Seminar to Advanced Macroeconomics

Economic Growth Reconsidered: Convergence and the Institutional Dimension

What to do if too many variables might be relevant?

Jaromír Baxa
jaromir.baxa@centrum.cz

Outline

- Too many variables and what to do
    - Principal components
- Application: Economic growth and Institutions
    - Institutions and catch-up
    - Measuring institutional quality
    - Application in empirical growth theory

Too many variables...

- The problem: too many variables suspect to be relevant for the process we want to explain.
- All into one regression?
  - many insignificant
  - degrees of freedom to low and parsimony is lost
  - multicollinearity might occur
- And what to do?
  - Any ideas?

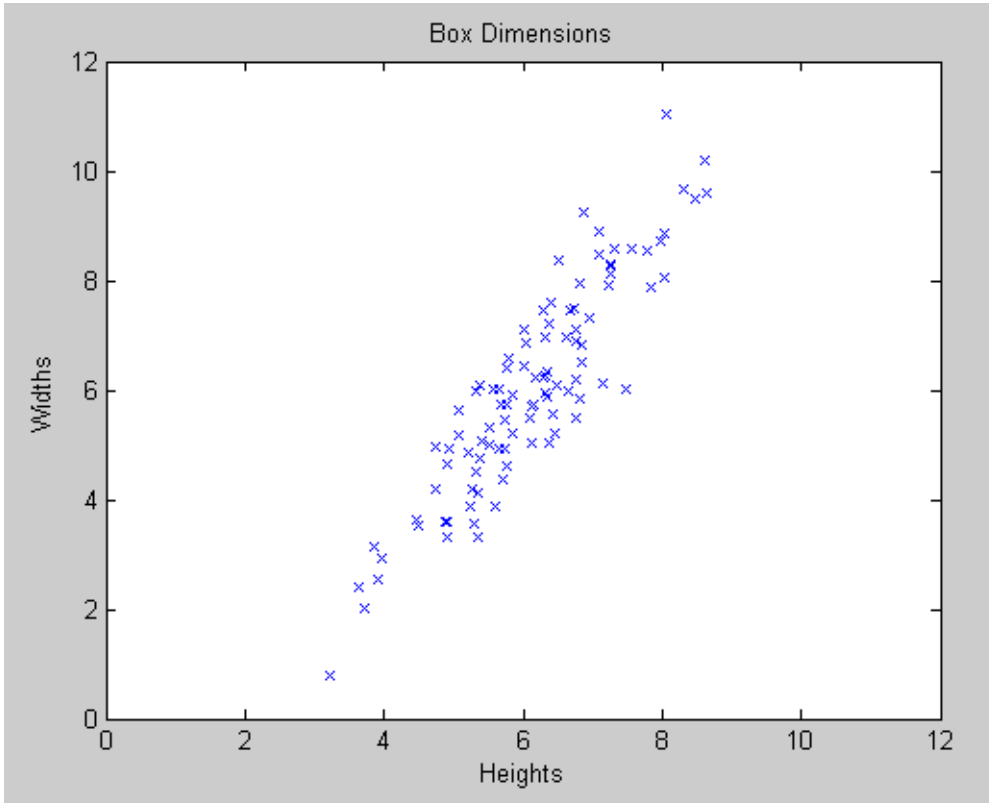
Too many variables...

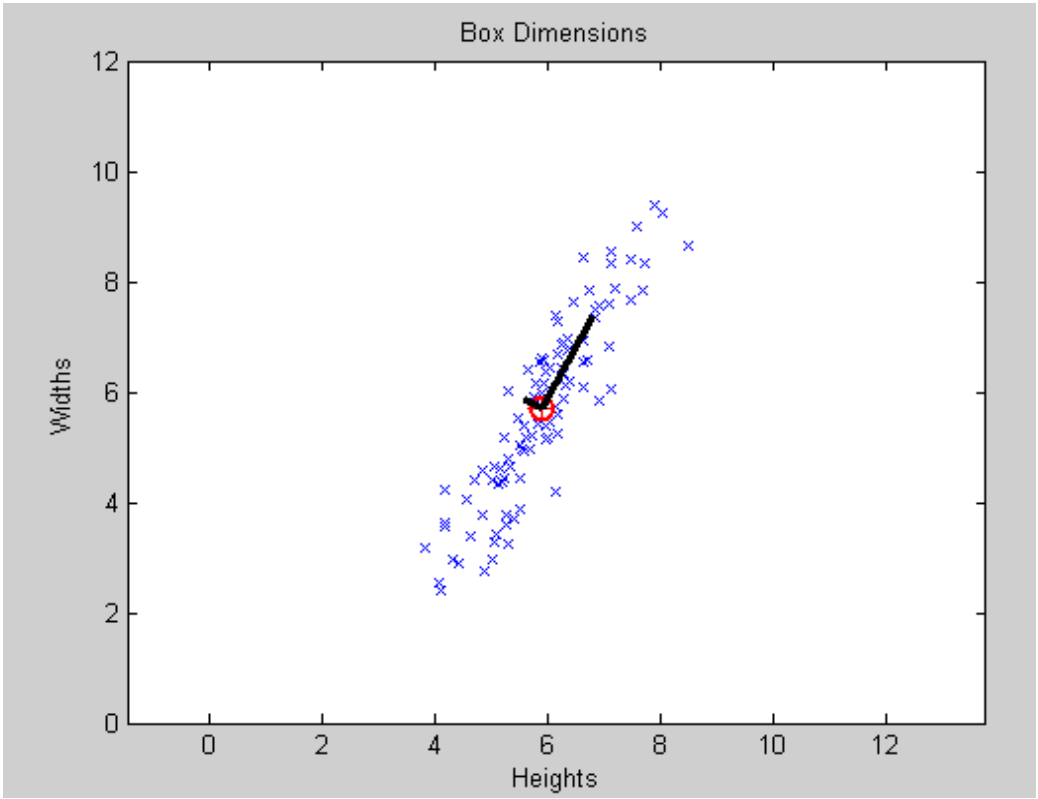
- And what to do?
  - Any ideas?
  - t-tests and F-tests
  - linear combination of variables
  - principal components method (*metoda hlavních komponent)*

Principal Components

- Linear transformation that converts data to the new coordinate system in a manner that the highest variance is on the first axis, the second highest on the second and so on.
- Each axis represents new, artificial variable that explains some part of the variance of original data.
- Then the most important ones are called the *principal components* of the system.

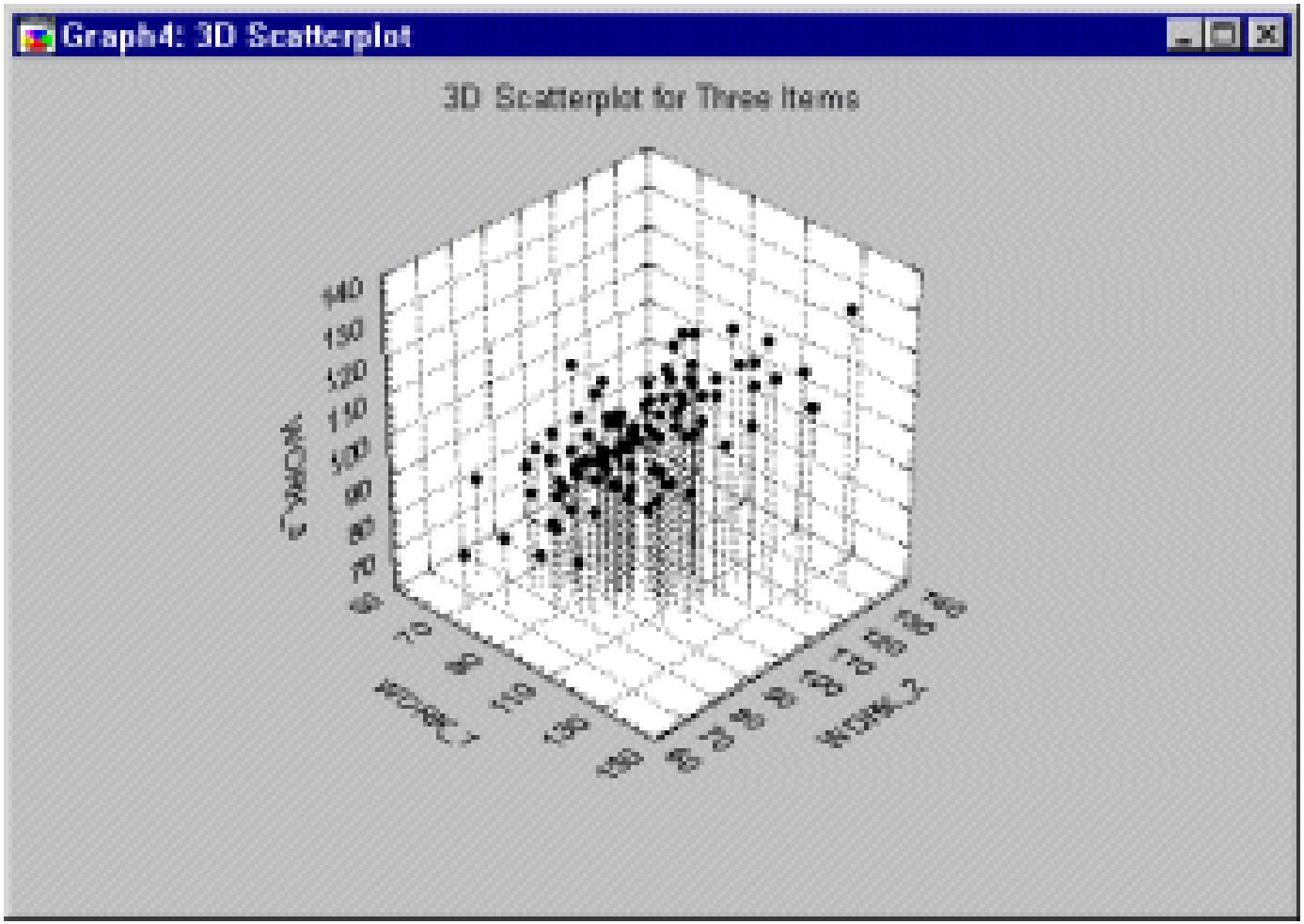
Principal Components





Source: http://brandon-merkl.blogspot.com/2006/04/principal-components-analysis.html

Principal Components



http://www.statsoft.com/textbook/stfacan.html

Principal Components

- Technically: based on spectral decomposition of the covariance/correlation matrix and their eigenvalues.
- Correlation matrix should be preferred for datasets with many different variables of different scales and units.

Economic Growth and Institutions

- Original hypothesis: Convergence poorer countries should grow faster than richer ones because:
  - they can take advantage of technological advances of the more developed countries (so called relative backwardness advantage; A. Gerschenkron (1952))
  - diminishing returns on capital
- Cons to the convergence hypothesis
  - different institutions and habits (Karl Polanyi, 1944)
  - low investments into both physical and human capital => Solow (1956)
  - Conditional convergence = if countries did not vary in their investment and population growth rates there would be a strong tendency to growth (Mankiw, Romer, Weil, 1992)

Institutions and Catch-Up

- Institutions: “rules of the game”;
  - structures and mechanisms of social order and cooperation
  - governing the behavior of a set of individuals
  - constraints, that shape the set of opportunities
  - decrease uncertainty (behavior of others is anticipated easily).
- Laws (written, formal, resulting from political process) & Norms, habits, tradition, culture (informal)
- Effect of “good” instiutions: increased investment incentives – possible to adopt routines maximizing long-run profit. Lower transaction costs thanks to culture of „trust“.
- Effects of “poor” institutions: degrade the security of property rights, not profitability but personal ties important...

Measuring Institutional Quality

- How to measure institutional quality?
  - *Ideal measurement:* objective evaluations comparable across countries and over time, indicators of security of property rights... And with long history of the data.
  - *Realistic approaches:* various indices: corruption (TI), trust (World Values Survey), freedom (Freedom House); costs of contract enforcement, firing and hiring people, starting and closing business (WB Doing business).
  - *Commercial datasets:* International Country Risk Guide, Business Environmental Risk Intelligence

Measuring Institutional Quality

- Property rights security: sometimes political stability as a proxy used. Better: risk of nationalization, contract enforceability, risk of repudiation of contracts by government.
- Rule of law: costs and delays in judiciary processes, costs of closing business and bancrupcies, stable political system and executive constraints
- Bureaucratical quality

Testing the Role of Institutions

- What we used: the World bank database Doing Business and the WDI database.
- Our specification follows Knack and Keefer 1998, however their dataset was to costly.
- Growth equation: whether convergence occurs and whether it can be explained using institutional factors (*not shown during the seminar – use the "growth" variable as dependent and add the "gap" into independent ones; however very poor results*)
- Level equation: level of income related to economic and institutionalist variables.

Testing the Role of Institutions

- Enforcement: costs of enforcement relatively to debt; proxy for transaction cost. Low level indicates high transaction costs.
- Protecting investors: proxy for property rights. The more the better.
- Closing variable: recovery rate when closing business. The more the better. Similar to the Enforcement variable
- Dbrank: ranking from the Doing business database.

What we did (gretl commands)

```
/* model 1 */
ols gdp2003 const popgrowth iratio
/* model 2 */
ols gdp2003 const popgrowth iratio PrimEnrollment SecondEnrollmen
TertEnrollment
/* model 3 */
ols gdp2003 const popgrowth iratio PrimEnrollment SecondEnrollmen
TertEnrollment EaseOfBusinessR StartNoProcedur StartTimeDays  StartingCostToI
RigidityoEmploy InvestorProtect PayingTaxes  EnforcementTime EnforcementCost
RecoveryWhenBan
/* Omit the insignificant variables */
omit StartNoProcedur StartTimeDays StartingCostToI RigidityoEmploy
InvestorProtect PayingTaxes EnforcementTime EnforcementCost
/* Principal component analysis on the institutional variables */
pca EaseOfBusinessR StartNoProcedur StartTimeDays StartingCostToI RigidityoEmploy
InvestorProtect PayingTaxes EnforcementTime EnforcementCost RecoveryWhenBan --save-all
/* model 4 */
ols gdp2003 const popgrowth iratio PrimEnrollment SecondEnrollmen
TertEnrollment PC1 PC2
```

NOTE: It seems gretl does not recognize the --save-all command at the end of the pca command, when it is copied from outside to the command or scipt window. Thus: 1) Copy the script line, 2) Delete --save-all, 3) Write --save-all on the same place (now it should be green - the correct syntax is highlited now)

Interpreting the PCA Output

- Eigenanalysis: List of PC's

Component	Eigenvalue	Proportion	Cumulativ
1	3.7883	0.3788	0.3788
2	1.3709	0.1371	0.5159
3	0.9975	0.0998	0.6157
4	0.9546	0.0955	0.7111
5	0.7927	0.0793	0.7904
6	0.6752	0.0675	0.8579
7	0.5177	0.0518	0.9097
8	0.4289	0.0429	0.9526
9	0.3319	0.0332	0.9858
10	0.1422	0.0142	1.0000

Based on these numbers we choose the appropriate number of principal components for the regression.
Rule: These, whose eigenvalue higher than 1.
Or up to variance we want to keep in our model.

Eigenvalue of PC

New components From the most important one to the least important one

Proportion of variance explained by each PC separately (1st one explains 38% of the o original variance, the 2nd the next 14% etc.)

Cumulative sum of explained variance: This says, that 4 PC's explain together 71%.

Interpreting the PCA Output

- Eigenvectors

Variable	PC1	PC2	PC3
EaseOfBusinessI	-0.471	0.082	-0.080
StartNoProcedur	-0.350	0.176	0.264
StartTimeDays	-0.302	-0.087	0.394
StartingCostToI	-0.343	-0.194	-0.505
RigidityoEmploy	-0.285	0.222	-0.273
InvestorProtect	0.295	-0.340	-0.236
PayingTaxes	-0.124	0.599	-0.167
EnforcementTime	-0.173	-0.352	0.499
EnforcementCost	-0.252	-0.500	-0.322
RecoveryWhenBan	0.411	0.142	-0.058

Original - "Old" - Variables

Correlations among the new and the old variables

New variables:
The principal components (only 3 included here)

This part says:
How much is each of our principal components correlated with the original variables.
The PC1 is mostly determined by EaseOfBusiness, StartingProcedures and Costs, RecoveryWhenBan
=> "Enterprise-friendly institutions"
The PC2: PayingTayes and Enforcement
=> "Rule of Law".
… (PC3? What name?)

Results:
2 PC's with eigenvalue higher than 1
Both explain 51% of the original variance
According to the correlations among old and new variables we found nice names for them.
We can run regression with 2 institutional variables – PC1 as Enterprise-Friendly, PC2 as RuleOfLaw – instead of 10 original variables.

Comparing with different results

- Generally not so good results as Keefer and Knack (1998) obtained
- Why: another proxies for institutions, broader set of countries, too short period (only 8 years comparing to almost 30 – perhaps the most important aspect)
- Or the effect is really very poor and the whole institutionalist stuff can be forgotten (despite at least 4 Nobel Prize winners since 1990).

TABLE II

Rule of Law, Contract Enforceability and Convergence

Dependent variable: average real per capita growth in GDP, 1960–1989

Variables	Rule of Law		Contract Enforceability	
	(1)	(2)	(3)	(4)
Constant	–8.014 (2.291)	–8.982 (2.360)	–3.890 (3.961)	–2.934 (3.976)
Primary Enrollment	3.205 (0.725)	3.028 (0.710)	3.255 (1.226)	1.716 (1.217)
Secondary Enrollment	3.433 (1.155)	3.045 (1.136)	2.958 (1.182)	3.344 (1.364)
Labor Force Growth	1.732 (1.033)	1.571 (0.730)	2.497 (1.377)	1.977 (1.009)
Price Changes	1.263 (0.551)	1.160 (0.538)	0.101 (0.858)	–0.277 (0.794)
Income Gap	1.101 (0.333)	1.329 (0.350)	1.223 (0.496)	1.487 (0.455)
Institutional Variable		0.467 (0.129)		0.826 (0.408)
Institution × Income Gap		0.266 (0.123)		1.119 (0.430)
N	97	97	47	47
Adj. *R*-Square	0.295	0.406	0.256	0.398

Note: White-corrected standard errors are in parentheses. Institutional variables are rule of law (columns 1 and 2) and contract enforceability (columns 3 and 4).

Key Points

1. Normalization of variable and linear combination of variables
2. Collinearity, significance, joint-significance
3. Principal component analysis: motivation and the intuition of the procedure
4. Using principal components: understanding the output