Lecture 25
AAE 374 Fall 2008

- *Announcements:*
  - Final exam: Tues (12/16) at 2:45 pm in Grainger 2120
  - Final format same as midterm
  - Can we schedule the discussion review this week and do a review on Sunday 12/14 (Andres available both times)
  - Next week: special guest Michael Carter

Overview

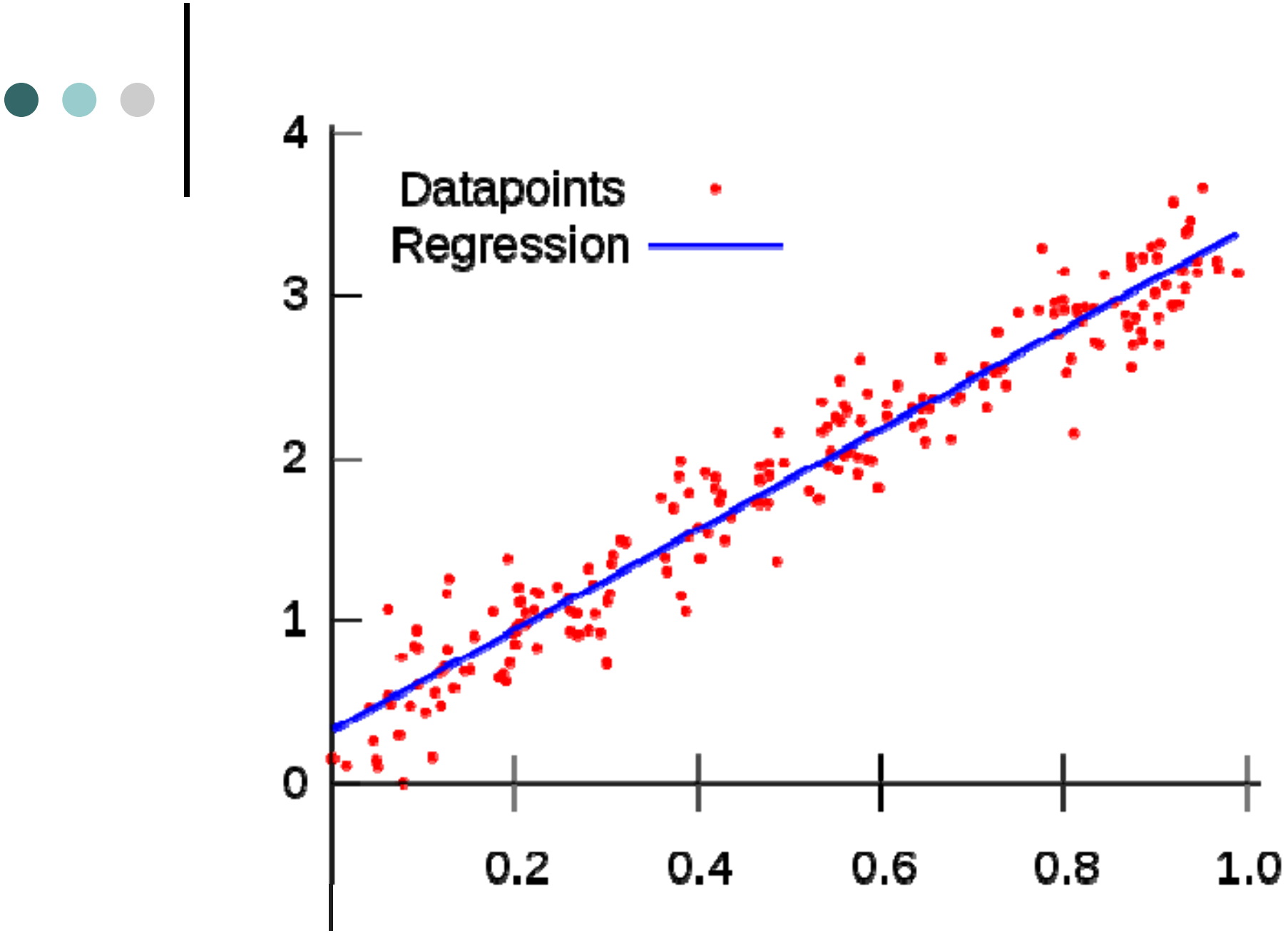
- Relative importance of Institutions, Integration and Geography (Rodrik, Subramanian & Trebbi)

- Before all that: Brief-up on Regression as an analysis tool

Regression Digression

(How to become a social scientist in under 10 minutes)

- Regression: The ubiquitous tool of social science research
- Many ways of characterizing regression as a statistical tool to "convert data into results"
- Geometrically the line of best fit, so captures relationships between variables


Datapoints
Regression
4
3
2
1
0
0.2
0.4
0.6
0.8
1.0

Regression Example

Simple example: Lets say Brad's basketball team plays pick-up games against other faculty teams

1. Each game is numbered by i, i from 1 to 30 and if Brad's team wins we record it as $Y_i = 1$ and if they lose $Y_i = 0$
2. Every once in a while Michael Jordan drives up to play on Brad's team, and when MJ plays we record it as $M_i = 1$ and if not $M_i = 0$
3. We want to access the effect of MJ playing on the win percentage and we can write a "regression model"

$$Y_i = \alpha + \beta M_i + \varepsilon_i$$

Interpreting a regression equation

$$Y_i = \alpha + \beta M_i + \varepsilon_i$$

here β captures the effect of MJ playing (expected sign?), α is the average win percentage w/o MJ and ε_i is some random noise (team had a really good/bad game that day, etc.)

Here the left hand side is called the dependent variable and the right hand side the independent variables or regressors, the idea being the outcome of a game "depends" on how good Brad's team is w/o MJ, whether MJ plays, etc.

Significance levels

The other dimension is statistical significance: if we see $\beta > 0$, is the value of β large enough (in magnitude) to be statistically different from zero? Typically papers report this as one to three stars (* to ***) in increasing significance – no stars means that we cannot conclude the parameter is different from zero at "conventional" levels of statistical confidence.

A growth example

6. Actually, we've been using regression results to talk about a lot of things, e.g. growth. We've seen the following growth equation:

$$\hat{y} = \hat{A} + \alpha\hat{k} + \beta\hat{s}$$

Elsewhere we've said that $\alpha = 0.39$ and that $\beta = 0.27$, essentially where this is coming from is taking country data on output growth, capital growth and human capital growth and running the following regression:

$$\hat{y}_t = \alpha\hat{k}_t + \beta\hat{s}_t + \hat{A}_t$$

There are popular packages in the labs (SAS, SPSS, STATA) you can also use Excel or the (free!) R package at home

Institutions Rule: Issues at stake

- 100:1 ratio between richest country, Luxemborg and Sierra Leone in 2000, or divergence in income per capita or diverse economic development outcomes.
- More specifically the **relative** importance of geography, integration (e.g. trade), and institutions in this outcome.
- Deeper drivers of physical and human capital accumulation and its endogenous growth variant, technological change.

Challenges to comparison

- Deeper determinants are endogenous to one another. Higher income could drive better institutions, better institutions could allow for more effective integration, and so-on.

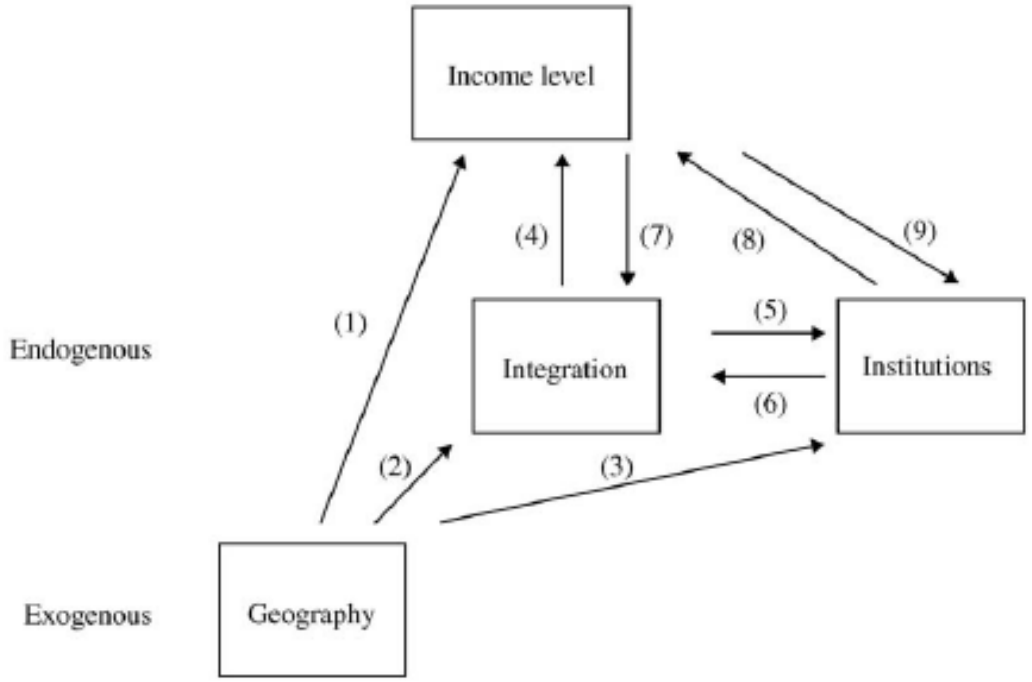


Figure 1. The ''deep'' determinants of income.

“Big Three”: Institutions, geography and integration

- Institutions: property rights, savings, and investment
- Institutions: technology transfer
- Institutions: innovation/R&D
- (Engerman and Sokoloff): factor endowments (soils, local conditions) shaped economic opportunities and institutions to perpetuate inequality.

“Big Three”: Institutions, geography and integration

- Geography: agricultural productivity differences in tropics and temperate areas.
- Geography: Natural resources and rents leading to Dutch Disease or Resource Curse
- Guns, Germs, and Steel:
  - East/West axis
  - Dynamic process of food surpluses and food storage to large, dense, sedentary, and stratified societies that lead to technological change.
  - Location and tech transfer

“Big Three”: Institutions, geography and integration

- Trade: comparative advantage allows specialization, leads to higher incomes which drive accumulation and growth.
- Trade: leads to more competition that stimulates innovation and efficiency.
- Trade and FDI lead to learning, technology transfer, and increased opportunities.

Institutions Rule: A first go

- Simple bivariate comparison of these three versus GDP per capita. All have a strong positive relationship, especially rule of law. (Left column is the smaller "good sample")

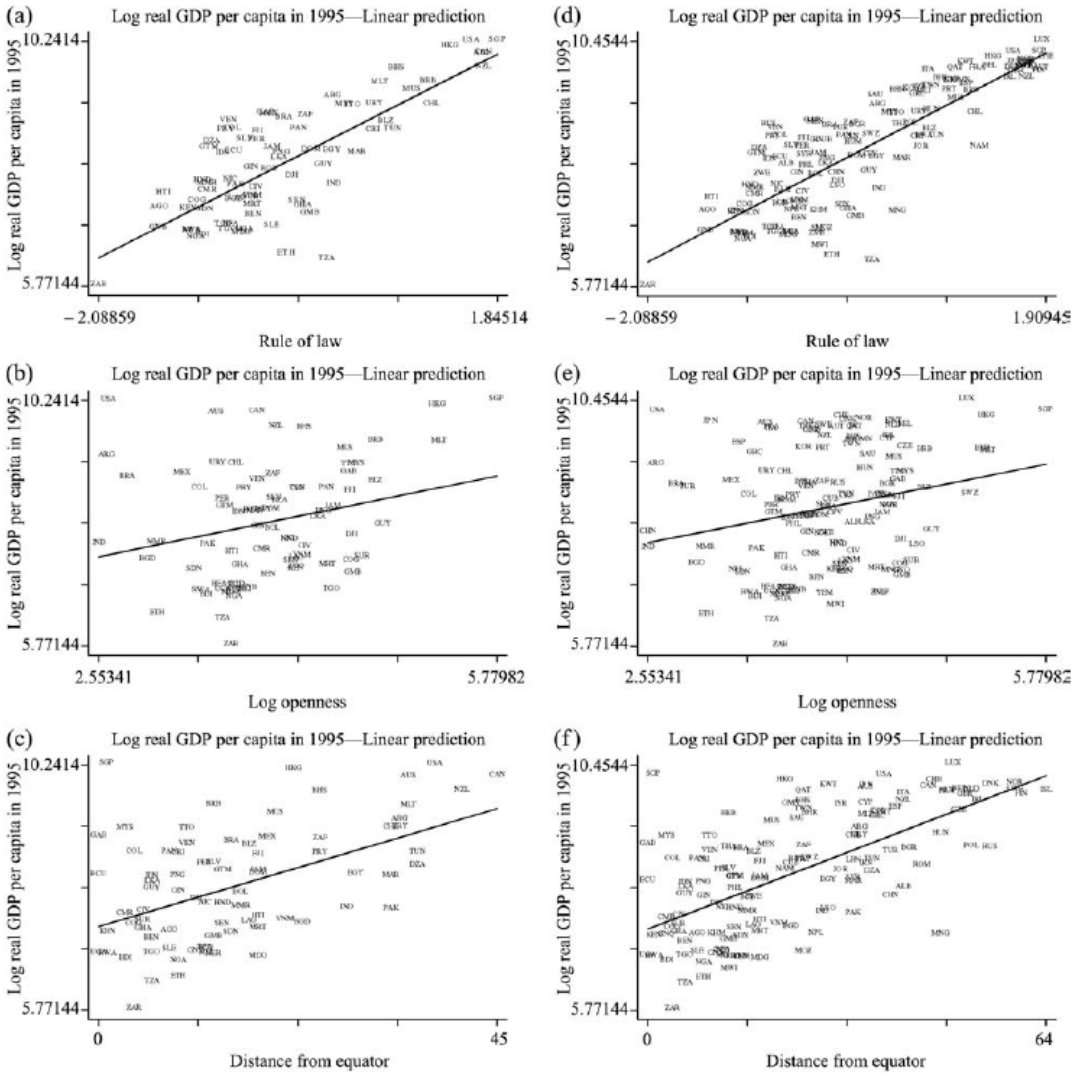


Figure 2. Simple correlations between income and its determinants (sample of 79 countries for (a)–(c); sample of 137 countries for (d)–(f)).

Institutions Rule: Empirical Strategy

- What Rodrik, Subramanian and Trebbi want to do is run a regression like this:

$$\log y_i = \mu + \alpha \mathrm{INS}_i + \beta \mathrm{INT}_i + \gamma \mathrm{GEO}_i + \varepsilon_i$$

While geography effects are probably not effected by institutions and integration, geography effects these two, which in turn effect each other.

For example, EU countries are fairly integrated, making measurements of INT high, so we might find beta is large compared to gamma, although really the geographical effect is driving integration and could be more important.

First, let's ignore these endogeneity issues, then revisit…

Measurements for institutions and integration?

$$\log y_i = \mu + \alpha \mathrm{INS}_i + \beta \mathrm{INT}_i + \gamma \mathrm{GEO}_i + \varepsilon_i$$

- Institution index: captures strength of rule of law and protection for property rights (Kaufmann, 2002)
- Integration: ratio of trade to gdp
- Geography: distance from the equator in degrees (i.e. N vs S in US, Brazil)

Table 2. Determinants of development: Core specifications, ordinary least squares estimates.

	Log GDP per capita								
	Acemoglu et al. Sample			Extended Acemoglu et al. Sample			Large Sample		
Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Geography (DISTEQ)	0.74 (4.48)*	0.20 (1.34)	0.32 (1.85)**	0.80 (5.22)*	0.22 (1.63)	0.33 (2.11)**	0.76 (10.62)*	0.20 (2.48)**	0.23 (2.63)*
Institutions (RULE)		0.78 (7.56)*	0.69 (6.07)*		0.81 (9.35)*	0.72 (6.98)*		0.81 (12.12)*	0.78 (10.49)*
Integration (LCOPEN)			0.16 (1.48)			0.15 (1.53)			0.08 (1.24)
Observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.57	0.59	0.26	0.61	0.62	0.42	0.71	0.71

Notes: The dependent variable is per capita GDP in 1995, PPP basis. There are three samples for which the core regressions are run: (i) the first three columns correspond to the sample of 64 countries in Acemoglu et al. (2001); (ii) columns (4)–(6) use a sample of 79 countries for which data on settler mortality (LOGEM4) have been compiled by Acemoglu et al.; and (iii) columns (7)–(9) use a larger sample of 137 countries. The regressors are: (i) DISTEQ, the variable for geography, which is measured as the absolute value of latitude of a country; (ii) Rule of law (RULE), which is the measure for institutions; and (iii) LCOPEN, the variable for integration, which is measured as the ratio of nominal trade to nominal GDP. All regressors are scaled in the sense that they represent deviations from the mean divided by the standard deviation. All regressors, except DISTEQ and RULE, in the three panels are in logs. See the Appendix for more detailed variable definitions and sources. *t*-statistics are reported under coefficient estimates. Significance at the 1, 5, and 10 percent levels are denoted respectively by *, **, and ***.

Institutions Rule: Endogeneity Fixes

$$\log y_i = \mu + \alpha \mathrm{INS}_i + \beta \mathrm{INT}_i + \gamma \mathrm{GEO}_i + \varepsilon_i$$

- Okay, but one thing might be causing the other so we don't really know the relative importance. All estimates are (+), although results suggest institutions dominate.

- RST try to estimate these interrelationships to capture endogenous relationships which lets us account for them and isolate the effects. Let's see how they did this…

Instruments for Institutions and integration

$$\mathrm{INS}_i = \lambda + \delta \mathrm{SM}_i + \phi \mathrm{CONST}_i + \psi \mathrm{GEO}_i + \varepsilon_{\mathrm{INS}_i},$$

$$\mathrm{INT}_i = \theta + \sigma \mathrm{CONST}_i + \tau \mathrm{SM}_i + \omega \mathrm{GEO}_i + \varepsilon_{\mathrm{INT}_i},$$

- The instruments are SMi (settler mortality) and CONSTi (constructed trade openness measure recovered from “gravity” model of trade).
- SMi: higher settler mortality makes institutions less robust, people are less likely to settle in and build good institutions. High settler mortality areas are ones that try to get riches out fast.
- CONSTi: trade openness is endogenous but you can predict it based on measures of country mass, distance between trade partners, and geographical variables.

Table 3. Determinants of development: Core specifications, instrumental variables estimates.

	Acemoglu et al. Sample			Extended Acemoglu et al. Sample			Large Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Second-stage: Dependent variable = Log GDP per capita									
Geography (DISTEQ)	0.74	−0.42	−0.56	0.80	−0.45	−0.72	0.76	−0.06	−0.14
	(4.48)*	(−1.19)	(−1.23)	(5.22)*	(−1.26)	(−1.38)	(10.62)*	(−0.5)	(−0.93)
Institutions (RULE)		1.68	1.78		1.75	1.98		1.19	1.30
		(4.29)*	(3.78)*		(4.42)*	(3.56)*		(8.02)*	(7.09)*
Integration (LCOPEN)			−0.18			−0.31			−0.15
			(−0.71)			(−1.10)			(−1.09)
No. of observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.54	0.56	0.26	0.51	0.52	0.417	0.51	0.56
Test for over-identifying restrictions (*p*-value)								(0.0089)	(0.0354)
Panel B: First Stage for Endogenous Variables (Institutions (RULE) and Integration (LCOPEN))									
Dependent variable	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN
Geography (DISTEQ)	0.41	0.47	−0.25	0.47	0.54	−0.18	0.67	0.66	−0.05
	(2.8)*	(3.21)*	(−2.00)**	(3.34)*	(3.87)*	(−1.37)	(10.81)*	(11.23)*	(−0.84)
Settler mortality	−0.39	−0.40	−0.30	−0.34	−0.34	−0.27			
(LOGEM4)	(−3.87)*	(−4.1)*	(−3.51)*	(−3.69)*	(−3.82)*	(−3.22)*			
Population speaking English (ENGFRAC)							0.19	0.18	0.17
							(2.69)*	(2.69)*	(2.65)*
Population speaking other European langages (EURFRAC)							0.14	0.17	−0.11
							(1.94)**	(2.55)**	(−1.67)**
Constructed openness	na	0.20	0.90	na	0.19	0.80	na	0.23	0.70
(LOGFRANKROM)		(1.95)**	(10.32)*		(2.16)**	(9.67)*		(3.99)*	(12.33)*
F-statistic	22.9	17.2	41.7	24	18.5	36.9	50.09	45.79	41.39
R-square	0.41	0.44	0.66	0.37	0.40	0.58	0.52	0.57	0.54
Partial *R*-square		0.16	0.58		0.12	0.51		0.18	0.52
corr(RULEFIT, LCOPENFIT)			0.14			0.21			0.27

Table 3. Determinants of development: Core specifications, instrumental variables estimates.

	Acemoglu et al. Sample			Extended Acemoglu et al. Sample			Large Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Second-stage: Dependent variable = Log GDP per capita									
Geography (DISTEQ)	0.74	−0.42	−0.56	0.80	−0.45	−0.72	0.76	−0.06	−0.14
	(4.48)*	(−1.19)	(−1.23)	(5.22)*	(−1.26)	(−1.38)	(10.62)*	(−0.5)	(−0.93)
Institutions (RULE)		1.68	1.78		1.75	1.98		1.19	1.30
		(4.29)*	(3.78)*		(4.42)*	(3.56)*		(8.02)*	(7.09)*
Integration (LCOPEN)			−0.18			−0.31			−0.15
			(−0.71)			(−1.10)			(−1.09)
No. of observations	64	64	64	79	79	79	137	137	137
R-square	0.25	0.54	0.56	0.26	0.51	0.52	0.417	0.51	0.56
Test for over-identifying restrictions (*p*-value)								(0.0089)	(0.0354)
Panel B: First Stage for Endogenous Variables (Institutions (RULE) and Integration (LCOPEN))									
Dependent variable	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN	RULE	RULE	LCOPEN
Geography (DISTEQ)	0.41	0.47	−0.25	0.47	0.54	−0.18	0.67	0.66	−0.05
	(2.8)*	(3.21)*	(−2.00)**	(3.34)*	(3.87)*	(−1.37)	(10.81)*	(11.23)*	(−0.84)
Settler mortality	−0.39	−0.40	−0.30	−0.34	−0.34	−0.27			
(LOGEM4)	(−3.87)*	(−4.1)*	(−3.51)*	(−3.69)*	(−3.82)*	(−3.22)*			
Population speaking English (ENGFRAC)							0.19	0.18	0.17
							(2.69)*	(2.69)*	(2.65)*
Population speaking other European langages (EURFRAC)							0.14	0.17	−0.11
							(1.94)**	(2.55)**	(−1.67)**
Constructed openness	na	0.20	0.90	na	0.19	0.80	na	0.23	0.70
(LOGFRANKROM)		(1.95)**	(10.32)*		(2.16)**	(9.67)*		(3.99)*	(12.33)*
F-statistic	22.9	17.2	41.7	24	18.5	36.9	50.09	45.79	41.39
R-square	0.41	0.44	0.66	0.37	0.40	0.58	0.52	0.57	0.54
Partial *R*-square		0.16	0.58		0.12	0.51		0.18	0.52
corr(RULEFIT, LCOPENFIT)			0.14			0.21			0.27

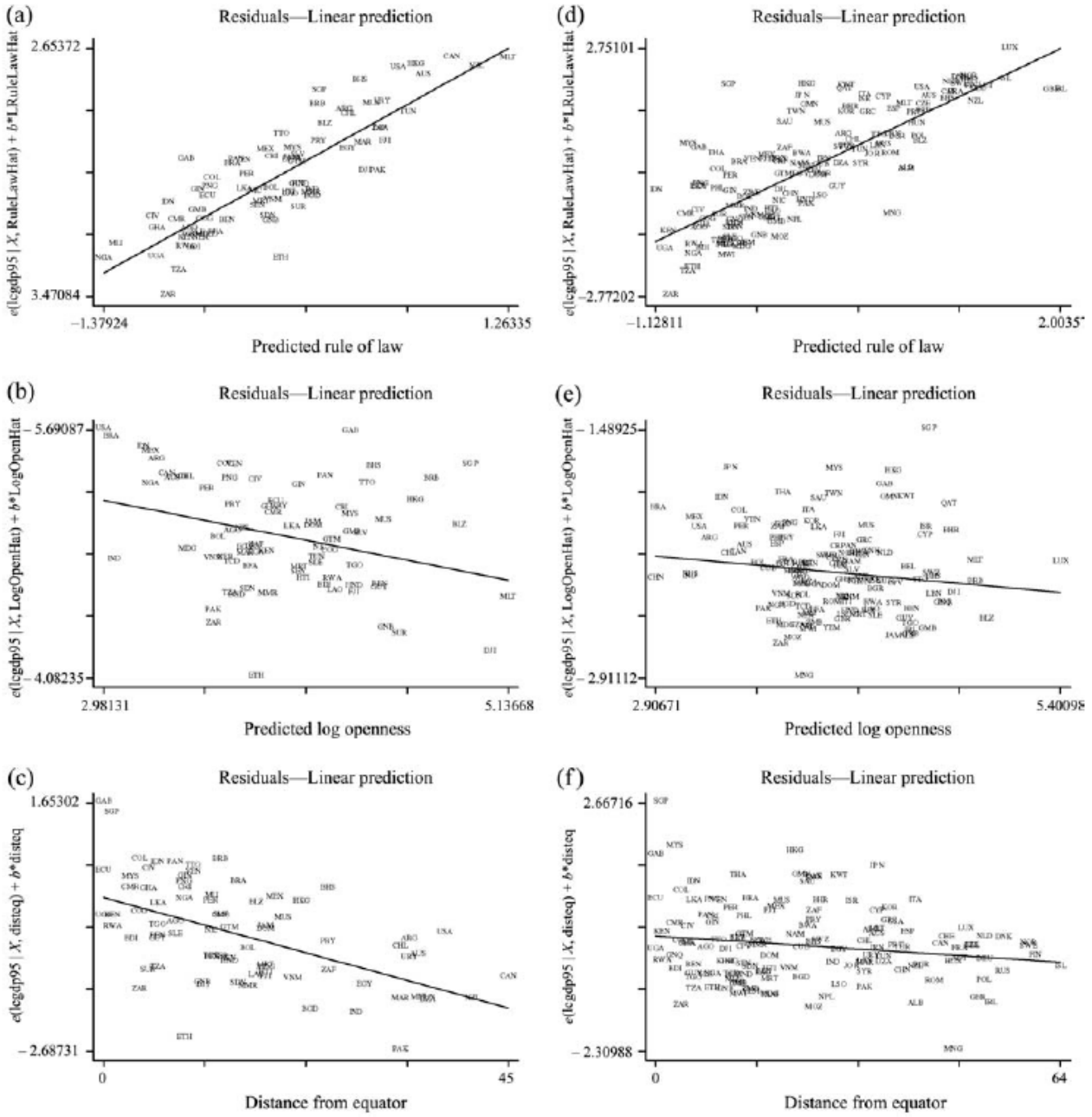


Figure 3. Conditional correlations between income and its determinants (sample of 79 countries for (a)–(c); sample of 137 countries for (d)–(f)).

Take Home Messages

- Institutions Rule but Instruments DO NOT MAKE A THEORY.
- What do they mean by that? Basic idea is that the instrument does not provide a basis for making cause and effect arguments, only that we are onto a key pathway for explanation but not how it works.
- They critique previous authors for stretching their argument.
- First lesson is thus what the paper *does not do*.

Policy implications?

- Institutional quality DOES NOT MEAN THAT POLICY CANNOT BE EFFECTIVE
- Many institutions have been successfully reformed. South Korea during the 1960s, and China since the 1980s. India may be another example.
- Why? Policy = Flow variable; Institutions = Stock variable
- As a reminder, income would be a flow and wealth would be a stock. Think of institutions as the cumulative outcome of past policy actions (both private and public?).
- Institutions do change and countries work on changing them.

So is this informative?

- Yes and No. Integration may not be the answer, geography is not necessarily fate.
- Which institutions? Take property rights: Great contrast between China and Russia since the 1980s.
  - China has maintained a socialist legal system and lots of collective property right regimes (leased not owned land).
  - Russia went private with the fall of Communism.
  - Investment has boomed in China and uncertainty plagues Russia.
- Successful institutions appear very context specific. E.g. North America, Western Europe, and Japan can have such different institutional contexts and yet do well.
- Need for more cross-national studies of what institutions work and don’t work.

Next time

- Milanovic
- Michael Carter
- One last thing…

Get well soon Brad!!!