Testing ELF as an Instrumental Variable

Prepared by: Aashish Panta ’26, Swarthmore College

Table of Contents

[Institutional Capacity Breakdown 3](#_Toc173751373)

[Cooperation and Coordination Breakdown 4](#_Toc173751374)

[Variables and Date Sources 5](#_Toc173751375)

[IV Outline by Prof O’Connell 7](#_Toc173751376)

[Rethinking Covariates 7](#_Toc173751377)

[Reference from other papers 9](#_Toc173751378)

[Understanding IV in context of this research 11](#_Toc173751379)

[Should local cooperation be logged? 11](#_Toc173751380)

[Caste/Ethnicity in Nepal 13](#_Toc173751381)

[Main regression with ELF 13](#_Toc173751382)

[Visuals of Political Affinity 14](#_Toc173751383)

[Reduced form Regression 15](#_Toc173751384)

[First Stage Regression 16](#_Toc173751385)

[IV Validity 17](#_Toc173751386)

[Two-stage least squares in Stata 18](#_Toc173751387)

[Way Forward 18](#_Toc173751388)

# Institutional Capacity Breakdown

* Local Government Institutional Capacity Self-Assessment (LISA):
  + An assessment metric introduced by Ministry of Federal Affairs and General Administration in 2020/21
  + Focused on day-to-day performance and overall quantitative accomplishments
* LISA Breakdown:

|  |  |  |
| --- | --- | --- |
| **SN** | **Category Title** | **Points** |
| 1 | Governance Management | 9 |
| 2 | Organization & Administration | 8 |
| 3 | Budget Plan Management | 11 |
| 4 | Fiscal Economic Management | 11 |
| 5 | Service Delivery | 16 |
| 6 | Judicial Execution | 9 |
| 7 | Physical Infrastructure | 13 |
| 8 | Social Inclusion | 10 |
| 9 | Environmental Protection and Disaster Management | 10 |
| 10 | Cooperation and Coordination | 6 |
|  | **Total** | **100** |

* Each of the LISA Category is weighted with sub-categories as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Indicators** | **Process Scenario** | **Quantitative Scenario** | **Total** |
| 21% | 34% | 45% | 100% |

* Cooperation and Coordination is calculated as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **10. Cooperation and Coordination Overall Scenario (6 points)** | | | | | |
| 10.1Indicators | | 10.2 Process Scenario | 10.3 Quantitative Scenario | | |
| 10.1.1 cooperation and coordination among federal, province & local governments | 10.1.2 Inter local governments cooperation and coordination | 10.2.1 Formation and operation of inter-local governments committee (more than two local governments) to address common interests. | 10.3.1 Inter-government (federal and provincial) coordination and facilitation | 10.3.2 Inter local governments partnership | 10.3.3 Coordination with District Development Committee |

* 6 out of 6 Cooperation and Coordination score can be achieved if the following conditions are met:
  + **10.1.1:** The local government does necessary coordination of policies and action with the federal and provincial government before setting its annual policies and programs and regularly provides tax revenue share to the provincial government.
  + **10.1.2:** The local government has policies and structures in place to collaborate and coordinate with other local governments in issues of common concerns and annually has at least one program to address common concerns with joint investment.
  + **10.2.1:** Joint committees are formed and are actively working among local governments on matters of common interests and concerns.
  + **10.3.1:** The local government is well informed about implementation and facilitation of federal and provincial level projects and has set up an active institution for it.
  + **10.3.2:** The local government has collaborated with other local governments to run one or more program to minimize costs, maximize the use of resources or provide effective service delivery.
  + **10.3.3:** In regards of development and construction activities, the local government has discussed the feedback, suggestion and monitoring reports of District Development Committee in its executive body and has taken necessary decision.
* For each score of LISA, I take the average of the FY 20/21 and FY 21/22 and convert them into **percentage** (for intuitive comprehension of the coefficients)

# Cooperation and Coordination Breakdown

* For my research, I group the 6 LISA categorizers into the following three variables:
  + **fed\_cop =** ((avg 10.1.1 + avg 10.3.1) / 2) \* 100
    - It measures the cooperation and coordination sub-score that relates with federal government and provincial government, represented by 10.1.1 and 10.3.1
  + **local\_cop =** ((avg 10.1.2 + avg 10.2.1 + avg 10.3.2) / 3) \* 100
    - It measures the cooperation and coordination sub-score that relates with other local governments, represented by 10.1.2, 10.2.1 and 10.3.1.
  + **ddc\_cop** = (avg 10.3.3) \* 100
    - It measures the cooperation and coordination sub-score that relates with district development committee, represented by 10.3.3.

# 

# Variables and Date Sources

1. **Cooperation and Coordination Breakdown**For my research, I group the 6 LISA categorizers into the following three variables and express them in percentage:
   * **fed\_cop =** ((avg 10.1.1 + avg 10.3.1) / 2) \* 100
     + It measures the cooperation and coordination sub-score that relates with federal government and provincial government, represented by 10.1.1 and 10.3.1
   * **local\_cop =** ((avg 10.1.2 + avg 10.2.1 + avg 10.3.2) / 3) \* 100
     + It measures the cooperation and coordination sub-score that relates with other local governments, represented by 10.1.2, 10.2.1 and 10.3.1.
   * **ddc\_cop** = (avg 10.3.3) \* 100
     + It measures the cooperation and coordination sub-score that relates with district development committee, represented by 10.3.3.
2. **rel\_chgnl**: ln (sum of nightlight of 2021) – ln (sum of nightlight of 2017)

*Extracted from* [*VIIRS Nighttime Lights dataset*](https://eogdata.mines.edu/products/vnl/) *using QGIS software*

1. **ln\_baseline\_nl\_km:** log [(average of sum of night light of 2014, 2015, 2016, 2017)/area of municipality]

*Extracted from* [*VIIRS Nighttime Lights dataset*](https://eogdata.mines.edu/products/vnl/) *using QGIS software*

1. **lisa\_avg**: average LISA score of 2020/21 and 2021/22 (In case, LISA score of 2020/21 was not available, LISA score of 20 21/22 is considered to be the average)

*Downloaded from* [*LISA website*](https://lisa.mofaga.gov.np/home) *of the government of Nepal*

1. **high\_school\_percentage:** percentage of population who have completed high school and above  
   (Higher Secondary level or equivalent + Graduate level or equivalent +Post Graduate level or equivalent)

*Manually constructed dataset from datasets of each province from* [*National Population and Housing Census 2021*](https://censusnepal.cbs.gov.np/results/literacy)

1. **ageatelection**: age of chairperson at election in 2017  
   *Manually constructed dataset from* [*the election result pdfs*](https://oldsite.election.gov.np/election/en/election-result-book.html) *published in Nepali*
2. **gov\_coalition**: dummy variable Government Coalition = 1 if the chairperson of the municipality is affiliated with the parties in the federal government coalition

*Manually constructed dataset from* [*the election result pdfs*](https://oldsite.election.gov.np/election/en/election-result-book.html) *published in Nepali and then coded in Stata*

1. **female**: dummy variable female = 1 if the chairperson is a female  
   *Manually constructed dataset from* [*the election result pdfs*](https://oldsite.election.gov.np/election/en/election-result-book.html) *published in Nepali and then coded in Stata*
2. **ln\_popn**: log of population as per census of 2021

*Extracted from the* [*Preliminary Data of National Population and Housing Census 2021*](https://opendatanepal.com/dataset/preliminary-data-of-national-population-and-housing-census-2021)

1. **urban\_num**: dummy variable = 1 if the local level is municipality, sub-metropolitan city or metropolitan city i.e. non rural municipality
2. **neighbour\_percentage**: (total no. of neighbors with chairperson belonging to same party/total no. of neighbors) \* 100
3. **Percentage of average points score in LISA subcategories in 2020/21 and 2021/22**

**(**I am converting the obtained scores in percentage for intuitive comprehension of the coefficients)

|  |  |  |  |
| --- | --- | --- | --- |
| **SN** | **Category Title** | **Points** | **Coded variable** |
| 1 | Governance Management | 9 | gov\_magm\_per |
| 2 | Organization & Administration | 8 | org\_admin\_per |
| 3 | Budget Plan Management | 11 | budg\_magm\_per |
| 4 | Fiscal Economic Management | 11 | fiscal\_magm\_per |
| 5 | Service Delivery | 16 | service\_dev\_per |
| 6 | Judicial Execution | 9 | jud\_exe\_per |
| 7 | Physical Infrastructure | 13 | phy\_infra\_per |
| 8 | Social Inclusion | 10 | soc\_inc\_per |
| 9 | Environmental Protection and Disaster Management | 10 | env\_protec\_per |
| 10 | Cooperation and Coordination | 6 | cop\_cor\_per |
|  | **Total** | **100** |  |

# IV Outline by Prof O’Connell

**Step 1:** Using OLS, estimate a **'reduced form'** growth equation that excludes cooperation, but includes your instrument (call it Z) and the other variables. Call the estimated coefficient on your instrument "a": this is the 'reduced form' impact of the instrument on growth.

**Step 2:** Using OLS, estimate a '**first stage'** regression of cooperation (dependent variable) on your instrument and the other variables in the structural equation. Let's call the estimated coefficient on Z "c": this is the ceteris paribus, causal impact of your instrument on cooperation.

**Step 3:** Calculate b = a/c. This is an IV estimate of the causal impact of one extra 'unit' of cooperation on growth!

Or can be implemented using two stages least squares using Stata, which yields a more complicated calculation for b, but one that follows the same intuition

In order for the IV estimate to be valid, you need to satisfy 2 conditions:

(1) RELEVANCE: Your instrument, Z, must have a strong effect in predicting cooperation. This condition is testable by looking at the t statistic on Z in the first-stage regression, which should be something like 3 or more for a good instrument.

(2) EXCLUSION: Your instrument cannot have any direct effect of its own on growth; it must affect growth only via its impact on cooperation. This requirement is not testable -- it is a maintained hypothesis (if you have multiple instruments, you can test the validity of each of them assuming the others are valid, but you cannot test the validity of the whole set).

# Rethinking Covariates

* Important criteria for deciding covariates:
  + prior research
  + theoretical relevance
  + own knowledge of the context
* Mapping associations between covariates in a diagram:

A diagram of a change in a company

Description automatically generated

* Questioning LISA sub-category:

|  |  |  |
| --- | --- | --- |
| **SN** | **Category Title** | **Relevance** |
| 1 | Governance Management | Maybe (Overlapping) |
| 2 | Organization & Administration | Maybe (Overlapping) |
| 3 | Budget Plan Management | Yes (Overlapping) |
| 4 | Fiscal Economic Management | Yes (Overlapping) |
| 5 | Service Delivery | No |
| 6 | Judicial Execution | No |
| 7 | Physical Infrastructure | Yes |
| 8 | Social Inclusion | Maybe |
| 9 | Environmental Protection and Disaster Management | Yes |
| 10 | Cooperation and Coordination | Yes |

* How do I test multicollinearity?
  + What is the rule of thumb with VIF?

# Reference from other papers

(from ECON31 slides of Professor Bronchetti)

1. Cigarettes taxes as IV:
   1. For relevance criteria, IV needs to have t statistic near 3.0 even with controls

A screenshot of a website

Description automatically generated

A table of numbers and numbers

Description automatically generated with medium confidence

1. Policy changes as an IV

A screenshot of a computer

Description automatically generated

1. Sex of first two child as IV

A person with a picture of his face

Description automatically generated

# Understanding IV in context of this research

* It gives us very specific estimate of the causal impact of an additional unit of **local cooperation** on **change in nightlight** for the local levels whose **local cooperation** is responsive to our IV (political distance, ethnic distance etc.)
* Should political distance be a 0 or 1 variable?
  + Because of how local cooperation score is measured, having a lot of neighbors from same party doesn’t necessarily increase cooperation score.
  + Assuming neighboring local levels with similar party affiliation highly cooperate with each other, a local level doesn’t necessarily need to be surrounded by 50% same party neighbors to get high cooperation score. Even less would suffice.
* Maybe political affiliation fails the relevance criteria because of its low t-statistics alongside the controls.

# Should local cooperation be logged?

* Same applies for other metrics of LISA that are used as controls



*Figure 1: scatter plot of local cooperation score average (expressed in percentage) over percentage of neighbors from the same party with a fitted line*

**

*Figure 2: scatter plot of log of local cooperation score average (expressed in percentage) over percentage of neighbors from the same party with a fitted line*

# Caste/Ethnicity in Nepal

* Ethnicity is defined by caste in Nepal.
* Historically, there has been a strict caste hierarchy, with Brahmin and Kshetri being the upper and privileged casts.

*Figure 3: Pie chart of caste composition in Nepal as per census of 2021*

* Tour of Dataset
  + It felt like the most elaborate one yet

# Ethnic Fractionalization Index (EFI)

* The ethnic fractionalization index corresponds to the probability that two randomly drawn individuals within a country are not from the same ethnic group.
* The ethnic fractionalization index is calculated as one minus the Herfindahl index (which in turn is calculated by squaring the proportion of the population represented by each ethnic group and then summing the resulting numbers)

A black text on a white background

Description automatically generated

* The EFI of Nepal is 0.936



*Figure 4: Box plot of ELF score of each municipality across seven provinces of Nepal*

* Montalvo and Reynal-Querol (2021) argue that “relationship between diversity and growth is positive for small geographical areas”[[1]](#footnote-1)



*Figure 5: scatter plot of change in nightlight from 2017 to 2021 for each municipality over its ELF index as per census of 2021*

# 

# Main regression

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| VARIABLES | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl | rel\_chgnl |
|  |  |  |  |  |  |  |  |  |  |  |
| local\_cop | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* | 0.001\*\*\* |
|  | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| ln\_baseline\_nl\_km |  | 0.077\*\*\* | 0.073\*\*\* | 0.060\*\*\* | 0.029\*\* | 0.030\*\* | 0.031\*\* | 0.030\*\* | -0.002 | -0.006 |
|  |  | (0.013) | (0.014) | (0.015) | (0.014) | (0.014) | (0.014) | (0.014) | (0.017) | (0.017) |
| gov\_magm\_per |  |  | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 | -0.000 |
|  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| org\_admin\_per |  |  | -0.001 | -0.001 | -0.000 | -0.000 | -0.000 | -0.000 | 0.000 | 0.000 |
|  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| budg\_magm\_per |  |  | -0.000 | -0.000 | -0.000 | -0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
|  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| fiscal\_magm\_per |  |  | -0.000 | -0.000 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 |
|  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| phy\_infra\_per |  |  | 0.001\* | 0.001 | 0.000 | 0.001 | 0.001 | 0.001 | 0.000 | 0.000 |
|  |  |  | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| soc\_inc\_per |  |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
|  |  |  | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| env\_protec\_per |  |  | -0.000 | -0.000 | -0.001\* | -0.001\* | -0.001\* | -0.001\* | -0.001\*\* | -0.001\*\* |
|  |  |  | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| elf |  |  |  | 0.124\*\*\* | 0.149\*\*\* | 0.151\*\*\* | 0.152\*\*\* | 0.152\*\*\* | 0.115\*\*\* | 0.123\*\*\* |
|  |  |  |  | (0.035) | (0.031) | (0.031) | (0.031) | (0.031) | (0.032) | (0.033) |
| high\_school\_percent |  |  |  |  | 0.009\*\*\* | 0.009\*\*\* | 0.009\*\*\* | 0.009\*\*\* | 0.009\*\*\* | 0.010\*\*\* |
|  |  |  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| ageatelection |  |  |  |  |  | -0.001 | -0.000 | -0.001 | -0.001 | -0.001 |
|  |  |  |  |  |  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| female |  |  |  |  |  |  | 0.068 | 0.068 | 0.062 | 0.059 |
|  |  |  |  |  |  |  | (0.049) | (0.051) | (0.051) | (0.051) |
| gov\_coalitiion |  |  |  |  |  |  |  | -0.019 | -0.019\* | -0.021\* |
|  |  |  |  |  |  |  |  | (0.012) | (0.012) | (0.012) |
| ln\_popn |  |  |  |  |  |  |  |  | 0.042\*\*\* | 0.054\*\*\* |
|  |  |  |  |  |  |  |  |  | (0.009) | (0.011) |
| urban\_num |  |  |  |  |  |  |  |  |  | -0.031\* |
|  |  |  |  |  |  |  |  |  |  | (0.017) |
| Constant | 0.303\*\*\* | 0.282\*\*\* | 0.322\*\*\* | 0.227\*\*\* | 0.189\*\*\* | 0.208\*\*\* | 0.203\*\*\* | 0.215\*\*\* | -0.161 | -0.267\*\* |
|  | (0.009) | (0.008) | (0.058) | (0.057) | (0.056) | (0.064) | (0.064) | (0.065) | (0.109) | (0.130) |
|  |  |  |  |  |  |  |  |  |  |  |
| Observations | 691 | 691 | 691 | 691 | 691 | 691 | 691 | 691 | 691 | 691 |
| R-squared | 0.042 | 0.113 | 0.120 | 0.132 | 0.191 | 0.192 | 0.195 | 0.198 | 0.217 | 0.222 |

Robust standard errors in parentheses

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

# First Stage Regression

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| VARIABLES | local\_cop | local\_cop | local\_cop | local\_cop | local\_cop | local\_cop | local\_cop | local\_cop |
|  |  |  |  |  |  |  |  |  |
| neighbor\_dummy | 6.998\*\*\* | 7.952\*\*\* | 5.407\*\* | 5.599\*\* | 5.523\*\* | 5.255\*\* | 5.226\*\* | 1.799 |
|  | (2.398) | (2.378) | (2.306) | (2.287) | (2.286) | (2.291) | (2.289) | (2.065) |
| elf |  | 27.865\*\*\* | 21.704\*\*\* | 12.102 | 11.511 | 10.225 | 10.427 | 11.992\* |
|  |  | (6.650) | (6.658) | (7.427) | (7.476) | (7.481) | (7.498) | (6.749) |
| high\_school\_percent |  |  | 1.543\*\*\* | 1.324\*\*\* | 1.270\*\*\* | 1.222\*\*\* | 1.214\*\*\* | 0.398\* |
|  |  |  | (0.225) | (0.237) | (0.244) | (0.244) | (0.245) | (0.220) |
| ln\_popn |  |  |  | 4.672\*\*\* | 3.638\* | 3.315\* | 3.288\* | 1.362 |
|  |  |  |  | (1.756) | (1.968) | (1.969) | (1.966) | (1.726) |
| urban\_num |  |  |  |  | 3.112 | 3.404 | 3.448 | 1.398 |
|  |  |  |  |  | (3.015) | (3.003) | (2.999) | (2.614) |
| ageatelection |  |  |  |  |  | 0.259\*\* | 0.265\*\* | 0.232\*\* |
|  |  |  |  |  |  | (0.126) | (0.127) | (0.108) |
| female |  |  |  |  |  |  | 4.812 | 3.580 |
|  |  |  |  |  |  |  | (8.237) | (6.988) |
| gov\_magm\_per |  |  |  |  |  |  |  | 0.122 |
|  |  |  |  |  |  |  |  | (0.102) |
| org\_admin\_per |  |  |  |  |  |  |  | 0.010 |
|  |  |  |  |  |  |  |  | (0.081) |
| budg\_magm\_per |  |  |  |  |  |  |  | 0.058 |
|  |  |  |  |  |  |  |  | (0.098) |
| fiscal\_magm\_per |  |  |  |  |  |  |  | 0.211\*\* |
|  |  |  |  |  |  |  |  | (0.099) |
| phy\_infra\_per |  |  |  |  |  |  |  | 0.275\*\*\* |
|  |  |  |  |  |  |  |  | (0.080) |
| soc\_inc\_per |  |  |  |  |  |  |  | 0.072 |
|  |  |  |  |  |  |  |  | (0.065) |
| env\_protec\_per |  |  |  |  |  |  |  | 0.452\*\*\* |
|  |  |  |  |  |  |  |  | (0.094) |
| Constant | 29.591\*\*\* | 8.165 | -1.853 | -40.265\*\* | -29.878 | -37.682\*\* | -37.863\*\* | -72.392\*\*\* |
|  | (1.470) | (5.153) | (5.453) | (15.603) | (18.183) | (18.579) | (18.577) | (16.962) |
|  |  |  |  |  |  |  |  |  |
| Observations | 691 | 691 | 691 | 691 | 691 | 691 | 691 | 691 |
| R-squared | 0.012 | 0.033 | 0.095 | 0.105 | 0.107 | 0.113 | 0.113 | 0.339 |

Robust standard errors in parentheses

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

IV Validity

* A statistical test for a good IV is its relevance: a t-statistic of 3 or more in “first-stage regression”.
* Initial t-static is 2.87, but it drops to 0.75 after adding all the other controls. So, political proximity might not be a good IV.

A table of numbers and letters

Description automatically generated

Ethnic Proximity as IV

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Municipality/Ethnicity | Caste1 | Caste2 | Caste 3 | EFI index | Prop with same ethnicity with A |
| A | 0.5 | 0.3 | 0.2 | 0.38 | - |
| B | 0.3 | 0.2 | 0.5 | 0.38 | 0.7 |
| C | 0.4 | 0.4 | 0.2 | 0.36 | 0.9 |
| D | 0.1 | 0.4 | 0.5 | 0.4 | 0.6 |

# Way Forward

* Week timeline:
  + This week: Wrap up the work so far
  + Next week: Vacation (trip to Canada)
  + The week after: Resident Assistant Training
* Can I communicate results via email?
* Two directions:
  + First, keep working in the IV and see if I can find a good IV and take it forward from there.
  + Second, draft a paper for an undergraduate journal/conference targeting for a fall submission
* How can I engage with SSQL?
* How can I engage with more research?

1. José G. Montalvo, Marta Reynal-Querol; Ethnic Diversity and Growth: Revisiting the Evidence. The Review of Economics and Statistics 2021; 103 (3): 521–532. doi: <https://doi.org/10.1162/rest_a_00901> [↑](#footnote-ref-1)