Social Networks and Economic Transformation: Evidence from a Resettled Village in Brazil *

Margherita Comola[†] Carla Inguaggiato[‡] Mariapia Mendola[§]

January 16, 2024

Abstract

We study the role of social learning in the diffusion of cash crops in a resettled village economy in Brazil. We combine detailed geo-coded data on farming plots with dyadic data on social ties among settlers, and we leverage the variation in network formation induced by the landless workers' movement land occupation. By using longitudinal data on farming decisions over 15 years, we find evidence of significant peer effects in the decision to farm new cash fruits (pineapple and passion fruit). Our results suggest that social diffusion is heterogeneous along observed plot and crop characteristics, i.e. farmers growing water-sensitive crop are more likely to respond to the actions of peers with similar water access conditions.

JEL codes: C45; D85; J15; O33; Q15

Keywords: Technology Adoption, Agrarian Reform, Social Networks, Peer Effects, Brazil

*We are grateful to numerous seminar participants for helpful comments and discussions, and to Isacco Rama for outstanding research assistance. Data used in this paper were collected by Carla Inguaggiato as part of her doctoral thesis at the University of Trento, with financial support from the Training Centre for International Cooperation, which she gratefully acknowledges. Margherita Comola aknowledges support of the grants ANR-17-EURE-0001 and ANR-21-CE26-0002-01. The usual disclaimer applies.

[†]University Paris-Saclay (RITM) and Paris School of Economics. Email: margherita.comola@universite-paris-saclay.fr.

[‡]Centre for Development and Environment, University of Bern. Email: carla.inguaggiato@cde.unibe.ch.

[§]Department of Economics, University of Milano–Bicocca and IZA. Email: mariapia.mendola@unimib.it.

1 Introduction

Knowledge diffusion and learning are major drivers of economic growth (Lucas, 1988; Acemoglu, 2009; Banerjee et al., 2019). Information constraints to innovation and technological change have long been studied in the literature on economic development, and a great deal of attention has been given to the role of social networks in the uptake and diffusion of new technologies, especially in agriculture (e.g. Feder and Zilberman, 1985; Bandiera and Rasul, 2006; Conley and Udry, 2010). Agricultural technology adoption involves a high degree of uncertainty and complexities that go beyond farm characteristics. The Green Revolution, which brought new high-yielding varieties (HYVs) to India in the 1960s, highlighted the importance of understanding why many farmers do not adopt new seeds despite suitable agronomic conditions and high returns (Foster and Rosenzweig, 1993; Munshi, 2004).[1]

While the importance of interpersonal networks and peer experience in shaping the adoption of new technologies is largely recognized, especially in developing contexts, the identification of social learning effects remains a major challenge because peers are usually chosen endogenously (Manski, 1993; Brock and Durlauf, 2001; Moffitt, 2001; Fafchamps and Lund, 2003). Indeed, when detecting a significant correlation between the outcomes of linked farmers, it is difficult to say whether this is due to authentic peer effects or to the well documented tendency to establish links with partners of similar characteristics (so-called network assortativity). Recently, field experiments have been used to generate exogenous variation in the timing and frequency of social interactions or in the degree of information supply about peers' behavior (see Feigenberg et al., 2010; Cai et al., 2015; Fafchamps and Quinn, 2018).

In this paper, we investigate the dynamics of social learning by exploiting the unique "quasi–experiment" of an agrarian reform settlement in the poorest state in Brazil, where stranger households found themselves living in a newly created 'village' ('assentamento') undergoing rapid economic and agricultural transformation they were not trained for. What makes this setting particularly attractive for our purposes is the variation in the farmers' social network induced by the land occupation movement. The latter organiza-

[1] By tracking farming households in rural India, Foster and Rosenzweig (1993) find that lack of knowledge is a main deterrent to adoption and that farmers whose neighbors had experience growing the new seeds adopted more HYVs on their own land. A large body of works have tested the role of other mechanisms such as credit and insurance constraints (e.g. BenYishay et al., 2019) but technology adoption oftentimes remains low even after some of these barriers are removed (e.g. Ambler et al., 2018; Karlan et al., 2014; Stifel and Minten, 2008).

tion maintains that it is legally justified in occupying unproductive land, hence creating informal settlements that turn into new villages where interpersonal ties are partly driven by the location of vacant land during the squatting process.[2] Specifically, we leverage the fact that proximity between farmers during the land squatting campaign fostered social links. This is so as the landless occupiers are externally recruited households, who do not know each other at the time of the encampment but enter together the targeted settlement and set up makeshift housing, which then become their permanent farm (Hammond, 2009). Remarkably, the speed of the recruiting and occupation processes did not allow for creating ties beforehand so that most occupiers were strangers to each other (Flynn, 2010). This enables us to address the endogeneity of social links and investigate peer effects in agricultural technology adoption in a unique link–formation setting. Our empirical strategy relies on a model of social diffusion along network lines (Bramoullé et al., 2009), which we extend to a longitudinal setting to exploit time variation in farmer technology adoption.[3] In addition, our estimation strategy includes household fixed effects, which control for time-invariant attributes at the time of the settlement (e.g. entrepreneurial ability).

Agrarian reform in Brazil, which has controversially involved both civil society and political leaders since the 1990s, is one of the most powerful and globally renowned intervention to reduce long-lasting land inequalities and extreme poverty (Pereira, 2003). The distinctive aspect of Brazilian agrarian reform is the active role played by the landless workers' movements – among which the most widely known is the *Movimento Sem Terra* (MST) – in pushing the state to expropriate unproductive land (Sigaud, 2004; Wolford, 2010). These movements mobilized alien households to occupy land in order to put pres-

[2]Community-based networks are active throughout the developing world and, unlike in our setting, they are typically organized around close-knit communities that have been in place for long periods of time. Depending on the context, these groups may be based on kinship (including castes in India or clans in sub-Saharan Africa) or on geographical proximity (villages or neighborhoods) (e.g. Munshi, 2004; Barr, 2003). However, communities based on either geography or kinship are likely to be formed endogenously (e.g. the former through location choice, the latter through marriage). For example, studying the American South in the decades of the late 19th century after Emancipation, Chay and Munshi (2014) show that Black spatial proximity varied substantially across southern counties, depending on the crops grown in the local area. Blacks worked (and lived) in close proximity to each other where labor-intensive plantation crops (such as tobacco, cotton and rice) were grown. They lived in a more dispersed fashion where less labor-intensive crops (such as wheat and corn) were more common.

[3]Most previous studies on peer effects use data where individuals are partitioned into mutually exclusive and fully overlapping reference groups (e.g., all farmers in the same village). In doing so, they assume that individuals are equally affected by all other individuals belonging to their group and by nobody outside their group. In our model the reference group has individual-level variation: if i and j are connected and j and k are connected, it does not necessarily imply that i and k are also connected.

sure on the governmental body (*Instituto Nacional de Colonizacao e Reforma Agraria*, or INCRA) to expropriate unproductive and vacant land.[4] After a process of occupation of unproductive land, settlers are officially given the 'right' to cultivate their own plots such that, for the first time in some rural areas in Brazil, household residency and property overlap (Bergamasco, 1997; Wanderley, 2000). This means that, for the first time, household farming has become a relevant exchange activity in these communities, as opposed to landless wage labor in large estates (*latifundios*). These agrarian reform settlements become centers for technological change and socio-economic development, whereby landless farmers can be the driving force of a newly created village economy.

Our study focuses on a resettled village in the northeast of Brazil. Following the land occupation process, this village went through an agrarian transformation that shifted labor from former enslavement in the sugar cane plantations to household-level commercial farming. The above transformation involved the adoption of a set of new crops, including, for the first time, perennial crops, such as pineapple (*abacaxi*) and passion fruit (*maracuja*). These are perishable commercial fruits with high commercial value, whose cultivation was forbidden before the agrarian reform.[5] Unlike subsistence farming, tropical plants rely on labor-intensive provision, efficient handling, water control and post-harvest conservation techniques.[6]

Our analysis combines dyadic data on social ties among farmers with geo-localized data on land squatting and longitudinal farm production information. We first provide evidence that the geographical location of settlers during the encampment period was not driven by assortativity along observables, and we then leverage the geographical proximity between farmers induced by the land-squatting process to instrument for the formation of social networks. By using network data on the entire village population, we can thus identify the role of social learning effects on households' farming choices. Our instrumentation strategy circumvents the difficulties related to the causal assessment of the role of interpersonal

[4]This is mostly dispossed land of sugar-cane factories that went bankrupt during the sugar crisis of the early 1990s.

[5]Sugar cane was the main plantation cultivated before the reform by landlords, who employed workers along slavery conditions. When sugar cane price was high, the cane was planted on every viable plot. When price was low, workers were allowed to plant garden crops in front of their houses. Yet, workers were prohibited from planting any tree or other species with 'long roots,' because these plants might give them a legal or moral claim on the land (Wolford, 2010).

[6]Northeastern Brazil is a drought-ridden region where rainfall, good soil and adequate infrastructure is much less well-distributed than in the southern region. Overall, after being the largest world producer of sugar cane, Brazil is increasingly populated by both subsistence farmers (cultivating traditional crops such cassava, peanuts, sweet potatoes, maize) and producers of forest products, cocoa and tropical fruits.

networks due to the endogeneity of real-life social interactions and network assortativity (Cai et al., 2015). Our results show a large positive peer effect on the adoption of new cash crops. Having one additional peer cultivating a given cash crop increases the probability of crop-specific adoption by 9 - 14 percentage points according to the most conservative estimate. Interestingly, the effect appears to be heterogeneous across plot and crop characteristics, i.e. conditional on the characteristics shared by farmers, the social learning process of water-needing crops appears to be mostly driven by peers whose plots have similar water access conditions.

Our paper relates to the large body of literature studying the role of social networks in shaping innovation and development.[7] By using a network approach Conley and Udry (2010) document the role of social learning through social ties in the diffusion of a new cash crop with high profit margins in Ghana. They use detailed information on communication patterns among farmers to show that new growers learn by observing the experimentation of their peers. A number of related contributions study social learning in technology adoption in different developing settings where markets are thin and information is incomplete (Bandiera and Rasul, 2006; Munshi, 2004; Foster and Rosenzweig, 2010; Maertens, 2017). By learning from others (typically neighbors), a new grower updates her priors about the unfamiliar technology and adds actual information on expected net returns in her optimal (profit-maximizing) problem of adoption choice (Foster and Rosenzweig, 1993; Conley and Udry, 2001). In this respect, the prospect of social learning declines with geographical distance (Glaeser, 1999; Fafchamps, 2010) as well as with socio-economic differences among peers (Fafchamps and Soderbom, 2011).[8] We add to this literature by exploiting the novel combination of a natural experiment and detailed network data, which allows us to identify the dynamics of diffusion along network lines.

Our findings also contribute to the vital debate on the role of grassroots innovation in the inclusive and sustainable development of rural communities (Mosse, 2001; Platteau, 2004; Rigon, 2014). Criticism of the 'participatory' agrarian reform refers to it as a social program mimicking a landlord's benevolent behavior towards the most marginalized groups

[7] Different studies have produced empirical evidence of peer effects in many areas, from school performances to female entrepreneurship, and from financial to insurance decisions (Case and Katz, 1991; Hoxby, 2000; Feigenberg et al., 2010; Banerjee et al., 2013; Cai et al., 2015). A related strand of literature has exploited the allocation of new immigrants in Scandinavian countries to assess their performance and labor market integration in the host society (Edin et al., 2003; Damm, 2009; Dahlberg et al., 2012).

[8] Important recent studies investigate the spatial structure of social interactions recorded by mobile phones or Facebook data in large cities and find consistent evidence that distance is costly to social networks (e.g. Bailey et al., 2020; Buchel and von Ehrlich, 2020).

in the population. The main reason lies in the significant focus on housing creation rather than technical assistance, agricultural training, facilities and credit. We document, instead, the extent to which agrarian reform settlements can be centers for technological change and socio-economic development. Indeed, our results point to the innovative element of Brazilian agrarian reform, which involves not only access to land but especially the mobilization of people with different backgrounds and geographical origin and the links among them. More in general, we show that in the difficult context of resettlement, where competition and 'gate keeping' may be likely to emerge, the structure of (extra-familial) social relations are essential in fostering learning, innovation and eventually structural transformation.

The rest of the paper is organized as follows. Section 2 describes the setting and the Brazilian context. Data are presented in Section 3 while Section 4 describes the empirical strategy. Results on social learning are reported in Section 5 and 6. Section 7 concludes.

2 Brazilian land reform and agrarian settlements

Land reform is a key policy tool for reducing poverty and inequality in Brazil, where the latter take the form of lack of access to land and socio–economic exclusion (Pereira, 2003). Agrarian reform has been included in the political agenda since the 1960s and social movements, including the MST, have played a major role in its implementation. This is because the mass mobilization of people to occupy unproductive *latifundios* has been crucial to the government's push for land expropriation.[9] Table 1 shows that social movement interventions peaked in the mid-1990s when the number of settled families increased significantly (almost 50 thousand households resettled).

Table 1: Number of families settled over time

Year	Government	Settlers per year
1964–1984	Military regime	3689
1985–1989	Josï¿œ Sarney	16737
1990–1992	Fernando Collor de Mello	14172
1993–1994	Itamar Franco	7183
1995–2002	Fernando Henrique Cardoso	48923
2003–2009	Luis Inacio Lula da Silva	7564

Source. L. S. de Medeiros (2013)

[9] Encampment was the main tool pushing governments to implement agrarian reform and MST has been able to make the 'landless' a new political force (Rosa, 2012).

The most innovative element of agrarian reform in Brazil is not access to land *per se* but rather the creation of a heterogeneous socio-economic space of professions and a novel autonomy in the use of working time (Inguaggiato, 2014). This is even more so in the northeastern region of Brazil, the poorest region with a tradition of large-estate cane-plantation agriculture combined with subsistence farming. Yet, in the early 1990s economic and political conditions, and especially the sugar-cane crisis, allowed social movements to mobilize people on a massive scale in order to occupy unproductive properties, hence pushing the government to expropriate land in those areas as well (Sigaud, 2004). When MST leaders received information about a possible land to be eligible for expropriation, they to act rapidly by asking all militants to recruit people to occupy the land as fast as possible both in urban peripheries and rural impoverished areas (Flynn, 2010). Consequently, in this setting the creation of new rural village economies and their structure are not driven by kinship or training. They are shaped, instead, by vacant land and land-scarcity conditions as well as the social movement and encampment processes. Eligibility criteria for beneficiary households of the agrarian reform include being landless, a smallholder, a wage worker or a land tenant.[10]

Importantly for our study, for a long time family farming was not considered a relevant economic activity in Brazil. This is due to the structure of land distribution on the one hand, and to the strict dependence of rural areas upon urban areas on the other hand. Before the reform in Brazil, almost 80 percent of rural properties were organized in *latifundios* and household residences and land property have typically been separated to a large extent.[11] Hence, the creation of new agrarian reform settlements introduced a novelty by generating household farming as a major activity where residence and property dimensions are combined.[12]

[10] Art. 5 Norma de Execucao No 45, de 25 de Agosto de 2005, DOU 166, de 29-8-2005, secao 1, p. 122 B.S. 35, de 29-8-2005.

[11] Historically, people used to reside in rural areas but everything they needed for their social and economic life, such as medical services, markets and banks, was located in town. Rural areas in Brazil were, therefore, not conceived as autonomous socio-economic spaces but rather as a periphery or appendix to an urban area (IBGE, 2009; Wanderley, 2000).

[12] The concept of family farming activities describes an estate that is directly and personally exploited by the farmer and his/her family, providing the family with subsistence and an economic and social livelihood. The extension area of family farming is ruled by regional law according to the type of production Tinoco and Julliatto (2008). The first definition of Familiar Property is included in the Land Statute (Estatuto da Terra, Law n. 4.504 30 November 1964, art. 4). The last Brazilian census (2010) also included settlers without title under the category 'family farmer', as opposed to the previous census in 1995. This was in addition to 'occupier', which was the only category of family involved in agrarian reform considered previously.

After land reform, household farming in Brazil represents 38% of the total value of agricultural production and employs 74% of the rural labor force (IBGE, 2018). The introduction of family farming entailed a change in crop production as well, in particular by allowing – for the first time – the cultivation of perennial crops with high profit margins. New crops (e.g. commercial fruits and cash crops) have been introduced in place of large plantations, requiring new farming techniques, a workforce over the production cycle and market access within the village.

2.1 The creation of the village settlement

The village studied in this paper was created during the Cardoso administration in the mid-1990s and is located in the north of Alagoas, in Maragogi municipality, an area that was historically – and is still today – dominated by sugar cane plantations. Alagoas is the poorest Brazilian state, with the highest land concentration rate (IBGE, 2018).

The process of settlement creation described here follows several steps, that are typical across all agrarian settlements in Brazil. The first phase entails the identification of unproductive land by a social movement. Starting from mid 1990's, in the municipality under study 18 *assentamentos* were created, which were sugarcane plantations (*fazendas*) before the agrarian reform. The sugarcane company, which had four productive units in Alagoas, was heavily indebted to the Bank of Brazil and had to give up several productive units to the bank in order to have access to new loans (Sigaud, 2004). Hence, landless mobilizers recruited households in towns and rural villages neighboring the city of Penedo (distance between different towns ranges between 50 and 300 kilometers). In the first week of January 1998, approximately eight buses with over 100 new households arrived in the village. At that time, only 15 households were living in the settlement (first settlers): the family of the only local shop tenant (*barraqueiro*) plus fourteen former workers of the just-bankrupted sugar-cane factory. The new settlers entered the abandoned *fazenda* area by building their encampment (i.e. tents made of black plastic) and occupying the land. Land occupation was fast and casual at that stage, while the whole period of encampment, which involved living in harsh conditions with no electricity, water or sanitation, lasted one year before the land was finally dispossessed. During the early months of occupation, many of the first settlers left the encampment and were replaced by a second round of new settlers, also recruited from neighboring area. This process increased the degree of heterogeneity across settled households even further. During the encampment, people lived off what they were able to grow and off food supply rations (*cesta basica*) provided by the municipality (Inguaggiato, 2014).

The next step included the checking of land eligibility for expropriation by INCRA and the division of land into parcels. Hence, INCRA measured the land and assigned a plot (*lote*) to each household. This was done on the basis of land occupation during the encampment, and on geo-morphological characteristics of the plots (e.g. slope/roughness, wood density, water access). This is the dimension we exploit, as encampment decisions were made permanent later on.[13]

Even though households are legally entitled to use their plots, the land is still today the property of the state, which means that no land market can officially occur. Technical assistance to promote agricultural production was very weak in agrarian reform settlements in the area (Leite et al., 2004). In the municipality we study, settlers are excluded from family farming state programs or any other governmental credit plan. Finally, as will occur in the entire state of Alagoas, the village will be officially turned into an entity autonomous from the state at some point in future, which will entail that households become owners of their own land.

3 Data description

Our study is based on a unique survey covering an entire village in the north of Alagoas State in northeastern Brazil, which was conducted in two waves (2012 and 2018).[14] Our data include 100 settler households, based on the complete list of village inhabitants (roster) provided by local official representatives.[15] The data collection followed an ethnographic study that gathered detailed information on the local community.[16]

As described above, the village's formation was the result of the stratification of different migration waves characterized by two main groups: first settlers and newcomers. Figure

[13]Since the INCRA objective was to assign land in fair proportions, small adjustments in the size of parcels were done according to the geographical position and morphological characteristics of each plot.

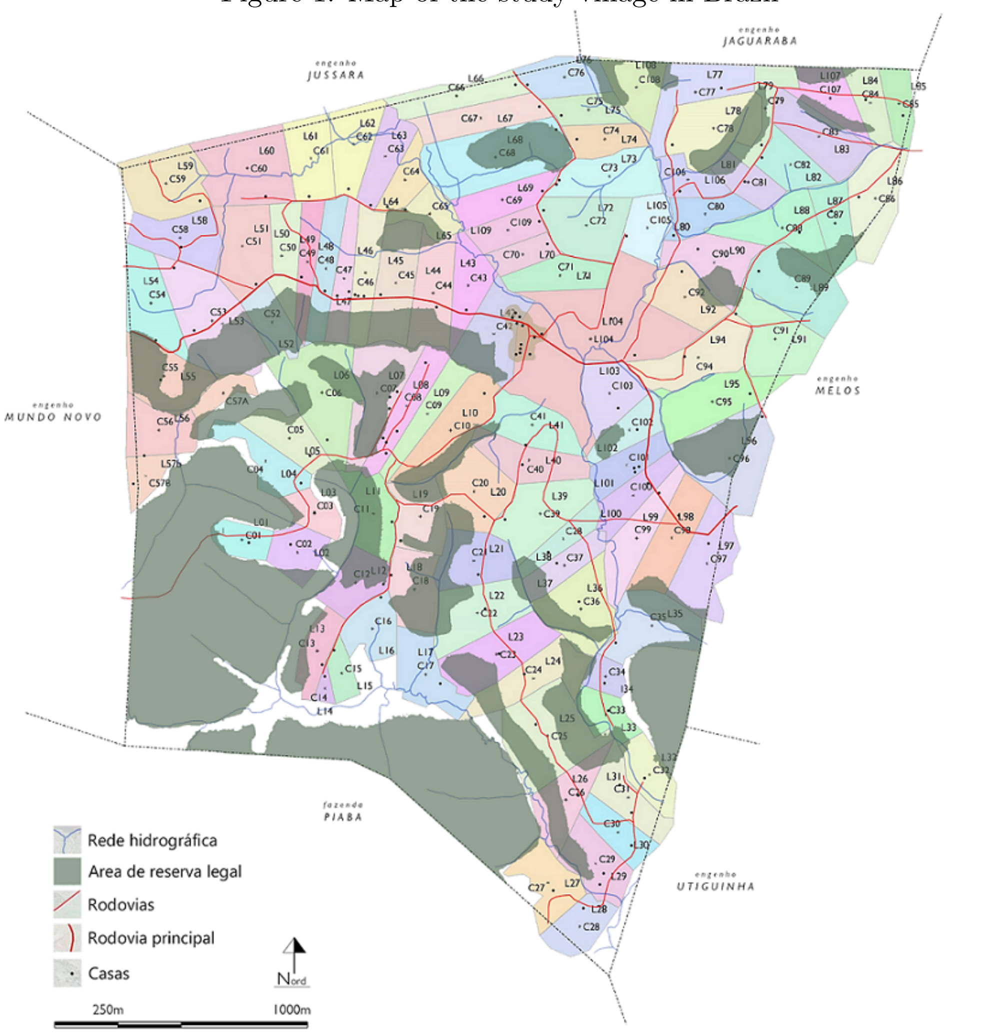
[14]A first wave of data collection took place between July and October 2012, gathering socio-demographic data, information about farming practices, and a census of links among settlers. A second visit took place in 2018, to collect retrospective farming information for the period 2012-2018.

[15]Local health service officers (*agentes de saude*) are hired by the municipality to provide basic health assistance on a weekly basis to all village households, and are the most reliable source of information. The official roster record is provided by INCRA, which registers households who were assigned a plot.

[16]Three months of fieldwork participatory observation and data collection preceded the actual survey (Inguaggiato, 2014). Face-to-face interviews were conducted in Portuguese, the only language used in the study context.

1 below depicts the division of plots within the village, along with protected (non-farmed) forest areas, roads and water sources. A detailed geo-morphological map of the study village along with village blueprint and geo-localized settlers are reported in Figure A1 and A2 in Appendix.

Figure 1: Map of the study village in Brazil



First settlers comprise 15 households already living in the village prior to the agrarian reform mobilization. Their houses are spread over the two main roads, where housing units were allocated by the *latifundio* administration at the time of the plantation (see black circles in Figure 2). 'Newcomers' refers to households mobilized by the agrarian reform and engaged in a social movement or in associations fighting for land rights. Their houses are spread along the plots, depending on their first spot of occupation in 1998. Among the newcomers, some households still have the same household composition as in 1998 ('occupiers', N=64) while other newcomer households experienced some rearrangement in their internal structure.[17] We call these latter newcomers 'substitutes' (N=21). Unfortunately our data do not provide insights on the reasons why some newcomers were replaced across time.[18] The large majority of households personally cultivate the plot they live on.[19]

Figure 2 reports the agrarian settlement land allocation map, as well as roads, rivers and the village center. It is worth noting the relatively equal distribution of land size (conditional on morphological characteristics) and, most importantly, the equally distributed locations of first settlers along the main village road. Figure 3 reports the geographical map of roads, river/water points and geo-coded settlers. In the next section we provide evidence that newcomers, who settled quickly during the encampment, may not have chosen their locations based on the other settlers at the site.

The social network information was collected at one point in time only (2012), and it maps all interpersonal links among all pairs of households (dyads) – this gives $(100*99)/2 = 4950$ undirected dyads (9900 directed dyads). Two types of network information were collected: social network, and labor exchange network. In order to elicit the social network, respondents were invited to nominate fellow inhabitants they meet on a regular basis by going to visit them.[20] We argue that this declared social relationship represents an opportunity for social learning among households. In what follows, this piece of information is used to define whether a link exists between farmers and to build the social network we use in our analysis.[21]

[17]Land cannot be sold, but over the years some original occupiers deceased or too old to work have been replaced as household heads by a family member, often a child raised in the village. In addition, some marriages have taken place among village residents, and some children have formed their own families while still cohabiting with their parents on the family plot.

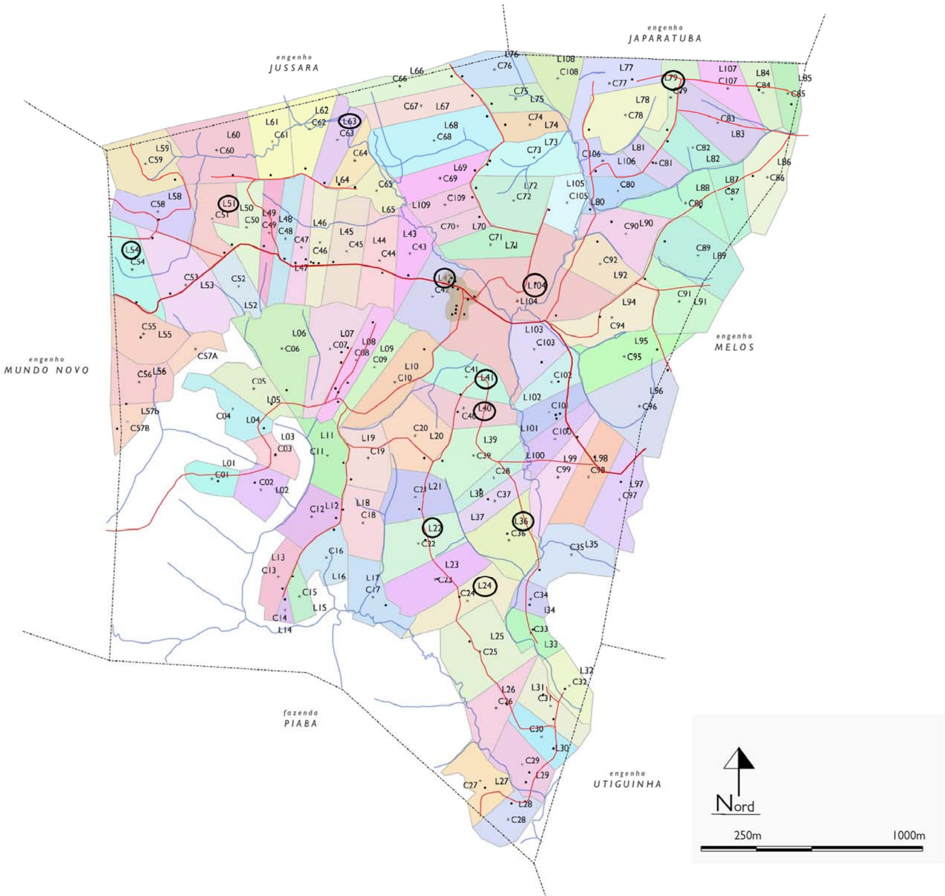
[18]Occupiers do not appear to have more kinship ties than other villagers, on average. This suggests that replacements were not related to the presence of blood links within the community, but the evidence we can gather from the data at hand remains largely suggestive.

[19]Some other households are primarily occupied in non-farming activities, leaving the plot uncultivated or giving it to others in exchange for compensation in cash or kind.

[20]The generator questions inserted in the interview were the following: *Are there some people that you frequently meet at home to talk? Who are they? Can I know their names?*. Respondents were asked to nominate other people residing in the village and there was no limit to the number of households that they could nominate.

[21]As a robustness check, we have also run the main estimates of Section 5 based on the (undirected) labor exchange network among villagers. Under the most conservative estimation strategy IV-B, our results

Figure 2: Agrarian settlement plot allocation map

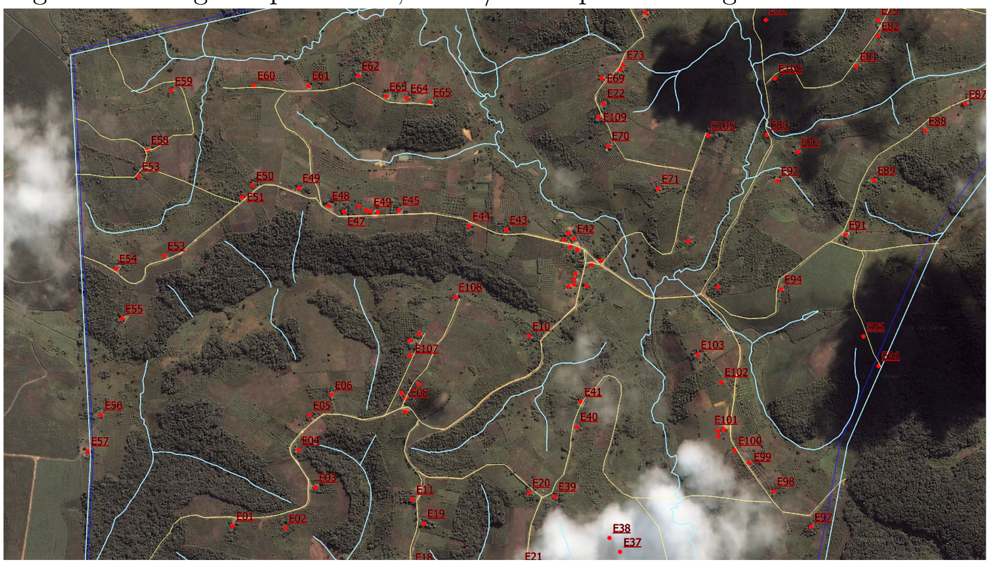


In addition, the survey collected detailed information on individual and household characteristics (e.g., demographics, age, education, religion), and occupational status. Importantly, a second wave of data collection took place in 2018 to gather information on adoption (and exit out of) cash fruit farming for all households since 2012. To do so, we re-interviewed all farmers and combined this piece of information with detailed registry records about quantities and price of transactions (by farmer and year). Hence, we have detailed farm production data (crop cultivated) for all settlers over a period of 15 years (2003 to 2017).[22] We combine this longitudinal data with detailed geo-localized and plot-

(available upon request) are in line with the findings of Table 5. However, our preferred specification relies on social ties because labor sharing arrangements are more likely to be contingent on the partners' farming background.

[22]The information about crops cultivated was collected from the registry of the local traders (e.g. the local

Figure 3: Village map of roads, rivers/water points and geo-localized settlers



level characteristics that allow us to map households' geographical positions with respect to the village center, water sources and roads throughout the area.

While focusing on our village household population, in Table 2 we report descriptive statistics for first settlers vs. newcomers (the latter, in turn, split between occupiers and substitutes). The last two columns report t-tests of the differences between groups. Unsurprisingly, first settlers appear to be different from newcomers along various dimensions (older male household head, less educated, more likely to be female-headed, more likely to already be in sugarcane business in 1996). However, there is no significant difference in the dimension of interest, namely geographical position with respect to the village center, water and roads, and the total area farmed. When we test the difference between first occupiers and substitutes, we see that the only significant difference is that substitute households are less likely to be female-headed, and when the head is a male, he is significantly younger. This is in line with anecdotal evidence suggesting the replacement of deceased first occupiers by younger family members. From now on, we pool together first occupiers and substitutes and consider all newcomers together (other than for robustness checks in Section 6).

The structure of the settlers' social network is depicted in Figures 4 and 5. Figure 4

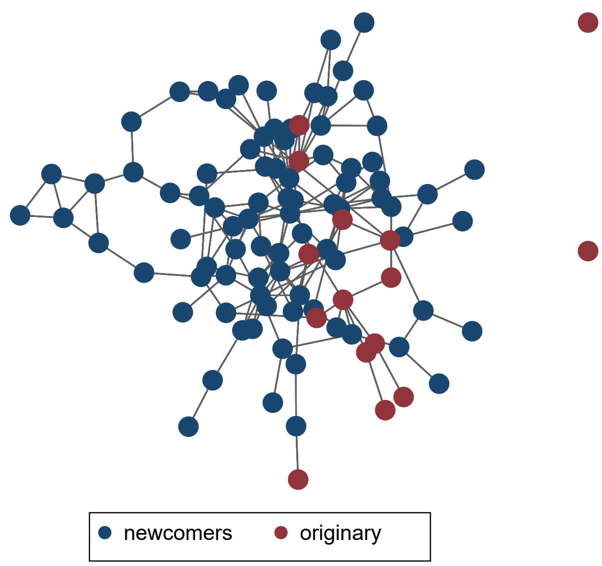
farming cooperative, operational since 2003) where settlers sell their products. Information information was complemented by self-declared information by settlers, and interviews with the local extension officers.

Table 2: Household descriptive statistics

	all (N=100)		Fist settlers (N=15)		newcomers (N=85)		occupiers (N=64)		substitutes (N=21)		t-test first settlers/ newcomers	t-test occupiers/ substitutes
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.		
age hh head	49.2	14.9	55.1	16.0	48.2	14.5	48.0	15.3	48.7	12.3	1.68*	0.17
female	0.1	0.3	0.3	0.5	0.1	0.3	0.0	0.2	0.2	0.4	3.11***	2.54**
income	814.8	525.9	661.7	376.3	841.8	545.3	839.2	501.3	849.8	676.4	-1.2	0.08
agr. income	466.3	422.9	314.8	325.4	493.1	433.9	472.2	435.7	556.8	432.7	-1.51	0.77
n. of years in village (male)	14.5	9.0	30.6	12.5	11.7	3.9	13.5	1.9	6.3	3.4	11.38***	-12.21***
years of schooling	4.3	3.4	2.5	2.3	4.6	3.5	4.6	3.7	4.4	2.5	-2.2**	-0.22
hh farming	0.8	0.4	0.7	0.5	0.8	0.4	0.8	0.4	0.8	0.4	-1.03	-0.34
in sugar cane since '96	0.3	0.5	0.5	0.5	0.3	0.5	0.3	0.5	0.3	0.5	1.94*	0.04
distance house center	972.9	501.3	805.3	498.0	1002.4	498.9	1003.1	480.2	1000.5	564.9	-1.41	-0.02
distance house water	336.7	249.8	258.3	158.2	350.5	260.9	362.5	255.0	314.0	281.5	-1.32	-0.74
distance house road	190.1	180.5	150.4	143.4	197.1	186.1	202.1	187.8	182.0	184.4	-0.92	-0.43
distance center	99.5	48.2	84.3	49.8	102.2	47.7	102.6	45.6	101.0	54.7	-1.33	-0.13
sum plots cultiv.	4.8	4.1	3.6	3.7	5.0	4.2	5.0	4.3	4.8	3.9	-1.2	-0.23

reports the undirected social networks among settlers based on the question about house visits (N=4950), with 168 links among 100 nodes (about 3 links per household on average).[23] Figure 5 plots the directed network of labor exchange (N=9900), with 118 links among 100 nodes. Both graphs reflect well-known stylized facts about social networks, known as 'small-world properties', which were initially pointed to by sociologists and physicists (Watts and Strogatz, 1998), namely: virtually all households are indirectly connected through a so-called 'giant component', the average number of contacts is limited, the average distance between two households (computed as the number of steps in the shortest path along the graph) is short, and two households are more likely to be linked if they share a common friend.[24]

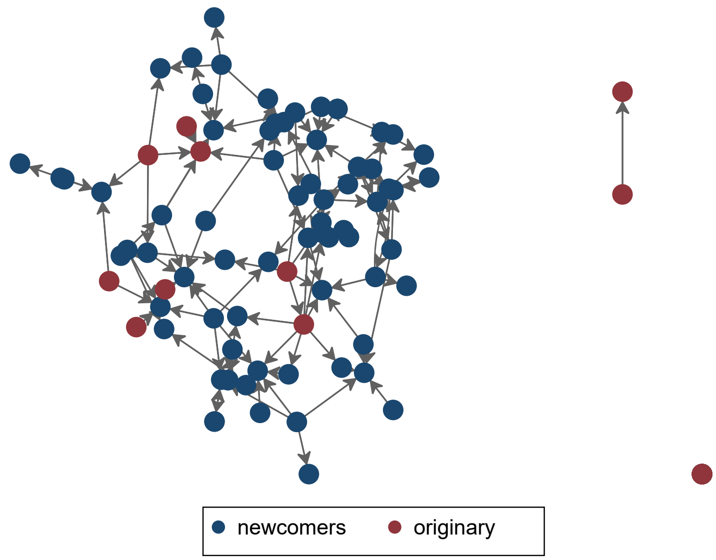
Figure 4: Undirected social network



[23]As all households were interviewed separately, we have two reports on the link between A and B (by A and B, respectively) that, in principle, may not coincide. When this is the case, we take the maximum report out of the two sides involved. This corresponds to an implicit assumption of under-reporting by one of the sides, because of omission and mistakes. This is a standard approach in the literature dealing with self-reported network data (Fafchamps and Soderbom, 2011; Liu et al., 2014; Banerjee et al., 2013; Comola and Fafchamps, 2014, 2017).

[24]The undirected contact graph of Figure 4 has one giant component with 98 nodes out of 100, density of 0.03, and average clustering coefficient of 0.18. The directed labor exchange graph of Figure 5 has one giant component with 81 nodes, density of 0.01, and an average clustering coefficient of 0.02.

Figure 5: Directed labour-exchange network



4 Location choice and social ties

Assortativity is the pervasive tendency to match with agents with similar characteristics, which plays a confounding role in the study of social diffusion. Imagine we observe social links among potential adopters of a new technology and notice that individuals tend to buy into the technology if their peers do so. This comes as no surprise and suggests a strong correlation between social networks and behavior. But we cannot straightforwardly conclude that these observed effects are causal, because peers share the same (observable or unobservable) attributes which could simultaneously drive link formation and adoption. When networks are formed spontaneously as a result of individual characteristics and tastes only, researchers may partially circumvent assortativity by using group-level fixed effects (Bramoullé et al., 2009) or by explicitly modeling the process of link formation (Goldsmith-Pinkham and Imbens, 2013). More convincing identification, however, can be achieved when the formation of interpersonal links depends on exogenous shocks or random allocations of individuals to locations or groups (e.g. forced migration following natural disasters, students allocated randomly to dorms, refugees placed in housing units in displacement camps, or newly recruited soldiers randomly attributed to military units).[25]

[25] Beaman (2012), Laschever (2009).

In our context, we argue that the resettlement process provides a source of variation in the location choice of households. In fact, the encampment phase, which was made perennial and legal later on, was hurried and mostly the result of chance. In this section, we provide evidence that location choices do not correlate with the observable attributes of settlers. Secondly, we show how location is a strong predictor of social network formation. Taken together, these results support the estimation strategy that we pursue in Section 5.

The process of land occupation by stranger households, as described in Section 2, in principle did not accommodate for location choice based on attitudes and tastes. However, one may still think that, to some extent, occupiers at the time of their arrival chose their location strategically in order to settle next to people with similar profiles. In order to convince the reader that this is not the case, we restrict the analysis to the dyads made by two occupiers (N=(64*63)/2=2016) and run the undirected dyadic regression[26]

$$proximity_{ij} = \alpha + \beta X_{ij} + \varepsilon_{ij}$$

where $proximity_{ij}$ is the geographical distance between two households (in km), and X_{ij} is a vector of undirected dyadic regressors that represents observable characteristics which could drive location choices.[27] Standard errors ε_{ij} are corrected for dyadic dependence (Fafchamps and Gubert, 2007). Dyadic regressors included in X_{ij} are a set of dummies equal to one if: the two households share the same hometown, were recruited by the same mobilizer,[28] had the same professional status before arrival in 1996 (agricultural sector, sugarcane, autonomous worker, unemployed), practice the same religion[29] respectively. Results are reported in Table 3, and show that proximity does not correlate significantly with any of these factors.

We then document the fact that location is a driving force in link formation. In table

[26]Location assortativity concerns do not apply to dyads composed by two households of first settlers as their housing units were allocated centrally by the administrative office of the *latifundo* well before the 1998 occupation. It also does not concern dyads made by first settlers and newcomer households, as there were no links between occupiers and plantation workers before the occupation.

[27]Since the dyadic relationship is undirected, the regressors must enter in a symmetric fashion (e.g. dummy variables, sum and/or absolute difference of continuous attributes).

[28]We have information on the identity of the person who recruited the settler candidates just prior to 1998.

[29]We say that two households practice the same religion if they both attend the catholic Church or an evangelical congregation at least once a month.

Table 3: Location Choice (proximity)

	(1)	(2)	(3)
both agriculture	-0.05	-0.03	-0.04
	(0.07)	(0.07)	(0.07)
both sugarcane	0.10	0.11	0.10
	(0.07)	(0.07)	(0.07)
both autonomous		-0.15	-0.16
		(0.12)	(0.12)
both unemployed		0.10	0.11
		(0.11)	(0.11)
same hometown		0.05	0.05
		(0.16)	(0.16)
same mobilizer		-0.20	-0.18
		(0.15)	(0.16)
same religion			-0.15
			(0.12)
const	1.34***	1.34***	1.36***
	(0.08)	(0.08)	(0.08)
Obs (# unique dyads)	2,016	2,016	2,016

Note: Dyadic s.e. are reported in brackets

*** p<0.01, ** p<0.05, * p<0.1

4 we report the results from the dyadic regression

$$link_{ij} = \alpha + \beta\, proximity_{ij} + \gamma Z_{ij} + \varepsilon_{ij}$$

where $link_{ij}$ is the undirected measure of a social link between two households and Z_{ij} is a vector of undirected dyadic controls representing socio-economic factors that have been shown to affect link formation, including: a dummy equal to one if the two households come from the same hometown, the absolute difference in the time of arrival,[30] absolute difference in age of household head, household size, years of schooling of the household head, and dummies if both households rely uniquely on farming for living and if they have the same religion, respectively. Standard errors ε_{ij} are corrected for dyadic dependence, as above.

The full sample (cols. 1–3) includes all dyads. In columns 4 to 6 we restrict the sample by dropping the dyads formed by two first settlers.[31] In columns 7 to 9, we restrict the sample further by dropping dyads including substitutes.[32]

Results show that geographical proximity is a strong predictor of link formation throughout. In addition, households are more likely to be linked if they come from the same locality, arrived at the same time, or practice the same religion.

5 Social learning and adoption

In this section, we present the main results of our analysis, which relate to the crop adoption behavior. We use longitudinal information on the adoption of two cash crops (pineapple and passion fruit), that were rarely farmed at the beginning of the period in the area, and have spread widely in the village community along the years of our study. They are both commercial tropical fruits with relatively high labor demand but high profit margins (as fresh whole fruit or for juice processing). Pineapple is a perennial plant that bears a single fruit growing in about 18 months, while passion fruit is a vine whose production cycle is faster and that can last up to seven years. For these reasons, tropical fruit adoption involves a longer-term investment and higher sunk costs than does the seasonal cultivation

[30] First settlers and substitutes households did not arrive in 1998.

[31] N=(100*99)/2-(15-14)/2=4845

[32] We have no information on the reasons why some replacements occurred within the 15-year span of our study (Section 3). However, results in columns 7 to 9 reassure the reader that the effect of geographical proximity is not driven by (unobserved heterogeneity of) substitute households.

Table 4: Link Formation

	full sample			no first settlers			no first settlers/substitutes		
	(1)	(2)	(3)	(4)	(5)	(3)	(7)	(8)	(9)
proximity	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***	-0.04***	-0.05***	-0.05***	-0.05***
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
same hometown		0.04***	0.04***		0.04***	0.04***		0.03**	0.03**
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
difference time arrival		-0.00***	-0.00***		-0.00***	-0.00***		-0.00***	-0.00***
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
difference age		-0.00	-0.00		-0.00	-0.00		-0.00	-0.00
		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
diff size			-0.00			-0.00			-0.00
			(0.00)			(0.00)			(0.00)
difference schooling			0.00			0.00			-0.00
			(0.00)			(0.00)			(0.00)
both farming			0.01*			0.01*			0.01
			(0.01)			(0.01)			(0.01)
same religion			0.04**			0.05**			0.07**
			(0.02)			(0.02)			(0.03)
constant	0.09***	0.09***	0.09***	0.09***	0.09***	0.08***	0.10***	0.10***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	4950	4950	4950	4845	4845	4845	2976	2976	2976

Note: Dyadic standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1

of grains, roots and tubers. Both fruits thrive in many different soil types but while tropical areas are most suitable for pineapple plantation also with low water availability, passion fruit is more sensitive to water stress and deficits, especially during leaf production and flowering (Carr, 2013).

Our dataset follows the villagers during a period of 15 years (2003 to 2017), and contains a binary indicator of whether the household has cultivated each of these two crop for each year of the panel. The aggregate adoption patterns for the crops are displayed in Figure 6. While the adoption of pineapple steadily rises over time from 1 to 15% (average of 5.6% over the years of study), passion fruit is increasingly farmed up to 2010 and dropped afterwards to its initial rate at the beginning of the period. Variation over time of farming choices is what we exploit in our panel estimation model below.

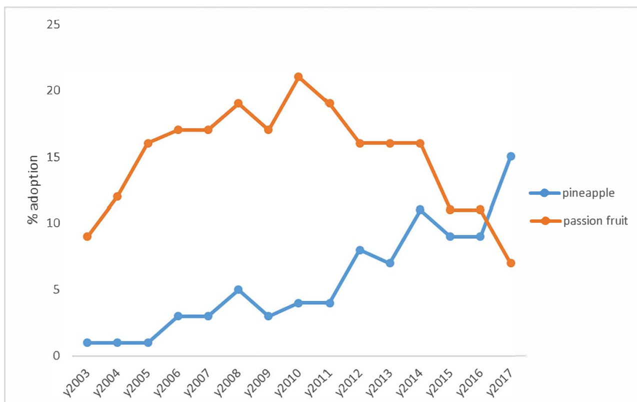


Figure 6: Adoption patterns

5.1 The model

Our aim is to document how adoption behavior spills over from peers, that is how the farming choices of peers have a positive impact on one's own choice. We frame our problem in the context of peer effects working through the network structure of social interactions. We leverage the improptu location choice at the time of settlement to address network assortativity, as proximity during the encampment has been shown to be a valid predictor for link formation.

In what follows, vectors are denoted with bold lower-case letters and matrices with bold capital letters. Let us consider a panel data with N households observed over multiple periods $t = 0, ..., 15$. Define $\mathbf{y}_c^t$ as the $n \times 1$ vector of binary outcomes representing the adoption outcome of crop c at time t, e.g. $y_{c,k}^t = 1$ if household k is cultivating crop c at time t.[33] Similarly, $\mathbf{x}^t$ contains time-varying household characteristics. Let the $n \times n$ interaction matrix $\mathbf{G}$ represent the social interaction within the sample. For the scope of our study, we assume that $\mathbf{G}$ is binary and undirected ($g_{ij} = g_{ji} = 1$ if households i and j are socially connected), and time-invariant across periods (i.e. $\mathbf{G}^t = \mathbf{G}$ for all t).[34] Our model of reference is represented by the linear equation

$$\mathbf{y}_c^t = \beta \mathbf{G}\mathbf{y}_c^t + \gamma \mathbf{x}_c^t + \delta \mathbf{G}\mathbf{x}_c^t + \boldsymbol{\lambda}_t + \boldsymbol{\mu} + \boldsymbol{\epsilon}_c^t \tag{1}$$

where the dependent variable $\mathbf{y}_c^t$ is binary,[35] and the 'first lag' of the dependent variable $\mathbf{G}\mathbf{y}_c^t$ is called 'endogenous' peer effect and represents the number of peers cultivating crop c at time t. Similarly, $\mathbf{x}_c^t$ represent household's exogenous attributes that could affect its adoption choice, and $\mathbf{G}\mathbf{x}_c^t$ represents the so-called 'contextual' peer effect, namely the (sum of) exogenous attributes of peers.[36] $\boldsymbol{\lambda}_t$ represents a set of year-level dummies, and we denote by $\boldsymbol{\mu} = (\mu_1, ..., \mu_N)'$ the vector of household-level effects, which we treat as fixed effects as they may be correlated with the regressors. Including both household and year fixed effects allow us to control for all confounding unobservables that are time-invariant within the duration of our study (including risk attitude and innovation propensity – all attributes that could lead to network assortativity if unaccounted for), as well as all unobservable time-varying shocks that are common to all farmers (e.g. climate, income or market-access shocks). The vector of disturbances $\boldsymbol{\epsilon}_c^t$ is auto-correlated across periods.

[33] We have also tried to single out peer effects on dis-adoption behavior by focusing on the sub-sample of households who already cultivated crop c in year $t-1$. Unfortunately, the small sample size did not allow us to recover significant peer-effect estimates.

[34] Choosing to use a binary interaction matrix implies that we estimate a linear-in-sums model, that is, an interaction model in which the household outcome is affected by the total number of peers who adopt (rather than their share). This is justified by a large literature on social adoption along with some game-theoretical reasoning (Calvó-Armengol et al., 2009; Liu et al., 2014). However, our estimation strategy could be applied to a peer effect model where the network matrix's rows sum up to unity (linear-in-means), obtaining results that are qualitatively similar.

[35] A linear probability model was also used by Fortin and Yazbeck (2015) in the context of binary-outcome network data.

[36] In the terminology of Manski (1993), $\mathbf{G}\mathbf{y}_c^t$ would be called 'endogenous social effects' and $\mathbf{G}\mathbf{x}_c^t$ 'exogenous social effects'. We include contextual effects for all exogenous attributes to avoid unjustified exclusion restrictions in relation with the instrumentation strategy.

5.2 Instrumentation strategy

All interaction models exploiting network data raise endogeneity concerns which relate to simultaneity, i.e. to the fact that the outcomes of partners may be jointly determined. Note that this concern is unrelated to network assortativity, and it would also arise if social links were distributed entirely at random. It simultaneity invalidates OLS inference to the extent to which the term $\mathbf{G}\mathbf{y}^t$ in Equation (1) is correlated with the disturbance vector.

The standard solution to address endogeneity stemming from simultaneity is to use 'lagged' partners' characteristics (that is, the exogenous attributes of the partners of one's partners) as instruments (e.g. Bramoullé et al., 2009; Calvó-Armengol et al., 2009; Drukker et al., 2013; Patacchini and Zenou, 2012; Comola and Prina, 2021). In fact, as long as there are individuals who are excluded from one's own reference group but are included in the reference group of one's partners, their exogenous characteristics can affect one's outcome only through the partners and thus are a natural set of instruments to address the reflection problem (Manski, 1993). In our context, this boils down to estimating Equation (1) via a 2SLS instrumental variable techniques where peer behavior $\mathbf{G}\mathbf{y}_c^t$ is taken as endogenous, and the exogenous attributes of peers of peers $\mathbf{G}^2\mathbf{x}_c^t$ (second-order lagged characteristics) and of their peers $\mathbf{G}^3\mathbf{x}_c^t$ (third-order lagged characteristics) are used as excluded instruments.[37] We call this instrumentation strategy 'IV-A'. Since model includes fixed effects at the level of the household, it addresses endogeneity concerns stemming from network assortativity as long as unobserved correlates are time-invariant across the spam of the panel (e.g., risk attitude).[38]

However, given that the social interaction matrix is measured at one point in time only, one can still be concerned about endogeneity stemming from time-varying unobserved shocks.[39] In order to address this additional concern, we augment our instrumentation

[37]Thus, the first-stage equation writes as $\mathbf{G}\mathbf{y}_c^t = \theta_1\mathbf{x}_c^t + \theta_1\mathbf{G}\mathbf{x}_c^t + \theta_3\mathbf{G}^2\mathbf{x}_c^t + \theta_4\mathbf{G}^3\mathbf{x}_c^t + \boldsymbol{\lambda}_t + \boldsymbol{\mu} + \boldsymbol{\epsilon}_c^t$. Instruments of higher order can be added at the cost of some additional requirements. See Bramoullé et al. (2009) for the identification condition and moment restrictions that also apply to our context.

[38]If we had cross-sectional data we could only condition exogeneity of the interaction matrix on network-level fixed effects as in Bramoullé et al. (2009), which is a valid identification strategy only to the extent that the correlated unobservables are common to the entire network. Thanks to the longitudinal nature of our data, we can actually condition exogeneity of the interaction matrix on the household-level effects and the time dummies (see Comola and Prina, 2021). Formally, this implies that we can assume $E[\boldsymbol{\epsilon}_c^{t_A}|\mathbf{G}, \mathbf{x}_c^{t_B}, \boldsymbol{\lambda}_t, \boldsymbol{\mu}] = 0 \quad \text{for all } t_A \neq t_B$. That is, even the (least demanding) empirical strategy 'IV-A' does not confound the effect of social links with household fixed characteristics, or time trends.

[39]This refers to cases where unobserved time-varying factors affects social interactions and individual outcomes. For instance, one can think of situations where farm-related shocks occurred before 2012 drove the observed linking patterns.

strategy to take into account the potential endogeneity of the network at the time it was measured. In Section 4 we argued that proximity between households resulting from the land squatting campaign was not driven by assortativity along observables. We now use geographical proximity between households to instrument for the interaction matrix G. The validity of this strategy relies on the identification assumption that the resettlement process was a source of exogenous variation in the location choice. Operationally, we proceed as follows: we predict social links on the basis of proximity, and we use the attributes of *predicted* lagged partners as an instrument for the *observed* behavior of partners. This estimation strategy, that we call 'IV-B', has been formally proved to be valid (i.e. consistent and asymptotically normal) for endogenous interaction matrices.[40] In practice, this requires computing the lagged partner characteristics as in strategy IV-A but replacing the observed network with its fitted version predicted on the basis of proximity.[41]

5.3 Main results

Table 5 reports our main results.[42] In our chosen specifications, the vector of time-varying attributes $\mathbf{x}_c^t$ includes three variables: the first variable is $\boldsymbol{distance\ road} \times \boldsymbol{price}_c^t$, where $distance\ road_i$ is the distance from the road of the plot of residence of farmer i,[43] and $price_c^t$ is the market price of crop c (pineapple and passion fruit respectively) for year t.[44] Similarly, we include $\boldsymbol{distance\ water} \times \boldsymbol{price}_c^t$, where $distance\ water_i$ is the distance from the plot of residence of farmer i to water. These two variables are meant to represent time-varying incentive factors shaping farming adoption decisions. While we expect farmers further away from either water or roads to be less likely to adopt cash crops (the distance, being time-invariant, is absorbed by fixed effects), we expect the differential effect of distance to change as long as crop-specific (time-varying) prices increase. Finally, the dummy $coop\ member_i^t$ equals one if i's head of household is member of the village cooperative in year t.

In columns (1) and (4) of Table 5 we report for reference the OLS estimates from

[40] Kelejian and Piras (2014); Hsieh and Lee (2016).

[41] We take the following steps: 1) run a cross-section dyadic regression of link formation on proximity (plus a constant term), and construct the 'fitted' interaction matrix $\widehat{\mathbf{G}}$; 2) compute the attributes of *predicted* lagged second- and third-order partners $\widehat{\mathbf{G}}^2\mathbf{x}_c^t$ and $\widehat{\boldsymbol{G}}^3\mathbf{x}_c^t$; 3) estimate the 2SLS model in Equation (1) with $\mathbf{G}\mathbf{y}_c^t$ as endogenous and $\widehat{\mathbf{G}}^2\mathbf{x}_c^t$ and $\widehat{\boldsymbol{G}}^3\mathbf{x}_c^t$ as excluded instruments. The first-stage equation now writes as $\mathbf{G}\mathbf{y}_c^t = \theta_1\mathbf{x}_c^t + \theta_1\mathbf{G}\mathbf{x}_c^t + \theta_3\widehat{\mathbf{G}}^2\mathbf{x}_c^t + \theta_4\widehat{\boldsymbol{G}}^3\mathbf{x}_c^t + \boldsymbol{\lambda}_t + \boldsymbol{\mu} + \boldsymbol{\epsilon}_c^t$.

[42] Descriptive statistics (for regressors and excluded instruments) are reported in the Appendix Table A7.

[43] Precisely, the distance is computed from the house located on the plot. We have recalculated all distances from plot centroids, for a virtually identical result.

[44] This is the yearly average selling price for raw (i.e. non-processed) fruit.

Table 5: Main results

	pineapple			passion fruit		
	(1) OLS	(2) IV - A	(3) IV - B	(4) OLS	(5) IV - A	(6) IV - B
$\mathbf{Gy}_c^t$	0.03**	0.09**	0.39***	0.07***	0.14***	0.25***
	(0.01)	(0.04)	(0.12)	(0.01)	(0.04)	(0.08)
$\boldsymbol{distance\ road} \times \boldsymbol{price}_c^t$	-0.18***	-0.18***	-0.16**	0.12	0.09	0.04
	(0.06)	(0.06)	(0.08)	(0.10)	(0.10)	(0.11)
$\boldsymbol{distance\ water} \times \boldsymbol{price}_c^t$	0.01	0.04	0.20**	-0.01	-0.03	-0.07
	(0.04)	(0.05)	(0.09)	(0.07)	(0.08)	(0.08)
$\boldsymbol{coop.\ member}^t$	0.22***	0.22***	0.23***	0.59***	0.59***	0.59***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
$\mathbf{G}\,(\boldsymbol{distance\ road} \times \boldsymbol{price}_c^t)$	0.06**	0.08***	0.15***	-0.04	-0.04	-0.04
	(0.03)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)
$\mathbf{G}\,(\boldsymbol{distance\ water} \times \boldsymbol{price}_c^t)$	-0.02	-0.04*	-0.13***	-0.01	0.01	0.03
	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)
$\mathbf{G}\,(\boldsymbol{coop.\ member}^t)$	-0.02	-0.03**	-0.09***	-0.04**	-0.09***	-0.15***
	(0.01)	(0.02)	(0.03)	(0.02)	(0.03)	(0.05)
household f.e.	yes	yes	yes	yes	yes	yes
year f.e.	yes	yes	yes	yes	yes	yes
R-squared	0.13	-	-	0.36	-	-
Weak identif. test	-	32.1	4.8	-	33	8.7
Observations	1,500	1,500	1,500	1,500	1,500	1,500
Number of id	100	100	100	100	100	100

NOTES: Autocorrelation-robust standard errors in parentheses (kernel=Bartlett; bandwidth=3). Cragg-Donald Wald F statistic for weak identification reported. Statistically significant coefficients are indicated as follows: * 10%, ** 5% and *** 1%.

the specification of Equation (1) for pineapple and passion fruit – following the discussion above, these estimates are biased. Columns (2) and (5) report instrumental variable results from estimation strategy IV-A, where $\mathbf{Gy}_c^t$ in Equation (1) is instrumented with the lagged-partner characteristics $\mathbf{G}^2\mathbf{x}_c^t$ and $\mathbf{G}^3\mathbf{x}_c^t$. Columns (3) and (6) report instrumental variable results from the estimation strategy IV-B, where $\mathbf{Gy}_c^t$ in Equation (1) is instrumented with the *fitted* counterparts of the excluded lagged-partner instruments, namely $\widehat{\mathbf{G}}^2\mathbf{x}_c^t$ and $\widehat{\boldsymbol{G}}^3\mathbf{x}_c^t$.

Results from Table 5 suggest a positive and significant peer effect in the adoption process of pineapple and passion fruit: according to specification IV-A having one additional peer cultivating the crop in year t increases the probability of adoption by 9 percentage points for pineapple and 14 percentage points for passion fruit. In specification IV-B the peer effect coefficient remains positively significant, and larger in magnitude (39 and 25 percentage points respectively). In models with spatial lags in the dependent variable, the interpretation of the estimated parameters is enriched (and complicated) by the structure of social interactions. Our results suggest that *ceteris paribus* a farmer would increase his propensity to crop pineapple by 3 - 12% if a shock were to double the adoption rate among her contacts. The corresponding number for passion fruit is 10 - 18%.[45] Cooperative membership has a positive effect in all specifications, while the remaining characteristics have an heterogeneous effect depending on the crop. In particular, we find a negative differential effect of market price for pineapple along distance to the road (but not to water source), pointing to a smaller price elasticity of pineapple supply the further away the farmer is from main connecting road. This turns out not to be the case for passion fruit. This is consistent with the delicate and costly endeavor of pineapple logistics, which need to be moved along the shortest possible transit time from farm to market in order to retain their freshness.[46] All specifications suggest that the contextual effects (i.e. the effect of peers characteristics) contribute to explain the process of social diffusion, and appear to be heterogeneous across crops.

The increase magnitude of estimates in columns (3) and (6) is in line with the empirical literature showing that the coefficient of endogenous peer effects tend to be larger when instrumented,[47] but it could also be partly due to the fact that the instrumentation strategy IV-B is extremely exigent in the context of a fixed-effect panel data with relatively few observations with respect to the time span, weakening the instruments power.[48]

---

[45]Farmers in our sample have on average 0.31 friends cropping pineapple over the whole span of the panel. This implies an increase of 0.03 for column 2 (IV-A) and of 0.12 for column 3 (IV-B). Similarly, since the average number of friends cropping passion fruit is 0.71 we get an increase of 0.10 for column 5 (IV-A) and of 0.18 for column 6 (IV-B).

[46]Unlike passion fruit, pineapples will not ripen anymore once harvested and after their prime time they degenerate quickly.

[47]Since assortativity with network data relates to the observed connectivity patterns of sub-groups of observations, the standard omitted-variable argument to predict the direction of the bias does not apply in a straightforward manner.

[48]The standard deviation of instruments under strategy IV-B is smaller than under strategy IV-A because the fitted matrices $\widehat{\mathbf{G}}^2$ and $\widehat{\mathbf{G}}^3$ are weighted (Table A7). Still, the Cragg-Donald statistics reported in Table 5 for IV-B are close to the critical values computed by Stock and Yogo (2005) for a scenario without social lags in the dependent variable. In fact, the statistic for column 3 is just below the threshold of 30% maximal bias of the IV estimator relative to OLS (cutoff at 5.15), and the statistics for column 6 meets the 20% threshold (cutoff at 6.76).

6 Alternative specifications

This section is devoted to discussing alternative specifications of our baseline model.

6.1 Lags in peer effects

Since there is no *a priori* motivation to believe that social effects are strictly contemporaneous, we have re-estimated the IV models of Equation (1) by assuming one lag in the peer effects. This corresponds to the estimating equation

$$\mathbf{y}_c^t = \beta \mathbf{G} \mathbf{y}_c^{t-1} + \gamma \mathbf{x}_c^t + \delta \mathbf{G} \mathbf{x}_c^{t-1} + \boldsymbol{\lambda}_t + \boldsymbol{\mu} + \boldsymbol{\epsilon}_c^t \tag{2}$$

As above, the instrumentation strategy IV-A relies on the characteristics at time $t-1$ of peers of peers ($\mathbf{G}^2\mathbf{x}_c^{t-1}$) and their peers ($\mathbf{G}^3\mathbf{x}_c^{t-1}$), and instrumentation strategy IV-B relies on the characteristics at time $t-1$ of expected peers of peers ($\widehat{\mathbf{G}}^2\mathbf{x}_c^{t-1}$) and their peers ($\widehat{\mathbf{G}}^3\mathbf{x}_c^{t-1}$). Estimated coefficient for β reported in Table 6–Panel A go in the same direction (for sign, significance and magnitude) as the main results of Table 5.

6.2 Heterogeneous social effects

Last but importantly, we explore the heterogeneity of social effects. In fact, one may imagine that, for the scope of social learning in agriculture, not all peers are the same. Munshi (2004), for instance, highlights the effectiveness of learning from others who share similar (observed and unobserved) characteristics in the context of the Green Revolution. To explore this hypothesis, we split the sample into two groups according to the ease of water access from the plot. On the one hand, we know for certain that water access is an important determinant of crop selection and agricultural success, especially in tropical and sub-tropical climates. On the other hand, as we can already see from the figures in Section 3, there is considerable heterogeneity in plot locations with respect to water, and two households located within a reasonable distance can have important differences in terms of water access. This is also reflected in our social network data: if we split the household sample in two groups according to water access (above/below the sample median of distance from water, which is 240 meters), we notice that out of the 168 (undirected) social links among farmers, 99 of them are between households of the same type (47 links between households both above the median, 52 links between households both below the median), while 69 links are between households of different types (one above and the other below the median).

In what follows, we test the hypothesis that social learning is stronger among farmers with similar plot characteristics in terms of water access. In order to do so, we generalize the linear social interaction model of Equation (1) by allowing for heterogeneous peer effects according to plot type (above/below the median distance to water). Our estimating equation becomes:

$$\mathbf{y}_c^t = \beta_S \mathbf{G}_S \mathbf{y}_c^t + \beta_D \mathbf{G}_D \mathbf{y}_c^t + \gamma \mathbf{x}_c^t + \delta_S \mathbf{G}_S \mathbf{x}_c^t + \delta_D \mathbf{G}_D \mathbf{x}_c^t + \boldsymbol{\lambda}_t + \boldsymbol{\mu} + \boldsymbol{\epsilon}^t \tag{3}$$

where we now define two distinct social interaction matrices representing links between peers of the same type ($\mathbf{G}_S$) versus peers of different types ($\mathbf{G}_D$).[49] By construction, $\mathbf{G}_S + \mathbf{G}_D = \mathbf{G}$. Thus, the so-called endogenous peer effects $\mathbf{G}_S \mathbf{y}^t$ ($\mathbf{G}_D \mathbf{y}_c^t$) represent the number of peers of the same (different) type that cultivate crop c at time t. Similarly, the contextual peer effects $\mathbf{G}_S \mathbf{x}^t$ ($\mathbf{G}_D \mathbf{x}_c^t$) represent the attributes of peers of the same (different) type at time t. As above, $\boldsymbol{\lambda}_t$ and $\boldsymbol{\mu}$ are year- and household-level fixed effects, and ${\boldsymbol{\epsilon}_c}^t$ is auto-correlated across periods. Both instrumentation strategies are expanded to take into account the split of the interaction matrices, namely, for IV-A we take $\mathbf{G}_S^2 \mathbf{x}^t, \mathbf{G}_D^2 \mathbf{x}_c^t, \mathbf{G}_S^3 \mathbf{x}^t, \mathbf{G}_D^3 \mathbf{x}_c^t$ and for IV-B we take $\widehat{\mathbf{G}}_S^2 \mathbf{x}^t, \widehat{\mathbf{G}}_D^2 \mathbf{x}_c^t, \widehat{\mathbf{G}}_S^3 \mathbf{x}^t, \widehat{\mathbf{G}}_D^3 \mathbf{x}_c^t$ as excluded instruments for the endogenous terms $\mathbf{G}_S \mathbf{y}_c^t, \mathbf{G}_D \mathbf{y}_c^t$.[50]

Estimated coefficient for β_S, β_D reported in Table 6 – Panel B show an interesting pattern of heterogeneity with respect to crops: for pineapple, both types of friends contribute to the same way to the social learning process (both endogenous peer effect coefficients are significant and of same magnitude which is comparable to the main results of Table 5). However, for passion fruit we notice that the peer effect reported in Table 5 flow through the network of peers of a similar type: the estimated coefficient for $\mathbf{G}_S \mathbf{y}^t$ is significant and slightly larger than the homogeneous effect reported in Table 5, while the coefficient for $\mathbf{G}_D \mathbf{y}_c^t$ appears non-significant. This is an interesting results, which is likely to relate to the different water needs of the two cash crops, lower for pineapple than for passion fruit (Carr, 2013; De Azevedo et al., 2007). Note that our estimation strategy includes

[49] $g_{S_{ij}} = g_{S_{ji}}$ equals one if households i and j are socially connected and they are of the same type, and zero otherwise. Conversely, $g_{D_{ij}} = g_{D_{ji}}$ equals one if households i and j are socially connected and they are of different types, and zero otherwise.

[50] This follows from the identification conditions and instrumentation strategy formalized by Comola et al. (2023) for the general heterogeneous social network model.

Table 6: Robustness checks

	(1)	(2)	(3)	(4)
	pineapple		passion fruit	
	IV - A	IV - B	IV - A	IV - B
	Panel A - Lagged social effects			
$\mathbf{Gy}_c^{t-1}$	0.10**	0.47***	0.13**	0.29***
	(0.04)	(0.16)	(0.05)	(0.11)
	Panel B - Hetorogenous effects			
$\mathbf{G}_S\mathbf{y}^t$	0.17***	0.36***	0.14***	0.35***
	(0.05)	(0.13)	(0.05)	(0.11)
$\mathbf{G}_D\mathbf{y}^t$	0.17*	0.39**	-0.04	0.06
	(0.09)	(0.17)	(0.08)	(0.12)

NOTES: Household and year fixed effects are included. Autocorrelation-robust standard errors in parentheses (kernel=Bartlett; bandwidth=3). Statistically significant coefficients are indicated as follows: * 10%, ** 5% and *** 1%.

household-level effects which account for time-invariant unobserved characteristics, such as farmer attitude and plot attributes. Once we control for these time-invariant factors, we still find that farmers learn crops which require high level of water from those peers with the same water access conditions. This provides evidence of the role of plot-dependent heterogeneity in the social diffusion of agricultural practices.

7 Conclusions

Social connectedness is instrumental for economic growth and development in that it eases information, collective action, investment and trade. In particular, while both innovation and good institutions have been emphasized as indispensable for economic efficiency and factor accumulation, social capital is the 'missing link' between institutional quality and economic performance, and even more so in transition and developing settings (Durlauf and Fafchamps, 2005; Guiso et al. 2010).

In this paper, we study the role of peers and social learning in the diffusion of cash crops with high profit margins in a newly created village in northeastern Brazil. The village settlement is the result of the Brazilian agrarian reform, where stranger households find themselves living next to each other during a major economic transformation, namely the shift from wage working to family farming. This setting allows us to circumvent the difficulties related to the causal assessment of the role of interpersonal networks: in fact, while individual- or household-level treatment are common target for field experiments, real-life social interactions can be hardly manipulated *via* randomized interventions. As a consequence, previous literature has mostly exploited data where exogenous variations in behavior was paired with endogenous social networks, overlooking network assortativity (Cai et al. 2015).

We combine geo-localized data on farming plots with dyadic data on social ties among old and new settlers, who moved into the village in a manner typical of the agrarian reform settlements and - without any formal training - from alien landless households turned into commercial farmers. We provide evidence that their location choice during the encampment period was not driven by assortativity along observables. We then leverage geographical proximity between farmers, induced by the land squatting process, to instrument for the formation of social networks.

Using longitudinal data on farming decisions over 15 years to estimate a model of social diffusion along network lines, we find consistent evidence of learning from peers in the decision to adopt new cash fruits (pineapple and passion fruit). Having one additional peer cultivating cash crops increases the probability of crop-specific adoption by 9 to 14 percentage points in the most conservative estimate. Conditional on unobserved household characteristics, our results suggest that social diffusion is heterogeneous along observed plot and crop characteristics, i.e. the decision to grow water-sensitive crops is strongly responsive to peers with similar water–access conditions.

Our findings offer new insights into the role of social learning in the transformation of rural societies. By using a network perspective, we examine a newly created community populated by people with different backgrounds engaged in a participatory development process, and provide evidence that peers influence drives technological change. This has implications for innovation policy in developing settings where the decentralized participation of citizens via networks can be key for economic growth and development.

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth.* Princeton, NJ: Princeton University Press. 1

Ambler, K., A. de Brauw, and S. Godlonton (2018). Agriculture support services in malawi: Direct effects, complementarities, and time dynamics. Technical Report 01725, IFPRI Discussion Paper. 1

Bailey, M., P. Farrell, T. Kuchler, and J. Stroebel (2020). Social connectedness in urban areas. *Journal of Urban Economics 118*, 103–264. 8

Bandiera, O. and I. Rasul (2006). Social Networks and Technology Adoption in Northern Mozambique. *The Economic Journal 116*(514), 869–902. 1

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(6144). 7, 23

Banerjee, A., G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trial. *The Review of Economic Studies 86*(6), 2453–2490. 1

Barr, A. (2003). Trust and expected trustworthiness: Experimental evidence from zibawean village. *The Economic Journal 113*, 614–630. 2

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the u.s. *The Review of Economic Studies 79*(1), 128–161. 25

BenYishay, A., M. Jones, F. Kondylis, and A. M. Mobarak (2019). Gender gaps in technology diffusion. *Journal of Development Economics 143*, 102–380. 1

Bergamasco, S. M. P. P. (1997). A realidade dos assentamentos rurais por detrás dos números. *Estudos Avancados 11*(31), 37–49. 1

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*(1), 41–55. 1, 4, 5.2, 37, 38

Brock, W. and S. N. Durlauf (2001). Discrete choice with social interactions. *Review of Economics Studies 68*, 235–260. 1

Buchel, K. and M. von Ehrlich (2020). Cities and the structure of social interactions: Evidence from mobile phone data. *Journal of Urban Economics 119*. 8

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics 7*(2), 81–108. 1, 7

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *76* (4), 1239–1267. 34, 5.2

Carr, M. K. V. (2013). The water relations and irrigation requirements of passion fruit (*Passiflora edulis sims*): A review. *Experimental Agriculture 49*(4), 585–596. 5

Case, A. and L. Katz (1991). The company you keep: The effects of family and neighborhood on disadvantaged youths. Technical Report 3705, National Bureau of Economic Research. 7

Comola, M., R. Dieye, and B. Fortin (2023). Heterogeneous peer effects and gender-based interventions for teenage obesity. Working Paper. 50

Comola, M. and M. Fafchamps (2014). Testing unilateral and bilateral link formation. *Economic Journal 124*, 954–976. 23

Comola, M. and M. Fafchamps (2017). The missing transfers: Estimating mis-reporting in dyadic data. *Economic Development and Cultural Change 65*(3), 549–582. 23

Comola, M. and S. Prina (2021). Treatment effect accounting for network changes. *Review of Economics and Statistics 103*(3), 597–604. 5.2, 38

Conley, T. G. and C. Udry (2001). Social learning through networks: The adoption of new agricultural technologies in ghana. *American Journal of Agricultural Economics 83*(3), 668–673. 1

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

Dahlberg, M., K. Edmark, and H. Lundqvist (2012, February). Ethnic diversity and preferences for redistribution. *Journal of Political Economy 120*(1), 41–76. 7

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics 27*(2), 281–314. 7

Drukker, D., P. Egger, and I. Prucha (2013). On two-step estimation of spatial autoregressive models with autoregressive disturbances and endogenous regressors. *Econometric Reviews 32*(5-6), 686–733. 5.2

Edin, P.-A., P. Fredriksson, and O. Ãslund (2003). Ethnic enclaves and the economic success of immigrants-evidence from a natural experiment. *Quarterly Journal of Economics 118*(1), 329–357. 7

Fafchamps, M. (2010). Vulnerability, risk management and agricultural development. *African Journal of Agricultural and Resource Economics 5*(1), 1–18. 1

Fafchamps, M. and F. Gubert (2007). The formation of risk-sharing networks. *Journal of Development Economics 83*(2), 326–350. 4

Fafchamps, M. and S. Lund (2003). Risk-sharing networks in rural ¶hilippines. *Journal of Development Economics 71*(2), 261–287. 1

Fafchamps, M. and S. Quinn (2018). Networks and manufacturing firms in africa: Results from a randomized field experiment. *World Bank Economic Review 32*(3), 656–675. 1

Fafchamps, M. and M. Soderbom (2011). Network proximity and business practices in african manufacturing. Technical report, CSAE Working Paper Series 2011–2008, University of Oxford. 1, 23

Feder, G., R. E. J. and D. Zilberman (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change 33*(2), 255–298. 1

Feigenberg, B., E. M. Field, and R. Pande (2010). Building social capital through microfinance. Technical report, CID Working Papers 209, Center for International Development at Harvard University. 1, 7

Flynn, A. J. (2010). *Pluralist Identities and Empowering 'the People': Brazil's Landless Workers' Movement (MST) at the Crossroads*. Ph. D. thesis, University of Manchester. 1, 2

Fortin, B. and M. Yazbeck (2015). Peer effects, fast food consumption and adolescent weight gain. *Journal of Health Economics 42*, 125–138. 35

Foster, A. D. and M. R. Rosenzweig (1993). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy 103*(6), 1176–1209. 1, 1

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of Technology Adoption. *Annual Review of Economics 2*(1), 395–424. 1

Glaeser, E. (1999). Learning in Cities. *Journal of Urban Economics 46*, 254–277. 1

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics 31*(3). 4

Hammond, J. L. (2009). Land occupations, violence, and the politics of agrarian reform in brazil. *Latin American Perspectives 36*(4), 156–177. 1

Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. Technical Report 7867, NBER Working Paper. 7

Hsieh, C.-S. and L.-F. Lee (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics 31*, 301–319. 40

IBGE (2009). Censo agropecuario 2009. Rio de Janeiro: IBGE. 11

IBGE (2018). Censo agropecuario 2018. Rio de Janeiro: IBGE. 2, 2.1

Inguaggiato, C. (2014). *Agrarian Reform, Social Movements and Community Based Organization: The Emergence of New Organizational Forms? A Case Study in Northeast Brazil*. Phd dissertation, University of Trento. 2, 2.1, 16

Karlan, D., I. Osei, Osei-Akoto, and C. Udry (2014). Agricultural Decisions after Relaxing Credit and Risk Constraints. *The Quarterly Journal of Economics 129*(2), 597–652. 1

Kelejian, H. and G. Piras (2014). Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarette. *Regional Science and Urban Economics 46*, 140–149. 40

Laschever, R. (2009). The doughboys network: Social interactions and the employment of world war i veterans. Working Paper. 25

Leite, S., B. Heredia, L. Medeiros, M. Palmeira, and R. Cintraao (Eds.) (2004). *Impactos dos assentamentos: um estudo sobre o meio rural brasileiro*. Rio de Janeiro, Brazil. 2.1

Liu, X., E. Patacchini, and Y. Zenou (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior and Organization 103*, 39–59. 23, 34

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics 22*(1), 3–42. 1

Maertens, A. (2017). Who cares what others think (or do)? social learning and social pressures in cotton farming in india. *American Journal of Agricultural Economics 99*(4), 988–1007. 1

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies 60*(3), 531–542. 1, 5.2

Moffitt, R. A. (2001). Policy interventions, low-level equilibria and social interactions. In S. N. Durlauf and H. Peyton Young (Eds.), *Social Dynamics*, pp. 45–82. Cambridge, MA: MIT Press. 1

Mosse, D. (2001). People's knowledge, participation and patronage: Operations and representations in rural development. In B. Cooke and U. Kothari (Eds.), *Participation: The new tyranny?*, pp. 16–36. London, New York, NY: Zed Books. 1

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. *Journal of Development Economics 73*(1), 185–213. 1, 2

Patacchini, E. and Y. Zenou (2012). Juvenile delinquency and conformism. *Journal of Law, Economics and Organization 28*(1), 1–31. 5.2

Pereira, A. (2003). Brazil's agrarian reform: Democratic innovation or oligarchic exclusion redux? *Latin American Politics and Society 45*(2), 41–65. 1, 2

Platteau, J.-P. (2004). Monitoring elite capture in community-driven development. *Development and Change 35*(2), 223–246. 1

Rigon, A. (2014). Building local governance: Participation and elite capture in slum-upgrading in kenya. *Development and Change 45*(2), 257–283. 1

Rosa, M. C. (2012). Landless: Meanings and transformations of a collective action category in brazil. *Agrarian South: Journal of Political Economy 2*, 205–31. 9

Sigaud, L. (2004). As condicoes de possibilidade das ocupaccioes de terra. *Tempo Social, Revista de Sociologia Da USP 17*(1), 256. 1, 2, 2.1

Stifel, D. and B. Minten (2008). Isolation and agricultural productivity. *Agricultural Economics 39*(1), 1–15. 1

Stock, J. and M. Yogo (2005). Testing for weak instruments in linear iv regression. In *Identification and Inference for Econometric Models*, pp. 80–108. Cambridge University Press. 48

Tinoco, S. and T. Julliatto (2008). Conceituaciones de agricultura familiar: Uma revisiones bibliografica. *Infobibos*. 12

Wanderley, M. d. N. B. (2000). A valorizacao da agricultura familiar e a reivindicacao da ruralidade no brasil? *Desenvolvimento e Meio Ambiente*. 1, 11

Watts, D. and S. Strogatz (1998). Collective dynamics of 'small-world' networks. *Nature 393*, 440–442. 3

Wolford, W. (2010). *This Land Is Ours Now: Social Mobilization and the Meanings of Land in Brazil.* Durham, NC: Duke University Press. 1, 5

A Appendix

Figure A1: Geo–morphological map of the study village in Brazil

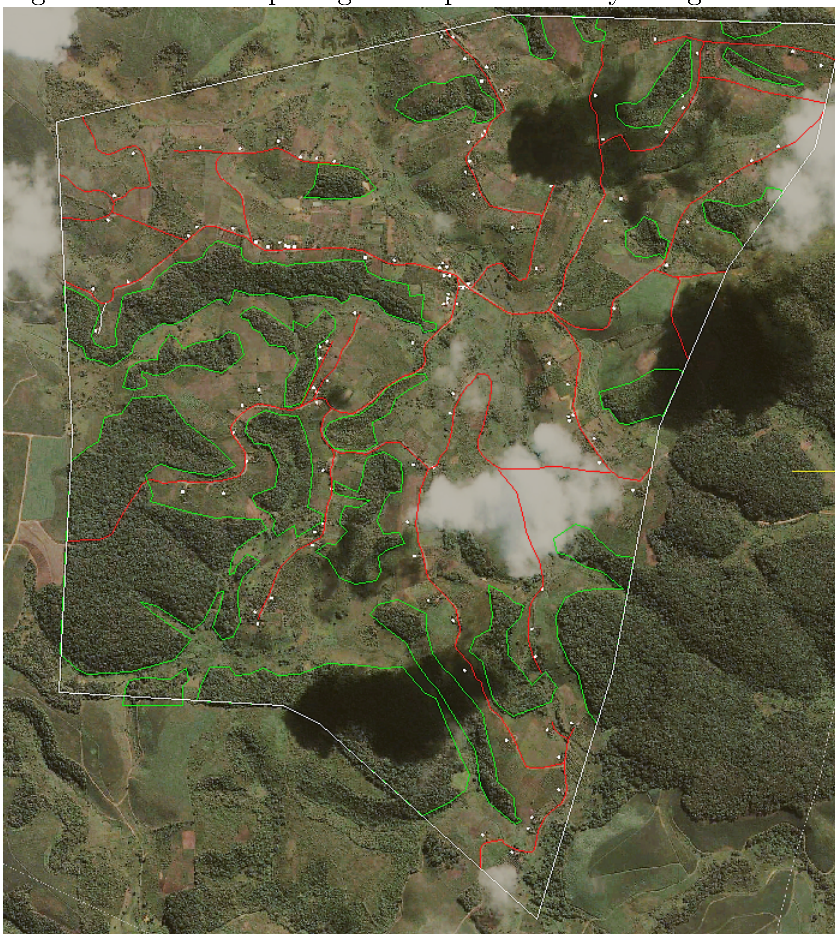


Figure A2: Village blueprint and geo-localized settlers

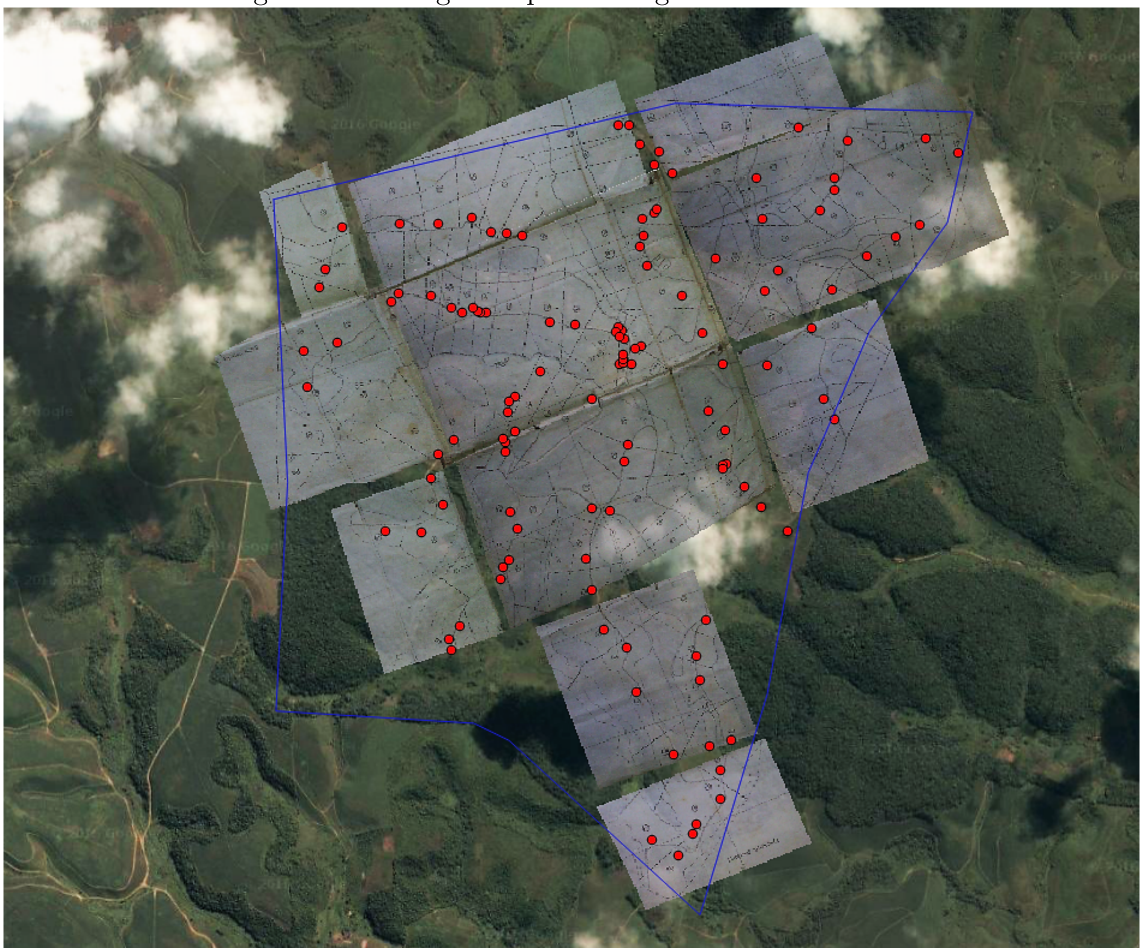


Table A7: Descriptive statistics (main results, N=1500)

		pineapple		passion fruit	
		mean	s.d.	mean	s.d.
	$\mathbf{y}^t$	0.06	0.23	0.15	0.36
	$\mathbf{Gy}^t$	0.31	0.62	0.71	1
distance road $\times$ ***price***$_c$	$\mathbf{x}^t$	0.2	0.23	0.15	0.17
	$\mathbf{Gx}^t$	0.59	0.63	0.45	0.46
	$\mathbf{G}^2\mathbf{x}^t$	2.63	2.34	2.02	1.71
	$\mathbf{G}^3\mathbf{x}^t$	11.41	11.81	8.76	8.7
	$\widehat{\mathbf{G}}^2\mathbf{x}^t$	2.49	1.58	1.91	1.12
	$\widehat{\mathbf{G}}^3\mathbf{x}^t$	9.6	6.07	7.37	4.32
distance water $\times$ ***price***$_c$	$\mathbf{x}^t$	0.35	0.34	0.27	0.25
	$\mathbf{Gx}^t$	1.2	1.24	0.92	0.92
	$\mathbf{G}^2\mathbf{x}^t$	5.68	5.66	4.36	4.16
	$\mathbf{G}^3\mathbf{x}^t$	26.48	29.58	20.33	21.87
	$\widehat{\mathbf{G}}^2\mathbf{x}^t$	4.94	3.17	3.8	2.26
	$\widehat{\mathbf{G}}^3\mathbf{x}^t$	19.09	12.1	14.66	8.61
coop member	$\mathbf{x}^t$	0.23	0.42	0.23	0.42
	$\mathbf{Gx}^t$	1.04	1.25	1.04	1.25
	$\mathbf{G}^2\mathbf{x}^t$	5.09	5.72	5.09	5.72
	$\mathbf{G}^3\mathbf{x}^t$	26.51	28.08	26.51	28.08
	$\widehat{\mathbf{G}}^2\mathbf{x}^t$	3.27	1.3	3.27	1.3
	$\widehat{\mathbf{G}}^3\mathbf{x}^t$	12.63	4.95	12.63	4.95