Peer effects in the adoption of formal property rights: experimental evidence from urban Tanzania

Matthew Collin*†

August, 2014

Preliminary draft (do not cite)

Abstract

This paper investigates the presence of endogenous peer effects in the adoption of formal property rights. Using data from a unique land titling experiment held in an unplanned settlement in Dar es Salaam, I find a strong, positive impact of neighbour adoption on the household's choice to purchase a land title. I also show that this relationship holds in a separate, identical experiment held a year later in a nearby community, as well as in administrative data for over 160,000 land parcels in the same city. While the exact channel is undetermined, evidence points towards complementarities in the reduction in expropriation risk, as peer effects are strongest between households living close to each other and there is some evidence that peer effects are strongest for households most concerned with expropriation. The results show that, for better or for worse, households will reinforce each other's decisions to enter formal tenure systems.

Keywords: Peer effects, Technology adoption, Land tenure, Tanzania, Unplanned settlements

JEL classification: P14, Q15

*Center for Global Development; Email: mcollin@cgdev.org

†I would like to thank Bet Caeyers, Stefan Dercon, Marcel Fafchamps, James Fenske, and Imran Rasul for their support, discussions and suggestions, as well as Stefano Caria, Martina Kirchberger and participants of Oxford's Gorman seminar for helpful comments and suggestions. I am also indebted to Daniel Ayalew Ali, Klaus Deininger, Stefan Dercon, Justin Sandefur and Andrew Zeitlin for their design and implementation of (and my involvement in) the land titling research project from which made this analysis possible. Finally, I am grateful to Andrew Zeitlin, who provided many helpful thoughts and discussions at the early stages of the analysis, as well as quick access to some of the administrative data presented in this paper.

1 Introduction

The formalisation of property rights is considered by many to be crucial to the the institutional development of societies as well as a path out of poverty for informal property owners (De Soto 2000). Land titling is seen as one of the most fundamental steps in this process, yet, despite mixed evidence of its immediate benefits (Field 2005; Galiani and Schargrodsky 2010). While these schemes seem particularly urgent in the face of massive levels of urban growth, particularly in sub-Saharan Africa, very little is known about how to successfully propagate new tenure regimes.

The context for this paper is Dar es Salaam, which throughout its history has been shaped by a constant battle between authorities desperate to maintain control over the city's development and the ongoing pressure of informal growth and migration from rural areas. This struggle has roots as far back as the times of British colonial rule, when the colonial authorities tried, but largely failed to introduce a formal land title system to help contain the expansion of a growing Indian population (Brennan 2007). Despite half a century of of large-scale urban planning and 'strict' government control over the allocation of land, Dar es Salaam remains largely informal today, with over 80% of land belonging to to informal, unrecognized settlements (Kombe 2005). It is hardly surprising then that the Tanzanian government, like many others dealing with rapid urban growth, is keen to find innovative ways to sustainably introduce a formal tenure system.

One facet of tenure adoption which is often overlooked is how individuals' decisions to enter the formal system might co-vary with one another. Here, I investigate whether the adoption of formal property rights is contagious, where the action of one agent adopting a new regime increases the chance that another does the same. In the peer effects literature these are known as *endogenous peer effects* (Manski 1993).

The discovery of endogenous peer effects in property rights adoption is useful for several reasons. First, the existence of adoption spillovers is informative as to whether or not property rights should be considered solely as a private good, or as one with substantial spatial externalities. Secondly, if the channel through which adoption peer effects operate can be identified, we might learn something more about the expected benefits of titling.

Finally, even if the exact mechanisms remain hidden, the presence of positive endogenous peer effects is interesting form a policy perspective, as interventions aimed at encouraging take up would have a subsequent knock-on effect on others, otherwise known as a social multiplier (Glaeser, Scheinkman, and Sacerdote 2003) effect.

Endogenous peer effects are notoriously difficult to identify, as they are subject to both 'reflection' bias (where the direction of causality cannot be determined) and correlated effects (where shared unobservable characteristics drives similar decisions). In this paper, I overcome these standard identification challenges by using exogenous variation in land title purchases resulting from a unique land titling experiment[1] in the unplanned settlements of Dar es Salaam. The experiment randomly allocated a subset of informal landowners to treatment groups which received massive subsidies to obtain a land title, leaving others excluded. I then combine this variation in the incentive to title with spatial information on the location and treatment status of each household's set of nearest-neighbours, allowing me to identify the impact of each neighbour's adoption decision on the probability that a given household will purchase a land title. This approach is similar to a number of studies which use randomised selection into a programme to identify peer effects (Duflo and Saez 2003; Lalive and Cattaneo 2009; Bobonis and Finan 2009; Oster and Thornton 2009).

My results suggest that there are strong, positive endogenous peer effects in land title adoption. In my main specification, the probability that a household chooses to purchase a land title increases by 8-15% for every neighbour that also chooses to purchase one, an effect equivalent in size to a 25-50% discount on the price of the land title. I also show that these results not only diminish with distance, but they appear to be operating primarily through physical proximity, rather than social proximity, and are not necessarily due to the exchange of information. Furthermore, I show that there is some evidence that households with a higher ex-ante perception of expropriation risk are more responsive to the behaviour of their neighbours, suggesting that there are strategic complementarities in adoption to those most fearful of expropriation. For robustness, I show that these results hold for some basic changes to the structure of the peer group. I then go on

[1]The experiment is described in detail in Ali, Collin, Deininger, Dercon, Sandefur, and Zeitlin (2014)

to show that these results remain roughly consistent for an identical experiment rolled out in a neighbouring community a year later. Finally, I turn to a database covering roughly 170,000 land parcels in Dar es Salaam, using popular non-experimental methods of identifying peer effects to show that positive effects also exist in this larger setting, albeit with a slightly different type of land title.

In the next section, I discuss the setting of urban Tanzania in more detail, as well as the types of land titles this paper will be covering. Section 3 covers some reasons why peer effects in land titling take-up are likely to exist. Section 4 outlines the randomised controlled trial which I will exploit to identify peer effects. Section 5 discusses identification and the empirical set up. Section 6 covers the main results of the paper, Section 7 covers the results from the second experiment and administrative data, and I conclude with Section 8.

2 Land tenure in urban Tanzania

In Tanzania, formal access to urban land is controlled exclusively by the government, as all land in the country is owned by the Office of the President (Kironde 1995). Given the rates of growth that Tanzania's cities experienced, the post-independence management and distribution of urban land has generally been haphazard and insufficient (Kombe 2005). Following the 1999 Land Act, the Tanzanian government introduced two new forms of land tenure in urban areas in an attempt to pave the way for more rapid formalisation of existing settlements. The first form of tenure was a temporary, two-year leasehold known as a residential license (RL), which had the benefit of being cheap and easy to implement, but lacked many of the features desired in full titles, such as perpetual security, transferability and collateralisability.

The second form of tenure has been considered to be much closer to a full land title: a certificate of right of occupancy (CRO) lasts 99 years, is transferable and is seen by many as reasonable proof of land ownership by credit providers. Despite the obvious appeal of the CRO, the Tanzanian government has largely failed to encourage urban land owners to purchase them.[2] The lack of progress has been principally due to the large practical and

[2]Records from the Kinondoni Municipality in Dar es Salaam indicate that a little over 2,000 applications

monetary hurdles that urban landowners face, including expensive prerequisites such as cadastral surveying and application fees (Collin, Dercon, Nielson, Sandefur, and Zeitlin 2012).

The benefits of CRO ownership

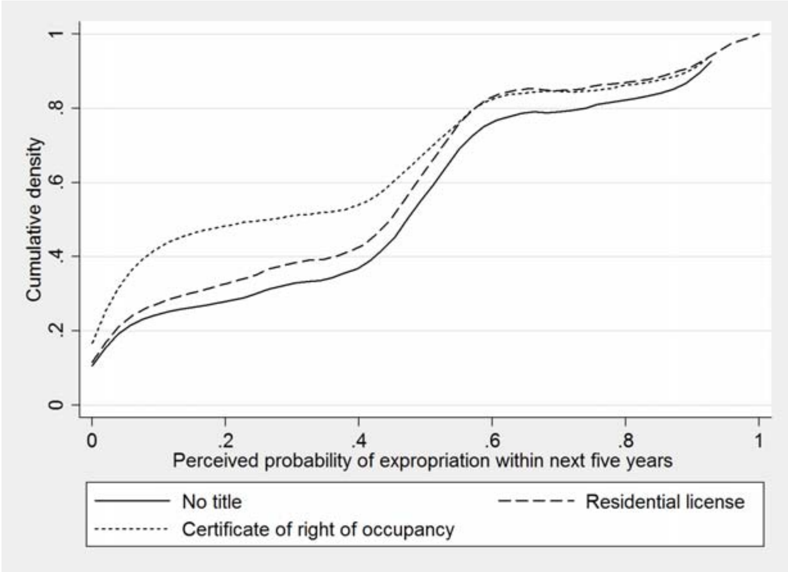
While the Land Act includes relatively straightforward provisions on the legality of using CROs to obtain credit or sell land, the interaction between CRO ownership, expropriation and compensation is less clear. The Land Acquisition Act of 1967 gives the Tanzanian Government broad powers to expropriate land for "any public purpose", even if the owner is in possession of a CRO. This includes government schemes, general public use, sanitary improvements, upgrading or planning, developing airfields or ports and uses related to mining or minerals. Indeed, recent history suggests such expropriation seems most likely to occur from government-driven development initiatives (Hooper and Ortolano 2012). While exact figures on government expropriation are not known, the practice seems frequent enough to elicit alarm in the media: Kironde (2009) found six expropriation-related stories in local newspapers in just one week.

While the Tanzanian government is legally obligated to relocate displaced residents and provide adequate compensation when it acquires land, case studies of recent land conflicts reveal that these efforts are at best mismanaged and at worst completely neglected (Kombe 2010). While a CRO does not legally protect a household from expropriation, it might very well *indirectly* protect a land parcel from government expropriation by raising the value of said land and making the compensation transfer more straightforward. The Land Acquisition Act only provides for compensation in the case where the owner can be identified (Ndezi 2009). Incidents of government expropriation of urban land reported in newspapers and in case studies suggest that informal settlements face the highest risk, so there is reason to believe that, when faced with a choice, governments will usually go for the low-hanging fruit of untitled land.

Even if CRO ownership had no discernable impact on the probability of expropriation, many residents still believe that it does. As part of the baseline data collection for

from CRO have been made, out of a total population of 60,000 land parcels.

Figure 1: Perceived impact of formal land tenure on expropriation risk



Note: Graph shows local-polynomial smoothed cumulative densities of self-reported perceived expropriation probability, conditional on (hypothetical) ownership of different forms of land titles. Data taken from baseline census of landowners in Kigogo Kati and Mburahati Barafu wards in Kinondoni Municipality, Dar es Salaam.

the experiment used in this paper, residents of two unplanned settlements were asked about their perceived probability of full expropriation in the next five years. Respondents were asked to condition their predictions on hypothetically having no title at all, having a residential license, or having a CRO.[3] Figure 1 displays local-polynomial smoothing estimates of the cumulative density function for each response[4] It is clear, at least for a substantial portion of residents, that CROs are *perceived* to be effective at mitigating expropriation risk. As mentioned before, the Land Act also establishes the legal basis for CROs to be used as collateral for loans. Anecdotal discussions with formal lenders in Dar es Salaam suggests that, while CROs are readily accepted as collateral, they do not necessarily offer a substantial benefit over than of a residential license. Still, evidence from the baseline survey used for the experiment described in this paper suggests that households, on average, also expect that CROs will lead to an increase in both credit supply and land values. While it is clear that households recognise a private benefit to

[3]The order of the conditional questions were randomised to avoid priming the respondents.

[4]While responses to the expropriation question were discrete bins, differences in perceived risk are easier to discern using this method. Paired Kolmorogov-Smirnov tests of equality of distributions (not shown) reject the null in every instance.

titling, what this fails to reveal is whether or not landowners perceive any *externalities* in adoption which would give rise to peer effects, a possibility I will explore further in Section 3.

Before describing the experimental setting where people have been induced to adopt this new form of land title, I will first consider the reasons why we might expect peer effects in CRO adoption to exist in this context.

3 Peer effects in land rights adoption

Most work on formal property rights bundles the benefits and expected impacts of titling into three broad categories, initially summarised by Besley (1995) and later expanded upon in Besley and Ghatak (2010). The first of these is through an (expected) reduction in expropriation risk: formalisation should, in theory, reduce the chance a landowner loses his or her land to to either the state or other individuals. In most theoretical contexts, the benefits of reducing expropriation risk are strictly private and positive. Tenure formalisation is also expected to make it easier for landowners access credit by giving them the ability to collateralize their property. Finally, formalisation is expected to increase the transferability of land, allowing landowners to take advantage of rising land prices and for ownership to shift to those who can use it most productively.

With the exception of general equilibrium credit market and land price impacts, which are often ambiguous (Besley and Ghatak 2010), many of these benefits are often modeled as private returns, with the act of one landowner having obtained formal land tenure having no impact on other landowners. There are a number of reasons why there might be immediate, direct spillovers from the decision to buy a land title, both of which have implications for the existence of peer effects. In this section I will consider the most plausible ones, given the context, and then discuss how, in this paper, I will attempt to discern between them.

Complementarity or substitutability in the returns to land title adoption

One particular area which remains understudied is whether or not there are spillovers in the *returns* to adopting formal property rights. Individual formalisation efforts, such as land titling, might not only result in a private benefit, but might also impact the returns to titling for other individuals. We might, for example, expect that the returns to titling would be increasing in the number of neighbours taking the same action. This is the classic case of *strategic complementarity*, when the private returns to an action are greater when other agents also take it (Schelling 1978; Bulow, Geanakoplos, and Klemperer 1985). In the above example, this would be the case when the cross-partial derivative is greater than zero, with household i's returns to titling *increasing* as more neighbours adopt.

Where might we see strategic complementarities in practice? For one, there might be a snowballing effect in the reduction in expropriation risk, with the government taking formal tenure more seriously as more people adopt it, possibly due to the rising implicit costs of paying out compensation or because a legal appeal against expropriation is increasingly more likely to succeed with each additional titled household. However, even in the presence of strategic complementarities, expropriation spillovers might not be entirely positive. If, for example, a government decides to expropriate land which has the lowest level of formal tenure, the act of land titling might just shift expropriation risk from one set of households to another. In this instance, households will be induced to title when their neighbours do the same, not because the decision leads to a net gain in welfare, but because they must do so to prevent a rise in their risk of expropriation. This implies that titling creates a 'race to the bottom,' where all households title in order to improve their security of tenure, but are no better off at the end of the titling scheme. This result is akin to De Meza and Gould's (1992) burglar alarm example: while there is a private benefit for a given household installing a burglar alarm, it increases the probability of neighbouring houses being burgled and hence a no-alarm equilibrium is preferable to an all-alarm one.

Complementarities might also exist in the other standard benefits of land titling. For example, banks may be more likely to accept land titles as form of collateral if they are widely used and accepted in a community (Fort, Ruben, and Escobal 2006) or the impacts of titling on land prices might increase as more neighbours adopt.[5]

[5]Note that both of these channels might also be subject to net negative impacts. If banks switch to a

Of course, titling decisions could also be substitutes: if the marginal utility from titling *decreases* as more neighbours take up, then household i will be more likely to opt out.[6] If the main benefits of titling are through reducing expropriation risk, this would reflect a context where low levels of titling are enough to deter a government from clearing an area, and so subsequent titling is less effective. Similarly, some have argued that the credit-supply effects of large scale titling will be smaller than individual titling: if lenders consider titling to be a signal of borrower quality, rather than as a collateralisable asset, then large-scale titling would imply a lower signal-to-noise ratio (Dower and Potamites 2012).

Strategic complementarity (substitutability) in the returns to titling would imply a positive (negative) endogenous peer effect, as the effect of neighbour take-up increases (decreases) the marginal benefit to titling for a given household. For most of the impacts discussed here, we would also expect these spillovers to be inherently spatial: both expropriation risk and land values are typically highly correlated across space (as might be collateral values, as lenders might be more confident in extending loans to areas they are already familiar with). Later on on this paper, I will take advantage of the spatial nature of the data to discern whether or not the observed endogenous peer effect varies with distance.

Learning and rule-of-thumb behaviour

Peer effects might also arise from learning behaviour: based on their peers' experience, individuals update their beliefs on the efficacy of a product. This 'social learning' behaviour has already been revealed in the decision of farmers to adopt new farming techniques or new types of crops (Bandiera and Rasul 2006; Conley and Udry 2010; Zeitlin 2012). This could equally apply to landowners in urban areas who observe their neighbours obtaining land titles and possibly being secure from expropriation, gaining access

regime where formal titles are the only legitimate form of collateral, non-adopting households might be rationed out of the market (Van Tassel (2004) shows a similar result might happen even if all households are given title) Similarly, if titles become the *de facto* means of transferring property, households relying on informal channels may feel the need to adopt if they are to sell in the future.

[6] This opens the door for standard public goods/collective action problems, as everyone has a private incentive to disinvest if they know their neighbour is investing.

to credit or selling at a high price. Yet, in the context of this study, the benefits of holding a land title would be impossible to measure: as I will discuss in the next section, land titles have yet to be issued for landowners involved in the field experiment. This prevents the sort of wait-and-see learning observed in previous studies.

However, if landowners believe that the adoption decisions of their peers reveal their *knowledge* about the benefits of land titling, high rates of peer adoption may act as a signal for high returns. Recent evidence suggest that peers' adoption decisions transmit important information, irrespective of actual adoption outcomes (Bursztyn et al. 2012). In this circumstance, any observed endogenous peer effects would be unambiguously positive, as take-up conveys a signal of high-returns to titling.

It is normally difficult to disentangle peer effects created by strategic complementarities from signaling/learning behaviour. However, we might expect peer effects determined by the latter to transmit through traditional social networks, as households observe the behaviour of not only their neighbours, but also their friends and acquaintances. Later in this paper, after establishing that that endogenous peer effects in land title take-up exist between spatially-proximate households, I will then take advantage of some basic social network data to investigate whether or not endogenous peer effects also exist across this alternate network structure, which would suggest that effects other than complementarity/substitutability spillovers are at play.

Other channels

Another concern might be strategic expropriation on the part of those obtaining a land title, with early-movers grabbing a portion of their neighbour's land by making an early claim. While this might be a concern in other settings, it is unlikely to be a factor here, as contiguous neighbours must sign off on the CRO application forms affirming the boundaries of the plot. Furthermore, these sorts of actions would still fall under the 'complementarity' channel: if adopting a CRO protects me from my neighbour's attempt to grab land, my neighbour's action increases the marginal gain from adopting that title.

Finally, there might be information-transfer peer effects, where households learn about the benefits of CRO adoption and share this information, then make entirely independent decisions to title. I will discuss this channel and my attempt to rule it out more in the next section.

4 Experimental design and data collection

As I described in the Section 2, most households in Dar es Salaam face formidable barriers to the adoption of formal land titles. In this section, I will describe an experimentally-provided land titling programme designed to overcome these barriers. The experiment, conceived as part of an impact evaluation of CRO adoption, is described in detail in Ali, Collin, Deininger, Dercon, Sandefur, and Zeitlin (2014). The random variation in CRO adoption induced by the experiment will then be used to identify the impact of a neighbour's adoption on a given household's propensity to adopt.

4.1 An experimental land titling programme

The setting is Kinondoni, which is the largest of the three municipalities which make up Dar es Salaam and houses approximately 50% of the city's population. The land titling programme was introduced in two adjacent neighbourhoods (known as sub-wards or *mitaa*), first in Mburahati Barafu then a year later in Kigogo Kati. Barafu will be the main focus of this paper, due to the completeness and robustness of its data, although I will be using the subsequent replication in Kati as a robustness check in Section 7.1.

Both neighbourhoods are located approximately five kilometers from the city centre. While there are a number of pre-planned parcels at the core of each settlement, each mtaa is primarily composed of unplanned, informal settlements. Table 1 displays some basic administrative data from both neighbourhoods alongside Kinondoni as a whole. Typical of most informal settlements, Barafu is a high density area with relatively low reported land values and a lack of access to public services and infrastructure. Informality is the norm here, with very few households holding formal tenure: estimates from a baseline census of Barafu put the total number of CRO owners at around 10 households, less than 1% of the community, and administrative data suggests less than 40% of households have ever purchased a residential license.

Table 1: Summary Statistics on Parcel Characteristics

	Kinondoni Municipality	Kigogo Kati	Mburahati Barafu
Formal employment	49.9%	44.6%	44.3%
Size and Value of Property			
Area in square meters	439	264	247
Property value in '000 TSh.	12,562	9,939	8,910
Land rent in TSh.	3,679	2,125	1,907
Accessibility to the Property			
No access	1.3%	1.1%	1.1%
Foot path	55.2%	71.3%	82.0%
Feeder road	36.4%	19.8%	16.2%
Main road	5.5%	6.6%	0.6%
Highway	1.6%	1.1%	0.0%
Access to Public Utilities			
Piped water (incl. public)	22.7%	22.0%	5.6%
Electricity connection	46.1%	38.6%	35.1%
Waste removal services			
Burn/buried on plot	35.4%	25.4%	55.7%
Gutter/river/street	20.0%	49.6%	35.4%
Collected by priv. company	40.8%	24.4%	8.4%
Collected by municipality	3.8%	0.7%	0.5%
Number of properties	65,535	1,474	990

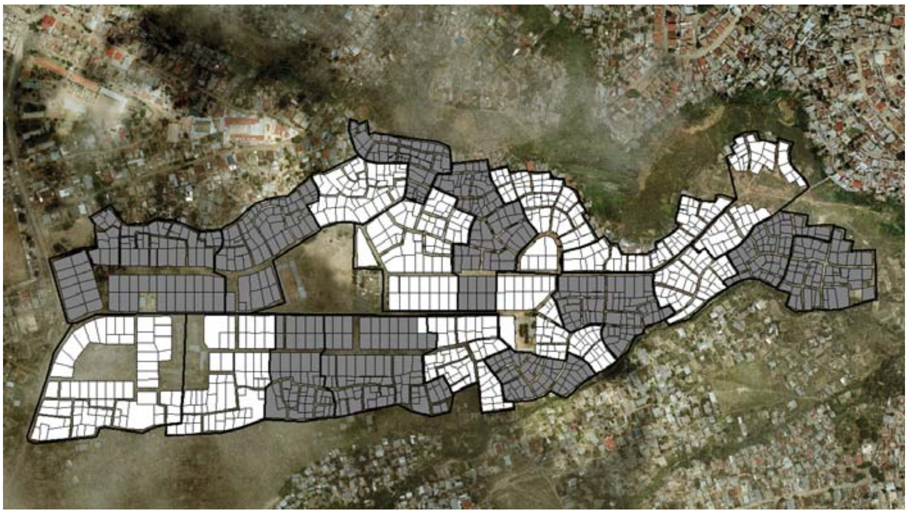
Source: Author's calculations based on the land registry maintained by Kinondoni Municipality.

In October, 2010, the University of Oxford and the World Bank began implementing a land titling programme ultimately aimed at identifying the impact of CRO adoption in Barafu. This was done in partnership with the Woman's Advancement Trust (WAT), a Tanzanian NGO which specialises in large-scale titling programmes. The intervention proceeded as follows: prior to launching the land titling programme, all land parcels in Barafu were identified via a household listing of the community and a recent map drawn up by a town-planning firm. This map was used to divide the community up into twenty 'blocks' of roughly 40-50 parcels each. Using a set of basic characteristics taken from the listing to establish balance, ten blocks were randomly allocated to a treatment group and ten to a control group. Balance results for the overall treatment are available in Table 11 in Appendix A.1. Figure 2 shows the map of Barafu with treatment and control blocks outlined. Parcels in treatment blocks (and their owners) were subject to several interventions:

1. *All* parcels in treatment blocks were subject to a cadastral survey (demarcation of boundaries using cement beacons), one of the prerequisites for applying for a CRO.
2. Parcels in treatment blocks were invited to meetings to discuss involvement in the land titling programme and the benefits of CRO ownership, run by WAT.
3. During these meetings and subsequent follow-up visits, treatment parcels were invited to pay 100,000 TSh to WAT (approximately $62, the average cost of the cadastral surveying plus application fees) over a period of about five months in return for a CRO. In exchange for this, WAT would manage the application process and any related fees.
4. Within treatment blocks, parcels were randomly allocated voucher discounts through a public lottery. Two types of vouchers were allocated: general vouchers, which were redeemable without condition, and conditional vouchers, which required that a female member of the household be included as an owner on the final documents. A parcel could receive both voucher types, just one type, or none at all, and vouchers could take on values of 20, 40, 60 or 80 thousand shillings.[7]

[7] Complete details of the voucher allocation process are discussed in Ali et al. (2014)

Figure 2: Treatment and control blocks in Mburahati Barafu



Note: Treatment blocks are shaded grey.

Following this, households in treatment blocks were free to sign up up and begin repayment. Through an agreement with the municipal government, treated households could not obtain a CRO through conventional means, only through the NGO. Households in control blocks were free to obtain CROs through the municipality, at the regular cost, although a subsequent review of municipal records revealed that none have done so to date. At the time of writing, the project is still underway, with no land titles having yet been issued, but with household decisions and payments having been completed.

4.2 Data sources

In this paper I use three primary sources of data from Barafu. The first was collected prior to the randomised intervention: in the summer of 2010, roughly six months before the start of the land titling programme, the University of Oxford conducted a complete census of all known parcels in Barafu, using records obtained from the Kinondoni Municipality. For each parcel, an owning household was identified and interviewed, resulting in a rich data set of owner and parcel characteristics. However, as this data was collected earlier and used a different sample frame than the administrative project data, there are a number of missing observations, mainly due to parcels which were missed during the baseline census or those that were sold to a new owner in the interim. Baseline data is available for roughly 92% of unplanned parcels in treatment blocks, but only 72% of control blocks have linked baseline data, due to a lack of project information for these households. I will use this data both for testing balance and for controls in my main specification.

The second is detailed parcel-level data taken from project records, including meeting attendance, sign-up and repayment information. As households in control blocks were excluded from participating in the project, this data only contains information on treated parcels. However, data obtained from the Kinondoni municipality reveals that no parcels in the control blacks purchased a CRO during the time frame of the project. Finally, the third source of data I use is detailed geographic information service (GIS) encoded data on the location, shape and size of each parcel in both treatment and control areas.[8] This will allow me to calculate 'nearest-neighbour' peer groups for every parcel and compare the probability of a household choosing to purchase a CRO with the average adoption rate of that household's nearest-neighbour set.

Throughout the remainder of the paper, I will primarily be using *unplanned* parcels in the analysis, as this group was the original target of the research. While there is complete project data and limited baseline data on planned parcels (those who were in a pre-planned area or had already obtained a cadastral survey prior to the intervention), I will only be using them as a robustness check for the main result.

5 Identification of peer effects

Consider a basic linear probability model for household i's decision to adopt a land title:

$$T_i = \alpha + \rho \overline{T}_{g(-i)} + x_i \beta + \overline{x}_{g(-i)} \delta + u_i + \epsilon_g + \varepsilon_i \tag{1}$$

Where T_i is the household's choice to adopt a land title, $\overline{T}_{g(-i)}$ is the average choice of the households group of neighbours g (excluding i), x_i is a vector of household characteristics, $\overline{x}_{g(-i)}$ is the same set averaged over the group, u_i is a household-specific effect and ϵ_g is a vector of group-specific characteristics. Using Manski's (1993) terminology, ρ is known as

[8] This data was taken from a 'town plan' of Barafu, the final planning document drawn up for a community before CROs can be provided

the *endogenous* effect, the impact of i's neighbours' choices on i's choice. The parameter δ represents a vector of effects stemming from i's neighbours' characteristics, known as *exogenous* or *contextual* effects. Finally, ϵ_g contains unobserved within-group *correlated* effects.

There are two primary challenges to the identification of ρ, the parameter of interest. The first is a result of Manski's 'reflection problem', where the direction of causality is difficult to discern. At first glance, we are unable to identify whether ρ captures aggregate effects of i's neighbours' adoption on i or vice versa. In the extreme case where peer groups are perfectly transitive,[9] it is difficult to separately identify endogenous peer effects ρ and the set of contextual effects δ_g.[10] However, when peer or neighbour groups are partially overlapping (i.e. when the neighbours of i's neighbours can reasonably be excluded from i's neighbour set) identification is made possible by exploiting variation in characteristics of these excluded neighbours (Bramoullé et al. 2009; De Giorgi et al. 2010), a popular method I will apply to a larger, non-experimental data set in Section 7.2.

The second concern is over conflating endogenous peer effects with correlated effects. The latter can arise when peer groups or neighbours are affected by common background characteristics or shocks which also predict adoption. For example, if land title adoption depends on unobserved (to the researcher) land quality, then adoption rates will be correlated across neighbours even in the absence of endogenous effects. Similarly, if the endogenous sorting of households into peer groups or neighbour sets is marked by *homophily*, then correlated adoption decisions might solely be the result of correlated individual characteristics, such as wealth or risk aversion.

In this paper, I use the random variation in the price and accessability of land title purchase to identify exogenous changes in $T_{g(-i)}$, allowing me to estimate (1) using two-stage least squares (2SLS) with reduced concerns for correlated effects and reflection. I do this using the percentage of household i's neighbours who were included in treatment blocks as well as their average voucher values[11] as instruments for the average adoption

[9] Transitivity implies that if i and j are peers and j and k are peers, then i and k must also be peers.

[10] Brock and Durlauf (2001) exploit nonlinearities in discrete choice models to identify linear-in-means models, yet identifying assumptions are heavily dependent on functional form, and do not allow for correlated effects.

[11] Averaged over included-neighbours. For precision I use both regular and conditional vouchers sepa-

of the neighbour set. Since households in control blocks were effectively excluded from purchasing CROs, the sample will only cover households in treated blocks (although I will consider neighbours from both treatment and control blocks). I will discuss the suitability of these instruments and possible reasons why identification might still fail in the next subsection.

While many studies have used random variation in *group* assignment to estimate peer effects (Sacerdote 2001; Guryan, Kroft, and Notowidigdo 2009), my approach in this paper is more similar to those which use random variation in *programme* assignment as an instrument for peer-level adoption. For example, both Lalive and Cattaneo (2009) and Bobonis and Finan (2009) use the random assignment of a conditional cash subsidy in PROGRESA villages to instrument for the school enrolment of a child's peer group. Similarly, Oster and Thornton (2009) use the random assignment of menstrual pads to Nepalese school girls to study the impact of group-level treatment on individual utilization of the pads. Both Godlonton and Thornton (2012) and Ngatia (2011) use randomized price incentives to get tested for HIV/AIDS in Malawi to instrument for peer group testing. In each of these studies, social interactions are treated as a specific type of treatment spillover: an individual's peer group is randomly shocked and the resulting change in behaviour affects the individual's adoption choice. This method was first laid out by Robert Moffitt as the *partial-population* approach (Moffitt et al. 2001).

There are a couple of caveats to the interpretation of ρ using the partial-population approach. First, while properly instrumenting $T_{g(-i)}$ solves the reflection problem and bypasses any group or individual-level unobservables, the resulting estimate of ρ is the endogenous peer effect, conditional on groups already having formed endogenously. These 'true' peer effects might be stronger or weaker for households which have chosen to live together as opposed to those randomly sorted into the same neighbourhood. For instance, households from the same religious background might be more likely to associate and share information about adoption decisions. In this instance, we might expect the estimate of ρ, post-endogenous sorting, to be higher than the estimate under random sorting.

Which estimate do we care about? While the 'randomly-assigned' endogenous peer

rately.

effect might be more appealing to those concerned with pure social interactions, in reality the policymaker has little control over the formation of these peer groups, in which case the 'post-sorting' endogenous peer effect is clearly the preferred parameter. In the context of urban formalisation, most policymakers are burdened with the significant task of getting large, informal settlements to take up formal property rights. As these settlements have not formed randomly, the post-sorting peer effect gives us an idea as to whether significant policy multipliers are present for property rights interventions.

Another issue follows directly from using 2SLS with an exogenous treatment instrument to identify peer effects. Under the assumption of heterogenous effects, instrumental variables regressions only allow the researcher to identify the local average treatment effect (LATE) (Imbens and Angrist 1994). The implications of this for the estimation of peer effects are nonnegligible. For example, when using the block-level treatment as an instrument for neighbour adoption, the effect identified ρ is only defined for compliers, households whose *neighbours* were induced to adopt from the treatment, but otherwise would have not done so. As mentioned in the previous section, there are no always-takers, so estimates of ρ using project treatment of an instrument will only be leaving out never-takers, those that do not respond to the treatment. If we have reason to believe that peer effects are heterogenous, then LATE estimates of ρ might deviate substantially from the average treatment effect estimate. The peer effects literature has largely been silent on this issue, with some exceptions.[12]

Finally, it should be emphasised that while the randomised control trial described above has generated geographic variation in the take-up of CROs, the block-level RCT itself was not designed for the purpose of of studying peer effects. Thus most[13] of the identifying variation in take-up will be generated by the large-scale block-level variation in treatment. While this is not as precise as a parcel-level treatment, identification will be possible as long as treated neighbours are not systematically different from untreated neighbours or households *with* treated neighbours are systematically different than those

[12]To date, only Dahl, Løken, and Mogstad (2012) and Ngatia (2011) have explicitly acknowledged that peer effects estimated using 2SLS are subject to a LATE interpretation. Ngatia (2011) explicitly models these heterogenous effects and estimates their effects by exploiting multiple instruments for adoption.

[13]Some of the variation will still be driven by variation in the voucher allocation received by treated neighbours.

without. To allay any concerns, I will show in Section 5.2 that when compared using baseline data, treated and untreated parcels are, on the whole, very similar.

5.1 Challenges to identification

Even though the instruments I use in this paper are randomly drawn, there are still a number of ways the above identification strategy might be undermined. For instance, despite the randomisation, a bad draw in assignment of treatment status or voucher values might have resulted in spurious correlation with relevant unobservable characteristics. Later, I will show that not only both treatment and voucher assignment are well-balanced across a range of observable characteristics obtained from the baseline census, but that the main results presented in Section 6, are unaffected by the inclusion of these characteristics. While balance and conditioning on observables does not guarantee identification (Bruhn and McKenzie 2009), randomisation is as close as we're ever likely to get, as in expectation the instruments should be uncorrelated with the error term in the main equation.

A more pertinent problem is the exclusion restriction. In order for the estimate of ρ to be interpreted solely as an endogenous peer effect, the instruments (being in a treatment block and the random voucher draw) must only affect a household's adoption of a land title through the adoption of its neighbours. There are a few reasons why this might not be the case:

One valid concern is that direct-adoption peer effects might be confused with information exchange. Prior to the intervention, most residents knew very little about CROs. Since households in treated blocks are invited to meetings in which they are given extensive information on the benefits of these titles, it is possible that attending households passed this information on to their non-attending neighbours. Thus the observed peer effect ρ might include the impact of this information transfer. To account for information in my main specification, I will use data on household and neighbour meeting attendance to proxy for knowledge of CROs.

Another potential problem is related to a second change in neighbour characteristics driven by the treatment. Recall that all parcels in treatment blocks are subject to a cadastral survey, even if the owners do not go on to purchase a land title. The act of surveying a neighbour's plot could have an independent effect on a household's decision to purchase a CRO if, for instance, being in a heavily-surveyed area affects the perceived value of a title. Recent evidence suggests that land demarcation has important implications for the function and growth of land markets (Libecap and Lueck 2011), so it is possible that a shift from the previous regime[14] to tightly-regulated cadastral surveying could have substantial impacts independent of land title adoption.

To deal with this, I first turn to data from the baseline census, which suggests that a household's perceived expropriation risk is unaffected by proximity to previously-surveyed parcels (these results are discussed in detail in Appendix A.4). Secondly, I also find that endogenous peer effects are of a similar magnitude when I include previous-surveyed parcels as neighbours.[15] Finally, the timing of the intervention suggests that adoption decisions might have been independent of surveying: while treatment and control blocks were decided at the beginning, actual cadastral surveying did not begin until several months following the initial sign-up period, and took over a year to complete, so the final surveying status of treated-neighbours would have been unconfirmed for most households.

Another assumption behind the exclusion restriction is that proximate neighbours have independent budget constraints. This would be undermined if two neighbours act as a single household or take part in risk-sharing groups.[16] However, while spontaneous risk-sharing groups have been observed in randomised controlled trials in the past,[17] the chances of such an arrangement existing in this context are slim, given that the households were presented with non-transferable vouchers which were tied to individual parcels.

Finally, the exclusion restriction might be undermined if households decide not to participate in the programme because of concerns for fairness (for their neighbours not being included) or if high/low voucher allocations to neighbours elicit feelings of envy or unfairness which stop them from adopting. However, anecdotally there is not much evidence that these sort of feelings are at play on the ground.

[14] Prior to the introduction of the town plan, parcels were delineated with hand-drawn maps produced using aerial photography.

[15] Results presented in the appendix

[16] Lalive and Cattaneo (2009) discuss this as a potential threat to identification, where sharing of PROGRESA transfers might lead to a spurious social interaction result.

[17] Blattman (2011) discusses difficulties with lottery recipients exchanging winnings. Similarly Angelucci and De Giorgi (2009) shows that ineligible households are affected by cash transfers to treatment households.

5.2 Empirical setup

Reconsider the empirical model presented in equation (1), which is presented as a linear probability model (LPM):

$$T_i = \alpha + \rho \overline{T}_{g(-i)} + x_i\beta + \overline{x}_{g(-i)}\delta + u_i + \epsilon_g + \varepsilon_i$$

While it is possible to estimate this using a nonlinear specification, such as a probit or logit, the LPM makes interpretation of the results relatively straightforward. The chief concern over the use of a LPM is over out-of-sample predictions and the potential bias which results from its use. In Table 16 in Appendix A.1, I show that the percentage of out-of-sample predictions is extremely low, which suggests that there is not much scope for bias in the LPM (Horrace and Oaxaca 2006).

A dummy variable equal to one if a household has fully paid for its CRO will be used as my main measure of title adoption T_i.[18] In my main specification, for household/parcel characteristics x_i, I will include the general and conditional voucher values that the household received and a control for whether or not that household attended the block-level meeting. In addition, I will include a series of baseline controls, including the natural log of the parcel's area, the year the parcel was obtained, the household's monthly income, total value of all assets (TSh), household size, average schooling and dummy variables for whether the parcel is rented out, the owner is resident on the parcel, and there has been recent investment in the parcel. Each of these controls is also averaged across the household's neighbour set and included in $\overline{x}_{g(-i)}$, with the exception of the neighbour's voucher values, which are used as instruments. I have also included a control for whether or not the household has neighbours outside of the treatment block, so as not to conflate differences in neighbour treatment with the household's relative location within the block.

Using GIS data to calculate distances between parcel borders, I construct peer groups using the n closest neighbours to household i. This approach allows for results which are intuitive and easy to understand, as each house has equal-sized peer groups. For robustness, I will also present results using fixed-distance neighbour sets (which include

[18]Results are also robust to using household sign-up as a measure of adoption instead of full purchase.

all neighbours within a certain distance d), but the differences are minor. As it offers a reasonable trade-off between proximity and the power of the instruments,[19] my main results will use the five nearest-neighbours, but extensions on the size of the neighbour sets are presented in Section 6.2.

Summary statistics for the main controls, as well as their balance across voucher allocations and the percentage of five nearest-neighbours treated, are shown in Table 2. Parcels which faced a high price were slightly less likely to be electrified and had slightly higher levels of schooling, but neither of these differences are substantial. Households with a high percentage of treated neighbours were more likely to attend meetings and were less likely to have purchased a residential license. Apart from these differences, the sample appears to be fairly well-balanced.

Finally, I will be using both average voucher values across the neighbour set and the percentage of treated neighbours as instruments for $\overline{T}_{g(-i)}$. While the results are robust to including these as separate instruments, the estimates are most precise when they are aggregated into a single instrument. This instrument is defined as the 'total' price of a CRO per household, which is set to TSh 500,000 for untreated neighbours (which is in line with previous estimates)[20] and set to the actual project price, net of vouchers, for treated neighbours.

To account for spatial dependence of observations, all standard errors are calculated using Conley's (1999) method, where the estimated covariance matrix is adjusted to allow for arbitrary spatial correlation between observations. The degree of correlation is allowed to decrease linearly with distance and is set at zero beyond a specified cutoff. For all nearest-neighbour specifications, cutoff values are set at the average distance of the fifth neighbour across observations. For distance-band specifications, cutoff values are set equal to the distance-band. In general, the results are not qualitatively different from standard heteroskedastic-robust estimates.

[19] The larger the neighbour set the greater the number of households which fall outside i's block and therefore have the potential to be treated.

[20] Average estimates put this at about $500-1000 per parcel.

Table 2: Summary statistics and balance (voucher distribution and treated neighbours)

	Mean/SD	Own price	% neighbours treated	Mean neighbour price
	(1)	(2)	(3)	(4)
Attended meeting	0.61	0.002	0.515	-.0008
	(0.488)	(0.0009)**	(0.165)**	(0.0004)**
Year parcel acquired	1992.126	0.014	6.454	-.016
	(12.307)	(0.023)	(4.228)	(0.009)*
Parcel rented out	0.4	0.001	-.040	0.00007
	(0.512)	(0.001)	(0.176)	(0.0004)
Owner resides on parcel	0.827	-.00007	0.05	-.00008
	(0.395)	(0.0007)	(0.136)	(0.0003)
Applied for CRO in past	0.014	-.00007	-.061	0.0001
	(0.124)	(0.0002)	(0.042)	(0.00009)
Applied for RL in past	0.386	0.0001	-.288	0.0006
	(0.508)	(0.001)	(0.175)*	(0.0004)
Parcel was inherited	0.107	0.0008	0.113	-.0002
	(0.322)	(0.0006)	(0.111)	(0.0002)
Parcel has electricity	0.408	-.002	0.033	-.0003
	(0.513)	(0.001)**	(0.177)	(0.0004)
# buildings on parcel	1.332	0.0007	-.065	0.0002
	(0.56)	(0.001)	(0.193)	(0.0004)
Invested in parcel	0.175	-.0006	0.06	-.00003
	(0.397)	(0.0007)	(0.137)	(0.0003)
Monthly income	356.346	0.477	-175.492	0.328
	(464.245)	(0.876)	(159.710)	(0.349)
Total assets (tsh 000')	4140.882	15.221	-1836.007	4.724
	(6848.912)	(12.908)	(2357.852)	(5.151)
Average schooling	12.263	0.009	1.479	-.002
	(2.783)	(0.005)*	(0.956)	(0.002)
Household size	4.716	-.007	-.669	0.0003
	(2.508)	(0.005)	(0.863)	(0.002)
Ln(area m^2)	5.096	0.001	0.146	-.0002
	(0.529)	(0.001)	(0.182)	(0.0004)
Obs	459	459	459	459

Column (1) displays the mean and standard deviation for each variable. Columns (2)-(4) display the mean and standard error of β_2 from the linear regression of each variable $var = \beta_1 + \beta_2 * Z$, where Z is overall price the household faced (2), the percentage of five-nearest neighbours who were in treatment blocks (3) and the average price faced by the household's five-nearest neighbours (setting p = 500,000 TSh for neighbours in control blocks)(4). Price measured in ('000 TSh).
$^{*}(p < 0.10),^{**}(p < 0.05),^{***}(p < 0.01)$

6 Main results

Table 3 shows the results from the estimation of equation (1) using the five nearest-neighbours as the relevant peer group. The first three columns display results from an OLS estimation of the probability that household i adopts a land title on the number of neighbours in the neighbour set also adopting.[21] In column (1), the controls included are household i's allocated vouchers, whether or not someone from the household attended the information/voucher distribution meeting held for the treatment block, the percentage of i's neighbours who attended a meeting and the percentage of neighbours who are in a different treatment/control block. Column (2) restricts the sample to households with baseline data and the nearest-neighbour set to neighbours with baseline data, but does not include baseline controls. These are introduced in column (3), so as not to conflate sample-selection differences with the changes induced by including controls. Also included are average values for these controls for household i's neighbour set. Columns (4), (5), and (6) repeat the same pattern, but using 2SLS to estimate equation (1), using the average 'total' price households in the neighbour set faced as an instrument.

OLS estimates of the endogenous peer effect ρ are positive and of similar size, even when including baseline controls, with the predicted probability that household i purchases a land title increasing by 7-8 percentage points with each neighbour that takes up. When instrumented, these estimates nearly double, with the probability that the household purchases a CRO increasing by 14-15 percentage points with each neighbour that takes up. In previous literature, IV estimates of peer effects are nearly always higher than the OLS estimates. In a pure Manski world, this is perplexing, as simultaneity bias and correlated effects should, on average, lead to bias away from zero, rather than towards it.

One possibility relies on the local average treatment effect interpretation of the estimated coefficient: as ρ is estimated using 2SLS, it is defined only over households whose neighbours were affected by the treatment, thus leaving out all households with neighbours who decided, despite facing large subsidies, not to purchase a land title. This decision might convey unobserved information which also interacts with the mechanisms

[21] This estimation is equivalent to using the average adoption rate for the neighbour set, multiplied by the size of the neighbour set, which is a constant. For results using distance bands instead of nearest-neighbour sets, I multiply by the average neighbour set size.

Table 3: Barafu - impact of neighbour's CRO take up on own take up - 5 closest neighbours

	OLS			2SLS		
	(1) Basic	(2) Restricted	(3) Restricted + Controls	(4) Basic	(5) Restricted	(6) Restricted + Controls
# of neighbours adopting	0.0773***	0.0836***	0.0835***	0.147***	0.137***	0.148***
	(0.0170)	(0.0173)	(0.0182)	(0.0409)	(0.0425)	(0.0396)
Voucher (tsh '000)	0.00386***	0.00318***	0.00389***	0.00290**	0.00244*	0.00301**
	(0.00113)	(0.00120)	(0.00120)	(0.00128)	(0.00136)	(0.00137)
Gender voucher ('000)	0.00388***	0.00394***	0.00424***	0.00294***	0.00316***	0.00342***
	(0.000946)	(0.000974)	(0.000994)	(0.00111)	(0.00116)	(0.00116)
Attended meeting	0.192***	0.126**	0.124**	0.203***	0.133**	0.129**
	(0.0514)	(0.0558)	(0.0550)	(0.0531)	(0.0567)	(0.0561)
% neighbours attended	-0.164**	-0.118	-0.119	-0.180**	-0.132	-0.128
	(0.0832)	(0.0854)	(0.0878)	(0.0850)	(0.0866)	(0.0895)
% neighbours out of block	0.0202	0.00630	0.00511	0.0499	0.0290	0.0359
	(0.0493)	(0.0503)	(0.0222)	(0.0518)	(0.0531)	(0.0516)
Constant	0.148*	0.176**	5.865	-0.00645	0.0559	0.00114
	(0.0798)	(0.0820)	(4.209)	(0.111)	(0.115)	(0.112)
Baseline controls	No	No	Yes	No	No	Yes
Adj. R-Square	0.110	0.106	0.121	0.0784	0.0865	0.0946
Obs	456	421	421	456	421	421
C-D Wald F-stat				84.52	67.75	75.94

Dependent variable is a dummy variable = 1 if the household purchases a CRO

Basic columns include only controls shown + # of neighbours attending meeting and a control for whether household has neighbours outside treatment block

Restricted columns are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data

Restricted + Controls columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence RL ownership, electricity access, number of buildings, recent parcel investment, monthly income, assets, average schooling and hh size

Instruments in 2SLS specification: average priced faced by neighbours (setting untreated parcels at price = tsh 500,000

Conley-adjusted standard errors in parentheses, $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

driving peer effects: for example, the choice of a neighbour not to purchase a title might reveal that expropriation complementarities are not expected to be particularly strong in a given location. Also, if some neighbors never intend to adopt CROs (even if they were to face a price of zero) their non-adoption might convey little-to-no information to other households, resulting in lower average peer effects when they are included.

The other possible reason why 2SLS results are higher than OLS is due to a mechanical downward bias in OLS estimates inherent in most endogenous peer effects models. Guryan, Kroft, and Notowidigdo (2009) show that when peer groups are constructed which *exclude* the household itself and peers are considered as observations as well, OLS estimates will be biased downward.[22] Guryan et al. (2009) also show that controlling for the average take-up of the pool from which a household's peers are selected corrects for this bias. However, in the current context, this 'pool' comprises all observations from Barafu except for the household of interest: as all variation in the pool average is being driven by variation in T_i, it is impossible to include it as a control. Caeyers (2013) shows that this bias is removed when using 2SLS, as valid instruments for $\overline{T}_{g(-i)}$ also side-step the mechanical bias, hence resulting in higher estimates under 2SLS than OLS.

Both types of vouchers have strong, significant effects on take up. Meeting attendance is correlated with higher take-up, although it is unclear if this due to the effect of the meeting or driven by unobserved demand for CROs. Interestingly, neighbour attendance of meetings is negatively correlated with CRO adoption, indicating that the direction of information channels is not straightforward. As meeting attendance is endogenous, Table 14 in the appendix shows the main results still hold when meeting attendance is excluded from the specification. The dummy indicating that the household has neighbours living outside the treatment block does not appear to be a significant correlate of adoption.

The voucher results give us a novel way to interpret the size of the peer effect results. In the 2SLS specification with baseline controls a 1,000 TSh voucher is associated with approximately a .03% increase in the predicted probability that a household purchases a CRO, the decision of a nearest-neighbour to purchase a CRO leads to approximately a

[22] The intuition is as follows: as households are being excluded from their own peer group, if the household had a high value of the outcome of interest Y_i then the resulting peer group will have, in expectation, a lower average outcome $\overline{Y}_i$. When, in turn a household from the same group with a low value of Y_i is considered, the constructed peer group will have a higher average value.

15% increase. Thus, the peer effect generated by a single neighbour adopting is roughly equivalent to a 50,000 TSh voucher transfer.

That peer effects are large and strictly positive suggests positive strategic complementarities in the purchase of CROs. I will investigate this further using a variety of robustness checks throughout this section. More substantial robustness checks are performed in Appendix A.2, where I show these results a robust to the inclusion of block fixed effects and controls for the take-up decisions of household's outside of the nearest-neighbour set.

6.1 Distance and social connections

To confirm that these results aren't isolated to a single specification, Table 4 shows estimates of ρ across different nearest-neighbour sets. In both the OLS and 2SLS specifications, peer effects are strong, positive and significant. Table 12, located in the appendix, shows these results to be similar when using distance-bands.

From these results, it is clear that peer effects are decreasing with distance. The average effect per-neighbour in the three-neighbour 2SLS specification is roughly seven times greater than the twenty-neighbour neighbour one (although this gradient is less steep for the OLS and distance-band specifications). Figure 3 shows the decrease in the effect for both nearest-neighbour and distance-band approaches as the number of neighbours included is increased. While this shows that peer effects in adoption are determined by distance, it doesn't suggest a direct mechanism. Although proximate geographic complementarities might be at play, physical distance might just be a convenient proxy for social distance, as those who live close to one another are more likely to interact on a day-to-day basis.

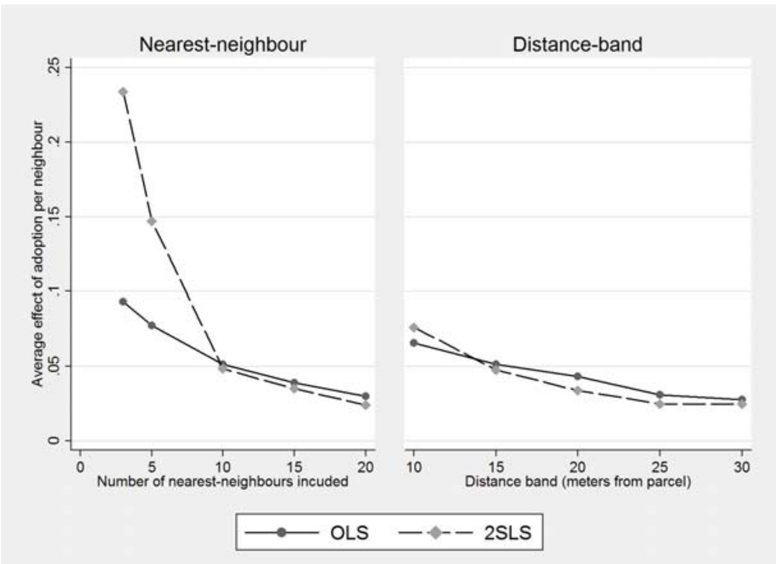
Data taken during the baseline survey might prove helpful in solving this conundrum. Prior to the baseline data collection, for each of fifteen administrative blocks of households (note that these blocks do not correspond to the blocks used for the experiment) a random sample of ten households were chosen to form a network questionnaire. During the baseline survey, each household was asked if they knew the head of each household from the network roster. For all households with baseline data, I have matched up those listed on the network roster with programme take-up data. Matching these responses in the

Table 4: Barafu - impact of neighbour's adoption for nth nearest-neighbour sets

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.0933**	0.0773**	0.0513**	0.039**	0.0299**
	(0.0232)	(0.017)	(0.0104)	(0.0086)	(0.0074)
Restricted	0.0944**	0.0836**	0.048**	0.0365**	0.0302**
	(0.0236)	(0.0173)	(0.0112)	(0.0088)	(0.0072)
Covariates	0.0914**	0.0835**	0.0448**	0.0301**	0.028**
	(0.0242)	(0.0182)	(0.0118)	(0.0101)	(0.0078)
2SLS					
Basic	0.2339**	0.147**	0.0483**	0.0349**	0.0239**
	(0.0593)	(0.0409)	(0.0194)	(0.0123)	(0.0094)
Restricted	0.1896**	0.1368**	0.0474**	0.0302**	0.0217**
	(0.0607)	(0.0425)	(0.0203)	(0.0129)	(0.01)
Covariates	0.2031**	0.1478**	0.0611**	0.0404**	0.0292**
	(0.0629)	(0.0396)	(0.0199)	(0.0127)	(0.0088)
# nearest neighbours =	3	5	10	15	20

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3 = 3 closest neighbours). Conley standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Figure 3: Average neighbour peer effect as neighbour set increases in distance



network questionnaire has allowed me to construct a limited dyadic sample of 402 parcels, each with 9.24 links on average, for a total of 3,718 observations. The i dimension of the dyad includes all treated households with responses to the network questionnaire. The j dimension includes all of those listed on the roster with take up data. This will allow me to investigate whether adoption peer effects are higher for households closer together, or those that know each other.

Table 5 shows the results from a regression of i's probability of take up on j's take up, including an interaction term if household i knows household j and a second interaction for the geographic distance between i and j in meters. Standard errors are clustered at both the i and j level using Cameron et al.'s (2011) method, which provides a good approximation of the dyad-specific approach proposed by Fafchamps and Gubert (2007). The first column of Table 5 shows the results using OLS, which show that j's purchase of a CRO is associated with a 10% increase in the probability that i purchases a CRO. This effect increases by roughly one percentage point if i knows j, but the effect is insignificant at the 10% level. However, the peer effect decreases with distance: the effect is 1% lower for every 15 meters of distance between the two households. Column (2) shows a 2SLS specification, again using aggregate price of a CRO as an instrument for j's take-up.[23] The coefficients in the 2SLS specification are very similar to those of OLS, with the negative coefficient on the distance interaction being nearly identical and still significant at the 10% level.

While the results here are based on a limited sample (those who answered the network questionnaire and those who were randomly selected to be on the network questionnaire), they do suggest that peer effects are primarily running through physical proximity, rather than ex-ante familiarity between households. Again, this points towards complementarities in the marginal gain from CRO adoption, rather than signaling or information flows.

[23] To instrument the interaction terms, I use interactions between the main instrument (average neighbour price) and the two dummies of interest, i knowing j and the distance between i and j.

Table 5: Impact of neighbour's CRO take up on own take up - matched network list

	(1) OLS	(2) 2SLS
Household j is adopting	0.103**	0.137**
	(0.0425)	(0.0615)
(j adopting) * (i knows j)	-0.00437	0.00769
	(0.0716)	(0.104)
(j adopting) * (i-j distance)	-0.000656**	-0.000709*
	(0.000274)	(0.000379)
Household i knows household j	0.0509	0.0399
	(0.0500)	(0.0647)
Distance between parcels i and j	0.000397***	0.000438***
	(0.000135)	(0.000156)
Unconditional voucher	0.00264**	0.00265**
	(0.00122)	(0.00122)
Conditional voucher	0.00469***	0.00466***
	(0.000974)	(0.000977)
Constant	0.317***	0.302***
	(0.0558)	(0.0574)
Adj. R-Square	0.0515	0.0506
Obs	3718	3718
C-D Wald F-stat		15.32

Dependent variable is a dummy variable = 1 if household i purchases a CRO
Instruments in 2SLS specification: j household in treatment block, (i knows j)*(j treated) and (i-j distance)*(j in treatment block). Robust standard errors in parentheses, two-level clustering at both i and j parcel level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

6.2 Distance or contiguity?

While the results in the previous subsection suggest that peer effects in land titling take-up are inherently spatial in nature, it is not yet possible to discern whether or not these effects are due to general spatial spillovers (where we would expect effects to diminish gradually with distance) or immediate neighbour effects (where we would expect a sharp 'drop' in the peer effect for non-contiguous parcels). While we would expect the former for spillovers in aggregate expropriation risk or, say land prices, the latter might be driven by concerns over losing land to titled neighbours. To investigate whether or not peer effects vary discontinuously, I have created a dyad of $i - j$ pairings of every parcel within 150 meters of each other in the neighbourhood, then repeated the specification used in Table

5, this time interacting the dummy for j's decision to purchase a CRO with a dummy equal to one if the two parcels are contiguous, allowing us to investigate whether or not contiguity implies a stronger peer effect.

The results are presented in Table 6. While in the basic OLS specification, the interaction term between contiguity and j's take-up decision is positive and significant, suggesting stronger peer effects when two parcels share a border, this effect disappears in the 2SLS specification. In both specifications, the interaction between j's adoption decision and the distance between the two parcels is negative and significant. Taken together, while there is still some scope for a 'contiguity effect', the peer effects observed here seem to be best captured by linear distance, implying general spillovers rather than titling driven by concerns over neighbour encroachment.

6.3 Peer effects and baseline perceptions of expropriation risk

In order to investigate further the role of expropriation risk in this context, I turn to baseline data on the parcel owner's perceived risk of expropriation (presented earlier in Figure 1). While respondents could choose probabilities anywhere from zero to one, predictions were clumped around zero, 0.5 and 1. To see if those who believe themselves to be at a higher risk of expropriation are more responsive to their neighbours' adoption, I create a dummy variable ($exprop_i$) which is equal to one if the house reported their perceived expropriation risk (conditioned on not having any form of title) to be greater than or equal to 50%. I then interact this dummy with $\overline{T}_{g(-i)}$, the adoption rate of the neighbours, and proceed with the same specification as before. For the 2SLS estimates, I retrieve predicted values of $\widehat{\overline{T}_{g(-i)}}$ from the first stage regression, then use these predicted values and their interactions $\widehat{\overline{T}_{g(-i)}} \times exprop_i$ as instruments for $\overline{T}_{g(-i)}$ and $\overline{T}_{g(-i)} \times exprop_i$.[24]

Table 7 shows the results from both the OLS and 2SLS estimation for three different neighbour sets, all with baseline controls included. In all OLS specifications, the interaction effect is significant and positive, where the 2SLS results show a significant effect

[24]Wooldridge (2010) suggests that there are efficiency gains when using predicted values as interaction terms, rather than the original instruments.

Table 6: Impact of neighbour's CRO take up on own take up - distance dyad

	(1) OLS	(2) 2SLS
Household j is adopting	0.0497**	0.211***
	(0.0197)	(0.0600)
(j adopting) * (i-j distance)	-0.000426*	-0.00132**
	(0.000234)	(0.000557)
(j adopting) * (i-j contiguous)	0.0849**	0.0245
	(0.0335)	(0.146)
Unconditional voucher value	0.00356***	0.00358***
	(0.00112)	(0.00111)
Conditional voucher value	0.00466***	0.00460***
	(0.000969)	(0.000969)
Distance between parcels i and j	-0.0000416	0.000226
	(0.000166)	(0.000211)
i and j are contiguous neighbours	-0.0565**	-0.0350
	(0.0235)	(0.0886)
Constant	0.636	0.527
	(0.619)	(0.727)
Baseline controls	Yes	Yes
Adj. R-Square	0.113	0.106
Obs	56386	55856
KP F-stat		6.770

Dependent variable is a dummy variable = 1 if household i purchases a CRO
Instruments in 2SLS specification: j household in treatment block, (i-j contiguous) *(j in treatment block) and (i-j distance)*(j in treatment block). Robust standard errors in parentheses, two-level clustering at both i and j parcel level.
$^{*}p < 0.10,^{**}p < 0.05,^{***}p < 0.01$

at the 10% level in the two largest-neighbour sets (columns (4) and (6)). These results suggest that the peer effect is stronger for those that had a higher ex-ante perceived probability of expropriation. The coefficient of the level effect of $exprop_i$ is consistently large, negative and significant in most specifications. It appears that while households with a higher ex-ante expropriation risk are more responsive to peer effects, they have a lower absolute level of take-up. This is consistent with a model in which households with a high perceived risk only bother to purchase if they observe others around them doing them same, suggesting that there are complementarities in the reduction of expropriation risk.

Table 7: Barafu - interaction between perceived expropriation risk and impact of neighbour's CRO take up

	5 nearest		10 nearest		15 nearest	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
# of neighbours adopting	0.0284	0.129**	0.000145	0.0310	-0.000481	0.0204
	(0.0350)	(0.0571)	(0.0197)	(0.0277)	(0.0161)	(0.0190)
High exprop risk × # adopting	0.0771*	0.0318	0.0682***	0.0527*	0.0464***	0.0337*
	(0.0393)	(0.0581)	(0.0228)	(0.0305)	(0.0176)	(0.0195)
High exprop risk	-0.260**	-0.111	-0.444***	-0.351*	-0.443***	-0.331*
	(0.131)	(0.186)	(0.141)	(0.184)	(0.158)	(0.174)
Voucher (tsh '000)	0.00404***	0.00301**	0.00451***	0.00423***	0.00464***	0.00442***
	(0.00125)	(0.00136)	(0.00124)	(0.00122)	(0.00122)	(0.00119)
Pink voucher (tsh '000)	0.00417***	0.00333***	0.00460***	0.00427***	0.00495***	0.00469***
	(0.00105)	(0.00115)	(0.00102)	(0.00101)	(0.00100)	(0.000971)
Attended meeting	0.122**	0.129**	0.133**	0.144***	0.122**	0.132**
	(0.0571)	(0.0560)	(0.0541)	(0.0531)	(0.0535)	(0.0522)
Constant	7.156	4.853	9.963	9.981	6.037	7.317
	(10.90)	(10.43)	(16.30)	(15.50)	(19.73)	(18.80)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-Square	0.127	0.0927	0.151	0.143	0.145	0.140
Obs	421	421	421	421	421	421
C-D Wald F-stat		32.48		62.17		144.3

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **High expropriation risk** a dummy = 1 if household's perceived probability of expropriation >= 50%. Specification includes main and baseline controls discussed in previous tables. **Instruments** in 2SLS specification: predicted values from first stage regression of 2SLS regression (using average priced faced by neighbours (setting untreated parcels at price = 500,000 TSh) as an instrument, interacted with high expropriation risk. Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

7 Additional results

In this section I will present results from two other sources of data, first from a replication of the above experiment in Kigogo Kati, an adjacent community, and then from administrative data from the Kinondoni Municipality.

7.1 A second experiment

Nearly a year following the allocation of treatment and control blocks and the voucher distribution in Mburahati Barafu, the same intervention was introduced in Kigogo Kati, which borders Barafu to the south. Due to the length of time between the two interventions, this provides us with an interesting replication of the Barafu experiment. Other than location, Kati differs from Barafu in two key ways which might affect estimates of peer effects. Firstly, shortly after the intervention began, Dar es Salaam was subject to some of the worst flooding in 60 years, with Kati being one of the areas which was affected the most. This subsequently depressed CRO adoption, as households were subject to a shortfall in income. While this should not necessarily dampen peer effects, the low levels of take-up (roughly 15% versus approximately 60% in Barafu), indicate that the instruments used to identify peer effects will be significantly weaker.

Secondly, Kati has been the recipient of a community infrastructure upgrading project (CIUP) for several years, which has led to a number of parcels being demolished to make way for road expansion and electrification. This increased probability of expropriation and the changes in the gains for land titling which might come from being in a heavily invested area are both likely to interact with peer effects. Furthermore, while the take up data from Barafu is considered complete, Kati is still in the process of collecting repayment and soliciting more participants, so these results should be considered preliminary.

Table 8 replicates the same specification seen in Table 3 for the five nearest-neighbours, first showing the results for OLS with and without baseline covariates and then using 2SLS. In order to maximize the explanatory power of the instrument, I use average voucher values and average assignment-to-treatment as individual instruments, rather than the composite price measure I used in the previous section. For the OLS specification, peer

Table 8: Kati - impact of neighbour's CRO take up on own take up - 5 closest neighbours

	OLS			2SLS		
	(1) Basic	(2) Restricted	(3) Restricted + Controls	(4) Basic	(5) Restricted	(6) Restricted + Controls
# of neighbours adopting	0.0944***	0.0980***	0.0788***	0.206***	0.225***	0.231**
	(0.0186)	(0.0192)	(0.0211)	(0.0737)	(0.0749)	(0.0915)
Voucher (tsh '000)	0.00114*	0.00117*	0.00123*	0.000819	0.000777	0.000953
	(0.000604)	(0.000653)	(0.000639)	(0.000649)	(0.000724)	(0.000688)
Gender voucher ('000)	0.00147**	0.00153**	0.00153**	0.00114	0.00116	0.00129*
	(0.000641)	(0.000672)	(0.000643)	(0.000695)	(0.000736)	(0.000700)
Attended meeting	0.133***	0.136***	0.141***	0.138***	0.142***	0.148***
	(0.0260)	(0.0274)	(0.0292)	(0.0274)	(0.0296)	(0.0321)
Constant	-0.0202	-0.00417	0.171	-0.0450	-0.0331	-0.0365
	(0.0461)	(0.0518)	(2.177)	(0.0458)	(0.0498)	(0.0542)
Baseline controls	No	No	Yes	No	No	Yes
Adj. R-Square	0.0892	0.0922	0.107	0.0135	-0.00721	-0.0217
Obs	684	615	615	684	615	615
C-D Wald F-stat				12.38	13.38	9.129

Dependent variable is a dummy variable = 1 if the household purchases a CRO

Basic columns include only controls shown + # of neighbours attending meeting and a control for whether household has neighbours outside treatment block

Restricted columns are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data

Restricted + Controls columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence RL ownership, electricity access, number of buildings, recent parcel investment, monthly income, assets, average schooling and hh size

Instruments in 2SLS specification: average programme treatment status and average priced faced by neighbours

Conley-adjusted standard errors in parentheses, $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

effects are of similar magnitude to the results from the Barafu experiment, with each neighbour adopting associated with a 8-9.5% increase in the probability the household will also adopt. As before, the 2SLS result is significantly higher, with each neighbour adopting associated with a 20% increase in the predicted probability (note that this is the maximum peer effect size allowed in this specification). Note that the Cragg-Donald Wald F statistic, reported at the bottom of the table, is quite low, so these results should be taken with some caution.[25]

In Appendix A.1, Table 15 shows the results from replication of the specification across different-sized neighbour sets. The results seem reasonably robust to variations of the peer group, again showing a decreasing effect as the neighbour set grows to include parcels which are further away. Not all results are significant at the 10% level, but the coefficients are on the same order of magnitude of the results from Barafu.

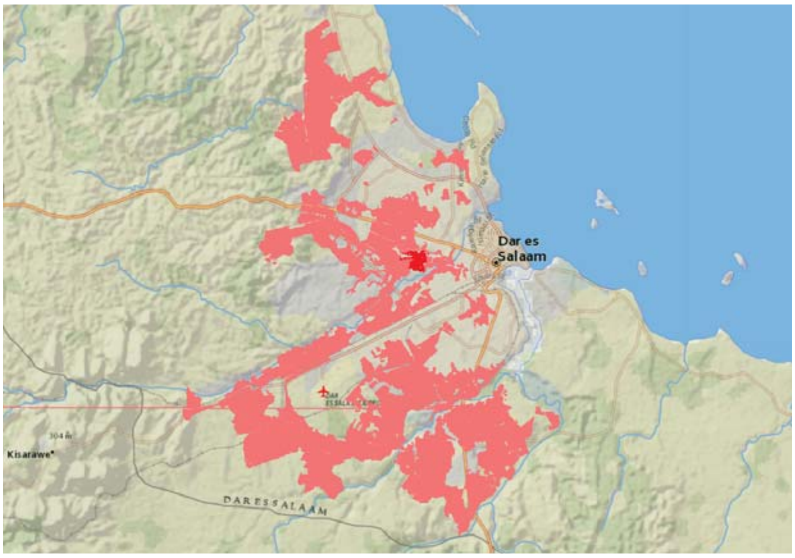
7.2 Peer effects in residential license take-up

While endogenous peer effects appear to be a determinant of take-up in both Mburahati Barafu and Kigogo Kati, it is not immediately clear that the results are generalisable to other settings or necessarily scalable. This is a common criticism of micro-empirical work, including most randomised controlled trials (Ravallion 2008; Deaton 2010) and one that is rarely dealt with.

Ideally, a replication of the experiments in Barafu and Kati at a larger level would show that these results are scalable. For lack of such an experiment, I turn to administrative data: in 2005, the Ministry of Lands embarked on a large-scale effort to register informally-held property in each of the three municipalities which make up Dar es Salaam (Kinondoni, Illala and Temeke). As part of this exercise, land officers and town planners used areal photographs of unplanned settlements to map out the size and location of approximately 219,000 land parcels. Together, these parcels comprise approximately 1.5 million residents, or roughly 50% of the population of Dar es Salaam at the time of the data collection. Figure 4 displays the location of these land parcels, including Barafu and Kati. Following this, the Ministry and the three municipal governments interviewed

[25] Also, this CDW F-stat has not been corrected for spatial correlation.

Figure 4: Coverage of DSM Municipality GIS data



Dark red shapes indicate Mburahati Barafu and Kigogo Kati

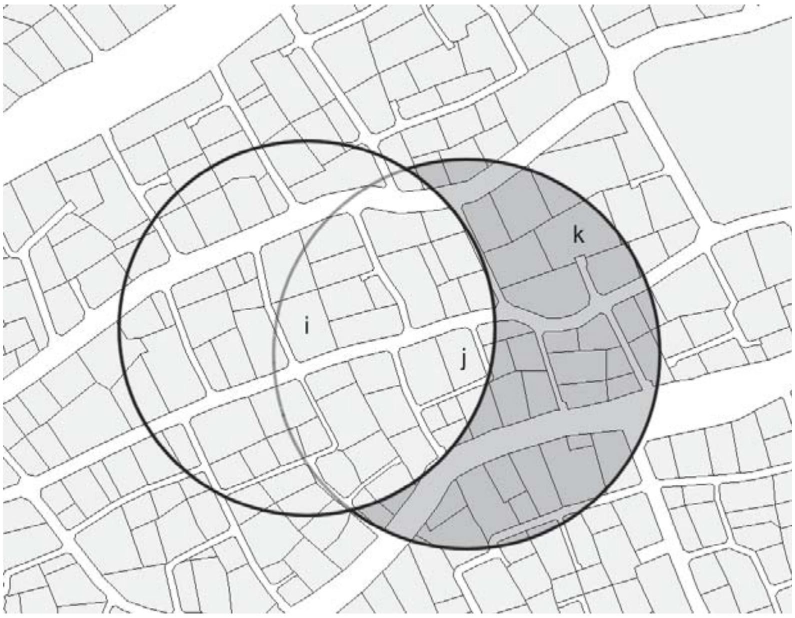
parcel owners to construct land registry, containing basic information on both parcel and owning-household characteristics.

Also available are records for every purchase of a residential license, the short-term land title mentioned in Section 2, from the time they first became available in mid-2005 until early 2013. By matching the GIS-coded map data to the municipal registry and residential license data, it is possible to investigate whether or not peer effects in residential license adoption exist at a larger scale.

While there is no experimental variation in residential license take up, the peer effects literature has developed several methods of identifying peer effects, given some limiting assumptions on how neighbours interact. A common method of overcoming the reflection problem is to take advantage of the structure of partially-overlapping peer groups. This is this case when peer group structures are not transitive; for example, when j being part of of i's peer group and k being part of j's peer group does not guarantee that k will be in i's peer group. When this is the case, there were will be characteristics in k's equation which can be used as instruments for j's adoption.

The intuition is this: k's exogenous characteristics affect k's adoption decision directly,

Figure 5: Example of excluded neighbours



Note: Black circles indicate the boundaries of i and j's neighbour set. Shaded area indicates *excluded* neighbours.

and thus j's adoption decision indirectly (through the endogenous peer effect).[26] Since i doesn't directly interact with k, the latter's exogenous characteristics *only* affect i's adoption through j's adoption. Thus, k's exogenous characteristics satisfy the exclusion restriction and are potential instruments for j's take up. This method was developed simultaneously by Bramoullé, Djebbari, and Fortin (2009) and De Giorgi, Pellizzari, and Redaelli (2010) and has since become a popular method of overcoming the simultaneity bias inherent in peer effects models.

Figure 5 shows a hypothetical case using a map of Mburahati Barafu. In this example, the shaded area indicates all parcels which are in j's peer group, but not in i's peer group, and therefore can be used as "excluded" neighbours. When peer groups are constructed spatially, the partially-overlapping requirement for intransitivity is usually met.

Letting $g(-i)$ indicates household i's neighbour set and $g^2(-i)$ indicate set of neigh-

[26] In the presence of exogenous (contextual) effects, k's characteristics will also affect j's adoption decision directly.

bours of i's neighbours (both discluding i), reconsider the empirical adoption equation:

$$T_i = \alpha + \rho \overline{T}_{g(-i)} + x_i \beta + \overline{x}_{g(-i)} \delta + u_i + \epsilon_g + \varepsilon_i$$

To identify ρ, we need to instrument $\overline{T}_{g(-i)}$ with $\overline{x}_{g^2(-i)}$, the average exogenous characteristics of the neighbours of i's neighbours.[27] The exclusion restriction, that $\overline{x}_{g^2(-i)}$ only affects T_i through $\overline{T}_{g(-i)}$, is heavily dependent on the assumption that household i doesn't interact with households outside of its designated peer group. While this assumption is more easily defended when peer groups are defined by a rigid structure (such as friends in a network roster), in a dense slum it is a little more precarious. Furthermore, while the neighbours-of-neighbours approach theoretically deals with the reflection problem, it does not eliminate correlated effects. For example, if wealth is positively correlated with unobserved land quality, but also affects residential license adoption, then the wealth of excluded neighbours might be correlated with unobserved land quality in i's equation as well.

My approach is as follows: first I restrict the sample and neighbour-sets to all non-empty parcels with non-missing observations for a set of characteristics, which reduces the total sample size to approximately 169,000 land parcels. This includes the log of the parcel's area, a dummy for the parcel being used for residential purposes, a dummy for the parcel being used for both residential and commercial purposes, whether the parcel is on hazard land or not (land deemed unlivable by the government), the number of rooms, the number of households living there, the number of people living there, a dummy equal to one if the parcel has a positive property value, and an interaction between the positive property value dummy and the natural log of the household's value in Tanzanian shillings.

This set of characteristics will be included both as a set of controls for the household/parcel in question i, averaged across the household's neighbour set. For instruments, I have taken a subset of these characteristics for the neighbour's excluded neighbours which are the most informative about $\overline{T}_{g(-i)}$ (the number of rooms, people living on the parcel, and the property value variables). To account for correlated effects, I have

[27] Recall that the identifying information is coming only from *excluded* neighbours. Those which are part of j's neighbour set but not part of i's neighbour set.

included first mtaa/ward fixed effects, then administrative block fixed effects.

Table 9 displays the results from the estimation of the five nearest-neighbours peer effects specification, using residential license take-up as the measure of property rights adoption. The first two columns show the OLS results, while controlling for neighbour characteristics, first with mtaa fixed effects then with block fixed effects. The second two columns show results from the 2SLS specification, when the neighbour-set's average take up of CROs is instrumented with the characteristics the neighbours of neighbours. Both OLS and 2SLS estimates are of similar magnitude to what was seen in both Kigogo Kati and Mburahati Barafu. However, the 2SLS estimates which incorporate controls for correlated effects, when administrative block fixed effects are included, are approximately 30% larger than the estimates for CRO adoption (11.8% in Table 9 versus 14.8% in Table 3).

Given that residential licenses only have a limited tenure value, as they must be renewed every five years, it is possible that there is less room for complementarities in reducing expropriation risk. Furthermore, given the differences in the two forms of tenure, and the fact that these results are based on a eight-year span of adoption, it is likely that the exogenous peer effects revealed here are operating through entirely different channels than in the field experiments.

Finally, Table 10 shows the results for other neighbour set sizes. Again, the results are of a very similar magnitude to what was seen in the experimental data, especially when the neighbour set is extended to the twenty nearest-neighbours. Peer effect estimates also seem to decline as the size of the neighbour set grows. It should be noted that most of these specifications seem to suffer from a weak instrument problem, as Kleibergen-Paap F statistic (standard errors are clustered at the block level, violating the homoscedasticity assumptions necessary for using standard Cragg-Donald test)r is very low. Also, some specifications also fail their overidentification tests, suggesting that the instruments here are not entirely valid.[28] Despite these problems, there is still evidence here that residential license uptake decisions are correlated, possibly as a result of strategic complementarities in their adoption.

[28] Although in a world of heterogenous peer effects, the Hansen J test may just be highlighting the local average treatment effect interpretation.

Table 9: Dar es Salaam - impact of neighbor's RL adoption on own adoption - 5 nearest neighbors

	OLS		2SLS	
	(1)	(2)	(3)	(4)
# neighbors adopting RL	0.0489***	0.0341***	0.193***	0.191***
	(0.00258)	(0.00160)	(0.0199)	(0.0268)
# households living on plot	-0.0146***	-0.0145***	-0.0141***	-0.0142***
	(0.000938)	(0.000807)	(0.000949)	(0.000846)
# rooms in house	0.0133***	0.0128***	0.0120***	0.0121***
	(0.000699)	(0.000590)	(0.000726)	(0.000622)
# people living on plot	6.22e-10***	8.32e-10***	3.57e-10***	3.89e-10***
	(5.97e-12)	(6.71e-12)	(3.64e-11)	(7.56e-11)
Plot is on hazard land	-0.488***	-0.487***	-0.482***	-0.482***
	(0.0143)	(0.00943)	(0.0156)	(0.0105)
Ln(parcel area)	0.0318***	0.0333***	0.0325***	0.0326***
	(0.00377)	(0.00253)	(0.00373)	(0.00269)
Mixed use plot	0.0474***	0.0475***	0.0488***	0.0483***
	(0.0138)	(0.0119)	(0.0145)	(0.0125)
Residential plot	0.00361	0.00530	0.0168	0.0167
	(0.0113)	(0.0100)	(0.0115)	(0.0106)
Property value > 0	-0.190***	-0.188***	-0.195***	-0.195***
	(0.0146)	(0.0106)	(0.0151)	(0.0115)
Ln(Prop value) * (value > 0)	0.0107***	0.0106***	0.0112***	0.0111***
	(0.000864)	(0.000633)	(0.000906)	(0.000698)
Constant	0.0000137	0.0000172	0.00000539	0.00000669
	(0.0000166)	(0.0000166)	(0.0000140)	(0.0000165)
Neighbor controls	Yes	Yes	Yes	Yes
Mtaa fixed effects	Yes	No	Yes	No
Block fixed effects	No	Yes	No	Yes
Adj. R-Square	0.0798	0.0608	-0.0350	-0.0677
Obs	168822	168822	168822	168822
C-D Wald F-stat			2119.2	7.122
Hansen J p-value			0.395	8.01e-10

Dependent variable is a dummy variable = 1 if the household purchases a RL in 2005-2013
Neighbor controls are average values of household/parcel characteristics for neighbor set
Instruments in 2SLS specification are average values of household/parcel characteristics for excluded neighbors of neighbors. Standard errors clustered at mtaa level in columns (1) and (3) and at block level in (2) and (4). $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table 10: Dar es Salaam - impact of neighbor's RL adoption on own adoption – different neighbour sets

	Nearest-neighbour groups			
	(1)	(2)	(3)	(4)
	5	10	15	20
# neighbours adopting RL	0.191***	0.0908***	0.0497***	0.0344***
	(0.0268)	(0.0129)	(0.00896)	(0.00741)
Neighbour controls	Yes	Yes	Yes	Yes
Admin block f.e.	Yes	Yes	Yes	Yes
Adj. R-Square	-0.0677	0.00959	0.0416	0.0490
Obs	168822	168822	168822	168822
C-D Wald F-stat	7.122	6.746	6.296	6.371
Hansen J p-value	8.01e-10	0.0885	0.190	0.440

Dependent variable is a dummy variable = 1 if the household purchases a RL in 2005-2013
Neighbor controls are average values of household/parcel characteristics for neighbor set
Instruments in 2SLS specification are average values of household/parcel characteristics for excluded neighbors of neighbors.
Standard errors clustered at block level $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

8 Conclusion

The options for the many developing countries grappling with high levels of urban growth could be boiled down to formality-by-force or formality-by-nudge. The former is characterized by high levels of urban planning and slum clearance, bringing new arrivals immediately into the formal system and dragging in older ones kicking and screaming. For the latter, the incentives to switch are introduced ex-post, through the introduction of simple, robust formal tenure systems and slum-upgrading. After advocating the former camp for decades following independence, the Tanzanian government has finally found itself pushing the latter. However, its efforts to entice informal settlements to shift to a new tenure system have broadly failed, partly due to the government's lack of knowledge of how to spur demand for land titles.

In this paper, I set out to determine whether or not endogenous peer effects in land titling adoption exist. Using the results from two randomised controlled trials in Dar es Salaam, I exploited random variation in the incentive to title in order to identify the impact of a neighbour's adoption on a household's propensity to adopt. The results suggest there are strong, positive endogenous peer effects, and these results are robust to different neighbour set specifications, as well as a replication of the main experiment in a second location. There is also evidence that positive exogenous peer effects are present at a much larger scale, as results from municipal records suggest that residential license take up show similar signs of being contagious. While the exact mechanism for these results is elusive, evidence strongly points towards geographic proximity as a determinant of the size of the peer effects. This, combined with evidence that households with a higher ex-ante perception of expropriation risk are more responsive to peer effects, suggests that perceived complementarities in risk-reduction are driving the result.

This paper has established that not only is encouraging take-up possible, as evidenced by the effectiveness of the land titling programme, but it also has positive spillovers which can encourage larger levels of adoption. This is encouraging not only from a narrow policy perspective, but it also suggests that landowners consider formal property rights to be complementary, that is, more useful to purchase if everyone else is doing the same.

If large externalities to land titling adoption do exist, then why haven't more communities embraced large scale formalisation, even without further government intervention? While the cost of an individual cadastral survey is prohibitively expensive, en-mass surveying can be considerably cheaper. Given that the demand for title has been shown to be substantial once these hurdles have been overcome (in this instance, by our intervention), the fact that households had not already coordinated to take advantage of these returns to scale suggests these communities already face significant barriers to collective action. These are not universally insurmountable, as there are a few examples of communities in Dar es Salaam coordinating to get the entire neighbourhood titled.[29] What remains to be seen is what policies best take advantage of this social multiplier effect and whether or not it is enough to ensure a full shift to a formal system.

[29] Magigi and Majani (2006) presents a case study of an informal community in Dar es Salaam with atypically high social capital organising a full cadastral survey of the entire unplanned settlement.

References

Ali, D. A., M. Collin, K. Deininger, S. Dercon, J. Sandefur, and A. Zeitlin (2014). The price of empowerment: Experimental evidence on land titling in tanzania. Technical report.

Angelucci, M. and G. De Giorgi (2009). Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? *The American Economic Review 99*(1), 486–508.

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

Besley, T. (1995). Property rights and investment incentives: theory and evidence from Ghana. *Journal of Political Economy 103*(5), 903–937.

Besley, T. and M. Ghatak (2010). Property rights and economic development. In D. Rodrick and M. Rosenzweig (Eds.), *Handbook of Development Economics*. Amsterdam: North-Holland.

Blattman, C. (2011). The trials of randomization *Chris Blattman* (blog), April 23, 2011. `http://chrisblattman.com/2011/04/23/the-trials-of-randomization/`.

Bobonis, G. and F. Finan (2009). Neighborhood peer effects in secondary school enrollment decisions. *The Review of Economics and Statistics 91*(4), 695–716.

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

Brennan, J. (2007). Between segregation and gentrification: Africans, Indians, and the struggle for housing in Dar es Salaam, 1920-1950. In J. Brennan, A. Burton, and Y. Lawi (Eds.), *Dar es Salaam: histories from an emerging African metropolis*, pp. 223–233. Mkuki Na Nyota Publishers.

Brock, W. and S. Durlauf (2001). Discrete choice with social interactions. *The Review of Economic Studies 68*(2), 235–260.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: randomization in practice in development field experiments. *American Economic Journal: Applied Eco-*

nomics 1(4), 200–232.

Bulow, J. I., J. D. Geanakoplos, and P. D. Klemperer (1985). Multimarket oligopoly: Strategic substitutes and complements. *The Journal of Political Economy 93*(3), 488–511.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2012, July). Understanding peer effects in financial decisions: Evidence from a field experiment. NBER Working Papers 18241, National Bureau of Economic Research, Inc.

Caeyers, B. (2013). The exclusion bias in social interaction models: cause, consequences and solution. Working paper.

Cameron, A., J. Gelbach, and D. Miller (2011). Robust inference with multiway clustering. *Journal of Business ands Economic Statistics 29*(2), 238–249.

Collin, M., S. Dercon, H. Nielson, J. Sandefur, and A. Zeitlin (2012). The practical and institutional hurdles to obtaining land titles in urban Tanzania. IGC report, International Growth Centre.

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

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

Dahl, G., K. Løken, and M. Mogstad (2012). Peer effects in program participation. Working Papers in Economics 12/12, University of Bergen, Department of Economics.

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

De Meza, D. and J. R. Gould (1992). The social efficiency of private decisions to enforce property rights. *Journal of Political Economy*, 561–580.

De Soto, H. (2000). *The mystery of capital: why capitalism succeeds in the West and fails everywhere else.* New York: Basic Books.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature 48*(2), 424–455.

Dower, P. and E. Potamites (2012, September). Signaling credit-worthiness: Land titles, banking practices and formal credit in Indonesia. Working Papers w0186, Center for Economic and Financial Research (CEFIR).

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

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

Field, E. (2005). Property rights and investment in urban slums. *Journal of the European Economic Association 3*(2-3), 279–290.

Fort, R., R. Ruben, and J. Escobal (2006). Spillover and externality effects of titling on investments: evidence from Peru. In *11th Annual Meeting of the Latin American and Caribbean Economic Association*.

Galiani, S. and E. Schargrodsky (2010). Property rights for the poor: effects of land titling. *Journal of Public Economics 94*(9), 700–729.

Glaeser, E., J. Scheinkman, and B. Sacerdote (2003). The social multiplier. *Journal of the European Economic Association*, 345–353.

Godlonton, S. and R. Thornton (2012). Peer effects in learning HIV results. *Journal of Development Economics 97*(1), 118–129.

Guryan, J., K. Kroft, and M. Notowidigdo (2009). Peer effects in the workplace: evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics 1*(4), 34.

Hooper, M. and L. Ortolano (2012). Confronting urban displacement social movement participation and post-eviction resettlement success in Dar es Salaam, Tanzania. *Journal of Planning Education and Research 32*(3), 278–288.

Horrace, W. and R. Oaxaca (2006). Results on the bias and inconsistency of ordinary least squares for the linear probability model. *Economics Letters 90*(3), 321–327.

Imbens, G. and J. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Kironde, J. (1995). Access to land by the urban poor in Tanzania: some findings from Dar es Salaam. *Environment and Urbanization 7*(1), 77–96.

Kironde, J. L. (2009). Improving land sector governance in Tanzania: implementation of the land governance assessment framework. Working paper, Ardhi University.

Kombe, W. (2005). Land use dynamics in peri-urban areas and their implications on the urban growth and form: the case of Dar es Salaam, Tanzania. *Habitat International 29*(1), 113–135.

Kombe, W. (2010). Land conflicts in Dar es Salaam: who gains? who loses? Cities and Fragile States Working Paper 82, Crisis States Research Centre, London School of Economics and Polical Science.

Lalive, R. and M. Cattaneo (2009). Social interactions and schooling decisions. *The Review of Economics and Statistics 91*(3), 457–477.

Libecap, G. and D. Lueck (2011). The demarcation of land and the role of coordinating property institutions. *The Journal of Political Economy 119*(3), 426–467.

Magigi, W. and B. Majani (2006). Community involvement in land regularization for informal settlements in Tanzania: a strategy for enhancing security of tenure in residential neighborhoods. *Habitat international 30*(4), 1066–1081.

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

Moffitt, R. et al. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, 45–82.

Ndezi, T. (2009). The limit of community initiatives in addressing resettlement in Kurasini ward, Tanzania. *Environment and Urbanization 21*(77), 77–88.

Ngatia, M. (2011). Social interactions and individual reproductive decisions. Working paper, Yale University.

Oster, E. and R. Thornton (2009). Determinants of technology adoption: private value and peer effects in menstrual cup take-up. NBER Working Papers 14828, National Bureau of Economic Research.

Ravallion, M. (2008). Evaluation in the practice of development. Policy Research Working Paper Series 4547, The World Bank.

Sacerdote, B. (2001). Peer effects with random assignment: results for Dartmouth roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Schelling, T. C. (1978). *Micromotives and macrobehavior*. WW Norton.

Van Tassel, E. (2004). Credit access and transferable land rights. *Oxford economic papers 56*(1), 151–166.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*, Volume 1. The MIT Press.

Zeitlin, A. (2012). Identification and estimation of peer effects on endogenous affiliation networks: an application to Ghanaian agriculture. Working paper.

A Appendix

A.1 Additional tables

Table 11: Barafu - main treatment and control balance

	Treatment	Control	T-test	OLS
	(1)	(2)	(3)	(4)
# Rooms	5.118	5.033	-0.560	0.085
	(2.604)	(2.175)	[0.575]	(0.149)
Electrical connection	0.739	0.733	-0.264	0.007
	(0.439)	(0.443)	[0.792]	(0.026)
Owner occupied	0.847	0.848	0.023	-0.000
	(0.360)	(0.360)	[0.982]	(0.022)
Tenants on parcel	0.636	0.602	-1.170	0.034
	(0.482)	(0.490)	[0.242]	(0.029)
Access to road	0.243	0.231	-0.471	0.012
	(0.429)	(0.422)	[0.638]	(0.025)
Log(parcel area m^2)	5.298	5.336	1.089	-0.038
	(0.605)	(0.594)	[0.276]	(0.035)

Notes: Columns (1)-(2) show means for treatment and control groups, standard deviations in ( ). Column (3) shows test statistic for t-test, p-values in [ ]. Column (4) shows coefficient and standard error for OLS regression of outcome variable on treatment.

Table 12: Barafu - impact of neighbour's adoption for neighbours within distance d

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.0657^{**}	0.0512^{**}	0.0432^{**}	0.0307^{**}	0.0277^{**}
	(0.0124)	(0.0101)	(0.0086)	(0.008)	(0.0068)
Restricted	0.0652^{**}	0.0458^{**}	0.0418^{**}	0.0294^{**}	0.0267^{**}
	(0.0136)	(0.0113)	(0.0094)	(0.0086)	(0.0075)
Covariates	0.0666^{**}	0.0445^{**}	0.039^{**}	0.0245^{**}	0.024^{**}
	(0.0139)	(0.0119)	(0.0103)	(0.0098)	(0.0088)
2SLS					
Basic	0.0762^{**}	0.0475^{**}	0.0337^{**}	0.0248^{**}	0.0247^{**}
	(0.0273)	(0.0178)	(0.0126)	(0.0099)	(0.0083)
Restricted	0.0806^{**}	0.0383^{*}	0.0338^{**}	0.025^{**}	0.0242^{**}
	(0.033)	(0.0211)	(0.0145)	(0.0113)	(0.0097)
Covariates	0.1003^{**}	0.0511^{**}	0.0429^{**}	0.0301^{**}	0.0302^{**}
	(0.0323)	(0.0202)	(0.0141)	(0.0116)	(0.0101)
Distance band length (m) =	10	15	20	25	30

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different distance band (i.e. 10b = all neighbours within 10 meters) Conley-adjusted standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table 13: Barafu - impact of neighbour's adoption for nth nearest-neighbour sets, including previously-surveyed neighbours

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.1094**	0.0795**	0.0521**	0.0433**	0.0341**
	(0.0235)	(0.0167)	(0.0101)	(0.008)	(0.0071)
Restricted	0.1045**	0.086**	0.0501**	0.0384**	0.0323**
	(0.0236)	(0.0169)	(0.0108)	(0.0085)	(0.0069)
Covariates	0.0995**	0.0801**	0.0436**	0.0301**	0.0296**
	(0.0242)	(0.0178)	(0.0117)	(0.0099)	(0.0081)
2SLS					
Basic	0.2276**	0.1395**	0.0563**	0.0396**	0.0316**
	(0.0665)	(0.0457)	(0.0222)	(0.0136)	(0.0105)
Restricted	0.2353**	0.1537**	0.0496**	0.0349**	0.0263**
	(0.06)	(0.0434)	(0.0212)	(0.013)	(0.0101)
Covariates	0.2268**	0.1513**	0.0616**	0.0362**	0.03**
	(0.0618)	(0.0414)	(0.0205)	(0.0133)	(0.0097)
# nearest neighbours =	3	5	10	15	20

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3n = 3 closest neighbours). Conley-adjusted standard errors in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table 14: Barafu - impact of neighbour's adoption for nth nearest-neighbour sets, without meeting controls

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.0917**	0.0735**	0.0476**	0.0351**	0.0246**
	(0.0232)	(0.0173)	(0.0107)	(0.0087)	(0.0075)
Restricted	0.0941**	0.0814**	0.0447**	0.0328**	0.0256**
	(0.0235)	(0.0173)	(0.0113)	(0.0091)	(0.0074)
Covariates	0.0919**	0.0822**	0.0417**	0.0267**	0.0248**
	(0.0242)	(0.0181)	(0.0121)	(0.0102)	(0.0078)
2SLS					
Basic	0.2478**	0.1421**	0.0392*	0.0262**	0.0147
	(0.0604)	(0.0435)	(0.0202)	(0.0132)	(0.0098)
Restricted	0.1911**	0.1309**	0.0352*	0.0211	0.0145
	(0.0611)	(0.0435)	(0.0213)	(0.0134)	(0.0097)
Covariates	0.2024**	0.1426**	0.0511**	0.0338**	0.0261**
	(0.0628)	(0.041)	(0.0213)	(0.0133)	(0.0088)
# nearest neighbours =	3	5	10	15	20

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3n = 3 closest neighbours). Conley-adjusted standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table 15: Kati - impact of neighbour's adoption for nth nearest-neighbour sets

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.1141**	0.0944**	0.063**	0.0456**	0.0379**
	(0.0247)	(0.0186)	(0.0116)	(0.0094)	(0.0071)
Restricted	0.1201**	0.098**	0.0662**	0.0431**	0.0361**
	(0.0261)	(0.0192)	(0.0121)	(0.0095)	(0.0073)
Covariates	0.1122**	0.0788**	0.0503**	0.0316**	0.0247**
	(0.0269)	(0.0211)	(0.0149)	(0.0107)	(0.0086)
2SLS					
Basic	0.1869	0.2061**	0.0957**	0.0605**	0.034
	(0.1293)	(0.0737)	(0.0468)	(0.0277)	(0.0242)
Restricted	0.2249*	0.2248**	0.1079**	0.0606**	0.0324
	(0.1239)	(0.0749)	(0.0423)	(0.0252)	(0.0218)
Covariates	0.1437	0.2311**	0.1217*	0.0454	0.0256
	(0.1301)	(0.0915)	(0.0727)	(0.0511)	(0.0484)
# nearest neighbours =	3	5	10	15	20

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3n = 3 closest neighbours). Conley-adjusted standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$

Table 16: Percentage of predictions outside of [0,1] in LPM model (Barafu)

# Neighbours	3	5	10	15	20
OLS					
Basic	0	0	.0022	0	0
Restricted	0	0	0	.0024	0
Restricted + Controls	.021	.038	.043	.036	.033
IV					
Basic	.036	.062	.04	.04	.036
Restricted	.02	.0022	.0022	0	0
Restricted + Controls	.019	0	0	0	0

Each cell shows the % of observations with predicted values which fall outside of the [0,1] interval for a given specification.
Results are from Barafu nearest-neighbour regressions.

A.2 Extra robustness: block fixed effects and outside-neighbour set adoption

To ensure that the main results in this paper are not being driven by block-level unobservables, I have re-run the main specification with treatment block fixed effects. The results are presented in Table 17. The results are broadly similar to what has been seen before, although they are less-precisely estimated.

Another concern is the identification of the correct neighbour set. Under the assumption that a nearest-neighbour set of size n is the "correct" neighbour set (i.e. it captures all neighbours relevant for the CRO adoption decision), then estimates of ρ using smaller nearest-neighbour sets may be biased upward. For instance, if the main specification (1) is estimated using the index household's five nearest-neighbours, the titling decisions of neighbours outside of this group will influence *both* the five nearest-neighbors and the household in question. Therefore, nearest-neighbour sets which are "too small" will also be proxying for the larger, "correct" neighbour sets, and so per-neighbour estimates will be biased upwards. This does not prevent identification of ρ, but it does complicated its interpretation. Instrumenting might take care of this problem, but the instruments used here (the percentage of neighbour who are treated) will also be correlated across different-sized neighbour sets.

To account for this, Table 18 re-runs the main specification whilst controlling for the percentage of neighbours outside of the chosen neighbour set (up to twenty-nearest neighbours) who have purchased a CRO. So, for example, when the specified neighbour set is the five-nearest neighbours, a control is included for the % of excluded neighbours (those between the sixth and twentieth-nearest neighbours) who have also adopted a CRO. The results indicate that specifications using smaller neighbour sets might be upward-biased (the 2SLS estimates for three-nearest neighbours are 0.159 for the full specification, versus 0.203 in Table 4). However, estimates of larger neighbour sets are very close to what was seen in previous results.

Table 17: Barafu - impact of neighbour's adoption - nearest neighbour - block fixed effects

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.077^{**}	0.0668^{**}	0.0489^{**}	0.0379^{**}	0.0328^{**}
	(0.0246)	(0.0182)	(0.0117)	(0.01)	(0.0098)
Restricted	0.0702^{**}	0.0687^{**}	0.0421^{**}	0.0341^{**}	0.0302^{**}
	(0.0251)	(0.0191)	(0.0127)	(0.011)	(0.0112)
Covariates	0.0642^{**}	0.068^{**}	0.0352^{**}	0.0189	0.0175
	(0.0253)	(0.0199)	(0.0139)	(0.013)	(0.0121)
2SLS					
Basic	0.2259^{**}	0.1501^{**}	0.0489^{*}	0.0363^{**}	0.0265^{*}
	(0.0715)	(0.0507)	(0.0261)	(0.0168)	(0.0144)
Restricted	0.1664^{**}	0.1398^{**}	0.0511^{*}	0.0314^{*}	0.0177
	(0.0714)	(0.0532)	(0.0267)	(0.0187)	(0.0179)
Covariates	0.1825^{**}	0.1446^{**}	0.0644^{**}	0.0479^{**}	0.0312^{**}
	(0.0734)	(0.0505)	(0.0276)	(0.0176)	(0.0156)
# nearest neighbours =	3	5	10	15	20

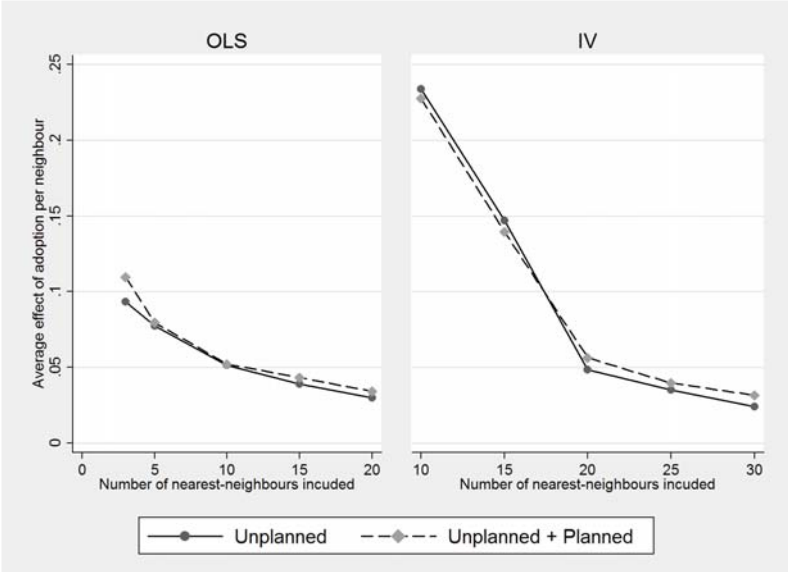
Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3 = 3 closest neighbours). Conley standard errors in parentheses. $^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$

Table 18: Barafu - nearest neighbour - controlling for adoption outside of neighbour set

	(1)	(2)	(3)	(4)	(5)
OLS					
Basic	0.0943**	0.0781**	0.0515**	0.0389**	0.0299**
	(0.0226)	(0.0167)	(0.0104)	(0.0086)	(0.0074)
Restricted	0.0911**	0.0815**	0.0472**	0.0366**	0.0302**
	(0.023)	(0.0171)	(0.0112)	(0.0089)	(0.0072)
Covariates	0.0879**	0.0821**	0.0447**	0.0301**	0.028**
	(0.0233)	(0.0178)	(0.0118)	(0.0101)	(0.0078)
2SLS					
Basic	0.1992**	0.1266**	0.0458**	0.0352**	0.0239**
	(0.0581)	(0.0397)	(0.0182)	(0.0118)	(0.0094)
Restricted	0.1524**	0.1171**	0.041**	0.0303**	0.0217**
	(0.062)	(0.0444)	(0.0202)	(0.0128)	(0.01)
Covariates	0.1597**	0.1299**	0.0555**	0.0402**	0.0292**
	(0.0615)	(0.0406)	(0.0194)	(0.0125)	(0.0088)
# nearest neighbours =	3	5	10	15	20

Dependent variable is a dummy variable = 1 if the household purchases a CRO. **"Basic"** rows include only controls shown & # of neighbours attending meeting and a control for whether household has neighbours outside treatment block. **"Restricted"** rows are the same as basic, except sample and neighbour sets are restricted to households with non-missing baseline data. **"Covariates"** columns include household and average neighbour set controls for Log(parcel area), year of purchase, rental status, owner residence. Each column represents a different nearest-neighbour set (i.e. 3 = 3 closest neighbours). For all nearest-neighbour sets < 20, a control is included for the take up of households outside of the neighbour set, but within the 20-nearest neighbours cut-off. Conley standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Figure 6: Average nearest-neighbour peer effects for both unplanned and all neighbours



A.3 Planned areas

So far, the results that I have presented are solely for unplanned areas of Barafu, with neighbour sets constructed only out of neighbours who have previously not had a cadastral survey. If the knowledge that a treated neighbour will be surveyed increases the perceived value of a CRO, then the instrument might have an effect on take up, outside of the peer effect. To determine whether or not this might be a problem, I first use the baseline data to check and see if there is any correlation between perceived expropriation risk and proximity to surveyed parcels, and I find none (these results are presented in Appendix A.4). Next, I re-run the main specification, this time including surveyed parcels as neighbours. If part of the observed peer effect is actually proxying for a surveyed-effect, then peer effects from already-surveyed neighbours should be lower. However, including already-surveyed neighbours does not seem to change the results in any meaningful way. Figure 6 shows a comparison between estimates of ρ from both before and after planned neighbours are included, using the basic specification without covariates. Table 13 in Appendix A shows the full set of results, revealing no substantial difference. Given that peer effects for already-surveyed parcels are of a similar magnitude to unplanned parcels, it does not appear that surveying of neighbours is a key factor in take-up decisions.

A.4 Cadastral survey proximity and perceived expropriation risk

To investigate this, I turn to the baseline data collected prior to the intervention. One of the questions asked in the survey requires the land owner to guess the probability that the parcel will be expropriated in the next five years, providing an excellent measure of self-perceived expropriation risk.[30] Using these responses by residents of unplanned areas, I have regressed the perceived probability of expropriation on several measures of proximity to already-surveyed parcels, distance to various geographic features in Barafu, and a set of household and parcel covariates. The results are displayed in Table 19.

The first two columns use two measures of proximity: distance of the household to the nearest cadastral-surveyed parcel and a dummy variable for whether or not the parcel is adjacent to a surveyed parcel. Neither is statistically significant at the 10% level, and while the coefficient on the adjacency dummy is of the "correct" sign, distance from a surveyed parcel seems to counter-intuitively reduce perceived expropriation risk. Columns (3) and (4) use the percentage of the nearest 20 neighbours who are surveyed as a measure, and while the coefficient is negative (indicating that parcels with more surveyed neighbours have lower perceived risk), it is not significant.

While one might be worried that measurement error in perceived expropriation risk might make it difficult to pick up *any* correlation, it is worth noting that many of these results do make sense. Proximity to "hazard land", areas which are deemed by the local government to be unsafe to build and are often subject to mass expropriation, is positive correlated with perceived expropriation risk. Similarly, proximity to the nearest primary road, where many parcels had already been marked for demolition to make way for expansion of existing infrastructure, is also correlated with higher perceived expropriation risk.

[30] Note that these data are only currently available for roughly 65% of the Barafu sample, so the following analysis might be subject to selection bias.

Table 19: Perceived expropriation risk at baseline and proximity to surveyed parcels

	Distance measures		Nearest-neighbor	
	(1)	(2)	(3)	(4)
Dist to nearest surveyed parcel (m)	-0.000103	0.000147		
	(0.000307)	(0.000430)		
HH is adjacent to surveyed parcel?	-0.0288	-0.0108		
	(0.0328)	(0.0343)		
% nearest 20 neighbours surveyed			-0.0967	-0.124
			(0.0898)	(0.110)
Distance to nearest church		0.000119		0.0000791
		(0.000654)		(0.000651)
Distance to nearest open field		0.000226		0.000209
		(0.000260)		(0.000256)
Distance to hazard land		-0.00191***		-0.00185***
		(0.000488)		(0.000464)
Distance to nearest mosque		0.00107*		0.00108*
		(0.000640)		(0.000638)
Distance to local govt office		0.000483		0.000537
		(0.000645)		(0.000642)
Distance to nearest walking path		0.000774		0.000699
		(0.000515)		(0.000498)
Distance to nearest primary road		-0.00207***		-0.00200***
		(0.000780)		(0.000762)
Distance to nearest river		0.000828		0.000784
		(0.000777)		(0.000773)
Distance to nearest secondary road		-0.000529		-0.000703
		(0.000444)		(0.000475)
Distance to nearest school		-0.00103*		-0.000953*
		(0.000530)		(0.000530)
Controls	No	Yes	No	Yes
R-Squared	0.000975	0.0979	0.00162	0.0991
Obs	755	755	755	755

Results are from an OLS regression of the household's self-reported perceived probability of expropriation within next five years on measures of proximity to parcels which have been cadastral surveyed. Sample is limited to households in un-surveyed areas. Robust standard errors in parentheses
$^{*}p < 0.10, ^{**}p < 0.05, ^{***}p < 0.01$