

Measuring the Independence of Regulatory Agencies

Chris Hanretty and Christel Koop

October 12, 2009

Introduction

Why do this research?

What data did we gather?

Measurement

How it's been done before

How we want to measure

Main (measurement-related) findings

The 'difficulty' of each item

Index items drop out

Why this matters

This changes our conclusions about independence

Outline

Introduction

- Why do this research?

- What data did we gather?

Measurement

- How it's been done before

- How we want to measure

Main (measurement-related) findings

- The 'difficulty' of each item

- Index items drop out

Why this matters

- This changes our conclusions about independence

Because independence matters to...

- Economists (independent agencies affect pricing regimes, market entry, etc.,
- Political scientists (because we're interested in consequences and effects of delegation)

Because independence matters to...

- Economists (independent agencies affect pricing regimes, market entry, etc.,
- Political scientists (because we're interested in consequences and effects of delegation)

Because there's a lot of work being done

- Gilardi's (2005) work on IRAs within Europe
- Jordana and Levi-Faur's (2010) extension of this worldwide

So let's get measurement right!

Because there's a lot of work being done

- Gilardi's (2005) work on IRAs within Europe
- Jordana and Levi-Faur's (2010) extension of this worldwide

So let's get measurement right!

Because there's a lot of work being done

- Gilardi's (2005) work on IRAs within Europe
- Jordana and Levi-Faur's (2010) extension of this worldwide

So let's get measurement right!

We looked at regulatory agencies in...

Competition	ICN	101
Financial markets	IOSCO	100
Communications	ITU	132
Energy	IERN	73
Pharmaceuticals	WHO	43
Food safety	WHO	37
Environment	IMPEL	15

We got good response rates...

Competition	58.4%
Financial markets	35.0%
Communications	23.4%
Energy	46.6%
Pharmaceuticals	39.5%
Food safety	27.0%
Environment	46.6%

...despite asking many questions

Agency head questions relating to appointment, term length, dismissal, incompatibility

Agency board as above

Accountability reporting requirements to executive, legislature

Operations who has final say over staffing, resourcing, audit, and internal org.?

Scope does the agency have exclusive competence?

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

The modal regulator...

- Has a CEO appointed by the executive
- for a renewable term of five years
- with similar arrangements for the board
- has a formal statement of independence
- receives government grants
- and must present annual reports to parl. & exec.

Outline

Introduction

- Why do this research?

- What data did we gather?

Measurement

- How it's been done before

- How we want to measure

Main (measurement-related) findings

- The 'difficulty' of each item

- Index items drop out

Why this matters

- This changes our conclusions about independence

Example index item

Agency head dismissal:

- Dismissal impossible: 1
- Dismissal possible for reasons unrelated to policy: .66
- There are no specific provisions: .33
- Dismissal at appointer's discretion: 0

Example index item

Agency head dismissal:

- Dismissal impossible: 1
- Dismissal possible for reasons unrelated to policy: .66
- There are no specific provisions: .33
- Dismissal at appointer's discretion: 0

Example index item

Agency head dismissal:

- Dismissal impossible: 1
- Dismissal possible for reasons unrelated to policy: .66
- There are no specific provisions: .33
- Dismissal at appointer's discretion: 0

Example index item

Agency head dismissal:

- Dismissal impossible: 1
- Dismissal possible for reasons unrelated to policy: .66
- There are no specific provisions: .33
- Dismissal at appointer's discretion: 0

Problems with this way

- Can we be sure that these intervals are meaningful?
- Can we be sure the order is right?
- Are we dealing correctly with the absence of provisions?

Problems with this way

- Can we be sure that these intervals are meaningful?
- Can we be sure the order is right?
- Are we dealing correctly with the absence of provisions?

Problems with this way

- Can we be sure that these intervals are meaningful?
- Can we be sure the order is right?
- Are we dealing correctly with the absence of provisions?

By using an item response model...

- We can model independence as a latent trait
- We can find out which items map on to the trait
- We can find out the intervals between items (the differences between response category thresholds)

By using an item response model...

- We can model independence as a latent trait
- We can find out which items map on to the trait
- We can find out the intervals between items (the differences between response category thresholds)

By using an item response model...

- We can model independence as a latent trait
- We can find out which items map on to the trait
- We can find out the intervals between items (the differences between response category thresholds)

The model we use...

... is Samejima's Graded Response Model.

i items, with

g for the response category ($g=1\dots n$), for which

β represents the 'difficulty' of the item,

α the degree to which the item discriminates, and

θ is the 'ability' of the respondent

Probability of responding in a given answer category

$$P_{ig}^* = \frac{e^{\alpha_i(\theta - \beta_{ig})}}{1 + e^{\alpha_i(\theta - \beta_{ig})}} \quad (1)$$

The model gives us...

- Difficulty thresholds for each item
- Ability estimates for each respondent

In this case, the 'difficulty threshold for each item' is the amount of the latent trait necessary to answer in that category, and the 'ability parameter for each respondent' is really 'how independent that regulator is'.

The model gives us...

- Difficulty thresholds for each item
- Ability estimates for each respondent

In this case, the 'difficulty threshold for each item' is the amount of the latent trait necessary to answer in that category, and the 'ability parameter for each respondent' is really 'how independent that regulator is'.

The model gives us...

- Difficulty thresholds for each item
- Ability estimates for each respondent

In this case, the 'difficulty threshold for each item' is the amount of the latent trait necessary to answer in that category, and the 'ability parameter for each respondent' is really 'how independent that regulator is'.

Example of thresholds

- Let's take the example of term lengths.
- Here we have five categories (≤ 3 yrs, 4, 5, 6, and >6 yrs)
- The GRM will give us four thresholds, which separate these categories

Example of thresholds

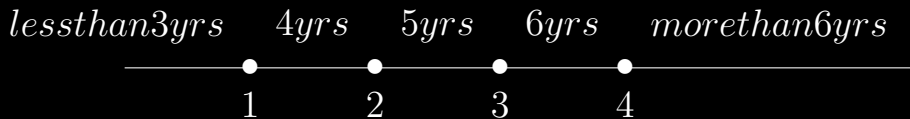
- Let's take the example of term lengths.
- Here we have five categories (≤ 3 yrs, 4, 5, 6, and >6 yrs)
- The GRM will give us four thresholds, which separate these categories

Example of thresholds

- Let's take the example of term lengths.
- Here we have five categories (≤ 3 yrs, 4, 5, 6, and >6 yrs)
- The GRM will give us four thresholds, which separate these categories

Four thresholds, five categories

Figure: Term length



Outline

Introduction

- Why do this research?

- What data did we gather?

Measurement

- How it's been done before

- How we want to measure

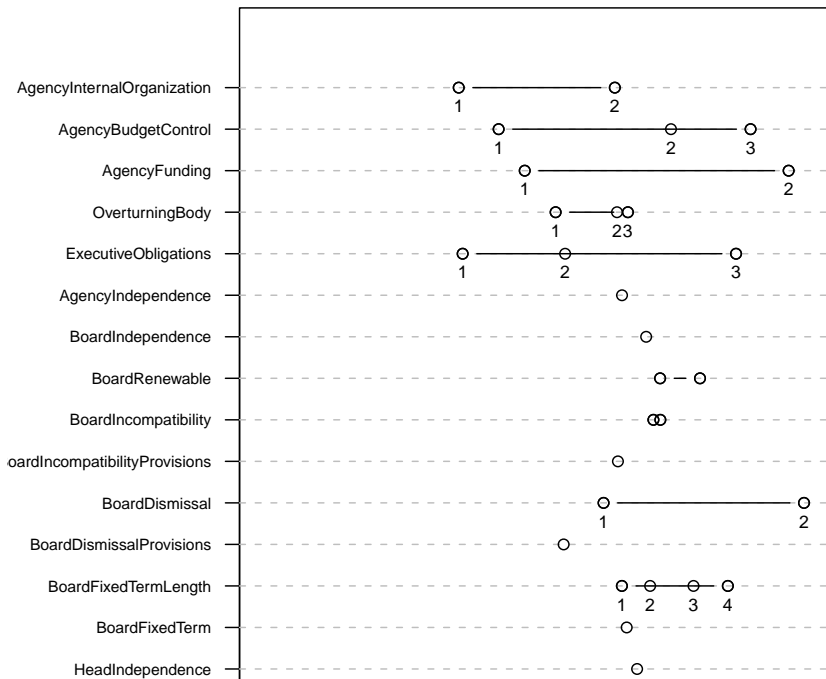
Main (measurement-related) findings

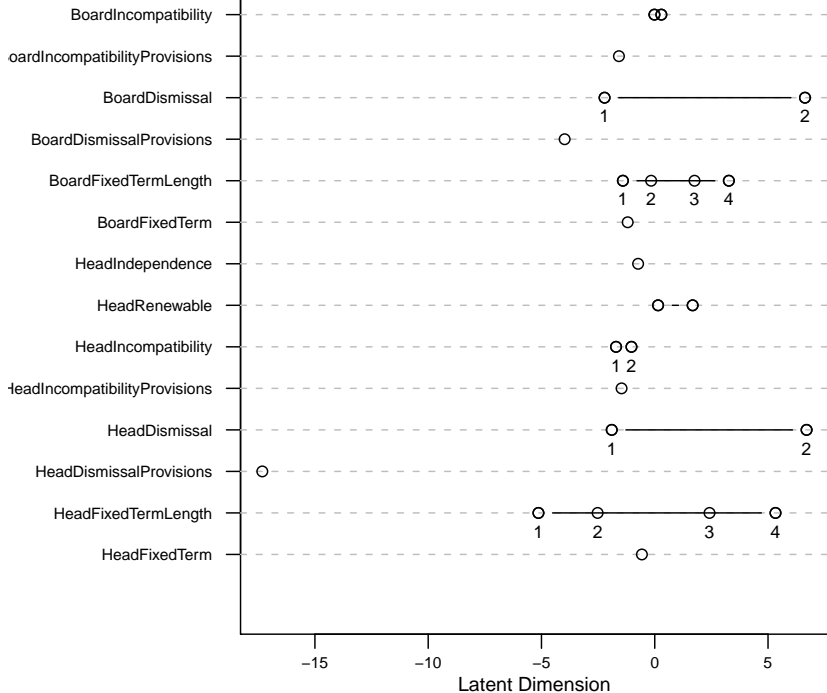
- The 'difficulty' of each item

- Index items drop out

Why this matters

- This changes our conclusions about independence





Some index items perform poorly

- We included all of Gilardi's original items
- But only some gave sensible parameters.
- Others, like parameters for personnel control and exclusive competence, were non-sensical

Some index items perform poorly

- We included all of Gilardi's original items
- But only some gave sensible parameters.
- Others, like parameters for personnel control and exclusive competence, were non-sensical

Some index items perform poorly

- We included all of Gilardi's original items
- But only some gave sensible parameters.
- Others, like parameters for personnel control and exclusive competence, were non-sensical

Some index items are the wrong way round

- *Worse*, some index items had negative discrimination parameters
- That is, the “higher-scoring” response categories were associated with lower independence
- These included the appointing actor, and parliamentary accountability

Some index items are the wrong way round

- *Worse*, some index items had negative discrimination parameters
- That is, the “higher-scoring” response categories were associated with lower independence
- These included the appointing actor, and parliamentary accountability

Some index items are the wrong way round

- *Worse*, some index items had negative discrimination parameters
- That is, the “higher-scoring” response categories were associated with lower independence
- These included the appointing actor, and parliamentary accountability

Some index items are the wrong way round (2)

- This is big news.
- The identity of the appointing actor is used in almost every index of independence...
- And if legislative accountability is unrelated, perhaps there's no antagonism between independence and accountability

Some index items are the wrong way round (2)

- This is big news.
- The identity of the appointing actor is used in almost every index of independence...
- And if legislative accountability is unrelated, perhaps there's no antagonism between independence and accountability

Some index items are the wrong way round (2)

- This is big news.
- The identity of the appointing actor is used in almost every index of independence...
- And if legislative accountability is unrelated, perhaps there's no antagonism between independence and accountability

Outline

Introduction

- Why do this research?

- What data did we gather?

Measurement

- How it's been done before

- How we want to measure

Main (measurement-related) findings

- The 'difficulty' of each item

- Index items drop out

Why this matters

- This changes our conclusions about independence

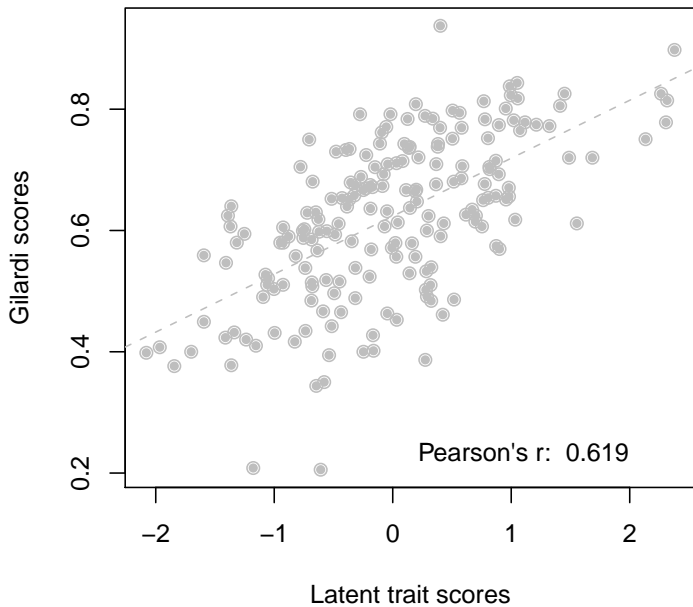
Our measurement differs from Gilardi's

- The two items are strongly correlated ($\rho=0.62$)
- But the divergence between the two measures affects our conclusions

Our measurement differs from Gilardi's

- The two items are strongly correlated ($\rho=0.62$)
- But the divergence between the two measures affects our conclusions

Comparison of the two scales



Our measure is better associated with de facto independence

- Maggetti has done work on the *de facto* independence of IRAs
- Gilardi's measure is not at all correlated with his de facto scores
- Our measure *is* (weakly, on a very small n of 16)

Our measure is better associated with de facto independence

- Maggetti has done work on the *de facto* independence of IRAs
- Gilardi's measure is not at all correlated with his de facto scores
- Our measure *is* (weakly, on a very small n of 16)

Our measure is better associated with de facto independence

- Maggetti has done work on the *de facto* independence of IRAs
- Gilardi's measure is not at all correlated with his de facto scores
- Our measure *is* (weakly, on a very small n of 16)

Our measure changes the determinants of de jure independence

- We took the measures from Gilardi's 2005 paper
- We calculated their 'new' independence scores using our method...
- and we re-ran the regression model.

Our measure changes the determinants of de jure independence

- We took the measures from Gilardi's 2005 paper
- We calculated their 'new' independence scores using our method...
- and we re-ran the regression model.

Our measure changes the determinants of de jure independence

- We took the measures from Gilardi's 2005 paper
- We calculated their 'new' independence scores using our method...
- and we re-ran the regression model.

Our measure changes the determinants of de jure independence (2)

- Gilardi originally claimed that replacement risk had a non-linear effect on degree of independence
- (Replacement risk is the risk of a government being replaced by an ideologically distant alternative government)

Our measure changes the determinants of de jure independence (2)

- Gilardi originally claimed that replacement risk had a non-linear effect on degree of independence
- (Replacement risk is the risk of a government being replaced by an ideologically distant alternative government)

Our measure changes the determinants of de jure independence (3)

- When we substitute our measure for his, the effect of replacement risk is linear.
- So even countries with really volatile governments create independent regulators.

Our measure changes the determinants of de jure independence (3)

- When we substitute our measure for his, the effect of replacement risk is linear.
- So even countries with really volatile governments create independent regulators.

Table: Comparison of models w/ different measurements of independence

	Original measurement	Revised measurement	Revised measurement, reduced model
Intercept	0.05 (0.06)	0.07 (0.06)	0.02 (0.05)
Utilities	0.32* (0.05)	0.33* (0.08)	0.32* (0.08)
Financial markets	0.22* (0.04)	0.21* (0.05)	0.21* (0.05)
Replacement risk	1.49* (0.43)	0.10 (0.48)	0.63* (0.16)
Replacement risk sq	-1.90* (0.80)	0.87 (0.80)	
Veto players	-0.06* (0.02)	-0.08* (0.03)	-0.09* (0.03)
<i>N</i>	106	106	106
<i>R</i> ²	0.41	0.41	0.40
adj. <i>R</i> ²	0.38	0.38	0.38
Resid. sd	0.19	0.22	0.22

Robust standard errors in parentheses

* indicates significance at $p < 0.05$