Network Structure and Public Good Provision

Evidence from Village Kinship Networks in the Gambia

Gift Tontarawongsa
Duke University

June 16rh, 2014

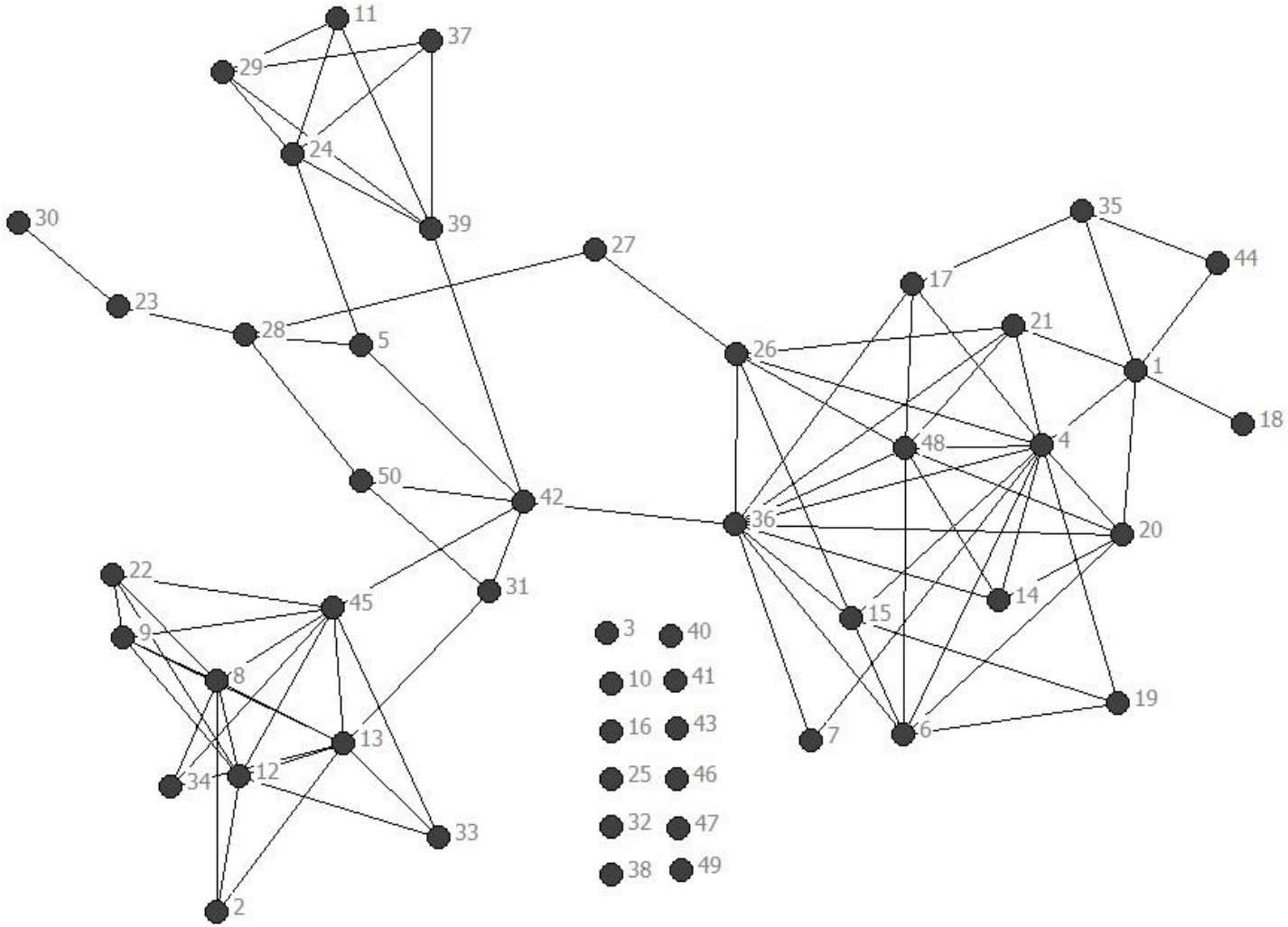
Overview

- Free-riding problem in public goods.
- Informal cooperation is prevalent especially in developing countries.
- Does social network actually play a role in sustaining this cooperation?
- Theory literature suggests:
  - Social capital helps mitigate this market failure.
  - Game theoretical models of cooperative games in networks.
- Lack of causal evidence due to limited data + endogeneity problem
- This paper provides field-based evidence of how networks affect decisions to contribute to public goods.

Research Questions

- How does social network position affect public good contribution?
  - Do better-connected agents in a network contribute more?
  - At the larger level, does a better-connected network facilitate a higher level of contribution?
- Estimate the effect of household network position on public good contribution using kinship network data from rural villages in the Gambia
- Address the endogenous network formation problem using the arguably exogenous variation in village ethnic composition as an IV.

Village Network Example



Theoretical motivation

- Game theory literature on cooperative games in networks
- No consensus on the predictions of the relationship
- Why might household network position affect public good contribution?
  - Facing higher social pressure, threat of sanction
  - Being monitored more, easier to get caught shirking
  - Position serves as a proxy for household valuation of public goods.
  - Private information about valuation is better known in the community
- Do not distinguish between these mechanisms in this paper

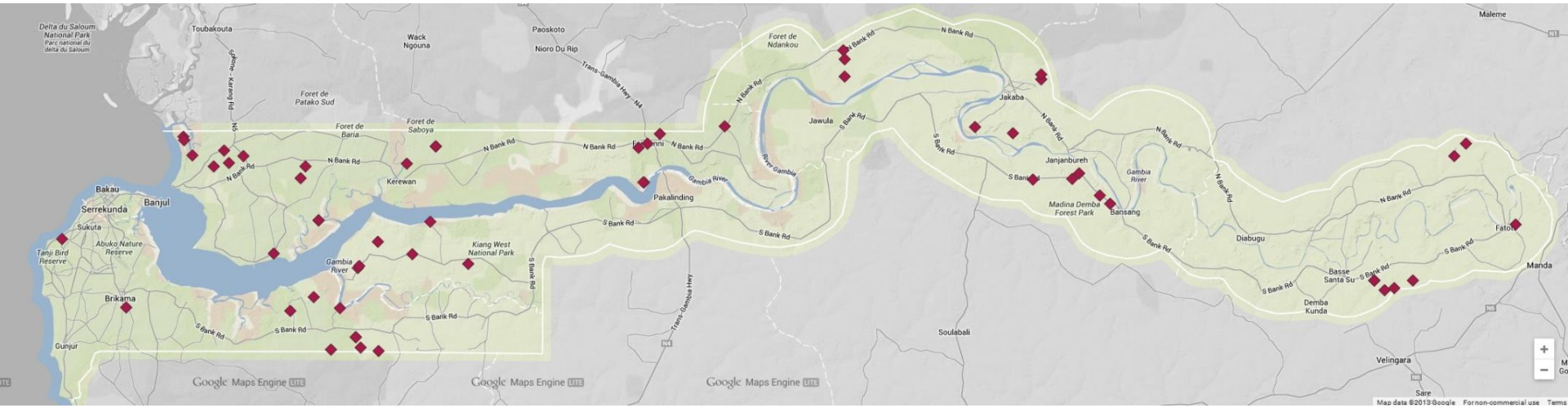
Network centrality measures

To capture the "position" in networks and relate it to public good contribution,

- Degree Centrality
  - Number of links
- Eigenvector centrality
  - Weighs in centrality of the connected nodes.
  - Idea: being connected to a well-connected person makes me *more central.*
  - Correspond to an eigenvector of a matrix representation of a network.

Data Overview

- Analysis sample: 2,110 households from 48 rural villages in the Gambia
- A baseline household survey for the impact evaluation of the Gambia Community-Driven Development Project (CDDP) in Feb-May of 2009.
- Shared with me by Dany Jaimovich of Goethe University Frankfurt



Public Goods in Rural Gambia

- "Did anyone in your household participate in the following *activities for the village* in the past one year?" – binary outcomes

Public good contribution	Fraction of households contributing	Std. Dev.
Road construction	58%	(0.49)
Communal building construction	71%	(0.46)
Water well construction	48%	(0.50)
Public health	67%	(0.47)
Public education	48%	(0.50)
Total contribution (sum all above)	3.1	(1.29)

- Form of contribution
  - in-kind contribution 99% and monetary contribution 1%

Public Goods in Rural Gambia

- Pattern of public goods
  - High correlation within the village
  - Cases where contribution to a particular public good is 0 within the village.
- Government and NGO involvement
  - A school building project
- Focus group interview report
  - Universal access at the village level

Village Kinship Networks

- **Household-level** relationships
  - "Kin network" – blood relatives of the household head and wives.

Variable	Mean	Std. Dev.	Min	Max
Degree centrality, Kin	3.99	(3.31)	0	15
Eigenvector centrality, Kin	0.76	(0.70)	0	5.09

Dist.

Endogeneity of network centrality

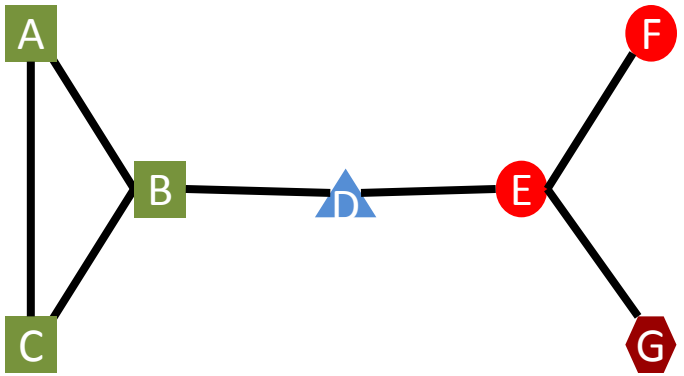
- Observed relationship could be driven by underlying endogenous network formation (Manski, 1993)
- Households may sort based on unobserved characteristics.
- Other potential sources of bias
  - Omitted variables - preferences for public goods.
  - Reverse causality

Identification Strategy

- Having more households of the same ethnic group makes it easier to form kinship networks because of the patrilineal culture and arranged marriages.
- IV for household centrality:

"household ethnic fraction"

  - The fraction of households within the village that belong to the same ethnic group as a given household.



ID	Ethnic fraction
A	3/7
B	3/7
C	3/7
D	1/7
E	2/7
F	2/7
G	1/7

Identification Strategy

- The Gambia is highly diverse.

Ethnicity	% of sample	% 2003 census
Mandinka	48	36
Fula	22	22
Wolof	10	14
Jola	7	11
Serahule	2	8
Serer	5	3
Manjago	1	2
Others	5	4

- Exploit the arguably exogenous variation in village ethnic composition determined by historical accident.
- Ethnic identity is not salient (Mwakikagile (2010)).
  - kin links (17% across ethnic groups)
  - Not a determinant of economic exchange relationships (Arcand and Jaimovich, 2013)

Identification Strategy

- **Within the same ethnicity, comparing households in villages with variation in ethnic fraction.**
- Concerns:
  1. Variation is not random and different types of households sort into villages based on this variation.
  2. Ethnic fraction affects public good contribution directly.

Validity of the IV - I

- Concern: different types of households sort into villages with large or small ethnic representation.
- Balance check.
- Obs = 2,110

Variable:	Household Ethnic fraction
Log income	-0.224
Household size	0.038
Head age	-0.042
Female head	-0.395
HH head illiterate	1.484
# sick members	-0.693
# workers	0.066
lands owned (hectares)	0.021
Self-reported high quality land	0.693
Role: village chief	1.762
Role: village development committee	1.700
Role: traditional healer	0.364
Role: Imam	4.060**
Role: Marabout	2.383
Farmer	1.079
Muslim	1.917
Marital status1: unmarried	2.171
Marital status2: monogamous	0.318
Marital status3: polygamous	0.339
Relative wealth: mid low	1.009
Relative wealth: mid high	0.372
Relative wealth: high	-1.975
Agricultural Income	-0.022
Number of corrugated huts	0.640
Missing number of corrugated huts	11.617
Number of grass-roof huts	-0.709
Missing number of grass-roof huts	-11.650

Validity of the IV - II

- Limited migration in rural Gambia – Migration Reports (GBoS, 1993, 2009)
  - Internal migration in urban Kanifing and Banjul areas (not in the data)
- Village-level ethnic composition data (GBoS census, 1993, 2003)
  - Ethnic fraction remain stable since 1993.
  - 8 out of the 60 villages had large changes in a few ethnic groups
    - Close to the Banjul areas
- No data before 1993, anthropological and historical evidence
  - Gamble (1949), Sonko-Godwin (1985, 1988)

Validity of the IV - III

Ethnic fraction should affect public good contribution only through networks.

- Direct ethnic conflict
  - Mwakikagile (2010), Sonko-Godwin (1985, 1988)
  - Arcand and Jaimovich (2012) – ethnicity is not related to econ-exchange.
- Diverged preferences over the type of public goods to be funded.
  - —Regressing different public goods on ethnicity – not significant

Validity of the IV - IV

Ethnic fraction should affect public good contribution only through networks.

$$\text{contribution}_j = \beta_0 + \beta_1 ELF_j + \gamma X_j + \varepsilon_j$$

Outcome variable	Roads	Commu. Buildings	Water wells	Public health	Public education	Total
Panel A: OLS estimates without controlling for network density						
Ethnic diversity index	-0.032	-0.037	0.127	-0.165	-0.113	-0.549
	(0.17)	(0.18)	(0.15)	(0.17)	(0.19)	(0.53)
Panel B: OLS estimates controlling for network density						
Ethnic diversity index	0.012	-0.059	0.043	-0.064	-0.104	-0.313
	(0.17)	(0.20)	(0.15)	(0.15)	(0.19)	(0.49)
Network density	0.801*	-0.383	-1.496***	1.818***	0.163	4.253**
	(0.47)	(0.82)	(0.43)	(0.55)	(0.92)	(1.98)
Mean outcomes	0.53	0.61	0.45	0.58	0.39	2.69

Observations = 60 villages. Controlling for: number of households, total population, areas, log per capita income, income Gini coeffcient, illiteracy rate, % of households with no electricity, ethnic diversity (Herfindahl index), religion diversity (Herndahl index), district FE.

Results: First stage

$$\text{centrality}_i = \alpha_0 + \alpha_1 \text{ethnic\_fraction}_i + \alpha_2 X_i + \delta_j + \phi_v + u_i$$

Endogenous variable:	Degree Cent.	Eigenvector cent.
IV: household ethnic fraction	3.956***	0.874***
	(0.506)	(0.090)
Constant	-1.293	-0.211
	(1.539)	(0.503)
First-stage F statistics	59.08	70.31
R-squared	0.388	0.172
Observations	2,110	

Robust standard errors clustered at the village level in parenthesis. Controlling for: ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, status as the village chief, traditional healer, village development committee, being a religious leader, religion, marital status, number corrugated huts, and number of grass-roof huts.

Graph

2nd Stage Specification

$$\text{contribution}_i = \alpha + \beta \widehat{\text{centrality}}_i + \gamma X_i + \delta_j + \theta_v + \varepsilon_i$$

- contribution = 5 public good types, total number of contribution
- centrality
  - Degree centrality
  - Eigenvector centrality
- X_i = household characteristics
- δ_j = ethnic group fixed effects
- θ_v = village/network fixed effects

Results: Degree centrality

$$contribution_i = \alpha + \beta \widehat{degree}_i + \gamma X_i + \delta_j + \theta_v + \varepsilon_i$$

Public good :	(1) Road	(2) Com. Buildings	(3) Water Well	(4) Public health	(5) Public educ	(6) Total cont.
OLS estimates						
Degree centrality	-0.000	0.005	0.016***	-0.002	0.009**	0.009
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.010)
IV estimates						
Degree centrality	0.006	0.016	0.028*	0.009	0.011	0.050**
	(0.012)	(0.013)	(0.014)	(0.009)	(0.011)	(0.023)
First stage F stat	59.08	59.08	59.08	59.08	59.08	59.08
Mean of outcomes	0.61	0.71	0.50	0.67	0.44	3.1

Sample size: 2,110, #Villages = 48

Robust standard errors clustered at the village level in parenthesis. Controlling for: ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, status as the village chief, traditional healer, village development committee, being a religious leader, religion, marital status, number of corrugated huts, number of grass-roof huts.

Probit

Results: Eigenvector centrality

$$contribution_i = \alpha + \beta \widehat{eig\_cent}_i + \gamma X_i + \delta_j + \theta_v + \varepsilon_i$$

Public good :	(1) Road	(2) Com. Buildings	(3) Water Well	(4) Public health	(5) Public educ	(6) Total cont.
IV estimates						
Eigenvector centrality	0.028	0.071	0.127*	0.040	0.048	0.227**
	(0.056)	(0.056)	(0.069)	(0.040)	(0.050)	(0.097)
Mean outcome	0.61	0.71	0.50	0.67	0.44	3.1

Sample size: 2,110, #Villages = 48

Robust standard errors clustered at the village level in parenthesis. Controlling for: ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, status as the village chief, traditional healer, village development committee, being a religious leader, religion, marital status, number of corrugated huts, number of grass-roof huts.

Interpreting the magnitudes

Public good :	(1) Road	(2) Com. Buildings	(3) Water Well	(4) Public health	(5) Public educ	(6) Total cont.
IV estimates						
Eigenvector centrality	0.028	0.071	0.127*	0.040	0.048	0.227**
	(0.056)	(0.056)	(0.069)	(0.040)	(0.050)	(0.097)
Interpret the magnitudes						
25 percentile	0.006	0.016	0.028	0.009	0.011	0.050
50 percentile	0.017	0.042	0.075	0.024	0.028	0.134
75 percentile	0.032	0.082	0.146	0.046	0.055	0.261
99 percentile	0.081	0.206	0.368	0.116	0.139	0.658
Mean outcome	0.61	0.71	0.50	0.67	0.44	3.1

Eigenvector centrality can be interpreted as a weighted sum of walks that starts from a given node, with longer walks discounted more.

Eigenvector centrality, mean = 0.76

Robustness check: additional controls

- Binary variables – Existence of relationships outside of the village.

Variable	Mean	Std. Dev.
Land exchange	0.14	(0.35)
Labor exchange	0.06	(0.25)
Input exchange	0.10	(0.31)
Credit exchange	0.15	(0.36)
Marriage	0.85	(0.35)

- Enumerator fixed effects
  - Address potential systematic measure errors

Robustness check: additional controls

$$contribution_i = \alpha + \beta \widehat{degree}_i + \gamma X_i + X_{2i} + \delta_j + \theta_v + \varepsilon_i$$

Public good contribution:	(1) Road	(2) Com. Buildings	(3) Water Well	(4) Public health	(5) Public educ	(6) Total cont.
Main specification						
Degree centrality, Kin	0.006	0.016	0.028*	0.009	0.011	0.050**
	(0.012)	(0.013)	(0.014)	(0.009)	(0.011)	(0.023)
Robustness check						
Degree centrality, Kin	0.012	0.021	0.029*	0.010	0.012	0.050**
	(0.014)	(0.013)	(0.015)	(0.010)	(0.010)	(0.023)
Mean of outcomes	0.61	0.71	0.50	0.67	0.44	3.1

Sample size: 2,110, #Villages = 48

Robust standard errors clustered at the village level in parenthesis. Controlling for: ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, status as the village chief, traditional healer, village development committee, being a religious leader, religion, marital status, number of corrugated huts, number of grass-roof huts.

Robustness check: expanded sample

Public good contribution:	(1) Road	(2) Com. Buildings	(3) Water Well	(4) Public health	(5) Public educ	(6) Total cont.
Stable villages (Obs = 2,110, village = 48)						
Degree centrality, Kin	0.006	0.016	0.028*	0.009	0.011	0.050**
	(0.012)	(0.013)	(0.014)	(0.009)	(0.011)	(0.023)
Mean of outcomes	0.61	0.71	0.50	0.67	0.44	3.1
Full sample (Obs = 2,697, Village = 60)						
Degree centrality, Kin	0.006	0.019*	0.031***	0.006	0.021**	0.058**
	(0.010)	(0.010)	(0.011)	(0.008)	(0.010)	(0.024)
Mean of outcomes	0.58	0.71	0.48	0.67	0.48	3.1

Robust standard errors clustered at the village level in parenthesis. Controlling for: ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, status as the village chief, traditional healer, village development committee, being a religious leader, religion, marital status, number of corrugated huts, and number of grass-roof huts.

Alternative specifications

- Possible nonlinear relationships
  - Bandiera and Rasul (2006) – inverse U shape
  - Specifications:
    1. add a squared term of network centrality.
    2. I also break down degree centrality into a set of dummy variables.
  - Little evidence of non-linear relationship.
- Linear peer effects specification (Manski, 1993)
  - Positive relationship between contributions of connected household.

Nonlinear *Peer*

Other types of networks

- Economic networks - four types of economic transactions in the past one year
  - Some positive correlation but the IV doesn't pass the weak IV test.
  - OLS estimates: small correlations with contribution. Only significant for communal building construction.
- Combined networks – both kin and economic networks
  - IV works well, possibly because of the kin networks.
  - IV results: similar patterns with larger magnitudes than the kin IV results.
  - Interpret with caution because of the weak IV problem for the economic networks.

Table

Conclusion

- Central agents contribute to public goods more.
  - Support the models in the game theory literature
  - Consistent with punishment in an experiment in Carpenter et al. (2012)
- Networks may explain why ethnic fragmentation lowers public goods
  - Unexplained cases like China, Haiti (Esteban et al., 2012)

Village-level