Journal of Urban Economics 83 (2014) 6–17

Contents lists available at ScienceDirect

Journal of Urban Economics

www.elsevier.com/locate/jue

Peer effects in the demand for housing quality ☆

Eleonora Patacchini [a,b,c,*], Giuseppe Venanzoni [d]

[a] Cornell University, United States
[b] EIEF, Italy
[c] CEPR, UK
[d] Sapienza University of Rome, Italy

ARTICLE INFO

Article history:
Received 9 February 2013
Revised 12 June 2014
Available online 24 June 2014

JEL classification:
A14
C21
D85
R21
Z13

Keywords:
Social networks
Linear-in-means model
Spatial autoregressive model
Social norms

ABSTRACT

Using detailed data on friendship networks within neighborhoods, we investigate the importance of social interactions in one's own residential neighborhood in the demand for housing quality. We find evidence consistent with the presence of peer effects, especially for households living in urban areas. Our findings are in line with the prediction of a model where conformity preferences underlie economic outcomes that involve interactions with peers.

1. Introduction

There is an increasing recognition in economics that social interactions play a major role in explaining a range of individual behaviors, as well as the individual's valuation of both the decision and the resulting outcome.[1] Peer effects have been indicated as important determinants of behavior in a variety of contexts. Examples include education, crime, labor market, fertility, obesity, productivity, participation in welfare programs, risky behavior, to mention a few (for surveys, see Glaeser and Scheinkman, 2001; Moffitt, 2001; Durlauf, 2004; Ioannides and Loury, 2004; Jackson, 2008; Ioannides, 2012). In many social phenomena peer effects stems from preferences for conformity. Conformism is the idea that the easiest and hence best life is attained by doing one's very best to blend in with one's surroundings, and to do nothing eccentric or out of the ordinary in any way. In an economy with conformity preferences peer effects are viewed as a social norm and individuals pay a cost from deviating from this norm. Different aspects of conformism and social norms have been explored from a theoretical point of view. To name a few, (i) peer pressures and partnerships (Kandel and Lazear, 1992) where peer pressure arises when individuals deviate from a well-established group norm, e.g. individuals are penalized for working less than the group norm, (ii) religion (Iannaccone, 1992; Berman, 2000) since praying is much more satisfying the more participants there are, (iii) social status and social distance (Akerlof, 1980, 1997; Bernheim, 1994, among others) where deviations from the social norm (average action) imply a loss of reputation and status, and (iv) crime (Glaeser et al., 1996; Patacchini and Zenou, 2012) where individual wants to minimize the social distance between her crime level and that of her reference group.

☆ We are grateful to the editor, William Strange, as well as three anonymous referees for helpful comments. We also thank Gary Engelhardt and Stuart Rosenthal for constructive suggestions. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

* Corresponding author at: Cornell University, United States.
E-mail addresses: eleonora.patacchini@cornell.edu (E. Patacchini), giuseppe.venanzoni@uniroma1.it (G. Venanzoni).

[1] The integration of models of social interactions within economic theory is an active and interesting area of research. See the recent *Handbook of Social Economics* (Benhabib et al., 2011).

http://dx.doi.org/10.1016/j.jue.2014.06.002
0094-1190/

In this paper we study whether conformist behavior affects the individual demand for housing quality. The literature on social interactions in the housing market is extremely limited (see Ioannides, 2012 for a critical survey)[2] and presents two important challenges: (i) to disentangle peer effects from neighborhood effects and (ii) to explain *how* peers influence each other, i.e. the mechanism generating such social interactions.

The study of peer effects in housing decisions is paramount for policy purposes. One of the reason suggesting government intervention in the housing market is inefficiency in housing consumption. Housing renovations improve not only one's own property but also neighbors' property values. However, this externality is not internalized in the individual's calculation of whether or not to undertake an improvement. As a results, the marginal social benefits of the improvement exceed the private marginal costs, and the property owner is likely to invest less than a socially efficient amount. Under this perspective, the existence of peer effects could overcome the underprovision of local public goods (Rosen, 1985).

Our analysis uses detailed data on friendship networks to measure peer groups more precisely than previous studies and elaborates on a conformism model, presented by Patacchini and Zenou (2012), to guide the interpretation of the results.[3] More precisely, borrowing from Patacchini and Zenou (2012), we first present a social network model of peer effects that show how conformism affects the demand for housing quality. We then take the model to the data by using the U.S. National Longitudinal Survey of Adolescent Health (AddHealth). This data contains unique information on friendship relationships among a representative sample of students from U.S. high school teenagers together with residential neighborhood identifiers. The survey design also includes a questionnaire administered to the interviewers which collects information on the type and quality of the respondent's residential building and area of residence. These questions are thus informative of each student's household decisions about house maintenance, repair and renovation. Under the assumption that the children's social contacts in the neighborhood are a good approximation of their parents' social contacts, these data are thus able to shed some light on the importance of social interactions in the demand for housing quality.

Empirical tests of models of social interactions are quite problematic. The issues that render the identification and measurement of peer effects quite difficult are well known: (i) reflection, which is a particular case of simultaneity (Manski, 1993) and (ii) endogeneity, which may arise for both peer self-selection and unobserved common (group) correlated effects.

In this paper, we exploit the architecture of social networks to overcome this set of problems and to achieve the identification of endogenous peer effects. More specifically, in social networks, each agent has a different peer-group, i.e. different friends with whom each teenager directly interacts. This feature of social networks guarantees the presence of excluded friends from the reference group (peer-group) of each agent, which are, however, included in the reference group of his/her best (direct) friends. This identification strategy is similar in spirit to the one used in the standard simultaneous equation model, where at least one exogenous variable needs to be excluded from each equation. In addition, because we observe precise patterns of social interactions, we can include network fixed effects in the empirical specification of the model. By doing so, we are thus able to disentangle peer effects from the presence of network unobserved factors affecting both individual and peer behaviors. Such factors might be important omitted variables driving the sorting of agents into networks. The application of this strategy in our context is based on the key premise that the children's social contacts in the neighborhood are a good approximation of their parents' social contacts. Indeed, the decisions about home repairs, maintenance and upkeep are taken by the parents. Evidence in support of the validity of this strategy is provided.

Our findings reveals statistically significant peer effects in the individual demand for housing quality. The analysis of peer effects is, however, a complex issue and our analysis has some limitations. First, our model is only one of the possible mechanisms generating such externalities. It is not, however, rejected by our data and it serves to highlight the importance of non-market interactions in explaining individual demand for housing quality. Second, in the absence of experimental data, one can never be sure to have captured all the behavioral intricacies that lead individuals to associate with others. In addition, our data provides an imprecise measure of the demand for housing quality. Finally, our friendship networks may be measured with error – we assume that the children's social contacts in the neighborhood are a good approximation of their parents' social contacts. Nevertheless, by using both within- and between-network variation and by taking advantage of the unusually large information on teenagers' behavior provided by our dataset, our analysis is a valid attempt to overcome the empirical difficulties.

The rest of the paper unfolds as follows. In the next section, we present the theoretical framework that helps us to understand how social contacts can influence individual demand for housing quality. Section 3 describes the data and the empirical strategy. We present our empirical results in Section 4, whereas Section 5 contains some robustness checks. Finally, Section 6 concludes.

2. Theoretical framework

Following Patacchini and Zenou (2012), we present a social network model of peer effects with conformity preferences for the demand of housing quality.

There are $N = \{1, \ldots, n\}$ individuals in the economy distributed among K networks. Let n_k be the number of individuals in the kth network, so that $N = \sum_{k=1}^{K} n_k$.

2.1. The network

The adjacency matrix $\mathbf{G} = [g_{ij}]$ of a network k keeps track of the direct connections in this network. Here, two players i and j are directly connected (i.e. best friends) in k if and only if $g_{ij,k} = 1$, and $g_{ij,k} = 0$, otherwise. Given that friendship is a reciprocal relationship, we set $g_{ij,k} = g_{ji,k}$.[4] We also set $g_{ii,k} = 0$. The set of individual i's best friends (direct connections) is: $N_i(k) = \{j \neq i | g_{ij,k} = 1\}$, which is of size $g_{i,k}$ (i.e. $g_{i,k} = \sum_{j=1}^{n} g_{ij,k}$ is the number of direct links of individual i). This means in particular that, if i and j are best friends, then in general $N_i(k) \neq N_j(k)$ unless the graph/network is complete (i.e. each individual is friend with everybody in the network). This also implies that groups of friends may overlap if individuals have common best friends. To summarize, the *reference group* of

[2] Most notably, Ioannides and Zabel (2003) consider the housing demand for a group of neighbors as a system of simultaneous equations. Ioannides and Zabel (2008) develop a model of housing demand with neighborhood effects and of neighborhood choice as a joint decision. Rossi-Hansberg et al. (2010) provide evidence that in neighborhoods targeted by the a revitalization program, sites that did not directly benefit from capital improvements nevertheless experienced considerable increases in land value relative to similar sites in a control neighborhood.

[3] The constraints imposed by the available disaggregated data force many studies to analyze peer effects at a quite aggregate and arbitrary level, such as at the neighborhood level (see, e.g., Durlauf, 2004; Ioannides and Topa, 2010; Ioannides, 2011).

[4] This is not an important assumption since all our theoretical results hold even when $g_{ij,k} \neq g_{ji,k}$. We discuss this issue in Section 5.

each individual i is $N_i(k)$, i.e. the set of his/her best friends, which does not include him/herself.

2.2. Preferences

Individuals in network k decide how much effort to exert in home maintenance, repair and renovation. We denote by $y_{i,k}$ the effort level of individual i in network k and by $Y = (y_{1,k}, \ldots, y_{n,k})'$ the population effort profile in network k. Denote by $\bar{y}_{i,k}$ the average effort of individual i's best friends. It is given by:

$$\bar{y}_{i,k} = \frac{1}{g_{i,k}} \sum_{j=1}^{n} g_{ij,k} y_{j,k} \tag{1}$$

Each agent i in network k selects an effort $y_{i,k} \geqslant 0$, and obtains a payoff $u_{i,k}(Y,k)$ that depends on the effort profile Y and on the underlying network k, in the following way:

$$u_{i,k}(Y,k) = (a_{i,k} + \eta_k + \varepsilon_{i,k}) y_{i,k} - \frac{1}{2} y_{i,k}^2 - \frac{d}{2} (y_{i,k} - \bar{y}_{i,k})^2 \tag{2}$$

where $d > 0$. The benefit part of this utility function is given by $(a_{i,k} + \eta_k + \varepsilon_{i,k}) y_{i,k}$ while the cost is $\frac{1}{2} y_{i,k}^2$; both are increasing in own effort $y_{i,k}$. In this part, $a_{i,k}$ denotes the agent's ex ante *idiosyncratic heterogeneity*, which is assumed to be deterministic, perfectly *observable* by all individuals in the network and corresponds to the observable characteristics of individual i (e.g. sex, race, age, parental education) and to the observable average characteristics of individual i's best friends, i.e. average level of parental education of i's friends, etc. (contextual effects). To be more precise, $a_{i,k}$ can be written as:

$$a_{i,k} = \sum_{m=1}^{M} \beta_m x_{i,k}^m + \frac{1}{g_i} \sum_{m=1}^{M} \sum_{j=1}^{n} \theta_m g_{ij}\ x_{j,k}^m \tag{3}$$

where x_i^m is a set of M variables accounting for observable differences in individual characteristics of individual i, and β_m, θ_m are parameters. In the utility function (2) η_k denotes the unobservable network characteristics and $\varepsilon_{i,k}$ is an error term, meaning that there is some uncertainty in the benefit part of the utility function. Both η_k and $\varepsilon_{i,k}$ are observed by the individuals but not by the researcher. The second part of the utility function $\frac{d}{2}(y_{i,k} - \bar{y}_{i,k})^2$ reflects the influence of friends' behavior on own action. It is such that each individual wants to minimize the *social distance* between herself and her reference group, where d is the parameter describing the *taste for conformity*. Here, the individual loses utility $\frac{d}{2}(y_{i,k} - \bar{y}_{i,k})^2$ from failing to conform to others. This is the standard way economists have been modelling conformity (see, among others, Akerlof, 1980; Bernheim, 1994; Kandel and Lazear, 1992; Akerlof, 1997; Fershtman and Weiss, 1998; Patacchini and Zenou, 2012). In the context of the demand for housing quality, a taste for conformity captures the idea of " keeping up with the Joneses," where individuals view their neighbors' decisions about maintenance, repair and renovation, and do their best to keep up by making similar decisions.[5] The social norm can be interpreted as friends' social status, as signalled by house quality.

Observe that the social norm here captures the differences between individuals due to network effects. It means that individuals have different types of friends and thus different reference groups $\bar{y}_{i,k}$. As a result, the social norm each individual i faces is endogenous and depends on her location in the network as well as the structure of the network.

2.3. Nash equilibrium

In this game where agents choose their effort level $y_{i,k} \geqslant 0$ simultaneously, there exists a unique Nash equilibrium (Patacchini and Zenou, 2012) given by:

$$y_{i,k}^* = \phi \frac{1}{g_i} \sum_{j=1}^{n_k} g_{ij} y_j^* + (1 - \phi)(a_{i,k} + \eta_k + \varepsilon_{i,k}) \tag{4}$$

where $\phi = d/(1+d)$. The optimal effort level depends on the individual ex ante heterogeneity ($a_{i,k}$), on the unobserved network characteristics (η_k) and it is increasing with the average effort of the reference group. This means that the more well kept the houses of one's friends are, the more the individual will provide effort in the upkeep of her own house.

3. Data and empirical strategy

3.1. Data

Our data source is the National Longitudinal Survey of Adolescent Health (AddHealth), which contains detailed information on a nationally representative sample of 90,118 students in roughly 130 private and public schools, entering grades 7–12 in the 1994–1995 school year. Every pupil attending the sampled schools on the interview day is asked to complete a questionnaire (*in-school survey*) containing questions on respondents' demographic and behavioral characteristics, education, family background and friendship. A subset of adolescents selected from the rosters of the sampled schools, about 20,000 individuals, is then asked to complete a longer questionnaire containing questions relating to more sensitive individual and household information (*in-home survey* and parental data). AddHealth contains unique information on friendship relationships, which is crucial for our analysis. The friendship information is based upon actual friends nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[6] The uniqueness of this information lies on the fact that by matching the identification numbers of the friendship nominations to respondents' identification numbers, one can also obtain information on the characteristics of nominated friends.[7] Importantly, these data also contain each respondent's residential neighborhood identifier.[8] Hence, it is possible to reconstruct the geometry of the friendship networks within each neighborhood. Neighborhoods are defined as census tracts. Our networks are constructed by considering that a link exists between two friends if at least one of the two individuals has identified the other as his/her best friend.[9]

Besides information on family background, school quality and area of residence, the AddHealth data enclose information on the interviewer's remarks after having visited the students' house for

[5] Morris and Winter (1975, 1978) introduced the notion of "housing deficit" to conceptualize residential (dis)satisfaction. In their housing adjustment model of residential mobility, they theorize that individuals judge their housing conditions according to predefined norms, which are dictated by societal living standards or rules.

[6] The limit in the number of nominations is not binding, not even by gender. Less than 1% of the students in our sample show a list of ten best friends, less than 3% a list of five males and roughly 4% name five females.

[7] The other existing survey data collecting information on social contacts (e.g. NSHAP, BHPS, GSOEP) are "ego networks". They contain a list of the contacts each respondent declares with few demographic characteristics (gender, relationship with respondent, education) of each contact, which are self-reported by the respondent. No extensive interview to each nominated contact is performed.

[8] The data also provide geo-coded information (latitude and longitude coordinates of residential address of each respondent). See Del Bello et al. (2014) for a paper using this information to disentangle the relative importance of geographical distance and social distance in peer effects.

[9] Note that, when an individual i identifies a best friend j who does not belong to the surveyed schools, the database does not include j in the network of i; it provides no information about j. However, in the large majority of cases (more than 94%), students tend to nominate best friends who are students in the same school and thus are systematically included in the network (and in the neighborhood patterns of social interactions).

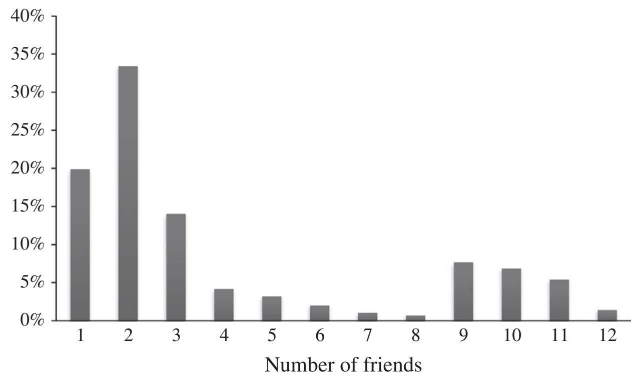

Fig. 1. Distribution of students by number of social contacts.

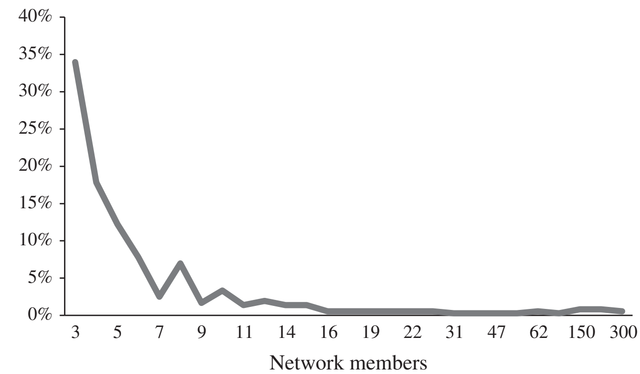

Fig. 2. Distribution of networks by network size.

the in-home interview. We use this information to construct our dependent variable $y_{i,k}$. Specifically, the interviewer is asked: " How well kept is the building in which the respondent lives?", with possible answers "very poorly kept (needs major repairs)", "poorly kept (needs minor repairs)", "fairly well kept (needs cosmetic work)" and "very well kept", coded 1–4.[10] Slightly more than 40% of the respondent answers "very well kept", about 28% "fairly well kept (needs cosmetic work)", roughly 20% "Poorly kept (needs minor repairs)", and 9% "very poorly kept (needs major repairs)". The interviewer questionnaire also asks to describe the immediate area or street (one block, both sides) where the respondent lives. We use this question to investigate whether peer effects in the demand for housing quality differ between urban and nonurban areas.[11] Using the corresponding information for nominated friends, we are able, for each individual i in network k, to calculate the average effort $\bar{y}_{i,k}$ of his/her peer group. Excluding the individuals with missing or inadequate information, we obtain a final sample of 3908 students distributed over 359 networks. This large reduction in sample size with respect to the original sample is mainly due to the network construction procedure – roughly 20% of the students do not nominate any friends and another 20% cannot be correctly linked.[12] In addition, we exclude networks at the extremes of the network size distribution (i.e. consisting of 2 individuals or more than 300) because peer effects can show extreme values (too high or too low) in these edge networks (see Calvó-Armengol et al., 2009).

Fig. 1 shows the distribution of students by number of friends. While on average AddHealth students name about 2.5 friends, there is a large dispersion around this mean value. Fig. 1 reveals that the distribution is bi-modal, with the large majority of students having between 1 and 3 friends, and a sizeable number of them with many friends, between 9 and 11. Fig. 2 shows the distribution of networks by network size. One can see that the social circles in our sample are quite small. Indeed, the large majority of social networks (more than 70%) have between 3 and 7 members.

Table 1 gives the definition of the variables used in our study as well as their descriptive statistics in our sample. Among the individuals selected in our sample, 47% are female and 21% are non-whites. The average parental education is more than high-school graduate. Roughly 10% have parents working in a managerial occupation, 8% in a professional/technical occupation, 25% in the office or sales sector, and 27% have parents in manual occupations. Roughly 20% of our individuals come from a household with only one parent. On average, our adolescents live in a household of about 3.5 people.

3.2. Empirical strategy

Guided by the behavioral mechanism formalized in Section 2, our aim is to assess the actual empirical relationship between the neighbors' effort $\bar{y}^*_{i,k}$ and individual effort level $y^*_{i,k}$.

The empirical equivalent of the first order conditions of our network model of peer effects (Eq. (4)) is given by:

$$y_{i,\kappa} = \phi \frac{1}{g_{i,k}} \sum_{j=1}^{n_\kappa} g_{ij,k}\, y_{j,k} + \sum_{m=1}^{M} \beta_1^m x_{i,\kappa}^m + \frac{1}{g_{i,k}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \theta_m g_{ij,k} x_{j,\kappa}^m + \eta_k + \varepsilon_{i,k} \tag{5}$$

where $y_{i,\kappa}$, is the housing quality of the household of student i in network k, $x_{i,\kappa}^m$ (for $m = 1, \ldots, M$) is the set of M control variables, $g_{i,k} = \sum_{j=1}^{n_\kappa} g_{ij,k}$ is the number of direct links of i, $\sum_{j=1}^{n_\kappa} \left(g_{ij,k} x_{j,\kappa}^m \right) / g_{i,k}$ is the set of the average values of the M controls of i's direct friends (i.e. contextual effects). As stated in the theoretical model, $\sum_{m=1}^{M} \beta_1^m x_{i,\kappa}^m + \frac{1}{g_{i,k}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \theta_m g_{ij,k} x_{j,\kappa}^m$ reflects the ex ante idiosyncratic heterogeneity of each individual i, and our measure of *taste for conformity* or *strength of peer effects* is captured by the parameter ϕ (in the theoretical model $\phi = d/(1+d)$). To be more precise, $\phi = d/(1+d)$ measures the taste for conformity relative to the direct, time or psychological costs of home repair and maintenance. Finally, η_k captures network specific unobserved factors (constant over individuals in the same network), which might be correlated with the regressors, and $\varepsilon_{i,k}$ is a white noise error.[13]

In model (5), ϕ represents *the endogenous effects*, where an agent's choice/outcome may depend on those of his/her friends on the same activity; and θ represents *the contextual effect*, where an agent's choice/outcome may depend on the exogenous characteristics of his/her friends. The vector of network fixed effects η_k captures *the correlated effect* where agents in the same network may behave similarly as they have similar unobserved individual characteristics or they face a similar environment.

A number of papers have dealt with the identification and estimation of peer effects in model (5) using network data (e.g. Clark and Loheac, 2007; Lee, 2007; Bramoullé et al., 2009; Liu and Lee, 2010; Calvó-Armengol et al., 2009; Lin, 2010; Lee et al., 2010; Patacchini and Zenou, 2012). The common strategy is to exploit

[10] The residential building concides with the residential house in the majority of the cases (more than 75% of the students live in semidetached or detached single family houses). The results remain largely unchanged if we exclude individuals living in apartment buildings.

[11] Urban areas mainly indicate residential only areas, whereas nonurban areas includes rural, suburban, mostly retail and mostly industrial areas.

[12] This is common when working with AddHealth data. The representativeness of the sample is preserved.

[13] In the spatial econometrics literature, model (5) is the so-called *spatial lag model* or *mixed-regressive spatial autoregressive model* (Anselin, 1988) with the addition of a network-specific component of the error term. Once the variables are transformed in deviations from the network-specific means, an IV or Maximum Likelihood approach (see, e.g. Anselin, 1988) allows us to estimate jointly $\hat{\beta}, \hat{\gamma}$, and $\hat{\phi}$.

Table 1
Description of data (3908 individuals, 359 networks).

	Variable definition	Mean	St. dev.
Residential neighborhood variables			
House well-kept	Interviewer response to the question "How well kept is the building in which the respondent lives", coded as 1 = very poorly kept (needs major repairs), 2 = poorly kept (needs minor repairs), 3 = fairly well kept (needs cosmetic work), 4 = very well kept	2.79	1.92
Peers' house well-kept	Average value among friends	2.59	1.69
Residential area urban	Interviewer's description of the immediate area or street (one block, both sides) where the respondent lives, coded as a dummy taking value 1 if the area is urban-residential only and 0 otherwise (i.e. if the area is rural, suburban, mostly retail, mostly industrial or other type)	0.43	0.49
Individual characteristics			
Female	Dummy variable taking value one if the respondent is female	0.47	0.50
Nonwhite	Race dummy. "White" is the reference group	0.21	0.41
Age	Grade of student in the current year	9.34	3.02
Mathematics score	Score in mathematics at the most recent grading period, coded as 1 = D or lower, 2 = C, 3 = B, 4 = A	2.12	1.20
Religion practice	Response to the question: "In the past 12 months, how often did you attend religious services", coded as 1 = never, 2 = less than once a month, 3 = once a month or more, but less than once a week, 4 = once a week or more	2.89	1.16
Household size	Number of people living in the household	3.55	1.91
Family income	Total family income in thousands of dollars, before taxes. It includes income of everybody in the household, and income from welfare benefits, dividends, and all other sources	49.40	52.77
Single parent family	Dummy taking value one if the respondent lives in a household with only one parent (both biological and non-biological)	0.17	0.38
Mother working	Dummy taking value one if the mother works for pay	0.67	0.47
Parent education	Schooling level of the (biological or non-biological) parent who is living with the child, distinguishing between "never went to school", "not graduate from high school", "high school graduate", "graduated from college or a university", "professional training beyond a four-year college", coded as 1–5. We considering only the education of the father if both parents are in the household	3.65	2.23
Parent occupation manager	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "Doesn't work without being disables" is the reference group	0.10	0.30
Parent occupation professional/technical	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "Doesn't work without being disables" is the reference group	0.08	0.27
Parent occupation office or sales worker	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "Doesn't work without being disables" is the reference group	0.25	0.43
Parent occupation manual	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "Doesn't work without being disables" is the reference group	0.27	0.44
Parent occupation other	Parent occupation dummies. Closest description of the job of (biological or non-biological) parent that is living with the child is manager. If both parents are in the household, the occupation of the father is considered. "Doesn't work without being disables" is the reference group	0.09	0.29
Peers' characteristics	Average values of all the individual characteristics among the respondent's friends (contextual effects)		
Network characteristics			
Network size	Number of network members	42.44	66.17
Number of social contacts	Number of friends	2.55	2.68

the architecture of network contacts to disentangle endogenous from exogenous (contextual) effects.[14] Network fixed effects act as a remedy for the selection bias that originates from the possible sorting of individuals with similar unobserved characteristics into a network. The underlying assumption is that such unobserved characteristics are common to the individuals within each network. This is reasonable in our case study where the networks are quite small (see Section 3.1, Fig. 2). However, the estimation of model (5) might still be flawed because of the presence of peer-group specific (rather than network specific) unobservable factors affecting both individual and peer behavior. For example, a correlation between the individual and the peer-school performance may be due to an exposure to common factors (e.g. having good teachers) rather than to social interactions. The way in which this has been addressed in the literature is to exploit the architecture of network contacts to construct valid IVs for the endogenous effect. Since peer groups are individual specific in social networks, the characteristics of indirect friends are natural candidates. Model (5) can then be estimated using an instrumental variable approach where the behavior of friends is instrumented with the characteristics of indirect peers, i.e. of friends of friends. The intuition is as follow. Consider a simple network with three individuals A, B and C (see Fig. 3). A and B play piano together and B and C swim together, but A and C have never met. Then, the only way C could influence A's behavior is through B. The characteristics of C are thus good instruments for the effect of the behavior of B on A because they certainly influence the behavior of B but do not directly influence the behavior of A.

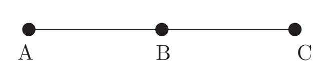

Fig. 3. Identification through intransitive triads.

The application of this strategy in our context is based on the key premise that the children's social contacts in the neighborhood are a good approximation of their parents' social contacts. Indeed, the decisions about home repairs, maintenance and upkeep are taken by the parents. The concern is that parents of friends of

[14] It is well-known that endogenous and contextual effects cannot be separately identified in a linear-in-means model due to *the reflection problem*, first formulated by Manski (1993). In social networks data, *the intransitivity in social connections* provides an exclusion restriction to identify endogenous and contextual effects (see, e.g. Bramoullé et al., 2009).

friends may be friends (even if their children are not friends) if parents sort in social circles according, for example, to their income, education, occupation. If this is the case (and those common factors are not controlled for), then peer effects can simply capture the effects of these common characteristics. In the following section, we provide evidence showing that this is not the case. In our context where networks are quite small, unobserved social circle characteristics are in fact likely to be captured by network fixed effects.

3.2.1. Evidence on the identification strategy

Following Bifulco et al. (2011), we investigate the validity of our identification strategy by performing two exercises.

First, we produce a table of "balancing tests" for our instruments. If the instruments are unrelated to a variety of pre-determined student characteristics, controlling for the fixed effects we use in the regression analysis, then the analysis provides evidence supporting the absence of sorting along observable dimensions. Further, if one uses the degree of selection on observables as a guide to the degree of selection on unobservables as suggested by Altonji et al. (2005), null results on the balancing tests would support the assumption that our model specification identifies variations unrelated to unobservables that determine outcomes. The results are contained in Table 2. We report the correlations when using family background characteristics, which are the most troubling factors in our case. Table 2 shows that none of the estimated coefficients appear to be significantly different from zero.

The second exercise consists in running placebo tests in which we replace the actual peers with simulated peers. We consider different types of simulated peers, that is we draw at random peers within the same family income, or parental education, or parental occupation, or cohort. More specifically, for each individual we draw at random a number of friends equal to the nominated one of a given type, i.e. belonging to a given social circle as defined by parental education, occupation, etc. If our estimates simply captures unobserved social circle characteristics, then these regressions should continue to show a statistical significant effect. If, on the other hand, our strategy is valid, then we should not find any effect of simulated peers behavior on one's own outcomes in these placebo regressions. The results are contained in Table 3. No evidence of significant correlation is revealed. Thus, this evidence provides further confirmation that our strategy, which is based on a large set of controls about individuals, peers and peers' parents, network fixed effects and IVs, is able to cope with sorting issues that could confound our estimates. In the next section, we will describe in greater detail the building blocks of our identification strategy while presenting our estimation results.[15]

4. Empirical results

We present the estimation results of model (5) using a wide range of models, with increasing sets of controls and various estimation strategies. They are reported in Table 4. The last row of this table shows the percentage of the variance which is explained by peer effects. We begin in column (1) by showing the raw correlation between individual and peers' behavior. It appears that such a correlation is quite high (about 40% in terms of explained variance). This suggests that friends' parents housing consumption is not a poor proxy for parents' friends' housing consumption. When conditioning on individual characteristics the percentage of the variance explained falls to about 25%. However, a correlation between individual and peers' behavior may be due to similar individual and peer characteristics, rather than to peer effects (i.e. endogenous effects). The uniqueness of our data where both respondents and friends are interviewed allows us to control for peers' characteristics, thus disentangling the effects of endogenous from exogenous effects. Observe that peers' characteristics also include the characteristics of the parents of the friends (in particular peers' family income, parental education, race). Column (3) shows that about 8% of effects attributed to peers' behavior is in fact due to peers' characteristics – the percentage of explained variance falls from 25% to 17% approximately. A remaining concern relates to the presence of unobserved factors. The observed characteristics of peers may not capture all the nuances of the social environment. There may be two types of unobservables: (i) unobservables that are common to all individuals in a (broadly defined) social circle and/or (ii) unobservables that are instead peer-group specific, i.e. shared by nominated friends only. The bi-dimensional nature of network data (we observe individuals over networks) allows us to control for the presence of unobserved factors of type (i) by including network fixed effects.[16] By doing so, we purge our estimates from the effects of unobserved factors that are common among directly and indirectly related individuals. Column (4) reports the results when network fixed effects are included in the model. It appears that the percentage of explained variance falls by about 10%, thus revealing the presence of important unobserved factors in each individual's social circle. The presence of type (ii) unobservables can instead be addressed by exploiting the architecture of networks through an IV strategy, as described in the previous section. Indeed, going back to our simple example in Fig. 3, it appears that individual C is not influenced by the possible unobserved factors affecting the peer group composed by A and B. The characteristics of indirect friends thus provide valid instrumental variables for the endogenous effect in presence of unobserved correlated effects between peers. First stage results are reported in Table 5. The first stage *F*-test of about 17 shows the relevance of the instruments. The second stage IV estimates are reported in column (5) of Table 4.[17] It appears that the percentage of variance captured by peers' behavior further drops from about 7–3.5%. This is a sizable reduction, but smaller than the one registered when including network fixed effects. In our context where networks are quite small (see Section 3.1, Fig. 2), it is in fact reasonable to think that most unobserved effects are likely to be captured at the network level.[18]

Let us now focus our attention on our preferred estimates (column (5)). The results show that the estimated coefficient of ϕ, which measures the *taste for conformity*, is statistically significant and has a positive sign. This evidence thus supports our theoretical framework predicting a relevant role of peers and conformity to peers' behavior in shaping housing-related decisions. Quantitatively, a standard deviation increase in the demand for housing

[15] A different concern about the use of the kids-parents proxy is the possibility that some social contacts relevant to the parents' decisions are missing, that is missing nodes rather than links in Fig. 3. This would not invalidate our identification strategy as the intransitive triad with the missing node would be broken and thus not used for identification. However, it would affect the magnitude of the estimated effects. Helmers and Patnam (2014) and Liu et al. (2013) investigate the bias of the IV estimator when misspecification of the social network structure is due to data missing at random. Their Monte Carlo experiments show that sampling induces a consistent downward bias in the estimates at all sample sizes. With downward bias, our analysis is likely to estimate a lower bound for the importance of social interactions in the demand for housing quality.

[16] This is a pseudo panel data within-group strategy, where the group mean (here network mean) is removed from each individual observation.

[17] Reduced form estimates are reported in Table A1 in Appendix A.

[18] When networks contains instead a large number of agents, the use of network fixed effects is certainly not an effective strategy as it is not reasonable to think that the unobserved factors are only variables which are common to all members. For example, networks of transactions in the housing market that involve a large number of properties may contain different types of unobservables for different types of properties, even though all the properties belong to the same network of buyers and sellers. In this cases, the use of network fixed effects would not eliminate endogeneity problems.

Table 2
Balancing tests.

X	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Instrument $\hat{G}2$ X											
(1) Parent education	0.0009	0.0118	0.0561	0.0376	0.0058	0.0049	0.0074	−0.1023	0.1331	0.1733	0.0121
	(0.0038)	(0.0088)	(0.1102)	(0.1098)	(0.0596)	(0.0597)	(0.0597)	(0.1682)	(0.1651)	(0.2348)	(0.1041)
(2) Single parent family	0.0031	0.0087	0.0186	0.0048	0.0014	0.0008	0.0025	−0.1063	0.1170	0.1555	−0.0156
	(0.0050)	(0.0114)	(0.0886)	(0.0883)	(0.0475)	(0.0475)	(0.0475)	(0.1338)	(0.1642)	(0.2338)	(0.0704)
(3) Mother working	−0.0004	−0.0007	0.0125	0.0030	−0.0024	−0.0025	−0.0019	−0.0682	0.0624	0.0817	−0.0265
	(0.0005)	(0.0011)	(0.0601)	(0.0598)	(0.0331)	(0.0331)	(0.0331)	(0.0934)	(0.1067)	(0.1503)	(0.1165)
(4) Parent occ. manager	−0.0069	−0.0094	−0.0149	−0.0196	−0.0110	−0.0111	−0.0103	−0.0663	0.0434	0.0639	−0.0453
	(0.0113)	(0.0260)	(0.0382)	(0.0380)	(0.0210)	(0.0210)	(0.0210)	(0.0593)	(0.0999)	(0.1418)	(0.0853)
(5) Parent occ. prof. tech.	0.0027	−0.0082	0.0184	0.0209	0.0013	0.0016	−0.0017	0.0056	0.0210	0.0116	0.0266
	(0.0073)	(0.0169)	(0.0244)	(0.0248)	(0.0132)	(0.0132)	(0.0130)	(0.0372)	(0.0212)	(0.0336)	(0.0342)
(6) Parent occ. manual	−0.0007	0.0066	−0.0099	−0.0217	0.0034	0.0038	0.0011	−0.0785	−0.0122	−0.0436	−0.0652
	(0.0160)	(0.0369)	(0.0400)	(0.0399)	(0.0183)	(0.0184)	(0.0183)	(0.0520)	(0.1237)	(0.1597)	(0.0820)
(7) Parent occ. sales	−0.0045	−0.0147	0.0381	0.0300	0.0275	0.0274	0.0265	0.0305	0.0286	−0.0123	0.0034
	(0.0089)	(0.0206)	(0.0419)	(0.0417)	(0.0170)	(0.0170)	(0.0169)	(0.0478)	(0.1265)	(0.1841)	(0.0085)
(8) Parent occ. other	−0.0010	−0.0039	0.0119	−0.0050	−0.0150	−0.0253	−0.0224	−0.0033	0.0248	0.0009	0.0310
	(0.0050)	(0.0123)	(0.0200)	(0.0088)	(0.0199)	(0.0535)	(0.0507)	(0.0050)	(0.0247)	(0.0150)	(0.0374)
(9) Family income	−0.0025	−0.0228	−0.0083	−0.0118	−0.0053	−0.0029	−0.0047	−0.0313	0.0148	0.0027	0.0181
	(0.0079)	(0.0182)	(0.0239)	(0.0238)	(0.0129)	(0.0141)	(0.0129)	(0.0367)	(0.0524)	(0.0682)	(0.0453)
(10) Household size	0.0017	0.0032	0.0053	0.0027	−0.0003	−0.0003	0.0002	−0.0006	0.0016	0.0018	0.0355
	(0.0013)	(0.0030)	(0.0060)	(0.0030)	(0.0017)	(0.0017)	(0.0016)	(0.0047)	(0.0026)	(0.0036)	(0.0298)
(11) Age	0.0014	0.0029	0.0162	0.0154	0.0075	0.0075	0.0068	0.0229	0.0033	−0.0032	0.0419
	(0.0027)	(0.0062)	(0.0143)	(0.0142)	(0.0075)	(0.0075)	(0.0075)	(0.0212)	(0.0356)	(0.0503)	(0.0500)
Network fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All regressions include other individual and peers' characteristics, which are listed in Table 1. The instruments $\hat{\mathbf{G}}\mathbf{2}\ \mathbf{X}$ are the characteristics X of the peers of peers listed in the first column. Precise definitions of variables are in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

Table 3
Placebo tests simulated peer effects.

Dep. Var. House Well-Kept	2SLS			
Simulated peers by	Family income	Parental education	Parental occupation	Student age
Peer effects ($\widehat{\phi}$)	0.0131	0.0172	0.0089	0.0098
	(0.0201)	(0.0288)	(0.0101)	(0.0118)
Individual characteristics	Yes	Yes	Yes	Yes
Peers' characteristics	Yes	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes	Yes
N. obs.	3908	3908	3908	3908
N. networks	359	359	359	359

Notes: For each individual, we draw at random a number of friends equal to the nominated one of a specific type, repeat 5000 times, and report mean estimates and standard errors of the empirical distributions. Control variables are those listed in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

Table 4
Estimation results.

Dep. Var. House Well-Kept	OLS				2SLS
	(1)	(2)	(3)	(4)	(5)
Peer effects ($\widehat{\phi}$)	0.3212***	0.2808	0.2124***	0.1785***	0.1196**
	(0.1069)	(0.0878)	(0.0771)	(0.0654)	(0.0577)
Individual characteristics	No	Yes	Yes	Yes	Yes
Peers' characteristics	No	No	Yes	Yes	Yes
Network fixed effects	No	No	No	Yes	Yes
N. obs.	3908	3908	3908	3908	3908
N. networks	359	359	359	359	359
% Variance explained by peer effects	39.1	25.2	17.4	7.1	3.5

Notes: Estimated standardized coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Control variables are those listed in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

quality of the peers translates into a roughly 12% increase of a standard deviation in the individual demand for housing quality. In order to better understand the magnitude of the effects, we provide a money metric evaluation of the individual response to peers' house quality variations. According to the American Housing Survey (AHS) data for the year 1993, a one standard deviation increase in house quality (about 0.49) raises house prices by 8.9%, which translates into about 0.5% with our data.[19] Therefore, our results reveal that in a group of two friends, if the friend invests in renovations so that her/his house changes from very poorly kept (needs major repairs) to very well kept (i.e., a roughly two standard deviation increase in house quality), then the individual house value would increase by roughly 1%. This is a small effect, as it is probably expected, especially given our long list of controls.

In order to further our understanding of the results, we estimate model (5) for individuals living in urban and nonurban areas separately. The results are collected in the last two columns of Table 6. The first column reports the results for the entire sample for comparison (results in column (5) of Table 4). The basic idea of our theoretical model is that agents' behavior in terms of housing quality choices is driven by their desire to reduce the discrepancy between their own house quality and that of their reference group (i.e. their best friends). Social interactions are the law of motion of this mechanism. If this is the behavioral mechanism at work, then we should observe in our data that peer effects are stronger in urban areas, i.e. where social interactions are more intense. Indeed, people living in urban areas have richer social opportunities than people living in nonurban areas, and they may get more benefit from conforming to the standards of their social group. It appears that peer effects are stronger for individuals living in urban areas. They are not even statistically significant for individuals living in nonurban areas. Hence, this evidence lends further support to the theoretical mechanism presented in Section 3.[20]

Our results, however, should be interpreted with caution. There are a variety of utility functions (or a variety of social processes) that can be consistent with our evidence, and it is extremely difficult to discriminate between the different mechanisms empirically (see Ghiglino and Goyal, 2010) for a discussion and a taxonomy of theoretical models).[21] Our analysis contributes to the empirics of those models by providing evidence of their main common feature – the decision of agents to consume a good cannot adequately be explained by the intrinsic utility derived from consuming it, while their utility is sensitive to the consumption of their neighbors.

5. Robustness checks

We perform two robustness checks. The first investigates the sensitivity of our estimation results to misspecification of the social network structure – mismeasurement of network links; the second relates to the possible presence of individual- (rather than peer-group- or network-) level unobserved factors affecting both network formation and outcome decisions.

5.1. Undirected versus directed networks

Our theoretical model and consequently our empirical investigation assume, so far, that friendship relationships are symmetric, i.e. $g_{ij} = g_{ji}$. In this section, we check how sensitive our results are to such an assumption, i.e. to a possible measurement error in the definition of the peer group. Indeed, our data make it possible to know exactly who nominates whom in a network and we find that 12% of relationships in our dataset are not reciprocal. Instead of constructing undirected network, we will now focus on the analysis of directed networks.

In the language of graph theory, in a directed graph, a link has two distinct ends: a head (the end with an arrow) and a tail. Each end is counted separately. The sum of head endpoints count toward the *indegree* and the sum of tail endpoints count toward the *outdegree*. Formally, we denote a link from i to j as $g_{ij} = 1$ if j has nominated i as his/her friend, and $g_{ij} = 0$, otherwise. The indegree of student i, denoted by g_i^+, is the number of nominations student i receives from other students, that is $g_i^+ = \sum_j g_{ij}$. The outdegree of student i, denoted by g_i^-, is the number of friends student i nominates, that is $g_i^- = \sum_j g_{ji}$. We can thus construct two types of directed networks, one based on indegrees and the other based on outdegrees. Observe that, by definition, while in undirected networks the adjacency matrix $\mathbf{G} = [g_{ij}]$ is *symmetric*, in directed networks it is *asymmetric*.

From a theoretical point of view, the symmetry of $\mathbf{G}$ does not play any explicit role and thus all the results remain valid with a non-symmetric $\mathbf{G}$ (Patacchini and Zenou, 2012). Turning to the empirical analysis, we report in Table 7 the results of the estimation of model (5) when the directed nature of the network data is taken into account. It appears that our results are only minimally affected in both tables. The estimated peer effects remain positive and statistically significant.

5.2. Endogenous network formation

Our identification assumption is based on the idea that the matrix of network links is exogenous conditional on individual characteristics, peer characteristics and network fixed effect. In other words, we assume that our large list of controls captures the characteristics that drive the sorting of individuals into groups (age, gender, race, parental education and occupation, etc.) and that any remaining unobserved characteristics are captured at the network level. Network fixed effects are thus a remedy for the selection bias that originates from the possible sorting of individuals with similar unobserved characteristics into a network. The underlying assumption is that such unobserved characteristics are common to the individuals within each network. This is reasonable in our case study where the networks are quite small (see Section 3.1, Fig. 2).

However, if there are student-level unobservables that affect both the propensity to engage in home upkeep and the likelihood to form friendship links, then this strategy fails. Recently, Goldsmith-Pinkham and Imbens (2013) highlight the fact that endogeneity of this sort can be tested. In this section, we borrow from this paper and document the extent to which, in our case, network structure is exogenous conditional on network fixed effects.

In order to understand the problem, let us include in model (5) an individual-level *unobserved characteristic,* ν.[22]

[19] The house quality questions asked to the interviewer in the AHS are more detailed and precise than the *house well-kept* index of our dataset. We mapped the AHS questions into the 1–4 categories of our index as follows. We code *house well-kept* equal to 4 if the interviewer does not mention needed repairs, equal to 1 if the mentioned repairs are major (hole in roof, foundation crumbling or with open crack, broken windows, etc.), equal to 2 if the mentioned repairs are minor (broken steps, railings not firmly attached, etc.), and equal to 3 for the residual category. We then run a simple hedonic regression – (log) house price on this indicator of house quality.

[20] A complete list of estimation results, including all the estimated effects of all individual controls and peers' characteristics is reported in Table A2 in Appendix A.

[21] Corneo and Jeanne (1997) show that conformity preferences under conspicuous consumption may result in an upward sloping demand of the conspicuous good.

[22] The model can also include more than one unobserved characteristic (i.e. one can consider a vector $(\nu_{1,i}, \ldots, \nu_{n,i})'$ of unobserved characteristics).

Table 5
2SLS estimation first-stage results.

Dep. Var. Peers' House Well-Kept	Own characteristics	Peers' characteristics	Peers of peers' characteristics (IVs)
Female	0.0086	−0.0468	−0.0458
	(0.0946)	(0.0734)	(0.0378)
Nonwhite	−0.0008	−0.2741*	−0.2404***
	(0.1932)	(0.1573)	(0.0635)
Age	0.0094*	0.0522***	0.0190***
	(0.0055)	(0.0163)	(0.0073)
Religion practice	−0.0019	0.0001	0.0015***
	(0.0011)	(0.0006)	(0.0004)
Mathematics score	0.1320*	0.0800**	0.0407
	(0.0717)	(0.0393)	(0.1040)
Family income(*1000)	0.0273*	0.0048**	0.0527***
	(0.0159)	(0.0022)	(0.0148)
Parent education	0.0476	0.1178***	0.0081**
	(0.0312)	(0.0224)	(0.0041)
Single parent family	−0.0108	−0.0455**	−0.0067
	(0.1290)	(0.0228)	(0.0408)
Mother working	0.1089	−0.0647	0.1238*
	(0.2403)	(0.1494)	(0.0700)
Parent occ. manager	0.1263	0.3080**	0.1715**
	(0.2899)	(0.1279)	(0.0816)
Parent occ. prof. tech.	0.0822	0.1854	−0.1343
	(0.2553)	(0.1804)	(0.1007)
Parent occ. manual	−0.0174	0.0963	−0.0164
	(0.2384)	(0.1720)	(0.0934)
Parent occ. sales	−0.0519	0.2413	0.0027
	(0.2470)	(0.1737)	(0.0938)
Parent occ. other	−0.0019	−0.0053	0.0110
	(0.1215)	(0.1212)	(0.1520)
Household size	−0.0376	−0.0289**	−0.0547***
	(0.0297)	(0.0127)	(0.0095)
F statistic			17.10

Notes: Estimated coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Network fixed effects are included. Precise definitions of variables are in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

Table 6
Estimation results urban versus nonurban areas.

Dep. Var. House Well-Kept	2SLS		
	All sample	Urban areas	Nonurban areas
Peer effects ($\widehat{\phi}$)	0.1196**	0.1665***	0.1041
	(0.0577)	(0.0596)	(0.0790)
First stage *F* statistic	17.10	19.12	10.93
Individual characteristics	Yes	Yes	Yes
Peers' characteristics	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes
N. obs.	3908	1680	2228
N. networks	359	359	359

Notes: Estimated standardized coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Control variables are those listed in Table 1.
* Statistical significance at the 10% level.
** Statistical significance at the 5% level.
*** Statistical significance at the 1% level.

Table 7
Alternative definition of network links.

Dep. Var. House Well-Kept	2SLS		
	All sample	Urban areas	Nonurban areas
Directed networks (outdegree)			
Peer effects ($\widehat{\phi}$)	0.1349**	0.1860***	0.1178
	(0.0643)	(0.0669)	(0.0870)
First stage *F* statistic	24.53	23.88	14.45
Directed networks (indegree)			
Peer effects ($\widehat{\phi}$)	0.1365**	0.1658**	0.0989
	(0.0605)	(0.0755)	(0.0857)
First stage *F* statistic	21.14	22.22	14.22
Individual characteristics	Yes	Yes	Yes
Peers' characteristics	Yes	Yes	Yes
Network fixed effects	Yes	Yes	Yes
N. obs.	3908	1680	2228
N. networks	359	359	359

Notes: Estimated standardized coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Control variables are those listed in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

$$y_{i,\kappa} = \phi \frac{1}{g_{i,k}} \sum_{j=1}^{n_\kappa} g_{ij,k} y_{j,k} + \sum_{m=1}^{M} \beta_1^m x_{i,\kappa}^m + \frac{1}{g_{i,k}} \sum_{m=1}^{M} \sum_{j=1}^{n_\kappa} \theta_m g_{ij,k} x_{j,\kappa}^m + \underbrace{\eta_k + \delta v_i + \varepsilon_{i,k}}_{e_{i,k}} \qquad (6)$$

Let us also consider a network formation model where the variables that explain the links between students i and j belonging to network κ, i.e. $g_{ij,\kappa}$, are the distances between them in terms of observed and unobserved characteristics:

$$g_{ij,\kappa} = \alpha + \sum_{m=1}^{M} \delta_m |x_{i,\kappa}^m - x_{j,\kappa}^m| + \theta |v_{i,\kappa} - v_{j,\kappa}| + \eta_\kappa + u_{ij,\kappa} \qquad (7)$$

A testable implication of the presence of unobserved factors affecting both outcome and network formation could be the existence of a statistically significant correlation between residuals of the

Table 8
Endogenous network formation.

Dep. Var.: Probability to form a link	OLS		
	All sample	Urban areas	Nonurban areas
No network fixed effects			
Difference in residuals ($\widehat{\theta}$)	−0.0043**	−0.0059***	−0.0039**
	(0.0018)	(0.0022)	(0.0020)
With network fixed effects			
Difference in residuals ($\widehat{\theta}$)	−0.0012	−0.0015	−0.0011
	(0.0011)	(0.0014)	(0.0012)

Notes: Observations are all pairwise combinations of students with complete data on covariates ($7{,}634{,}278 = N * (N-1)/2$ observations where $N = 3908$). Linear probability model (6) estimated via least squares. Control variables are those listed in Table 1.
*,**,*** Denote statistical significance at the 10, 5 and 1% level.

outcome Eq. (6) and the probability of observing a link (model (7)). One can thus replace $|v_{i,\kappa} - v_{j,\kappa}|$ in model (7) with the difference in the residuals $|\hat{e}_{i,\kappa} - \hat{e}_{j,\kappa}|$ from model (6) and estimate model (7) where the dependent variable is a dummy variable that takes a value of 1 if there is a link between *i* and *j* and 0 otherwise. Evidence supporting network exogeneity would be the finding of $\hat{\theta} = 0$. [23] Our results are contained in Table 8. They reveal that if we do not include network fixed effect, then we have a statistically significant correlation between the probability to form a link and unobserved similarity in pairs. The sign is negative and is in line with an homophily behavior in the unobserved characteristics, i.e. that the closer two individuals are in terms of unobservable characteristics, the higher the probability that they are friends. When we include network fixed effects, the correlation disappears, i.e. $\hat{\theta} = 0$. Therefore, conditional on the large set of controls provided by the AddHealth data, peer characteristics, and network fixed effects, we find no evidence of additional individual-level unobserved characteristics that may bias our results.

6. Concluding remarks

Housing is a composite commodity that satisfies dwelling needs, but it also provides other intangibles such as security, access to jobs and social status. The diversity in individual preferences along these different dimensions leads to a large heterogeneity in the revealed behavior, that is, the demand for housing quality. An understanding of the importance of non-market factors in housing related decisions is crucial to design more effective housing programs. There is a vast literature providing estimates of the distributional impact of providing subsidies to owner-occupied housing. There seems to be a widespread consensus that the income tax treatment of owner-occupied housing has substantially increased the consumption of housing services (see, e.g., Poterba, 1992; Poterba and Sinai, 2008, 2011; Albouy and Hanson, 2014; Hanson, 2014). Our evidence on the existence of peer effects in the demand for housing quality suggests that these quantifications could potentially be different.

Although our results are not conclusive on the determinants of nonfunctional demand for housing services, they suggest that social comparisons originated in one's own residential neighborhood are important in shaping the demand for housing quality. There is little doubt that individuals' satisfaction with a given behavior also depends on what one achieves in relative terms, i.e. compared to other individuals. A "conspicuous consumption" (Veblen, 1899) or a "bandwagon effect" are cases where the commodity serves the purpose of social belonging or status definition (Leibenstein, 1950). For most individuals, housing is the largest consumption and investment item of their life. A discrepancy between current and desired housing needs may create stress or dissatisfaction through migration or remodelling and thus distraction of resources from alternative investments such as education. Individuals' subjective evaluation of their housing forms the basis of demand for public action. This suggests that an effective policy should take into account the group interactions it stimulates.

[23] On the other hand, homophily behavior in the unobserved characteristics would imply $\theta < 0$, i.e. that the closer two individuals are in terms of unobservable characteristics, the higher is the probability that they are friends. Dissortative behavior in the unobserved characteristics instead implies that $\theta > 0$, i.e. that the closer two individuals are in terms of unobservable characteristics, the lower is the probability that they are friends.

Table A1
2SLS estimation reduced form.

Dep. Var. House Well-Kept Own characteristics		Peers' characteristics	Peers of peers' characteristics
Female	0.0040	0.0120	0.0145
	(0.0594)	(0.0273)	(0.0303)
Nonwhite	−0.0055	−0.0182**	−0.0311**
	(0.0109)	(0.0099)	(0.0151)
Age	0.0133**	0.0755**	0.0219***
	(0.0069)	(0.0360)	(0.0079)
Religion practice	0.0101	0.0004	0.0011*
	(0.0111)	(0.0010)	(0.006)
Mathematics score	0.0530	0.0710	0.0240
	(0.0671)	(0.0739)	(0.0410)
Family income(*1000)	0.0427**	0.0509**	0.0275**
	(0.0210)	(0.0252)	(0.0128)
Parent education	0.0455***	0.0718***	0.0110**
	(0.0125)	(0.0204)	(0.0050)
Single parent family	0.0028	−0.0149*	−0.0045
	(0.0129)	(0.0088)	(0.0174)
Mother working	0.0889**	0.0655	0.0888
	(0.0433)	(0.0714)	(0.0870)
Parent occ. manager	0.2635*	0.2308*	0.2701**
	(0.1541)	(0.1358)	(0.1552)
Parent occ. prof. tech.	0.1082	0.1338	−0.1553
	(0.2445)	(0.3107)	(0.2020)
Parent occ. manual	−0.0775	−0.1245	−0.0102
	(0.2222)	(0.1788)	(0.10998)
Parent occ. sales	−0.0579	−0.0879	−0.0269
	(0.2035)	(0.1072)	(0.1934)
Parent occ. other	0.0114	−0.0075	−0.0099
	(0.0772)	(0.1013)	(0.0877)
Household size	−0.0075**	−0.0112**	−0.0054**
	(0.0036)	(0.0051)	(0.0022)
Network fixed effects	Yes		
N. obs.	3908		
N. networks	359		

Notes: Estimated coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Precise definitions of variables are in Table 1.
* Statistical significance at the 10% level.
** Statistical significance at the 5% level.
*** Statistical significance at the 1% level.

Appendix A. Additional results

See Tables A1 and A2.

Table A2
2SLS estimation results complete list of controls.

Dep. Var. House Well-Kept	2SLS		
	All sample	Urban areas	Nonurban areas
Peer effects ($\widehat{\phi}$)	0.1196**	0.1665***	0.1041
	(0.0577)	(0.0596)	(0.0790)
Individual characteristics			
Female	−0.0960	−0.1002	−0.1023
	(0.1682)	(0.1682)	(0.1682)
Nonwhite	−0.1088**	−0.1247**	−0.1063
	(0.0459)	(0.0608)	(0.1338)
Age	0.0666***	0.0680**	0.0682*
	(0.0300)	(0.0334)	(0.0378)
Religion practice	0.0644	0.0759*	−0.0266
	(0.0593)	(0.0419)	(0.0793)
Parent education	0.0910***	0.0814***	0.0569**
	(0.0360)	(0.0265)	(0.0270)
Mathematics score	0.1604	0.1510	0.1505
	(0.1891)	(0.1999)	(0.2200)
Single parent family	−0.0918*	−0.0866*	−0.0785
	(0.0500)	(0.0501)	(0.0520)
Mother working	0.0864*	0.0879*	0.0905*
	(0.0476)	(0.0476)	(0.0478)
Parent occ. manager	0.0328	0.0344	0.0313
	(0.0363)	(0.0363)	(0.0367)
Parent occ. prof. tech.	0.0010	−0.0005	−0.0006
	(0.0046)	(0.0047)	(0.0047)
Parent occ. manual	0.0212	0.0229	0.0229
	(0.0212)	(0.0212)	(0.0212)
Parent occ. sales	−0.0361	−0.0371	−0.0366
	(0.0364)	(0.0364)	(0.0364)
Parent occ. other	0.0149	0.0176	0.0202
	(0.0201)	(0.0269)	(0.0312)
Family income(*1000)	0.0284***	0.0291***	0.0279**
	(0.0101)	(0.0122)	(0.0129)
Household size	−0.0460	−0.0464	−0.0461
	(0.0353)	(0.0353)	(0.0353)
Peers'characteristics			
Female	0.0009	0.0039	0.0119
	(0.2150)	(0.2123)	(0.2040)
Nonwhite	−0.0283	−0.0297	−0.0163
	(0.0795)	(0.0771)	(0.0704)
Age	0.0296**	0.0312	0.0266
	(0.0134)	(0.0215)	(0.0265)
Religion practice	0.0509	0.0508	0.0453
	(0.0880)	(0.0868)	(0.0835)
Parent education	0.0201**	0.0492**	0.0273**
	(0.0103)	(0.0232)	(0.0126)
Mathematics score	0.1320	0.1325	0.1692**
	(0.0935)	(0.0882)	(0.0820)
Single parent family	−0.1603*	−0.1520*	−0.1341
	(0.0888)	(0.0875)	(0.0858)
Mother working	−0.0324	−0.0336	−0.0480
	(0.0507)	(0.0503)	(0.0474)
Parent occ. manager	0.0216**	0.0180	0.0181
	(0.0105)	(0.0256)	(0.0653)
Parent occ. prof. tech.	0.0353	0.0368	0.0355
	(0.0309)	(0.0310)	(0.0298)
Parent occ. manual	0.0463	0.0421	0.0419
	(0.0530)	(0.0523)	(0.0500)
Parent occ. sales	0.0228	0.0194	0.0187
	(0.0474)	(0.0465)	(0.0444)
Parent occ. other	0.0042	0.0106	0.0087
	(0.0099)	(0.0112)	(0.0144)
Family income(*1000)	0.0381**	0.0419**	0.0410*
	(0.0179)	(0.0206)	(0.0226)
Household size	−0.0135	−0.0130	−0.0105
	(0.0734)	(0.0717)	(0.0688)
Network fixed effects	Yes	Yes	Yes
N. obs.	3908	1680	2228
N. networks	359	359	359

Notes: Estimated standardized coefficients and standard errors (in parentheses) are reported. Regressions include weights to control for the AddHealth survey design. Control variables are those listed in Table 1.
* Statistical significance at the 10% level.
** Statistical significance at the 5% level.
*** Statistical significance at the 1% level.

References

Akerlof, G.A., 1980. A theory of social custom of which unemployment may be one consequence. Quarterly Journal of Economics 94, 749–775.
Akerlof, G.A., 1997. Social distance and social decisions. Econometrica 65, 1005–1027.
Albouy, D., Hanson, A.R., 2014. Tax benefits to housing and inefficiency in location and housing consumption. Tax Policy and the Economy (forthcoming).
Altonji, J.G., Elder, Todd E., Tabor, C.R., 2005. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. Journal of Political Economy 113, 151–184.
Anselin, L., 1988. Spatial Econometrics, Methods and Models. Kluwer, Dordrech.
Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), 2011. Handbook of Social Economics. Elsevier Science, Amsterdam.
Berman, E., 2000. Sect, subsidy, and sacrifice: an economist's view of ultra-orthodox Jews. The Quarterly Journal of Economics 115 (3), 905–953.
Bernheim, B.D., 1994. A theory of conformity. Journal of Political Economy 102, 841–877.
Bifulco, R., Fletcher, J.M., Ross, S.L., 2011. The effect of classmate characteristics on post-secondary outcomes: evidence from the Add Health. American Economic Journal: Economic Policy 3, 25–53.
Bramoullé, Y., Djebbari, H., Fortin, B., 2009. Identification of peer effects through social networks. Journal of Econometrics 150, 41–55.
Calvó-Armengol, A., Patacchini, E., Zenou, Y., 2009. Peer effects and social networks in education. Review of Economic Studies 76, 1239–1267.
Clark, A.E., Loheac, Y., 2007. It wasn't me, it was them! Social influence in risky behavior by adolescents. Journal of Health Economics 26, 763–784.
Corneo, G., Jeanne, O., 1997. Conspicuous consumption, snobbism and conformism. Journal of Public Economics 66, 55–71.
Del Bello, C., Patacchini, E., Zenou, Y., 2014. Social and Geographical Distance in Peer Effects, unpublished manuscript.
Durlauf, S.E., 2004. Neighborhood effects. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier Science, Amsterdam, pp. 2173–2242.
Fershtman, C., Weiss, Y., 1998. Social rewards, externalities and stable preferences. Journal of Public Economics 70, 53–73.
Ghiglino, C., Goyal, S., 2010. Keeping up with the neighbors: social interaction in a market economy. Journal of the European Economic Association 8, 90–119.
Glaeser, E.L., Scheinkman, J., 2001. Measuring social interactions. In: Durlauf, S.N., Young, H.P. (Eds.), Social Dynamics. MIT Press, Cambridge, pp. 83–132.
Glaeser, E.L., Sacerdote, B., Scheinkman, J., 1996. Crime and social interactions. Quarterly Journal of Economics 111, 508–548.
Goldsmith-Pinkham, P., Imbens, G.W., 2013. Social networks and the identification of peer effects. Journal of Business and Economic Statistics 31, 253–264.
Hanson, A.R. 2014. Limiting the mortgage interest deduction by size of home: effects on user cost across metropolitan areas. Journal of Housing Research (forthcoming).
Helmers, C., Patnam, M., 2014. Does the rotten child spoil his companion? Spatial peer effects among children in rural India. Quantitative Economics 5 (1), 67–121.
Iannaccone, L.R., 1992. Sacrifice and stigma: reducing free-riding in cults, communes, and other collectives. Journal of Political Economy 100 (2), 271–291.
Ioannides, Y.M., 2011. Neighborhood effects and housing. In: Benhabib, J., Bisin, A., Jackson, M.O. (Eds.), Handbook of Social Economics, vol. 1B. Elsevier Science, Amsterdam, pp. 1281–1340.
Ioannides, Y.M., 2012. From Neighborhoods to Nations: The Economics of Social Interactions. Princeton University Press.
Ioannides, Y.M., Loury, L.D., 2004. Job information networks, neighborhood effects, and inequality. Journal of Economic Literature 42 (4), 1056–1093.
Ioannides, Y.M., Topa, G., 2010. Neighborhood effects: accomplishments and looking beyond them. Journal of Regional Science 50, 343–362.
Ioannides, Y.M., Zabel, J.E., 2003. Neighbourhood effects and housing demand. Journal of Applied Econometrics 18 (5), 563–584.
Ioannides, Y.M., Zabel, J.E., 2008. Interactions, neighborhood selection and housing demand. Journal of Urban Economics 63, 229–252.
Jackson, M.O., 2008. Social and Economic Networks. Princeton University Press, Princeton.
Kandel, E., Lazear, E.P., 1992. Peer pressure and partnerships. Journal of Political Economy 100, 801–817.
Lee, L.-F., 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. Journal of Econometrics 140, 333–374.
Lee, L-F., Liu, X., Lin, X., 2010. Specification and estimation of social interaction models with network structures. Econometrics Journal 13, 145–176.
Leibenstein, H., 1950. Bandwagon, snob and Veblen effects in the theory of consumer demand. Quarterly journal of Economics, 183–207.
Lin, X., 2010. Identifying peer effects in student academic achievement by a spatial autoregressive model with group unobservables. Journal of Labor Economics 28, 825–860.
Liu, X., Lee, L-F., 2010. GMM estimation of social interaction models with centrality. Journal of Econometrics 159, 99–115.
Liu, X., Patacchini, E., Rainone, E., 2013. The Allocation of Time in Sleep: a Social Network Model with Sampled Data. CEPR Discussion paper n. 162.
Manski, C.F., 1993. Identification of endogenous effects: the reflection problem. Review of Economic Studies 60, 531–542.

Moffitt, R., 2001. Policy interventions low-level equilibria, and social interactions. In: Durlauf, S., Young, P. (Eds.), Social Dynamics. MIT Press, Cambridge, MA, pp. 45–82.
Morris, E.W., Winter, M., 1975. A theory of family housing adjustment. Journal of Marriage and the Family 37, 79–88.
Morris, E.W., Winter, M., 1978. Housing, Family and Society. John Wiley and Sons, New York.
Patacchini, E., Zenou, Y., 2012. Juvenile delinquency and conformism. Journal of Law, Economic, and Organization 28, 1–31.
Poterba, J., 1992. Taxation and housing: old questions, new answers. American Economic Review 82, 237–242.
Poterba, J., Sinai, T., 2008. Tax expenditures for owner-occupied housing: deductions for property taxes and mortgage interest and the exclusion of imputed rental income. American Economic Review 98, 84–89.
Poterba, J., Sinai, T., 2011. Revenue cost and incentive effects of income tax provisions for owner-occupied housing. National Tax Journal 64 (2), 531–564.
Rosen, H., 1985. Housing subsidies: effects on housing decisions, efficiency and equity. In: Auerbach, A., Feldstein, M. (Eds.), Handbook of Public Economics, vol. 1. North-Holland, Amsterdam, pp. 375–420.
Rossi-Hansberg, E., Sarte, P.-D., Owens, R., 2010. Housing Externalities. Journal of Political Economy 118 (3), 485–535, 06.
Veblen, T., 1899. The Theory of Leisure Class. Modern Library, New York.