DISCUSSION PAPERS IN ECONOMICS

W Pa e N . 23-07

! "#$%&'($&)$*#+,-(,&./$
$0$1#.#2-3,4#5$67-(,-3$08(&2#92#'',+#$: &5#3!

!

!

!

"#$%&$'()!*+,(!
-(,.$/0,1&!'2!3'4'/#5'!6'745$/!

!

891':$/!;<=!;>;?!

!

!

!

!

•

Depar men of Economics

!

!

!

!

Uni ersi of Colorado Bo lder
Bo lder, Colorado 80309

@891':$/!;<=!;>;?=!"#$%&$'()!*+,(!

!

The Cost of Deviation: A Generalized Spatial Autoregressive Model

`90c´0njX ~_j

mH§n@0} ©£J ©·©¨

Abstract

A generalized spatial autoregressive model that bridges the complementarity and the conformity peer efect is proposed. A weight matrix is defned as an adjacency matrix minus a diagonal degree matrix multiplied by a conformity parameter between zero and one. This conformity parameter identifes the relative magnitude of one's complementary beneft and deviation cost from his or her own peer group. The social multiplier efect arises only when the complementary beneft overwhelms the deviation cost, and the threshold for the positive multiplier is lower for more centralized networks. The model is applied to the microfnance data from Karnataka, India. In contrast to the common belief, evidence of strong conformity is found by utilizing multiple dimensions of the rǾ¨ Indi formÄ ! a-d̯ ג ˜ Ǿя

people enjoy extra utility by engaging in the same action with their peer groups. On the other hand, under conformity, people get disutility by deviating from their peers. Although these motives' consequences appear similar, their policy implications are vastly different. For instance, suppose that an organizer of a microfinance program wants to introduce it to a specific village. The program is expected to be more beneficial when people are tied together and encourage each other to set up a plan for the responsible use of money. Then, the organizer may successfully spread the program if she can encourage a core group of people to join it first. As they keep participating in the program, their friends or relatives will find it even more beneficial due to the complementary benefit arising from the already established basis and soon want to join the program altogether. As a result, the organizer can achieve her goal more effectively at a smaller cost. This scenario is an example of what is called the "social multiplier" effect, and this multiplier effect is working in a positive direction. The organizer's job is to carefully choose the target group based on criteria such as network centralities. A pitfall in this picture is that the peer effect is implicitly assumed to be complementarity. If the initially targeted group is under the conformity motive rather than complementarity, the target group's participation will soon dwindle because they would not want to behave differently from the other villagers. Their peers might also be willing to join in conforming with them, but the overall impact will be much lower than in the former case. Under this circumstance, the social multiplier effect is working in a negative direction, and the organizer should significantly increase the size of the target group or give them stronger incentives to compensate for this.

Due to this policy implication, several attempts have been made to distinguish and estimate these motives within the SAR model. Unfortunately, as pointed out by Boucher and Fortin (2016), the two motives are not identified with the ordinary SAR model that uses an adjacency matrix with zero diagonal elements as a weight matrix. This negative identification result is attributed to the fact that the underlying microeconomic foundations for the two motives are observationally equivalent. To address this, the "local-average" model is suggested by (Ushchev and Zenou 2020; Patacchini and Zenou 2012). In these papers, the authors model conformity as an additional utility from the average effort of one's peer group. Since the weight matrix is row-sum normalized, the conformity peer effect can be distinguished from the complementarity in which the additional utility comes from the aggregated effort ("the local-aggregate"). This approach is further sophisticated by Liu, Patacchini, and Zenou (2014). They present a microfoundation for the local-average model and perform the J-test to choose a suitable peer effect for a given dataset. This branch of literature, however, has a limitation that the social multiplier arises even under the conformity motive. A noticeably different approach is taken by Boucher (2016) by using the graph Laplacian as a weight matrix for the pure conformity peer effect. In his model, conformity is described as a disutility from the difference between one's own and peers' actions. This specification is a distinguished feature from the previous literature, where there is no such disutility. As a result, there is no multiplier effect as expected.

This paper departs from the existing literature by considering the two pure types of peer effects as two extrema of a single peer effect model that differ by the relative magnitude of the complementary benefit and the deviation cost. In the generalized spatial autoregressive model (GSAR), a *conformity parameter* is introduced to measure such magnitude. The most straightforward advantage of this approach is that it is a more realistic characterization of the peer effect. Frequently, one's true motive of peer effect is more complicated. Back to the previous example, the villagers may join the program partly because of the benefit of cooperating in the microfinance groups and, at the same time, partly because of peer pressure. The GSAR model can capture such situations properly with the flexible conformity parameter. Another benefit of this flexibility is that no model selection is required. A common approach for researchers who want to study peer effect has been arbitrarily assuming the motive they believe to be true or relying upon model selection for choosing the correct one. There are two issues with these strategies. First, there is no guarantee that one of the pure models will be selected. In other words, when each specification is tested against the other, both or none of the nulls may be rejected. In this case, the researchers are forced to draw another arbitrary conclusion to interpret the result. Even after that, the direction of the social multiplier effect is still unclear because both positive and negative multipliers cannot coexist. Second, the pure conformity model requires considerably larger samples due to the smaller variance caused by using the difference of outcome variables as an explanatory variable. The insignificant estimates from small samples may prohibit a researcher from performing model selection reliably.

To support the GSAR model, a microeconomic foundation is built upon the network utility and the private utility, where the network utility comprises the complementary benefit and the deviation cost. Furthermore, it is shown that the parameters for each component are identified by the peer effect and the conformity parameters in the econometrics model. This flexibility introduces a question on the nature of the social multiplier effect. Because the two motives are now on contiguity, the multiplier must also be a continuous function. To answer this, the concept of the social multiplier is extended to positive and negative multipliers. Under the positive multiplier effect, the complementarity motive dominates, and the aggregated output of individuals will be greater than the sum of isolated ones. On the other hand, with the negative multiplier effect, the deviation cost will be overwhelming, and individuals will tend to move toward their social norm. As a next step, a threshold for the positive multiplier is obtained as a function of both peer effect and conformity parameters. To have a positive multiplier, the peer effect must be sufficiently higher than the conformity. This threshold is characterized by a decreasing function of the sum of squared degrees divided by the sum of degrees, which is parallel to the "tendency to make hubs" centrality of Saberi et al. (2021). The intuition is that it is easier to have a positive multiplier if people are clustered around a small number of individuals.

The GSAR model is extended to the rational expectation model of Lee, Li, and Lin (2014) and applied to the case of a microfinance program in Indian villages collected and studied by Banerjee et al. (2013). Implementing the GSAR model with binary outcome variables and multiple networks, several results distinguished from the original study are drawn. Under the assumption of the simultaneous-move game, the estimation shows that there is a significant peer effect on the equilibrium, and its dominating motive is conformity. Moreover, few dimensions of the social network actually transmit the peer effect. By estimating the conformity parameter separately for each relevant network, it is found that different dimensions of an individual's social network may have different underlying motives, meaning that people have distinctive tendencies to conform depending on their reference groups.

The rest of the paper is organized as follows. Section 2 reviews related literature. Section 3 presents the GSAR model with a microeconomic foundation with the extended social multiplier effect, and the baseline model is extended to the binary outcome model with rational expectation. Section 4 applies the model to the microfinance data of Banerjee et al. (2013), and empirical results are shown. Lastly, section 5 provides concluding remarks.

2 Related Literature

The infection problem of Manski (1993) draws attention to the identification of the "endogenous" and the "contextual" effect because only the former generates a positive social multiplier. He shows that those two effects are not identified with the linear-in-means model under complete networks. As a solution, utilizing more detailed specification of networks is suggested (Kelejian and Prucha 1998, 2010; Lee 2003), and the existence of the multiplier effect is also identified as a result. Nevertheless, the endogenous effect of the standard SAR model is still based on the complementarity peer effect, and the multiplier effects under different motives are not fully explained.

The existence of the social multiplier effect based on the linear-in-means model is empirically studied by Glaeser, Sacerdote, and Scheinkman (2003). In their work, the authors find evidence of the multiplier in the form of $1/(1-x)$ from multiple datasets. The limitation of this conventional form is that, as in Manski (1993), a network is assumed to be complete. Its result is a significantly higher multiplier for most cases, considering social networks are generally sparse. Indeed, (Giorgi, Frederiksen, and Pistaferri 2019) shows that such a measure can be misleading under more realistic network structures. The definition of the social multiplier effect proposed in this paper is rooted in the marginal effect analysis employed by many social network studies (Liu, Patacchini, and Rainone 2017; Chomsisengphet, Kiefer, and Liu 2018). This common notion of the multiplier under general networks is extended to incorporate the distinct impacts of the complementarity and conformity motives with a single metric.

This paper connects to the SAR model with binary outcomes of Lee, Li, and Lin (2014). It is assumed that the players have complete information on network structures and attributes of their peers. Because the GSAR model does not depend on the assumption or the form of the outcome variable, it can be applied to the case of incomplete information (Yang and Lee 2017) or the Tobit model (Yang, Lee, and Qu 2018).

For the application to the microfnance program, the GSAR model is also extended for the higher-order networks. A particular line of literature on geographical relationships employs a convex combination to utilize multiple networks. Debarsy and J. LeSage (2018) concerns the convex combination model to address the scaling issue between multiple networks derived from continuous measures. While it has the advantage that it is easier to include networks with diferent distance measures, more complicated estimation techniques are required (Debarsy and J. P. LeSage 2022). In contrast, networks are defned as binary relationships in many social networks, making distance measures less problematic. Therefore, the higher-order specifcation of this paper follows the simpler form of Blommestein (1983) and Huang (1984). It is worth noting that the method used in Hsieh and Lin (2017) can estimate the peer efects across networks. For the case of this study, however, the peer efect between villages is believed to be minimal.

¨ g nLOe

Consider a set of samples with N individuals who are linked to each other by a non-stochastic, exogenous social network. The network is defned by a graph on a set of vertices, V, and a nonempty set of edges, E, which is denoted as $G_N = G_N(V; E$

Definition 1.

$$A = \begin{cases} [a_{ij}] & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Adjacency matrix)}$$

$$D = \begin{cases} [d_i] & \text{if } j = i; \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Degree matrix)}$$

In the literature of the spatial autoregressive model (SAR), the connectivity of individuals is represented by a weight matrix. Throughout this paper, it is defined as a combination of the adjacency and the degree matrices multiplied by *a conformity parameter*, $\ \ 2\ [0;1] \ \ \mathrm{R}$.

Definition 2. *A generalized weight matrix derived from the network structure, G_N, is an $N \quad N$ square matrix defined as follows:*

$$W(\ ) = A \qquad D;$$

This definition encompasses both the ordinary SAR model ($= 0$) and the graph Laplacian variant of Boucher (2016) ($= 1$) as special cases. The proposed weight matrix is a type of the negative generalized graph Laplacian, where its diagonal may be any nonnegative numbers.

The generalized spatial autoregressive model (GSAR) is proposed as follows:

$$y = \quad W(\ )y + X \quad + '';\tag{1}$$

where $\ > 0$ and X is an $N \quad K$ covariate matrix. The outcome variable, y, is an $N \quad 1$ vector of real numbers. The vector of coefficients, , is a $K \quad 1$ vector and measures the direct effect of one's attributes. The idiosyncratic error, $''$, is a random variable independent of the individual characteristics and the network structure, and no particular distribution is imposed. The parameter of interest is $\ = (\ ;\ ;\ ^0)^0$.

Assumption 1. *(Exogenous network) The network structure, G_N, is independent of the idiosyncratic error, $''$.*

Assumption 2. *(Model stability) The domain of the peer effect parameter is bounded by the spectral radius of $W(\ )$. That is,*

$$0 < \quad <$$

costs.

$$u_i(y_i; y_{-i}; x_i; \varepsilon_i) = \underbrace{\underbrace{\phi \sum_{j \neq i} a_{ij} y_j y_i}_{\text{Complementary benefit}} - \underbrace{\frac{\gamma}{2} \sum_{j \neq i} a_{ij} (y_j - y_i)^2}_{\text{Deviation cost}}}_{\text{Network net benefit}} + \underbrace{(x_i \beta + \varepsilon_i) y_i - \frac{1}{2} y_i^2}_{\text{Private net benefit}}; \tag{3}$$

where $\phi > 0$, $\gamma \geq 0$ and a_{ij} is the $(i; j)$ element of A. Given this, player i chooses an action, y_i, to maximize his or her utility, u_i. The network net benefit of i is maximized for the set of actions of the other players, y_{-i}. The private net benefit is determined by his or her own attributes, x_i, and an idiosyncratic shock, ε_i. The benefit arises from complementarity, which has been focused by many studies (Ballester, Calvó-Armengol, and Zenou 2006; Bramoullé, Kranton, and D'Amours 2014). The players benefit from each others' engaging in the same activity if they are linked by an edge, a_{ij}, in a network. On the other hand, the deviation cost obtained that $= {}^{0}$, which is a contradiction and proves the uniqueness of . Trivially, is also uniquely

Figure 3: The pure conformity peer effect

$$A - D = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}; \quad X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}; \quad (I - 0.15(A - D))^{-1} X = \begin{bmatrix} 0.80 \\ 0.10 \\ 0.10 \end{bmatrix}$$

A simple star network with three vertices is focused to elucidate further the difference between the existing

For $= 0$, $|I \quad W(0)|$ is equal to the difference between 1 and the sum of paths of all lengths discounted by . Therefore, $0 < |I \quad W(0)| < 1$ by Assumption 2. □

As the conformity parameter is flexible in the GSAR model, the social multiplier is hinged upon the relative magnitudes of the two motives. Accordingly, a threshold where the social multiplier is exactly equal to 1 can be found.

Proposition 3. *The threshold for the positive social multiplier is a solution of* $(\ ; \ ) = 1$*, which is*

$$(\ ) = \frac{2}{1 + {}^2 \sum_{i=1}^N d_i^2 / \sum_{i=1}^N d_i}.$$

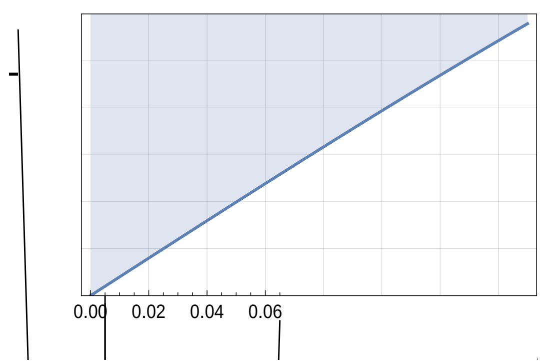
If $>$ *, there exists a positive social multiplier effect, and* $<$ *, there exists a negative social multiplier effect.*

Proof. See Appendix A. □

This result suggests that, for the positive social multiplier effect, the complementary benefit must be sufficiently higher than the conformity motive of individuals. Note that the coefficient, $\sum_{i=1}^N d_i^2 / \sum_{i=1}^N d_i$, is parallel to the "tendency to make hubs centrality" proposed by Saberi et al. (2021). The implication is that the more vertices are linked with a specific "hub" vertex, the easier for the social multiplier effect to arise.

The threshold is based on the sparse approximation of matrix determinants proposed by Ipsen and Lee (2011). The performance of approximation depends on the sparsity of non-diagonal elements of a matrix, and, indeed, most of social networks are qualified as sparse. A graphical presentation of a numerical and an approximated solution for a random graph with $N = 10$ and a density of 10% is shown in Figure 5. The

Figure 5: The numerical (left) and the approximated solution (right)



Note: The graph is generated randomly with $N = 10$ and density of 10%.

shaded areas represent where the multiplier effect is positive. The example shows that the approximated solution is tracking the actual solution relatively well, even for the higher density than the graphs in the real world.

¨u© N²§Oj _njI ?_j9}´ G\n_HO gnLOe °_§\ |9§_nj9e N²qOH§9§_nj

In this section, the GSAR model is extended to the binary outcome variables with rational expectation. Individuals are considered as players who maximizes their utility without directly observing the others'

Proposition 4. *Under Assumption 8 and 9, the set of parameter,* $\quad = (\; ; \; ; \; {}^{0})^{0}$, *is identifed.*

Proof. See Appendix A. □

V 8qqe_H9§_njl ¦\O pOO} N¸OH§ _j §\O g_H}n»j9jHO p}nX}9h

Vuo K9§9 KO H}_q§_nj

This study utilizes the data set collected by Banerjee et al. (2013) in collaboration with Bharatha Swamukti Samsthe (BSS), a non-proft organization that conducted the microfnance program in Karnataka, India. The authors had conducted a household-level survey prior to the entry of BSS into 43 rural villages in the target districts including household characteristics such as roof materials, ownership of houses, access to electricity, religion, and castes. An individual-level survey was conducted for the members of households that are randomly selected among eligible ones to collect more detailed information. As a result, additional data were obtained for about 46% of the sample households. BSS had reported the take-up rate of the program periodically.

The main interest of the individual survey was the social network information. Respondents were asked to provide the names of their peers who belong to a total of twelve dimensions of their peer network respectively as follows: (1) from whom they would borrow money, (2) to whom they would lend money, (3) to whom they give advice, (4) from whom they fnd help for important decision making, (5) from whom one would borrow kerosine and rice, (6) to whom one would lend kerosine and rice, (7) to whom they visit for free time, (8) who visits them for free time, (9) from whom they seek medical advice, (10) relatives in the village, (11) non-relative people with whom they socialize and (12) people with whom they go to temples together. Note that some networks, such as (1) and (2), are examined bidirectionally. The average overlapping rate of the are considered leader households.

The entries without additional details such as caste or religion are removed to control the household characteristics.[3] As a result, a total of 7,919 households are used. The descriptive statistics can be found in Figure 1. There is no noticeable difference in the characteristics depending on the sample sizes. The additional characteristics are crucial due to the potential endogeneity between the socioeconomic status variables and the peer networks. In the original paper, the individual-level demographics are not used for the analysis. In this paper caste and religion are controlled by imposing the attributes of the heads of households on each sample. This is not a significant issue, as the other member of the households are likely to be identical with their heads'.

As pointed out by many studies, random removal of nodes on a network does not guarantee unbiasedness and consistency. If such an issue cannot be avoided, it is advisable to clarify its expected impacts. Comparing the density and the spectral radius of the networks under different sample sizes, it appears that the density increases while the radius decreases. It may be interpreted as a result of more nodes with smaller degrees and some bridging the others being removed. In this case, a downward bias is expected on the estimates of the peer effects. Consequently, if any estimates of peer effects are found significant under this smaller sample, it can be believed as a lower bound. Also, all 1,672 of the excluded samples are only from the first two villages, so its impact will be contained in those two.[4]

The dummy variables for the income proxies are constructed to present lower qualities, considering that the most demands are from underprivileged households.

i O§°n}c G\9}9H§O}_ §_H

Definitions and graph statistics of the surveyed networks are summarized in Table 2. For the individual networks, the average degree is 3.284, and the average density is 0.023, which is common sparsity for social networks.[5][6] The union network is constructed by the villagers who are connected through any of the individual dimensions.

A well-known issue in social network studies is the endogeneity of networks, which is the case where the networks are correlated with unobserved attributes of individuals. Although this cannot be completely eliminated, it can be mitigated by investigating the source of correlation and finding proper control variables. In Table 3, the relationships between the chosen attributes of the respondents and those of those who were

3. Christianity is excluded due to its extremely small sample size
4. See Table A1 for full descriptive statistics.
5. The respondents were asked to report up to four persons for each network. It can potentially cause a mismeasurement problem suggested by Griffith (2022). Its potential impact, however, is only an underestimation of peer effect, not an overestimation. If a peer effect turns out to be positive, it is not an issue anymore, which is the case in this study.
6. The degrees over four is due to the directionally surveyed networks merged into a single one.

Table 1: Descriptive statistics

Variables	Description	Mean	S.D.
Outcome variable			
Microfnance take-up	Yes = 1, No = 0	0.178	0.382
Income proxies			
Roof material: tile	Yes = 1, No = 0	0.37	0.48
Number of rooms		2.41	1.3
Number of beds		0.82	1.23
No latrine in house	Yes = 1, No = 0	0.73	0.45
House not owned	0, if the house is privately owned, 1, otherwise	0.1	0.3
No access to electricity	Yes = 1, No = 0	0.38	0.48
Social Status			
Leader group	1, if the household belongs to the leader group, 0, otherwise	0.13	0.33
Other backward castes	Base caste variable	0.53	0.5
General caste	Yes = 1, No = 0	0.12	0.33
Minority castes	Yes = 1, No = 0	0.2	0.13
Scheduled caste	Yes = 1, No = 0	0.28	0.45
Scheduled tribe	Yes = 1, No = 0	0.05	0.22
Religion			
Hinduism	Base religion variable	0.95	0.21
Islam	Yes = 1, No = 0	0.05	0.21
Village			
Village Population		184.16	86.08

N = 7,919
Number of villages = 43

chosen as their peers are shown. The two attributes are the number of beds per capita and rooms per capita, which are the proxies for income. According to the numbers shown, they tend to report those who are wealthier as peers.[7] Since the income proxies are included in the regression, it can be expected that the network endogeneity is controlled by the explanatory variables.

The other potential source of endogeneity is networks related to outcome variables. However, this possibility can be precluded for this study, as it is not conceivable for the villagers to adjust their social network

Table 2: Networks description

Network	Description	Degree	Density	Spectral radius
Money	From/to whom they may borrow/lend money	4.244		

peer effect is provided in the original research.[8] As a related study, an attempt is made by Chandrasekhar and Lewis (2011), where the issue is attributed to a sampled network and consequent missing network links.

In this paper, however, a distinct approach is taken with the specification of a network. As mentioned in the previous section, they surveyed a total of twelve dimensions of social networks and took a union of them for their study. By doing so, it is ignored that an individual's social network has multiple dimensions, and one

Table 4: Ordinary logit and SAR with the union network

	Model 1		Model 2	
	Mean	SD	Mean	SD
Union			0.125***	0.018
Constant	-1.376***	0.289	-1.864***	0.223
Roof material: tile	0.105	0.076	0.099	0.076
No. of rooms	-0.026	0.032	-0.053*	0.032
No. of beds	-0.018	0.034	-0.025	0.034
No latrine in house	0.353***	0.085	0.401***	0.085
House not owned	0.020	0.103	0.031	0.101
No access to electricity	0.190***	0.073	0.212***	0.072
Leader group	1.098***	0.167	0.572***	0.087
General caste	-0.312**	0.134	-0.254***	0.125
Minority castes	-0.045	0.235	0.228	0.239
Scheduled caste	0.493***	0.083	0.395***	0.076
Scheduled tribe	0.406***	0.146	0.372***	0.138
Islam	0.627***	0.087	0.935***	0.147
Multiplier			1.40	
Likelihood	3451.06		3431.58	
AIC	6928.12		6891.16	
BIC	7018.82		6988.84	

Significant at *10%, **5%, ***1%.

with = 0 (Model 3), the pure conformity model with = 1 (Model 4), and the GSAR model with 2 [0; 1] (Model 5). Model 3 reveals that only three of the networks, Relatives, Temple, and Visit, transmit significant peer effects. Note that the first four networks highly correlated with the income proxies turn out to be insignificant, which suggests that they are properly controlled by the covariates. In Model 4, no significant peer effect is found, and the coefficients of the household characteristics are similar to Model 1.

The main result of this paper is Model 5. In addition to the peer effects, the conformity parameter, , is estimated as 0.284 and significant at 95% confidence level. Although the number is closer to zero than one, this does not imply that complementarity is the dominant motive. The social multiplier computed from the estimates is 0.009, which means that the complementary benefit is not high enough to overwhelm the deviation cost for the villagers. This becomes more apparent with the marginal effects that will be presented later.

Despite the existence of a significant conformity motive, the pure conformity model does not capture such peer effect from the data. One of the possible explanations is the higher requirement of the pure complementarity model. Due to the smaller variance in the difference of choice probabilities, p_j p_i, the pure conformity model needs more samples to achieve the same level of performance as the pure complementary

Table 6: Network specifc conformity parameters

	Model 6		
	Mean		SD
Relatives	0.410	**	0.179
Visit	0.250	**	0.121
Relatives	0.644	**	0.256
Visit	0.091		0.367
Constant	-1.769	***	0.229
Roof material: tile	0.137	*	0.082
No. of rooms	-0.079	**	0.034
No. of beds	-0.027		0.037
No latrine in house	0.413	***	0.094
House not owned	0.058		0.107
No access to electricity	0.230	***	0.078
Leader group	0.836	***	0.143
General caste	-0.234	*	0.121
Minority castes	0.308		0.257
Scheduled caste	0.335	***	0.074
Scheduled tribe	0.340	**	0.135
Islam	0.667	***	0.108
Multiplier	0.007		
Likelihood	-3411.4		
AIC	6856.80		
BIC	6975.41		

An interesting observation is that the indirect efects of Model 5 and 6 are higher than that of Model 3. Again, consider the leader group as an example. Even though those who are appointed as leaders might be less willing to join the program compared to those under complementarity, the other people's take-up rate can be increased even more by the same conformity motive. Considering the social status of the leaders, this interpretation is not inconceivable.

Table 7: Marginal efects: Model 1-3

	Model 1	Model 2			Model 3		
		Naive	Direct	Indirect	Naive	Direct	Indirect
Roof material: tile	1.43	1.34	1.34	0.002	1.69	1.72	0.004
No. of rooms	-0.35	-0.71	-0.71	-0.001	-0.86	-0.88	-0.002
No. of beds	-0.14	-0.35	-0.35	-0.001	-0.26	-0.27	-0.001
No latrine in houses	2.72	5.13	5.16	0.008	5.09	5.17	0.012
House not owned	0.16	0.42	0.42	0.001	0.69	0.70	0.002
No access to electricity	1.49	2.91	2.92	0.005	2.94	2.99	0.007
Leader group	4.83	8.68	8.73	0.013	9.14	9.31	0.021

Table 8: Marginal efects: Model 4-6

	Model 4			Model 5			Model 6		
	Naive	Direct	Indirect	Naive	Direct	Indirect	Naive	Direct	Indirect
Roof material: tile	1.83	1.55	0.002	1.99	1.85	0.007	1.86	1.71	0.005
No. of rooms	-0.65	-0.55	-0.001	-1.14	-1.06	-0.004	-1.07	-0.99	-0.003
No. of beds	-0.28	-0.23	0.000	-0.25	-0.23	-0.001	-0.36	-0.33	-0.001
No latrine in houses	5.52	4.72	0.005	5.28	4.93	0.019	5.30	4.89	0.016
House not owned	0.28	0.24	0.000	0.73	0.67	0.002	0.80	0.74	0.002
No access to electricity	2.66	2.24	0.002	3.08	2.85	0.011	3.17	2.90	0.009
Leader group	14.31	11.66	0.012	9.96	9.14	0.033	10.30	9.34	0.029
General caste	-2.73	-4.09	-0.004	-2.97	-2.77	-0.010	-3.01	-2.79	-0.009
Minority castes	-2.01	-1.72	-0.002	6.80	6.24	0.023	4.53	4.12	0.013
Scheduled caste	7.11	5.94	0.006	4.22	3.91	0.014	4.72	4.32	0.014
Scheduled tribe	6.55	5.39	0.006	5.12	4.72	0.017	5.00	4.55	0.014
Islam	21.49	17.05	0.018	10.92	9.97	0.036	13.73	12.35	0.038
Money	0.50			-0.79					
Advice	1.62			-1.47					
Kerosine	1.19			0.65					
Medical	0.96			1.33					
Non-relatives	-1.26			-0.56					
Relatives	2.46			2.43			5.56		
Temple	3.43			7.06					
Visit	0.41			5.15			3.39		

The numbers are in percentage points.

Table 9: Comparison with the relevant networks

	Aggregated		Union		Selected		Individual	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Degree	6.36	1.27	9.22	1.69	2.36	0.16	3.00	0.24
Density	0.04	0.02	0.06	0.04	0.02	0.01	0.03	0.01
Spectral Radius	9.78	2.35	14.07	3.20	4.12	0.54	5.22	0.66

Note: the selected networks are Relatives, Temple and Visit. The statistics are averaged over the three networks.

of the union network, this is a signifcantly lower number. Network researchers are often concerned about network misspecifcation, especially with missing links. However, this can be considered as evidence that suggests that misspecifcation in the other direction, or overspecifcation, could also be a problem.

U GnjHe« _nj

In this study, a generalized SAR model that unifes and estimates the complementarity and the conformity peer efects is proposed. The distinguished feature of the model is its weight matrix that is a generalized graph Laplacian with its diagonal elements multiplied by the conformity parameter. With the microeconomic

References

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou. 2006. “Who’s Who in Networks. Wanted: The Key Player.” *Econometrica* 74 (5): 1403–1417.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2013. “The Diffusion of Microfinance.” *Science* 341 (6144): 1236498–1236498.

Blommestein, Hans J. 1983. “Specification and estimation of spatial econometric models A discussion of alternative strategies for spatial economic modelling.” *Regional Science and Urban Economics* 13 (2): 251–270.

Boucher, Vincent. 2016. “Conformism and self-selection in social networks.” *Journal of Public Economics* 136:30–44.

Boucher, Vincent, and Bernard Fortin. 2016. “The Oxford Handbook of the Economics of Networks,” 276–302.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. “Identification of peer effects through social networks.” *Journal of Econometrics* 150 (1): 41–55.

Bramoullé, Yann, Rachel Kranton, and Martin D’Amours. 2014. “Strategic Interaction and Networks.” *American Economic Review* 104 (3): 898–930.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. “Understanding Mechanisms Underlying Peer Effects: Evidence From a Field Experiment on Financial Decisions.” *Econometrica* 82 (4): 1273–1301.

Chandrasekhar, Arun G., and Randall Lewis. 2011. “Econometrics of Sampled Networks.” *Working Paper.*

Chomsisengphet, Souphala, Hua Kiefer, and Xiaodong Liu. 2018. “Spillover effects in home mortgage defaults: Identifying the power neighbor.” *Regional Science and Urban Economics* 73:68–82.

Debarsy, Nicolas, and James LeSage. 2018. “Flexible dependence modeling using convex combinations of different types of connectivity structures.” *Regional Science and Urban Economics* 69:48–68.

Debarsy, Nicolas, and James P. LeSage. 2022. “Bayesian Model Averaging for Spatial Autoregressive Models Based on Convex Combinations of Different Types of Connectivity Matrices.” *Journal of Business & Economic Statistics* 40 (2): 547–558.

Giorgi, Giacomo De, Anders Frederiksen, and Luigi Pistaferri. 2019. “Consumption Network Effects.” *The Review of Economic Studies* 87 (1): 130–163.

Glaeser, Edward L., Bruce I. Sacerdote, and Jose A. Scheinkman. 2003. “The Social Multiplier.” *Journal of the European Economic Association* 1 (2-3): 345–353.

Griffith, Alan. 2022. “Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects

Liu, Xiaodong, and Jiannan Zhou. 2017. "A social interaction model with ordered choices." *Economics Letters* 161:86–89.

Manski, Charles F. 1993. "Identifcation of Endogenous Social Efects: The Refection Problem." *The Review of Economic Studies* 60 (3): 531.

Patacchini, Eleonora, and Yves Zenou. 2012. "Juvenile Delinquency and Conformism." *The Journal of Law, Economics, and Organization* 28 (1): 1–31.

Saberi, Majid, Reza Khosrowabadi, Ali Khatibi, Bratislav Misic, and Gholamreza Jafari. 2021. "Topological impact of negative links on the stability of resting-state brain network." *Scientifc Reports* 11 (1): 2176.

Ushchev, Philip, and Yves Zenou. 2020. "Social norms in networks." *Journal of Economic Theory* 185:104969.

Yang, Chao, and Lung-fei Lee. 2017. "Social interactions under incomplete information with heterogeneous expectations." *Journal of Econometrics* 198 (1): 65–83.

Yang, Chao, Lung-fei Lee, and Xi Qu. 2018. "Tobit models with social interactions: Complete vs incomplete information." *Regional Science and Urban Economics* 73:30–50.

8 p}nnT

The following Lemma will be used for the identifcation result for both continuous and binary outcome models.

Lemma 1. *Let* $Q_A = A_s T^{1}$ *and* $Q_D = D_s T^{1}$*, where* $T = I \quad {}_0 W(\; {}_0)$*. Then,* $(\quad {}_0) Q_A X_s \; {}_0 \quad (\; {}_0 \; {}_0) Q_D X_s \; {}_0 = 0$ *if and only if* $= {}_0$ *and* $= {}_0$*.*

Proof. Suppose that $\neq {}_0$. Then, $(\quad {}_0) A_s T^{1}$

$(\tilde{};\tilde{};\tilde{})$. Then, for each set, there exists an equilibrium that satisfes

$$p_i(\ ) = F_i(\ a_i \mathrm{p} \qquad d_i \mathrm{p} \ + x_i\ )$$

$$\tilde{p}_i\ (\tilde{}) = F_i(\tilde{a}_i \mathrm{p} \qquad \tilde{d}_i \mathrm{p} \ + x_i \tilde{});$$

where a_i and d_i are i-th row of A and D. Since $p_i(\ )$ and $\tilde{p}$

Table A1: Full descriptive statistics

			Whole samples		Participating samples		Participating samples w/ caste data	
			N=14890 Number of villages = 75		N=9591 Number of villages = 43		N=7919 Number of villages = 43	
	Min	Max	Mean	SD	Mean	SD	Mean	SD
Outcome variable								
Microfnance take-up	0	1			0.182	0.386	0.178	0.382
Income proxies								

G K_ H« _nj nT ?9jO}aOO O§ 9eu r©·o¨s

This section discusses the methodological diferences between the GSAR model and the difusion model of Banerjee et al. (2013), BCDJ henceforth. The main fnding of BCDJ is the role of non-participants in the information difusion process. Although the non-participants are less likely to pass information to others, they play a larger role in the overall take-up rate due to their larger number. Meanwhile, the authors do not fnd any evidence of the "endorsement efect," which corresponds to the peer efect of this paper, in their study. In other words, not only the probability of participation does not increase with the number of already participating peers, but even slightly decreases. Their paper does not fully explain this negative result, but some hints are suggested by Chandrasekhar and Lewis (2011). In that paper, the insignifcant peer efect is attributed to the missing links in sampled networks. As mentioned in Chapter 4, the organizers surveyed

2) estimate the peer effect with the entire sample using the estimated coefficients. The main assumption implicitly made here is that the coefficients do not vary across the rest of the individuals outside of the leader group. As the leaders are chosen by the organization based on their characteristics, it is reasonable to assume that there is no significant social link between them. Therefore, the estimated coefficients will be

of the take-up rate over time. Instead, estimation is easier to perform and does not require panel data. Meanwhile, the difusion model explains the difusion pattern over time better but the probability of passing information must be estimated as well. The additional computational burden and theoretical assumptions will be the cost that must be incurred.

Table A3: The two-step and the SAR estimation with additional covariates

	Two-step estimation	SAR
	N=995	N=7919
Constant	-1.186***	-2.050***
	(.260)	(.116)
Roof material	.251	.143**
	(.166)	(.060)
Number of rooms	-.131**	-.063**
	(.067)	(.030)
Number of beds	-.063	-.032
	(.076)	(.030)
Islam	1.290***	1.046***
	(.404)	(.138)
No private latrine	.337*	.200**
	(.187)	(.080)
House not owned	-.159	.057
	(.308)	(.096)
House not owned	-.159	.057
	(.308)	(.096)
No electricity	.022	.214***
	(.190)	(.067)
General caste	-.153	-.409***
	(.221)	(.113)
Minority caste	.559	.153
	(.659)	(.231)
Scheduled caste	.385*	.396***
	(.198)	(.068)
Scheduled tribe	-.390	.223*
	(.448)	(.131)
Leader		.564***
		(.085)
	N=7919	N=7919
Peer efect	-.122***	.082***
	(.015)	(.012)