Codebook for *Mining and local corruption in Africa*

April 21, 2016

A   Overview

This codebook described the different variables employed in "Mining and local corruption in Africa". All variable names (as written in the Stata .do file) are listed in italics (in parenthesis) directly after the label of the variable. Regarding the survey items drawn from the Afrobarometer, each survey question is written exactly as it is described in the Afrobarometer codebook(s), as a direct quote (using quotation marks). In the Afrobarometer codebook, some variables (those coded by the interviewer, such as whether the sampling unit is a rural or urban area) lack the actual question text. In these cases, we simply print their labels as given in the Afrobarometer codebooks.

A.1   Main dependent variables (all variables are from the Afrobarometer Data (N.d.))

- **Bribes to police (*bribe_police*):**

    - **Question/label:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Avoid a problem with the police (like passing a checkpoint or avoiding a fine or arrest)?"

    - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes for government permit (*bribe_permit*):**

    - **Question/label:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Get a document or a permit?"

    - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Perceptions of corruption (local police) (*corruption_police*):**

– **Question/label:** "How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: Police?"

– **Values:** 0=None, 1=Some of them, 2=Most of them, 3=All of them

- **Perceptions of corruption (local government councilors) (*local_corruption*):**

  – **Question/label:** "How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: elected local government councilors?"

  – **Values:** 0=None, 1=Some of them, 2=Most of them, 3=All of them

A.2 Control variables (all variables are from the Afrobarometer):

- **Urban area (*urban*):**

  – **Question/label:** (This question is answered by the interviewer): "Urban or Rural Primary Sampling Unit"

  – **Values:** 1=urban, 0=rural

- **Education (*education*):**

  – **Question/label:** "What is the highest level of education you have completed?"

  – **Values:** 0=No formal schooling, 1=Informal schooling (including Koranic schooling), 2=Some primary schooling, 3=Primary school completed, 4=Some secondary school/ High school, 5=Secondary school completed/High school, 6=Post-secondary qualifications, other than university e.g. a diploma or degree from a technical/polytechnic/college, 7=Some university, 8=University completed, 9=Post-graduate

- **Gender (*female*):**

- **Question/label (answered by interviewer):** "This interview, gender"

  - **Values:** 1=female, 0=male

- **Age (*age*):**

  - **Question/label:** "How old are you?"

  - **Values:** Numeric age

- **Age squared (*age2*):**

  - **Question/label:** "How old are you?"

  - **Values:** Numeric age squared

A.3   Other survey variables used (from Afrobarometer):

- **Police station (*police_station*):**

  - **Question/label:** (answered by interviewer): "Were the following services present in the primary sampling unit/enumeration area: Police station?"

  - **Values:** 1=yes, 0=no

- **Police in area (*police*):**

  - **Question/label:** (answered by interviewer): "In the PSU/EA, did you (or any of your colleagues) see: Any policemen or police vehicles?"

  - **Values:** 1=yes, 0=no

- **National-level corruption (president) (*corruption_president*):**

  - **Survey question:** "How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: The President/Prime Minister and Officials in his Office?"

– **Values:** 0=None, 1=Some of them, 2=Most of them, 3=All of them

- **National-level corruption (national government officials) (*corruption_national_officials*):**

  – **Survey question:** "How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: National government officials?"

  – **Values:** 0=None, 1=Some of them, 2=Most of them, 3=All of them

- **Bribes for school placement (*bribe_school*):**

  – **Survey question:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Get a child into school?"

  – **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes to tax officials (*bribe_tax*):**

  – **Survey question:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: avoid problem with tax official?"

  – **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes for government services (*bribe_service*):**

  – **Survey question:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Get a household service (like piped water, electricity, or phone)?"

  – **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes for border crossing (*bribe_border*):**

- **Survey question:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Cross a border (like a customs or immigration post)?"
- **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes for medical attention (*bribe_medical*):**

  - **Survey question:** "In the past year, how often (if ever) have you had to pay a bribe, give a gift, or do a favor to government officials in order to: get medicine or medical attention?"
  - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribe offered in elections (*bribe_elections*):**

  - **Survey question:** "And during the [20xx] election, how often (if ever) did a candidate or someone from a political party offer you something, like food or a gift, in return for your vote? Variable Label: Election incentives offered"
  - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Bribes for water (*bribe_water*):**

  - **Survey question:** "In the past year, how often, if ever, have you had to pay a bribe, give a gift, or do a favor to government officials in order to: Get water or sanitation services?"
  - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

- **Roadblocks by police or army (*police_roadblock*):**

  Question/label: (answered by interviewer): "Were the following present in the primary sampling unit/enumeration area: Roadblock by police or army?"
  - **Values:** 0-3, 0=Never, 1=Once or Twice, 2=A Few Times, 3=Often

B   Main independent/mining variables (all variables draw from information in the RMD Database, SNL Metals and Mining (2014):

- **Active 50 km (*active50*):**

  - **Measure:** Is there an active mine within a 50km buffer zone centered on the respondent's location?
  - **Values:** 1=Yes, 0=No

- **Inactive 50 km (*inactive50*):**

  - **Measure:** Is there an inactive mine (but no active mine) within a 50km buffer zone centered on the respondent's location?
  - **Values:** 1=Yes, 0=No

- **Active 25 km (*active25*):**

  - **Measure:** Is there an active mine within a 25km buffer zone centered on the respondent's location?
  - **Values:** 1=Yes, 0=No

- **Inactive 25 km (*inactive25*):**

  - **Measure:** Is there an inactive mine (but no active mine) within a 25km buffer zone centered on the respondent's location?
  - **Values:** 1=Yes, 0=No

- **Above median production volume (*high_prod*):**

- **Measure:** Dummy variable for mine with above median production volume

  - **Values:** 1=above median, 0=otherwise

- **Above median production levels (*high_value*):**

  - **Measure:** Dummy variable for mine with above median production value

  - **Values:** 1=above median, 0=otherwise

- **Number of active mines conditional on the area being a mining area (*nactive50*):**

  - **Measure:** Number of active mines within a 50km buffer zone centered on the respondent's location, conditional on the area being a mining area

  - **Values:** Numeric value

- **Log number of active mines conditional on the area being a mining area (*lognactive50*):**

  - **Measure:** Log number of active mines within a 50km buffer zone centered on the respondent's location, conditional on the area being a mining area

  - **Values:** Numeric value

- **Suspended 25 km (*suspended25*):**

  - **Measure:** Is there a suspended mine within a 25km buffer zone centered on the respondent's location?

  - **Values:** 1=Yes, 0=No

- **Suspended 50 km (*suspended50*):**

- **Measure:** Is there a suspended mine within a 50km buffer zone centered on the respondent's location?

- **Values:** 1=Yes, 0=No

- **Active 100 km (*active100*):**

  - **Measure:** Is there an active mine within a 100km buffer zone centered on the respondent's location?

  - **Values:** 1=Yes, 0=No

- **Inactive 100 km (*inactive100*):**

  - **Measure:** Is there an inactive mine (but no active mine) within a 100km buffer zone centered on the respondent's location?

  - **Values:** 1=Yes, 0=No

- **Active 200 km (*active100*):**

  - **Measure:** Is there an active mine within a 100km buffer zone centered on the respondent's location?

  - **Values:** 1=Yes, 0=No

- **Inactive 200 km (*inactive100*):**

  - **Measure:** Is there an inactive mine (but no active mine) within a 100km buffer zone centered on the respondent's location?

  - **Values:** 1=Yes, 0=No

- **ID of first mine within 50km (*minefixed50*):**

– **Measure:** Is there an inactive mine (but no active mine) within a 100km buffer zone centered on the respondent's location?

– **Values:** 1=Yes, 0=No

C   Nighttime light variables and mining variables drawing from sources other than RMD:

- **Median light (*median*):**

  – **Source:** US Airforce

  – **Measure:** Yearly median level of night light emissions in a 50km buffer zone centered on the respondent's location, as measured by satellite retrieved nightlight measures ranging from zero (no light) to 63.

  – **Values:** Interval scale from 0 to 63

- **Average light (*mean*):**

  – **Source:** US Airforce

  – **Measure:** Yearly average level of night light emissions in a 50km buffer zone centered on the respondent's location, as measured by satellite retrieved nightlight measures ranging from zero (no light) to 63.

  – **Values:** Interval scale from 0 to 63

- **Diamond mines (*dia50*):**

  – **Source:** Gilmore et al. (2005)

  – **Measure:** A variable registering the presence of at least one diamond deposit within a 50km buffer zone centered on the respondent's location

– **Values:** 1=diamond deposit present, 0 otherwise

- **USGS mines (*usgs50*):**

  – **Source:** USGS

  – **Measure:** A variable registering the presence of at least one mine based on USGS survey within a 50km buffer zone centered on the respondent's location

  – **Values:** 1=mine present, 0=otherwise

- **Oil deposit (*oil_cluster*):**

  – **Source:** Lujala, Rød and Thieme (2007)

  – **Measure:** A variable registering the presence of a petroleum deposit polygon within a 50km buffer zone centered on the respondent's location

  – **Values:** 1=petroleum deposit present, 0=otherwise

D Country-level variables

- **Corrupt country (*corrupt_country*):**

  – **Source:** (Teorell et.al., 2015)

  – **Measure:** Dummy variable registering whether the country records above the median level of corruption as measured by the world bank's World Governance Indicators, taken from the QOG dataset.

  – **Values:** 1=above median corruption, 0=otherwise

- **Rich country (*rich_country*):**

  – **Source:** World Development Indicators

- **Measure:** Dummy variable registering whether the country records above median per capita income, measured using PPP adjusted GDP per capita, taken from the WDI.

- **Values:** 1=above median income, 0=otherwise

- **Democratic country (*democratic_country*):**

  - **Source:** (Marshall, n.d.)

  - **Measure:** Dummy variable registering whether the country records above the median score on the Polity index.

  - **Values:** 1=above median democracy, 0=otherwise

- **National level corruption (*wbgi_cee*):**

  - **Source:** (Teorell et.al., 2015)

  - **Measure:** "Control of corruption" index from the World Bank's World Governance Indicators measuring national-level corruption, taken from the QOG dataset.

  - **Values:** Ordinal scale

E   Instrumental variables

With the exception of the instrument using mineral prices, the instruments are created using the mining variables from RMD. The price information comes from the US geological survey (USGS).

- **Mines within 100-200km *(nwithin100_200)*:**

  - **Measure:** A variable counting the number of mines within a 100–200km band from the respondent.

   – **Values:** Numeric value

• **Dummy for any mine within 100-200km *(nwithin100_200)*:**

   – **Measure:** Dummy for any mine within a 100–200km band from the respondent.

   – **Values:** 1 =any mine present, 0 otherwise

• **Mineral-year specific conditions interacted with mineral presence *(sactive_nwithin)*:**

   – **Measure:** A composite measure, described most clearly by the following summation: For each respondent in cluster i and year t, the instrument IV_{it} is given by:

$$IV_{it} = \sum_{j=1}^{n} \left(\frac{nactive_{jt}}{nmines_j} \times nwithin100\_200_{ij} \right), \tag{1}$$

where the sum runs over mineral j in the set of n distinct minerals in our mining data. *nactive* is the number of active mines with mineral j as the main mineral in year t, and *nmines* is the total number of mines of mineral j ever recorded in our data.

We note that since year fixed effects are included in the specifications, this fraction is a proxy for the conditions for production and export of a given mineral in a given year. This could potentially capture mineral-specific technology shocks, price expectations, or other aspects related to the profitability of being an active producer of mineral j. When interacted with the mine "presence" variable (*nwithin100_200*) and summed over all minerals, this provides composite predictor for the probability of there being an active mine nearby in a given year.

   – **Values:** Numeric value

- **Mineral presence interacted with price** *(priceXpresence)*:

  - **Measure:** 5 year moving average price for a given mineral interacted with a dummy registering the presence of at least 1 mine extracting that mineral within the 100-200km band.
  - **Values:** Numeric value

F   Other variables (based on RMD data)

- **Active mine within 200km but not within 100 km** *(active200_not100)*:

  - **Measure:** A variable registering the presence of an active mine within 200km, but not within 100km, band.
  - **Values:** 1=active mine within 200km, 0 otherwise

- **Inactive mine within 200km but not within 100 km** *(inactive200_not100)*:

  - A variable registering the presence of an inactive mine within 200km, but not within 100km, band.
  - **Values:** 1=inactive mine within 200km, 0 otherwise

G   Other variables (general)

- **Afrobarometer respondent interview year** *(int_year)*:

  - **Measure:** Afrobarometer respondent interview year
  - **Values:** Numerical

- **Country** *(Country)*:

  - **Measure:** Country

– **Values:** Categorical

• **Country code** *(countrycode)*:

  – **Measure:** Afrobarometer numeric country code, ISO 3166-1

  – **Values:** numeric

• **Afrobarometer cluster** *(cluster)*:

  – **Measure:** Afrobarometer cluster (same latitude/longitude)

  – **Values:** integer

• **Name of closest mine** *(closestmine)*:

  – **Measure:** Name of closest mine

  – **Values:** categorical

• **Afrobarometer wave** *(wave)*:

  – **Measure:** Round of afrobarometer survey $(2, 3, 4$ etc.$)$

  – **Values:** integer

• **Region** *(region)*:

  – **Measure:** Sub-national region or province

  – **Values:** categorical

• **Region id** *(region_id)*:

  – **Measure:** group id variable (country, region)

  – **Values:** categorical

- **District id** *(district_id)*:

  - **Measure:** group id variable (country, region, district)

  - **Values:** categorical

- **Year** *(year)*:

  - **Measure:** year

  - **Values:** integer

- **Year** *(year)*:

  - **Measure:** year

  - **Values:** integer

- **Years since first mine opened, 50 km** *(yearssince_active50*:

  - **Measure:** year

  - **Values:** integer

- **Bribes to police dummy** (*bribe_police_dummy*):

  - **Question/label:** Bribes to police dummy

  - **Values:** 1=paid bribe to police, 0 otherwise

- **Bribes for government permit dummy** (*bribe_permit_dummy*):

  - **Question/label:** Bribes for permit dummy

  - **Values:** 1=paid bribe to police, 0 otherwise

- **Perceptions of corruption (local police) dummy** (*corruption_police_dummy*):

- **Question/label:** Perceptions of local police dummy

- **Values:** 1=local police corrupt, 0 otherwise

- **Perceptions of corruption (local politicians) dummy (*local_corruption_dummy*):**

  - **Question/label:** Perceptions of local corruption dummy

  - **Values:** 1=local politicians corrupt, 0 otherwise

- **Distance to closest mine in kilometers (*kilometers*):**

  - **Question/label:** kilometer distance to closest mine

  - **Values:** numeric

- **Dummy for >=2 active mines within 50km (*active50_2*):**

  - **Question/label:** Dummy for >=2 active mines within 50km

  - **Values:** 1 >=2 active mines within 50km, 0 otherwise

- **Dummy for >=3 active mines within 50km (*active50_3*):**

  - **Question/label:** Dummy for >=3 active mines within 50km

  - **Values:** 1 >=3 active mines within 50km, 0 otherwise

- **Dummy for >=4 active mines within 50km (*active50_4*):**

  - **Question/label:** Dummy for >=4 active mines within 50km

  - **Values:** 1 >=4 active mines within 50km, 0 otherwise

- **Dummy for >=5 active mines within 50km (*active50_5*):**

  - **Question/label:** Dummy for >=5 active mines within 50km

  - **Values:** 1 >=5 active mines within 50km, 0 otherwise

- **Observation id (*obs_id*):**

  - **Question/label:** Observation id

  - **Values:** Integer

- **Country-dummies (*_Icountry1*-*_Icountry33*):**

  - **Question/label:** Dummies for each country code

  - **Values:** Binary

- **Year-dummies (*_Yyear1*-*_Yyear10*):**

  - **Question/label:** Dummies for each year

  - **Values:** Binary

References

Afrobarometer Data. N.d. Round 2–5, 2002, 2003, 2004, 2005, 2006, 2008, 2009, 2011, 2012, 2013.

  URL: *http://www.afrobarometer.org*

Gilmore, Elisabeth, Nils Petter Gleditsch, Päivi Lujala and Jan Ketil Rød. 2005. "Conflict diamonds: A new dataset." *Conflict Management and Peace Science* 22(3):257–272.

Lujala, Päivi, Jan Ketil Rød and Nadja Thieme. 2007. "Fighting over oil: Introducing a new dataset." *Conflict Management and Peace Science* 24(3):239–256.

Marshall, Monty G. n.d. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2009.".

  URL: *http://www.systemicpeace.org/polity/polity4.htm. Accessed January 16, 2014.*

SNL Metals and Mining. 2014. "Raw Materials Database.".

Teorell, Jan and Stefan Dahlberg and Soren Holmberg and Bo Rothstein and Felix Hartmann and Richard Svensson. 2015. "The Quality of Government Standard Dataset, version Jan15.".