**Tribe or Nation? Nation Building and
Public Goods in Kenya versus Tanzania**

Edward Miguel   Replication Materials - Supplement

This serves as a supplement to the published paper by Edward Miguel, *Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania,* World Politics, (2004). It outlines discrepancies found in data replication which do not substantially change reported data analysis as published in the paper, as a whole. While some estimates change, statistical significance levels for the main estimates are unchanged. Most discrepancies arise in total local public expenditures.

In the process of organizing replication files, we noticed a discrepancy in the number of observations in the dataset for Tanzania. In this set of replication files, we have a dataset with 65 observations as opposed to 66 observations as was used in publication. We attribute this to disorganization of replication files over time, and believe that over time an observation was omitted or dropped out of the dataset and at this point is not recoverable. We have taken extensive steps to attempt to reconcile this, as follows: 1) We duplicated each observation in the current dataset to replicate analysis to see if analysis was replicated as published which would have told us that one observation was duplicated in original analysis and that there was only 65 observations in the original dataset. Upon doing this, none of the analysis replicated the exact same as is published. 2) We traced back steps in data cleaning to reconcile any loss of observations over time but were unable to recover the observation.

In all, we have the original log file from analysis that shows analysis run with 66 observations that we present in our replication file set in this folder. The original log file is in our replication folder, titled **"tz-ethnic_jun04_Original-Log-File.txt"** which shows output of regression analysis with full 66 observations.

In an effort to reconcile discrepancies, we present the differences in analysis from the replication files in this folder as opposed to those published in the tables listed below. The published analysis is highlighted, and replicated analysis listed in red nearby.

The data replication generated discrepancies in **Table 1:** *Descriptive Statistics.*

TABLE 1
DESCRIPTIVE STATISTICS

	Meatu District, Tanzania			Busia District, Kenya		
Panel A: Data for Tanzania and Kenya	Mean	Std Dev.	Obs. Villages	Mean	Std Dev.	Obs. Schools
Local ethnic fractionalization (ELF)	0.13	0.16	66 65	0.23	0.14	84
Average years of education	4.1	1.1	66 65	7.4	1.3	84
Proportion formal sector employment	0.06	0.07	66 65	0.23	0.10	84
Proportion of homes with iron roofs	0.26	0.20	66 65	0.24	0.10	84
Proportion households grow cash crops	0.61	0.25	66 65	0.40	0.24	84
Proportion households own cattle	0.47	0.17	66 65	0.60	0.19	84
Proportion Catholic	0.17	0.12	66 65	0.58	0.22	84
Annual expenditures on local primary school projects per pupil (U.S.$)	4.88 4.28	3.90 2.72	66 65	3.45	2.23	84
Desks per primary school pupil	0.19	0.09	66 65	0.21	0.12	84
Latrines per primary school pupil	0.011	0.005	66 65	0.016	0.013	84
Classrooms per primary school pupil	0.013	0.005	66 65	0.031	0.014	84
Teachers per primary school pupil	0.013	0.004	66 65	0.039	0.015	84
Proportion wells with "normal water flow"	0.57	0.37	66 65	0.56	0.14	84

Panel B: Data for Tanzania						
Number of households per village	413.4	178.8	66 65	–	–	–
Annual per capita consumption expenditures (U.S.$)	198.4	81.2	66 65	–	–	–
Gini coefficient of annual per capita consumption expenditures (at village level)	0.36	0.15	66 65	–	–	–
Annual local expenditures on all public goods projects, per household (U.S.$)	8.65 7.77	6.39 4.53	66 65	–	–	–
Annual local expenditures on health and water well projects, per household (U.S.$)	1.51 1.48	1.78 1.74	66 65	–	–	–
Annual local tax collection, per household (U.S.$)	2.14	3.47	66 65	–	–	–
Average number of completed local public goods project, per year	0.67	0.40	66 65	–	–	–
Average household spending on local taxes and school expenses (U.S.$) [HH Survey]	12.3	14.5	66 65	–	–	–
Wells with normal water flow, per household	0.008	0.009	66 65	–	–	–
Average road quality (scale 1–4)	2.5	0.8	65 64	–	–	–
Total community groups, per household	0.026	0.017	66 65	–	–	–

TABLE 1 (cont.)

Panel B: Data for Tanzania	Meatu District, Tanzania			Busia District, Kenya		
	Mean	Std Dev.	Obs. Villages	Mean	Std Dev.	Obs. Schools
Community group memberships [HH survey]	2.25	0.93	66 65			
Proportion survey respondents who are community group members [HH Survey]	0.33	0.23	66 65	–	–	–
Village meeting attendance, per household	1.22	1.03	66 65	–	–	–
Proportion households attending a village meeting [HH Survey]	0.94	0.07	66 65			
In general, can you trust people in other tribes?	0.83	0.11	66 65			
Spirit of cooperation across tribes in village, proportion stating "above average"	0.59	0.11	66 65			
Village unity, proportion stating "above average"	0.65	0.14	66 65			
Nation is more important than tribe to respondent	0.40	0.14	66 65			

The data replication generated discrepancies in **Table 2:** *Ethnic Diversity and Local Public Goods: Kenya and Tanzania.* Most discrepancies are found in the Seemingly Unrelated Regression (SUR), found as "Prob > F" in regression output. There are some differences between the original log file, new log file, and published results, which we attribute to rounding. As the reader can examine, the rounding differences can account for many of the differences observed in the presented analysis across files.

TABLE 2
ETHNIC DIVERSITY AND LOCAL PUBLIC GOODS: KENYA AND TANZANIA[a]

Explanatory Variable	Annual School Spending/ Pupil, U.S.$ (1)	Desks/ Pupil (2)	Latrines/ Pupil (3)	Classrooms/ Pupil (4)	Proportion Wells with Normal Flow (5)	$H_0: \beta = 0$ F-statistic p-value (SUR)[b]
Socioeconomic Controls						
Average years	0.52	0.013	0.0013***	0.0013*	−0.083**	0.08*
of education	(0.55)	(0.011)	(0.0004)	(0.0007)	(0.037)	0.15
Proportion formal	−11.0	0.30*	0.015**	0.016	−0.31	0.38
sector employment	(9.0)	(0.17)	(0.006)	(0.010)	(0.58)	0.60
Proportion homes	−1.9	−0.05	−0.006***	−0.002	0.12	0.82
with iron roofs	(2.3)	(0.07)	(0.002)	(0.004)	(0.18)	
Proportion households	−0.8	−0.03	0.000	−0.001	−0.12	0.96
grow cash crops	(2.2)	(0.04)	(0.002)	(0.002)	(0.18)	0.36
Proportion households	−2.6	−0.05	0.011***	−0.002	−0.27	0.37
own cattle	(2.5)	(0.05)	(0.003)	(0.005)	(0.24)	0.31
Proportion Catholic	1.9	−0.06	−0.003	−0.011	−0.64	0.37
	(3.3)	(0.09)	(0.003)	(0.009)	(0.52)	
Socioeconomic controls* Kenya Indicator	Yes	Yes	Yes	Yes	Yes	
R^2	0.15	0.19	0.13	0.41	0.19	
Root MSE	3.07	0.098	0.011	0.011	0.25	
Number of observations	150	150	150	150	150	
Ethnic diversity	−3.6*	−0.32**	−0.007	−0.008	−0.06	0.02**
effect, Kenya	(2.0)	(0.12)	(0.012)	(0.010)	(0.19)	

The data replication generated discrepancies in **Table 3:** *Local Public Finance, Collective Action, and Social Capital Outcomes: Tanzania.*

TABLE 3
LOCAL PUBLIC FINANCE, COLLECTIVE ACTION, AND SOCIAL CAPITAL
OUTCOMES: TANZANIA[a]

Dependent Variable	Coefficient Estimate on Local ELF[b]	
Panel A: Public Finance Outcomes, 2001–2 Village Council, Household Data		
Annual total local expenditures on all public goods projects, per household (U.S.$)	7.0 (8.3)	3.7 (5.4)
Annual local expenditures on health and water well projects, per household (U.S.$)	0.5 (1.3)	0.6 (1.2)
Total annual local tax collection, per household	0.6 (2.7)	
Average number of completed local public goods project, per year	–0.33 (0.42)	-0.43
Average household spending on local taxes and school expenses (U.S.$)	7.3 (11.0)	11.2 (11.3)