CSAE Working Paper WPS/2014-11

Peer effects in development programme awareness of vulnerable groups in rural Tanzania

February 2014

Bet Caeyers[1]

Abstract

A pre-condition for grassroots participation, key for community-based development success, is widespread programme knowledge among the eligible population. The current literature on local participatory institutions mainly focuses on village meetings and media campaigns as a means to strengthen community awareness. The role played by social interactions in this process has received little attention to date. In this paper I use Manski's (1993) standard linear-in-means model to estimate endogenous peer effects on the awareness of vulnerable groups on Tanzania Social Action Fund II (TASAF II), i.e. Tanzania's flagship community-driven development programme. I employ a popular 2SLS estimation strategy developed by Bramouille et al. (2009) and De Giorgi et al. (2010) on a unique spatial household dataset from Tanzania to eliminate both the 'reflection bias' (Manski, 1993) and the 'exclusion bias' (Caeyers, 2014). Denoting the geographically nearest neighbours set as the relevant peer group in this context, I identify significant average and heterogeneous endogenous social interaction effects in the diffusion of information about TASAF II. The findings of this paper inform the design of effective sensitisation campaigns.

[1] University of Oxford, UK; email: betcaeyers@gmail.com; I thank the World Bank for sharing with me the baseline data collected for the impact evaluation of TASAF II. Specifically, I gratefully appreciate the support I received from Sarah Baird, Craig McIntosh and Berk Ozler. I also thank the Department of Economics at the University of Oxford, UK, for funding an additional trip to Tanzania to collect qualitative data for this project. I very much benefited from the support and assistance I received from Lies Caeyers during my last field trip to Tanzania. Finally, I thank Oriana Bandiera, Matthew Collin and Marcel Fafchamps for their constructive feedback on an earlier draft of this paper.

Centre for the Study of African Economies
Department of Economics · University of Oxford · Manor Road Building · Oxford OX1 3UQ
T: +44 (0)1865 271084 · F: +44 (0)1865 281447 · E: csae.enquiries@economics.ox.ac.uk · W: www.csae.ox.ac.uk

1. Introduction

There is a growing literature on the effectiveness of community-based development (CBD) projects (for a recent review, see Mansuri and Rao, 2013). A pre-condition for grassroots participation, key for CBD success, is widespread programme knowledge among the eligible population (Khemani, 2007). Transparency allows community members to participate and to hold those in charge of resource spending to account. Asymmetric information, with poor and illiterate households having less access to information than richer and more powerful ones, may result in elite capture in decentralised public service delivery (Bardhan and Mookherjee, 2006; Conning and Kevane, 2002). The current literature on local participatory institutions focuses on village meetings and media campaigns as a means to strengthen community awareness (Goetz and Jenkins, 2001; Banerjee et al., 2010; Bjorkman and Svensson, 2007; Khemani, 2007; Olken, 2007; Shankar et al., 2010). The role played by peer effects in this process has received little attention to date. This paper investigates the extent to which well-informed peers influence the development programme awareness of vulnerable groups in rural Tanzania, the elderly and the disabled in particular.

Recent years have seen a growth in social interaction literature, especially in the area of education (Borjas, 1995; Gaviria and Raphael, 2001; Sacerdote, 2001; Hanushek et al., 2003; Cipollone and Rosolia, 2007), but also in other areas such as technology adoption (Bandiera and Rasul, 2006; Conley and Udry, 2010; Van den Broeck and Dercon, 2011) and criminal activity (Glaeser et al., 1996; Bayer et al., 2009). In this literature, individuals' outcomes are influenced not only by their own characteristics but also by exogenous characteristics of their peers (*contextual* peer effects) or by the prevalence of this outcome among their reference groups (*endogenous* peer effects).

There is some evidence from the developed world to indicate that endogenous peer effects may also be present in the diffusion of information on welfare programmes (Duflo and Saez, 2003). The evidence from developing countries is limited. Information sharing among peers may well be crucial for a vulnerable household's awareness of, and hence participation in, CBD programmes. Most community-based initiatives rely heavily on village meetings and media campaigns (e.g. radio, television and newspapers) to sensitise communities (Mansuri and Rao, 2013). However, as a result of their physical and/or socioeconomic conditions, various disadvantaged demographic groups, such as the elderly and disabled, have only restricted access to such traditional sources of information about development programmes. To the extent that these groups have only limited direct access to news sources, social

interactions with well-informed peers may be crucial for their inclusion in bottom-up development approaches.

In this paper, we examine whether endogenous peer effects were present in the public diffusion of information about the Tanzania Social Action Fund II (TASAF II), one of the flagship community-driven development (CDD) projects of the government of Tanzania.[2] Baird et al. (2013) find that 49% of the eligible non-beneficiary households had never heard about the programme.[3] Their study identified the household's direct access to information (such as education and village meeting attendance rate) and its direct political connectedness (such as being blood related to the village elite) as dominant determinants of being informed. However, Baird et al. (2013) do not consider the potential role of social interactions with informed peers in the information diffusion process. In view of the earlier discussion, the current paper concentrates on two specific vulnerable groups (VGs) targeted by TASAF II, namely the elderly (defined, in our base model, as those older than 55 years) and the (mentally or physically) disabled. The aim of this paper is to measure the extent to which these VGs in rural Tanzania benefited from having well-informed individuals living in their neighbourhood, in terms of information exchange about TASAF II.

The household data that we use to answer this empirical question come from a substantial baseline listing exercise organised by the World Bank in April–November 2008 to stratify and sample households to be followed up for an impact evaluation of the VG component of TASAF II. This unique dataset captures details of 30,339 households, including the 19,916 households of *all* elderly and disabled people, in 100 villages in rural Tanzania. The dataset includes, among others, global positioning system (GPS) coordinates of each of these households and information about these households' knowledge of TASAF.[4] The unit of study in this paper is a household in which there is at least one elderly and/or disabled person (henceforth 'elderly/disabled household').

Our identification strategy employs Manski's standard linear-in-means model (Manski, 1993). A set of empirical problems related to identifying peer effects have been extensively discussed in the literature, including the *reflection problem* and the presence of *correlated effects* (e.g. Manski, 1993; Brock and Durlauf, 2001; Moffitt, 2001). More recently, Caeyers (2014) has formally shown that the OLS estimate of endogenous peer effects in standard

[2] Whereas CBD is an umbrella term for projects that engage communities in their design and management, CDD refers to CBD projects in which communities have control over key project decisions, including management of funds.

[3] The VG component of this programme was targeted at the elderly, the disabled, people infected with HIV/AIDS, orphans, widow(ers) and unemployed youth.

[4] The data were collected using advanced computer-assisted personal interviewing (CAPI) techniques.

linear-in-means models suffers from a downward *exclusion bias* whenever households are excluded from their own peer groups, even when peers are randomly assigned.

To deal with these empirical challenges, this paper employs a two-stage least squares (2SLS) estimation strategy developed by Bramouille et al. (2009) and De Giorgi et al. (2010). This strategy uses exogenous characteristics of peers' peers, i.e. *excluded* peers, as instruments for the peer group's average outcome. In addition, we control for *hamlet* fixed effects (something that is rarely feasible in practice due to data limitations) and a set of observed neighbours set and individual household characteristics.[5]

Our findings are consistent with a significant and large endogenous peer effect: *each* additional informed neighbour in the set of 10 nearest sampled neighbours raises an elderly/disabled household's probability of being informed by 8 percentage points (pp). This compares to an increase by 14 pp of the probability of being informed if a household attended a village meeting (i.e. TASAF's most important sensitisation medium). Importantly, the unique size of the dataset also allows us to identify the type of elderly/disabled households that are most/least receptive to information in the neighbourhood, and those households (vulnerable and non-vulnerable) that are most effective in transmitting information to others. From a policy perspective, the finding of such a significant endogenous peer effect is of primary importance, in view of its social multiplier effect. Further, these results provide development intervention planners, targeting elderly and disabled households, with a better understanding of which individuals in the village network to target in their sensitisation campaigns.

This paper is organised as follows. Section 2 summarises the TASAF II programme objectives, resource allocation procedures and awareness-building strategies. Section 3 subsequently describes the data and motivates the peer group measure used in this study. Before moving on to the empirical analysis, section 4 carefully discusses the empirical issues involved and the identification strategy employed. The empirical results of the paper are presented in section 5. The main results of the study are tested against robustness to alternative specifications in section 6. Section 7 concludes.

[5] The hamlet (*kitongoji*) is the smallest administrative unit in rural Tanzania.

2. Background of Tanzania Social Action Fund II

2.1. Programme objectives and resource allocation process

Tanzania is currently implementing TASAF II, a large-scale CDD fund financed by the World Bank. Launched in May 2005 as an important component within the national framework of the Tanzanian National Strategy for Growth and Reduction of Poverty (popularly known as MKUKUTA), TASAF II follows its predecessor TASAF I (2000–2005) in the aim of empowering local communities to participate in development initiatives..

Through an elaborate screening process, eligible communities throughout Tanzania could collectively apply for TASAF funding to implement projects corresponding to their most urgent needs. The benefiting community could be an entire ward, an entire village or a subgroup of the village. Among the communities explicitly targeted by TASAF II were the vulnerable groups (VGs), comprising orphans, disabled, elderly, widows/widowers and those infected by HIV/AIDS.[6] These VGs could collectively apply for funding to start up income-generating group activities (Tanzania Social Action Fund, 2011a).

In accordance with the principles of CDD, TASAF's institutional structure is highly decentralised. Officials at the national level allocate funds to districts, and within those districts resource allocation follows a bottom-up approach (Tanzania Social Action Fund, 2011a). The village council was responsible for sensitising all members of its village about TASAF and for encouraging all eligible communities to participate in the programme. Any interested community could indicate its interest in a particular project and apply for TASAF funding. TASAF projects were selected through pairwise voting at a public village meeting attended by at least 50% of the targeted community members.

At the outset, it was expected that one-third of all 11,000 Tanzanian villages would have at least one project funded by TASAF II by the end of 2010, targeting 5950 projects in total.[7] In June 2008, at the time of baseline data collection for an impact evaluation (see below), a total of 4582 subprojects had been selected for funding by TASAF II, of which 1091 (23.8%) were for VGs (Tanzania Social Action Fund, 2008).[8]

[6] TASAF II also targeted (i) service-poor communities, which could apply for projects improving basic social and market services such as education, water and sanitation, roads, banking and markets, and (ii) communities wishing to apply for funding to start up public work programmes targeted to the food insecure, that is, able-bodied members of food-insecure households.

[7] In June 2011, the number of projects funded amounted to 10,526, far exceeding the original target of 5950 projects, thanks to additional resources obtained (Tanzania Social Action Fund, 2011b).

[8] Note that, in June 2011, the distribution of subprojects across beneficiary groups had moved to 41% service poor, 50% VGs and 9% food insecure (Tanzania Social Action Fund, 2011b).

2.2. Sensitisation strategies and the scope for peer effects

TASAF II's flagship feature is its CDD development approach. Successful CDD approaches are expected to be inclusive, that is, to empower all minority groups to identify and express their priorities. To achieve this objective, it is important that programme awareness reaches all members of the community.

One of the main activities implemented by TASAF during the first 2 years of implementation (2005–2007) was, therefore, building programme awareness through various sensitisation methods (Tanzania Social Action Fund, 2006, 2007). Information about TASAF was disseminated through various media, such as national and regional radio programmes, newspapers, newsletters and television. Technical launch workshops were organised for politicians and district technical staff to inform them about the delivery details of TASAF II. Communication and promotional materials, such as brochures, posters, t-shirts and caps, were distributed to districts, which in turn were responsible for further distribution to their respective communities. Every one of Tanzania's 11,000 villages were supposed to be visited by a district official and given details about the programme.

The village council members in turn were in charge of raising awareness to the larger public through village meetings. The extent to which this exercise was successful varies from district to district, and from village to village. Annual TASAF reports highlight that village leaders played a crucial role in the sensitisation process by organising village meetings and that information dissemination was challenging in large districts with sparse populations.

In sum, the TASAF information campaigns relied heavily on village meetings, radio and television programmes and newspapers to reach eligible communities. However, elderly and disabled individuals' access to these primary sources of information is likely to be restricted in rural Tanzania, owing to their physical and socioeconomic constraints. Physical disabilities may prevent individuals from regularly attending village meetings. Even if they do attend those meetings (village meeting attendance is compulsory in most of rural Tanzania) [9], a lack of education may prevent them from effectively capturing information during those meetings[10]. To the extent that elderly/disabled individuals have higher illiteracy rates, they are also less likely to read newspapers and newsletters. Poor economic conditions may deny

[9] In many villages, village meeting attendance is compulsory for all citizens older than 18 years. Some demographic groups are exempt from this norm, such as women taking care of children and very old people.
[10] Given that village meetings in Tanzania are always conducted in Swahili, those not familiar with this language are unlikely to benefit from them.

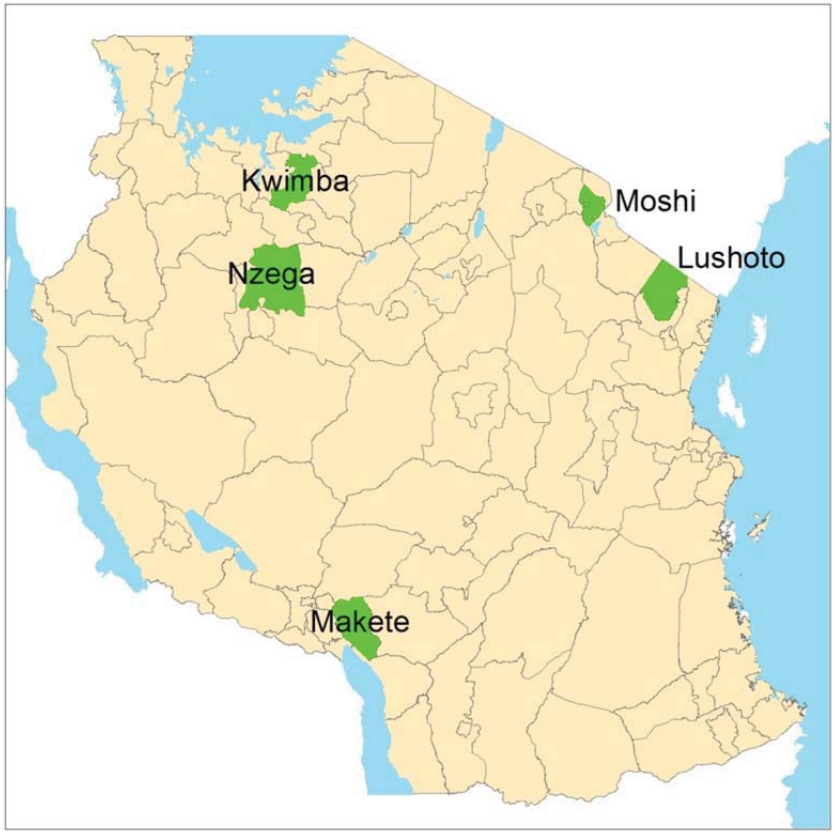
them access to radios and televisions. Social interactions with well-informed peers may well compensate for this lack of direct access to information.

3. Data

3.1. Data source

The data in this paper come from a listing exercise organised at baseline (April–November 2008) of an evaluation of the VG component of TASAF II, conducted by the World Bank. Twenty villages of five Tanzanian districts (100 villages in total) were randomly selected for the evaluation. The five districts visited were Moshi, Lushoto, Makete, Kwimba and Nzega (Figure 1).

Figure 1 Sampled districts (each containing 20 sampled villages)



Source: Baird et al. (2013).

Each of the 100 sampled villages had at least one VG selected for TASAF funding.[11] The deposit of the funds was pending until after the baseline survey visit and therefore no TASAF VG activities had been undertaken at the time of our survey visit.

Visits of *all* households living in these 100 villages resulted in a dataset of 61,610 households. A short listing survey was conducted of a subset of these households, namely *all* village elites (village chairperson and village executive officer) *all* VG households (including the elderly/disabled households) and a randomly sampled subset of the non-VG households. The resulting sample contains 35,870 households. For 15% of these households, we have missing data because of the absence of household members during listing.[12] The resulting dataset of 30,339 households contains 23,865 VG households in total, of which 19,916 are elderly and/or disabled households. Appendix A.1 summarises the population and sample characteristics.

The data on this sample include details on household head characteristics (e.g. education), basic consumption data, asset ownership and social/political connectedness of the household members. They also include the answer to the question 'Have you personally ever heard of a development programme called TASAF?' We use the answer to this question (yes or no) as the main proxy for a household's TASAF awareness. In section 6.2 we discuss the limitations of this measure and test the robustness of our results to an alternative proxy.

We also collected GPS data for all dwellings (latitude, longitude and altitude), which allows us to map all households and to calculate distance measures. Automatic capturing of the GPS data ensures high quality of these data (Caeyers and De Weerdt, 2012).

3.2. Peer group measure

Nearest neighbours set

In the social interaction literature, peer groups have been defined in various ways, such as farmers living in the same district (Case, 1992), households living in the same village

[11] Less than 5% of the households in the sample contain TASAF VG beneficiary members. This relatively low number of TASAF beneficiaries does not mean that TASAF II was not considered an important development programme. By sampling design, all villages in the sample are those that ended up selecting VG activities as development initiatives to be funded by TASAF. The latter activities tend to benefit only a relative small number of households. However, the decision to select TASAF VGs as the type of activities to be funded by TASAF was the result of a long decision process, in which VG activities were balanced against projects improving basic social and market services, such as building schools and roads and public work programmes. The stakeholders of the last two project types are likely to be higher in number than those of VG activities. It was therefore in any community member's interest to be informed about TASAF.

[12] This attrition rate is unlikely to be random, because the chances for mobile households to be absent at an unexpected visit are high relative to less mobile households. This means that attrition naturally leaves our sample with the least mobile vulnerable households in the village, which – for reasons discussed below – is in fact beneficial for the estimation strategy used in this study.

(Munshi, 2004; Wydick et al., 2011), people living in the same neighbourhood (Glaeser et al., 1996; Aaronson, 1998; Bayer and Ross, 2009; Wydick et al.), students in the same school (Evans et al., 1992) and roommates living in the same dormitory (Sacerdote, 2001).

Most of these definitions are ad hoc and rely on strong assumptions about the way people interact with each other (Manski, 2000). Without asking individuals directly about the set of people they interact with, as is done in Conley and Udry (2010), De Weerdt and Dercon (2006) and Van den Broeck and Dercon (2011), it is generally difficult to convincingly define peer groups. Reasonable exceptions to this are situations in which individuals are somehow constrained in their interactions. An example of this is the study by Bayer et al. (2009) on the influence of juvenile offenders on their peers in juvenile institutions. Given the legal constraint on the freedom of movements of the subjects in this study, it is reasonable to define peers as individuals who serve time in the same correctional facility, despite the lack of detailed social network data available. In a similar vein, the peer group definition in this paper relies on the physical and socioeconomic constraints faced by elderly/disabled households in rural Tanzania. We argue that such households are restricted to their geographically nearest neighbours for information on development initiatives.

In the developed world, there is a legitimate claim that geographic distance has become less restrictive for social relationships. In places such as rural Tanzania, however, geographic location matters, especially for people with physical disabilities and socioeconomic constraints. In our sample of 19,916 elderly/disabled households, only 2% own a motorbike, car or truck. Despite the proliferation of mobile phones throughout the continent in the previous decade, 76% of the elderly/disabled sampled households do not own a mobile phone; 44% of them do not own a radio; and 43% of them are headed by a person who has not had any form of education, which makes them less likely to listen to radio programmes or to read newsletters and leaflets about TASAF. As a result, the majority of these groups heavily rely on their peers within the village for access to information.

If we further assume that elderly/disabled households need frequent interactions with well-informed people for information to be effectively transmitted to them, the assumption underlying the nearest neighbours set definition is not that stringent. Its validity only requires elderly/disabled households to *mainly* interact with their nearest neighbours about development programmes such as TASAF. Note that we do allow for some form of interactions with individuals other than neighbours, such as irregular meetings at the village market or greetings on the street, as long as these do not result in effective information exchange about TASAF.

As land traditionally is passed on from father to son in these areas, people tend to live near to their relatives. Therefore, using the nearest neighbours set as reference group has the additional advantage of including most of the household's blood relatives as peers. This is important, as qualitative evidence suggests that people in rural Tanzania mainly discuss welfare programmes with their blood relatives.[13] The fact that peer groups systematically include blood relatives may cause problems of unobserved correlated effects, however. Our identification strategy has to take this into account (see below). To allow for relatives living further away, we control for whether or not the household has blood relatives living in the village.

To address the concern that the respondent who answered the question 'Have you personally ever heard of a development programme called TASAF?' (our measure of TASAF awareness) was not necessarily the vulnerable person (and therefore not necessarily less mobile), in section 6.2 we test whether our results are robust to restricting the sample to those households for which the survey respondent was the household head.[14] This test also addresses the possibility that elderly/disabled individuals may have heard about TASAF from more mobile members within their household. Since information tends to reach households via the household heads in Tanzania, restricting the sample to these individuals limits the likelihood of contamination by mobile members obtaining the information elsewhere.

Distances are calculated using three-dimensional Euclidean distance formulae between GPS locations of households, in which we also consider the elevation level of the household's dwelling. Given the specific survey sampling strategy applied during collection of data used in this study (see section 3.1), in which only a fraction of non-VG households were sampled whereas all VG households were sampled, we give each neighbour j in household i's neighbours set an importance weight, $\omega_{i,j}$, equal to the inverse probability of being sampled for the survey. When computing means of variables at the peer group level, we normalise the weights to sum to 1 within each group.

Base specification: 10 nearest neighbours

For reasons outlined next, in our base model we define the peer group by the set of 10 nearest (sampled) neighbours, restricted to those living within a maximum distance threshold of 1

[13] These contextual data were collected in November 2012 during an additional field trip organised in preparation of this paper.

[14] The data do not allow us to identify the elderly/disabled individuals within the household and therefore we cannot restrict the sample to such respondents. However, given that the elderly members tend to be the household heads in Tanzania, and given that the elderly also tend to have disabilities, we consider household head to be a reasonable proxy for the elderly/disabled individual in the household.

kilometre.[15] This neighbours set may include both vulnerable and non-vulnerable households. In the sample, the maximum distance of the tenth nearest sampled household ranges from 92 to 1513 metres (see Appendix A.1). By using the 10 nearest neighbours set, we risk excluding relevant households living nearby, or risk including households living too far. Our measure, just like any other peer group measure, is likely to suffer from measurement error. It is important to test the robustness of our results to alternative nearest neighbours set definitions, which we do in section 6.2.

Choosing the 10 nearest neighbours, instead of, for instance, the five, 15 or 20 nearest neighbours, minimises the sample size reduction. In our analysis, we need to drop households for which not all *k* nearest neighbours live within the 1-kilometre range. Moreover, the 2SLS estimation strategy that we use requires households with no *excluded* neighbours to be dropped (see below). The extent to which these restrictions lead to a reduction in the total sample size depends on the number of nearest neighbours considered. When we consider the 10 nearest neighbours, the sample reduction is minimal – less than 3%.[16] In the choice of the number of nearest neighbours, *k*, we also took into consideration the fact that controlling for hamlet fixed effects in our main regression models makes it in our interest to keep *k* small.

Finally, we should note that only a random subset of neighbours is captured in the sample. The nearest *sampled* tenth neighbour lives, on average, 310 metres away, which is approximately the same distance as the average nearest *actual* 20th neighbour lives in the true population, that is, 335 metres away (see Appendix A.1). This means that the set of 10 sampled neighbours is in fact a proxy for the set of 20 actual neighbours (this will be important when interpreting the results). Also note that 306 metres is similar to the 300-metre band limit that Van den Broeck and Dercon (2011) consider in their neighbours set definition using another Tanzanian dataset.

3.3. Data description

This study is concerned with some of the most vulnerable groups in Tanzania, namely the elderly and the disabled. Table 1 compares these VGs with other households, in terms of their

[15] We choose 1 kilometre as a threshold in view of the evidence provided in the longer household questionnaire that many vulnerable individuals find it difficult to walk more than 1 kilometre, and hence are less likely to interact with households living further away.

[16] This compares with a drop of 6% and 11% of sampled households for, respectively, the 15 and 20 nearest neighbours set peer group definitions. For the five nearest neighbours set, the drop is 2%, but, as we will highlight below, the exclusion restriction of the instruments we use is unlikely to be satisfied for such a small number of households.

baseline characteristics.[17] The elderly constitute 35% of the population, whereas households with disabled individuals make up 6% of the population.[18] Overall, elderly and disabled households are significantly worse off than non-vulnerable households. Although their housing conditions are similar, these VGs own significantly fewer assets (radio, phone, vehicle and watch or clock), eat meat and eggs on a less frequent basis, are significantly less educated, and the elderly are less likely to have heard about TASAF.

Almost half of the elderly/disabled households attended a village meeting at least once in the 3 months prior to the survey.[19] This number is not significantly different from that of the non-vulnerable group. As mentioned above, this may be the result of the fact that village meeting attendance is not necessarily voluntarily in Tanzania. Moreover, the data were collected a few months after TASAF VG beneficiaries had been selected. As explained in section 2.1, the beneficiary selection could only go ahead if at least 50% of the VG community members were present at the village meeting. Qualitative discussions with villagers indicate that village officials made significant efforts (sometimes using force) to mobilise the VG population in order to reach that quota. As we will see next, despite the relatively high number of elderly/disabled households that attended village meetings, many of those were not informed about TASAF, possibly as a result of socioeconomic constraints (as discussed in section 2.2).

At the time of the survey, only 53% of elderly households and 57% of disabled households were informed about TASAF (Table 1). Whereas 60% of the elderly/disabled households that attended the meetings were informed, only 47% of the absentees were informed (not reported in the tables). We make two observations here, as follows. (i) Village meetings are not the only means of informing households, given that almost half of the households that did not attend were also informed. This finding suggests some scope for social interactions, with people attending village meetings passing on information to the absentees. (ii) Village meetings are not entirely successful in informing households, as 40% of those attending were not informed. This is possibly the result of a lack of education and related languages problems.[20] This means that, even for those attending village meetings,

[17] Note that the 'other households' include other types of vulnerable households, such as widow(ers).
[18] Note that these vulnerable groups are not mutually exclusive. That is, the elderly group contains disabled households and the disabled group contains elderly households.
[19] Most village meetings concerning TASAF II took place within the 6-month period prior to the survey visit.
[20] When we look at the proportion of informed village meeting attenders by education category, we find, indeed, that the uneducated households are least likely to benefit from attendance at village meetings. Other reasons provided by the villagers during qualitative discussions include reporting bias, that is, some people may lie about village meeting attendance because they are afraid of being punished if they are not attending. Another reason provided is that some of them are not at all interested in the meetings, and only attend because attendance

there may be some scope for regular neighbourhood interactions to enhance the transmission of information.

Table 2 compares the characteristics of informed and uninformed elderly/disabled households. The first group is, on average, wealthier in terms of assets, amenities and consumption patterns. A higher proportion of them have an educated head (71% versus 42%); they have more household members; and they are less likely to be female headed (25% versus 39%). They are almost twice as likely to be blood related to the village elite and more than four times as likely to hold political office. On average, compared with their uninformed counterparts, informed households live significantly closer to the village leader (by almost 400 metres; note from Appendix A.1 that the average distance between sampled households in the village is 1800 metres), and they have a higher proportion of informed neighbours (six out of 10 compared with five out of 10).

4. Empirical strategy

4.1. Empirical specification

Since the pioneering work of Manski (1993), the probability of observing an individual behaving in one way or another is modelled as a function of her own characteristics (*individual effects*), of the exogenous characteristics of her peers (*contextual peer effects*), the prevalence of this behaviour among her peers (*endogenous peer effects*) and some unobserved characteristics of the individual's environment (*correlated effects*). Here, endogenous peer effects refer to the influence of one household's TASAF information set on its neighbouring household's TASAF awareness, through word of mouth. This is the effect this study aims to identify. *Contextual peer effects* refer to the direct effect of neighbours' characteristics on a household's probability of being TASAF informed. For instance, if neighbours have a radio that broadcasts information about TASAF, and if neighbours listen to that radio together, then the correlation in TASAF awareness would be the result of contextual social effects rather than of being driven by endogenous peer effects. Lastly, *correlated effects* capture the possibility that households living in the same neighbourhood may behave in the same way simply because they have similar characteristics, or because they face similar conditions. Such correlated effects include sorting of similar households into specific neighbourhoods. Correlated effects are discussed more extensively below.

---

is compulsory. As a consequence, 'they are drunk and/or they fall asleep and/or chat with their friends during the meetings' (according to respondents to qualitative surveys).

Formally, we follow Manski's (1993) popular linear-in-means model of social interactions, incorporating these effects into one empirical framework:

$$I_{lji} = \alpha + \beta \bar{I}_{P_{lji}} + x_{lji}\gamma + \bar{x}_{lji}\delta + \varepsilon_{lji} \quad (1)$$

where I_{lji} is a dummy for whether or not an elderly/disabled household *i* in hamlet *j* in village *l* had ever heard of TASAF at the time of the survey. Assume that each household has a neighbours set P_{lji}, defined by the set of neighbours potentially affecting household *i*'s TASAF awareness, with $i \notin P_{lji}$. The neighbours set P_{lji} may include both vulnerable and non-vulnerable households. The term $\bar{I}_{P_{lji}}$ is the weighted average information set of the households in P_{lji}.[21] Its coefficient β measures the endogenous peer effect of interest in this study. Vector x_{lki} contains a set of socioeconomic individual household characteristics, described below. Vector $\bar{x}_{lji}$ is the weighted average of *x* over the neighbours in P_{lki}. Its coefficient δ measures the contextual peer effects. Finally, parameter ε_{lji} captures the unobserved correlated effects, which we discuss in more detail below.

Vector x_{lki} controls for the following household characteristics: asset ownership (radio, phone, vehicle, watch/clock, sofa); amenities (improved roof,[22] improved latrine[23]); basic consumption (whether or not the household had eaten meat or drunk milk in the 7 days prior to the survey); whether or not the household head had had primary education; the household's degree of political connectedness (whether or not any household member held political office in the village, and whether or not any household member was related to the political elite); its degree of political activity (two dummies for whether or not any household member attended a village meeting in the previous 6 months and previous 3 months prior to the survey); its family ties in the village (whether or not the household had any relatives living in the village); and a set of household demographics (household size, age and gender of the household head, a dummy for whether the head belonged to the main tribe and main religion of the village, and a dummy for whether the main occupation of the household head was in the agricultural sector).

Given that the household information was collected during a listing exercise, in which households were visited with no prior notification, in some incidences information was provided by a household member other than the household head, if the latter was not present

[21] As noted in section 3.2, the observations are weighted by the probability of being sampled.
[22] Improved roofs are made of iron, concrete, cement, roofing tiles or asbestos.
[23] Improved latrines include covered or ventilated latrines.

during the interview. The other member was either the spouse of the head or another adult household member. This happened for 40% of the elderly/disabled households. Given the possibility that the household head is better informed about welfare programmes, we control for whether or not the survey respondent was the household head.

4.2. Empirical challenges and identification strategy

Besides the difficulty in defining relevant peer groups (see section 3.2), the identification of social interactions encounters other well-known challenges extensively discussed in the literature (Manski, 1993; Brock and Durlauf, 2001, 2007; Moffitt, 2001; Graham and Hahn, 2005; Soetevent, 2006; Graham, 2008). This section discusses these problems and explains how this paper deals with this.

Reflection problem

The reflection problem, first highlighted by Manski (1993) and further discussed by e.g. Moffitt (2001) and Blume et al. (2010), is the challenge of disentangling contextual and endogenous effects in a linear-in-means model of social interactions. The problem arises because a household and its peers are likely to simultaneously affect each other. Given that outcomes are observed in equilibrium, it is difficult to know whether an individual's outcome is the cause or the effect of the peer's outcome.

Bramouille et al. (2009) and De Giorgi et al. (2010) developed a strategy that allows one to estimate endogenous and contextual peer effects in the context of 'partially overlapping reference groups'. Since its development, this strategy has been frequently used in the non-experimental empirical literature on peer effects (Guiso and Schivardi, 2007; Bayer et al., 2009; Calvo-Armengol et al., 2009; Mas and Moretti, 2009; Helmers and Patnam, 2012; Collin, 2013). Peer groups of two observations are partially overlapping if they do not perfectly coincide. Such social networks ensure the occurrence of *intransitive triads,* which are sets of three households i, j and k, such that i is affected by j and j is affected by k but i is not affected by k. Bramouille et al. (2009) show how the presence of such intransitive triads allows one to deal with the reflection problem: the characteristics of 'excluded peers', that is, peers of peers who are neither one's own peers nor oneself, can be used as instruments for the average outcome of the peers. In the example of the intransitive triad, the characteristics of k can be used as instruments for the outcome of j, because they only have an indirect effect on the outcome of i through the effect on the outcome of j.

In our context, the set of nearest neighbours of household A is unlikely to perfectly coincide with the set of nearest neighbours of neighbouring household B. The presence of such intransitive triads allows us to use the method suggested by Bramouille et al. (2009) and De Giorgi et al. (2010). Specifically, we use the average exogenous characteristics of a household's neighbours' neighbours set as identifying instruments for $\bar{I}_{P_{lji}}$.[24] The validity of the exclusion restrictions will be discussed below.

Correlated effects

We rewrite equation (1) to specify the different types of unobserved effects that are likely to be present in our particular study:

$$I_{lji} = \alpha + \beta \bar{I}_{P_{lji}} + x_{lji}\gamma + \bar{x}_{lji}\delta + \eta_l + \chi_{lj} + \theta_{P_{lji}} + \mu_{lji} + \omega_{lji} \qquad (2)$$

where ω_{lji} is assumed to be a random independently and identically distributed component, that is, $E\left[\omega \mid \bar{I}_P, x, \bar{x}\right] = 0$. It is reasonable to assume that unobserved covariates at the household level, μ_{lji} (such as a household's interest in politics), are not correlated with observed outcomes at the neighbours set level, and therefore we do not expect the presence of μ_{lji} in equation (2) to bias the results. However, error components η_l, χ_{lj} and $\theta_{P_{lji}}$ capture three types of unobserved covariates that are potentially correlated with the peer group outcome $\bar{I}_{P_{lji}}$:

1 η_l: unobserved village covariates – residents of a specific village are expected to be jointly affected by various village-specific factors that contribute to TASAF awareness. As highlighted in section 2.2, village leaders played a crucial role in the dissemination of information about TASAF procedures. Some village officials may have been more effective in organising village meetings than others. Moreover, the village officials were informed by district officials, whose work may have been more productive in some villages than in others. Such intravillage differences are unobserved to the researcher and are captured by η_l.

[24] Note that, in contrast to the claim made by De Giorgi et al. (2010), even when the SEM is identifiable thanks to the presence of excluded peers, it is still necessary to instrument for the endogenous effect, even in the absence of correlated effects. De Giorgi et al. (2010, page 18) state that '[…] I will discuss how to exploit such excluded peers in an IV strategy that addresses the potential endogeneity due to correlated effects, but it is important to clarify now that […] we would not need to instrument if the only identification issue were reflection'. This is because, even if the model is identifiable using excluded peers as instruments, it can easily be shown by writing out the reduced forms that the error term in the structural equation is correlated with the endogenous explanatory variable in SEM and that OLS would yield biased and inconsistent estimates.

2 χ_{lj}: unobserved hamlet covariates – within each village, households living in the same hamlet may be subject to similar information shocks. There are 705 hamlets in total across the 100 sampled villages, with an average of seven hamlets per village. The average hamlet in the sample contains 28 elderly/disabled households (see Appendix A.1). Some hamlets are more remote than others, making it more difficult for their inhabitants to attend village meetings. Most hamlets in Tanzania have a hamlet leader, who is empowered to organise meetings at the hamlet level, or to go from door to door to inform people within his/her hamlet. Also, it is possible that households with similar unobserved characteristics sort themselves into specific hamlets. Such common unobserved characteristics are also captured by χ_{lj}. We asked all households in the sample whether or not they were blood related to specific village elite. These data allow us to look at the spatial spread of blood relatives of the village elite across the village.[25] Analysing these data, we find evidence in support of residential sorting based on family ties: 50% of the local blood relatives of a village elite kin live within the same hamlet as that kin. Given that there are on average seven hamlets per village, this is a remarkable result.

3 $\theta_{P_{lji}}$: unobserved neighbours set covariates – households may have characteristics in common with the households in their neighbours set that affect both their own degree of TASAF information and that of their neighbours. For instance, households living near a noticeboard may each be more likely to be informed about TASAF, independently of the presence of peer effects. To the extent that households sort themselves into neighbours sets *within* the hamlet, vector $\theta_{P_{lji}}$ also includes common characteristics that are driven by residential sorting. In the context of Tanzania's villagisation movement in the 1970s, most rural land was assigned to households by the government. Since then, plots have been passed on from generation to generation, usually from father to son. Most of the older rural villagers therefore have little choice as to who their neighbours are, except by donating parts of their plots to their children. This means that if at all households sort themselves within hamlets, this is likely to be at the nearest neighbours set level. Using again the GPS data on the

[25] On average in a village, there are four (sampled) households blood related to the village leader and two (sampled) households blood related to the Village Executive Officer. On average per hamlet, there are three (sampled) households blood related to the hamlet leader.

dwellings of the blood relatives of the village elite, we find that on average two out of 10 of their nearest (sampled) neighbours are, indeed, blood related to themselves.

To deal with correlated effects at the village level (η_l) we include village fixed effects in our empirical specifications. The unique size of the dataset also allows us to deal with correlated effects at the hamlet level, χ_{lj}, by controlling for hamlet fixed effects (in specifications in which we drop the village fixed effects), something that is rarely feasible in other studies owing to insufficient variation within hamlets.[26]

Focusing on within-hamlet variation in households' TASAF awareness also deals with the bias caused by residential sorting into hamlets. Assuming that controlling for hamlet fixed effects captures all shocks simultaneously faced by a household and its excluded peers, instrumenting $\bar{I}_{P_{lji}}$ by the average characteristics of the excluded peers also deals with correlated effects at the neighbours set level. These latter effects include those resulting from possible residential sorting *within* hamlets that is driven by passing on land from generation to generation.

Our 2SLS identification strategy crucially relies on the validity of the exclusion restrictions. Specifically, it relies on the assumption that excluded neighbours do not affect an elderly/disabled household's awareness of TASAF other than through passing such information on to their own neighbours. Whether or not this identifying assumption is valid depends on whether or not the peer group is appropriately defined. We refer back to section 3.2 for a more detailed motivation for the choice of the *k* nearest neighbours as the elderly/disabled households' relevant peers for effective TASAF information exchange. In sum, we argued that effective transmission of information about uncommon activities such as new CDD programmes (as opposed to more familiar matters such as family or agriculture) requires frequent interactions with well-informed individuals, whom one can trust. Constrained by their physical and socioeconomic conditions, most elderly and disabled in rural Tanzania rely on their geographically nearest neighbours for repeated information exchange. This is especially true in the context of rural Tanzania, where people's blood relatives tend to live nearby. Although VGs surely also engage in interactions of some sort with other villagers, such as chats on their monthly visits to the village markets, we do not expect such infrequent interactions to significantly affect their knowledge about TASAF.

[26] Controlling for hamlet fixed effects is more effective in our study than adding component fixed effects (which is commonly done in the literature) since the average component in our study is much larger than the average hamlet.

Of course, the validity of these assumptions depends on the actual degree of mobility of the elderly and disabled observations in our sample. In section 6.2 we test the robustness of our results to the dropping of observations that are possibly less geographically restricted. Moreover, in section 5.2 we test the results against various neighbours sets sizes. This is important, since the validity of the exclusion restrictions depends on the choice of *k*: defining the neighbours set as the 10 nearest neighbours when the 15 nearest neighbours set is in fact the relevant neighbours set would invalidate the identifying assumptions. Section 6.2 addresses a few other identification concerns.

Exclusion bias

In the analysis below, a household is systematically excluded from its peer group's average outcome calculations. In view of the discussion in Caeyers (2014), we therefore expect a downward exclusion bias to manifest itself in the OLS estimation results. Furthermore, as peer groups in this application are defined based on geographic distance, their formation is bound to be correlated with hamlet clustering. Caeyers (2014) demonstrates how such empirical structure significantly amplifies the extent to which the exclusion bias affects the OLS results in models adding cluster fixed effects. Her exposition also shows that the 2SLS estimation strategy outlined above (using exogenous characteristics of excluded peers as instruments while controlling for the household's own exogenous characteristics) successfully eliminates this bias.

Spatial dependence of observations

Results of spatial econometrics teach us to adjust standard errors for spatial dependence of observations. However, generalized method of moments (GMM) estimations such as Conley's routine (1999), which allows for arbitrary spatial correlation between observations, are computationally very cumbersome for datasets with 19,392 observations. Therefore, in our main specifications, we ignore spatial dependence within hamlets and report standard errors that are clustered at the hamlet level. In section 6.1 we test the robustness of our results to the controlling for spatial dependence within hamlets, by randomly selecting 20 out of the 100 villages in the sample and running regressions using Conley's (1999) GMM estimation on the reduced sample. We refer to section 6.1 for more details about this method.

5. Results

5.1. Average peer effects

Table 3 presents the OLS estimation results of the basic model (equation (1)), in which we define the peer group as the set of 10 nearest neighbours. Column (1) shows the results of a regression without any controls. In each of the remaining columns (2)–(4), the regressions include the full set of a household's own characteristics, vector *x*. Column (2) shows the OLS estimation results of a village fixed effects regression excluding the contextual effects. All coefficient estimates of the household's own characteristics have the expected signs. Households that attended at least one village meeting in the previous 6 months are 8 percentage points (pp) more likely to be informed. Households that attended a village meeting in the previous 3 months are an additional 4 pp more likely to be informed. Moreover, educated households are, *ceteris paribus*, 13 pp more likely to have heard about TASAF, and those that hold political office are 16 pp more likely to be informed. Female-headed households are 8 pp less likely to be informed about TASAF. Interestingly, survey respondents who are head of the household, as opposed to the head's spouse or any other adult household member, are 14 pp more likely to have heard about TASAF. This confirms our earlier conjecture that household heads are generally more informed about development programmes than any other adult members of the household. Intra-household communication about development programmes clearly is imperfect. The OLS estimate of the endogenous peer effect in column (2) suggests that, for each additional informed neighbour in its set of 10 nearest sampled neighbours, a household is approximately 3.9 pp more likely to be informed about TASAF. Controlling for (observable) contextual social interaction effects in column (3) slightly increases the OLS estimate of the endogenous social interaction effect to 4.2 pp.

Controlling for hamlet fixed effects in column (4), in addition to controlling for contextual effects, significantly reduces the endogenous peer effect to approximately 1 pp per additional informed sampled neighbour.

As discussed in section 4, we have good reasons to believe that the OLS estimates are biased and inconsistent. Table 5 below addresses this bias by instrumenting the endogenous average information set of the peers by the exogenous characteristics of the excluded neighbours. Although the results remain similar when we include the complete set of characteristics of excluded neighbours as instruments (not reported), the table shows the results for regressions only including the instruments that are significant at the 5% level in the 2SLS first-stage regression. The resulting set of four instruments includes: a dummy for

whether the household had an improved roof; whether the household was female headed; whether the household had relatives in the village; and whether the respondent to the survey was the household head.

Table 4 shows the first-stage regression results for the 2SLS regression controlling for hamlet fixed effects. Each instrument is significant at either the 1% or 5% significance level. The current literature lacks concrete results for testing for weak instruments when errors are not identically and independently distributed: on the one hand, the Craig–Donald (1993) (C-D) Wald F-statistic is not appropriate, as it relies on the assumption of homoscedastic errors; on the other hand, when using the Kleibergen–Paap (2006) (K-P) rank Wald F-statistic, although it does adjust for heteroscedasticity, autocorrelation and clustering of standard errors, it is difficult to interpret its magnitude, as its critical values have not yet been defined.[27]

For these reasons, Table 4 shows both the C-D and the K-P test statistics. The C–D F-statistic is 26.54, which is above the Stock–Yogo (2005) critical value for 5% maximal instrumental variables (IV) bias toleration relative to OLS and above the Stock–Yogo (2005) critical value for 10% maximal IV size distortion. Using the same Stock–Yogo (2005) critical values, the K-P statistic is above the 20% maximal IV relative bias and just below the 25% maximal IV size threshold. A *p*-value of 0.00 associated with a Kleibergen–Paap rank LM statistic of 29.03 implies a strong rejection of the null of underidentification of the model.

In our preferred specification, column (2) in Table 5, the estimated endogenous peer effect indicates an increase of 7.7 pp in the probability of being TASAF informed, per additional informed sampled neighbour in the neighbours set.

This means that an elderly/disabled household that has five out of 10 sampled neighbours who are informed (e.g. because they were able to attend village meetings) is, *ceteris paribus*, 38 pp more likely to be informed itself than households living among uninformed households. This endogenous peer effect is quite significant in magnitude, especially when compared with the effect of a household's own characteristics on TASAF programme awareness. For instance, education and village meeting attendance increase the probability of being TASAF informed, *ceteris paribus*, by, respectively, 12 and 14 pp.

Up to now, we have interpreted the social interaction results in terms of the effect of the average characteristics of the 10 nearest *sampled* neighbours. In section 3.2, we discussed why the set of 10 nearest *sampled* neighbours is actually a proxy for the 20 *actual* nearest

[27] Its critical values have not been computed as they depend on the type of violation of the independent and identically distributed assumption, which in turn depends on the nature of the study (Bazzi and Clemens, 2009).

neighbours in the true population. This means that, in terms of the true population, we should actually interpret the results in column (2) of Table 5 as follows: each additional informed household in a vulnerable household's 20 nearest neighbours set increases its probability of being informed about TASAF by almost 4 pp. However, given that the main purpose of this paper is to identify the importance of neighbourhood peer effects, rather than determining the exact number of nearest neighbours that matters, this difference in interpretation is not all that crucial for the main findings of the study.

Moreover, it should be noted that Table 5 presents average peer effects. In spatial social interactions models, however, the individual peer effect is expected to depend on the geographical distance between the residence of the household and that of its peers. In our dataset, this distance ranges from 16 meters up to 969 meters (note that we only consider peers who live within 1 km range in our analysis), with an average distance of 186 meters. In general, we expect the peer effects to die out with distance. In Appendix A.2, we present the results of a regression adding an interacted term between the peer effect of the ten nearest neighbours and the average distance of those ten nearest neighbours. The estimation results confirm, indeed, that the peer effect becomes significantly smaller as the distance of the peers grows larger.

5.2. Varying the size of the neighbours set

Table 6 presents the 2SLS coefficient estimates of the endogenous peer effect equivalent to regression (2) in Table 5, for the five, 10, 15 and 20 nearest neighbours set specifications. As noted above, we drop households for which not all k neighbours live in the 1-kilometre range and also those that do not have any excluded neighbours. For comparison purpose, we run the four different regressions on the same sample, that is, one that contains only households with suitable neighbours sets for each value of k.

Whereas the magnitude of the endogenous peer effect (per neighbour) is fairly similar for the ten, fifteen and twenty nearest neighbours specifications, the effect more than quadruples when we consider the five nearest neighbours. Although this result could possibly indicate that the five nearest neighbours are most effective in transmitting information about development programmes (especially given that they tend to include blood relatives of the household), the Hansen J p-value of 0.075 and the negative adjusted R^2 suggest that the exclusion restrictions may not be valid in the five nearest neighbours specification. As discussed in section 4.2, if peer groups are too narrowly defined, the identifying assumptions

may be violated. For specifications assuming larger social interaction bands (i.e. k = 10, 15 or 20), the Hansen J p-values are consistent with valid exclusion restrictions.

Unsurprisingly, the larger the neighbours set we consider, the smaller the C-D Wald F-statistic and the K-P Wald F statistic.[28] Weak instruments lead to lower precision of coefficient estimates, which may explain the finding that the significance level decreases for higher values of k, despite the magnitude of the effect remaining similar.

5.3. Heterogeneous peer effects

This section examines the presence of heterogeneity in endogenous peer effects on programme awareness. Specifically, the aim is to identify (i) the type of disabled/elderly households that are most and least responsive to information in the neighbourhood and (ii) the type of well-informed neighbours who are most effective in transmitting information to disabled/elderly households.

The answers to both these questions possibly depend on the degree of socioeconomic similarity between a household and its neighbours. The literature on peer group formation has identified gender, age, tribe, religious affiliation, wealth, occupation, education level and kinship as characteristics other than proximity that matter for the formation of information links (Conley and Topa, 2002; Jackson, 2008; Conley and Udry, 2010; Van den Broeck and Dercon, 2011). Although this literature has focused on other types of information links, such as relationships among African farmers sharing information about agricultural technologies, it is useful to start from the same set of characteristics and investigate which ones also matter in the diffusion of information about development programmes. We also consider the extent to which the holding of political positions in the village matters for information diffusion through word of mouth.

Table 7 reports on the estimation results when interaction terms are added to the main specification in column (2) of Table 5.[29] In panel A, we interact the endogenous peer effect with disabled/elderly household i's *own* socioeconomic characteristics. Panel B shows the results of regressions adding interactions of the endogenous peer effect with the same set of socioeconomic characteristics but now with those that apply to the average of the 10 nearest neighbours. In Panel C, we add a triple interaction term, in addition to the bivariate

[28] The larger the neighbours sets, the higher the probability that the instruments rely on excluded neighbours for which the information transmission effect is relatively weak.

[29] Instruments for the interaction terms are computed by taking the product of each of the main instruments and the interaction term of interest. Table 7 shows the results when the interaction terms are added individually to the base regression. The results remain similar, however, in a model jointly adding all interaction terms (not reported).

interactions added in panels A and B (controlling for the interaction between the household's own characteristic and that of the average neighbour; not reported).

Households most responsive to information in the neighbourhood

The results in panel A of Table 7 suggest that gender, tribe, phone ownership and the holding of political positions in the village do not matter for the extent to which a household benefits from the availability of information in the neighbourhood. However, the results in panel C show that gender does matter when you take the characteristics of the neighbourhood into account: female-headed households benefited more from the presence of well-informed neighbours if they lived in an area with a relatively high proportion of female-headed households.

The results also indicate that elderly/disabled households that are headed by very old members are less likely to effectively capture information about TASAF through endogenous peer effects.[30] This may be due to memory problems related to very old age, or it may be due to very old people being less likely to go out and interact with other people in the neighbourhood.[31] Moreover, the results show that, in poorer neighbourhoods, information on TASAF was less likely to reach wealthier households through social interactions than their non-wealthy counterparts (with wealth being proxied by phone ownership and the number of times a household had eaten meat).[32] It is possible that wealthy individuals living in economically deprived areas prefer to obtain information through other means (e.g. by attending village meetings or reading newspapers). In predominantly wealthy neighbourhoods, on the other hand, wealthy households seem to have been as likely to obtain information on TASAF through interactions with neighbours as their more deprived counterparts.[33]

Households that belong to the main religious group in the village and/or that are headed by educated and/or employed individuals were significantly more likely to benefit from the presence of TASAF-informed households in the neighbourhood. This result is independent of the area in which these households lived. Similarly, households with blood relatives living in

[30] The interaction term in this column makes use of the continuous age of the household head. Recall that the sample mainly consists of households with heads who are older than 55 years; therefore, this result applies to the oldest of the elderly.
[31] Whereas the sampled units of interest mainly consist of elderly households only, their neighbours may include both young and elderly headed households. This means that an interaction with a dummy for whether or not the neighbours are also headed by elderly households, the results of which are presented in column (2) of panel B in Table 7 and discussed below, are more meaningful than adding triple interactions. We therefore omit the respective triple interactions in panel C of Table 7.
[32] Similar results apply when we use sofa ownership or an improved roof as proxies for wealth (not reported).
[33] This is tested by using a joint F-test on the sum of the estimates of the three interaction terms.

the village were more likely to be informed through neighbourhood effects. The latter finding may be a result of residential sorting based on family ties (see above). That is, people who have relatives living in the village are more likely to have relatives among their neighbours, and, as a result, are more likely to interact with them. If this is true, we would expect this result to apply only to households whose neighbours *also* claimed to have blood relatives living in the village. This can be verified by adding triple interactions, which we do in panel C of Table 7. The results, shown in column (9), are, indeed, consistent with this hypothesis.

Neighbours most effective in the transmission of information

The results in column (9) in panel B of Table 7 show that households that have blood relatives living in the same village were more effective in disseminating TASAF information to their neighbours. As discussed above, this result is conditional on the household itself having blood relatives living in the village. As expected, households holding a political position in the village were significantly more likely to spread information in their neighbourhood (column (10)). Each additional political official in the neighbours set adds 1 pp to the endogenous social interaction effect.[34] However, the results in column (13) in panel C show that this only holds for households that do not hold a political office themselves (possibly because those households have other means of being informed and do not need to rely on local social interactions).

6. Robustness tests

6.1. Correcting for spatial dependence of standard errors within hamlets

In section 4, we explained that the use of estimation methods that control for spatial dependence within hamlets, such as the routine suggested by Conley (1999), becomes computationally cumbersome when the sample size is very large. However, the presence of such spatial dependence of standard errors may affect our inference results. To address this concern, we randomly select 20 out of 100 villages in the sample and run regressions using Conley's (1999) GMM estimation to adjust for spatial dependence in the reduced sample of 3685 observations. The Conley (1999) method allows for arbitrary spatial correlation between observations that is decreasing with distance and is zero beyond a specified cut-off point. We use the average distance to the tenth neighbour as the cut-off value.

[34] In the data, the number of political neighbours within the 10 nearest neighbours set ranges from zero (57% of the households) to six (less than 1% of the households), with 10% of the households having two political neighbours.

Using the reduced sample, Table 8 compares the results of a model in which we do correct for spatial dependence in standard errors within hamlets using the Conley (1999) method with a model in which we do not. In the latter model, we adjust the standard errors for clustering at the hamlet level. We find that the adjusted standard error of the coefficient estimate of the endogenous peer effect is almost identical in both cases. This suggests that, by clustering standard errors at the hamlet level in our main specifications, we are dealing with most of the spatial dependence of observations.

6.2. Other identification challenges

As discussed in section 4, for our 2SLS strategy to be valid, excluded neighbours should not affect a household's information set about TASAF other than through passing information on to their own neighbours. However, one might consider that not all elderly and disabled households are constrained to their nearest k neighbours for effective information exchange about development programmes such as TASAF. Some of them may be more mobile and frequently interact with people in the village living further away. Such observations would invalidate our IV instruments and consequently bias our 2SLS estimates upwards.

To address this concern, we test whether our results hold up to the dropping of observations for which the exclusion restriction is less likely to hold. As mentioned above, village meeting attendance in most villages in rural Tanzania is compulsory for all those older than 18 years. Elderly and disabled individuals who are physically too weak to attend are usually exempted. This means that restricting the sample to those households that did not attend village meetings is expected to retain the least mobile households. Similarly, the exclusion restriction is more likely to hold for the elderly/disabled households that do not own a vehicle (bicycle, car or truck) and those whose head is relatively old (e.g. older than 65 years). Moreover, as discussed in section 3.1, the data for this paper come from a listing exercise in which we visited all households in 100 villages in Tanzania. Because these listing visits were generally not announced to villagers, in 40% of the elderly and disabled households the household head was not at home at our survey visit (instead, we interviewed the spouse or another adult household member). Heads who were not at home during the surprise visit are expected to be relatively more mobile.

Column (1) in Table 9 copies the results from the unrestricted sample given in column (2) in Table 5. Columns (2)–(4) show the results for the set of samples restricted to those households we expect to be least mobile for reasons described in the previous paragraph. If the failure of the exclusion restriction, owing to the presence of mobile households, led to an

upward bias in our estimate of the endogenous peer effect in Table 5, we would expect the estimate of the endogenous peer effect in Table 9 to go down. Instead, we find that the estimates slightly increase, which is consistent with less mobile households being more reliant on neighbourhood interactions for access to information than mobile households.

Another concern is whether the answer to the question 'Have you personally ever heard of a development programme called TASAF?' is a good proxy for a household's degree of TASAF awareness. It could be argued that people, although knowing the activities of a programme, may have forgotten its name.[35] To address this worry, we use an alternative measure of TASAF awareness that is more specific about the activities that TASAF VGs supported. Specifically, we use as a dependent variable the answer to the question of whether the VG respondent had ever heard of any of the TASAF groups in the village, where we listed the names under which the groups had been registered when asking this question. The names of these groups usually referred to the activity (such as 'pig keeping for the elderly'). Column (6) in Table 9 shows that the estimate of the peer effect increases to almost 16 percentage points when using this alternative measure.

Lastly, our base measure of TASAF awareness is predicated on whether the respondents themselves had personally ever heard about the programme. If the respondent had not heard about TASAF, this does not mean that no one else in the household had, assuming imperfect intra-household transmission of information. As mentioned earlier, the household head is usually the best informed person in the household, and his/her information set is therefore the best proxy for the household's neighbours set. We already noted that the estimated endogenous peer effect is not reduced when we restrict the sample to those households for which the head was the survey respondent (see column (3) in Table 9).

7. Conclusion

For targeting of CBD programmes to be effective, it is crucial that information about the programmes reaches all members of the eligible communities. In practice, however, many development programmes suffer from low public awareness. In rural Tanzania, VGs such as the disabled and the elderly often have limited access to primary sources of information (such as village meetings and newspapers) as a result of their physical and socioeconomic

[35] Our survey visit took place immediately after TASAF groups had been formed, when the programme was still a very hot topic. When people discussed the programme, they would always use its name. It is therefore very unlikely that respondents who had never heard about the programme would not know the programme's name. Probing more about the programme in qualitative discussions with respondents who had never heard about the programme suggested that they did, indeed, not know anything about the programme.

conditions. The aim of this paper was to analyse the extent to which the presence of well-informed neighbours may compensate for this lack of direct access to information. We find that, on average, for each additional informed neighbour in a set of 10 sampled nearest neighbours, a disabled or elderly household was approximately 8 percentage points more likely to be informed about TASAF. To put this effect into context, it is useful to compare it with the impact of attending village meetings (i.e. the medium most heavily relied on by TASAF to sensitise the population): having two informed neighbours has a similar impact on the probability of being informed as has attending a village meeting (15% and 14%, respectively).

The finding of such significant endogenous peer effects implies the presence of large social multiplier effects in TASAF programme awareness. This result is important for anyone involved in development programme sensitisation campaigns. To illustrate, consider the effectiveness of a sensitisation intervention that distributes radios to a small subset of households in a village. In the absence of any peer effects, this policy would only affect the awareness of the beneficiary households. In the presence of contextual peer effects but no endogenous peer effects, the policy would not only affect the direct beneficiaries but also their direct peers, who benefit by listening to the radio of their peers. In the presence of an endogenous peer effect in addition to the contextual peer effects, the intervention would reach not only the beneficiaries and their direct peers but also the peers of these peers through word of mouth. Given the presence of endogenous peer effects, it becomes a more reasonable expectation for relatively cheap interventions, targeted only at subsets of a community, to substantially improve the awareness level of the entire population.

This paper not only identified the presence of endogenous peer effects. Through the unique size of the dataset, it also allowed us to gauge the profile of elderly and disabled households that are most and least responsive to information present in the neighbourhood. We found that, on average, the economically worse-off households seem to be those that benefit most from the presence of well-informed neighbours. This is possibly because better off households have better access to primary sources of information, such as radio and newsletters and, therefore, rely less on peers for information. The paper also identified specific groups that are less likely to be informed through social interactions and therefore should be targeted separately by complementary – specifically designed – sensitisation strategies. These groups include the oldest of the elderly members, female-headed households living in a male-dominated neighbourhood, those who do not have blood relatives living in the village, those who are uneducated and/or unemployed and/or those who do not belong to

the main religion of the village. On the other hand, the study identified households with members holding political positions as being relatively more effective in the information diffusion process.

We should note that the external validity of the findings in this paper is likely to be restricted to those CBD programmes that are targeted at elderly and disabled individuals. More generally, in addition to informing those in charge of designing sensitisation campaigns, the study may also be of interest to those involved with programme monitoring. It highlights the importance of regularly sampling households from different parts of a community during the programme sensitisation phase, to evaluate the extent to which information effectively spreads to the various sections of the community and to adjust sensitisation programmes accordingly.

References

Aaronson, D., 1998. Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *Journal of Human Resources* 33 (4), 915–946.

Baird, S., McIntosh, C. and Ozler, B., 2013. The regressive demands of demand-driven development. *Journal of Public Economics*, Elsevier, vol. 106(C), 27-41.

Bandiera, O. and Rasul, I., 2006. Social networks and technology adoption in northern Mozambique. *Economic Journal* 116 (514), 869–902.

Banerjee, A.V., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S., 2010. Pitfalls of participatory programs: evidence from randomized experiments in education in India. *American Economic Journal: Economic Policy* 2 (1): 1-30.

Bayer, P. and Ross, S., 2009. Identifying individual and group effects in the presence of sorting: a neighborhood effects application. Economic Research Initiatives at Duke (ERID) Working Paper No. 51. Duke University, Durham, NC. Available at http://ssrn.com/abstract=1602011 or http://dx.doi.org/10.2139/ssrn.1602011.

Bayer, P., Hjalmarsson, R. and Pozen, D., 2009. Building criminal capital behind bars: peer effects in juvenile corrections. *Quarterly Journal of Economics* 124 (1), 105–147.

Bazzi, S. and Clemens, M., 2009. Blunt instruments: a cautionary note on establishing the causes of economic growth. Working Paper 171. Center for Global Development, Washington, DC.

Bjorkman, M. and Svensson, J., 2007. Power to the people: evidence from a randomized field experiment of a community-based monitoring project in Uganda. Policy Research Working Paper Series 4268, The World Bank.

Blume, L.E., Brock, W.A., Durlauf, S.N. and Ioannides, Y.M., 2010. Identification of social interactions. Economies Series 260. Institute for Advanced Studies, Vienna.

Borjas, G.J., 1995. Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review* 85 (3), 365–390.

Bramouille, Y., Djebbari, H. and Fortin, B., 2009. Identification of peer effects through social networks. *Journal of Econometrics* 150 (1), 41–55.

Brock, W. and Durlauf, S., 2001. *Handbook of Econometrics*. Elsevier Science, Amsterdam, pp. 3299–3380.

Brock, W.A. and Durlauf, S.N., 2007. Identification of binary choice models with social interactions. *Journal of Econometrics* 140 (1), 52–75.

Caeyers, B., 2014. Exclusion bias in empirical social interaction models: causes, consequences and solutions. CSAE Working Paper WPS/2014-05. University of Oxford, UK.

Calvo-Armengol, A., Patacchini E. and Zenou, Y., 2009. Peer effects and social networks in education. *Review of Economic Studies* 76 (4), 1239–1267.

Case, A., 1992. Neighborhood influence and technological change. *Regional Science and Urban Economics* 22, 491–508.

Cipollone, P. and Rosolia, A., 2007. Social interactions in high school: lessons from an earthquake. *American Economic Review* 97 (3), 948–965.

Collin, M., 2013. Peer effects in the adoption of property rights: experimental evidence from urban Tanzania. Unpublished paper presented at a conference held at the Centre for the Study of African Economies, University of Oxford, UK.

Conley, T.G., 1999. GMM estimation with cross sectional dependence. *Journal of Econometrics* 92 (1), 1–45.

Conley, T.G. and Topa, G., 2002. Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Economics* 17, 303–327.

Conley, T.G. and Udry, C.R., 2010. Learning about a new technology: pineapple in Ghana. *American Economic Review* 100 (1), 35–69.

Conning, J. and Kevane, M., 2002. Community-based targeting mechanisms for social safety nets: a critical review. *World Development* 30 (3), 375–394.

De Giorgi, G., Pellizzari, M. and Redaelli, S., 2010. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics* 2 (2), 241–275.

de Melo, G., 2011. Peer effects identified through social networks. Evidence from Uruguayan schools. Working Paper, Department of Economics, University of Siena, 627, Department of Economics, University of Siena, Italy. Available at http://www.econ-pol.unisi.it/dipartimento/it/node/1606.

De Weerdt, J. and Dercon, S., 2006. Risk-sharing networks and insurance against illness. *Journal of Development Economics* 81 (2), 337–356.

Duflo, E. and Saez, E., 2003. The role of information and social interactions in retirement plan decisions: evidence from a randomized experiment. *Quarterly Journal of Economics* 118 (3), 815–842.

Evans, W.N., Oates, W.E. and Schwab, R.M., 1992. Measuring peer group effects: a study of teenage behavior. *Journal of Political Economy* 100 (5), 966–991.

Gaviria, A. and Raphael, S., 2001. School-based peer effects and juvenile behavior. *Review of Economics and Statistics* 83 (2), 257–268.

Glaeser, E.L., Sacerdote, B. and Scheinkman, J.A., 1996. Crime and social interactions. *Quarterly Journal of Economics* 111 (2), 507–548.

Goetz, A.M. and Jenkins, R., 2001. Hybrid forms of accountability: citizen engagement in institutions of public-sector oversight in India. *Public Management Review* 3 (3), 363-383.

Graham, B.S., 2008. Identifying social interactions through conditional variance restrictions. *Econometrica* 76 (3), 643–660.

Graham, B.S. and Hahn, J.Y., 2005. Identification and estimation of the linear-in-means model of social interactions. *Economics Letters* 88 (1), 1–6.

Guiso, L. and Schivardi, F., 2007. Spillovers in industrial districts. *Economic Journal* 117 (516), 68–93.

Hanushek, E.A., Kain, J.F., Markman, J.M. and Rivkin, S.G., 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics* 18 (5), 527–544.

Helmers, C. and Patnam, M. 2012. Does the rotten child spoil his companion? Spatial peer effects among children in rural India. Working Paper WPS/2010-13, Centre for the Study of African Economies, University of Oxford, UK.

Jackson, M.O., 2008. *Social and Economic Networks.* Princeton University Press, Princeton, New Jersey.

Khemani, S., 2007. Can information campaigns overcome political obstacles to serving the poor? In: Devarajan, S. and Widlung, I. (eds), *The politics of service delivery in democracies: better access for the poor*. Stockholm, Ministry of Foreign Affairs, Government of Sweden, pp. 56-69.

Kleibergen, F. and Paap, R., 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133 (1), 97–126.

Manski, C., 1993. Identification of endogenous social effects: the reflection problem. *Review of Economic Studies* 60 (3), 531–542.

Manski, C.F., 2000. Economic analysis of social interactions. *Journal of Economic Perspectives* 14 (3), 115–136.

Mansuri, G. and Rao, V., 2013. Localizing development: does participation work? Policy Report. World Bank, Washington, DC.

Mas, A. and Moretti, E., 2009. Peers at work. *American Economic Review* 99 (1), 112–145.

Moffitt, R.A., 2001. Policy interventions, low level equilibria, and social interactions. In: Durlauf, S.N. and Young, H.P. (eds), *Social Dynamics*. MIT Press, Cambridge, MA, pp. 45–82.

Munshi, K., 2004. Social learning in a heterogeneous population: technology diffusion in the Indian green revolution. *Journal of Development Economics* 73 (1), 185–213.

Olken, B., 2007. Monitoring corruption: evidence from a field experiment in Indonesia. *Journal of Political Economy* 115 (2), 200–249.

Sacerdote, B., 2001. Peer effects with random assignment: results for Dartmouth roommates. *Quarterly Journal of Economics* 116 (92), 681–704.

Shankar, S., Gaiha, R. and Jha, R., 2010. Information and corruption: the National Rural Employment Guarantee Scheme in India. ASARC Working Paper 32-32, Australia South Asia Research Centre, Australian National University, Canberra.

Soetevent, A.R., 2006. Empirics of the identification of social interactions: an evaluation of the approaches and their results. *Journal of Economic Surveys* 20 (2), 193–228.

Stock, J.H. and Yogo, M., 2005. Testing for weak instruments in linear IV regression. In: Stock, J.H. and Andrews, D.W.K. (eds), *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, Chapter 5. Cambridge University Press, Cambridge.

Tanzania Social Action Fund, 2006. *TASAF Annual Report 2005–2006*. Dar es Salaam, Tanzania.

Tanzania Social Action Fund, 2007. *TASAF Annual Report 2006–2007*. Dar es Salaam, Tanzania.

Tanzania Social Action Fund, 2008. *TASAF Annual Report 2007–2008*. Dar es Salaam, Tanzania.

Tanzania Social Action Fund, 2011a. *TASAF II Operational Manual*. Dar es Salaam, Tanzania.

Tanzania Social Action Fund, 2011b. *TASAF 4th Quarterly Report April–June 2011*. Dar es Salaam, Tanzania.

Van den Broeck, K. and Dercon, S., 2011. Information flows and social externalities in a Tanzanian banana growing village. *Journal of Development Studies* 47 (2), 231–252.

Wydick, B., Karp Hayes, H. and Hilliker Kempf, S., 2011. Social networks, neighborhood effects, and credit access: evidence from rural Guatemala. *World Development* 39 (6), 974–982.

Tables

Table 1 Elderly/disabled versus other households: a comparison in characteristics

	Other households	Elderly households	Disabled households
Proportion of population (%)	54.9	35.1	6.0
Percentage that owns:			
Radio	66.7	55.7***	52.5***
Phone	36.4	23.6***	24.1***
Vehicle	49.0	38.0***	40.3***
Watch/clock	40.8	37.6**	34.4***
Sofa	21.0	15.9*	14.5***
Average HH size	4.9	5.3***	6.1***
Percentage that is female headed	12.1	31.9***	28.7***
Average age of HH head (years)	36.9	64.4***	56.5***
Percentage with head with primary education	85.8	56.2***	62.6***
Percentage with head = member main tribe of village	78.4	85.2**	84.3***
Percentage with head = member main religion of village	53.9	61.6***	56.8
Percentage of dwellings with improved:			
Floor	29.5	22.7*	21.5**
Roof	63.2	62.4	59.1*
Walls	26.0	20.6	20.9
Water source	44.9	44.0	46.8
Latrine	62.0	60.0	57.0**
Energy source	11.3	8.0	7.1*
Average number of days that the HH ate [ITEM] during the previous week:			
Meat	1.0	0.7***	0.7***
Eggs	0.3	0.2***	0.2***
Milk	1.8	1.9	1.7
Percentage who attended village meeting in previous 3 months	46.7	47.2	47.6
Percentage who attended village meeting in the previous 6 months	68.7	68.5	69.5
Percentage with blood relative of village elite	6.8	7.0	7.0
Percentage with a member who holds any village office	5.4	5.4	5.2
Percentage TASAF informed	59.2	52.6***	57.4
N	6474	18,762	3303

Notes: Observations are weighted by inverse sampling probability. Significance of difference from non-vulnerable group is indicated by asterisks. Vehicle includes bicycle, car, motorbike or truck. See section 4.1 for a description of the other variables. ***$p<0.01$; **$p<0.05$; *$p<0.10$.

HH, household; TASAF, Tanzania Social Action Fund.

Table 2 Characteristics of informed and uninformed elderly/disabled households

	TASAF-uninformed elderly/disabled households	TASAF-informed elderly/disabled households	Statistical significance level of difference
Proportion of sample (%)	46.73	53.27	–
Percentage that owns:			
Radio	45.9	64.4	***
Phone	16.6	30.1	***
Vehicle	39.1	37.6	
Watch/clock	28.3	45.4	***
Sofa	10.6	20.5	***
Average HH size	5.1	5.7	***
Percentage that is female headed	39.0	24.7	***
Average age of HH head (years)	64.9	61.4	***
Percentage with head with primary education	42.3	70.8	***
Percentage with head = member main tribe of village	83.3	86.6	***
Percentage with head = member main religion of village	60.7	61.4	
Percentage of dwellings with improved:			
Floor	16.4	28.1	***
Roof	53.1	70.0	***
Walls	15.8	24.8	***
Water source	39.4	48.2	***
Latrine	54.9	63.2	***
Energy source	5.7	10.0	***
Average number of days that the HH ate [ITEM] during the previous week:			
Meat	0.6	0.8	***
Eggs	0.2	0.3	***
Milk	1.5	2.2	***
Percentage that attended village meeting in previous 3 months	40.1	53.5	***
Percentage that attended village meeting in the previous 6 months	60.4	76.0	***
Percentage with blood relative of village elite	4.7	9.1	***
Percentage with a member who holds any village office	1.7	8.6	***
Percentage of nearest 10 neighbours who are TASAF informed	47.6	64.0	***
Average distance from village leader (metres)	1938	1547	***
N	9306	10,610	

Notes: Significance level of difference is indicated by asterisks. ***$p<0.01$; **$p<0.05$; *$p<0.10$.
HH, household; TASAF, Tanzania Social Action Fund.

Table 3 Determinants of being informed: 10 nearest neighbours specification (OLS)

Dependent variable = 1 if informed	(1)	(2)	(3)	(4)
Endogenous peer effect: number of nearest 10 neighbours who are informed	0.045*** (0.027)	0.039*** (0.003)	0.042*** (0.003)	0.007* (0.004)
HH's own characteristics:				
1 = owns a radio		0.058*** (0.009)	0.058*** (0.009)	0.057*** (0.009)
1 = owns a mobile phone		0.042*** (0.009)	0.044*** (0.009)	0.044*** (0.009)
1 = owns a vehicle		–0.005 (0.010)	–0.005 (0.010)	–0.004 (0.011)
1 = owns a watch or clock		0.031*** (0.007)	0.032*** (0.007)	0.035*** (0.007)
1 = owns a sofa		0.031*** (0.009)	0.032*** (0.009)	0.028*** (0.009)
1 = dwelling has improved roof		0.011 (0.012)	0.004 (0.013)	0.005 (0.013)
1 = dwelling has improved latrine		0.028*** (0.008)	0.032*** (0.009)	0.035*** (0.009)
Household size		0.009*** (0.001)	0.009*** (0.001)	0.009*** (0.001)
Age head		–0.003*** (0.000)	–0.003*** (0.000)	–0.003*** (0.001)
1 = employed		0.040*** (0.012)	0.040*** (0.012)	0.041*** (0.012)
1 = member main tribe village		–0.020* (0.010)	–0.024** (0.012)	–0.022* (0.012)
1 = member main religion village		–0.007 (0.009)	–0.009 (0.009)	–0.008 (0.009)
1 = has eaten meat in past 7 days		0.031*** (0.009)	0.028*** (0.009)	0.026*** (0.009)
1 = has drank milk in past 7 days		0.013 (0.008)	0.014* (0.008)	0.011 (0.008)
1 = head with primary education		0.131*** (0.008)	0.130*** (0.009)	0.126*** (0.009)
1 = female headed		–0.084*** (0.010)	–0.082*** (0.010)	–0.085*** (0.010)
1 = attended village meeting in previous 3 months		0.040*** (0.010)	0.055*** (0.011)	0.061*** (0.011)
1 = attended village meeting in previous 6 months		0.082*** (0.013)	0.089*** (0.013)	0.085*** (0.014)
1 = blood related to village elite		0.045*** (0.012)	0.045*** (0.013)	0.046*** (0.014)
1 = holds political office		0.156*** (0.015)	0.154*** (0.015)	0.152*** (0.015)
1 = has relatives in village		0.017** (0.007)	0.016** (0.007)	0.020** (0.007)
1 = respondent household head		0.137*** (0.011)	0.138*** (0.011)	0.141*** (0.011)
Contextual peer effects	No	No	Yes	Yes
Hamlet fixed effects	No	No	No	Yes
N	19,392	19,392	19,392	19,392
Adjusted R^2	0.122	0.244	0.246	0.143

Notes: Dependent variable is a dummy = 1 if HH was informed. Standard errors (shown in parentheses) are adjusted for clustering of observations at village level (as some networks extend to individuals outside the hamlet). (1)-(3) include village fixed effects. Regressions with contextual effects include the full set of neighbours set characteristics. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure of the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. HH, household. ***$p<0.01$, **$p<0.05$, *$p<0.1$.

Table 4 Two-stage least squares first-stage regression

	Proportion of 10 nearest neighbours who are informed
Excluded peers' characteristics:	
1 = dwelling has improved roof	0.030*** (0.011)
1 = female headed	0.024** (0.011)
1 = has relatives in village	0.024** (0.009)
1 = respondent household head	0.031*** (0.010)
Household's own characteristics	Yes
Contextual peer effects	Yes
Hamlet fixed effects	Yes
N	19,392
Joint F-statistic	83.48
Adjusted R^2	0.16
Craig–Donald Wald F-statistic	26.54
Kleibergen–Paap Wald F-statistic	8.10
Kleibergen–Paap LM statistic	29.03
p-value for null of underidentification	0.00
Hansen J *p*-value	0.23

Notes: Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level.
Contextual effects include the full set of neighbours set characteristics.
TASAF, Tanzania Social Action Fund.
*p<0.10; **p<0.05; ***p<0.01.

Table 5 Determinants of being informed: 10 nearest neighbours specification (2SLS)

Dependent variable = 1 if informed	(1)		(2)	
Endogenous peer effect: number of nearest 10 neighbours who are informed	0.092***	(0.027)	0.077**	(0.030)
Household's own characteristics:				
1 = owns a radio	0.060***	(0.009)	0.057***	(0.008)
1 = owns a mobile phone	0.047***	(0.009)	0.049***	(0.009)
1 = owns a vehicle	–0.005	(0.010)	–0.005	(0.009)
1 = owns a watch or clock	0.031***	(0.007)	0.034***	(0.008)
1 = owns a sofa	0.031***	(0.010)	0.031***	(0.011)
1 = dwelling has improved roof	–0.000	(0.013)	–0.000	(0.010)
1 = dwelling has improved latrine	0.033***	(0.009)	0.033***	(0.009)
Household size	0.009***	(0.001)	0.009***	(0.001)
Age head	–0.003***	(0.000)	–0.003***	(0.000)
1 = employed	0.042***	(0.012)	0.043***	(0.012)
1 = member main tribe village	–0.026**	(0.012)	–0.025**	(0.011)
1 = member main religion village	–0.011	(0.009)	–0.010	(0.009)
1 = has eaten meat in past 7 days	0.027***	(0.009)	0.027***	(0.008)
1 = has drank milk in past 7 days	0.013*	(0.008)	0.012	(0.008)
1 = head had at least 1 year of primary Education	0.123***	(0.009)	0.122***	(0.009)
1 = female headed	–0.087***	(0.010)	–0.090***	(0.009)
1 = attended village meeting in previous 3 Months	0.058***	(0.011)	0.062***	(0.010)
1 = attended village meeting in previous 6 months	0.085***	(0.013)	0.083***	(0.011)
1 = blood related to village elite	0.040***	(0.013)	0.042***	(0.013)
1 = holds political office	0.153***	(0.015)	0.152***	(0.013)
1 = has relatives in village	0.013*	(0.007)	0.014*	(0.008)
1 = respondent household head	0.137***	(0.011)	0.136***	(0.008)
Contextual peer effects	Yes		Yes	
Hamlet fixed effects	No		Yes	
N	19,392		19,392	
Adjusted R^2	0.217		0.101	

Notes: Dependent variable is a dummy = 1 if household was informed about TASAF. Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level (as some networks extend to individuals outside the hamlet). Regressions include the full set of neighbours set characteristics. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. The four characteristics of excluded peers used as instruments for 2SLS are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head.

*p<0.10; **p<0.05; ***p<0.01.

Table 6 Endogenous peer effect for varying neighbours set sizes

Dependent variable = 1 if informed	$k = 5$ (1)	$k = 10$ (2)	$k = 15$ (3)	$k = 20$ (4)
Endogenous social interaction effect: number of nearest k neighbours who are informed about TASAF	0.284*** (0.086)	0.064** (0.032)	0.053* (0.028)	0.051 (0.032)
Household's own characteristics	Yes	Yes	Yes	Yes
Contextual peer effects	Yes	Yes	Yes	Yes
Hamlet fixed effects	Yes	Yes	Yes	Yes
N	17,495	17,495	17,495	17,495
Adjusted R^2	–0.168	0.117	0.114	0.100
Craig–Donald Wald F-statistic	7.23	24.44	15.18	11.27
Kleibergen–Paap Wald F-statistic	3.89	10.01	3.63	2.24
Hansen J p-value	0.075	0.635	0.448	0.298

Notes: Dependent variable is a dummy = 1 if household was informed about the Tanzania Social Action Fund (TASAF). Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level. Contextual effects include the full set of neighbours set characteristics. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. The four characteristics of excluded peers used as instruments for two-stage least squares are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head.
*p<0.10; **p<0.05; ***p<0.01.

Table 7 Heterogeneous peer effects in the probability of being informed (2SLS)

Panel A: Interactions with household's *own characteristics*										
	Female headed (1)	Age head (continuous) (2)	Main religion (3)	Main tribe (4)	Primary education (5)	Head is employed (6)	Eaten meat in past 7 days (7)	Owns phone (8)	Relatives in village (9)	Political office (10)
Endogenous peer effect	0.070***	0.159***	0.070**	0.078***	0.064**	0.030	0.086***	0.078***	0.063***	0.079***
	(0.030)	(0.033)	(0.029)	(0.033)	(0.029)	(0.034)	(0.029)	(0.029)	(0.030)	(030)
Interaction effect	–0.009	–0.001***	0.027***	0.005	0.024**	0.036**	–0.019**	–0.011	0.020**	–0.027
	(0.009)	(0.000)	(0.010)	(0.012)	(0.010)	(0.015)	(0.009)	(0.009)	(0.008)	(0.018)
N	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392

Panel B: Interactions with household's *neighbours set characteristics*										
	Female headed (1)	Dummy age head >55 (2)	Main religion (3)	Main tribe (4)	Primary education (5)	Head is employed (6)	Eaten meat in past 7 days (7)	Owns phone (8)	Relatives in village (9)	Political office (10)
Endogenous peer effect	0.085***	0.063**	0.080***	0.085***	0.070**	0.148**	0.087***	0.079***	0.055*	0.072**
	(0.026)	(0.027)	(0.031)	(0.045)	(0.032)	(0.070)	(0.028)	(0.028)	(0.029)	(0.030)
Interaction effect	–0.012	0.012	0.019	–0.002	0.003	–0.077	0.055	0.008	0.042**	0.123**
	(0.026)	(0.008)	(0.020)	(0.031)	(0.025)	(0.059)	(0.022)	(0.022)	(0.018)	(0.054)
N	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392

Notes: Each regression controls for a household's own characteristics, household's neighbours set characteristics and hamlet fixed effects. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level. In all columns except for (1) and (6) the interacted variable is binary (0 or 1). The four characteristics of excluded peers used as instruments for two-stage least squares are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head. TASAF, Tanzania Social Action Fund.

*p<0.10; **p<0.05; ***p<0.01.

Table 7 *Continued*

Panel C: Triple interactions with household's own and household's neighbours set characteristics										
	Female headed (1)	Age head (2)	Main religion (3)	Main tribe (4)	Primary education (5)	Head is employed (6)	Eaten meat in past 7 days (7)	Owns phone (8)	Relatives in village (9)	Political office (10)
Endogenous peer effect	0.092*** (0.027)	–	0.064** (0.030)	0.070 (0.057)	0.062* (0.037)	0.094 (0.229)	0.095*** (0.029)	0.082*** (0.028)	0.063** (0.029)	0.068** (0.029)
Interaction with HH's own characteristic	–0.033* (0.012)	–	0.019 (0.025)	–0.005 (0.041)	0.007 (0.031)	0.036 (0.242)	–0.037* (0.020)	–0.046*** (0.016)	–0.017 (0.024)	–0.023 (0.024)
Interaction with HH's neighbours set characteristic	–0.043 (0.032)	–	–0.051 (0.028)	–0.004 (0.056)	–0.020 (0.036)	–0.062 (0.239)	–0.005 (0.027)	–0.017 (0.025)	–0.003 (0.030)	0.139** (0.059)
Triple interaction (own + neighbours set)	0.078 (0.051)	–	0.020 (0.035)	0.018 (0.054)	0.065 (0.044)	–0.004 (0.253)	0.028 (0.037)	0.098** (0.040)	0.057 (0.037)	–0.118 (0.164)
Significance level of sum of three interaction terms	>10%	–	>10%	>10%	>10%	>10%	>10%	>10%	10%	>10%
N	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392	19,392

Notes: Each regression controls for a household's own characteristics, household's neighbours set characteristics and hamlet fixed effects. In addition, regressions in panel C control for the interaction between a household's own and a household's neighbours set characteristics. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level. In all columns except for (1) and (6) the interacted variable is binary (0 or 1). The four characteristics of excluded peers used as instruments for two-stage least squares are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head.
*p<0.10; **p<0.05; ***p<0.01.

Table 8 Robustness to Conley's (1999) correction of standard errors

Dependent variable = 1 if informed	2SLS (1)	2SLS (2)
Endogenous peer effect: number of nearest *k* neighbours who are informed	0.119*** (0.046)	0.119*** (0.047)
Adjusting for spatial dependence of standard errors within hamlet (using Conley's method, 1999)	No	Yes
N (reduced to 20 randomly sampled villages out of 100)	3685	3685

Notes: Dependent variable is a dummy = 1 if household was informed about the Tanzania Social Action Fund (TASAF). Standard errors are shown in parentheses. Standard errors in (1) are adjusted for clustering of observations at the village level. Each regression controls for contextual effects, individual household characteristics and hamlet fixed effects. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. The four characteristics of excluded peers used as instruments for 2SLS are: dummy for improved roof, female-headed households, whether the household has relatives in the village, whether the respondent was the household head.

2SLS, two-stage least squares.

*p<0.10; **p<0.05; ***p<0.01.

Table 9 Samples restricted to the least mobile households (HHs) (2SLS)

Dependent variable	= 1 if informed about TASAF					= 1 if informed about VG group
Sample restricted to:	All HHs (1)	Not attending village meetings (2)	Survey respondent was head (3)	Does not own a vehicle (4)	Head aged >65 years (5)	All households (no restrictions) (6)
Endogenous peer effect	0.077** (0.030)	0.074** (0.044)	0.098*** (0.035)	0.115*** (0.041)	0.087** (0.042)	0.157***
N	19,392	10,228	11,591	12,099	9,438	19,392
Kleibergen–Paap Wald F-statistic	8.095	5.397	6.772	5.333	5.899	3.811

Notes: Endogenous peer effect = number of nearest *k* neighbours who are informed. All regressions control for a household's own characteristics, contextual effects and hamlet fixed effects. Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. The four characteristics of excluded peers used as instruments for two-stage least squares (2SLS) are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head.

TASAF, Tanzania Social Action Fund.

*p<0.10; **p<0.05; ***p<0.01.

Appendix

A.1. Population and sample information

Number of sampled districts		5	
Number of sampled villages per district		20	
Total number of sampled villages		100	
Total number of hamlets in sampled villages		705	
Number of hamlets per sampled village	Min.	2	
	Max.	15	
	Mean	7	

			True population	Sample
Total number of households (number in parentheses shows fraction of true population)	All		61,610	30,339 (49%)
	Elderly/disabled		22,957	19,916 (87%)
Number of households per village	All	Min.	116	72
		Max.	3761	964
		Mean	616	303
	Elderly/disabled	Min.	42	27
		Max.	641	580
		Mean	230	199
Number of households per hamlet	All	Min.	17	3
		Max.	655	196
		Mean	114	43
	Elderly/disabled	Min.	8	1
		Max.	114	135
		Mean	42	28
Distance between households within village (metres) (table shows *average* min., max. and mean of 100 villages)		Min.	2	4
		Max.	7662	8900
		Mean	1867	1800
Distance to the *k*th nearest household within village (table shows *average* min., max. and mean of 100 villages)	$k = 1$	Min.	2	4
		Max.	2471	922
		Mean	54	72
	$k = 5$	Min.	26	45
		Max.	2839	1296
		Mean	141	200
	$k = 10$	Min.	55	92
		Max.	2976	1513
		Mean	216	310
	$k = 15$	Min.	81	134
		Max.	3083	1681
		Mean	278	402
	$k = 20$	Min.	106	173
		Max.	3188	1852
		Mean	335	488

A.2. Interaction between peer effect and geographical distance of peers

Dependent variable = 1 if informed		
Endogenous peer effect: number of nearest 10 neighbours who are informed	0.086***	(0.028)
Interaction of peer effect with average geographical distance of ten peers (in meters)	-0.00002***	(0.0001)
Household's own characteristics:		
1 = owns a radio	0.057***	(0.008)
1 = owns a mobile phone	0.046***	(0.009)
1 = owns a vehicle	–0.005	(0.011)
1 = owns a watch or clock	0.034***	(0.007)
1 = owns a sofa	0.028***	(0.011)
1 = dwelling has improved roof	–0.003	(0.013)
1 = dwelling has improved latrine	0.032***	(0.009)
Household size	0.010***	(0.001)
Age head	–0.003***	(0.001)
1 = employed	0.043***	(0.012)
1 = member main tribe village	–0.024**	(0.012)
1 = member main religion village	–0.010	(0.009)
1 = has eaten meat in past 7 days	0.027***	(0.009)
1 = has drank milk in past 7 days	0.013	(0.008)
1 = head had at least 1 year of primary Education	0.122***	(0.009)
1 = female headed	–0.089***	(0.009)
1 = attended village meeting in previous 3 Months	0.062***	(0.011)
1 = attended village meeting in previous 6 months	0.084***	(0.014)
1 = blood related to village elite	0.040***	(0.014)
1 = holds political office	0.152***	(0.015)
1 = has relatives in village	0.013*	(0.007)
1 = respondent household head	0.138***	(0.011)
Contextual peer effects	Yes	
Hamlet fixed effects	Yes	
N	19,392	

Notes: Dependent variable is a dummy = 1 if household was informed about TASAF. Standard errors (shown in parentheses) are adjusted for clustering of observations at the village level (as some networks extend to individuals outside the hamlet). Regressions include the full set of neighbours set characteristics. Although the regressions include the proportion of 10 nearest neighbours being informed as a measure for the endogenous social interaction effect, the table presents the marginal endogenous social interaction effect of one additional neighbour being informed. The four characteristics of excluded peers used as instruments for 2SLS are: dummy for improved roof, female-headed households, whether the household has relatives in the village and whether the respondent was the household head.

*p<0.10; **p<0.05; ***p<0.01.