

Farm size shapes friend choice amongst rice producers in China: Some evidence for the theory of network ecology*

Cohen R. Simpson^{a,b,*}

^a Oxford Internet Institute, University of Oxford, United Kingdom

^b Nuffield College, University of Oxford, United Kingdom

ARTICLE INFO

Keywords:

Ecology
Network formation
Friendship
China
Farming

ABSTRACT

Multiple dynamics jointly determine who we befriend, however, researchers have failed to systematically assess which processes matter most under different circumstances. Here I draw on observations around how the demands of paddy rice cultivation shape social interaction to demonstrate that the relative importance of reciprocity, transitivity and generalised exchange to who rice producers choose as friends varies with the amount of agricultural land under their control. In doing so, I use unique data on farm size and friendship amongst 4713 rice-growing smallholders in 162 rural villages in Jiangxi, China along with a new technique for measuring the relative importance of effects in Stochastic Actor-Oriented Models. In line with the micro-level component of the recently advanced “theory of network ecology”, results suggest that features of an individual’s proximal environment can powerfully moderate the relative expression of network-formation mechanisms such that for some individuals, a dynamic may be expected to hold substantial sway over the process of choosing social contacts and, for others, no sway at all.

Introduction

To answer the question “Where do social networks come from?”, scholars have invoked diverse generative mechanisms — i.e., dynamics such as reciprocity, transitivity and homophily that are presumed to govern the establishment of social ties at the micro-level and, in turn, shape macro-level network structure (Robins et al., 2005). Each of these mechanisms has animated its own robust stream of research across the sciences (Rivera et al., 2010). However, network analysts have firmly concluded that the process by which individuals choose their social contacts is multi-mechanistic (Lusher et al., 2012; Monge and Contractor, 2003; Robins et al., 2005) such that networks in diverse modern and non-modern contexts are characterised by the joint operation of several regularities in tie formation (Apicella et al., 2012; Block, 2015; Heidler et al., 2014).

Here my concern is an important new line of sociological inquiry contending that this multi-mechanistic process systematically varies with circumstances. Specifically, McFarland et al. (2014) build on arguments around non-human animal behaviour and the interaction between psychology, action and habitat (Oishi, 2014) to advance a

“theory of network ecology” which posits that: (i) features of the environment within which a network is embedded moderate the expression of generative mechanisms; such that (ii) the same basic micro-level dynamics can produce different macro-level network topologies depending on conditions. For example, in the expansive analysis of adolescent friendship in 131 U.S. secondary schools that McFarland et al. (2014) use to test their theory, the authors find that variability in features of educational institutions (the titular network ecologies) such as their size, the manner in which teaching is organised and their demographic composition all influence the operation of oft-discussed tie-formation processes (e.g., reciprocity, transitivity, and homophily).

Prima facie, the theory of network ecology is an elegant and plausibly general socioecological framework for making sense of the relationship between the structure of face-to-face networks and features of social, cultural and physical environments. Yet the analysis McFarland et al. (2014) use to test their theory — currently the only large-scale test in existence — has a shortcoming that I address here in order to build on the authors’ pioneering work. Specifically, McFarland et al. (2014) rely exclusively on the Exponential-family Random Graph Model (ERGM; Amati et al., 2018; Lusher et al., 2012) in order to

* Code used for the preparation of data and the execution of my analyses and as the saved model objects for the analyses reported herein are available via the Open Science Framework (OSF): <https://osf.io/r3n6e/>. The data from China are made available to all by Jing Cai via the American Economic Association: <https://doi.org/10.1257/app.20130442>. However, for convenience, the version of J. Cai’s dataset used for my analyses is included on the OSF portal for this paper.

* Corresponding author at: 1 St Giles, Oxford OX1 3JS, United Kingdom.

E-mail address: cohen.simpson@oii.ox.ac.uk.

<https://doi.org/10.1016/j.socnet.2019.10.001>

explore how the attributes of secondary schools covary with the magnitude of ERGM parameter estimates (conditional log odds ratios). Unfortunately, however, this approach raises the spectre of a key methodological constraint: at present, there are no formal techniques for directly comparing effects within or across ERGMs (Amati et al., 2018). And although contrasting the magnitude of ERGM parameter estimates is a common workaround (e.g., see Wimmer and Lewis, 2010), such comparisons are ill-advised when using logistic regression-type models due to the counterintuitive properties of odds ratios and their great dependency on data and model specification (Norton and Dowd, 2017; Norton et al., 2018).

Consequently, the theory of network ecology leads one to expect variation in the expression of those mechanisms repeatedly found to play a role in the emergence of real social networks but the analysis currently buttressing this theory cannot tell us whether the *relative* power of tie-formation processes is also ecologically contingent. This is a substantial research gap as scholars have failed to systematically assess which dynamics matter most to an individual when choosing a social contact under different conditions despite: (i) decades of sociological work asserting the power of schools, organisations, neighbourhoods and culture or, more generally, “context”, over network formation (Chan Tack and Small, 2017; Doreian and Conti, 2012; Entwisle et al., 2007; McLean, 2016; Small and Adler, 2019); and (ii) a growing body of anthropological and psychological research detailing how the dynamics around friendship strongly depend on sociocultural ecology (Adams and Plaut, 2003; Hruschka, 2010; Oishi and Kesebir, 2012; Thomson et al. 2018). Accordingly, here I test an extension of the micro-level component of McFarland et al.’s (2014) theory that is encapsulated in a straightforward argument — network ecology moderates the relative importance of generative mechanisms.

The current study

To judge both the plausibility of my conjecture and the potential generalisability of the theory of network ecology to adults in non-educational scenarios, I carried out a large-*N* study of choice of social contacts amongst “smallholders” in the People’s Republic of China — individuals with farms less than two hectares in size but who are still vital to food security for the 1.37 billion people that call their country home (Cohn et al., 2017; Cui et al., 2018; Fan and Chan-Kang, 2004; Guo et al., 2018; Wu et al., 2018; Zhang et al., 2016). More specifically, I analysed distinctive cross-sectional data documenting asymmetric “close friendships” amenable to the discussion of “rice production or financial issues” amongst the heads of virtually all rice-producing households in 162 rural villages in Jiangxi Province (*N* = 4713) with the aid of a new technique for measuring the relative importance of effects in Stochastic Actor-Oriented Models (SAOMs).

As for moderators of the relative expression of generative mechanisms, dimensions of variation germane to agriculture and rural China were not immediately obvious when relying solely on the work of McFarland et al. (2014) as the authors develop their propositions around network ecology in direct relation to U.S. secondary education. Therefore, I turned to cognate but broader socioecological research wherein anthropologists and psychologists have actively explored how diverse dispositions and behaviours are shaped, in part, by the environments within which humans are embedded (Oishi, 2014).

An important subtheme of this research is the downstream implications of “subsistence style” — i.e., the primary manner in which humans interact with their physical surroundings to make their living (Sutton and Anderson, 2009); where the interdependence of these survival strategies — that is, the degree to which one’s livelihood requires them to depend on others (Glowacki and Molleman, 2017) — is regarded as a key behavioural determinant (Glowacki and Molleman, 2017; Talhelm, 2020; Talhelm and Oishi, 2018; Talhelm et al., 2014; Thomson et al., 2018; Uchida et al., 2018). For example, in an experiment with 75 Nyangatom men in Ethiopia, Glowacki and Molleman

(2017) demonstrate that pastoralists (i.e., nomadic individuals who repeatedly move across large distances between camps) and urban dwellers (i.e., wage labourers embedded in a market economy) relied on social learning more than the comparatively independent horticulturalists in their sample who meet their subsistence needs at the individual and household levels through fishing and small-scale farming with little help from other families or need for dispersed information. Furthermore, in an analysis of data from 39 countries, Thomson et al. (2018) find that lower levels of “relational mobility” (i.e., the freedom and opportunity to establish and terminate interpersonal relationships based on personal preference) are associated with more interdependent survival strategies (i.e., herding [relatively mobile and independent] versus wheat farming [more settled and interdependent] versus paddy rice farming [most settled and interdependent]). And in a study of 408 farming, fishing and urban communities in Japan, Uchida et al. (2018) show that participation in collaborative public projects such as festivals and infrastructure maintenance on the part of farming and non-farming residents is associated with having a greater proportion of farmers (predominantly rice-growers) in their community, leading the authors to conclude that economic activity shapes social interaction.

Inspired by these lines of enquiry, here I focus on the implications of paddy rice production for how the farmers in Jiangxi establish friendships with one another. To be clear, my goal is not to compare and contrast network formation in the face of different agricultural regimes common in China (e.g., rice farming versus wheat farming; see Hu and Yuan, 2015). Instead, I am interested in rice production in China in and of itself and what the qualities of this mode of subsistence might mean for how resident rice producers build friendships with their agricultural counterparts.

With that in mind, I investigate, in particular, variation in the relative importance of generative mechanisms in relation to the amount of land that an individual farmer has that is devoted to producing rice (i.e., rice production area or simply “farm size”).¹ Although it may seem rather arbitrary, agricultural land has clear behavioural implications. Indeed, the field of rural sociology has long recognised the broad social consequences of farm size and the impact of the structure and the scale of agricultural production on farmers’ values and behaviour in the United States (Finn and Buttel, 1980) and the United Kingdom (Gasson, 1977), where research focused specifically on life in rural China has extensively detailed how land underpins farmers’ social, economic and political decisions and outcomes (Chen et al., 2010; Deininger et al., 2014; Cai, 2016; Oi and Rozelle, 2009; van der Ploeg et al., 2014). That said, farm size is most appropriately viewed as a proximal, micro-level feature of a rice producer’s network ecology, whereas past socioecological research — including the analysis provided by McFarland et al. (2014) and scholarship on the interplay between physical space and network formation (Small and Adler, 2019) — generally focuses on more distal factors (see Oishi, 2014). Nevertheless, intentionality in face-to-face social networks rests with the individual (Robins et al., 2005) such that it is not unreasonable to expect that features of a rice producer’s immediate physical environment may play a role in how he manages his portfolio of personal social ties.

Before advancing, it is worth acknowledging that some readers may question the broad relevance of an investigation of network formation

¹ Although no definitive definition of “smallholder” exists (Cohn et al., 2017), here I follow Lowder et al. (2016) to view these individuals as farmers with land areas that are less than two hectares — a metric that reflects the tiny size of farms in China relative to those across the world (see Figure I of Cohn et al., 2017). Amongst the 4713 farmers in the dataset analysed here, over 95% have rice production areas less than 32 *mu* (i.e., 2.13 hectares). Accordingly, the vast majority are “smallholders” whereas the remaining farmers are not under a strict reading of Lowder et al.’s (2016) definition. Seeking to use as much information as possible for my analyses, I instead take a relaxed view whereby farmers with relatively-larger rice production areas are retained.

amongst rural, relatively impoverished, generally land-poor farmers in what the World Bank presently classifies as an upper-middle income country by gross national income per capita. This is especially so as the scientific literature at the intersection of friendship networks and multi-mechanistic models overwhelmingly focuses on children and young people in schools in Western societies such as Germany, the Netherlands, the United Kingdom and the United States (c.f. An, 2015). However, critique of my chosen case as overly idiosyncratic is arguably misplaced as around 80% of farms across the globe are estimated to belong to smallholders and are typically in the low- and middle-income countries containing much of the world's poor (Cohn et al., 2017; Lowder et al., 2016). Moreover, there is great concern about the farming practices of Chinese smallholders, particularly around inefficiency in production, overreliance on environmentally-damaging chemical fertilisers and pesticides, and pathways to the adoption of more sustainable techniques. Scholars of agriculture have underscored the importance of easing farmers' access to knowledge in addressing these issues (Cui et al., 2018; Guo et al., 2018; Wu et al., 2018; Zhang et al., 2013, 2016), making the study of how Chinese smallholders build friendship networks wherein they stand to learn from one another of practical importance. Nevertheless, as a result of my empirical focus, my findings may not generalise: (i) beyond small to medium-sized groups; (ii) to non-agricultural scenarios characterised by greater levels of economic development; (iii) beyond rice areas in East Asia; (iv) to more gender-heterogeneous networks (91% of the rice farmers in the sample are men); or (v) to networks not composed of “friendly” relations.

The social organisation of rice farming in rural China

Let us now turn solely to the question of social structure and land — a vital resource collectively owned at the village level and contracted to individual households (Brandt et al., 2002; Chen et al., 2010; Ho, 2001; Cai and Sun, 2018; Oi and Rozelle, 2009; Vendryes, 2010; Wu et al., 2018) that is both in short supply in China (Zhang et al., 2013) and strongly linked to the security of incomes and equitable living in the country's rural areas (Brandt et al., 2002; Chen et al., 2010; Cai, 2016; van der Ploeg et al., 2014). In order to understand what land might mean for a farmer's decisions about his choice of social contacts, we must first consider the nature of rice farming itself in East Asia.² As Talhelm et al. discuss while advancing their subsistence style “rice theory” of culture and cooperation (Talhelm, 2020; Talhelm and Oishi, 2018; Talhelm et al., 2014), paddy rice production (i.e., standing-water/submerged rice production) is a high-yield, labour-intensive endeavour. By far the most common method of rice growing in Jiangxi (Guo et al., 2018; Tan et al., 2008), paddy-based cultivation requires farmers to make substantial individual physical investments³ and to extensively collaborate with their agricultural counterparts in order to: build and manage irrigation systems; share water; coordinate planting dates and harvesting schedules; and exchange labour (Aoki, 2001; Talhelm, 2020; Talhelm and Oishi, 2018; Talhelm et al., 2014; Uchida

et al., 2018). The practical effect of paddy rice cultivation is thus the entanglement of farmers' lives and livelihoods, a situation succinctly captured by Smith (as cited in Aoki, 2001, p. 46) who states in relation to irrigation systems in the context of traditional Japanese village life that “a rice farmer never owned or controlled all of the essential means of production himself, and he could not individually make all of the critical decisions of farming. He might wish, for instance, to turn an unirrigated field into a paddy, but he would not be allowed to do so if this would impair the water supply of others.”

To the extent that a farmer's portfolio of occupational friends is mismatched to his need to cope with the demands of paddy rice production by relying on colleagues in his village, it is conceivable that he will move to build a personal network amenable to this aim. More precisely, the need to foster a social environment conducive to co-operation in rice production — an economic necessity that is not inevitable (Talhelm and Oishi, 2018) — should incentivise rice farmers to build strong positive social relationships in order to avoid conflict, avert negative reputations and enforce norms against free-riding (Aoki, 2001; Talhelm, 2020; Talhelm and Oishi, 2018; Talhelm et al., 2014; Uchida et al., 2018; Wu and Pretty, 2004). That said, this more instrumental concern is underpinned by a righteous imperative. As van der Ploeg et al. (2014) persuasively detail in their ethnographic study of inter-generational land use, endemic to rural Chinese farming culture is an intermingling of kinship with a set of complementary farming-related rights, duties, tasks, and expectations. This coupling of family and farming ultimately creates a moral obligation whereby smallholders are expected to rely on (i.e., draw sustenance from), tend to, and actively develop their land in order to ensure a level of food production capable of: (i) supporting family consumption; and (ii) providing an income in the present and for future generations (see also Cai, 2016 on land as socioeconomic insurance). Presumably, Chinese smallholders — individuals who suffer gaps in both economic resources and agricultural knowledge (Cui et al., 2018; Fan and Chan-Kang, 2004; Wu et al., 2018; Zhang et al., 2013, 2016) — will work to build beneficial farming-relevant relationships with their counterparts to the extent that these ties allow them to fulfil their duty to the “familial land-labour institution” (van der Ploeg et al., 2014) through agricultural productivity; a duty that should be salient given sociological research pointing to strong cultural norms against failing obligations to kin (Peng, 2004).

As rice farmers construct social networks more conducive to meeting their cooperative needs, oft-discussed tie-formation mechanisms are plausibly implicated. Of particular interest here are direct reciprocity (i.e., responding in kind), transitivity (i.e., befriending friends-of-friends) and generalised exchange (i.e., the unilateral provision of support, here information on rice and finance, amongst three or more actors). I limit my attention to these specific dynamics as the collectivism from which cooperation emerges in East Asian rice areas is intertwined with concerns about reputation (Uchida et al., 2018) and characterised by “reciprocal obligations and tight [i.e., high-cost, duty-laden] relationships with trusted ties” (Talhelm and Oishi, 2018, p. 63; see also work on *guanxi* [Aoki, 2001; Burt et al., 2018; Peng, 2004; Xiao and Tsui, 2007]). Furthermore, these three mechanisms are ubiquitous in their operation across social contexts (Apicella et al., 2012; Block, 2015; McFarland et al., 2014; Rivera et al., 2010), with network analysts and theorists of social exchange historically linking each to trust, positive sentiment, cohesion and group solidarity in diverse scenarios (Baldassarri and Diani, 2007; Burt et al., 2018; Molm, 2010; Moody and White, 2003; Simpson et al., 2017). For example, evidence from Uganda (Baldassarri, 2015) indicates that direct reciprocity during repeated interaction induces cooperation amongst coffee-growing smallholders and evidence from Tanzania (Apicella et al., 2012) suggests that co-operation amongst Hadza hunter gatherers stems from social proximity in networks (i.e., being friends and friends-of-friends). Therefore, direct reciprocity, transitivity, and generalised exchange are context-relevant vantage points from which to explore links between proximal features of Chinese smallholders' network ecology — here, the amount of

² This is an important distinction. As Hu and Yuan (2015) stress, the production costs of rice and other crops are likely to vary in different parts of the world such that the sociocultural implications of growing a certain crop stem from the level of collaboration necessary for that crop's production *given local conditions* as opposed to anything inherent to the crop itself.

³ For example, there is older ethnographic evidence to suggest that traditional rice farming requires double the labour as that which is required by other forms of agricultural production such as for wheat (Fei and Chang, 1949, p. 144–147), where the cropping of double-season rice is comparatively more labour-intensive than single-season rice (Fengbo et al., 2013). That said, corn may also see a level of labour intensity that is similar to rice in modern China as is highlighted by Hu and Yuan (2015) in their discussion and critique of Talhelm's “rice theory” of culture. For a sketch of traditional agricultural production in China, see Sutton and Anderson (2009, p. 278–294).

agricultural land under their control — and shifts in the relative importance of the rules by which these individuals form social ties.

That said, I stress that my intention is not to detail the emergence of cooperation — an impossible task here given the data at my disposal. Instead, I move to simply assess how three generative mechanisms conducive to the diverse cooperative dynamics necessary for paddy rice cultivation (e.g., tangible aid, information exchange, mutual investment, norm enforcement and the proliferation of trust) might vary in their relative power over friend choice with the scale of one's agricultural production.

As for what this variation might look like, here I start with the assumption that farmers who control different amounts of rice production land are likely to have distinct agricultural needs, face unique agricultural constraints — particularly around access to labour and the efficiency of land use (Brandt et al., 2002; Otsuka et al., 2016; Van den Berg et al., 2007) — and, arguably, ascribe dissimilar subjective meanings to their immediate agricultural environment (Stedman, 2003; van der Ploeg et al., 2014). Indeed, scholars have linked farm size in China to variation in agricultural practices (e.g., fertiliser use, shifts in cropping patterns to single-season rice from the comparatively labour-intensive double-season rice) and productivity (Brandt et al., 2002; Chen et al., 2010; Fan and Chan-Kang, 2004; Fengbo et al., 2013; Otsuka et al., 2016; Tan et al., 2008; Van den Berg et al., 2007; Wu et al., 2018), suggesting that the scale of rice production is substantively linked to farmers' behaviours.⁴ Along this line, growth in the extent of rice production should see the demands and potential economic rewards of paddy rice redouble (Van den Berg et al., 2007), making breakdowns in cooperation costlier to the smallholder responsible for a relatively larger farm and to the broader familial land-labour institution that he is charged with nurturing (van der Ploeg et al., 2014).⁵ This is especially so as other farmers in a larger-scale rice producer's village are presumably amongst his more likely sources of non-mechanical agricultural support due to the finite nature of family labour and historic weaknesses in the markets for both hired farm help and the rental of excess farmland that make it difficult to redress large land-to-labour ratios (Brandt et al., 2002; Fan and Chan-Kang, 2004; Otsuka et al., 2016; Van den Berg et al., 2007). Consequently, I expect rice producers with increasingly larger farms to assign a relatively higher priority to building social ties in a manner that is conducive to cooperation with their agricultural counterparts, resulting in the following three hypotheses:

Hypothesis 1. Farm size will be positively associated with the relative importance of direct reciprocity to a rice producer's friend choices.

Hypothesis 2. Farm size will be positively associated with the relative importance of transitivity to a rice producer's friend choices.

Hypothesis 3. Farm size will be positively associated with the relative importance of generalised exchange to a rice producer's friend choices.

Note that these hypotheses raise intriguing questions around the precise trade-offs between mechanisms. For example, if direct reciprocity were to take on a relatively greater role in the friend-choice process, would it be at the expense of transitivity, preferential attachment or homophily? Nevertheless, theorisation and empirical investigation of which dynamics may be expected to “give way” to reciprocity, transitivity and generalised exchange as their relative importance increases is left to future research.

Data summary

Data from the People's Republic of China were collected by Cai et al.

(2015; see also Cai, 2012) during the roll out of a rice-production insurance programme designed to bolster food security and resilience to weather-related shocks (e.g., floods, extreme temperatures) led by the People's Insurance Company of China (PICC). Partnering with the PICC, Cai et al. (2015) randomly selected villages included in the 2010 expansion of the programme for a field experiment designed to assess the impact of farmers' social networks on their decisions to adopt rice insurance. Fieldwork took place in 185 villages in Jiangxi province which is located in the southeast of China below the Yangtze river. The province is composed of an arable land area of approximately 28,000 km² with paddy rice production covering approximately 22,000 km² of this land and yielding approximately 23 of the 208 million tons of China's total rice yield in 2014 (Guo et al., 2018). Consequently, Jiangxi is a major “rice bowl” sitting at the core of the country's agricultural production. Note that the 185 villages are “natural villages” as opposed to “administrative villages” which, in the context of rural China, typically subsume several natural villages. In line with how Cai et al. (2015) collected their data, my analyses are conducted at the level of the natural village which may be expected to broadly correspond to the smallest production unit or team within the “commune” farming system that was dismantled in the wake of the Great Leap Forward (see Ho, 2001, p. 404–405 and Footnote 38 on collective ownership).

Data collection and exclusion criteria

Cai et al. (2015) invited all rice-producing households in each village to participate in their experiment, achieving a response rate of approximately ninety percent. Rice-producing households were identified using a household roster created with the aid of village leadership (Cai, 2012), where around eighty percent of households in each village were rice-producing (Personal communication with J. Cai, October 30, 2018). Attribute and relational data were elicited using a household survey administered to the heads of a total of 5335 households. More specifically, Cai et al. (2015) collected information on: farmers' experiences with natural disaster, their perceived probability of a future disaster, their aversion to risk, the percentage of their total income from rice production, their basic demographics (i.e., age, gender, education and household size) and, most importantly here, their amount of rice-production land and their friendships. To measure social ties, Cai et al. (2015) used a single name generator which asked respondents to “list five close friends, either within or outside the village, with whom they most frequently discuss rice production or financial issues” (Cai et al., 2015, p. 88).⁶ Unfortunately, monadic data on farmers' ethnicity, household income and religion and dyadic data on joint membership in agricultural cooperatives, labour exchange, kinship and geographic proximity within each village are not available.

Here I work with the version of Cai et al.'s (2015) dataset made available for public use. This version includes information detailing the networks of rice-producing household heads in 173 of the 185 natural villages (47 administrative villages; 4902 heads). From the 173 natural villages, I dropped two from my analysis as they have insufficient information. Specifically, networks in these two villages are virtually

⁶ As is standard amongst network analysts who adopt a sociological perspective and who employ multi-mechanistic models, I operationalise friendship as an asymmetric relation. Furthermore, I view friendship as a vehicle by which individuals invest emotional, informational and/or tangible resources in those they perceive to be a “friend” in the traditional (i.e., mutual) sense with the expectation that investments will be returned (see Block (2015) on asymmetric friendship in relation to social exchange theory). Accordingly, one individual may befriend another without the latter responding in kind at the risk of the tie being withdrawn. Although there is an ongoing debate about the true nature of one-sided friendship (Block, 2015), for my purposes, an analysis of asymmetric friends is preferable as I am fundamentally interested in how individuals select their social contacts and whether these choices are beholden to network ecology.

⁴ See also cognate work on variation in farm size and farmer cooperation in England (Gasson, 1977).

⁵ Although note economies of scale (Otsuka et al., 2016; Tan et al., 2008).

empty, with one having eight farmers and zero arcs (i.e., directed friendships) and the other having three farmers and three arcs. Additionally, another nine natural villages were dropped from my analysis due to issues around the convergence of the estimation algorithm for the Stochastic Actor-Oriented Model related to multicollinearity (see Duxbury, 2019) when using the preferred global model specification (detailed below). The remaining 162 networks have between 23 and 206 arcs, range in size from 8 to 55 household heads and, in total, are constituted by 4713 farmers resident in 47 administrative villages.

Data limitations

There are three important limitations to the Jiangxi data that must be discussed. First, networks only concern rice-producing household heads as opposed to all adults in each natural village. Therefore, they are most appropriately viewed as virtually complete networks for the sub-population of residents in each natural village who are both household heads and rice farmers. By extension, the “village networks” analysed here should not be conflated with their corresponding “village-wide networks” constituted by all adult residents of each natural village. And, somewhat relatedly, the portfolio of social contacts belonging to household heads should not be assumed to either fully or partially reflect the sets of social contacts held by other members of a head’s home (Wencélius et al., 2016), especially as the relationships elicited are presumably subject to respondents’ personal preferences.

Second, Cai et al. (2015) used a fixed-choice design to elicit friendships — i.e., the number of individuals or “alters” that could be reported by those who were administered surveys was artificially capped at five. Accordingly, out-degrees were constrained to be five or less during estimation of/simulation from all Stochastic Actor-Oriented Models in order to respect this aspect of the data-generating process (Ripley et al., 2019). Note that in a pilot study of two villages wherein farmers could name an unlimited number of friends, Cai et al. (2015) found that 82 percent of respondents nominated five alters and 14 percent of respondents nominated four alters (pilot sample size not reported). Nevertheless, because of the artificial truncation of out-degrees, the networks of rice-farming household heads in each of the natural villages may simply be incomplete such that results may not be representative of scenarios wherein more exhaustive data are collected — especially for villages with more rice farmers wherein we can generally expect the truncation of out-degrees to result in a less accurate picture of connectivity. Indeed, network size has a strong inverse relationship with connectivity in terms of density for the graphs analysed here (Pearson’s $r = -0.83$; $df = 160$; $t = -18.6$; $p < 0.01$).

And third, friendship is a multidimensional construct which Cai et al. (2015) arguably measured in a narrow manner by using just one name generator. Although use of a single sociometric task is a standard practice across the social sciences, the wording of the name generator used by Cai et al. (2015) unfortunately hybridises a question about friendship with a second question about sources of information such that the elicited social tie is most appropriately regarded as the intersection of the two — i.e., a rice-producing friend who frequently provides occupation-relevant information as opposed to a rice-producing friend who fulfils material and/or emotional needs.⁷ That said, the name generator is transparent in its emphasis on what a friend actually does for the respondent and “close friendships” amenable to the exchange of information around farming and finance should be especially germane to the cooperative dynamics vital to paddy rice production as

Chinese smallholders face agricultural information gaps (Cui et al., 2018; Fan and Chan-Kang, 2004; Wu et al., 2018; Zhang et al., 2013, 2016) they presumably hope to fill. In fact, Cai et al.’s (2015) experiment on the adoption of weather insurance empirically demonstrates that the friendships analysed here are amenable to the diffusion of information related to rice production. Nevertheless, the dynamics around tie formation that I have inferred may differ entirely for friendship measured in a more exhaustive manner.

Despite these shortcomings, I stress that the data from Jiangxi are not without unique benefit. Specifically, they present an unusual opportunity to analyse a substantial number of face-to-face, annotated networks measured in the same fashion outside of educational institutions in societies that are Western, educated, industrialised, rich and democratic — i.e., WEIRD (Henrich et al., 2010). As the acronym suggests, WEIRD societies are globally atypical settings which currently typify empirical applications of multi-mechanistic models across the social sciences (Block, 2015; Heidler et al., 2014; Indlekofer and Brandes, 2014; Kretschmer et al., 2018; Leszczensky and Pink, 2015; McFarland et al., 2014; Raabe et al., 2019; Schaefer et al., 2010; Smith et al., 2016; Wimmer and Lewis, 2010). As a result, WEIRD networks underpin much of our more sociological knowledge about how humans establish their social ties, where the fitting of multi-mechanistic models to non-WEIRD data, especially data concerning adults, is rare (see Matous and Wang, 2019; Nolin, 2010 and Power, 2017 for laudable examples). This led me to select a case to study that is in stark contrast to those typically found in the literature on social networks.

It is also worth mentioning that although it has become fashionable to focus on large-scale networks defined by ephemeral exchange via social media or mobile phones, more traditional data enable assessment of the establishment of social ties in relatively closed systems wherein individuals depend on one another for vital support. Accordingly, the offline intra-village networks that I rely on here are ideal for an investigation of mechanisms and motivations as it is somewhat safe to assume the existence of ties that require investment on the part of those who establish and maintain them compared to, for example, “friendship” on Facebook or “population-scale” networks measured using text messages or mobile phone calls. Similarly, it is relatively safe to assume that members of small-to-medium-sized groups of rice producers will generally be aware of their counterparts and the social dynamics of the wider agricultural community in their natural village, especially given the history of agricultural production, land use and tensions around land ownership at this level (Ho, 2001). This is critical as the multi-mechanistic models I use here are built on the simplifying assumption that actors have “complete information” such that they can meaningfully decide on one friendship over another.

Analytical strategy and methods

Part I: meta-analysis of Stochastic Actor-Oriented Models

To test my hypotheses, I carried out an empirical analysis in three parts — the first of which consisted of fitting 162 identically-specified Stochastic Actor-Oriented Models (one for each village’s network of rice farmers). Briefly, the SAOM is a continuous-time multi-mechanistic model that simulates the process of network evolution in accordance with one or more empirical observations or “snapshots” of a network x (Block et al., 2019; Snijders, 1996, 2001; Snijders and Steglich, 2015). Emphasising the deliberation underpinning an actor’s choice of social contacts, SAOMs model “change” as asynchronous and sequential “mini-steps” whereby just one actor has the opportunity to alter just one of her outgoing ties (i.e., creating or “sending” a tie if it does not exist or dropping a tie if it exists), where “no change” is an option and the opportunity to make a change is governed by a rate function λ (here, fixed at a constant). Using multinomial logistic regression, actors’ decisions are modelled within a discrete-choice framework such that they are assumed to myopically modify their ties in order to maximise

⁷ Also note that some network analysts have rightfully expressed concerns about the validity of sociometric data obtained by simply asking respondents to list their “(close) friends” or “confidants” or with whom they discuss “important matters” — approaches which are reflective of an understanding of friendship in Western societies that may not be appropriate in a more global context (Hruschka, 2010; Shakya et al., 2017).

the value of a utility function that may generally be regarded as encoding the “attractiveness” of tie changes (Snijders, 1996, 2001; Ripley et al., 2019). Formally, the utility or “evaluation” function driving the friend choices of actor i (i.e., Which of the $N-1$ other actors j will i modify their relation with?) given the state of the network x takes the form of the linear predictor

$$\sum_{k=1}^L \hat{\beta}_k s_{k,i}(x) + U_i(x, j). \quad (1)$$

Here, $U_i(x, j)$ is a Gumbel-distributed random variable used to capture uncertainty in actor's decision-making in relation to their tie variable with some actor j (Snijders, 1996, 2001; Ripley et al., 2019, p. 177). Moreover, are the statistics for the L researcher-specified effects k that individually or collectively reflect various mechanisms of interest at the monadic, dyadic and triadic levels — each calculated from the perspective of/centred on actor i given the state of the network x . Furthermore, $\hat{\beta}$ are the estimated parameters (conditional log odds ratios) which weight these effects. Positive estimates $\hat{\beta}_k$ suggest a dynamic tendency for actors to establish ties to others in a such a way that leads to more instances of a given statistic (e.g., an increase in reciprocated dyads), with the converse for negative $\hat{\beta}_k$. Note that I have used the cross-sectional SAOM whereby the network x is assumed to be in short-term dynamic equilibrium (Block et al., 2019; Snijders and Steglich, 2015). See Appendix A for additional details on the cross-sectional SAOM and the fixing of λ .

Results from the individual SAOMs were summarised using a Frequentist meta-analysis procedure under the assumption that the 162 village networks are random samples from a larger population of networks. In turn, it is assumed that the estimated SAOM parameters are themselves random samples from a distribution of possible parameters for this same population of networks (An, 2015; Ripley et al., 2019).⁸ More specifically, I performed a multivariate meta-regression with random (i.e., network/SAOM-specific) effects (Gasparrini et al., 2012). Importantly, this style of meta-analysis does not assume that the estimated parameters for a given SAOM are independent — a dubious assumption as network statistics are typically functions of one another such that we can reasonably expect their associated parameters to be highly correlated.

Formally, let $\hat{\beta}_i$ represent a set of estimated parameters corresponding to β — the set of L effects used to specify the aforementioned utility function for the i th network/SAOM — and let S_i represent the $L \times L$ covariance matrix for these estimated parameters. For the meta-regression, $\hat{\beta}_i$ is assumed to be sampled with error from a multivariate normal distribution of dimension L centred on $X_i\theta$ or

$$\hat{\beta}_i \sim \text{Normal}_L(X_i\theta, S_i + \psi). \quad (2)$$

Here X_i is a $L \times Lp$ block-diagonal design matrix for p study-level “meta-predictors” associated with each of the i networks/SAOMs and θ is a Lp -length vector of unknown coefficients to be estimated describing the association between the magnitude of each of the L SAOM parameters and the p meta-predictors — one of which is the intercept indicating the population mean of each of the L SAOM parameters when the other meta-predictors are zero. The marginal covariance matrix Σ_i is given by the sum of the “within-study” (i.e., SAOM-specific) covariance matrix S_i and the “between-study” covariance matrix ψ , where the former is known and the latter, which is defined by $L(L+1)/2$ covariance parameters, is estimated. Note that the only other meta-predictor I use outside of the intercept is network size given its implications for: (i) the structural properties of graphs at both the macro and micro-levels (Anderson et al., 1999; Faust, 2007); and thus (ii) the estimated parameters of multi-mechanistic models (Faust, 2007; Krivitsky et al., 2011; McFarland et al., 2014).

Under the assumption that variability around the population-mean SAOM parameters is due to both between-network heterogeneity and uncertainty in the estimated parameters underlying the meta-analysis, the meta-regression adjusted for network size is used to “update” each of the SAOM-specific parameters in $\hat{\beta}_i$ in an attempt to improve their quality. Following Gasparrini et al. (2012; see also Gasparrini, 2018), these updated parameters $\hat{\beta}_{b(i)}$ are best linear unbiased predictions of $\hat{\beta}_i$ and they are derived by summing: (i) the expected mean SAOM parameters from the fixed (i.e., non-varying) part of the meta-regression for a given set of meta-predictor values (here, the intercept and network size); and (ii) the network/SAOM-specific deviations from the expected mean SAOM parameters, where these deviations are weighted by the relative size of the components of Σ_i (i.e., the within-study and between-study covariance matrices S_i and ψ). More formally,

$$\hat{\beta}_{b(i)} = X_i\hat{\theta} + \hat{\psi}\hat{\Sigma}_i^{-1}(\hat{\beta}_i - X_i\hat{\theta}), \quad (3)$$

where $\hat{\Sigma}_i^{-1}$ is the inverse of $\hat{\Sigma}_i = S_i + \hat{\psi}$ and $\hat{\theta}$ and $\hat{\psi}$ are, respectively, the estimated meta-regression coefficients and the estimated between-study covariance matrix. With respect to interpretation, the updated SAOM parameters $\hat{\beta}_{b(i)}$ may be regarded as versions of $\hat{\beta}_i$ that are “shrunk” towards their expected population means in an attempt to improve their quality by combining information about the independently-estimated SAOM parameters across the 162 networks based on their assumed adherence to a multivariate normal distribution (Gasparrini et al., 2012; Gasparrini, 2018). See also empirical Bayes meta-analysis and conditional shrinkage estimators (Raudenbush, 2009; Raudenbush and Bryk, 1985).

Individual SAOMs were estimated using the R package “RSiena” (Ripley et al., 2019) whereas I relied on the R package “mvmeta” (Gasparrini et al., 2012; Gasparrini, 2018) to carry out the meta-regressions using restricted maximum likelihood estimation. See Gasparrini et al. (2012) and Gasparrini (2018) for additional formalism around multivariate meta-regression and An (2015) and Ripley et al (2019) for general discussions on the meta-analysis of SAOMs.

Part II: The Indlekofer-Brandes measure of relative importance

For part two of my analysis, I measured the relative influence of the effects in each of the 162 SAOMs for each of the 4713 farmers. Like ERGMs (Amati et al., 2018), directly comparing effects in SAOMs is bedevilled by a number of challenges. These include: parameter estimates that are unstandardised (making naked comparisons of effect sizes highly suspect); correlations between effects associated with statistics that are functions of/nested within one another (e.g., two-paths $[i \rightarrow k \rightarrow j]$, in-two-stars $[k \rightarrow j \leftarrow i]$, out-two-stars $[k \leftarrow i \rightarrow j]$, and transitive triplets $[i \rightarrow k \rightarrow j \leftarrow i]$); the multiple possible outcomes for any one tie change; and, potentially, the longitudinal nature of the analysed data. These issues are adroitly addressed by Indlekofer and

⁸ Although meta-analyses of parameter estimates from multiple SAOMs are regularly performed, there is perhaps some question as to the validity of this practice following Norton et al.'s (2018) recent recommendations against synthesising odds ratios from logit models fit to different data due to potential differences in the scaling factors of the estimated coefficients; where the concern here is the standard deviation of the unobserved component of the utility/evaluation function (see Train, 2009, pp. 23–25, 34–36 and 40–42 on the scale of utility functions). Nevertheless, I have carried out a meta-analysis to ensure that my analytical strategy is consistent with previously published analyses of medium to large samples of networks using SAOMs and ERGMs (e.g., see An, 2015; Block, 2015; Kretschmer et al., 2018; Lubbers and Snijders, 2007; Smith et al., 2016; Snijders and Baerveldt, 2003 and McFarland et al., 2014 as well as the recommendations of Ripley et al., 2019). As with other core aspects of statistical modelling that are better understood in the context of traditional regression models (e.g., multicollinearity (Duxbury, 2019)), more research and practical guidance is needed around the comparability of odds ratios from multi-mechanistic models fit to binary network data in light of concerns raised by Norton and colleagues (Norton and Dowd, 2017; Norton et al., 2018). See also Kuha and Mills (2018) for a counter-argument via-à-vis traditional logit models.

Brandes (2014) who recently proposed a measure for the relative importance of the estimated effects in SAOMs that accounts for: the complete model specification and possible correlations between effects; the size of parameter estimates; and the analysed network data. As a result, their measure can be used to compare the relative importance of effects within and across SAOMs fit to the same or different data.

In line with the foundational sociological assumption of the SAOM (i.e., individuals have agency and thus control their outgoing ties), Indlekofer and Brandes (2014) define the importance of an effect as its influence on actors' choices throughout the simulated tie changes underpinning the SAOM estimation procedure. More formally, for some individual i that has the opportunity to amend her portfolio of outgoing relations by changing a single tie variable and all other individuals h' in network x , let $\hat{\beta}_b$ represent a set of SAOM-specific parameter estimates that have been updated/shrunk in accordance with the procedure described in the preceding section and that correspond to β — the set of L effects used to specify the SAOM's aforementioned utility function. Furthermore, let π_i represent the probability distribution implied by $\hat{\beta}_b$ that assigns to each of i 's potential alters $j \in h'$ a value $\pi_i(j)$ — the probability that i will change the value of her tie variable with j by creating a tie if it does not exist or dropping a tie if it exists; all depending on the state of the network x — such that $\sum_{j=1}^{h'} \pi_i(j) = 1$. For some effect $k \in \beta$, its importance is defined to be the sum of the absolute values of the pointwise differences between π_i and $\pi_i^{(-k)}$, the latter of which is implied by $\hat{\beta}_b$ when only $s_{k,i}(x)$ is set to zero. As such, k 's importance is its direct contribution to π_i according to $\hat{\beta}_b$ given x and thus its impact on i 's inferred relational decisions.⁹ In this respect, construction of k 's importance is akin to assessing the amount of change in a dependent variable (here, modulation of π_i) associated with a change in some independent variable (i.e., setting $s_{k,i}(x)$ equal to zero). Moreover, by setting $s_{k,i}(x)$ equal to zero — as opposed to excluding β_k from the model specification — the complete model specification, the magnitude of parameter estimates and the correlations between effects (as manifest in the set of estimated parameters $\hat{\beta}$ and, by extension, their shrunk versions $\hat{\beta}_b$) are respected, which would not be the case were β_k to be removed and the SAOMs re-estimated (Indlekofer and Brandes, 2014, p. 300–301).¹⁰ Last, the expected relative importance of k for the relational decisions of individual i in departure from the observed state of the network x , or $I_k(x, i)$, is relative as it is normalised using the sum of the expected importance of each effect in β to reflect its proportional contribution to π_i such that $\sum_{k=1}^L I_k(x, i) = 1$.

Part III: Multilevel beta regression models of actor-level shares of influence

The third and final component of my analysis consisted of using multilevel models to summarise the relationship between

⁹ This is of course distinct from: (i) the magnitude of the estimated SAOM parameter $\hat{\beta}_k$ associated with an effect; and (ii) the count for the network statistic $s_{k,i}(x)$ associated with an effect. Relative to my hypotheses, assessment of the former implies a meta-regression of the estimated SAOM parameters for direct reciprocity, transitivity and generalised exchange wherein the average farm size for each natural village is used as a meta-predictor (i.e., the strategy of McFarland et al. (2014)) or, possibly, village-specific SAOMs employing interactions between farm size and the effects used to capture direct reciprocity, transitivity and generalised exchange. On the other hand, assessment of the latter implies modelling the relationship between farm size and counts of the number of reciprocated dyads, transitive triads and cyclic triads centred on some focal actor i in the observed network x .

¹⁰ A major shortcoming of the measure of Indlekofer and Brandes (2014) is that it is not designed to account for the uncertainty of the estimated parameters $\hat{\beta}$ and, by extension, their updated versions $\hat{\beta}_b$. More specifically, their measure only considers the parameter estimates (irrespective of their standard errors or “statistical significance”), the model specification and the state of the observed network x .

$I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ and the amount of land a farmer has that is devoted to rice production. Recall that the multivariate meta-regression with random effects is only used to summarise the dynamics driving the friend choices of the Jiangxi farmers. More specifically, it yields the expected average SAOM parameter estimate associated with each effect across the 162 networks. Accordingly, it is conceivable that farmers in some villages will exhibit tendencies *against* direct reciprocity, transitivity and generalised exchange despite any general tendency for these dynamics indicated by the estimated population parameters $\hat{\theta}_{\text{Direct Reciprocity (Intercept)}}$, $\hat{\theta}_{\text{Transitivity (Intercept)}}$ and $\hat{\theta}_{\text{Generalised Exchange (Intercept)}}$. By extension, farmers in villages where the updated/shrunk SAOM parameters $\hat{\beta}_{b\text{Direct Reciprocity}}$, $\hat{\beta}_{b\text{Transitivity}}$ and $\hat{\beta}_{b\text{Generalised Exchange}}$ are negative may see values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ that systematically differ from those associated with farmers in villages where $\hat{\beta}_{b\text{Direct Reciprocity}}$, $\hat{\beta}_{b\text{Transitivity}}$ and $\hat{\beta}_{b\text{Generalised Exchange}}$ are positive as the relative shares of influence in the two scenarios relate to opposite dynamics. As a result, multilevel models are fit using information only from those villages for which $\hat{\beta}_{b\text{Direct Reciprocity}}$, $\hat{\beta}_{b\text{Transitivity}}$ and $\hat{\beta}_{b\text{Generalised Exchange}}$ are positive in line with the implicit assumptions underlying my three hypotheses; namely that farmers will unambiguously exhibit tendencies to respond in kind (direct reciprocity), form cohesive local social groups (transitivity) and establish friendships in an egalitarian manner (generalised exchange).

As for the specifics of the multilevel models themselves, here the measured actor-level shares of influence $I_k(x, i)$ assume values within the mixed interval [0,1]. Accordingly, I relied on a mixture of the beta distribution, a flexible distribution for bounded response variables (Smithson and Verkuilen, 2006), and a degenerate distribution for true zeros, which the standard beta distribution cannot accommodate, in order to fit zero-inflated beta regression models (Ospina and Ferrari, 2010, 2012). That is, the set of actor-level shares of influence for some effect/mechanism k for the N farmers with complete data or $\{I_k(x, 1), \dots, I_k(x, N)\}$ is assumed to adhere to a mixed continuous-discrete distribution capable of accommodating excess zeros — i.e., the zero-inflated beta distribution.

The zero-inflated beta regression model itself is piecewise in that it is comprised of a set of sub-models used to separately estimate the probability of a response being 0 and where the response falls between 0 and 1. Specifically, the response y (i.e., $\{I_k(x, 1), \dots, I_k(x, N)\}$) is modelled as a function of three distributional parameters — μ , the mean of the beta distribution (i.e., $0 < I_k(x, i) < 1$), ϕ , the precision of the beta distribution, and π_0 , the probability of the response being 0 (i.e., $\Pr(I_k(x, i) = 0)$). Formally,

$$f(y; \pi_0, \mu, \phi) = \begin{cases} \pi_0, & \text{if } y = 0 \\ (1 - \pi_0)f(y|\mu, \phi), & \text{if } y \in (0, 1), \end{cases} \quad (4)$$

where $f(y|\mu, \phi)$ is the density function of the beta distribution or

$$f(y; \mu, \phi) = \text{Beta}(\mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, \quad (5)$$

with $\Gamma(\cdot) = (\cdot-1)!$ being the gamma function, $0 < \mu < 1$ and $\phi > 0$.

The overall mean of the response $E(y)$ is given by $(1 - \pi_0)\mu$ whereas $E(y|y \in (0,1))$ is simply μ (Ospina and Ferrari, 2010, 2012). For a fixed value of μ , the larger the value of the precision parameter ϕ , the smaller the variance of the response y — where ϕ itself can be modelled as a function of a set of correlates that may or may not be the same as those used in the sub-models for μ and π_0 in order to relax the assumption that precision is constant (Smithson and Verkuilen, 2006).

The relationships between μ , ϕ and π_0 and their linear predictors inclusive of the correlates of interest and, possibly, random effects capturing the hierarchical structure of the data are modelled using logit (μ and ϕ) and log (π_0) link functions — standard choices for beta regression (Liu and Eugenio, 2016; Smithson and Verkuilen, 2006).

Formally,

$$\begin{aligned}\text{logit}(\mu) &= \mathbf{X}\beta_1 + \mathbf{Z}\gamma_1 \\ \ln(\phi) &= \mathbf{X}\beta_2 + \mathbf{Z}\gamma_2 \\ \text{logit}(\pi_0) &= \mathbf{X}\beta_3 + \mathbf{Z}\gamma_3\end{aligned}\quad (6)$$

Ignoring the index for each sub-model, \mathbf{X} is the design matrix corresponding to the fixed (i.e., population-level, non-varying or “common”) parameters β and \mathbf{Z} is the design matrix corresponding to the random (i.e., group-level or “varying”) parameters γ which capture group-specific deviations from their corresponding fixed effect parameters representing population averages. The group-level parameters γ are assumed to come from a multivariate normal distribution with a mean equal to zero and an unknown covariance matrix Σ to be estimated.

To estimate the multilevel zero-inflated beta regression models, I took a Bayesian approach with the aid of the R package “brms” or Bayesian Regression Models using “Stan” (Bürkner, 2018, 2019), where Stan is a probabilistic programming language. I relied on Bayesian estimation as the Frequentist estimation of beta regression models may suffer a number of issues relative to my need to model diverse clustered outcomes with zero-inflation (e.g., large biases, poor coverage probability, poor handling of outliers; see Liu and Eugenio (2016) and Niekerk et al. (2018)). A complete rationale behind my choice of prior distributions appears in Appendix B. However, here it suffices to say that fixed effects in all components of the beta regression models for $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ appearing below were given Student- t prior distributions with seven degrees of freedom, a mean of zero and a scale of 2.5, save the intercepts which were given a scale of 10. Similarly, the standard deviation of the random effects was given the same prior distribution regardless of sub-model, namely a truncated or “half” Student- t prior with three degrees of freedom and a scale of ten.

With respect to estimation settings, all beta regression models were run using twelve parallel chains and the No-U-Turn Sampler — a type of Markov Chain Monte Carlo algorithm based on Hamiltonian Monte Carlo (Hoffman and Gelman, 2014). Each chain is comprised of 7000 iterations, where the first 1000 iterations were devoted to “warm-up” resulting in a total of 72,000 post warm-up iterations (i.e., the posterior sample size) for each parameter in each model. To ensure the stability of the limits of the 95% Highest Density Intervals (HDI) for all parameters in each model, the length of chains and the warm-up period were chosen through trial and error to achieve an effective sample size of at least 10,000 (Kruschke, 2015).¹¹ The convergence of chains was diagnosed using \hat{R} — the potential scale reduction factor for rank-normalised split chains (Vehtari et al., 2019). Values of \hat{R} less than 1.01 are generally regarded as indicative of convergence of the chains. As with the effective sample size, the length of chains and the warm-up period were chosen through trial and error to ensure \hat{R} was less than 1.01 for all parameters in each model.

¹¹ The 95% Highest Density Interval (HDI) indicates the values bounding 95% of a parameter’s posterior distribution, where all values inside this interval have a higher probability than those outside. The effective sample size (ESS) is the posterior sample size in absence of correlation between values of each chain (i.e., an estimate of the amount of independent information in a chain). In light of findings by Vehtari et al. (2019), here I report the minimum standard ESS, the minimum “Bulk” ESS and the minimum “Tail” ESS across all parameters in each multilevel model, where the latter two effective sample sizes are derived using rank-normalised values of the chains. Respectively, Bulk ESS and Tail ESS summarise the amount of independent information in the centre and the tails of a posterior distribution, with the former approximating the standard ESS.

Model specifications

Rationale for the global SAOM specification β

The following three criteria were used to construct the global SAOM specification β . First, the specification needed to be parsimonious as the relative importance of effects may generally be expected to shrink as the size of a model specification grows. Second, the specification was required to perform well in terms of goodness-of-fit across the large majority of the analysed networks. And third, the specification needed to be amenable to the convergence of the SAOM estimation algorithm for as many of the 173 networks as possible — a vexing issue when fitting SAOMs to small graphs due to the risk of model instability brought about by multicollinearity (Duxbury, 2019). Given these criteria, I arrived at the following global specification which was found to be suit the convergence of SAOMs for 162 of the village networks.

First to consider is the endogenous component of the model which needed to minimally include effects reflecting direct reciprocity, transitivity and generalised exchange. Capturing the former was straightforward. However, the precise manner in which the latter two mechanisms are accounted for deserves some comment.

When analysing directed networks with SAOMs and when not using geometrically-weighted counts of triads, “Transitivity” (i.e., a tendency towards establishing balanced triads; Krackhardt and Handcock, 2007) is typically captured with the effect *Transitive Triplets*. This effect jointly accounts for: (i) the closure of two-paths (i.e., $[i \rightarrow k \rightarrow j]$ leading to the tie $i \rightarrow j$); and (ii) the closure of in-two-stars (i.e., $[i \rightarrow j \leftarrow k]$ leading to the tie $i \rightarrow k$). Respectively, these dynamics reflect transitive closure and closure induced by approximate structural equivalence based on outgoing ties. I used a single effect to capture both of these dynamics as my aim is simply to model any preference rice producers have for the formation of transitive groups.¹²

Moreover, I used the *Three Cycles* effect to account for the establishment of cycles of friendship of length three through the closure of the two-path $[i \leftarrow k \leftarrow j]$ with the tie $i \rightarrow j$. This dynamic represents the simplest form of chain-generalised exchange wherein resources — here, the social and informational returns to friendship — flow along the same pattern of ties and circle back to the initiator (Molm, 2010; Simpson et al., 2017). Crucially, the *Three Cycles* effect is accompanied by the *Transitive Reciprocated Triplets* effect which captures whether actors respond in kind to ties embedded inside of transitive triads. As Block (2015) has shown, the *Transitive Reciprocated Triplets* effect is a necessary addition to SAOM specifications for the unbiased estimation of the parameter for the *Three Cycles* effect.

To fully represent tendencies for or against transitive closure and cyclic closure when using the *Transitive Triplets* and *Three Cycles* effects, models must also include effects that account for the creation of relevant two-paths and two-stars. Most critical is the *Out-degree Popularity* effect which I used to capture any tendency for actors who send many friendships to receive a one-sided friendship from actor i (i.e.,

¹² Recall that network statistics in the SAOM are calculated from the perspective of/centred on the focal actor i as opposed to being calculated in the “tie-oriented” way of the ERGM wherein the positions of actors in sub-structures are irrelevant (Block et al., 2019). Consequently, there is no statistic for the closure of the out-two-star $[k \leftarrow i \rightarrow j]$ as this action would be taken by k or j and not i . Of course, one is able to account for the closure of the out-two-star $[i \leftarrow k \rightarrow j]$ when using the SAOM. However, only the ties controlled by i that are capable of establishing a transitive triad (i.e., $i \rightarrow j$ and $i \rightarrow k$ in the triad $[i \rightarrow k \rightarrow j \leftarrow i]$) contribute to the change statistic (i.e., the change in the value of a network statistic associated with toggling a single tie variable) used to calculate the probabilities of transitioning between adjacent network states when using the *Transitive Triplets* effect (see Block et al., 2019, p. 222–229 on asymmetric transition dependence and elementary effects). Along this line, it is generally understood that the *Transitive Triplets* effects can adequately account for transitivity when using the SAOM (Ripley et al., 2019).

establishment of the two-path $[i \rightarrow k \rightarrow j]$ in the case of *Transitive Triplets* or $[i \rightarrow j \rightarrow k]$ in the case of *Three Cycles*). Furthermore, the formation of in-two-stars (i.e., $[i \rightarrow j \leftarrow k]$), the other basis of the *Transitive Triplets* effect, is accounted for with the *In-degree Popularity* effect which may be interpreted here as capturing the formation of in-stars over and above those embedded in transitive triplets (i.e., more global returns to popularity in the form of *Preferential Attachment/The Matthew Effect* [Schaefer et al., 2010]). Last, the *Out-degree Activity* effect is used to account for any tendency for actors with many outgoing ties to send more ties (i.e., the creation of the out-two-star $[k \leftarrow i \rightarrow j]$).

The primary purpose of the exogenous component of the model specification is to account for trait-based assortative mixing as it can confound evidence of transitivity (Goodreau et al., 2009). Accordingly, the global SAOM specification includes effects for homophily around age, education, rice production area, the percentage of a farmer's income from rice production and household size.¹³

As there is a need to keep the global model specification parsimonious, particularly for the smallest villages with around 10 farmers, models were adjusted for homophily in a fashion that is interpretable in the absence of complementary covariate-related activity (“ego”) and covariate-related popularity (“alter”) effects. Specifically, I used a series of covariate-related similarity effects designed to capture whether actors tend to befriend individuals with similar covariate values. This is in contrast to using interaction effects created by multiplying the covariate values of ego and alter. That said, I included covariate-related activity and covariate-related popularity effects for age to make basic adjustments to the SAOMs for variation in the degree to which rice producers send and receive ties in line with their stage in the life course and their farming experience (for which age has been used as a proxy [Chen et al., 2010]). As mentioned above, the sample of farmers is virtually homogenous with respect to gender such that networks have very few or no women members. Consequently, homophily, ego and alter effects for gender were not included in the model specifications in order to keep them identical across the 162 villages.¹⁴

Formulae used to calculate the statistics $s_{k,i}(x)$ for each effect in the global SAOM specification, along with short descriptions to aid reader interpretation, are found Table 1. Descriptive statistics for the monadic covariates for the 4713 farmers appear in Table 2. As SAOMs treat missing attribute data as non-informative (Ripley et al., 2019), all 4713 rice producers are retained for the first two parts of my analysis despite

¹³ Cai et al. (2015) measured education as an ordinal variable (i.e., Illiterate, Primary, Secondary, High School, College). The inclusion of this variable in both the SAOMs and the beta regression models was done under the assumption that its categories are equidistant in order to treat it as a continuous. This is in contrast to discarding information contained in the variable related to its order by using a set of binary “dummy” variables for each education level. Neither strategy is ideal. However, my preference is for the former as it allowed me to use more parsimonious model specifications.

¹⁴ Cultural sociologists have argued that worldviews are logically “prior to” network structure such that they shape how individuals choose their social contacts (Vaisey and Lizardo, 2010). Accordingly, there is some justification for including covariate-related similarity effects for the disaster-perception and risk-aversion variables measured by Cai et al. (2015) in the SAOM specification. Nevertheless, these two variables are omitted from the models in line with a more traditional understanding of networks, on the one hand, and perceptions/opinions/values, on the other hand, as mutually constitutive and thus likely to co-evolve. Given this perspective and my reliance on cross-sectional data, inclusion of the disaster-perception and risk-aversion variables into the model specification would induce serious concerns around reverse causality as I cannot untangle social selection from social influence. And if one were to have to choose *a priori*, social influence seems most likely in light of: (i) Cai et al.' (2015) experiment indicating that the Jiangxi farmers' social ties are instrumental in the diffusion of information about agricultural insurance; and (ii) research on the social and demographic determinants of risk perception amongst smallholders which include access to information in personal networks (Cullen et al., 2018).

the small amount of incomplete information summarised in Table 2.

Note that I rely on versions of *Out-degree Popularity*, *In-degree Popularity* and *Out-degree Activity* that are centred using the mean in/out-degree for each network as this was found to greatly aid convergence of the SAOM estimation algorithm. Similarly, for each of the 162 SAOMs, age, education, rice production area, the percentage of a farmer's income from rice production and household size are centred at their within-network means and then divided by their within-network standard deviations to convert to Z-scores which was also found to aid convergence. As with standard treatments of changes in income, relative/proportional increases in rice production area (here measured in *mu*), as opposed to absolute changes, are assumed to capture salient shifts in proximal physical environment and are thus most relevant to the behaviour of farmers (e.g., increases of 2 *mu* to 4 *mu* and 50 *mu* to 52 *mu* as respective increases of 50% and 4% versus a constant increase of simply 2 *mu*). Accordingly, I first log transform (natural) rice production area and then convert this transformed variable to a Z-score.

Rationale for the specification of the beta regression models

In specifying the mean, precision and zero-inflated components of the beta regression models, I used identical sets of covariates due to the absence of any theoretical rationale for divergence in the composition of the linear predictors. Specifically, I reproduced the exogenous component of the global SAOM specification by adjusting the mean, precision and zero-inflated sub-models for age, education, the percentage of a farmer's income from rice production and household size alongside network size and gender. With regard to the non-land determinants of how farmers might go about building social ties with their agricultural counterparts, the logic behind my use of age, education and gender is self-evident. However, it is useful to summarise why network size, household size and the percentage of a farmer's income from rice farming are necessary covariates for appropriately testing the hypothesised relationships between log rice production area (i.e. farm size) and the relative importance of *Direct Reciprocity*, *Transitivity* and *Generalised Exchange*.

First, it has long been known that structure (i.e., the arrangement of ties) varies in accordance with network size (B. S. Anderson et al., 1999). Indeed, the basic features of social interaction in diverse scenarios has been linked to the number of members of a group (Mayhew and Levinger, 1976). As actor-level shares of influence $I_k(x, i)$ strongly depend on the state of the analysed network data, it seems reasonable to conclude that they may also be a function of network size to the extent that it determines the opportunities network members have to establish ties under various constraints. For example, consider how the number of ties capable of being reciprocated or the number of open triads one can close might scale with group size (Anderson et al., 1999; Faust, 2007).

As for household size, this variable is used to capture access to relatively unqualified labour from kin (Brandt et al., 2002; Lowder et al., 2016; van der Ploeg et al., 2014). Consequently, it provides some protection against the confounding of the relationship between farm size and relative importance as household size could plausibly impact: (i) how land is allocated to a farmer by local officials (Brandt et al., 2002; Chen et al., 2010; Deininger et al., 2014; Tan et al., 2008; Wu et al., 2018); and (ii) the extent to which a farmer may need to rely on his agricultural counterparts and thus how much he prioritises building farming-relevant friendships in a manner conducive to cooperation. Similarly, the percentage of a farmer's total income from rice production — i.e., an indicator of his dependence on rice versus alternative sources of revenue — also provides some protection against the confounding of the relationship between farm size and relative importance. This is because reliance on non-farm work also stands to impact how land is allocated by local officials (Brandt et al., 2002; Deininger et al., 2014) and, of course, the extent to which a rice producer may need to rely on agricultural counterparts.

Table 1
Global SAOM Specification β and Formulae for Effects.

Level	Effect	Network Dynamic (Do farmers tend to...)	$s_{k,i}(x) = \dots$
Triadic	Transitive Triplets (Transitivity)	...befriend friends-of-friends?	$\sum_{j,h} x_{ij} x_{jh} x_{hj}$
	Transitive Reciprocated Triplets	...respond in kind when befriended by an indirect contact?	$\sum_{j,h} x_{ij} x_{ji} x_{jh} x_{hj}$
	Three Cycles (Cyclic Closure/Generalised Exchange)	...choose as friends those who choose their friends as friends?	$\sum_{j,h} x_{ij} x_{jh} x_{hi}$
Dyadic	Direct Reciprocity	...respond in kind when befriended?	$\sum_j x_{ij} x_{ji}$
	Age Homophily (Covariate-Related Similarity)	...befriend those who are of a similar age?	$\sum_j x_{ij} (sim_{ij}^v - \hat{sim}^v)$
	Education Homophily (Covariate-Related Similarity)	...befriend those with a similar level of education?	
	Rice Production Area Homophily (Covariate-Related Similarity)	...befriend those with a similar amount of farmland?	
	% Income from Rice Homophily (Covariate-Related Similarity)	...befriend those with a similar level of rice dependency?	
	Household Size Homophily (Covariate-Related Similarity)	...befriend those with a similar number of coresidents?	
Monadic	In-degree Popularity (Preferential Attachment)	...befriend popular villagers?	$\sum_j x_{ij} (x_{+j} - d_{+Avg})$
	Out-degree Popularity	...befriend those who send many friendships?	$\sum_j x_{ij} (x_{j+} - d_{Avg+})$
	Out-degree Activity (Cumulative Activity)	...befriend others when sending many friendships?	$x_{i+} (x_{i+} - d_{Avg+})$
	Out-degree (Sociality)	...befriend others? (akin to an intercept)	x_{i+}
	Age (Covariate-Related Activity/Ego)	...be more sociable when older?	$v_i \sum_j x_{ij}$
	Age (Covariate-Related Popularity/Alter)	...be more popular when older?	$\sum_j x_{ij} v_j$

$x_{ij} = 1$ indicates the presence of a tie from actor i to actor j ; $x_{ij} = 0$ indicates the absence of a tie.

$x_{i+} = \sum_j x_{ij}$ = out-degree of actor i ; $x_{+i} = \sum_j x_{ji}$ = in-degree of actor i ; the same for x_{j+} and x_{+j} .

d_{Avg+} = Average out-degree across all network waves; d_{+Avg} = Average in-degree across all network waves.

v_i = the value of the monadic covariate of interest for actor i ; the same for v_j .

\hat{sim}^v is equal to the mean of all similarity scores for all dyads where $sim_{ij}^v = \frac{\Delta - |v_i - v_j|}{\Delta}$ and $\Delta = \max_{ij} |v_i - v_j|$ = the observed range of the covariate v .

Table 2
Global Descriptive Statistics.

Variable	Description	Mean	SD	Median	Min	Max	Missing	Levels (%N)
Age	Years of age	51.52	12.00	50	15	89	9	—
Gender	0 = Male 1 = Female	—	—	—	0	1	8	0 (91.49%) 1 (8.34%)
Education	0 = Illiteracy; 1 = Primary School 2 = Secondary School 3 = High School 4 = College	1.20	0.85	1	0	4	59	0 (21.11%) 1 (42.58%) 2 (29.28%) 3 (4.88%) 4 (0.59%)
Household Size	Number of coresidents in the farmer's home	4.91	2.13	5	1	21	9	—
Percent (%) Income from Rice Production	Share of a farmer's total income that is rice income	74.14	27.71	80	10	100	108	—
Rice Production Area (i.e., Farm Size)	Amount of land devoted to rice production in μ [traditional unit of land area in China] where 1 μ = 1/15 hectare	13.71	19.73	10	0.30	650	18	—
Network Size	Total number of surveyed rice farmers in the village/the size of the network	32.44	9.66	32	8	55	0	—
$I_{Direct\ Reciprocity}(x, i)$	The relative importance of direct reciprocity	0.078	0.060	0.069	0.00 (N = 468)	0.419	0	—
$I_{Transitive\ Closure}(x, i)$	The relative importance of transitivity	0.203	0.114	0.203	0.00 (N = 399)	0.736	0	—
$I_{Generalised\ Exchange}(x, i)$	The relative importance of generalised exchange	0.020	0.018	0.016	0.00 (N = 564)	0.126	19	—

$N_{Farmers} = 4713$.

In addition to the fixed effects, the mean and zero-inflated sub-models include random intercepts for natural village in order to account for the hierarchical nature of the data (i.e., farmers nested in villages/networks) and heterogeneity in the response due to village-specific social dynamics and unobserved conditions of rice production linked to geographic location (e.g., temperature, rainfall, soil quality, etc.) that may have implications for farmers' agricultural needs and thus how they go about establishing their social ties. That said, I excluded random intercepts from the precision sub-model and I did not include random slopes for rice production area in any of the sub-models. This is because ensuring a reasonable number of observations per estimated parameter is a concern due to my need to jointly estimate the three sub-models. More specifically, were I to have fit a beta regression model using data from all 162 natural villages and random intercepts in all sub-models, it would have required the estimation of 514 parameters —

i.e., the common intercept, 162 varying intercepts, the parameter summarising the variance of the random intercepts, and seven non-varying coefficients in each of the three sub-models plus the log of the posterior density for the entire model. Assuming that this hypothetical model was fit to data from all 4713 farmers, there would be roughly nine observations per parameter — where the use of random slopes for rice production area would see the addition of another 486 varying coefficients and result in roughly five observations per parameter. Consequently, the model specification favoured here reflects the minimal variation demanded by the structure of the data given my substantive interest (i.e., the varying intercepts for the mean and zero-inflated sub-models; 351 parameters in total) with the necessary assumption then being that the group-level variances for the estimated coefficients for the intercept in the precision sub-model, rice production area and the other covariates across the sub-models are zero. Note

Table 3
Multivariate Meta-Regression of Stochastic Actor-Oriented Models.

Meta-Predictor	Mechanism	Baseline Model				Extended Model			
		$\hat{\theta}$	$se_{\hat{\theta}}$	p	RAND SD	$\hat{\theta}$	$se_{\hat{\theta}}$	p	RAND SD
Intercept	Sociality	−1.632	0.030	0.000	0.269	−1.590	0.025	0.000	0.155
	Direct Reciprocity	0.733	0.029	0.000	0.194	0.691	0.028	0.000	0.133
	Transitivity	0.420	0.015	0.000	0.129	0.403	0.014	0.000	0.107
	Transitive Reciprocated Triplets	−0.171	0.026	0.000	0.140	−0.136	0.026	0.000	0.102
	Generalised Exchange	0.095	0.017	0.000	0.065	0.081	0.017	0.000	0.056
	Preferential Attachment	0.032	0.003	0.000	0.016	0.030	0.003	0.000	0.016
	Out-degree Popularity	−0.091	0.009	0.000	0.033	−0.085	0.009	0.000	0.035
	Out-degree Activity	0.118	0.011	0.000	0.098	0.118	0.011	0.000	0.094
	Age (Alter)	0.004	0.006	0.513	0.009	0.004	0.006	0.498	0.009
	Age (Ego)	0.005	0.008	0.495	0.014	0.005	0.009	0.577	0.013
	Age Homophily	0.440	0.027	0.000	0.070	0.424	0.028	0.000	0.060
	Education Homophily	0.041	0.021	0.045	0.016	0.041	0.021	0.054	0.014
	Rice Production Area/Farm Size Homophily	0.082	0.029	0.005	0.065	0.082	0.030	0.007	0.065
	% Income from Rice Production Homophily	0.039	0.020	0.046	0.031	0.028	0.021	0.186	0.039
	Household Size Homophily	0.033	0.027	0.228	0.022	0.031	0.029	0.280	0.021
Network Size*	Sociality	—	—	—	—	−0.228	0.024	0.000	—
	Direct Reciprocity	—	—	—	—	0.150	0.028	0.000	—
	Transitivity	—	—	—	—	0.080	0.014	0.000	—
	Transitive Reciprocated Triplets	—	—	—	—	−0.098	0.025	0.000	—
	Generalised Exchange	—	—	—	—	0.034	0.017	0.040	—
	Preferential Attachment	—	—	—	—	0.002	0.003	0.536	—
	Out-degree Popularity	—	—	—	—	−0.010	0.009	0.275	—
	Out-degree Activity	—	—	—	—	0.012	0.011	0.262	—
	Age (Alter)	—	—	—	—	−0.003	0.006	0.580	—
	Age (Ego)	—	—	—	—	0.004	0.008	0.620	—
	Age Homophily	—	—	—	—	0.068	0.028	0.015	—
	Education Homophily	—	—	—	—	−0.003	0.021	0.871	—
	Rice Production Area/Farm Size Homophily	—	—	—	—	0.025	0.030	0.407	—
	% Income from Rice Production Homophily	—	—	—	—	0.026	0.020	0.191	—
	Household Size Homophily	—	—	—	—	0.019	0.028	0.489	—
	Multivariate Cochran's Q Test								
					Q= 2654.95 (df= 2415; p< 0.001)		Q= 2411.76 (df= 2400; p= 0.429)		
	I^2				9%		1%		
	AIC				−1127.55		−1124.30		
	BIC				−345.97		−256.82		

RAND SD = Standard deviation of the random effects for each SAOM parameter reflecting village/network/SAOM-specific deviations from the intercept.

p = Two-sided.

$N_{SAOM\ Parameters}$ = 2430.

N_{SAOMs} = 162.

See Supplementary File 1 for parameters $\hat{\beta}$ and the $se_{\hat{\beta}}$ for each effect in the 162 underlying SAOMs.

* Z-score (Mean network size for the 162 villages = 29; SD = 9.9; Range = 8–55).

that random intercepts for administrative village were not used as the number of natural villages nested within administrative villages is typically small (Mean = 3.5), with some administrative villages subsuming just one or two natural villages in this particular sample (Range = 1–8).

Descriptive statistics for the variables used to fit the beta regression models are found in Table 2. Missing data are handled with listwise deletion following Pepinsky (2018), where farmers in villages for which $\hat{\beta}_{Direct\ Reciprocity}$, $\hat{\beta}_{Transitivity}$ and $\hat{\beta}_{Generalised\ Exchange}$ are negative are given missing values for $I_{Direct\ Reciprocity}(x, i)$, $I_{Transitivity}(x, i)$ and $I_{Generalised\ Exchange}(x, i)$. Before fitting the beta regression models, all explanatory variables were centred using their global means (i.e., across all farmers with non-missing data) and then divided by their global standard deviations to convert to Z-scores to help improve the efficiency of the Markov Chain Monte Carlo sampler following Kruschke (2015, p. 484–485). See Appendix C for a brief summary of weaknesses

of the model specification related to the measurement of rice production area.

Results

Results from the meta-analysis of the estimated parameters $\hat{\beta}$ from each of the 162 SAOMs appear in Table 3. On average, the assumptions underlying my three hypotheses hold. Specifically, the estimated population-mean parameters $\hat{\theta}$ summarising the inferred dynamics of the Jiangxi farmers as they choose their “close friends” clearly indicate that the rice producers tend to respond in kind when befriended (Positive $\hat{\theta}_{Direct\ Reciprocity(Intercept)}$), befriend friends-of-friends (Positive $\hat{\theta}_{Transitivity(Intercept)}$) and establish one-sided friendships in an egalitarian manner (Positive $\hat{\theta}_{Generalised\ Exchange(Intercept)}$). Critically, results from the baseline, intercept-only model hold when using the Z-score of network

Table 4
Bayesian Multilevel Beta Regression Models with Zero-inflation.

Sub-Model	Covariate	$I_{\text{Direct Reciprocity}}(x, i)$			$I_{\text{Transitivity}}(x, i)$			$I_{\text{Generalised Exchange}}(x, i)$		
		PMEAN	HDI-L	HDI-U	PMEAN	HDI-L	HDI-U	PMEAN	HDI-L	HDI-U
Mean	Intercept	−2.358	−2.381	−2.336	−1.299	−1.358	−1.240	−3.854	−3.927	−3.779
	Rice Production Area/Farm Size	0.076	0.050	0.102	0.003	−0.016	0.022	0.102	0.073	0.130
	Network Size	−0.040	−0.062	−0.018	−0.070	−0.125	−0.015	0.200	0.128	0.271
	Age	0.019	−0.006	0.045	−0.014	−0.032	0.005	0.030	0.003	0.057
	Female	−0.040	−0.125	0.041	−0.037	−0.102	0.026	−0.032	−0.128	0.060
	Education	0.051	0.027	0.074	0.003	−0.013	0.021	0.047	0.023	0.071
	% Income from Rice Production	−0.017	−0.042	0.008	−0.003	−0.020	0.015	−0.024	−0.049	0.002
	Household Size	−0.003	−0.025	0.019	0.006	−0.010	0.023	−0.003	−0.025	0.018
	RAND SD (Natural Village)	0.024	0.000	0.054	0.349	0.307	0.391	0.433	0.380	0.488
Zero-Inflated	Intercept	−2.334	−2.456	−2.214	−2.747	−2.930	−2.568	−2.164	−2.304	−2.028
	Rice Production Area/Farm Size	−0.227	−0.343	−0.109	0.507	0.362	0.651	−0.184	−0.297	−0.072
	Network Size	0.094	−0.014	0.203	0.068	−0.083	0.218	0.105	−0.023	0.232
	Age	−0.042	−0.153	0.072	0.465	0.334	0.596	−0.014	−0.120	0.094
	Female	0.519	0.194	0.829	0.544	0.140	0.934	0.469	0.161	0.776
	Education	−0.075	−0.187	0.037	0.045	−0.080	0.165	−0.094	−0.199	0.010
	% Income from Rice Production	−0.065	−0.172	0.041	−0.511	−0.636	−0.385	−0.063	−0.165	0.039
	Household Size	−0.137	−0.244	−0.031	−0.262	−0.379	−0.145	−0.142	−0.241	−0.040
	RAND SD (Natural Village)	0.226	0.000	0.399	0.620	0.439	0.804	0.502	0.358	0.657
Precision	Intercept	3.254	3.207	3.300	3.189	3.143	3.235	4.625	4.576	4.676
	Rice Production Area/Farm Size	−0.006	−0.060	0.051	−0.043	−0.094	0.008	−0.069	−0.131	−0.009
	Network Size	−0.096	−0.143	−0.051	−0.104	−0.150	−0.060	−0.289	−0.339	−0.238
	Age	0.010	−0.044	0.062	−0.019	−0.071	0.032	−0.016	−0.073	0.041
	Female	−0.005	−0.175	0.166	−0.133	−0.298	0.036	−0.062	−0.251	0.116
	Education	−0.032	−0.081	0.016	0.067	0.015	0.117	0.014	−0.039	0.065
	% Income from Rice Production	0.046	−0.006	0.099	0.011	−0.038	0.061	0.082	0.028	0.137
	Household Size	0.005	−0.040	0.052	−0.007	−0.053	0.039	0.036	−0.012	0.082
	N_{Farmers}	4526			4526			4507		
	$N_{\text{Natural Villages}}$	162			162			160		
	Minimum ESS	13,938			13,381			14,731		
	Minimum Bulk ESS	13,933			13,396			14,777		
	Minimum Tail ESS	20,659			25,403			26,910		

PMEAN = Posterior Mean.

HDI = 95% Highest Density Interval (Lower & Upper).

RAND SD = Standard deviation of the random intercepts capturing village/network/SAOM-specific deviations from the overall intercept.

Minimum ESS/Bulk ESS/Tail ESS = The smallest effective sample size amongst all varying and non-varying parameters in all sub-models.

size as a meta-predictor in an extended model. Furthermore, the estimated associations between network size and the magnitude of the SAOM parameter estimates for *Direct Reciprocity*, *Transitivity* and *Generalised Exchange* are all positive — where those for *Direct Reciprocity* and *Transitivity* are consistent with the meta-regressions of ERGM parameter estimates presented in Table 4 of McFarland et al. (2014; see estimates for “mutuality” and “closure” on p. 1107). Moreover, note that there is, on average, a tendency against the reciprocation of friendship inside of transitive triads (Negative $\hat{\theta}_{\text{Transitive Reciprocated Triplets (Intercept)}}$). And as $\hat{\theta}_{\text{Direct Reciprocity (Intercept)}}$ is positive, this suggests that the farmers in this sample prefer mutual friendships outside of transitive triplets, corroborating emerging evidence from analyses of adolescents in Western societies that one-sided friendship is stabilised when embedded in a group (e.g., see Block, 2015; Kretschmer et al., 2018 and Raabe et al., 2019).

Of course, these population parameter estimates tell us nothing about the relative importance of the three mechanisms of interest. Respectively, the updated/shrunk parameters $\hat{\beta}_{\text{Direct Reciprocity}}$, $\hat{\beta}_{\text{Transitivity}}$ and $\hat{\beta}_{\text{Generalised Exchange}}$ are positive for 162, 162 and 160 of the villages. And across the farmers in these villages, the median values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ are, respectively, 0.069, 0.203 and 0.016. Accordingly, *Transitivity* appears to generally make a much larger proportional contribution to the friend choices of the farmers in this particular sample compared to *Direct*

Reciprocity and *Generalised Exchange*. That said, what is far more interesting is the distribution of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ across the farmers which suggests that the three mechanisms of interest can take on wide-ranging roles in the friend-choice process. Specifically, $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ assume an array of values — respectively, 0 to 0.419 (468 zeros), 0 to 0.736 (399 zeros) and 0 to 0.126 (564 zeros) — raising the question of what might drive this variation.

Parameter estimates from the mean, zero-inflated and precision components of the beta regression models for $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ appear in Table 4. However, what is perhaps more useful is a series of graphs depicting the regression curves based on the fixed (i.e., non-varying) effects from the mean and zero-inflated sub-models for the typical farmer in the sample who is embedded in a network of a typical size. Specifically, I limit my attention to a male in an average-sized network (global) of 32 rice farmers. This individual is 51 years of age and he has roughly a primary school education and derives 74% of his income from rice production whilst leading a household composed of five people. For this individual, Figs. 1–3, respectively depict the expected values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ alongside the expected probabilities of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ being zero in relation to log rice production area using the three models in Table 4.

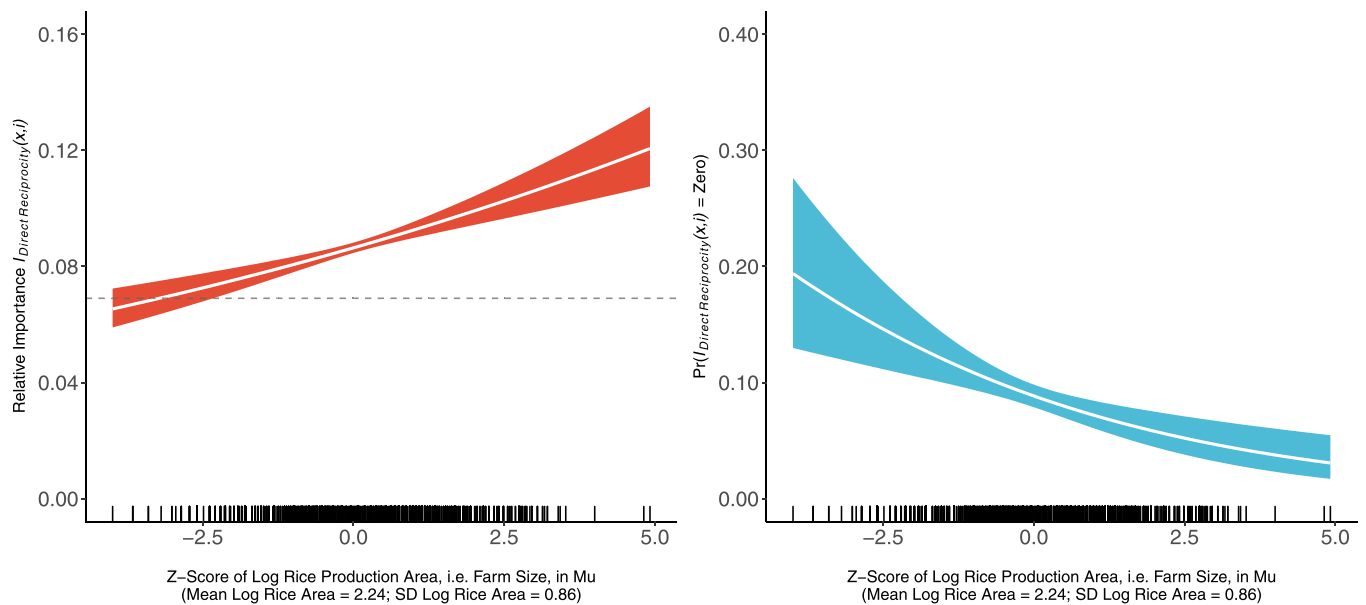


Fig. 1. Expected values of the proportional contribution of *Direct Reciprocity* to the inferred friend choices of the Jiangxi farmers from the Bayesian multilevel beta regression model alongside the expected probabilities of these contributions being zero in relation to the size of farmers' rice production areas. The expected values from the mean sub-model (left) and the expected probabilities from the zero-inflated sub-model (right) are based on the fixed (i.e., non-varying) effects with all covariates held at their global average values (see Table 2). Regression lines are in white with the coloured bands representing the 95% Credible Interval (i.e., the 2.5% and the 97.5% quantiles). Regression lines are constructed using 100 support points (i.e., their resolution). Rug plots (bottom) depict the distribution of log rice production area across the sample of 4526 farmers with complete data used to fit the beta regression model. The dashed grey lines (left) indicates the median value of $I_{\text{Direct Reciprocity}}(x, i)$ for these farmers.

Consistent with my supposition that growth in the demands of rice production will lead smallholders to prioritise building strong ties conducive to cooperation at the dyadic level (Hypothesis 1), as farm size increases, the expected proportional contribution of *Direct Reciprocity* to a smallholder's inferred friend choices also increases (Fig. 1; Left). More precisely, the expected magnitude of

$I_{\text{Direct Reciprocity}}(x, i)$ rises from 0.065 to 0.121 between those with the smallest farms and those with the largest. Further still, the negative relationship between farm size and the probability that $I_{\text{Direct Reciprocity}}(x, i)$ will be zero serves as additional support for Hypothesis 1 as smallholders with increasingly larger rice production areas are quite unlikely to be unmoved by considerations of mutuality

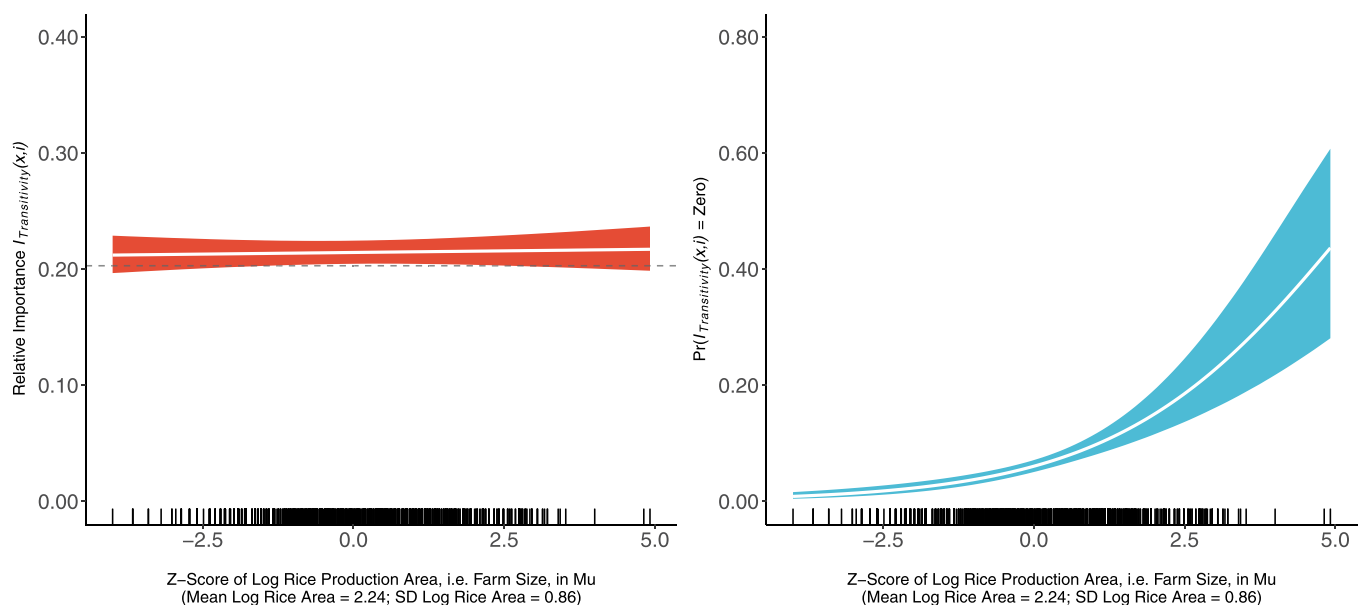


Fig. 2. Expected values of the proportional contribution of *Transitivity* to the inferred friend choices of the Jiangxi farmers from the Bayesian multilevel beta regression model alongside the expected probabilities of these contributions being zero in relation to the size of farmers' rice production areas. The expected values from the mean sub-model (left) and the expected probabilities from the zero-inflated sub-model (right) are based on the fixed (i.e., non-varying) effects with all covariates held at their global average values (see Table 2). Regression lines are in white with the coloured bands representing the 95% Credible Interval (i.e., the 2.5% and the 97.5% quantiles). Regression lines are constructed using 100 support points (i.e., their resolution). Rug plots (bottom) depict the distribution of log rice production area across the sample of 4526 farmers with complete data used to fit the beta regression model. The dashed grey lines (left) indicates the median value of $I_{\text{Transitivity}}(x, i)$ for these farmers.

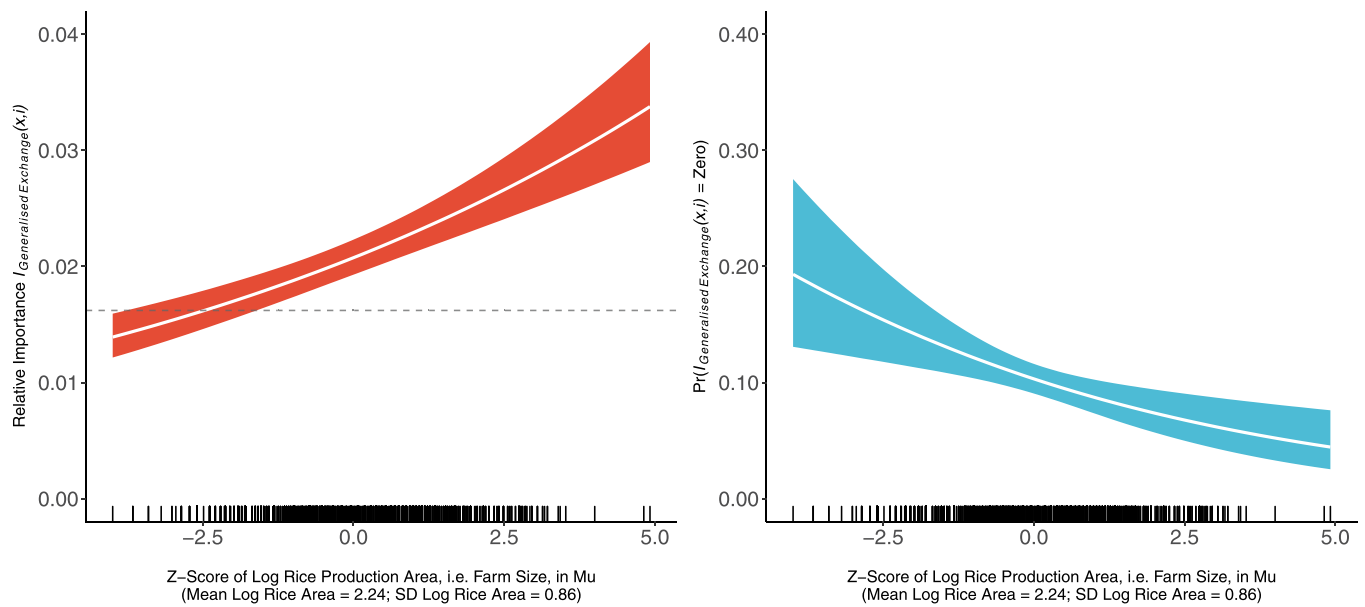


Fig. 3. Expected values of the proportional contribution of *Generalised Exchange* to the inferred friend choices of the Jiangxi farmers from the Bayesian multilevel beta regression model alongside the expected probabilities of these contributions being zero in relation to the size of a farmer's rice production area. The expected values from the mean sub-model (left) and the expected probabilities from the zero-inflated sub-model (right) are based on the fixed (i.e., non-varying) effects with all covariates held at their global average values (see Table 2). Regression lines are in white with the coloured bands representing the 95% Credible Interval (i.e., the 2.5% and the 97.5% quantiles). Regression lines are constructed using 100 support points (i.e., their resolution). Rug plots (bottom row) depict the distribution of log rice production area across the sample of 4507 farmers with complete data used to fit the beta regression model. Dashed grey lines (left) indicate the median value of $I_{Generalised\ Exchange}(x, i)$ for these farmers.

compared to their counterparts with smaller-scale operations. More specifically, the $Pr(I_{Direct\ Reciprocity}(x, i) = 0)$ falls from 0.194 to 0.031 between those with the smallest farms and those with the largest.

On the other hand, and contrary to my first triadic supposition (Hypothesis 2), the expected proportional contribution of *Transitivity* to a farmer's inferred friend choices appears to have no substantive association with farm size (Fig. 2; Left). More precisely, the expected magnitude of $I_{Transitivity}(x, i)$ rises by a miniscule amount from 0.212 to 0.217 between those with the smallest farms and those with the largest. What is more striking, however, is that the expected probability of $I_{Transitivity}(x, i)$ being zero — i.e., that *Transitivity* will play no role at all in who a smallholder chooses as a friend — rises markedly with farm size. More specifically, the $Pr(I_{Transitivity}(x, i) = 0)$ grows from 0.008 to 0.437 between those with the smallest farms and those with the largest.

As for my second triadic supposition (Hypothesis 3), results are similar to those seen in the model for $I_{Direct\ Reciprocity}(x, i)$ in that they are consistent with my argument that the rising demands of rice production will lead smallholders to prioritise building ties with their agricultural counterparts in a manner that is conducive to cooperation. Accordingly, as farm size increases, the expected proportional contribution of *Generalised Exchange* to a smallholder's inferred friend choices grows (Fig. 3; Left). More precisely, the expected magnitude of $I_{Generalised\ Exchange}(x, i)$ increases from 0.014 to 0.034 between those with the smallest farms and those with the largest farms. And unlike the results from the beta regression model of $I_{Transitivity}(x, i)$, the association between farm size and the expected probability that $I_{Generalised\ Exchange}(x, i)$ will be zero is negative. More specifically, the $Pr(I_{Generalised\ Exchange}(x, i) = 0)$ falls from 0.193 to 0.044 between those with the smallest farms and those with the largest.

Model validity

By this point, discerning readers will undoubtedly wonder about the overall quality of the models that I have discussed — especially as the validity of my attempt to measure relative importance is directly linked to the accurate representation of the farmers' observed friend choices

on the part of each of the underlying Stochastic Actor-Oriented Models. As is standard when using multi-mechanistic models, correspondence between model and the observed data was judged with multi-dimensional auxiliary statistics. These statistics are summaries of macro-level features of the observed network that should be faithfully reflected in networks simulated in accordance with: (i) the micro-level choice preferences implied by the estimated (here, un-shrunk/non-updated) parameters $\hat{\beta}$ (Snijders and Steglich, 2015); and, in my particular case, (ii) the additional constraint that actors' out-degrees be five or less. In selecting auxiliary statistics, I took a maximal approach by using a battery of metrics often discussed by network analysts. Namely, I focused on the distribution of in-degrees, out-degrees, geodesic distances, and eigenvector centralities as well as the clique census and the triad census.

The difference between the distribution of an auxiliary statistic across the observed network and the distribution of that same statistic across the simulated networks was assessed using a test of Mahalanobis distance (Ripley et al., 2019). The null hypothesis is that these distributions are the same making its rejection undesirable. Although this formal test is secondary to visual inspection of the distributions, use of p -values (one-sided) alone allows one to efficiently judge goodness-of-fit when working with a large number of networks for the purposes of a meta-analysis (e.g., see Block, 2015).

Along this line, the proportion of the 162 SAOMs with p -values greater than 0.05 for the tests of Mahalanobis distance used to compare the observed and simulated distributions of in-degrees, out-degrees, geodesic distances and eigenvector centralities, as well as the triad census and the clique census is, respectively, 0.98, 0.88, 0.94, 0.96, 1.00 and 0.94. As crude summaries of model validity, these proportions indicate that fit is, on the whole, excellent such that the global SAOM model specification generates synthetic networks that are a good approximation of the networks that were observed in the vast majority of the villages.

Somewhat akin to the SAOMs, I judged the fit of the beta regression models with graphical posterior predictive checks (Gabry et al., 2019) whereby the distribution of the measured values of $I_{Direct\ Reciprocity}(x, i)$,

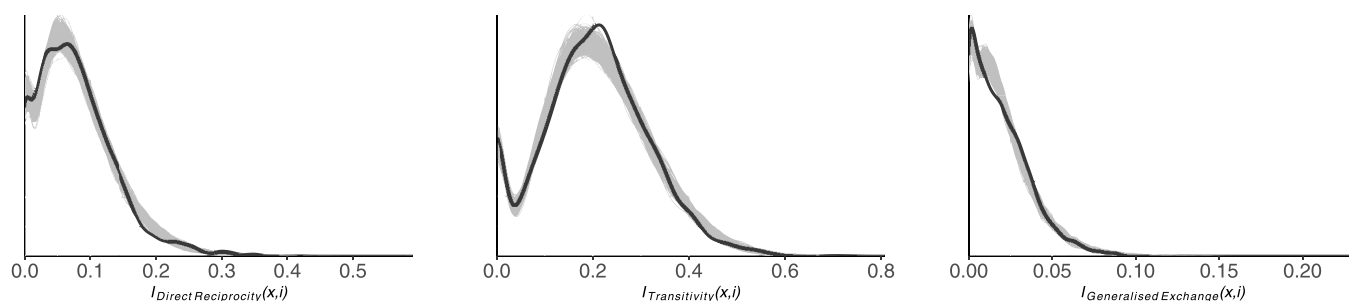


Fig. 4. Kernel density estimates for the distributions of the measured values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ (i.e. the dark line) against distributions of simulated values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ (i.e., light thin lines) using 500 samples from the posterior predictive distribution based on each mechanism's corresponding zero-inflated beta regression model in Table 4.

$I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ are qualitatively compared against simulated values of $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ using a large number of samples from the posterior predictive distribution (here, arbitrarily, 500). Correspondence between these distributions is generally excellent (Fig. 4), suggesting that the specification used for the beta regressions generates relative shares of influence that are a good approximation of what was measured. Furthermore, the results depicted in Table 4 are robust to an alternative specification of the prior distributions for the overall intercepts and the non-varying effects wherein the Student-*t* distributions are replaced with Gaussian distributions (see Appendix B and Table A1).

Discussion and conclusion

In this paper I have attempted to test the micro-level component of McFarland et al.'s (2014) theory of network ecology in relation to the relative importance of generative mechanisms. Inspired by socio-ecological research on the behavioural implications of subsistence style and drawing heavily on scholarship around the social organisation of rice farming in China, I have explored in particular the association between the amount of land that a farmer controls that is devoted to rice production and the relative power of reciprocity, transitivity and generalised exchange over who this rice-producer chooses as a friend. With respect to my hypotheses, findings are best described as mixed. Specifically, estimates from a set of multilevel zero-inflated beta regression models indicate that farm size is positively associated with the relative importance of *Direct Reciprocity* and *Generalised Exchange*, as was anticipated in my first and third hypotheses, but negatively associated with the relative importance of *Transitivity*, contra to my second hypothesis. There are perhaps two plausible explanations for these unexpected results to be clarified with further study.

On the one hand, direct reciprocity is a quite an easy process to act upon — a farmer only needs to know who engages with him in a friendly manner in order to respond in kind. In contrast, transitivity requires two unconnected farmers to be exposed to and repeatedly interact with one another by way of their intermediaries (Schaefer et al., 2010) and management of a large land area, which is likely divided into small, non-contiguous parcels given a history of land fragmentation in rural China (Chen et al., 2010; Tan et al., 2008; Wu et al., 2018), may simply curtail opportunities to meet friends-of-friends and, in turn, undermine this particular integrative process.

On the other hand, the positive relationship between farm size and the probability that $I_{\text{Transitivity}}(x, i)$ will be equal to zero may be a manifestation of the more instrumental concerns of larger-scale rice producers. More specifically, disregarding normative pressure to build “redundant” ties — i.e., the friendship $i \rightarrow j$ between the information seeker i and information provider j in the face of the two-path $[i \rightarrow k \rightarrow j]$ or the in-two-star $[i \rightarrow k \leftarrow j]$ — would allow rice producers with larger farms to avoid superfluous information around finance and cultivation and access new agricultural knowledge to the extent that it

is associated with having an open local network (Aral, 2016). Given recognised gaps in Chinese rice producers' agricultural know-how (Cui et al., 2018; Wu et al., 2018; Zhang et al., 2013, 2016) and research suggesting that new, environmentally-beneficial agricultural knowledge originates outside of rural farming communities (Matous, 2015; Matous and Wang, 2019), the avoidance of redundant ties stands to be advantageous. Indeed, new research by Wu et al. (2018) using nationally representative data from China suggests that farmers with more land use less environmentally-damaging agricultural chemicals and are more productive — outcomes that the authors attribute to the possession of better agricultural knowledge and management skills on the part of larger-scale producers of rice and other major crops (e.g., wheat and maize).

How then should one make sense of the positive relationship between farm size and *Generalised Exchange* which, qualitatively speaking, mirrors that which is observed for farm size and *Direct Reciprocity*? On the face of it, this triadic result seems inconsistent with land fragmentation derailing meeting opportunities for friends-of-friends (i.e., $[i \leftarrow k \leftarrow j]$ in the three cycle $[i \leftarrow k \leftarrow j \leftarrow i]$), although this regrettably cannot be tested due to my lack of more detailed data on geography and the farmers' plots (see Appendix C). Yet three cycles, which represent a very specific kind of chain-generalised exchange, are arguably also poor vehicles for farmers to use to access new agricultural knowledge through friendship due to the information redundancies inherent to these sub-graphs. That said, chain-generalised exchange is a form of structured reciprocity that is indirect in nature (Molm, 2010; Simpson et al., 2017). And as Molm (2010) underscores, “...the act of reciprocity itself has value — expressive value — over and above the instrumental value of benefits obtained through exchange. Consequently, maintaining or increasing the frequency of ... reciprocity may be a stronger motivation than maximizing the value of the benefits obtained from exchange” (p. 129; see also p. 127–128 on actors' motivations and short-term costs in return for long-term gains). As a result, it is possible that any perceived cooperative returns to generalised friendliness amongst the farmers — namely, trust and group solidarity — outweigh the costs of being exposed to redundant information, leading farmers who are engaged in larger-scale rice production to assign a higher priority to direct and indirect reciprocity when engaging with their agricultural counterparts.

Implications of findings

Although my results concern rice farming in rural China, they have sufficiently general implications to be relevant to all researchers with an interest in choice of social contacts amongst humans and notions of universality. Specifically, the association between farm size and the relative power of *Direct Reciprocity*, *Transitivity* and *Generalised Exchange* — particularly the impact of rice production area on whether or not each of these mechanisms play any role in friend choice — is straightforward and compelling evidence that network ecology can

moderate the relative importance of generative mechanisms. And while definitive conclusions cannot be made due to the observational and case-based nature of my research, to say nothing of the limitations of the analysed data (e.g., the artificial constraining of farmers' out-degrees), the evidence I have presented also hints that the micro-level component of McFarland et al.'s (2014) theory of network ecology should be generalised with possible wide relevance to adults and to non-educational social settings. Together, these findings have two non-trivial consequences for how we should study the formation of face-to-face networks going forward: one methodological; the other, theoretical.

First, analysts ought to formally assess relative importance, especially when using widely-popular multi-mechanistic models like the SAOM and the ERGM, before they draw conclusions about the role some generative mechanism plays in how individuals choose their social contacts. This is made plain when considering the great variation in the relative importance of *Direct Reciprocity* and *Transitivity* across the 4713 farmers (Table 2) — a simple but telling result that is inconsistent with the presumed “baseline” level of influence of both dynamics which, along with homophily, have captured the imagination of students of networks for many years (Block, 2015; Rivera et al., 2010; Wimmer and Lewis, 2010). Although the precise degree to which direct reciprocity and transitivity should overshadow other mechanisms is left unspecified in the academic literature, analysts regularly position both as the most central aspects of tie formation without question. Implicit in this positioning, however, is the hypothesis that *Transitivity* and/or *Direct Reciprocity* should be pre-eminent in their relative importance (i.e., in the present parlance, $I_{\text{Transitivity}}(x, i) + I_{\text{Direct Reciprocity}}(x, i) \geq 0.5$). Using extensive data, I have found that this kind of blanket supposition is perhaps unfounded as the degree to which either of these mechanisms dominates friend choice appears to be contingent on circumstances such that for some individuals, they may be expected to hold substantial sway over the process of choosing social contacts and, for others, no sway at all. Practically speaking, researchers may be overlooking important and interesting heterogeneity in the relative power of tie-formation processes (see also generative network models with actor-level random effects (Box-Steffensmeier et al., 2018; Minhas et al., 2019)).¹⁵

And second, analysts making theoretical claims around how individuals might be expected to choose their social contacts must, at the very least, carefully consider and clearly communicate through scope conditions (Harris, 1997) how network ecology might foster or restrict the *generative salience* of tie-formation processes — i.e., the degree to which some posited mechanism is likely to be top of mind relative to other dynamics from the perspective of those tasked with making a relational decision. To be clear, this should not be construed as an argument for what Henrich et al. (2010, p. 62) call “radical versions of interpretivism and cultural relativity” in an attempt to deny shared commonalities amongst humans. There may well be consistencies in the way that oft-discussed mechanisms have over the process of choosing social contacts across social, cultural and physical environments. However, the current dearth of large-scale assessments of the relative importance of these dynamics, alongside the overwhelming focus on the networks of children and young people in educational settings in WEIRD societies, makes the extent of these consistencies unclear. Pending the aggregation of additional evidence from multi-mechanistic

models fit using more diverse samples, explicit qualification of *a priori* claims around the relative power of tie-formation processes is prudent.¹⁶

Limitations and directions for future research

At this point my results must be qualified with respect to two key limitations linked to clear directives for future research designed to further test the theory of network ecology with an eye to relative importance.

First, I have exclusively relied on cross-sectional data. The insightful work of Indlekofer and Brandes (2014) and Schaefer et al. (2010) suggests that the roles played by generative mechanisms can vary over time, raising the question of the temporal invariance of values of $I_k(x, i)$ and whether they may systematically differ for the creation of wholly new ties and the maintenance of old connections (i.e., reciprocation with a new friendship versus the stability of a reciprocated dyad). Moving forward, comprehensive analyses of the interplay between relative importance and network ecology will ultimately require longitudinal data for large samples of face-to-face networks. Along this line, assessments of $I_k(x, i)$ before and after land reallocation (Brandt et al., 2002) or land expropriation (Cai, 2016; Cai and Sun, 2018) — events that stand to shape rice producers' social and economic behaviours (Brandt et al., 2002; Deininger et al., 2014) — strike me as a powerful means of leveraging longitudinal designs to more definitively assess how proximal physical environment moderates the relative importance of tie-formation mechanisms in the rural Chinese agricultural context.

And second, I have tested the relationship between network ecology and relative importance at the individual level. Yet, more distal aspects of village geography and sociocultural ecology beyond network size also stand to moderate relative importance over and above factors at the level of individuals, at least in East Asian rice areas (Uchida et al., 2018). In addition to those characteristics of communities that may make rice farming more or less difficult (e.g., heterogeneity in soil acidification (Guo et al., 2018)), of particular interest in the rural Chinese agricultural context is the total amount of arable land in a village and how it is distributed amongst resident farmers. This is because great inequity in land holdings may adversely impact the relative importance of direct reciprocity, transitivity and generalised exchange by undermining collective sentiment and a cooperative ethos to the extent that egalitarian land control is linked to social and economic security (Brandt et al., 2002; Chen et al., 2010; Cai, 2016) and the struggle for autonomy and sense of belonging amongst the Chinese agrarian class (see van der Ploeg et al., 2014).

Limitations notwithstanding, here I have taken a critical first step in demonstrating the necessity of the principled comparative analysis of

¹⁵ As mentioned in the introduction, there are currently no formal measures of the relative importance of effects in ERGMs. Moreover, the ERGM and the SAOM make distinct assumptions about how ties come into being (Block et al., 2019; see also Appendix A). Nevertheless, key similarities between the SAOM and the ERGM (e.g., use of conditional log odds ratios; effects that are functions of one another, multiple possible outcomes (Amati et al., 2018)) give the strong impression that comparing the magnitude of parameter estimates from the latter model is also likely to betray variation in the relative importance of generative mechanisms.

¹⁶ I of course acknowledge that collecting face-to-face data in diverse locals is likely to be resource intensive and pose unique challenges for network analysts (see Perkins et al., 2015). Furthermore, my comments should not be taken as a dismissal of the value of network data concerning children and young people in WEIRD locales. Scientists, including many of those cited here, have made a number of exciting discoveries using new data collected in WEIRD educational settings and through the secondary analysis of older samples of networks from, for example, the National Longitudinal Study of Adolescent to Adult Health (Add Health) and the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU). Although I have attempted to interrogate variation in the relative power of mechanisms in relation to the theory of network ecology, perhaps one efficient way forward is the execution of more studies that eschew theoretical development to instead conduct simple descriptive analyses (Besbris and Khan, 2017) of relative importance using existing datasets, especially as interesting variation within WEIRD societies is also to be expected (e.g., see Henrich et al., 2010, p. 74–78 on differences amongst North Americans). These descriptive analyses would also raise awareness/illustrate the use of the measure of Indlekofer and Brandes (2014) or some new measure of relative importance for ERGMs.

the relative importance of generative mechanisms whilst reaffirming the necessity of an explicitly ecological perspective on networks, showing that this kind of research stands to complicate commonly held assumptions around how humans establish non-romantic social ties.

Acknowledgements

I thank Jing Cai for graciously fielding my queries about her dataset and, for their participation, I am deeply grateful to the farmers of

Jiangxi Province. Furthermore, I thank the two anonymous reviewers for their formative critique as well as Per Block, David S. Kirk, Jeremy Koster, P.M. Krafft, Emiel De Lange, Eleanor Power, Marc Sarazin, Tom A. B. Snijders, Felix C. Troup, Joachim De Weerd and Dingeman Wiertz for providing guidance around theory, data and/or method in relation to the various incarnations of this paper. Finally, this research would not have been possible without the generous support of a British Academy Postdoctoral Fellowship (Grant Number: pf170158).

Appendix A. The cross-sectional Stochastic Actor-Oriented Models (SAOM)

SAOMs were originally developed for the longitudinal investigation of network dynamics. However, their cross-sectional variant has recently been detailed and shown to hold great potential for the investigation of singly-observed networks when a more explicit focus on actors and their choices between competing ties is appropriate. Although the cross-sectional SAOM is new (Block et al., 2019; Snijders and Steglich, 2015) — indeed there exists no published extensive applications or simulation-based assessments of the method — the foregrounding of agency and choice alongside the ability to measure the relative importance of effects (Indlekofer and Brandes, 2014) makes this version of the SAOM the best available means of analysing the data from China in accordance with my research aims.

Nevertheless, the implications of this modelling decision deserve mention. Specifically, SAOMs assume that individuals compare ties based on their rewards and costs such that choosing a relatively more rewarding tie is a choice against a less attractive other. Consequently, in this modelling framework, actors cannot establish (i.e., “send”) an unlimited number of ties due to the zero-sum comparison of costs that relations incur (Block et al., 2019) — an appropriate constraint given the resource demands of offline, face-to-face friendship (see Dunbar, 2018). This is in contrast to the Exponential-family Random Graph Model (ERGM). Unlike the SAOM, which is defined at the level of actors, the ERGM is defined at the level of dyads and it is not driven by the mutual comparison of ties. Instead, the cost of each tie is considered in isolation which leads to the implicit assumption that there are no constraints on the number of ties actors may have (Block et al., 2019).

The matter of tie cost and mutual comparison points to the rather different dependence hierarchies of the SAOM and the ERGM. Most important to mention in this brief treatment is that a SAOM with only a density parameter assumes dependence between all ties that share a sender whereas an ERGM with only a density parameter assumes that all ties are independent (i.e., the Bernoulli random graph model (Amati et al., 2018)). This assumption of sender dependence makes explicit a sociological perspective on how agentic individuals might manage their relations (Snijders, 1996). However, it is important to stress that *ERGMs are not incompatible with notions of agency and action, nor are they in some way deficient and deserving of replacement by the SAOM*. The ERGM is simply more general in the sense that it suits multiple perspectives on how ties come into being, whereas the actor-oriented perspective favoured here is but one viewpoint amongst others found in the scientific literature (see Block et al., 2019; Butts, 2017 and Lusher et al., 2012). Nevertheless, practically speaking, the differences in dependence hierarchies make the models distinct such that the SAOM posits some dependence from the start. Block et al. (2019) provide an excellent extended discussion of differences between SAOMs and ERGMs with respect to their dependence hierarchies, amongst other features, and the implications that these differences have for empirical applications and the interpretation of results. See also Lusher et al. (2012) who comprehensively discuss the dependency assumptions of the ERGM.

With regard to the actual estimation of the cross-sectional SAOM, the summaries provided by Snijders and Steglich (2015) and Block et al. (2019) are most instructive. Specifically, the cross-sectional SAOM enables investigation of the formation of a singly-observed graph under the assumption that parameter estimates are for a network in short-term dynamic equilibrium. Indeed, this version of the SAOM may be defined as the stationary distribution of a Markov chain that is defined by the probabilities of transitioning between adjacent network states (i.e., structures differing by at most one tie) for a given set of actors N in a network x and a given set of parameters β associated with effects k (see Block et al., 2019, p. 206–209). Short-term dynamic equilibrium itself may be defined as follows (Snijders and Steglich, 2015): beginning with the observed network state and ending with the observed network state, i.e., $[x(t_{obs}), x(t_{obs})]$, the simulation of the stochastic and sequential micro-steps is run for a period of time allowing actors to make some average number of changes in their outgoing ties where this average is equal to the rate function λ . This yields a distribution of networks that has, for all aggregated micro-statistics $\sum_{i=1}^N s_{k,i}(x)$ corresponding to the model parameters, an average that is equivalent to their empirically observed values in x . Put simply, the singly-observed network is used as the starting and ending point for the SAOM estimation procedure, such that the simulation has the opportunity to start at, move away from, and, ultimately, reach an equilibrium state for which the observed network $x(t_{obs})$ is assumed to be representative. Although extensive investigation is lacking, it is understood that λ should be fixed at a high value, with too high a value perhaps incurring model degeneracy as λ approaches infinity (Snijders and Steglich, 2015). This is especially so when using linear counts of network statistics, as I do here to aid interpretability, versus their geometrically-weighted variants (Ripley et al., 2019).

The rate parameter λ was fixed at 20 for each of the 162 SAOMs. As maximum out-degrees were limited to five during data collection in Jiangxi, $\lambda = 20$ should, on average, afford farmers with the maximum number of outgoing ties the opportunity to turn over their entire portfolio of connections during the simulation four times which seems sufficient for my purposes. Cross-sectional SAOMs were estimated using the R package “RSiena” (Ripley et al., 2019) whilst following the coding directives offered by Block et al. (2019) and Snijders and Steglich (2015). With respect to settings for the estimation algorithm, all 162 SAOMs were fit using six sub-phases for Phase 2 and 6000 iterations for Phase 3. The overall maximum convergence ratio for all 162 SAOMs is less than 0.2. More information on these estimation settings are provided in the RSiena User’s manual (Ripley et al., 2019).

Appendix B. Rationale for the specification of priors for the beta regressions

Bayesian inference requires one to specify prior distributions (i.e., “priors”) which summarise information about each of the unknown model parameters that is external, at least in part, to the data themselves. Given the newness of the theory of network ecology and the dearth of assessments of the relative importance of generative mechanisms, I favour conservative choices for prior distributions. Specifically, I used priors that can be broadly described as “weakly informative” in that they contain just enough information to set reasonable bounds for the probability distribution for each parameter (i.e., the “posterior distribution” or simply “the posterior”) and yield stable parameter estimates (Gelman et al., 2013). Although choice of priors is ultimately subjective (Gelman et al., 2008), I note that relatively uninformative priors should more readily “let the data speak” by

playing a smaller role in influencing the estimated posterior distributions given the larger sample size (Kruschke, 2015) with the caveat that larger datasets are not necessarily informative.

As for the priors themselves, in older research, Gelman et al. (2008) recommend the use of a Cauchy prior (i.e., a Student-*t* distribution with one degree of freedom) that is centred at zero and that has a scale parameter of 2.5 for fixed/non-varying coefficients and a scale of 10 for the intercept for logistic regression models and, more generally, for models with logarithmic links. Practically speaking, these Cauchy priors “shrink” (i.e., place a larger probability mass on) coefficients near to zero whilst occasionally allowing for coefficients with large magnitudes due to their “heavy tails” (i.e. the prior distributions have gentle slopes that distribute probability mass across a wide range of values (Gelman et al., 2008, 2013)). Moreover, centring the prior distributions on zero reflects the assumption that there are no strong *a priori* reasons to expect that the relationships summarised by the coefficients for the fixed effects are positive or negative, where the larger scale parameter flattens/widens the distribution considerably in order to accommodate intercepts that are positive or negative with a magnitude that may assume a large range of values. In this latter respect, the Cauchy prior with a scale of ten is much “weaker” than the one with a scale of 2.5.

That said, more recently, Ghosh et al. (2018) recommend that applied researchers consider lighter-tail priors in the wake of their comparison of the performance of Cauchy, Gaussian (i.e., Normal) and Student-*t* priors for the purposes of logistic regression. Specifically, the authors raise concerns about the thickness of the Cauchy distribution’s tails leading to inefficient sampling from the posterior and poor shrinkage when the data in question are not informative — concerns echoed by the Stan Development Team (<https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations>; Accessed June 2019). Accordingly, to fit the models appearing in the main text, I adhered to Ghosh et al.’s (2018) suggestion that one use Student-*t* priors, as opposed to Gaussian priors, when the robustness of results to outliers is a concern; guidance that is also consistent with the recommendations of the Stan Development Team. However, to judge the sensitivity of my results to prior specification, I also estimated ancillary models for $I_{\text{Direct Reciprocity}}(x, i)$, $I_{\text{Transitivity}}(x, i)$ and $I_{\text{Generalised Exchange}}(x, i)$ wherein the Student-*t* priors are replaced with Gaussian priors.

More precisely, priors for the intercepts in the mean, precision and zero-inflated sub-models appearing in Table 4 of the main text take the form of Student-*t* distributions with seven degrees of freedom, a mean of zero and scale of ten as in Ghosh et al. (2018). Additionally, parameters for all other fixed/non-varying effects are given Student-*t* priors with seven degrees of freedom, a mean of zero and scale of 2.5, again following Ghosh et al. (2018). In the ancillary models, which appear in Table A1, the Student-*t* priors for the intercepts are replaced with Gaussian priors with a mean of

Table A1
Ancillary Bayesian Multilevel Beta Regression Models with Zero-inflation.

Sub-Model	Covariate	$I_{\text{Direct Reciprocity}}(x, i)$			$I_{\text{Transitivity}}(x, i)$			$I_{\text{Generalised Exchange}}(x, i)$		
		PMEAN	HDI-L	HDI-U	PMEAN	HDI-L	HDI-U	PMEAN	HDI-L	HDI-U
Mean	Intercept	−2.358	−2.381	−2.336	−1.299	−1.359	−1.242	−3.854	−3.931	−3.780
	Rice Production Area/Farm Size	0.076	0.049	0.102	0.003	−0.016	0.023	0.102	0.073	0.130
	Network Size	−0.040	−0.062	−0.019	−0.070	−0.126	−0.015	0.200	0.129	0.273
	Age	0.019	−0.007	0.045	−0.014	−0.033	0.005	0.030	0.004	0.057
	Female	−0.041	−0.122	0.044	−0.037	−0.101	0.026	−0.032	−0.128	0.062
	Education	0.051	0.027	0.073	0.004	−0.014	0.020	0.047	0.023	0.071
	% Income from Rice Production	−0.017	−0.042	0.008	−0.003	−0.021	0.015	−0.024	−0.049	0.002
	Household Size	−0.003	−0.025	0.019	0.006	−0.010	0.023	−0.003	−0.025	0.018
	RAND SD (Natural Village)	0.024	0.000	0.054	0.349	0.308	0.392	0.433	0.380	0.488
Zero-Inflated	Intercept	−2.334	−2.456	−2.213	−2.747	−2.936	−2.574	−2.165	−2.305	−2.030
	Rice Production Area/Farm Size	−0.227	−0.344	−0.109	0.507	0.364	0.651	−0.184	−0.297	−0.073
	Network Size	0.094	−0.013	0.202	0.067	−0.080	0.221	0.106	−0.022	0.233
	Age	−0.042	−0.153	0.073	0.464	0.332	0.595	−0.014	−0.122	0.093
	Female	0.519	0.204	0.837	0.544	0.148	0.940	0.469	0.153	0.773
	Education	−0.075	−0.187	0.037	0.045	−0.082	0.163	−0.094	−0.197	0.013
	% Income from Rice Production	−0.065	−0.174	0.040	−0.511	−0.637	−0.387	−0.063	−0.166	0.039
	Household Size	−0.137	−0.246	−0.031	−0.262	−0.380	−0.144	−0.142	−0.244	−0.041
	RAND SD (Natural Village)	0.226	0.000	0.398	0.620	0.447	0.808	0.503	0.351	0.651
Precision	Intercept	3.254	3.206	3.299	3.189	3.144	3.235	4.625	4.577	4.675
	Rice Production Area/Farm Size	−0.006	−0.063	0.049	−0.043	−0.096	0.008	−0.069	−0.130	−0.007
	Network Size	−0.096	−0.142	−0.050	−0.104	−0.150	−0.060	−0.289	−0.339	−0.238
	Age	0.010	−0.043	0.061	−0.019	−0.070	0.033	−0.016	−0.072	0.041
	Female	−0.005	−0.177	0.162	−0.134	−0.303	0.030	−0.062	−0.252	0.114
	Education	−0.032	−0.081	0.016	0.067	0.016	0.118	0.014	−0.037	0.066
	% Income from Rice Production	0.047	−0.006	0.099	0.011	−0.038	0.062	0.082	0.026	0.136
	Household Size	0.005	−0.040	0.051	−0.007	−0.052	0.040	0.037	−0.011	0.083
N_{Farmers}		4526			4526			4507		
$N_{\text{Natural Villages}}$		162			162			160		
Minimum ESS		14,707			12,307			11,949		
Minimum Bulk ESS		14,712			12,335			12,006		
Minimum Tail ESS		21,361			24,539			23,769		

PMEAN = Posterior Mean.

PSD = Posterior Standard Deviation.

HDI = 95% Highest Density Interval (Lower & Upper).

RAND SD = Standard deviation of the random intercepts capturing village/network/SAOM-specific deviations from the overall intercept.

Minimum ESS/Bulk ESS/Tail ESS = The smallest effective sample size amongst all varying and non-varying parameters in all sub-models.

zero and a scale of ten and the Student-*t* priors for the fixed effects are replaced with Gaussian priors with a mean of zero and a scale parameter of 2.5, again following Ghosh et al. (2018).

Note that the shape of the Student-*t* distribution departs from the shape of the Cauchy distribution and approaches the shape of the Gaussian distribution as the degrees of freedom increase. Consequently, in the present scenario, the Student-*t* priors are less informative than the Gaussian priors, the latter of which make a comparatively stronger statement that parameter estimates should be nearer to zero (i.e., no effect) in that they *a priori* rule against extreme parameter values (given their scale). That said, the Student-*t* priors are themselves more informative than the Cauchy priors such that they occupy a kind of “middle-ground”.

Choice of a prior for the parameters summarising the variation of the random intercepts in the mean and zero-inflated sub-models of the beta regressions appearing in Tables 4 and A1 was much more straightforward. In line with Gelman's (2006) general recommendations and Liu and Eugenio's (2016) specific recommendations for zero-inflated beta regression models, the standard deviations of the random intercepts for natural village/network were restricted to be non-negative and given half Student-*t* priors (i.e., a Student-*t* distribution “folded” at zero) which is the default in Bürkner's “brms” package (Bürkner, 2018, 2019). These half Student-*t* prior distributions have three degrees of freedom and a scale of ten.

Appendix C. Limitations of the specification of the beta regression models related to farm size

Beyond the constrained variances brought about by not using random slopes for rice production area, there are a series of weaknesses of the specification of the beta regression models related to the available information on farm size that should be mentioned.

First, in addition to lacking information on rice yield, the dataset collected by Cai et al. (2015) does not distinguish between types of rice production land. A 1992 survey of land tenure and land rights across rural China suggests that the rice areas of the Jiangxi farmers will primarily take the form of “responsibility land” (*zeren tian*) allocated for the purposes of quota-based production (see Brandt et al., 2002, p. 75), with more recent data from 2015 indicating that current farm sizes reflect those of the early 2000s due to enduring constraints around land transfer in China (Wu et al., 2018, p. 7011). Nevertheless, there are other types of land subject to different expectations, pressures and constraints (see Brandt et al., 2002, p. 73–74) that may have unique implications for the need to cooperate for the purposes of rice production and thus how farmers build relationships with their agricultural counterparts (e.g., grain ration land (*kouliang tian*) vs. private family plots (*ziliu di*) vs. contracted plots (*chengbao tian*)).

Furthermore, despite information on land use in Jiangxi suggesting that the vast majority of the land cultivated by the province's farmers will be devoted to rice production (Guo et al., 2018; Hu and Yuan, 2015; Tan et al., 2008), farmers may still have land earmarked for other crops (e.g., fruits, vegetables and corn (Hu and Yuan, 2015; Tilt, 2008)). Controlling for the percentage of a farmer's income that is derived from rice crudely incorporates information on non-rice reliance into the models. Unfortunately, however, unambiguous information on the production of other agricultural items is not available.

Last, rice production areas may be fragmented, with a 2000 survey of plot cultivation amongst a sample of 331 households in three natural villages in Jiangxi by Tan et al. (2008, Footnote 4) finding that homes had an average of 7.4 plots (Range = 1–17). Regrettably, however, I lack geospatial information on how the land under the control of each farmer is divided such that the role of farm size and the role of how a farm is distributed across physical space cannot be distinguished. Accordingly, the observed relationships between farm size and relative importance may in part stem from the management of non-contiguous land parcels which seems likely to impact how rice producers traverse their immediate physical environment and thus their opportunities for social interaction (see Small and Adler (2019) on tie formation and the configuration of space).

Appendix D. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.socnet.2019.10.001>.

References

- Adams, Glenn, Plaut, Victoria C., 2003. The cultural grounding of personal relationship: friendship in North American and West African worlds. *Pers. Relationsh.* 10 (3), 333–347.
- Amati, Viviana, Lomi, Alessandro, Mira, Antonietta, 2018. Social network modeling. *Annu. Rev. Stat. Appl.* 5 (1), 343–369.
- An, Weihua, 2015. Multilevel meta network analysis with application to studying network dynamics of network interventions. *Soc. Netw.* 43, 48–56.
- Anderson, Bringham S., Butts, Carter T., Carley, Kathleen, 1999. The interaction of size and density with graph-level indices. *Soc. Netw.* 21 (3), 239–267.
- Aoki, Masahiko, 2001. *Toward a Comparative Institutional Analysis*. MIT Press.
- Apicella, Coren L., Marlowe, Frank W., Fowler, James H., Christakis, Nicholas A., 2012. Social networks and cooperation in hunter-gatherers. *Nature* 481 (7382), 497–501.
- Aral, Sinan, 2016. The future of weak ties. *Am. J. Sociol.* 121 (6), 1931–1939.
- Baldassarri, Delia, 2015. Cooperative networks: altruism, group solidarity, reciprocity, and sanctioning in Ugandan producer organizations. *Am. J. Sociol.* 121 (2), 355–395.
- Baldassarri, Delia, Diani, Mario, 2007. The integrative power of civic networks. *Am. J. Sociol.* 113 (3), 735–780.
- Besbris, Max, Khan, Shamus, 2017. Less theory. More description. *Sociol. Theory* 35 (2), 147–153.
- Block, Per, 2015. Reciprocity, transitivity, and the mysterious three-cycle. *Soc. Netw.* 40, 163–173.
- Block, Per, Stadtfeld, Christoph, Snijders, Tom A.B., 2019. Forms of dependence: comparing SAOMs and ERGMs from basic principles. *Sociol. Methods Res.* 48 (1), 202–239.
- Box-Steffensmeier, Janet M., Christenson, Dino P., Morgan, Jason W., 2018. Modeling unobserved heterogeneity in social networks with the frailty exponential random graph model. *Polit. Anal.* 26 (1), 3–19.
- Brandt, Loren, Huang, Jikun, Li, Guo, Rozelle, Scott, 2002. Land rights in rural china: facts, fictions and issues. *China J.* 47, 67–97.
- Burt, Ronald S., Bian, Yanjie, Oppen, Sonja, 2018. More or less guanxi: trust is 60% network context, 10% individual difference. *Soc. Netw.* 54, 12–25.
- Bürkner, Paul-Christian, 2018. Advanced bayesian multilevel modeling with the R package Brms. *R J.* 10 (1), 395–411. Retrieved. <https://journal.r-project.org/archive/2018/RJ-2018-017/index.html>.
- Bürkner, Paul-Christian, 2019. Package ‘Brms’: Bayesian Regression Models Using ‘Stan’ (v. 2.10.0). Retrieved. <https://cran.r-project.org/web/packages/brms/brms.pdf>.
- Butts, Carter T., 2017. Comment: actor orientation and relational event models. *Sociol. Methodol.* 47 (1), 47–56.
- Cai, Jing, 2012. Social networks and the decision to insure: evidence from randomized experiments in China (Job Market paper). Goldman School of Public Policy Working Paper 1–65. Retrieved. <https://gspp.berkeley.edu/research/working-paper-series/social-networks-and-the-decision-to-insure-evidence-from-randomized-experim>.
- Cai, Jing, Janvry, Alain D., Sadoulet, Elisabeth, 2015. Social networks and the decision to insure. *Am. Econ. J. Appl. Econ.* 7 (2), 81–108.
- Cai, Meina, 2016. Land for welfare in China. *Land Use Policy* 55, 1–12.
- Cai, Meina, Sun, Xin, 2018. Institutional bindingness, power structure, and land expropriation in China. *World Dev.* 109, 172–186.
- Chan Tack, Anjanette, Small, Mario L., 2017. Making friends in violent neighborhoods: strategies among elementary school children. *Sociol. Sci.* 4, 224–248.
- Chen, Zhuo, Huffman, Wallace E., Rozelle, Scott, 2010. Inverse relationship between productivity and farm size: the case of China. *Contemp. Econ. Policy* 29 (4), 580–592.
- Cohn, Avery S., et al., 2017. Smallholder agriculture and climate change. *Annu. Rev. Environ. Resour.* 42 (1), 347–375.
- Cui, Zhenling, et al., 2018. Pursuing sustainable productivity with millions of smallholder farmers. *Nature* 555 (7696), 363–366.
- Cullen, Alison C., Anderson, C.L., Biscaye, Pierre, Reynolds, Travis W., 2018. Variability in cross-domain risk perception among smallholder farmers in Mali by gender and other demographic and attitudinal characteristics. *Risk Anal.* 38 (7), 1361–1377.
- Deininger, Klaus, Jin, Songqing, Xia, Fang, Huang, Jikun, 2014. Moving off the farm: land institutions to facilitate structural transformation and agricultural productivity growth in China. *World Dev.* 59(C), 505–520.

- Doreian, Patrick, Conti, Norman, 2012. Social context, spatial structure and social network structure. *Soc. Netw.* 34 (1), 32–46.
- Dunbar, Robin I.M., 2018. The anatomy of friendship. *Trends Cogn. Sci.* 22 (1), 32–51.
- Duxbury, Scott W., 2019. Diagnosing multicollinearity in stochastic actor-oriented models. *Open Sci. Framew. (OSF)* 1–56. <https://doi.org/10.17605/OSF.IO/ZA2TS>. Retrieved.
- Entwistle, Barbara, Faust, Katherine, Rindfuss, Ronald R., Kaneda, Toshiko, 2007. Networks and contexts: variation in the structure of social ties. *Am. J. Sociol.* 112 (5), 1495–1533.
- Fan, Shenggen, Chan-Kang, Connie, 2004. Is small beautiful? Farm size, productivity, and poverty in Asian agriculture. *Agric. Econ.* 32 (s1), 135–146.
- Faust, Katherine, 2007. Very local structure in social networks. *Sociol. Methodol.* 37 (1), 209–256.
- Fei, Hsiao-Tung, Chang, Chih-I., 1949. *Earthbound China: A Study of the Rural Economy of Yunnan*, revised edition. Routledge.
- Fengbo, Chen, Pandey, Sushil, Shijun, Ding, 2013. Changing rice cropping patterns: evidence from the Yangtze River Valley, China. *Outlook Agric.* 42 (2), 109–115.
- Finn, William L., Buttel, Frederick H., 1980. Sociological aspects of farm size: ideological and social consequences of scale in agriculture. *Am. J. Agric. Econ.* 62 (5), 946–953.
- Gabry, Jonah, Simpson, Daniel, Vehtari, Aki, Betancourt, Michael, Gelman, Andrew, 2019. Visualization in bayesian workflow. *J. R. Stat. Soc. Ser. A* 182, 1–14.
- Gasson, Ruth, 1977. Farmers' participation in cooperative activities. *Sociol. Ruralis* 17 (1), 107–123.
- Gasparrini, Antonio, Armstrong, Ben, Kenward, Mike, 2012. Multivariate meta-analysis for non-linear and other multi-parameter associations. *Stat. Med.* 31 (29), 3821–3839.
- Gasparrini, Antonio, 2018. Package “Mvmeta”: Multivariate and Univariate Meta-Analysis and Meta-Regression (v. 0.4.11). Retrieved. <https://cran.r-project.org/web/packages/mvmeta/mvmeta.pdf>.
- Gelman, Andrew, 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* 1 (3), 515–534.
- Gelman, Andrew, et al., 2013. *Bayesian Data Analysis*, 3rd ed. CRC Press.
- Gelman, Andrew, Jakulin, Aleks, Pittau, Maria G., Su, Yu-Sung, 2008. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* 2 (4), 1360–1383.
- Ghosh, Joyee, Li, Yingbo, Mitra, Robin, 2018. On the use of Cauchy prior distributions for bayesian logistic regression. *Bayesian Anal.* 13 (2), 359–383.
- Glowacki, Luke, Molleman, Lucas, 2017. Subsistence styles shape human social learning strategies. *Nat. Hum. Behav.* 1 (5) 0098–21.
- Goodreau, Steven M., Kitts, James A., Morris, Martina, 2009. Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks. *Demography* 46 (1), 103–125.
- Guo, Xi, et al., 2018. Drivers of spatio-temporal changes in paddy soil pH in Jiangxi Province, China from 1980 to 2010. *Sci. Rep.* 8 (1), 1–11.
- Harris, William A., 1997. On ‘Scope Conditions’ in sociological theories. *Soc. Econ. Stud.* 46 (4), 123–127.
- Heidler, Richard, Gamper, Markus, Herz, Andreas, Eßer, Florian, 2014. Relationship patterns in the 19th century: the friendship network in a german boys' school class from 1880 to 1881 revisited. *Soc. Netw.* 37, 1–13.
- Henrich, Joseph, Heine, Steven J., Norenzayan, Ara, 2010. The weirdest people in the world? *Behav. Brain Sci.* 33 (2–3), 61–83.
- Ho, Peter, 2001. Who owns China's land? Policies, property rights and deliberate institutional ambiguity. *China Q.* 166, 394–421.
- Hoffman, Matthew D., Gelman, Andrew, 2014. The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 15 (1), 1593–1623.
- Hruschka, Daniel J., 2010. *Friendship: Development, Ecology, and Evolution of a Relationship*. University of California Press.
- Hu, Shihu, Yuan, Zhiguo, 2015. Commentary: ‘Large-scale psychological differences within China explained by rice vs. wheat agriculture’. *Front. Psychol.* 6, 1–3.
- Indlekofer, Natalie, Brandes, Ulrik, 2014. Relative importance of effects in stochastic actor-oriented models. *Netw. Sci.* 1 (3), 278–304.
- Krackhardt, David, Handcock, Mark S., et al., 2007. Heider vs simmel: emergent features in dynamic structures. In: Airoldi, E. (Ed.), *Statistical Network Analysis: Models, Issues, and New Directions*, Vol. 4503, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 14–27.
- Kretschmer, David, Leszczynski, Lars, Pink, Sebastian, 2018. Selection and influence processes in academic achievement—more pronounced for girls? *Soc. Networks* 52, 251–260.
- Krivitsky, Pavel N., Handcock, Mark S., Morris, Martina, 2011. djusting for network size and composition effects in exponential-family random graph models. *Stat. Methodol.* 8 (4), 319–339.
- Kruschke, John K., 2015. *Doing Bayesian Data Analysis*, 2nd ed. Elsevier Inc.
- Kuha, Jouni, Mills, Colin, 2018. On group comparisons with logistic regression models. *Sociol. Methods Res.* 3 004912411774730–28.
- Leszczynski, Lars, Pink, Sebastian, 2015. Ethnic segregation of friendship networks in school: testing a rational-choice argument of differences in ethnic homophily between classroom- and grade-level networks. *Soc. Netw.* 42, 18–26.
- Liu, Fang, Eugenio, Evercita C., 2016. A review and comparison of bayesian and likelihood-based inferences in Beta regression and zero-or-one-inflated beta regression. *Stat. Methods Med. Res.* 16 096228021665069–21.
- Lowder, Sarah K., Skoet, Jakob, Raney, Terri, 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Dev.* 87(C), 16–29.
- Lubbers, Miranda J., Snijders, Tom A.B., 2007. A comparison of various approaches to the exponential random graph model: a reanalysis of 102 student networks in school classes. *Soc. Networks* 29 (4), 489–507.
- Lusher, Dean, Koskinen, Johan, Robins, Garry, 2012. In: Lusher, D., Koskinen, J., Robins, G. (Eds.), *Exponential Random Graph Models for Social Networks: Theory, Methods, and Application*. Cambridge University Press.
- Matous, Petr, 2015. Social networks and environmental management at multiple levels: soil conservation in Sumatra. *Ecol. Soc.* 20 (3), art37–17.
- Matous, Petr, Wang, Peng, 2019. External Exposure, boundary-spanning, and opinion leadership in remote communities: a network experiment. *Soc. Netw.* 56, 10–22.
- Mayhew, Bruce C., Levinger, Roger L., 1976. Size and the density of interaction in human aggregates. *Am. J. Sociol.* 82 (1), 86–110.
- McFarland, Daniel A., Moody, James, Diehl, David, Smith, Jeffrey A., Thomas, Reuben J., 2014. Network ecology and adolescent social structure. *Am. Sociol. Rev.* 79 (6), 1088–1121.
- McLean, Paul D., 2016. *Culture in Networks*. Polity.
- Minhas, Shahryar, Hoff, Peter D., Ward, Michael D., 2019. Inferential approaches for network analysis: AMEN for latent factor models. *Polit. Anal.* 12, 1–15.
- Molm, Linda D., 2010. The structure of reciprocity. *Soc. Psychol. Q.* 1–13.
- Monge, Peter R., Contractor, Noshir S., 2003. *Theories of Communication Networks*. Oxford University Press.
- Moody, James, White, Douglas R., 2003. Structural cohesion and embeddedness: a hierarchical concept of social groups. *Am. Sociol. Rev.* 68 (1), 103–126.
- Niekerk, Janet V., Bekker, Andriette, Arashi, Mohammad, 2018. Beta regression in the presence of outliers—a wieldy bayesian solution. *Stat. Methods Med. Res.* 9 096228021881457–12.
- Nolin, David A., 2010. Food-sharing networks in Lamalera, Indonesia. *Hum. Nat.* 21 (3), 243–268.
- Norton, Edward C., Dowd, Bryan E., 2017. Log odds and the interpretation of logit models. *Health Serv. Res.* 53 (2), 859–878.
- Norton, Edward C., Dowd, Bryan E., Maciejewski, Matthew L., 2018. Odds ratios—current best practice and use. *J. Am. Med. Assoc.* 320 (1), 84–85.
- Oi, Jean C., Rozelle, Scott, 2009. Elections and power: the locus of decision-making in Chinese villages. *China Q.* 162, 513–527.
- Oishi, Shigehiro, 2014. Socioecological psychology. *Annu. Rev. Psychol.* 65 (1), 581–609.
- Oishi, Shigehiro, Kesebir, Selin, 2012. Optimal social-networking strategy is a function of socioeconomic conditions. *Psychol. Sci.* 23 (12), 1542–1548.
- Ospina, Raydonal, Ferrari, Silvia L.P., 2010. Inflated beta distributions. *Stat. Pap.* 51 (1), 111–126.
- Ospina, Raydonal, Ferrari, Silvia L.P., 2012. A general class of zero-or-one inflated beta regression models. *Comput. Stat. Data Anal.* 56 (6), 1609–1623.
- Otsuka, Keijiro, Liu, Yanyan, Yamauchi, Futoshi, 2016. The future of small farms in Asia. *Dev. Policy Rev.* 34 (3), 441–461.
- Peng, Yusheng, 2004. Kinship networks and entrepreneurs in China's transitional economy. *Am. J. Sociol.* 109 (5), 1045–1074.
- Pepinsky, Thomas B., 2018. A note on listwise deletion versus multiple imputation. *Polit. Anal.* 1–9.
- Perkins, Jessica M., Subramanian, S.V., Christakis, Nicholas A., 2015. Social networks and health: a systematic review of sociocentric network studies in low- and middle-income countries. *Soc. Sci. Med.* 125(c), 60–78.
- Power, Eleanor A., 2017. Social support networks and religiosity in Rural South India. *Nat. Hum. Behav.* 1 (3), 1–6.
- Raabe, Isabel J., Boda, Zsófia, Stadtfeld, Christoph, 2019. The social pipeline: how friend influence and peer exposure widen the STEM gender gap. *Sociol. Educ.* 92 (2), 105–123.
- Raudenbush, Stephen W., 2009. Analyzing effect sizes: random-effects models. In: Cooper, H., Hedges, L.V., Valentine, J. (Eds.), *Handbook of Research Synthesis and Meta-Analysis*. The Russel Sage Foundation.
- Raudenbush, Stephen W., Bryk, Anthony S., 1985. Empirical bayes meta-analysis. *J. Educ. Stat.* 10 (2), 75–98. <https://doi.org/10.3102/10769986010002075>. Retrieved.
- Ripley, Ruth M., Snijders, Tom A.B., Boda, Zsófia, Vörös, András, Preciado, Paulina, 2019. *Manual for RSiena (v. 1.2-17)*. Retrieved. University of Oxford: Department of Statistics, Nuffield College and University of Groningen: Department of Sociology. https://www.stats.ox.ac.uk/~snijders/siena/RSiena_Manual.pdf.
- Rivera, Mark T., Soderstrom, Sara B., Uzzi, Brian, 2010. Dynamics of dyads in social networks: assortative, relational, and proximity mechanisms. *Annu. Rev. Sociol.* 36 (1), 91–115.
- Robins, Garry, Pattison, Philippa, Woolcock, Jodie, 2005. Small and other worlds: global network structures from local processes. *Am. J. Sociol.* 110 (4), 894–936.
- Schaefer, David R., Light, John M., Fabes, Richard A., Hanish, Laura D., Martin, Carol L., 2010. Fundamental principles of network formation among preschool children. *Soc. Netw.* 32 (1), 61–71.
- Shakya, Holly B., Christakis, Nicholas A., Fowler, James H., 2017. An exploratory comparison of name generator content: data from rural India. *Soc. Netw.* 48, 157–168.
- Simpson, Brent, Harrell, Ashley, Melamed, David, Heiserman, Nicholas, Negraia, Daniela V., 2017. The roots of reciprocity: gratitude and reputation in generalized exchange systems. *Am. Sociol. Rev.* 83 (1), 88–110.
- Small, Mario L., Adler, Laura, 2019. The role of space in the formation of social ties. *Annu. Rev. Sociol.* 45 (1) annurev-soc-073018-022707–22.
- Smith, Sanne, McFarland, Daniel A., van Tubergen, Frank, Maas, Ineke, 2016. Ethnic composition and friendship segregation: differential effects for adolescent natives and immigrants. *Am. J. Sociol.* 121 (4), 1223–1272.
- Smithson, Michael, Verkuilen, Jay, 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychol. Methods* 11 (1), 54–71.
- Snijders, Tom A.B., 1996. Stochastic actor-oriented models for network change. *J. Math. Sociol.* 21 (1–2), 149–172.
- Snijders, Tom A.B., 2001. The statistical evaluation of social network dynamics. *Sociol. Methodol.* 31 (1), 361–395.

- Snijders, Tom A.B., Baerveldt, Chris, 2003. A multilevel network study of the effects of delinquent behavior on friendship evolution. *J. Math. Sociol.* 27 (2–3), 123–151.
- Snijders, Tom A.B., Steglich, Christian E.G., 2015. Representing micro–macro linkages by actor-based dynamic network models. *Sociol. Methods Res.* 44 (2), 222–271.
- Stedman, Richard C., 2003. Is it really just a social construction?: the contribution of the physical environment to sense of place. *Soc. Nat. Resour.* 16 (8), 671–685.
- Sutton, Mark Q., Anderson, Eugene N., 2009. *Introduction to Cultural Ecology*. Rowman & Littlefield.
- Talhelm, Thomas, et al., 2014. Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science* 344 (6184), 603–608.
- Talhelm, Thomas, 2020. Emerging evidence of cultural differences linked to rice versus wheat agriculture. *Curr. Opin. Psychol.* 32, 81–88.
- Talhelm, Thomas, Oishi, Shigehiro, 2018. How rice farming shaped culture in Southern China: how rice farming shaped culture in Southern China. In: Üskül, A.K., Oishi, S. (Eds.), *Socio-Economic Environment and Human Psychology Social, Ecological, and Cultural Perspectives*, pp. 53–76.
- Tan, Shuhao, Heenrink, Nico, Kruseman, Gideon, Qu, Futian, 2008. Do fragmented landholdings have higher production costs? Evidence from rice farmers in Northeastern Jiangxi Province, P.R. China. *China Econ. Rev.* 19 (3), 347–358.
- Thomson, Robert, et al., 2018. Relational mobility predicts social behaviors in 39 countries and is tied to historical farming and threat. *Proc. Natl. Acad. Sci. U. S. A.* 7 201713191–83.
- Tilt, Bryan, 2008. Smallholders and the ‘Household responsibility system’: adapting to institutional change in Chinese agriculture. *Hum. Ecol.* 36 (2), 189–199.
- Train, Kenneth, 2009. *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge University Press.
- Uchida, Yukiko, et al., 2018. Farming cultivates a community-level shared culture through collective activities: examining contextual effects with multilevel analyses. *J. Pers. Soc. Psychol.* 1–15.
- Vaisey, Stephen, Lizardo, Omar, 2010. Can cultural worldviews influence network composition? *Soc. Forces* 88 (4), 1595–1618.
- Van den Berg, M.M., et al., 2007. The impact of increasing farm size and mechanization on rural income and rice production in Zhejiang Province, China. *Agric. Syst.* 94 (3), 841–850.
- van der Ploeg, Jan D., Ye, Jingzhong, Pan, Lu, 2014. Peasants, time and the land: the social organization of farming in China. *J. Rural Stud.* 36(C), 172–181.
- Vehtari, Aki, Gelman, Andrew, Simpson, Daniel, Carpenter, Bob, Bürkner, Paul-Christian, 2019. Rank-Normalization, Folding, and Localization: An Improved \hat{R} for Assessing Convergence of MCMC. *arXiv* 1–23. Retrieved. <https://arxiv.org/abs/1903.08008>.
- Vendryes, Thomas, 2010. Land Rights in Rural China Since 1978: reforms, successes, and shortcomings. *China Perspect.* 2010 (4), 1–14.
- Wencélius, Jean, Thomas, Mathieu, Barbillon, Pierre, Garine, Eric, 2016. Interhousehold variability and its effects on seed circulation networks: a case study from Northern Cameroon. *Ecol. Soc.* 21 (1) 44–12.
- Wimmer, Andreas, Lewis, Kevin, 2010. Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *Am. J. Sociol.* 116 (2), 583–642.
- Wu, Bin, Pretty, Jules, 2004. Social connectedness in marginal rural China: the case of farmer innovation circles in Zhidan, North Shaanxi. *Agric. Human Values* 21 (1), 81–92.
- Wu, Yiyun, et al., 2018. Policy distortions, farm size, and the overuse of agricultural chemicals in China. *Proc. Natl. Acad. Sci. U. S. A.* 115 (27), 7010–7015.
- Xiao, Zhixing, Tsui, Anne S., 2007. When brokers may not work: the cultural contingency of social capital in Chinese high-tech firms. *Adm. Sci. Q.* 52 (1), 1–31.
- Zhang, Fusuo, Chen, Xinping, Vitousek, Peter, 2013. Chinese agriculture: an experiment for the world. *Nature* 497 (7447), 33–35.
- Zhang, Weifeng, et al., 2016. Closing yield gaps in China by empowering smallholder farmers. *Nature* 537 (7622), 671–674.