \beta_{m'} \sum_{j \neq n} g_{nj} P_{j,m'}^*)} \end{pmatrix} \quad (25)$$

Collecting the left-hand side terms, we obtain an $n \times M$ matrix $P^* = (P_1^*, \dots, P_M^*)$ that we can estimate by iteratively solving for the fixed-point of $P^* = F(L\alpha^*, X^*\gamma, GP^*\beta^*)$, with $\beta^* = (\beta_1, \dots, \beta_M)$ being a $M \times 1$ vector of parameters.

⁶We are thankful to an anonymous reviewer for suggesting this extension.

Table 5: Estimation of peer effects on transportation mode choice using non-linear models

	Linear model	Binary logit (eq. 19)	Binary logit RE (eq. 21)	MN logit (eq. 23)	MN logit RE (eq. 25)
Peer effects on:					
Active transportation choice					
Point estimate	0.763**	1.156**	1.921*	1.021*	1.326
Standard errors	(0.185)	(0.411)	(0.834)	(0.457)	(0.985)
Marginal effects †	0.763**	0.201**	0.338*	0.177*	0.232
Public transportation choice					
Point estimate				0.347	1.632
Standard errors	–	–	–	(0.482)	(1.255)
Marginal effects †	–	–	–	0.064	0.299
Car choice					
Point estimate	–	–	–	0.150	-0.316
Standard errors	–	–	–	(0.517)	(1.348)
Marginal effects †	–	–	–	0.021	-0.044
Individual variables	Yes	Yes	Yes	Yes	Yes
Contextual effects	No	No	No	No	No
Labs' Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.51	–	–	–	–
AIC	–	390.8	393.4	651.2	653.1

Notes: RE stands for rational expectations, indicating that we assume in these models that individuals do not observe peers' choices but form expectation of these choices. MN stands for multinomial, i.e., individuals can choose between three modal choices instead of choosing only between an active or a non-active mode. † computed as the average over the sample of the partial derivatives of the regression equation with respect to each variable in the model for each observation in the data.

*p-value < 0.05; **p-value < 0.01; ***p-value < 0.001.

The results from the nonlinear model of peer effects offer supporting evidence for the existence of significant estimated peer effects in transportation mode choice (Table 5). Indeed, the point estimates for peer effects are positive and significant in the binary and multinomial logit models assuming observed peers' modal choices (logit and MN logit). These models yield considerable estimated marginal effects of peers' choices of active transportation on individual probabilities of choosing an active mode (20.1pp and 17.7pp, respectively). The estimated peer effects are also very close, indicating that a simple binary outcome is a satisfying simplification of multinomial modal choices when one is interested in active transportation mode.

The nonlinear models assuming that peers' modal choices are not perfectly observed by individuals and thus subject to rational expectations (logit RE and MN logit RE) yield stronger estimated marginal effects of peers' active transportation on individual probabilities although the estimated peer effects are not significant in the multinomial logit model.

Although our sample size might be too small to confidently estimate complex multinomial logit models, we do not find any evidence of peer effects in public transportation or car modal choices, although we find evidence of peer effects in active transportation in the MN logit model. This finding corroborates previous research on modal choices that also uncovers peer effects for bicycle use but not for public transport, for which spatial neighborhood effects seem more relevant (Pike, 2014). The difference in the absolute values of the estimated marginal effects between

1 the nonlinear models and the linear model is similar to the findings of Boucher and Bramoullé
 2 (2020), who estimate peer effects on (binary) smoking behavior with a linear model using a 2SLS
 3 estimator and compare the results to those obtained by Lee et al. (2014) with their nonlinear
 4 model. Even if the magnitude of the estimated marginal effects of the endogenous peer effects
 5 are different in the linear and nonlinear models, they are both significant and have the same sign.
 6
 7

8 However, they have the disadvantage of requiring a larger sample (the rule of thumb in the lit-
 9 erature is of at least ten observations per outcome type per independent variable in the model
 10 (Peduzzi et al., 1996)), which also easily leads to incidental parameter problems when lab fixed
 11 effects are added, increasing the number of explanatory variables in the model. Moreover, logit
 12 models do not directly yield econometric parameters that can be interpreted as marginal ef-
 13 fects. Instead, one has to estimate marginal effects as the average of individual increases in
 14 the probability of choosing an active mode for a marginal increase in each independent vari-
 15 able. Furthermore, nonlinear models do not have linear models' property of being robust to
 16 the assumption that peers' outcomes are based on individuals' expectations rather than directly
 17 observed by individuals (see [Identification via instrumental variables](#) for more details about this
 18 property in linear models instrumenting peers' behavior).
 19
 20
 21
 22
 23
 24
 25
 26

27 4.3 Alternative specifications of the linear local-average model of peer effects

28 **Sustainable transportation mode as dependent variable.** To assess whether peer effects
 29 are specific to active transportation modes, we estimate the linear models of peer effects with
 30 the choice of a sustainable transportation (combining active mode and public transportation)
 31 as the dependent variable. The estimation of the linear local-aggregate, local-average and gen-
 32 eralized models of peer effects on sustainable transportation mode reveals relatively strong but
 33 not significant estimated peer effects (Online Appendix [A.1](#)).
 34
 35
 36
 37
 38

39 **Alternative specifications of the individual payoff π_i .** The individual payoff component
 40 of the utility function (Equation [3](#)) can accommodate a variety of specifications, leading to vari-
 41 ous alternatives for the local-average models of peer effects. The estimation of these alternative
 42 specifications confirm that our preferred specification, including exogenous peer effects and lab
 43 fixed effects and accounting for network endogeneity, is the only one for which all instrument
 44 tests are satisfactory while the estimated peer effects are similar in magnitude in all specifica-
 45 tions (Online Appendix [A.2](#)).
 46
 47
 48
 49
 50

51 **Sociodemographic variables.** We also test several alternative specifications of the sociode-
 52 mographic variables (quadratic terms, interactions, thresholds) in the simplest specification of
 53 the linear local-average model. The estimated endogenous peer effects remain consistent across
 54 all different specifications of the sociodemographic variables (Online Appendix [A.3](#)).
 55
 56
 57

58 **Local nonlinearities in the linear local-average model of peer effects.** We conduct an
 59
 60
 61
 62
 63
 64
 65

1 additional estimation of local-average linear model, introducing nonlinearity in the endogenous
 2 peer effects through the control function approach proposed by Wooldridge (2015) to the 2SLS.
 3 We find that the linear local-average model is generally a satisfying specification, confirming
 4 our interpretation of the results of the nonlinear model of peer effects presented in Table 5.
 5 However, note that for extreme values of the instrumented peers' shares of active transporta-
 6 tion mode, i.e., when none or all of the peers use an active mode, the linear local-average model
 7 overestimates the endogenous peer effects (Online Appendix A.4).
 8
 9

10
 11 **Heterogeneous peer effects.** We also examine an alternative specification of the linear local-
 12 average model, aiming to introduce heterogeneity in social relationships and peer effect, i.e.,
 13 some peers may have more influence than others on one's behavior, building on Dieye and
 14 Fortin (2017) and Beugnot et al. (2019). Our hypothesis is that peer effects in transportation
 15 mode choice might be heterogeneous if individuals are influenced differently depending on the
 16 fact that they or their peers live close or far from their work place, have children or practice
 17 pro-environmental behaviors. Our analyses show that accounting for heterogeneous peer effects
 18 in the linear local-average model yields nonsignificant peer effects (Online Appendix A.5).
 19
 20
 21
 22
 23

24 4.4 Robustness assessments of the collected network data.

25
 26 **Weighted adjacency matrix.** Similarly to Lin (2014), we estimate the linear local-average
 27 model with a weighted specification of the adjacency matrix A , where a_{ij} is not binary (0 or
 28 1) but may take several values, representing the strength of the relationship. Weighting the
 29 resulting social interaction matrix weakens the significance and diminishes the strength of the
 30 estimated endogenous peer effects across all specifications of the linear local-average model (On-
 31 line Appendix A.6).
 32
 33
 34
 35

36 **Missing links in the network data.** Additionally, we adapt a method introduced by Pat-
 37 acchini et al. (2017) to address possible missing links in the adjacency matrix arising because
 38 the participants in our study could nominate a maximum of six peers, among the list of study
 39 participants, which might bias our results due to measurement error. To assess the existence
 40 and direction of such bias, we simulate adjacency matrices with a varying number of additional
 41 virtual links. Simulating new social relationships between individuals that are peers of peers
 42 yields a virtually identical estimate of endogenous peer effects (Online Appendix A.7). How-
 43 ever, simulating links between unrelated individuals (without common friends) decreases the
 44 magnitude of the estimated peer effects.
 45
 46
 47
 48
 49

50
 51 **Falsification test of the adjacency matrix.** Finally, we conduct a falsification test to ascer-
 52 tain whether the estimated peer effects might be a result of measurement errors in the network
 53 data, in which case they would not truly reflect social influence on modal choice (An, 2015;
 54 De Giorgi et al., 2010). To explore this, we perform a placebo test by randomly constructing
 55 alternative adjacency matrices with a similar density to the observed adjacency matrix. If the
 56 estimated peer effects under these artificial and random adjacency matrices turn out to be vir-
 57
 58
 59
 60
 61
 62
 63
 64
 65

1 tually zero and nonsignificant, it would suggest that the peer effects observed in our main model
2 are relatively robust and not merely driven by chance or measurement errors. Encouragingly,
3 the estimated endogenous peer effects are virtually 0 and nonsignificant (Online Appendix [A.8](#)),
4 indicating that the peer effects we observe in our sample are unlikely to be attributed to random
5 sample variation or measurement errors, supporting the validity and robustness of our findings.
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

5 Discussion and conclusion

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

Our results provide important evidence on the role of social interactions in adopting an active transportation mode. First, using a generalized model of peer effects including both conformism and strategic complementarity, we show that the desire to conform to peers' behavior drives most of the endogenous peer effects. Our preferred empirical model is thus the linear local-average model, which supports the tendency in the network econometrics' literature to use this model rather than the local-aggregate model in empirical research (Bramoullé et al., 2019). We provide evidence that social relationships in the workplace have a strong influence on individual behavior, even outside the workplace. We estimate that an increase of 33pp in the fraction of peers of a given individual (i.e., one peer out of three shifting to an active transportation mode) using an active transportation mode results in a 24pp increase in the own probability of choosing an active transportation mode. The estimated local-average peer effects are close to the effects found in a study of carpooling and car commuting in the US military, which reports that an increase of 20pp in the fraction of all coworkers from the same site (without defined social relationships) that carpool increases the probability of carpooling by 10pp (Morrison and Lawell, 2016). Similarly, Pike and Lubell (2018) identify a 85pp estimated marginal effect of the share of college friends biking on one's probability of biking to campus, instrumenting friends' modal choices by their neighborhoods' bike density. This indicates that an individual who has 100% of her peers commuting with a bicycle has a 85pp higher probability of also using a bicycle to commute than an individual with 0% of her peers commuting with a bicycle, *ceteris paribus*. However, the estimation strategies of these two studies do not take advantage of the network data to instrument peers' modal choice by intransitive relationships and only use information averaged at the workplace level or friends' neighborhoods. In Morrison and Lawell (2016)'s study, the source of social influence in the average modal choice of car and carpooling at the site level is instrumented using the share of workers born in Latin America and the share of workers that immigrated recently, respectively. In Pike and Lubell (2018), the source of peer effects is elicited friends' modal choice (bicycle or not) which is instrumented by friends' neighborhoods' bicycle density. Although they do not provide extensive testing of their instruments, we might expect that the instruments are still endogenous because social link formation is likely to be influenced by the origin and immigration status or by the neighborhood environment.

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Second, we show that the linear local-average model of peer effects is an adequate modeling strategy in the presence of binary outcomes, here active transportation mode choice. Indeed, to the best of our knowledge, the linear local-average model has been applied only twice to binary dependent variables, specifically in the context of alcohol consumption (Fletcher, 2012) and smoking (Boucher and Bramoullé, 2020), each time using the "Add Health" data and thus friendship relationships among students. We find endogenous peer effects of similar magnitude and significance as in the above-cited articles, despite that we use a smaller and less exhaustive data set and coworker networks, which tends to show that the local-average linear model with instrumental variables is an appealing estimation strategy of endogenous peer effects when col-

lection of network data is costly or nonexhaustive.

These results on the existence and strength of endogenous peer effects in modal choice call for increased consideration of social influence in policy interventions. Public policies could especially be more efficient if they incorporate the role of conformism into their design. Specifically targeting the peers of individuals who have small social networks in the workplace, among which the social norm would be more easily shifted, could be particularly efficient when conformism is the main driver of peer effects.

We also extend our results by estimating a nonlinear alternative model à la Lee et al. (2014). In contrast to the linear local-average model, whether one assumes that individuals observe or not their peers' modal choices has consequences for the modeling strategy and thus the estimated peer effects. The estimated marginal effects of peers' active modal choice in the nonlinear models of active transportation choice amount to 20.1pp and 33.8pp when peers' choices are assumed to be observable and when they are based on rational expectations respectively. Our results thus support the findings of Boucher and Bramoullé (2020): both linear and nonlinear models of peer effects with binary outcomes provide similar evidence on the direction and significance of peer effects, although the marginal effects estimated in the linear model are also three times larger than those in the nonlinear alternative.

In addition, we estimate a multinomial logit model to complement the nonlinear model of peer effects. Modeling modal choice as a discrete choice among three alternative (active modes, public transportation and car), we find that peer effects are significant if we assume that peers' modal choices are observed by individuals but not otherwise. In the case of modal choice in the workplace, we can legitimately assume that such behaviors are salient and well known among coworkers. The nonsignificance of peer effects in active transportation choice in the multinomial logit model with rational expectations about peers' modal choice is probably due to the fact that there are too many parameters involved in the fixed-point iteration required to solve the system of nonlinear equations.

This latter point underlines the data intensity of nonlinear models which is the major drawback of such models of peer effects in applied analyses with primary collected network data such as ours. The fact that peer effects' parameters are qualitatively similar in both nonlinear and linear models of peer effects is promising for future applications of network econometrics to studies of social influence on pro-environmental behaviors. Future research could attempt, for instance, to compare the two modeling strategies with larger samples and different contexts. Following the estimation strategy we developed, the role of conformism and strategic complementarity in other pro-environmental behaviors is also worth exploring.

We also provide various alternative specifications of our preferred empirical model of peer effects – the linear local-average model – and robustness assessments of the collected network data. We

1 show that accounting for exogenous peer effects and correlated effects helps to identify a stronger
2 and more significant endogenous peer effects. The main results are quantitatively robust to dif-
3 ferent specifications of the independent variables and qualitatively robust to weighting the social
4 interaction matrix by the intensity of the social relationships. However, accounting for the het-
5 erogeneity of peer effects based on spatial or familial constraints does not yield significant esti-
6 mated peer effects. Moreover, modifying the dependent variable to include public transportation
7 reduces the strength and significance of peer effects. This dilution of the endogenous peer effects
8 may be caused by the fact that a majority (76%) of respondents in our study use a sustainable
9 transportation mode while only 43% use an active transportation mode. Identifying peer ef-
10 fects is intuitively more difficult when the behavior in question is already the norm in the social
11 groups under scrutiny. In addition, social influence and peer effects might be stronger for active
12 transportation modes, as they are related to athletic motivations and (friendly) competition,
13 which may create stronger social emulation and intrinsic motivation than public transportation
14 (Rejeb et al., 2023). However, public transportation might be preferred by individuals with
15 higher gregariousness, as this transportation mode facilitates social interactions. Pike (2014)
16 also suggests that explicit peer networks play a greater role in explaining active transportation
17 choice whereas spatial neighborhoods are more relevant to analyze public transportation choice.
18
19
20
21
22
23
24
25

26 Nevertheless, these results must be interpreted with caution as we faced some empirical chal-
27 lenges. First, our data feature a rather small size, compared with other applications in network
28 econometrics, which usually have access to thousands of individuals and larger networks. A
29 small sample size increases the risk of type II errors, and more sociodemographic variables
30 could be found significant in a larger study. Second, the dataset we have is not randomized,
31 and a selection bias may arise because the individuals who responded to the survey might be
32 the most socially integrated in the research labs and thus might have a denser social network
33 or a higher propensity to conform to social norms. Third, the data we collected on network
34 are partial. We do not observe the full network because, on average, only 35% of the labs'
35 personnel answered the survey. As these 334 participants could only nominate peers who also
36 agreed to participate in the study, there is certainly a substantial number of missing nodes
37 and edges in the induced networks. While our robustness analyses account for possible missing
38 edges and measurement errors in the networks, the biases of the estimated coefficients due to
39 the missing nodes remain (Chandrasekhar and Lewis, 2016). Furthermore, the sample we use
40 is not representative of the French population, as only 24% of the respondents in our study use
41 private cars to commute, whereas the average in France is at least two to three times as large
42 (Brutel and Pages, 2021). Nonetheless, the average transportation modes in our sample are in
43 line with the specific situation of Grenoble, which is one of the cities where inhabitants use the
44 most active and public transportation in France because of political and topographic reasons
45 (the city has an environmentalist mayor and is extraordinarily flat). Similarly, we surveyed
46 research labs, in which the workers are generally more educated, have higher incomes and are
47 more aware of the environmental and health externalities of the different transportation modes
48 than the general population. Social relationships among research labs could also differ in nature
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 with relationships in private companies where hierarchy may play a larger role. The strong peer
2 effects we estimate are likely upper bounds of the peer effects in active transportation modal
3 choice. They thus should not be generalized to the whole population or to other contexts, and
4 further research is necessary to externally validate our empirical results.
5

6
7 Collecting larger network data using the recently developed strategy of Breza et al. (2020) with
8 *aggregated relational data* could resolve some of these empirical issues. Indeed, collecting such
9 data does not require eliciting peers by name and can be used to reconstitute an entire network
10 through a sample of respondents. Similarly, gathering panel data or information on a larger
11 variety of trips per individual – not only commuting – might refine our results (Kim et al., 2017).
12 Indeed, individuals plan their trips as a tour, and commuting trips are not always independent
13 of trips planned after or before work, which might constrain individuals' modal choices. In that
14 case, encouraging someone to change her commuting mode would also indirectly yield a modal
15 shift for other activities, increasing health and environmental benefits. Moreover, gathering
16 panel data can reveal additional information, such as the estimation of long-term peer effects
17 after a policy or intervention (Alacevich et al., 2021) or the dynamic evolution of the networks
18 when individuals change their behaviors (Badev, 2021; Boucher, 2016). Collecting panel data on
19 the networks would allow researchers to directly observe the marginal impact of link formation
20 or dissolution on individual behavior as well as the impact of individual modal choices on these
21 processes. The development of smartphones and tracking apps also offers new opportunities
22 to capture individual habits and influence mobility behavior (Hintermann et al., 2021). Future
23 research could integrate tracking and social gamification to explore the direct effect of social
24 influence in mobility behavior.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

6 References

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Akerlof, G. (1997), ‘Social distance and social decisions’, *Econometrica* **65**(95), 1005.
- Alacevich, C., Bonev, P. and Söderberg, M. (2021), ‘Pro-environmental interventions and behavioral spillovers: Evidence from organic waste sorting in sweden’, *Journal of Environmental Economics and Management* **108**.
- An, W. (2015), ‘Instrumental variables estimates of peer effects in social networks’, *Social Science Research* **50**, 382–394.
- Badev, A. (2021), ‘Nash equilibria on (un)stable networks’, *Econometrica* **89**(3), 1179–1206.
- Bajari, P., Hong, H., Krainer, J. and Nekipelov, D. (2010), ‘Estimating static models of strategic interactions’, *Journal of Business & Economic Statistics* **28**(4), 469–482.
- Beugnot, J., Fortin, B., Lacroix, G. and Villeval, M. C. (2019), ‘Gender and peer effects on performance in social networks’, *European Economic Review* **113**, 207–224.
- Blume, L. E., Brock, W. A., Durlauf, S. N. and Ioannides, Y. M. (2011), Identification of social interactions, in ‘Handbook of social economics’, Vol. 1, Elsevier, pp. 853–964.
- Boucher, V. (2016), ‘Conformism and self-selection in social networks’, *Journal of Public Economics* **136**, 30–44.
- Boucher, V. and Bramoullé, Y. (2020), ‘Binary outcomes and linear interactions’, *AMSE Working Papers 2038*.
- Bouscasse, H., Gabet, S., Kerneis, G., Provent, A., Rieux, C., Ben Salem, N., Dupont, H., Troude, F., Mathy, S. and Slama, R. (2022), ‘Designing local air pollution policies focusing on mobility and heating to avoid a targeted number of pollution-related deaths: Forward and backward approaches combining air pollution modeling, health impact assessment and cost-benefit analysis’, *Environment International* **159**.
- Bouscasse, H., Mathy, S., Rejeb, R. and Treibich, C. (2023), ‘Is the impact of transport modes on health an individual determinant of transport mode choices?’.
- Bramoullé, Y., Djebbari, H. and Fortin, B. (2009), ‘Identification of peer effects through social networks’, *Journal of Econometrics* **150**, 41–55.
- Bramoullé, Y., Djebbari, H. and Fortin, B. (2019), ‘Peer effects in networks: a survey’, *Annual Review of Economics* **12**, 603–619.
- Breza, E., Chandrasekhar, A. G., McCormick, T. H. and Pan, M. (2020), ‘Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data’, *American Economic Review* **110**(8), 2454–2484.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Brock, W. A. and Durlauf, S. N. (2001), 'Discrete Choice with Social Interactions', *The Review of Economic Studies* **68**(2), 235–260.
URL: <https://academic.oup.com/restud/article-lookup/doi/10.1111/1467-937X.00168>
- Brock, W. and Durlauf, S. N. (2003), 'Multinomial choice with social interactions'.
- Brutel, C. and Pages, J. (2021), La voiture reste majoritaire pour les déplacements domicile-travail, même pour de courtes distances, Technical Report 1835.
- Chandrasekhar, A. G. and Lewis, R. (2016), 'Econometrics of sampled networks'.
- Cialdini, R. B. and Goldstein, N. J. (2004), 'Social influence: Compliance and conformity', *Annu. Rev. Psychol.* **55**, 591–621.
- Clark, A. E. and Oswald, A. J. (1998), 'Comparison-concave utility and following behaviour in social and economic settings', *Journal of Public Economics* **70**(1), 133–155.
- De Giorgi, G., Pellizzari, M. and Redaelli, S. (2010), 'Identification of social interactions through partially overlapping peer groups', *American Economic Journal: Applied Economics* **2**(2), 241–75.
- De Groot, J. I. and Schuitema, G. (2012), 'How to make the unpopular popular? policy characteristics, social norms and the acceptability of environmental policies', *Environmental Science & Policy* **19**, 100–107.
- De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M. and Macharis, C. (2013), 'Linking modal choice to motility: A comprehensive review', *Transportation Research Part A: Policy and Practice* **49**, 329–341.
- Dieye, R. and Fortin, B. (2017), 'Gender peer effects heterogeneity in obesity', *SSRN Journal* .
- Dugundji, E. and Gulyás, L. (2008), 'Sociodynamic discrete choice on networks in space: impacts of agent heterogeneity on emergent outcomes', *Environ. Plann. B: Plann. Des.* **35**(6), 1028–1054.
- Falk, A. and Ichino, A. (2006), 'Clean evidence on peer effects', *Journal of labor economics* **24**(1), 39–57.
- Farrow, K., Grolleau, G. and Ibanez, L. (2017), 'Social norms and pro-environmental behavior: A review of the evidence', *Ecological Economics* **140**, 1–13.
- Fletcher, J. (2012), 'Peer influences on adolescent alcohol consumption: evidence using an instrumental variables/fixed effect approach', *Journal of Population Economics* **25**(4), 1265–1286.
- Goetzke, F. and Andrade, P. (2010), 'Walkability as a summary measure in a spatially autoregressive mode choice model: An instrumental variable approach', *Spatial Statistics Methods and Applications* .

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Goetzke, F. and Rave, T. (2011), ‘Bicycle use in germany: Explaining differences between municipalities with social network effects’, *Urban studies* **48**(2), 427–437.
- Graham, B. S. (2017), ‘An Econometric Model of Network Formation With Degree Heterogeneity’, *Econometrica* **85**(4), 1033–1063.
- Guerra, J.-A. and Mohnen, M. (2022), ‘Multinomial choice with social interactions: Occupations in victorian london’, *Review of Economics and Statistics* **104**(4), 736–747.
- Hintermann, B., Schoeman, B. M., Molloy, J., Schatzmann, T., Tchervenkov, C. and Axhausen, K. W. (2021), The impact of covid-19 on mobility choices in switzerland, Technical report, WWZ Working Paper.
- Houndetoungan, E. A. (2020), ‘A Count Data Model with Social Interactions’, *SSRN Electronic Journal*.
- Hsieh, C.-S., Lin, X. and Patacchini, E. (2020), Social Interaction Methods, in K. F. Zimmermann, ed., ‘Handbook of Labor, Human Resources and Population Economics’, Springer International Publishing, Cham, pp. 1–30.
- IPSOS and BCG (2017), Observatoire européen des mobilités, Technical report.
- Johnsson, I. and Moon, H. R. (2021), ‘Estimation of peer effects in endogenous social networks: control function approach’, *Review of Economics and Statistics* **103**(2), 328–345.
- Kelejian, H. H. (2008), ‘A spatial j-test for model specification against a single or a set of non-nested alternatives’, *Letters in Spatial and Resource Sciences* **1**(1), 3–11.
- Kim, C., Parent, O. and vom Hofe, R. (2017), ‘The role of peer effects and the built environment on individual travel behavior’, *Environment and Planning B: Urban Analytics and City Science* **45**, 452–469.
- Kormos, C., Gifford, R. and Brown, E. (2015), ‘The influence of descriptive social norm information on sustainable transportation behavior: A field experiment’, *Environment and Behavior* **47**, 479–501.
- Lee, L.-f., Li, J. and Lin, X. (2014), ‘Binary Choice Models with Social Network under Heterogeneous Rational Expectations’, *Review of Economics and Statistics* **96**(3), 402–417.
URL: <https://direct.mit.edu/rest/article/96/3/402-417/58168>
- Lin, X. (2014), ‘Peer effects in adolescents’ delinquent behaviors: Evidence from a binary choice network model’, *Regional Science and Urban Economics* **46**, 73–92.
- Lin, Z., Tang, X. and Yu, N. N. (2021), ‘Uncovering heterogeneous social effects in binary choices’, *Journal of Econometrics* **222**(2), 959–973.
- Lindquist, M. J., Sauermann, J. and Zenou, Y. (2022), ‘Peer effects in the workplace. a network approach’.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Liu, X., Patacchini, E. and Zenou, Y. (2014), 'Endogenous peer effects: local aggregate or local average?', *Journal of Economic Behavior & Organization* **103**, 39–59.
- Maness, M., Cirillo, C. and Dugundji, E. R. (2015), 'Generalized behavioral framework for choice models of social influence: Behavioral and data concerns in travel behavior', *Journal of Transport Geography* **46**, 137–150.
- Manski, F. C. (1993), 'Identification of endogenous social effects: The reflection problem', *Review of Economic Studies* **60**, 531–542.
- McFadden, D. L. (1984), 'Econometric analysis of qualitative response models', *Handbook of econometrics* **2**, 1395–1457.
- Morgan, M., Young, M., Lovelace, R. and Hama, L. (2019), 'Opentripplanner for r', *Journal of Open Source Software* **4**(44), 1926.
- Morrison, G. M. and Lawell, C.-Y. C. L. (2016), 'Driving in force: The influence of workplace peers on commuting decisions on us military bases', *Journal of Economic Behavior & Organization* **125**, 22–40.
- Patacchini, E., Rainone, E. and Zenou, Y. (2017), 'Heterogeneous peer effects in education', *Journal of Economic Behavior & Organization* **134**, 190–227.
- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. and Feinstein, A. R. (1996), 'A simulation study of the number of events per variable in logistic regression analysis', *Journal of clinical epidemiology* **49**(12), 1373–1379.
- Phithakkitnukoon, S., Sukhvibul, T., Demissie, L., Smoreda, Z., Natwichai, J. and Bento, C. (2017), 'Inferring social influence in transport mode choice using mobile phone data', *EPJ Data Science* **6**, 11.
- Pike, S. (2014), 'Travel mode choice and social and spatial reference groups: Comparison of two formulations', *Transportation Research Record* **1**(75-81), 531–542.
- Pike, S. and Lubell, M. (2016), 'Geography and social networks in transportation mode choice', *Journal of Transport Geography* **57**, 184–193.
- Pike, S. and Lubell, M. (2018), 'The conditional effects of social influence in transportation mode choice', *Research in Transportation Economics* **68**, 2–10.
- Pustejovsky, J. E. and Tipton, E. (2018), 'Small-sample methods for cluster-robust variance estimation and hypothesis testing in fixed effects models', *Journal of Business & Economic Statistics* **36**(4), 672–683.
- URL:** <https://doi.org/10.1080/07350015.2016.1247004>
- Rejeb, R., Bouscasse, H., Chalabaev, A. and Mathy, S. (2023), 'What is the role of active mobility habits in the relationship between self-determination and modal shift intentions? a mediation analysis', *Submitted to Traffic Psychology and Behaviour*.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Robertson, J. L. and Barling, J. (2013), 'Greening organizations through leaders' influence on employees' pro-environmental behaviors', *Journal of organizational behavior* **34**(2), 176–194.
- Rojas-Rueda, D., de Nazelle, A., Teixidó, O. and Nieuwenhuijsen, M. (2013), 'Health impact assessment of increasing public transport and cycling use in barcelona: A morbidity and burden of disease approach', *Preventive Medicine* **57**(5), 573–579.
- Sheeran, P. and Webb, T. L. (2016), 'The Intention-Behavior Gap', *Social and Personality Psychology Compass* **10**(9), 503–518.
- Sherwin, H., Chatterjee, K. and Jain, J. (2014), 'An exploration of the importance of social influence in the decision to start bicycling in england', *Transportation Research Part A: Policy and Practice* **68**, 32–45.
- Smith, J. R., Louis, W. R., Terry, D. J., Greenaway, K. H., Clarke, M. R. and Cheng, X. (2012), 'Congruent or conflicted? the impact of injunctive and descriptive norms on environmental intentions', *Journal of environmental psychology* **32**(4), 353–361.
- Videras, J., Conover, E. and Wu, S. (2011), 'The influence of social relationships on pro-environment behaviors', *Journal of Environmental Economics and Management* **63**.
- Walker, J. L., Ehlers, E., Banerjee, I. and Dugundji, E. (2005), 'Correcting for endogeneity in behavioral choice models with social influence variables', *Transportation Research Part A: Policy and Practice* **45**(4), 362–374.
- Wang, H., Akar, G. and Guldman, J. (2015), 'Do your neighbors affect your bicycling choice? a spatial probit model for bicycling to the ohio state university', *J. Transp. Geogr.* **42**, 122–130.
- Wesselink, R., Blok, V. and Ringersma, J. (2017), 'Pro-environmental behaviour in the workplace and the role of managers and organisation', *Journal of cleaner production* **168**, 1679–1687.
- Wolske, K. S., Gillingham, K. T. and Schultz, P. (2020), 'Peer influence on household energy behaviours', *Nature Energy* **5**(3), 202–212.
- Wooldridge, J. M. (2015), 'Control function methods in applied econometrics', *The Journal of Human Resources* .