Maddy Butts
Econ 428
Data Appendix

Countries in dataset:

Algeria	Ethiopia	Namibia
Angola	Gabon	Niger
Benin	Gambia	Nigeria
Botswana	Ghana	Rwanda
Burkina Faso	Guinea	Senegal
Burundi	Guinea-Bissau	Sierra Leone
Central African Republic	Kenya	Swaziland
Chad	Madagascar	Tanzania
Cote D'Ivoire	Malawi	Togo
Democratic Republic of the Congo	Mali	Uganda
Eritrea	Mauritania	Zambia

Variable overview:

Variable name	**Variable description**	**Original downloaded unit of measure**	**Changes to data**	**Final unit of measure**	**Data Source**
lmal	Total Malaria cases	Total suspected cases	Divide by 1000	Cases per 1000 people	WHO
agr	Agricultural area	Area per 1000 ha	logged all data points for normality	Logged area per 1000 ha	FAO
lfor	Forest area	Area per 1000 ha	logged all data points for normality	Logged area per 1000 ha	FAO

lirr	Total area equipped for irrigation	Area per 1000 ha	logged all data points for normality	Logged area per 1000 ha	FAO
cer	Total cereal yield	Hectograms/hectare	Divide all data points by 10,000 to convert from hectograms to tons	Tons/hectare	FAO
lngdp	GDP per capita, PPP	(Constant 2005 international $)	logged all data points		World Bank
sqcement	CO2 emissions from cement production	(thousand metric tons)	Square rooted all data points for normality	Square root(thousand metric tons)	World Bank
rural	Rural population	% of total population	n/a	% of total population	World Bank
lnpop	Population	Total people	Logged all data points	Logged total population	World Bank
young	Population ages 0-14	% of total population	n/a	% of total population	World Bank
lnterrha	Terrestrial protected areas	sq km	Divide all data points by 10 to convert to area per 1000 hectare and logged	Logged area per 1000 ha	World Bank
life	Life expectancy at birth	total years	n/a	total years	World Bank
health	Health expenditure, public	% of GDP	Took square root for normality	Square rooted % of GDP	World Bank
inv	Gross public investment	% of GDP	lagged by one year	% of GDP lagged one year	World Bank
labor	Labor force participation rate, total	% of total population ages 15-64	n/a	% of total population ages 15-64	World Bank
govt	Government Effectiveness	Estimate of governance	n/a	Estimate of governance	Worldwide Governance Indicators from World Bank

reg	Regulatory quality	Estimate of governance	n/a	Estimate of governance	Worldwide Governance Indicators from World Bank
percoast	Population living 100 km from coast	total population	Divided all data points by total population	% of population living 100 km from coast	UN Environmental Program
mallab	lmal*labor	Interaction term			
malinv	lmal*inv	Interaction term			

Histograms:

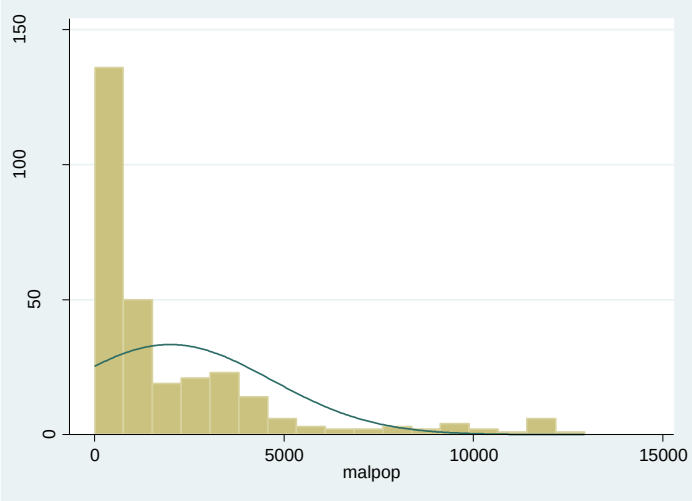
I) Malaria

World Health Organization. World malaria report 2013, Appendix 6D . World Health Organization, 2013. Retrieved from http://www.who.int/malaria/publications/world_malaria_report_2013/en/

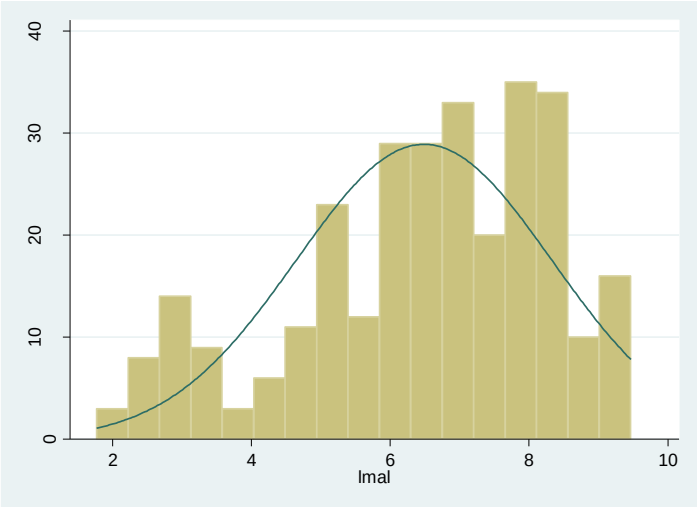
Description: Logged total suspected malaria cases per 1,000 people

Missing observations: 2

Variable	Obs	Mean	Std. Dev.	Min	Max
lmal	295	6.486733	1.843131	1.771727	9.467617



Histogram of malaria cases per 1,000 people



Histogram of logged malaria cases per 1,000 people

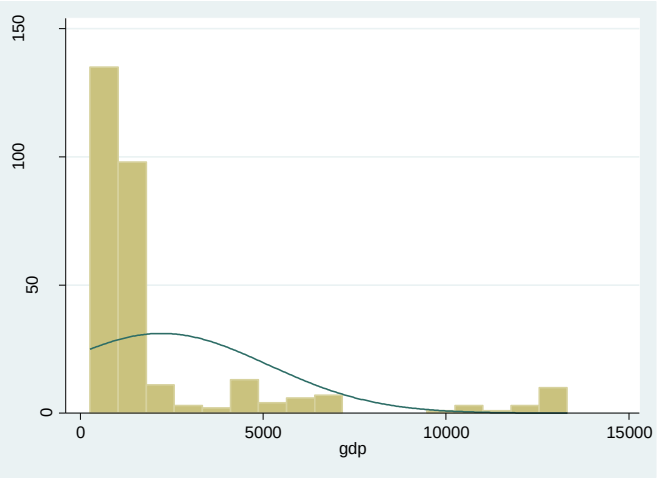
II) GDP data

The world bank, World Development Indicators (2014). *GDP per capita, PPP (Constant 2005 international $)*. Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

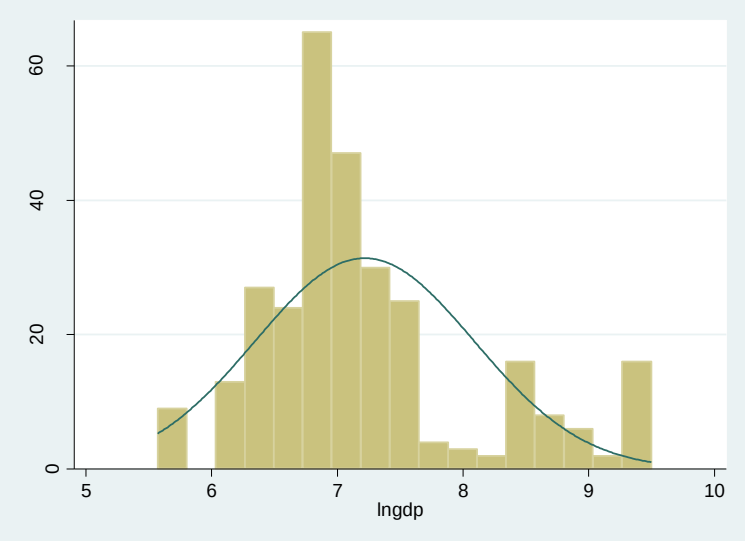
Description: Logged GDP per capita, PPP (Constant 2005 international $).

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
lngdp	297	7.216821	.8722221	5.569924	9.496856



Histogram of per capita GDP



Histogram of logged per capita GDP

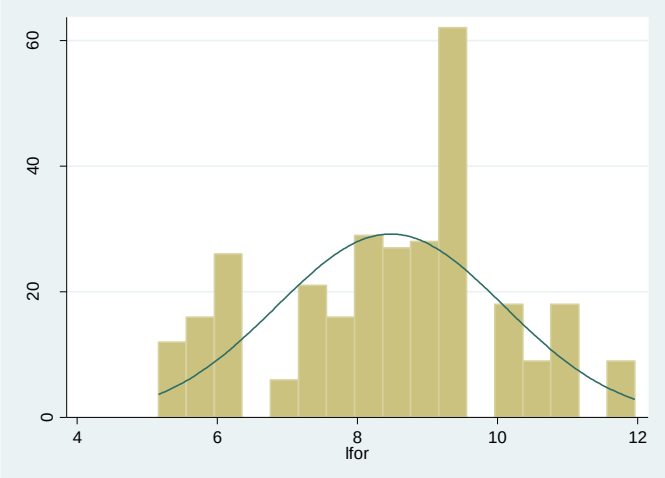
III) Forest area

FAOSTAT, Food and Agriculture Organization of the United Nations (2014). *Forest area*. Retrieved from http://faostat3.fao.org/faostat-gateway/go/to/download/R/RL/E

Description: Logged forest area per 1000 hectare

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
lfor	297	8.473378	1.624956	5.157905	11.9636



Histogram of loggest forest area

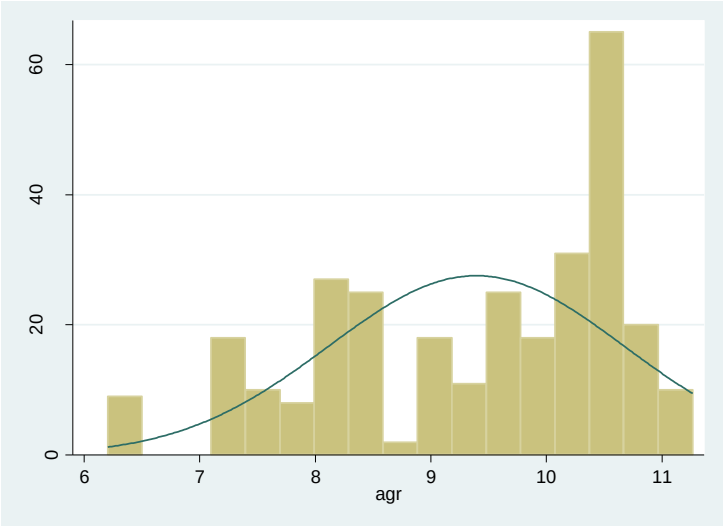
IV) Agricultural area

FAOSTAT, Food and Agriculture Organization of the United Nations (2014). *Forest area*. Retrieved from http://faostat3.fao.org/faostat-gateway/go/to/download/R/RL/E

Description: logged agricultural area per 1,000 hectares

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
agr	297	9.393908	1.279902	6.204558	11.26446



Histogram of logged agricultural area

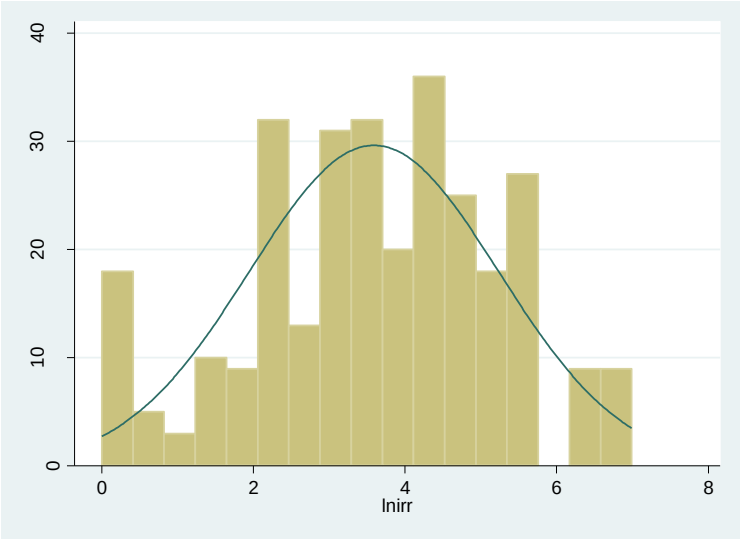
V) Irrigation

FAOSTAT, Food and Agriculture Organization of the United Nations (2014). *Total area equipped for irrigation*. Retrieved from http://faostat3.fao.org/faostat-gateway/go/to/download/R/RL/E

Description: Logged total area equipped for irrigation per 1,000 hectares

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
lnirr	297	3.589905	1.643967	0	6.990256



Histogram of logged area equipped for irrigation

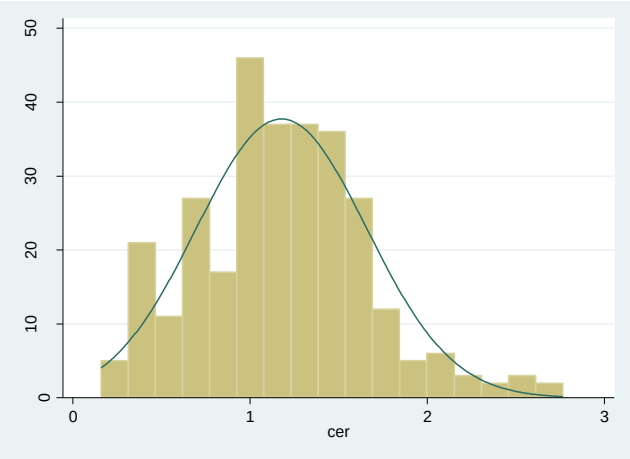
VI) Cereal data

FAOSTAT, Food and Agriculture Organization of the United Nations (2014). *Cereals, Total + (Total)* . Retrieved from http://faostat3.fao.org/faostat-gateway/go/to/download/R/RL/E

Description: Total yield of cereals in tons per hectare

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
cer	297	1.177754	.4818472	.158229	2.766479



Histogram of cereal production

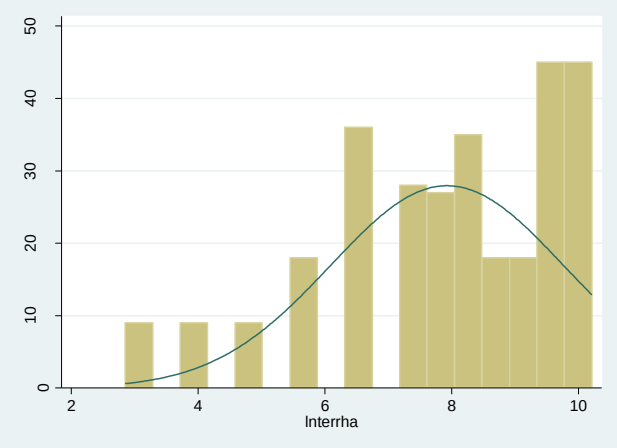
VII) Protected terrestrial area

The world bank, World Development Indicators (2014). *Terrestrial protected areas (sq km)*. Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Logged protected terrestrial areas per 1000 hectares

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
lnterrha	297	7.925282	1.834999	2.851458	10.2083



Histogram of logged terrestrial protected area

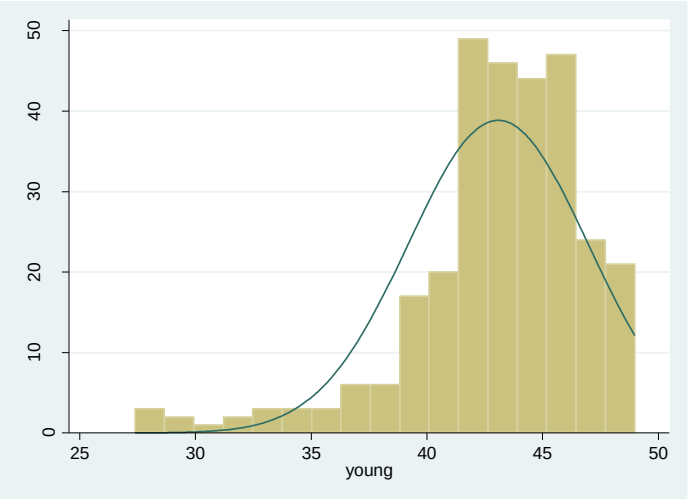
VIII) Young: Population under 14 years old

The world bank, World Development Indicators (2014). *Population ages 0-14 (% of total)* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Percentage of population between 0-14 years old

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
young	297	43.0794	3.874518	27.39492	48.99195



Histogram of percentage of population under 14 years old

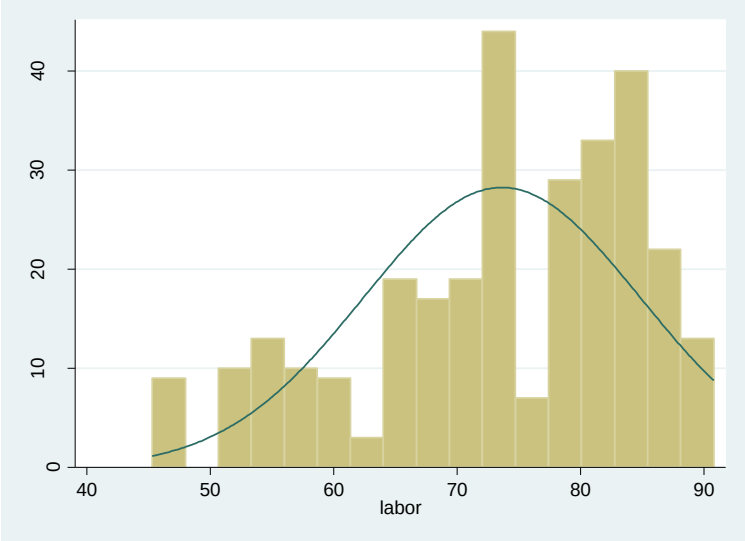
IX) Labor force

The world bank, World Development Indicators (2014). *Labor force participation rate, total (% of total population ages 15-64).* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Labor force participation rate, total (% of total population ages 15-64)

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
labor	297	73.63098	11.22167	45.3	90.8



Histogram of labor participation rate

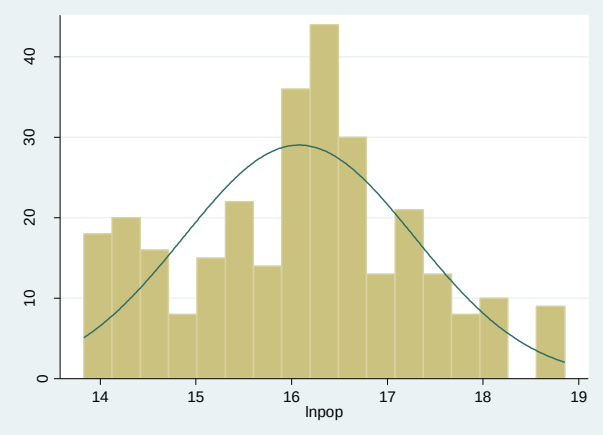
X) Population

The world bank, World Development Indicators (2014). *Population, total.* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Logged total population

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
lnpop	297	16.07744	1.20649	13.82755	18.85563



Histogram of logged population

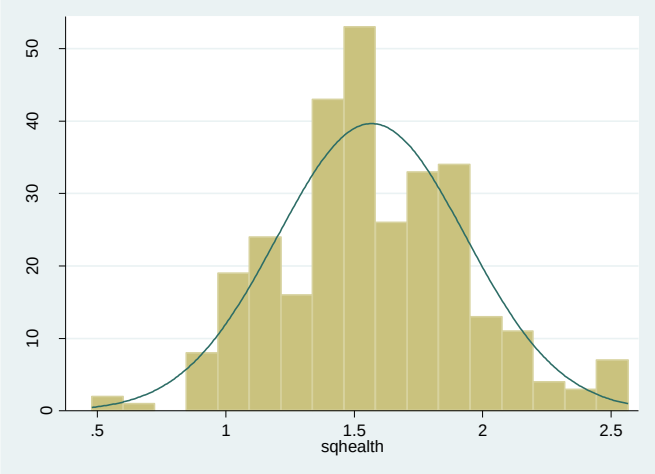
XI) Health expenditure

The world bank, World Development Indicators (2014). *Health expenditure, public (% of GDP).* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Square rooted public health expenditure (% of GDP)

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
sqhealth	297	1.567647	.3668554	.4780961	2.565595



Histogram of square rooted public health expenditure

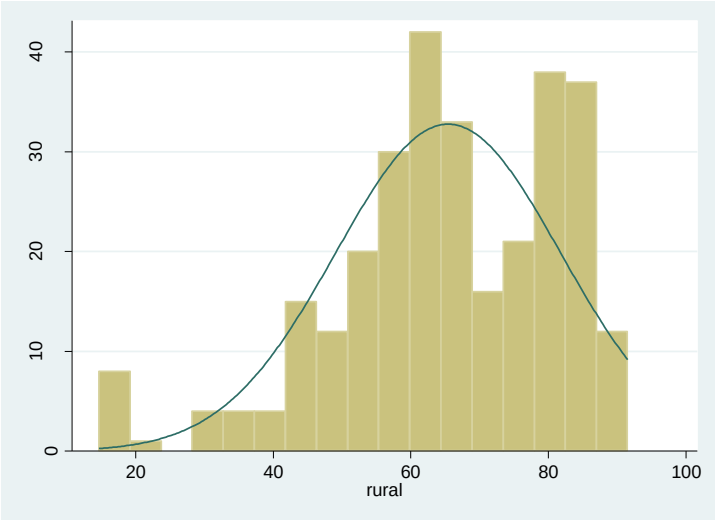
XII) Rural population

The world bank, World Development Indicators (2014). *Rural population (% of total).* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Percent of population that lives in rural areas

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
rural	297	65.42327	16.35679	14.6308	91.5282



Histogram of rural population

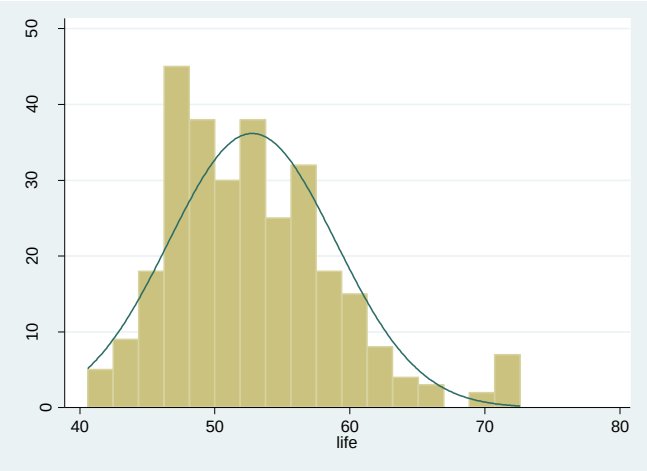
XIII) Life expectancy

The world bank, World Development Indicators (2014). *Life expectancy at birth.* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description**:** Life expectancy at birth in years

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
life	297	52.76819	6.174681	40.57768	72.62276



Histogram of life expectancy at birth

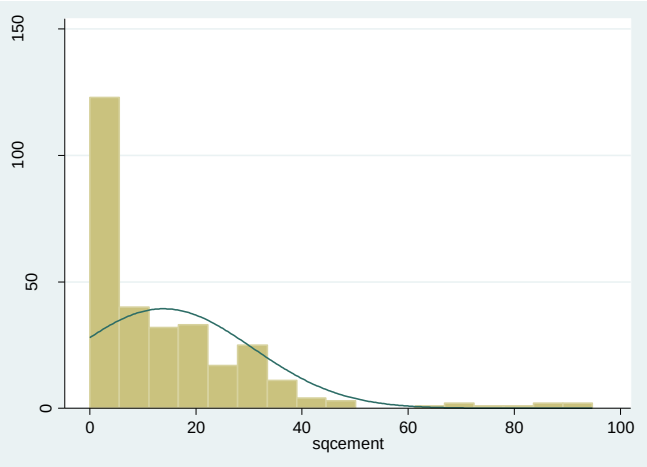
XIV) Cement production

The world bank, World Development Indicators (2014). *Life expectancy at birth.* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Square rooted C02 emissions from cement production in thousand metric tons

Missing observations: 0

Variable	Obs	Mean	Std. Dev.	Min	Max
sqcement	297	13.91092	16.77433	0	94.74606



Histogram of CO2 emissions from cement production

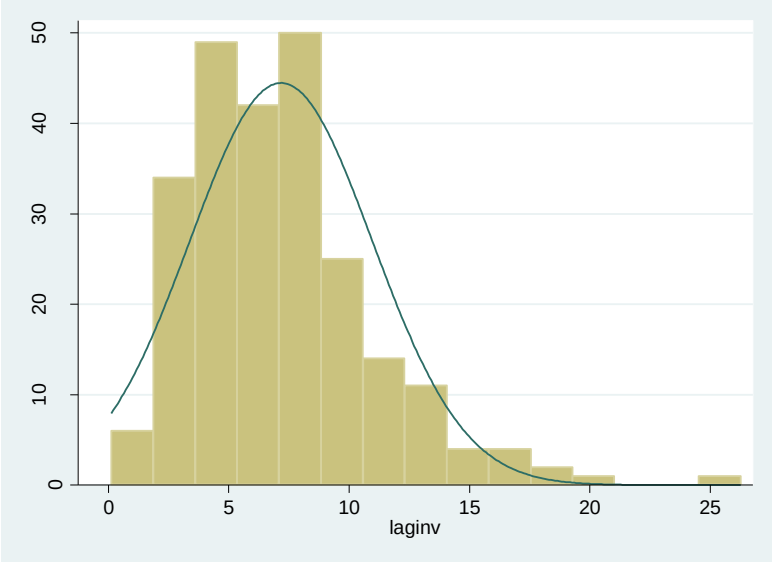
XV) Investment

The world bank, World Development Indicators (2014). *Gross public investment (% of GDP).* Retrieved from http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=africa-development-indicators

Description: Lagged gross public investment as % of GDP

Missing observations: 54

Variable	Obs	Mean	Std. Dev.	Min	Max
laginv	243	7.17092	3.79793	.1203258	26.25196



Histogram of lagged public investment

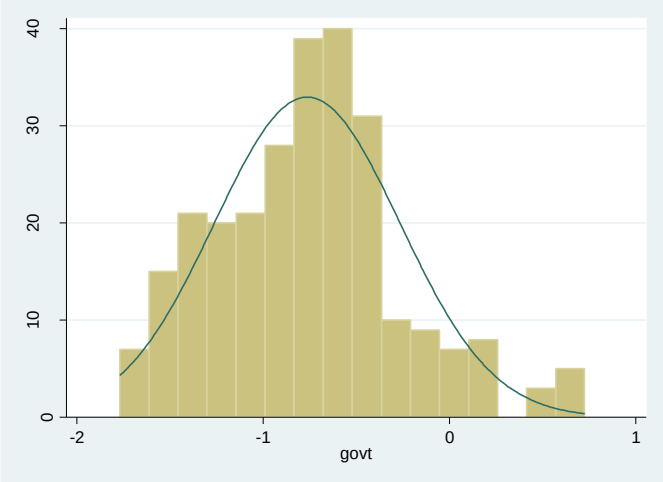
XVI) Government effectiveness

Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi (2010). Worldwide Governance Indicators. The World Bank. Retrieved from http://info.worldbank.org/governance/wgi/index.aspx#home

Description: Estimate of governance, Reflects perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.

Missing observations: 33 (Missing 2001 data)

Variable	Obs	Mean	Std. Dev.	Min	Max
govt	264	-.7633795	.4985083	-1.768885	.7272412



Histogram of government effectiveness

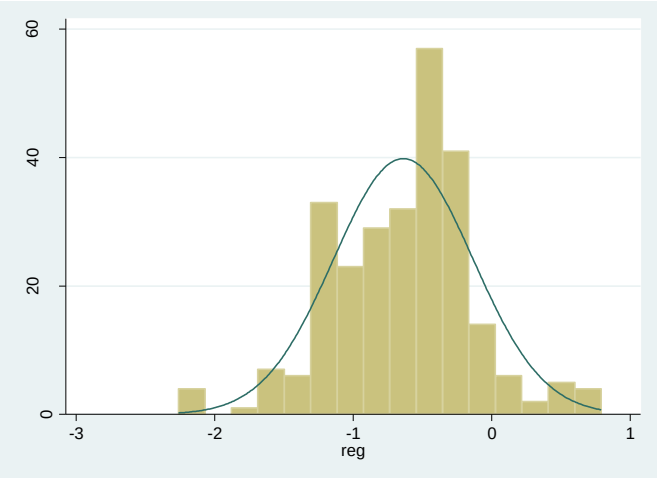
XVII) Regulatory quality

Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi (2010). Worldwide Governance Indicators. The World Bank. Retrieved from http://info.worldbank.org/governance/wgi/index.aspx#home

Description: Estimate of governance, Reflects perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development.

Missing observations: 33 (missing 2001 data)

Variable	Obs	Mean	Std. Dev.	Min	Max
reg	264	-.6376965	.5042844	-2.260322	.7914705



Histogram of regulatory quality

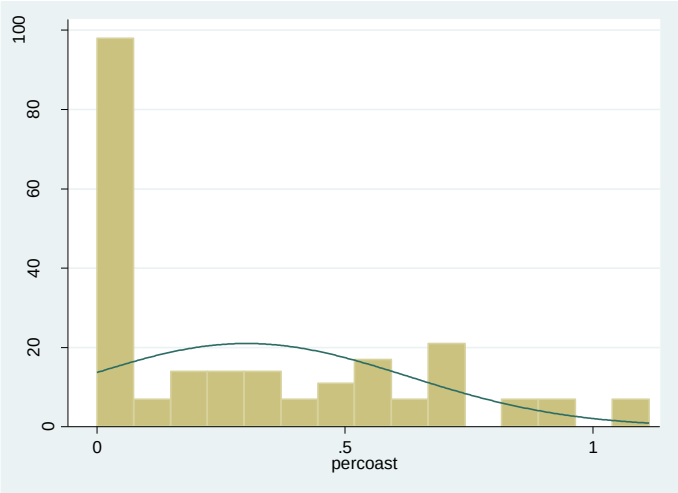
XVIII) Population within 100 km of coast

UNEP (2014): The UNEP Environmental Data Explorer, as compiled from UNEP/DEWA/GRID-Geneva . United Nations Environment Programme. Retrieved from http://geodata.grid.unep.ch/mod_download/download.php

Description: Percentage of population living within 100 km of coast

Missing observations: 66 (Missing 2008 and 2009 data; however, this variable is not expected to change very must year to year, as landlocked countries generally have 0% of population within 100 miles of coast)

Variable	Obs	Mean	Std. Dev.	Min	Max
percoast	231	.3013986	.3255677	0	1.113264



Histogram of percent of population living within 100 km of coast

n.b. *percoast* is not normal because landlocked countries have 0% of population within 100 km of coast. This makes *percoast* similar to a dummy variable.

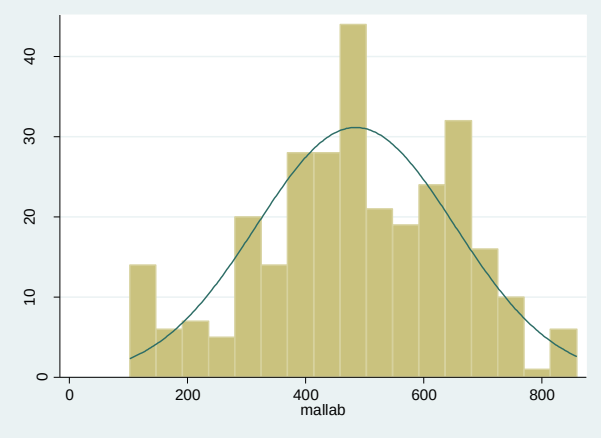
Interaction terms

XIX) Malaria Labor interaction term

Description: interaction of *lmal* and *labor* variables

Missing observations: 2

Variable	Obs	Mean	Std. Dev.	Min	Max
mallab	295	484.5047	168.0152	102.2286	858.7128



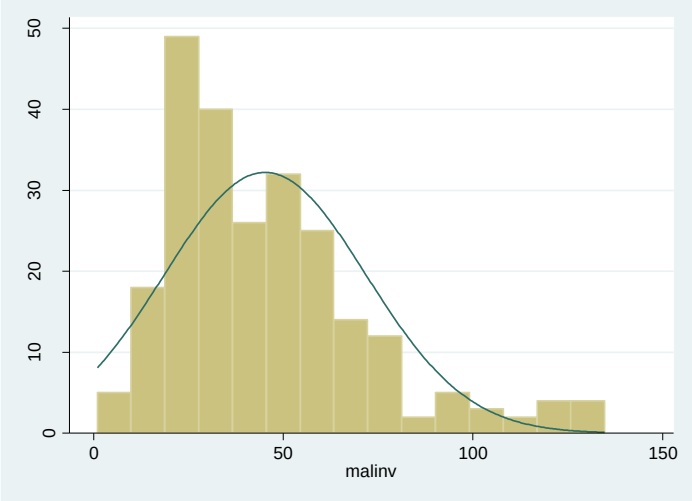
Histogram of *mallab*

XX) Malaria investment interaction term

Description: interaction of *lmal* and *laginv* variables

Missing observations: 56

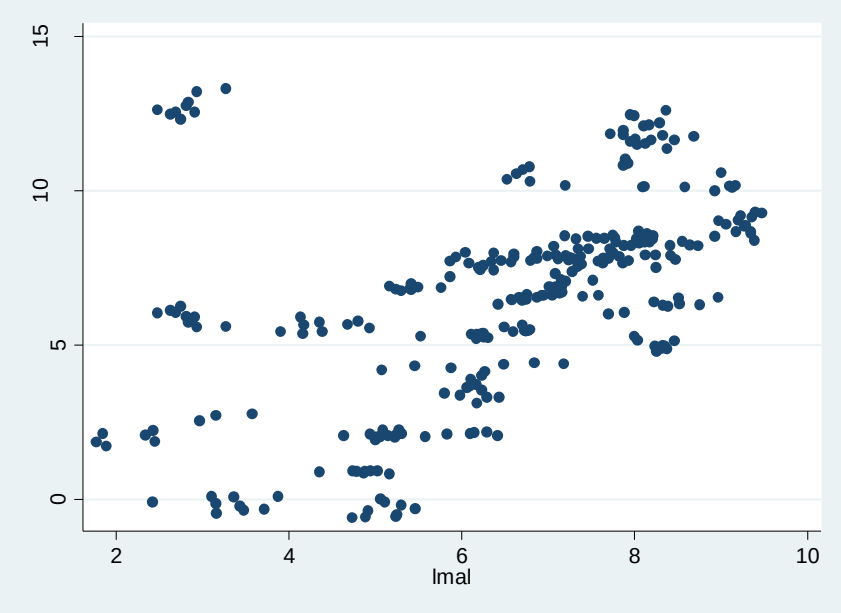
Variable	Obs	Mean	Std. Dev.	Min	Max
malinv	241	45.20664	26.62587	.9479987	134.7341



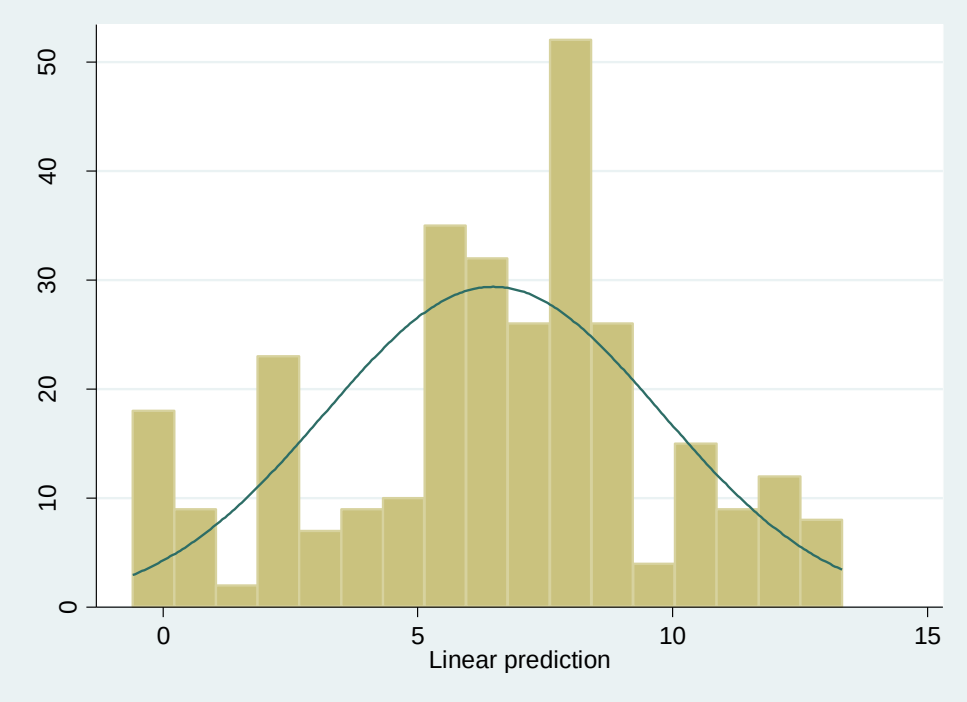
Histogram of *malinv*

Model Specification:

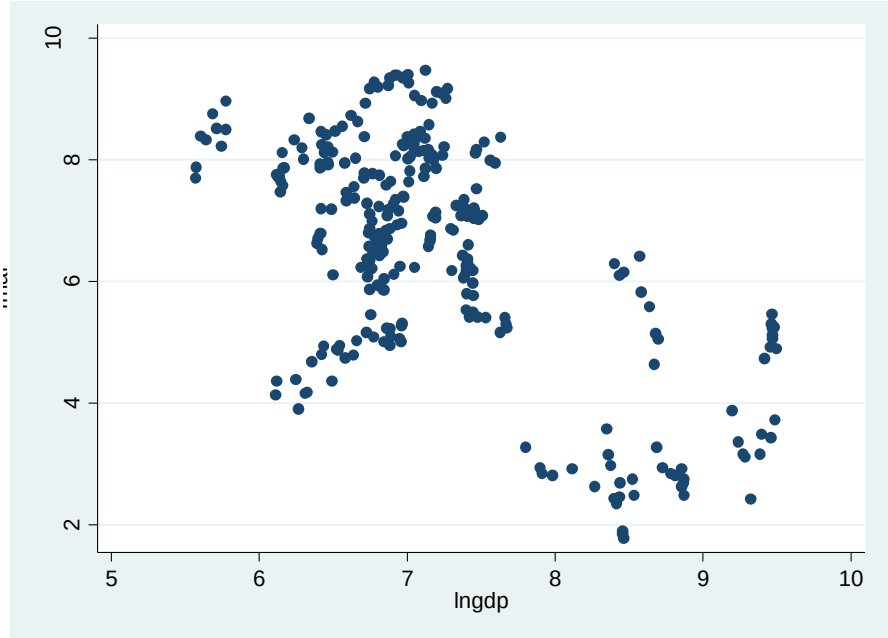
MODEL I:



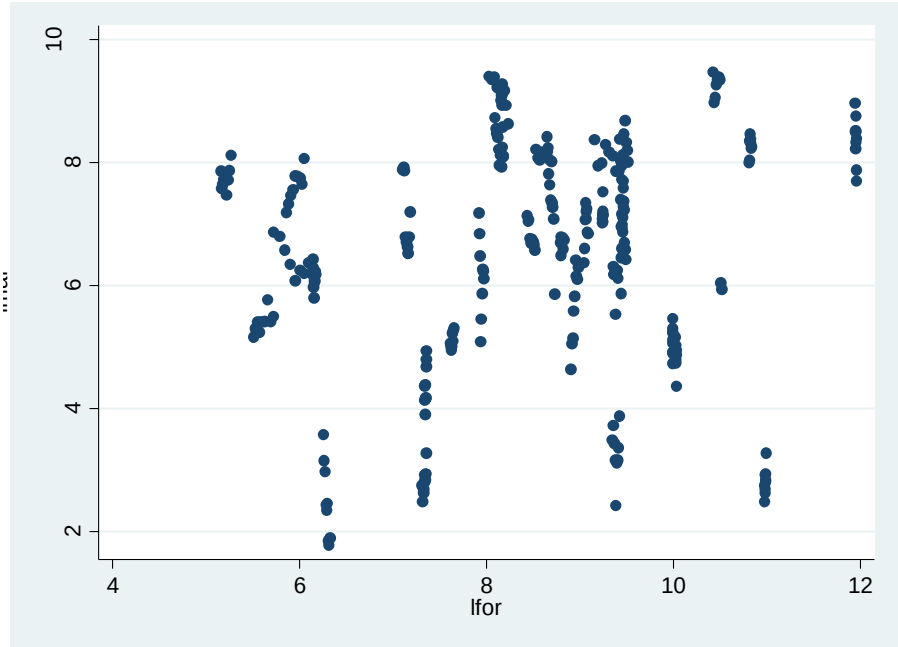
Two way scatterplot of residuals vs lmal



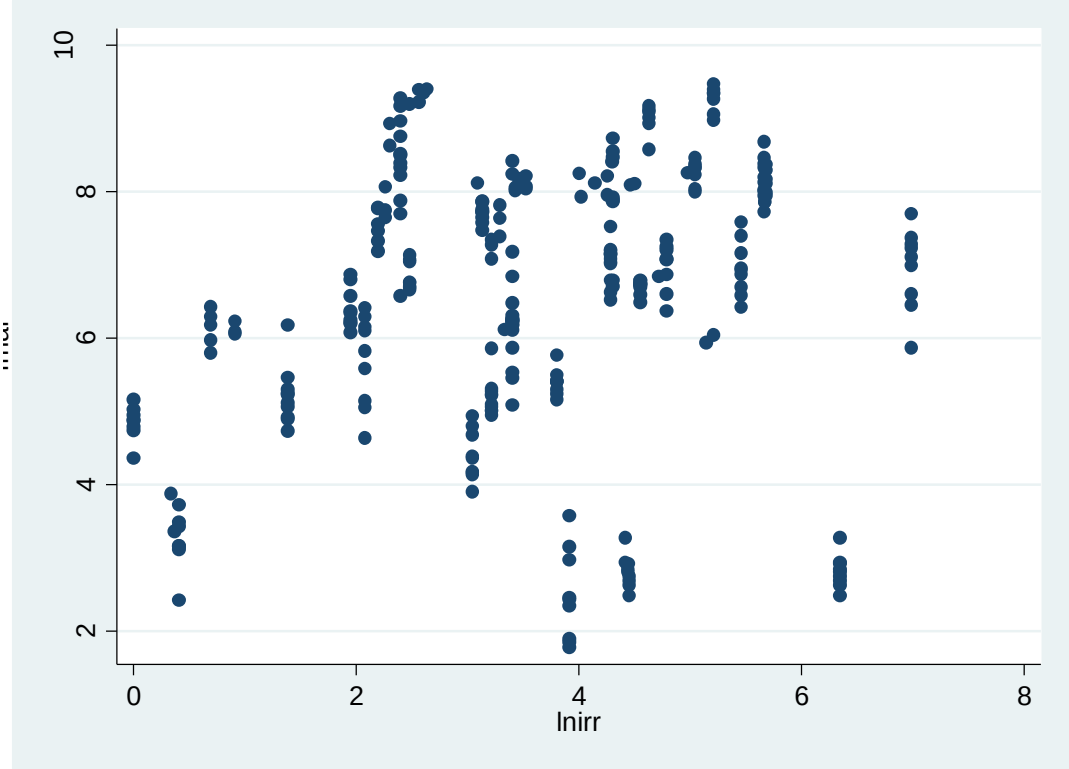
Histogram of residuals for model 1



Two way scatterplot of lmal vs lngdp



Two way scatterplot of lmal vs lfor

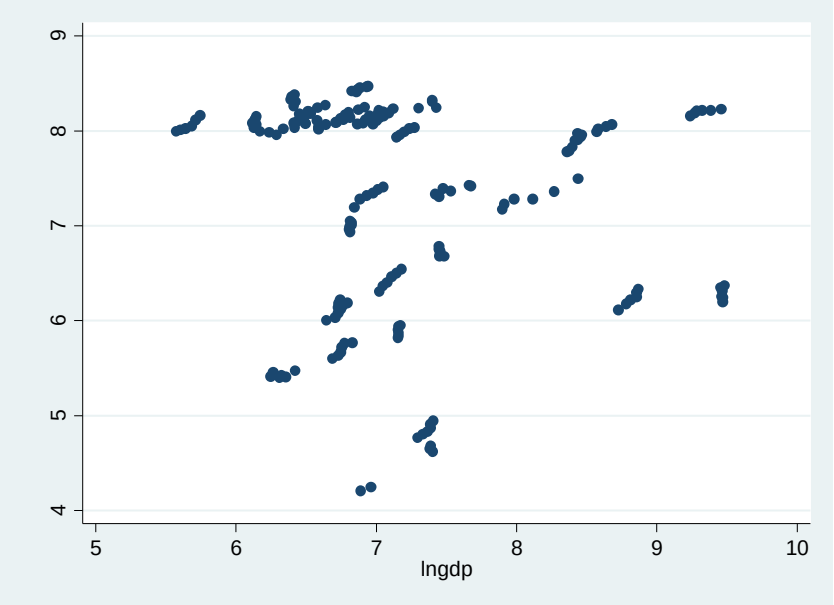


Scatterplot of malaria vs irrigation

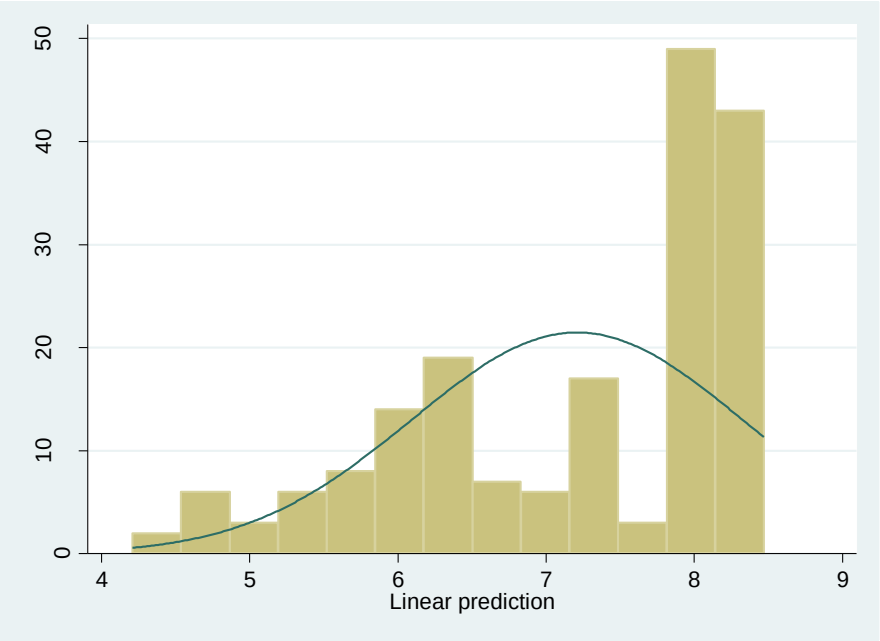
Ran a Breusch-pagan test for heteroskedasticity of fitted values of lmal

```
. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lmal

         chi2(1)      =     1.34
         Prob > chi2  =   0.2479
```

MODEL II:



Two way scatter of uhat vs lngdp



Histogram of residuals for model 2