The Decade of Roma Inclusion: A Unifying Framework of Progress Measurement

Martin Kahanec
Institute for the Study of Labor, Bonn

March 31, 2009
Belgrade

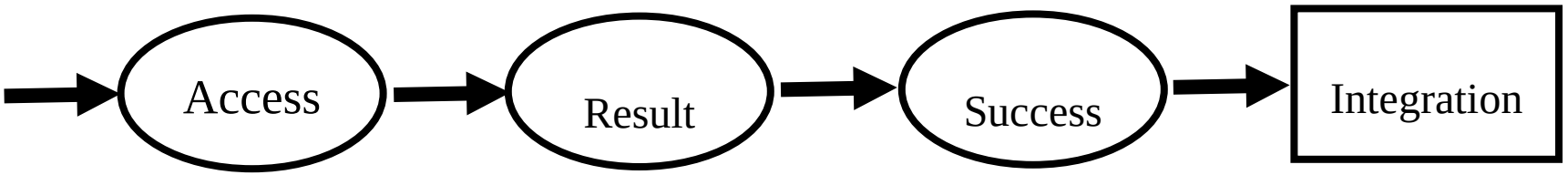
DECADE OF ROMA INCLUSION 2005-2015

Objectives

- Propose a mechanism to allow Decade countries to track and report on the results of Roma inclusion policies in 2015. Measure changes in the lives of people
- Propose a measurement methodology and a set of indicators covering education, employment, health and housing
- Propose data collection mechanisms
- Propose first and second best options

The First Best: What is integration?

- Integration: full participation in terms of social and economic life of the broader society, i.e. achieving outcomes comparable to the majority
- Three measured stages of the integration process
  - Opportunity to **access** a particular institution or service
  - Access provided, ability to realize a positive **result**
  - Realization provided, the chances to achieve **success**



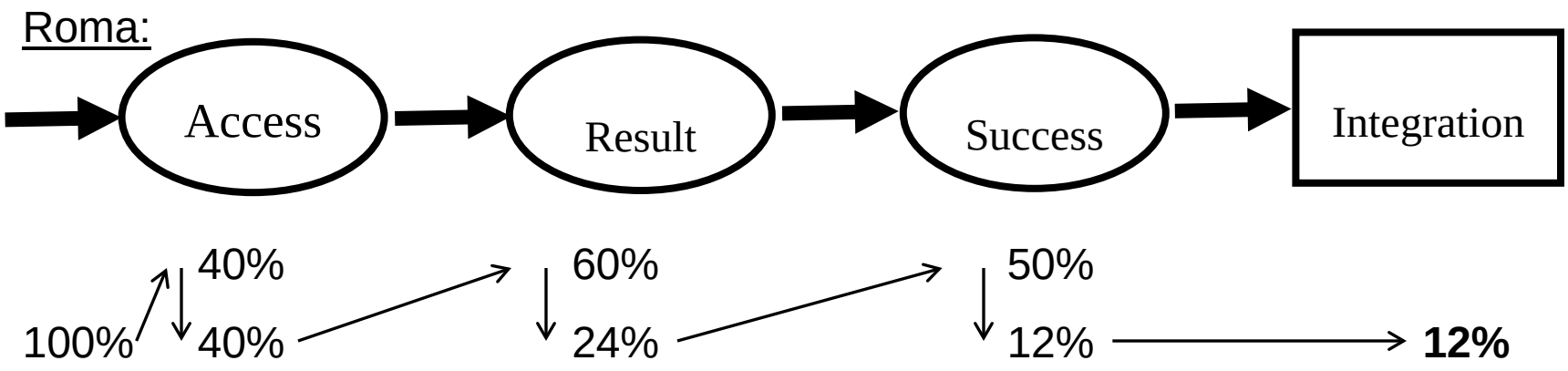
The First Best: Measuring integration

- No data problems assumed
- Access and result measured by respective chances (rates) of achieving a positive outcome
- Success measured at the group level:
  - expected outcome (e.g. population average hourly wage)
  - **chance to achieve some “decent outcome" (e.g. 3 EUR an hour)**
  - chance to achieve an outcome similar to the majority. (e.g. the median earnings of the majority)
- Ratio of minority and majority chances is our key value

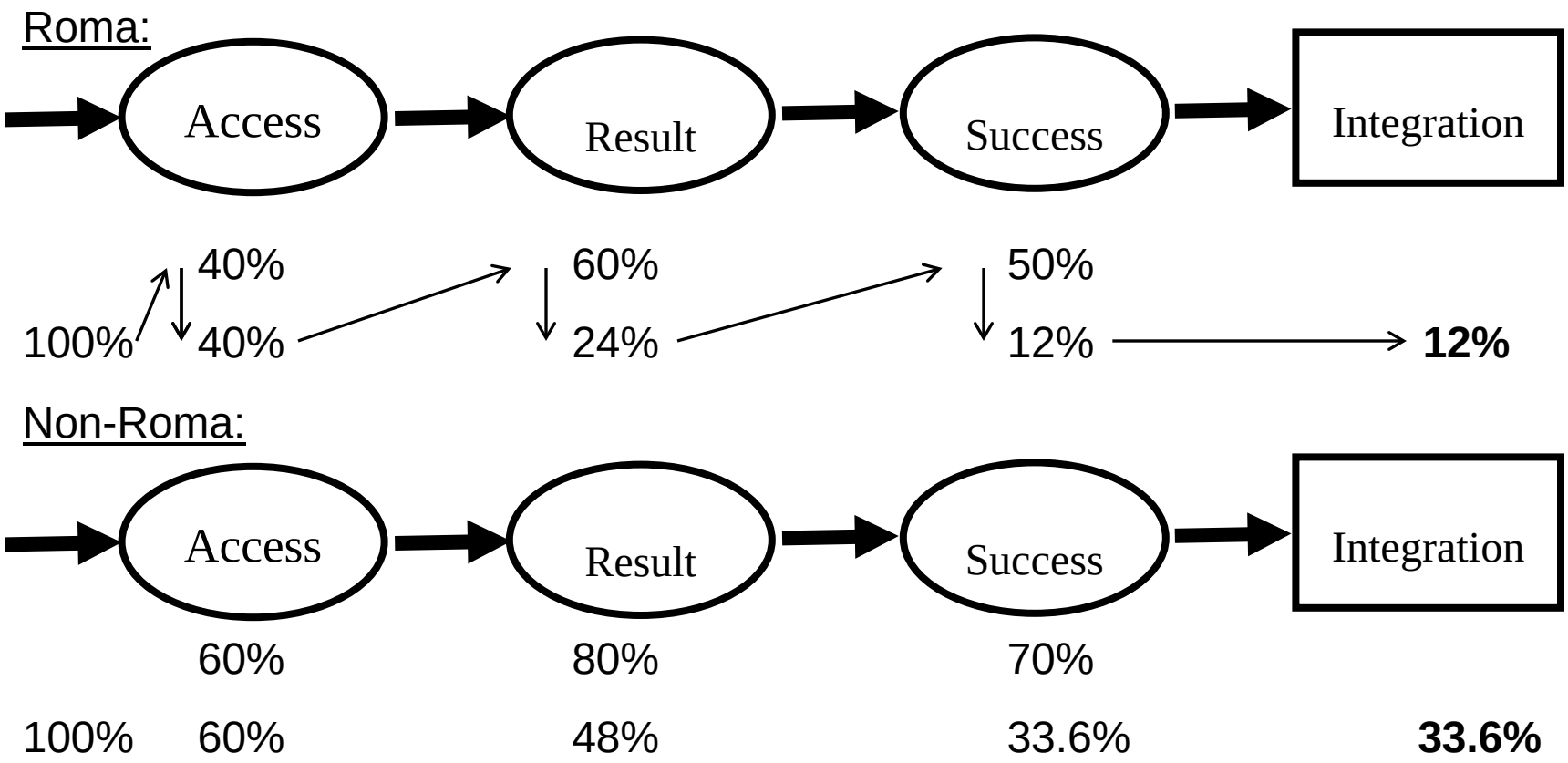
$$\sigma \equiv p_R / p_N$$

- This ratio can be calculated for every stage, and also every dimension (employment, housing, etc)

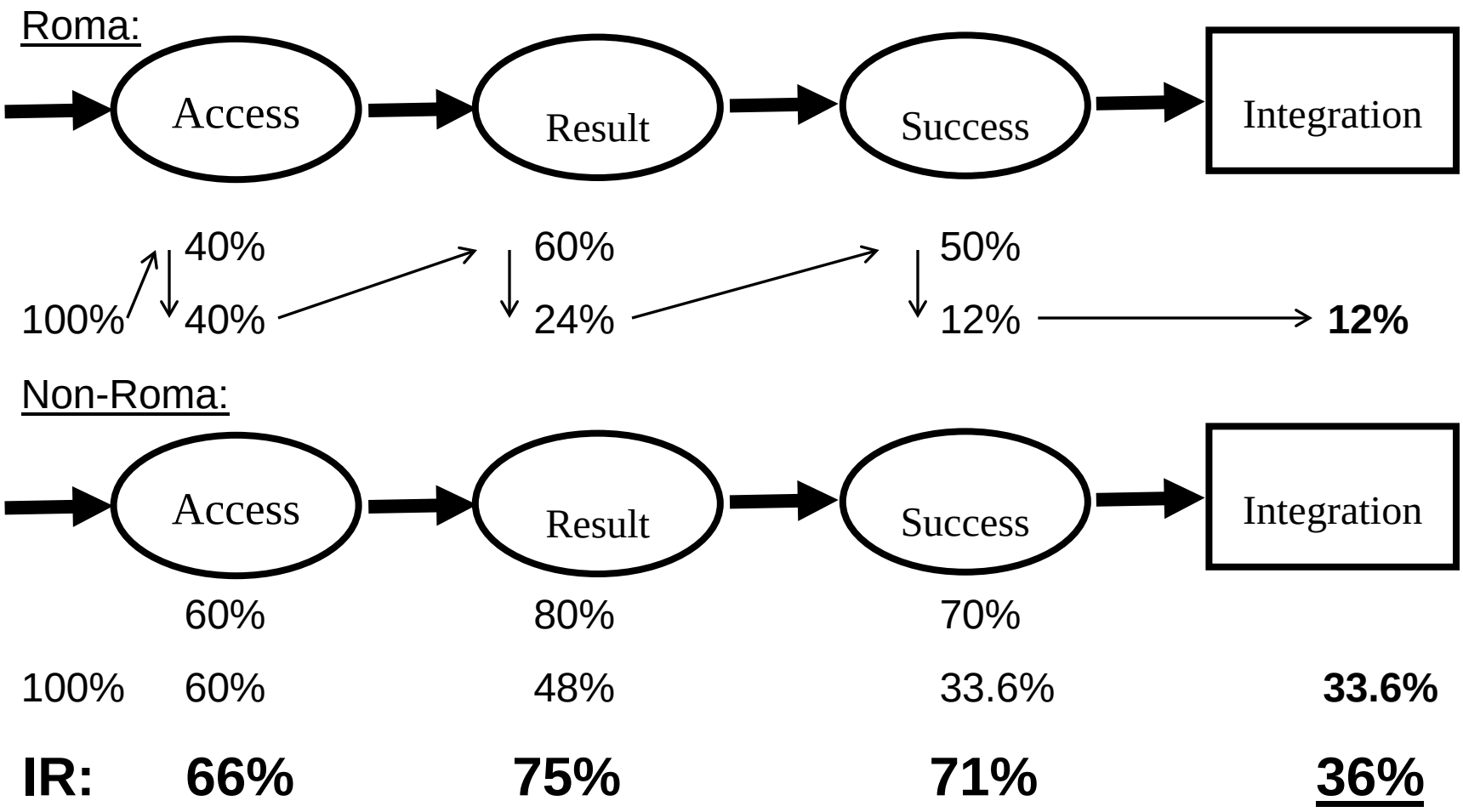
Example: Roma



Example: Roma and non-Roma



Example: Roma vs. non-Roma



The First Best: Issues

- Many possible indicators
  - Wage employment vs. self-employment
  - Hourly wage or occupational status
- Often ambiguity wrt appropriate target population
  - Age
  - Gender
  - Location
- Solution:
  - Provide core and secondary indicators
  - Provide indicators for the general population, and if possible report indicators for subpopulations (women, youth…)

A Unifying Framework: Integration indicators

ble 1: *Three-stage Integration Indicators (Core indicators bold)*

	Employment	Education	Health	Housing
Access	**Labor force participation rate**	**Enrolment rate in primary education**, Enrolment rate in pre-primary education	**Possession of health insurance (rate)**	**Legal housing in a segregated neighborhood (as opposed to illegal housing) (rate)**
Result	**1 - unemployment rate (including self-employment),** 1 - unemployment rate (excluding self-employment)	**Integration at classroom level in primary education (index)**, 1- Special school incidence	**Registration with a general practitioner (rate)**, Registration with a gynecologist (rate), Vaccination rate	**Legal housing in a non-segregated neighborhood (rate)**
Success	**Average hourly wage**, Occupational status (ISCO-88)	**Share with (upper) secondary or tertiary education (ISCED 3+)**, Share with tertiary education (ISCED 5+), Mean educational achievement in standardized screenings and tests, Mean length of stay in pre-preprimary education	**Life expectancy at birth**, Infant mortality rate	**Mean net floor area (in m^2) per inhabitant (in legal housing in a non-segregated neighborhood),** Mean number of rooms per inhabitant (in legal housing in a non-segregated neighborhood)

But Still Problems: The Data Issues

- General lack of data and severe measurement problems
  - No indicators of ethnicity or missing variables in the existing data
  - Where ethnicity indicated, extreme measurement error due to low self-identification.
  - Restrictions on data availability
  - Restrictive questionnaires: no room for complex (i.e. normal) ethnicities
  - Confusion: ethnicity, nationality, citizenship
  - Sometimes negative associations with Roma ethnicity

Recommendations

- The long run
  - Include ethnicity questions in the regularly collected data
  - Apply broad measures of ethnicity and ethno-cultural background in the questionnaires
  - Remove social and psychological barriers to self-identification (generally an din data collection)
  - Remove excessive restrictions on data availability
- The medium run
  - Small-scale collection of dedicated data
    - dedicated mini-surveys,
    - Roma boosters or ethnicity supplements in existing surveys
    - community surveys providing aggregated data for well defined Roma communities
    - custom surveys collecting data form social service recipients on voluntary basis
  - Problems
    - costs (time and money), representativeness, and subjectivity

A Unifying Framework: Feasible *Short Run* Solutions?

- Can we apply the first best methodology using imperfect data?
- Use existing markers of ethnicity other than self identification?
  - Language or mother tongue? No.
  - Religion? No.

A Unifying Framework: A Feasible Second Best

- But perhaps we should look at what we have!
  - A: Detailed high-quality datasets without (reliable) ethnicity variables (LFS)
  - B: Various datasets targeting the Roma population such as neighborhood-level mappings

  - What is necessary is that the primary dataset (A) contains a variable that is correlated with ethnicity
  - The auxiliary dataset (B) provides information about the link between this variable and ethnicity

- But we often do have such a possibility: Geographical segregation!
  - Location
  - Neighborhood level segregation and info on the share of the Roma

  - The idea is very general, but an extreme case to illustrate the idea: If we have a dataset with the information about the neighborhood of the respondent, and we know which neighborhoods are “Roma” and which “non-Roma”, we know who is Roma and who not.

A Unifying Framework: A Feasible Second Best - Steps

- Step 1 (Partition)
  - Distinguish "segregated" and "integrated" neighborhoods by the share of the Roma
- Step 2 (Measurement)
  - Measure the outcome variable in segregated and integrated neighborhoods
  - Estimate the total numbers of Roma and non-Roma
  - Estimate the shares of Roma and non-Roma in the two types of neighborhoods
  - Estimate relative integration of Roma and non-Roma within segregated and integrated neighborhoods
- Step 3 (Calculation)
  - A well defined formula equal to first best if perfect measurement

Second Best: Evaluation

- Permits combining information from a detailed dataset (census, LFS…) with high quality data on outcome variables plus location AND inputs from other statistics/datasets (neighborhood mapping, mini-survey) that are much less demanding
- Equal to first best in the limit
- Proper incentives (for policy makers)
- Does not eliminate the measurement problem, but offers a flexible framework to address it
- Offers a workable easy-to-implement alternative with acceptable properties $\sigma^s = 1 = \sigma^i$

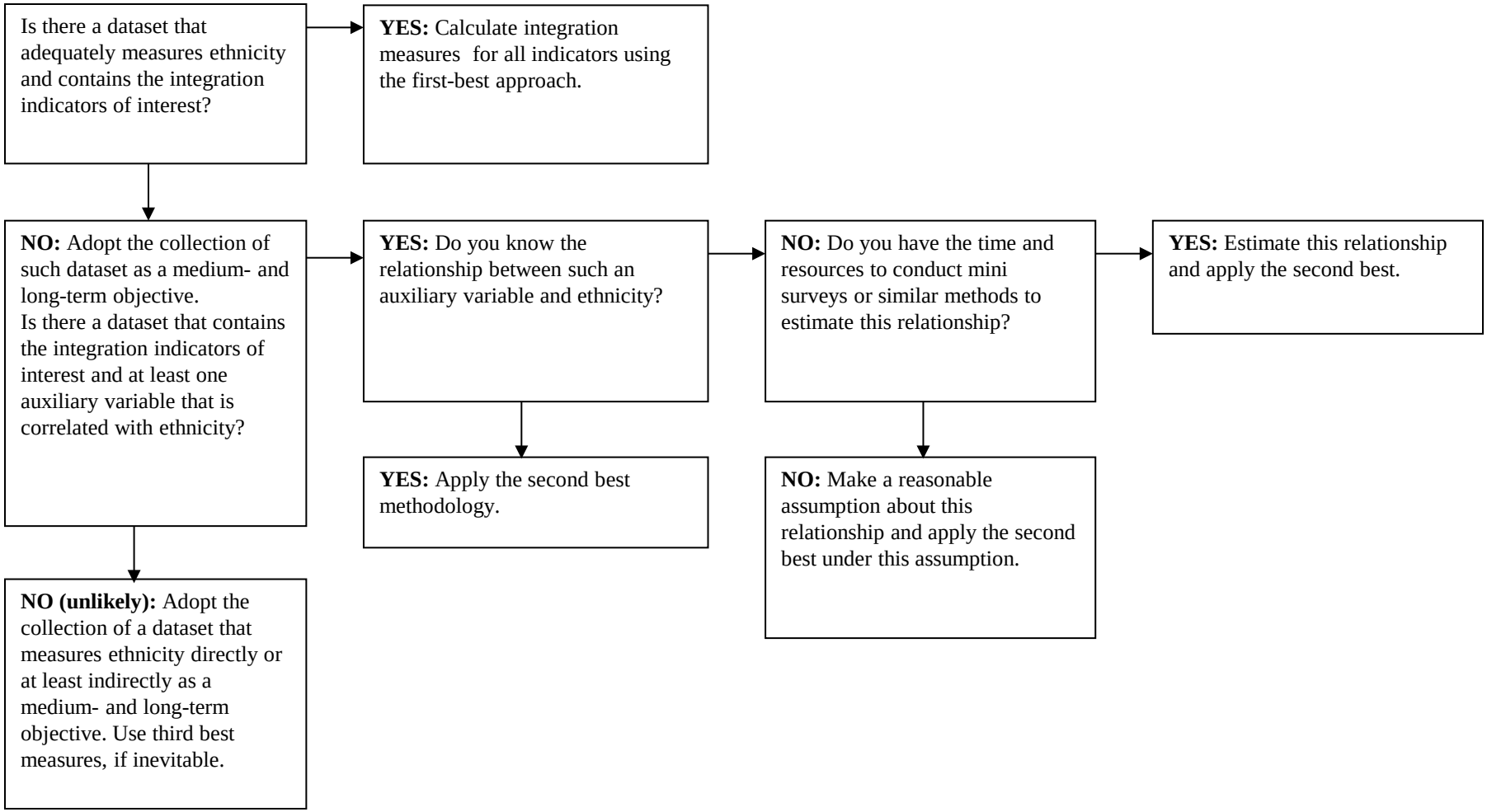
A Unifying Framework: Some Third Bests

- Alternatives based on the assumption that bad outcomes are correlated with ethnicity
- Shares of the total population
  - In poverty (e.g. below 1-2-3$/day)
  - In long term unemployment
  - Lacking education (or bad in PISA), health care, housing
- Advantages:
  - Readily available data
- Problems:
  - Dependent on the share of Roma
  - Dependent on the non-Roma's outcomes in additive way
  - Not really integration measures: not benchmarked
  - Unclear policy makers' incentives
- NOT RECOMMENDED

Table 2: *Data options*

Methodological approach	Data requirements	Data options
First-best	Contain integration indicators of interest and it is possible to distinguish Roma and non-Roma	Living Standard Measurement Surveys and Multi-Topic Household Surveys of the World Bank; the UNDP data covering vulnerable groups in Central South-Eastern Europe; and the Multiple Indicator Cluster Survey collected by UNICEF in Serbia
Second-best	Core data: Contain integration indicators of interest and an auxiliary variable that is correlated with ethnicity	National censuses; micro-censuses; labor force surveys; administrative data from employment offices, labor agencies, or the records of educational, health, and other register offices; Eurostat data such as the European Community Household Panel (ECHP); the EU Survey of Income and Living Conditions (SILC); and the European Social Survey (ESS); PISA data
	External data: Facilitate identification of the relationship between ethnicity and the auxiliary variable from the core data	Sociographic Mapping of Roma Communities in Slovakia; Living Standard Measurement Surveys and Multi-Topic Household Surveys of the World Bank; the UNDP data covering vulnerable groups in Central South-Eastern Europe; and the Multiple Indicator Cluster Survey collected by UNICEF in Serbia
Third best	Any dataset that contains integration indicators of interest	Any of the above

A Policy Chart



Conclusions

- We are facing a serious measurement challenge.
- There are solutions.
  - Long term: Improve standard data
  - Medium term: Collect own data
  - Short term: A feasible and valid second best solution that reduces the measurement problem, albeit it does not quite eliminate it
  - Immediate possibilities: Third best alternatives seriously flawed, second best with appropriate assumptions on σ^I and σ^S preferable.

Martin Kahanec
Institute for the Study of Labor (IZA)

IZA, P.O. Box 7240,
53072 Bonn, Germany
Phone: +49 (0) 228 - 38 94 - 529
Fax: +49 (0) 228 - 38 94 180
E-mail: kahanec@iza.org

http://www.iza.org