The Consistency of Cross-sectional and Longitudinal Data in EU-SILC Countries

when Measuring Income Levels, Inequality, and Mobility

Joachim R. Frick & *Kristina Krell*

<jfrick@diw.de> <kristina.krell@googlemail.com>

(SOEP, DIW Berlin)

2nd EU-Microdata User-conference, Mannheim, March 31 - April 1, 2011

Motivation

Explore problematic features of EU-SILC income data and find suggestions for further quality research

- Comparison of data for Germany with alternative German household survey (SOEP) shows the extensive influence of survey methods on income measures (Frick & Krell 2010)
- Unexpected developments of income levels, income mobility, and inequality in some countries
- Inconsistencies in cross-sectional and longitudinal data

SOEP The German Socio-Economic Panel Study

Structure of this presentation

1. Basics of EU-SILC (and SOEP)
2. Definitions
3. Comparative Analyses
   - Income levels, Inequality, Poverty
   - Cross-sectional vs. longitudinal datasets
   - Income & Poverty Mobility
4. Conclusion & recommendations

1. Basics of EU-SILC (and SOEP)

http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/eu_silc

- Types of interview: CAPI, PAPI, CATI, self-administred, proxy-interview
- Cross-sectional and 4-wave longitudinal sample in most countries
- Rotational Panel design: 25% of households replaced each year
- German SOEP: started 1984 including ~ 11.000 households with 25.000 persons; pure panel, personal interviews, no proxy-interviews (http://www.diw.de/en/soep)
- Country groups:
  - Former Transition countries: Poland, Czech Republic, Slovakia, Slovenia, Hungary, Estonia, Latvia, Lithuania
  - Mediterranean countries: Spain, Portugal, Italy, Greece, Cyprus
  - Continental and liberal countries: Austria, France, Netherlands, Belgium, Luxembourg, Germany, United Kingdom, Ireland
  - Scandinavian countries: Sweden, Norway, Finland, Denmark, Iceland

2. Definitions

Population:

- Individuals in private households

Observation years and periods:

- cross-sectional: 2005-2008, longitudinal: 2005/2006, 2006/2007, 2007/2008 (income reference period usually observation year-1)

Income:

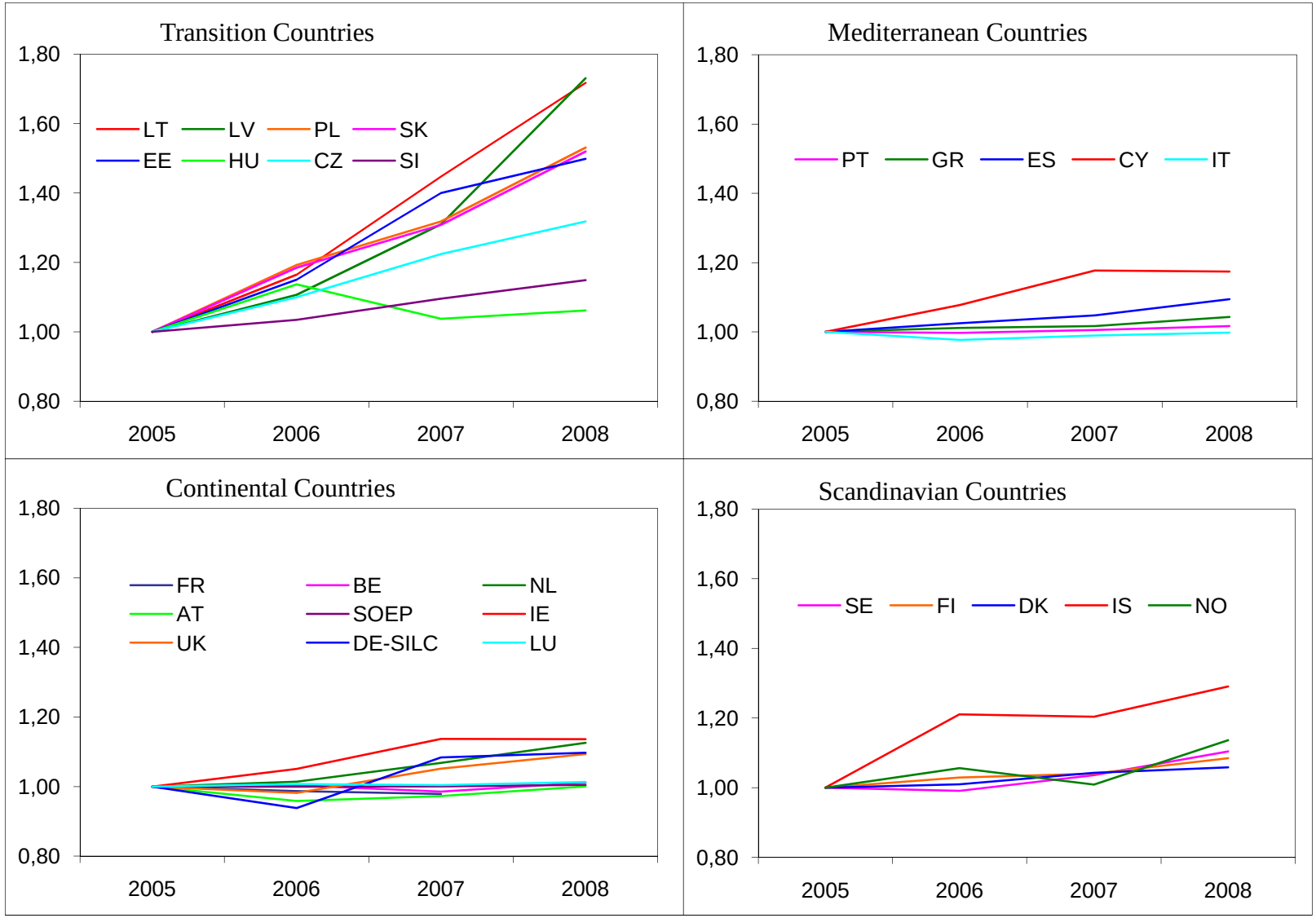
Equivalent disposable household income

- Sum of gross income components of all household members (income from employment and capital, private and public transfers, private and public pension) minus direct income taxes and social security contributions (without Imputed Rent) using the modified OECD-scale (1; 0.5; 0.3)
- Income in prices of 2005
- Imputation of missing income components

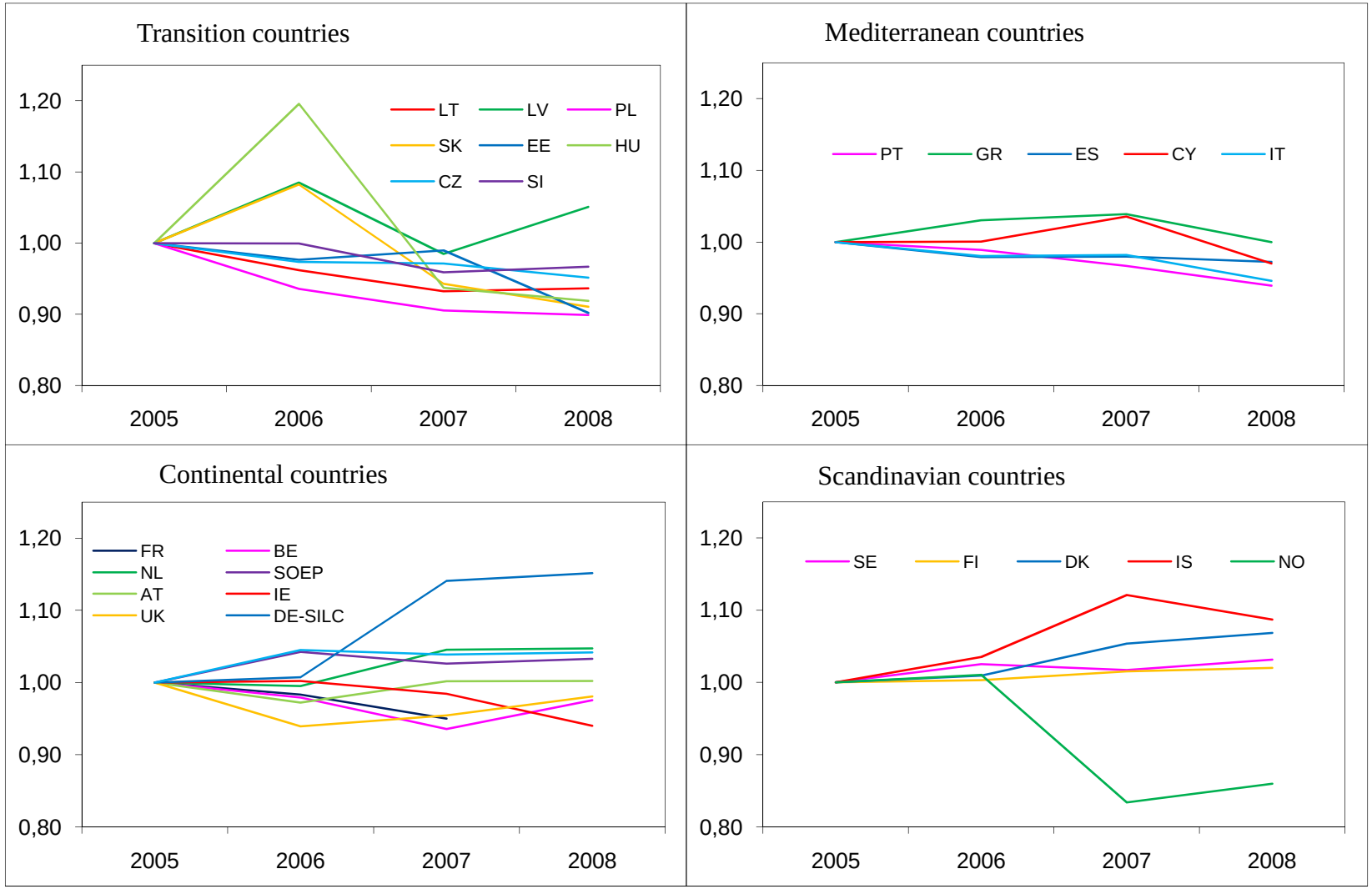
3. Comparative Analyses

- Cross-sectional perspective: Income Levels, Inequality, Poverty
- Cross-sectional vs. longitudinal datasets
- Longitudinal perspective: Income & Poverty Mobility

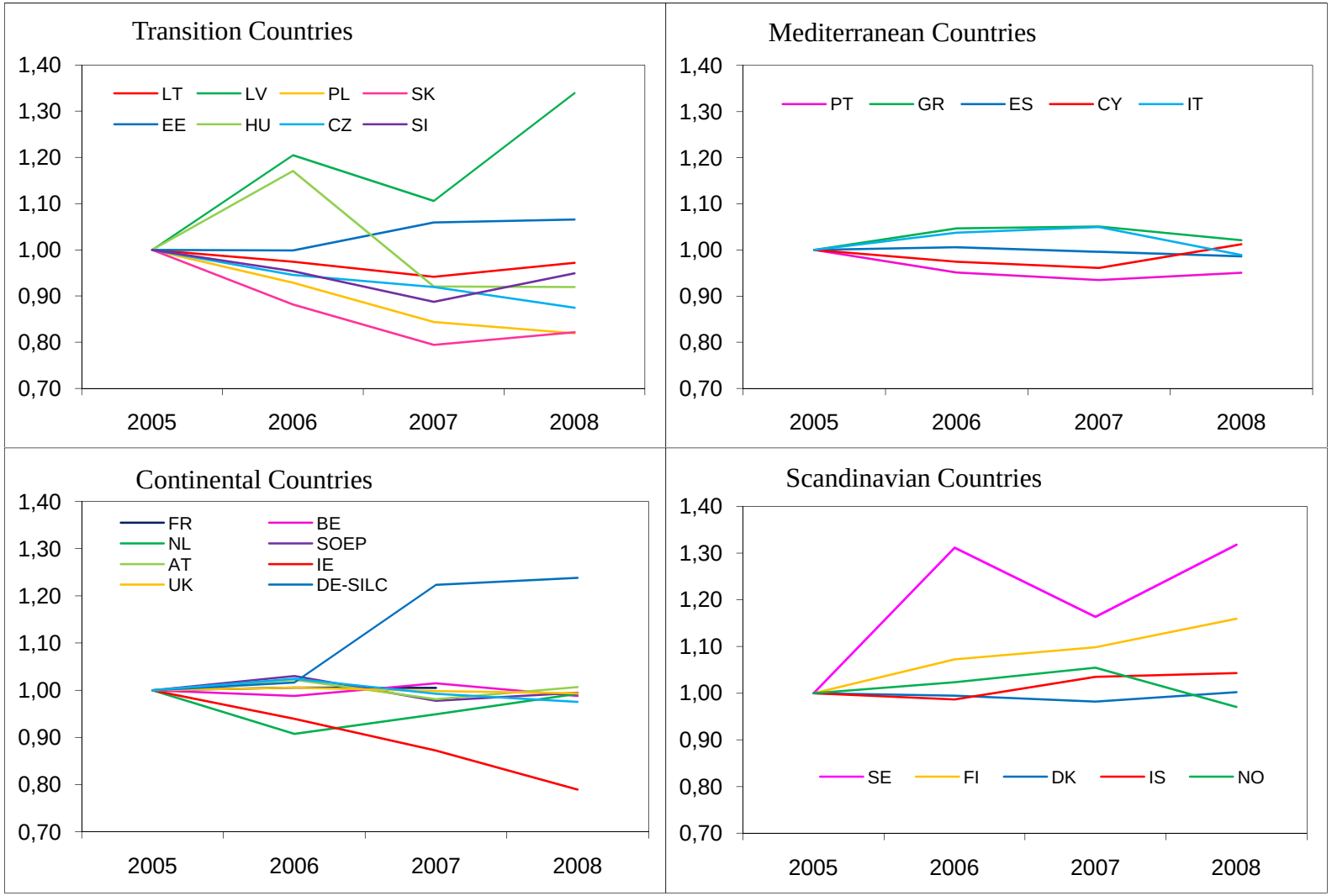
Income levels



Income inequality – Gini coefficient



Relative Income Poverty rate*

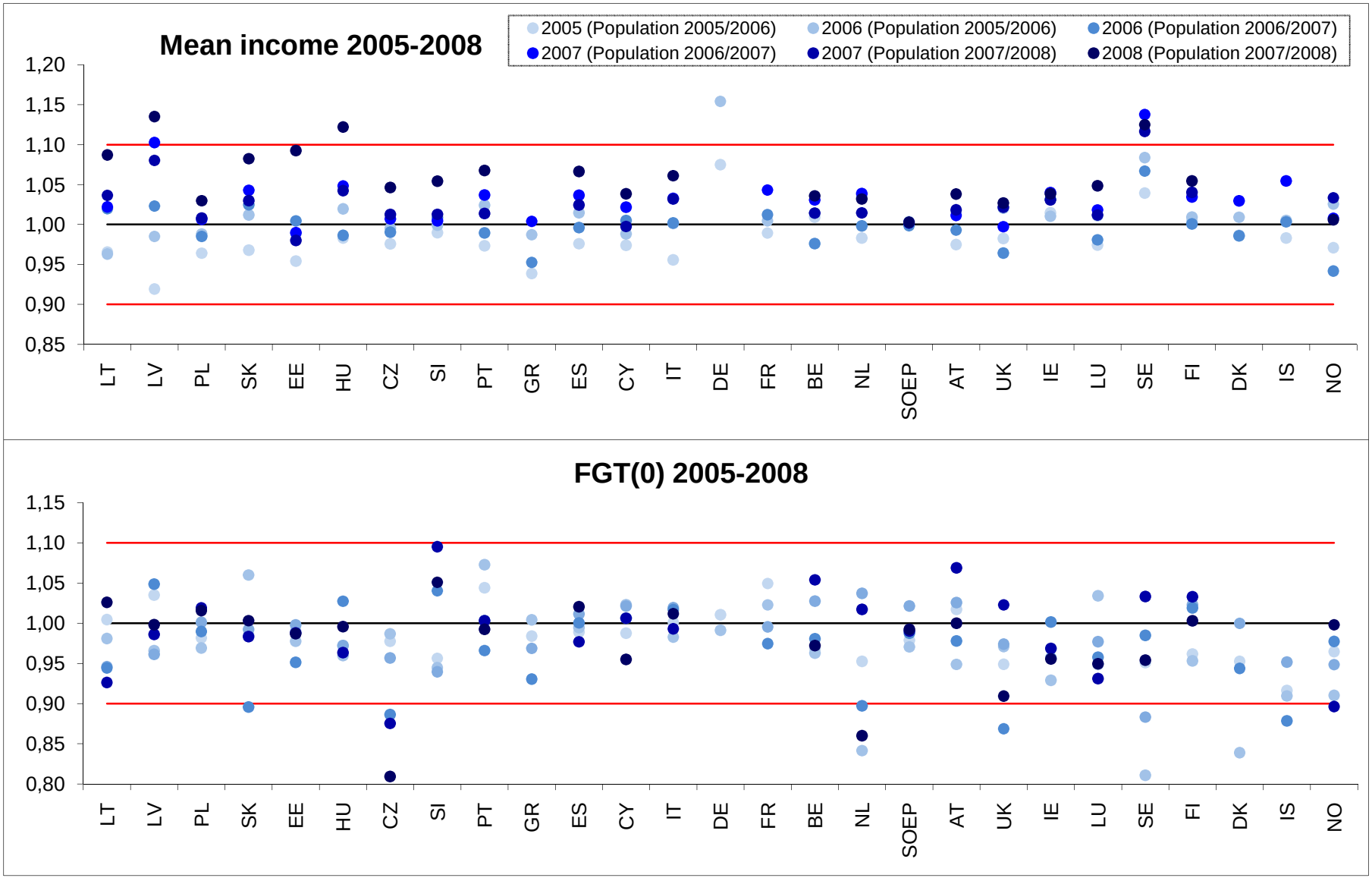


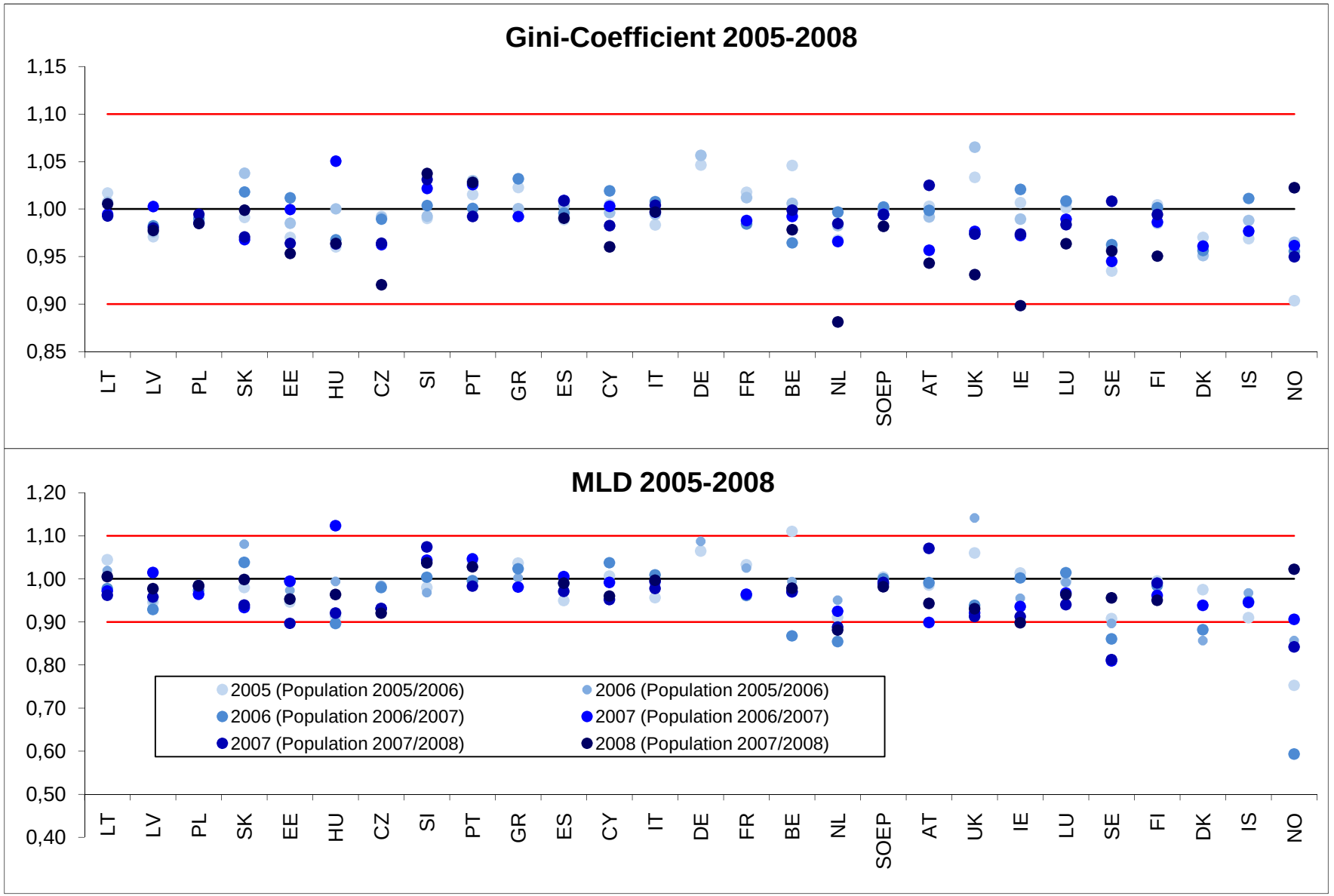
* Poverty threshold given at 60% of Median National Equivalent Income.

3. Comparative Analyses

- Cross-sectional perspective: Income Levels, Inequality, Poverty
- Cross-sectional vs. longitudinal datasets
- Longitudinal perspective: Income & Poverty Mobility

Indicator: fraction LT/CS




Gini-Coefficient 2005-2008
1,15
1,10
1,05
1,00
0,95
0,90
0,85
LT LV PL SK EE HU CZ SI PT GR ES CY IT DE FR BE NL SOEP AT UK IE LU SE FI DK IS NO
MLD 2005-2008
1,20
1,10
1,00
0,90
0,80
0,70
0,60
0,50
0,40
2005 (Population 2005/2006)
2006 (Population 2005/2006)
2006 (Population 2006/2007)
2007 (Population 2006/2007)
2007 (Population 2007/2008)
2008 (Population 2007/2008)
LT LV PL SK EE HU CZ SI PT GR ES CY IT DE FR BE NL SOEP AT UK IE LU SE FI DK IS NO

Why do cross-sectional and longitudinal results differ?

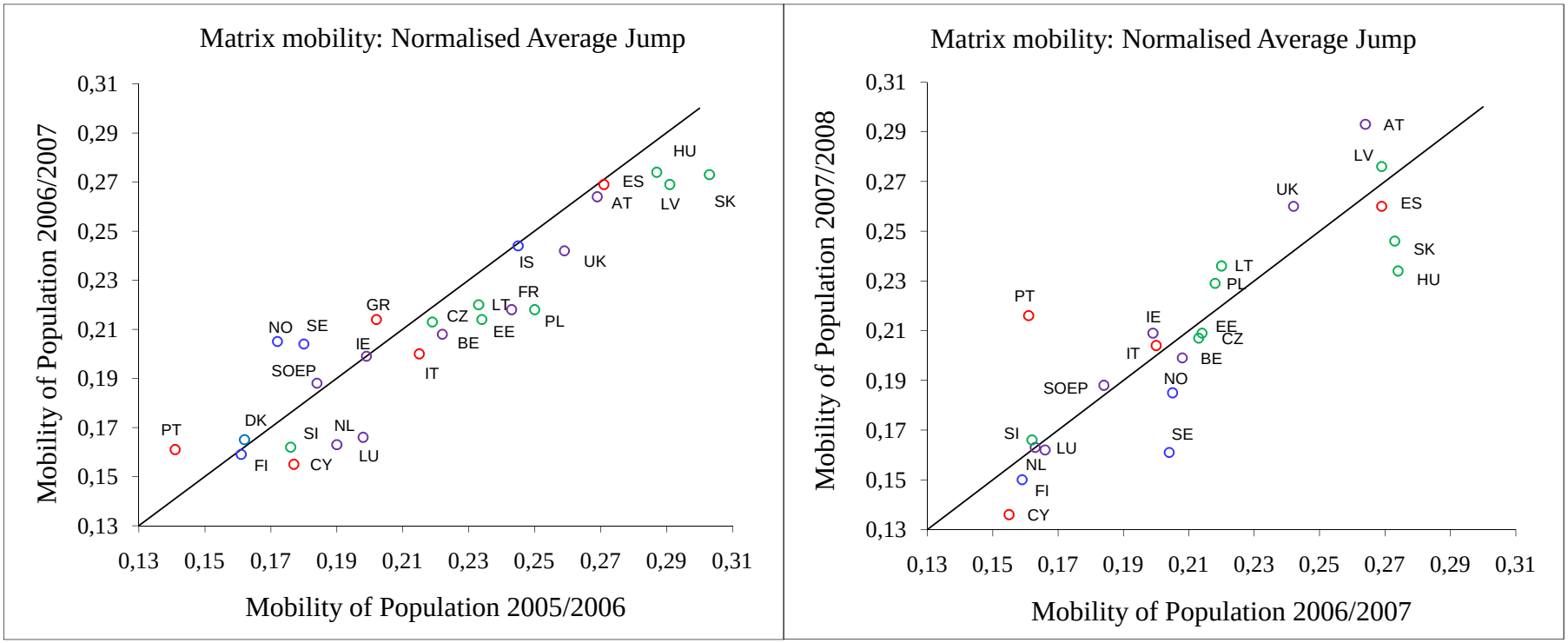
- Balanced Panel Design causes underrepresentation of poorer population subgroups (deceased, births, migrants)
- Attrition bias (extremes more likely to leave the panel)
- Non-inclusion of households with Partial Unit Non-Response (CZ, LU, IT, IE, UK, HU + register countries): maybe biased (Verma & Betti 2010)

→ No sufficient correction of attrition or achievement of representetiveness via weighting factors in some countries?

3. Comparative Analyses

- Cross-sectional perspective: Income Levels, Inequality, Poverty
- Cross-sectional vs. longitudinal datasets
- Longitudinal perspective: Income & Poverty Mobility

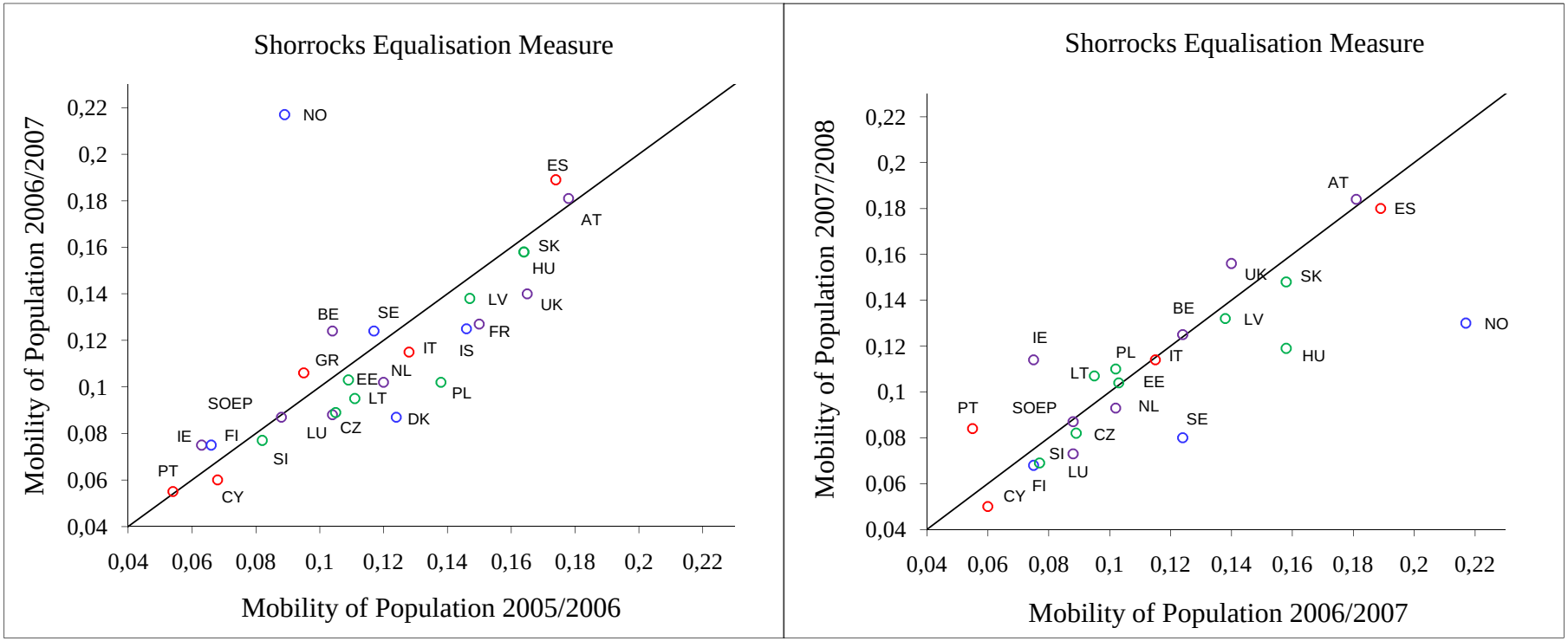
Income Mobility: Comparison of longitudinal populations 2005/2006, 2006/2007, and 2007/2008



Continental countries Transition countries

Scandinavian countries Mediterranean countries

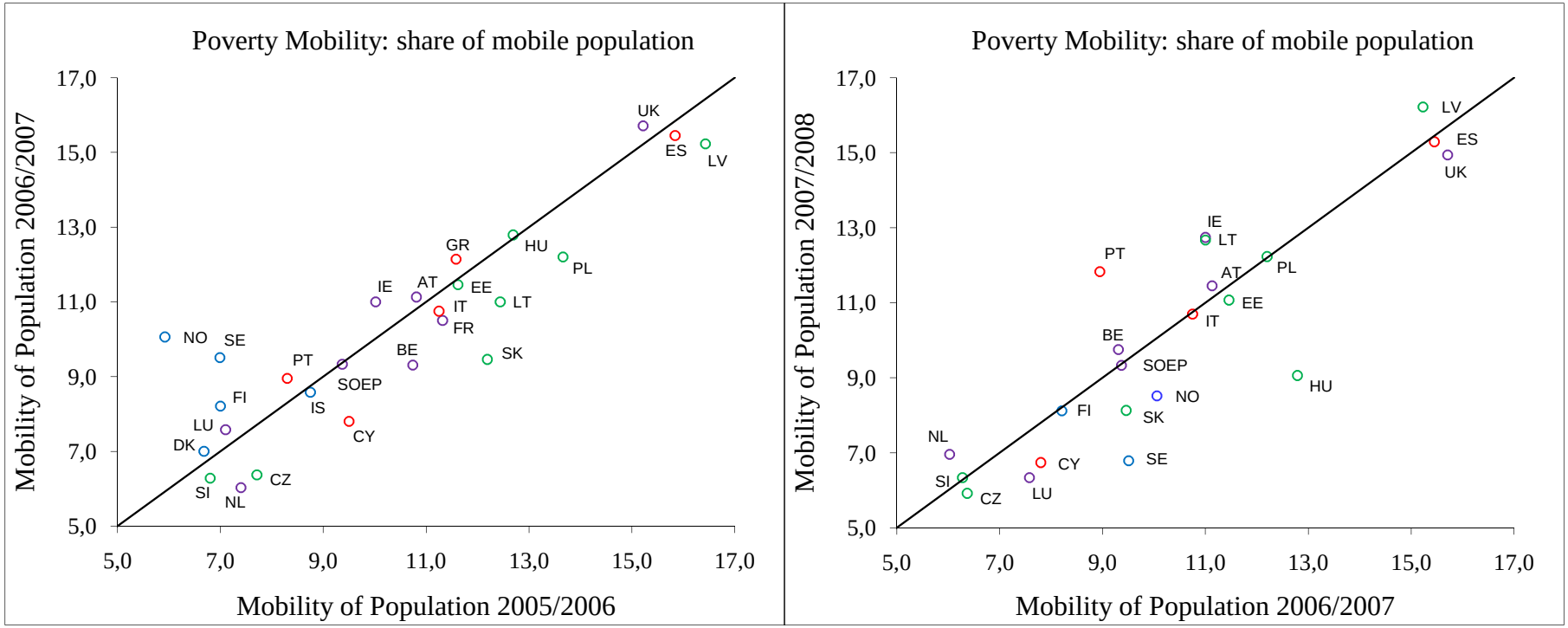
Income Mobility: Comparison of longitudinal populations 2005/2006, 2006/2007, and 2007/2008



Continental countries Transition countries

Scandinavian countries Mediterranean countries

Poverty Mobility: Comparison of longitudinal populations 2005/2006, 2006/2007, and 2007/2008



Continental countries Transition countries

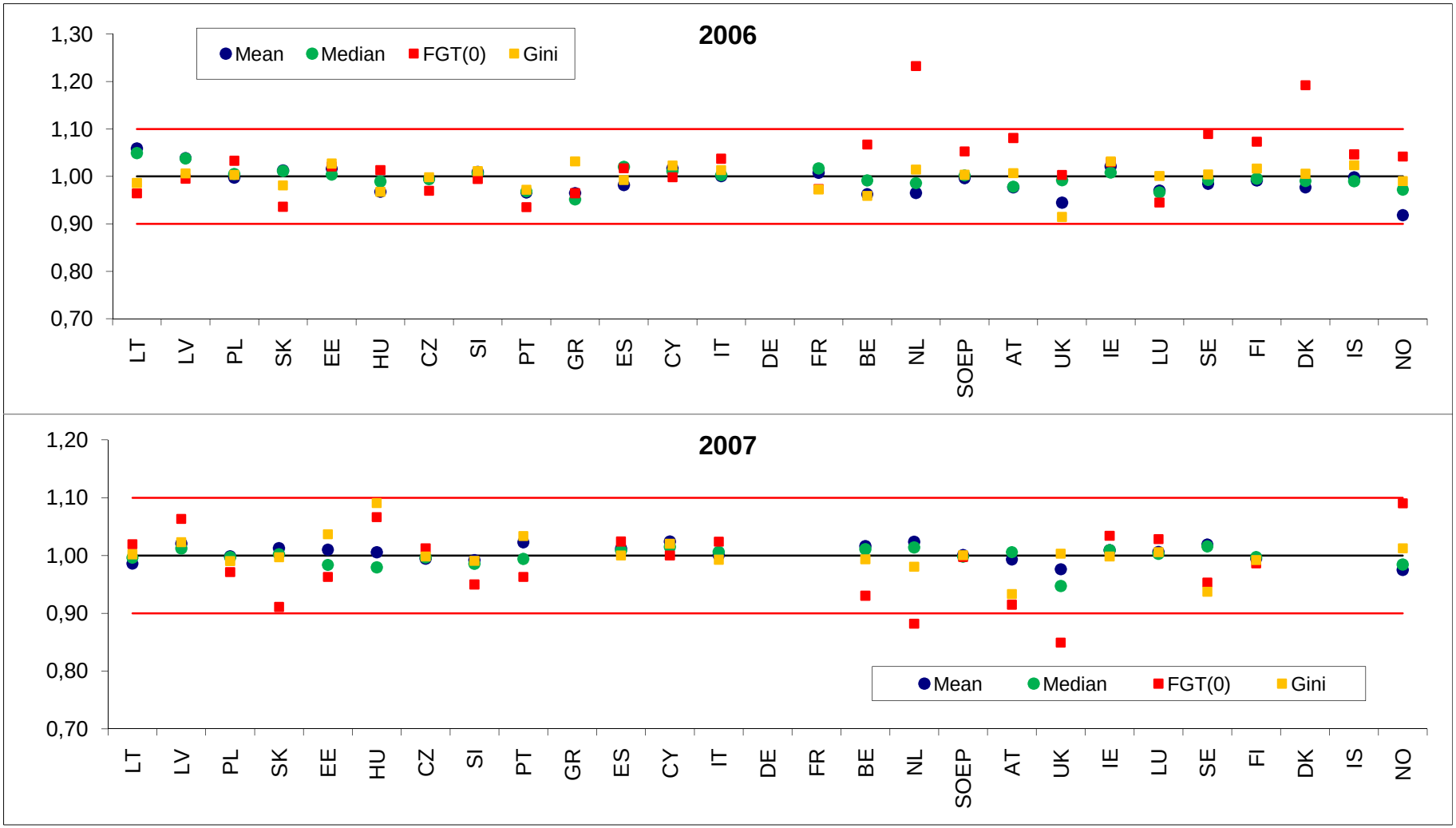
Scandinavian countries Mediterranean countries

Factors influencing mobility

- Type of interview: personal interview, self-administred by respondent (no interviewer assistance), telephone interview, proxy interview
- Imputation method: use of cross-sectional or cross-sectional and longitudinal information (Rässler & Riphahn 2006, Starick & Watson 2009, Spieß & Goebel 2005)
- Rotational panel design: income of recent surveys often more volatile due to the lack of experience in answering income questions (Frick et al. 2006)

The German Socio-Economic Panel Study

Comparison of income, poverty, and inequality in two longitudinal populations for one year



4. Conclusion & Recommendations

- Overall: Large variance in development of income levels, inequality, and mobility
- Countries with inconspicuous values in cross-sectional data can be conspicuous in the comparison of cross-sectional and longitudinal data
- All such surveys have to deal with difficulties in measuring economic outcomes (eg. Income) due to non-response, measurement error (self-administration, proxy interviews), …
- Imputation of missing data is based on assumptions and normative decisions and can have great effects on mobility

4. Conclusion & Recommendations

- Further check of data quality (esp. weighting?) in countries with problematic measures of income inequality and mobility in cross-sectional and longitudinal datasets – robustness check via alternative surveys
- Better, transnationally and intertemporarily comparable documentation of imputation methods and extent of imputation
- Use of longitudinal data for imputation in all countries
- Improvement of infrastructure for substantive and survey methodology research (provision of link between cross-sectional and longitudinal data)
- Continuous ex-post quality check (incl. eventual revisions for all waves to ensure consistency across time)

References

Frick, J. R., Goebel, J., Schechtmann, E., Wagner, G. G., Yitzhaki, S. (2006): Using Analysis of Gini (ANoGi) for detecting whether two sub-samples represent the same universe: The German So- cio-Economic Panel Study (SOEP) Experience, in: Sociological Methods & Research, 34 (4) pp. 427-468.

Frick, J.R. and Krell, K. (2010): Measuring Income in Household Panel Surveys for Germany: A Comparison of EU-SILC and SOEP. SOEPpaper 265, Berlin: DIW.

Lelkes, O., Medgyesi, M., Tóth, I. and Ward, T. (2009): Income distribution and the risk of poverty. In: Ward et al (2009), Chapter 1.

Rässler, S., Riphahn, R. T. (2006): Survey Item Non-Response and its Treatment, in: Allgemeines Statistisches Archiv, 90, pp. 213-228.

Starick, R., Watson, N. (2007): Evaluation of Alternative Income Imputation Methods for the HILDA Survey. HILDA Project Technical Paper Series No. 1/07, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Spieß, M., Goebel, J. (2005): On the Effect of Item Nonresponse on the Estimation of a Two-Panel-Waves Wage Equation, in: Allgemeines Statistisches Archiv, 1, pp. 63-74.

Verma, V. and Betti, G. (2010): Data accuracy in EU-SILC. In: Atkinson, A. and Marlier, E. (2010): Income and living conditions in Europe. Eurostat Statistical Book.

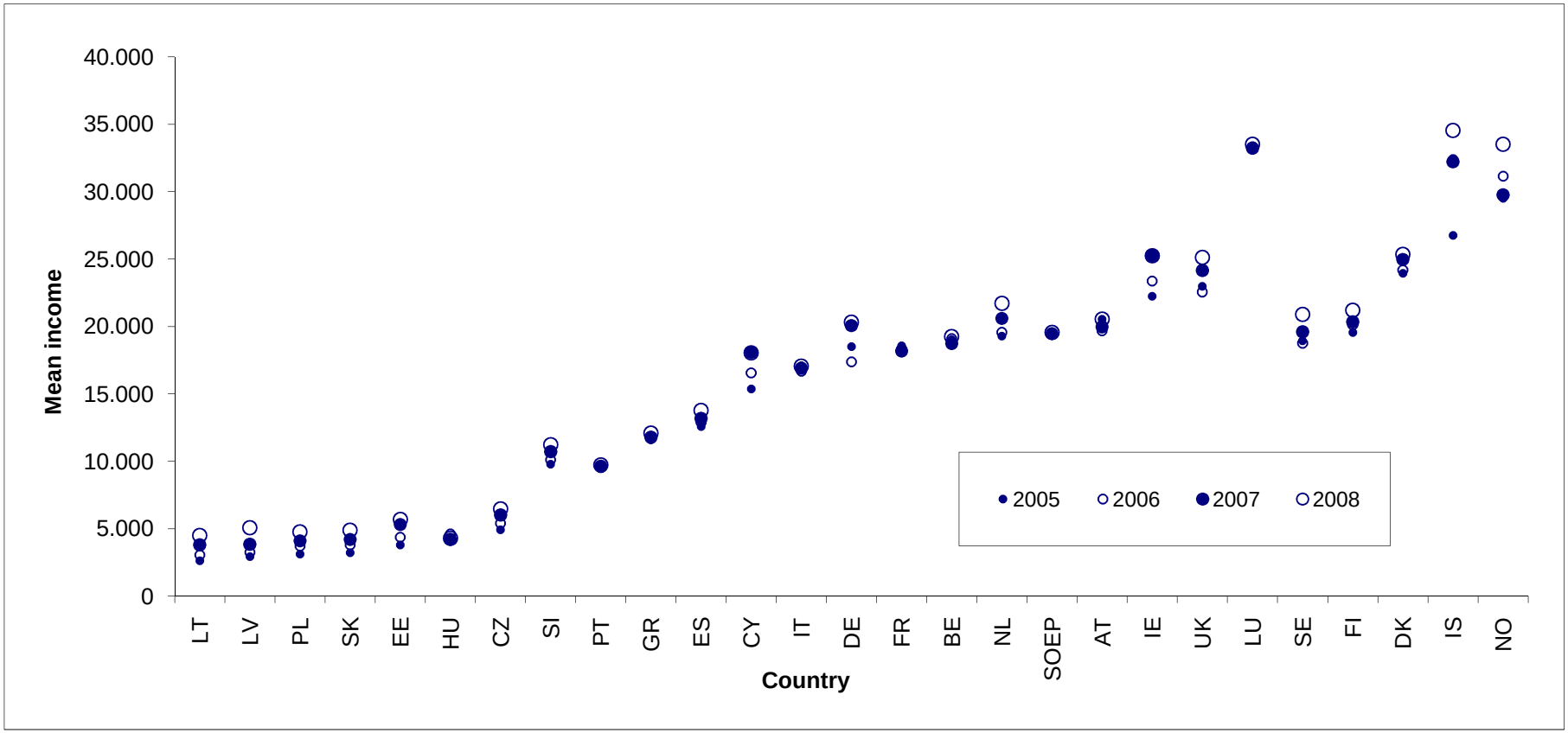
Thank you for your attention.

Kristina Krell

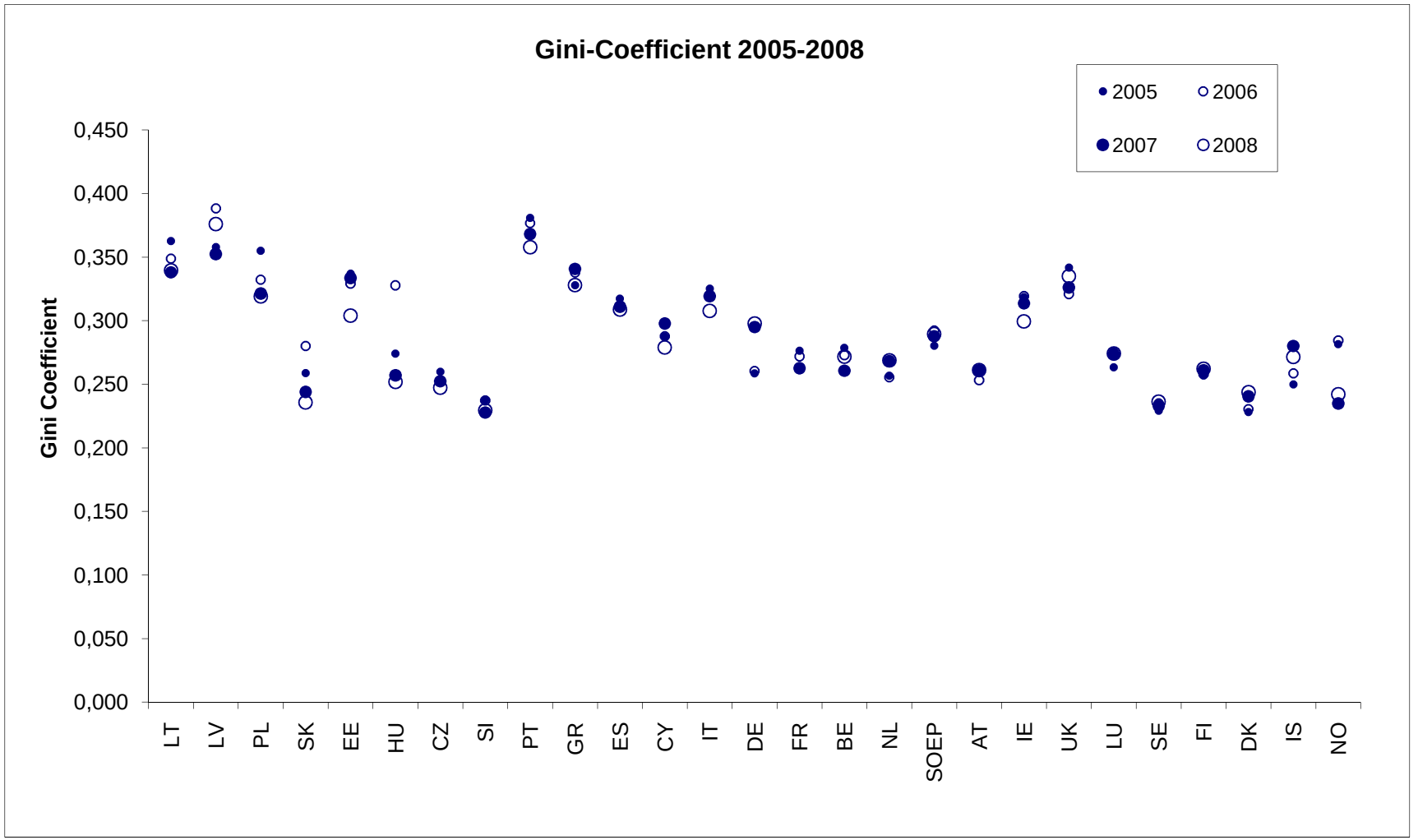
kristina.krell@googlemail.com

SOEP-homepage www.diw.de/gsoep

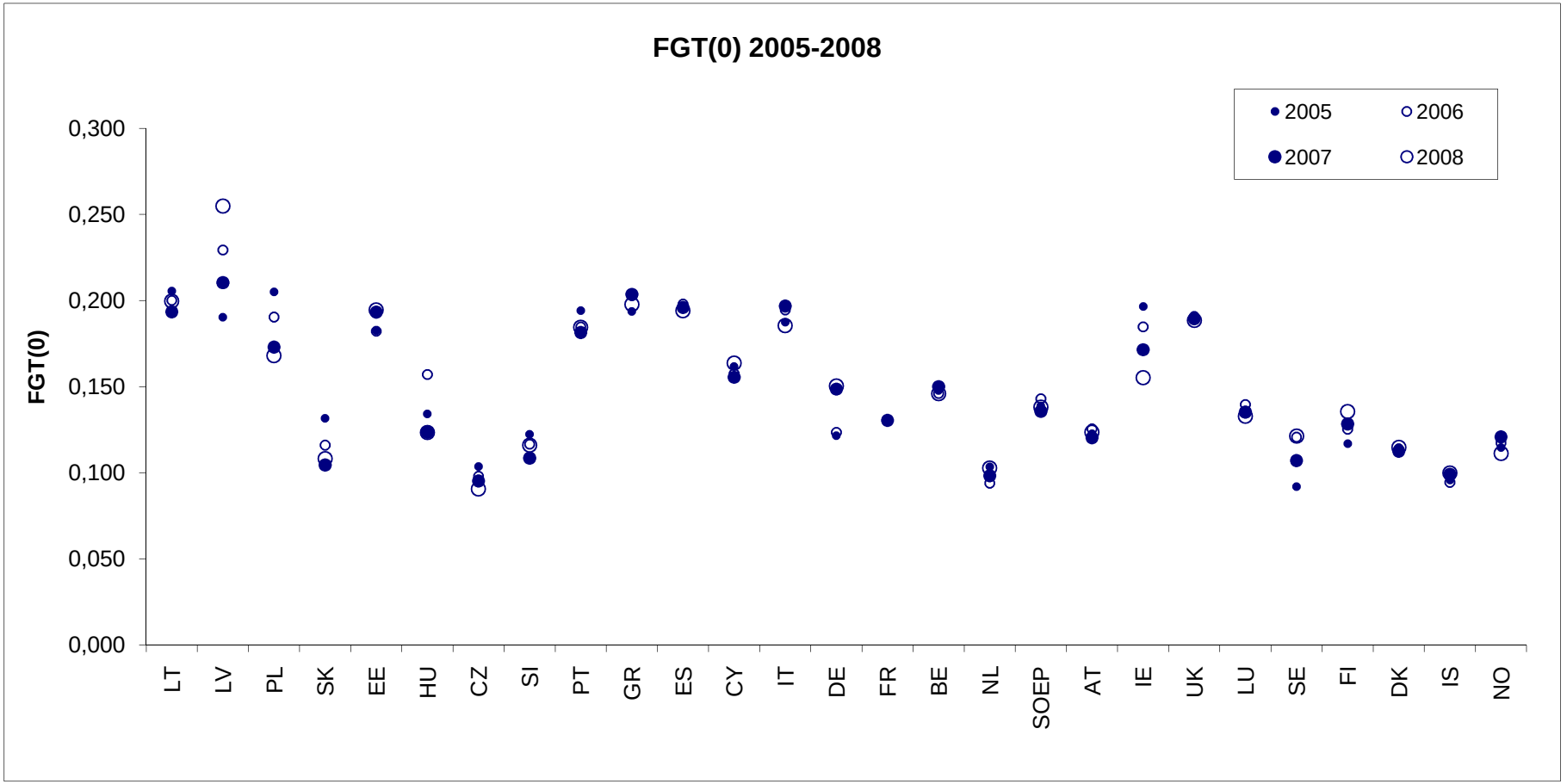
Mean income 2005-2008



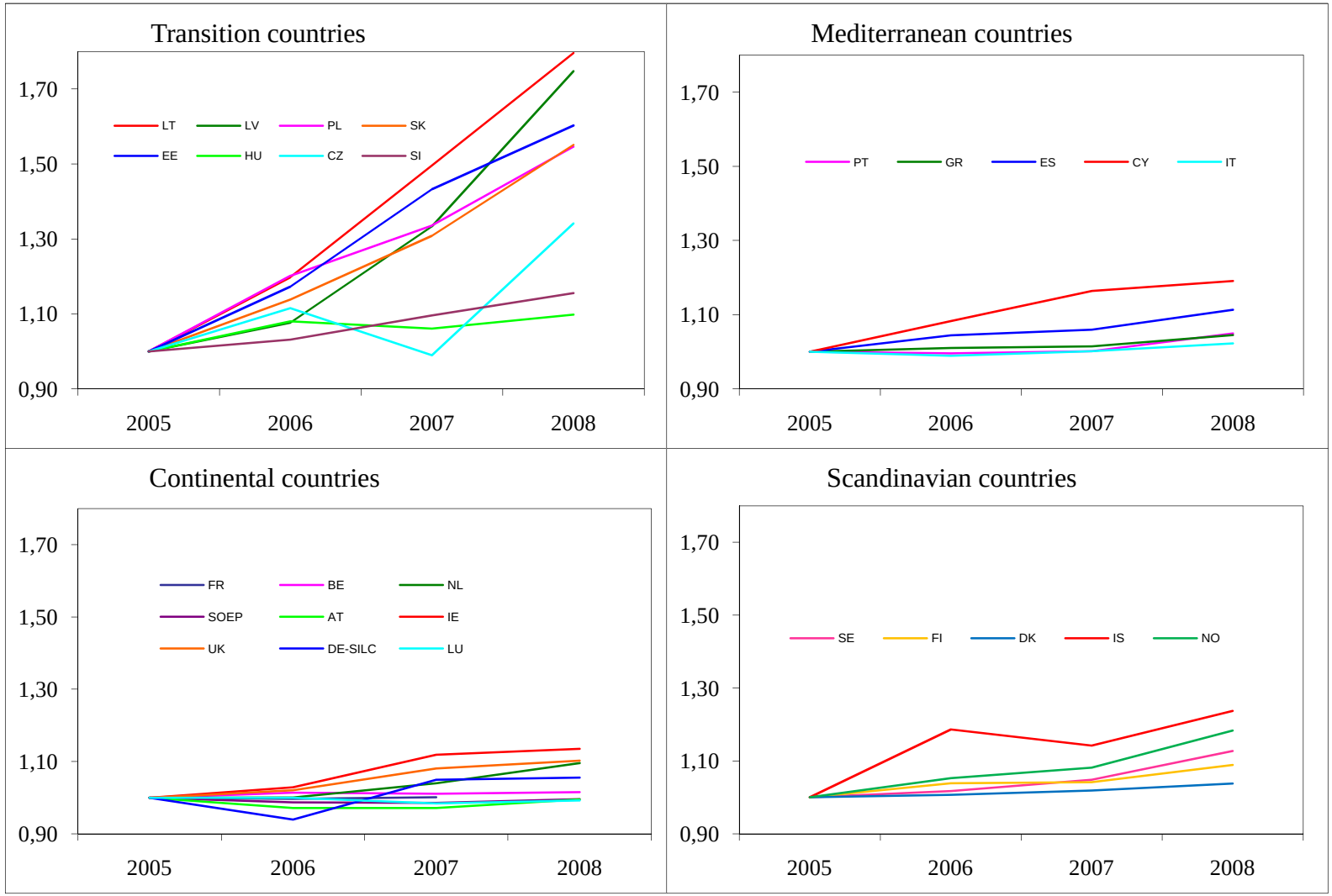
Inequality 2005-2008



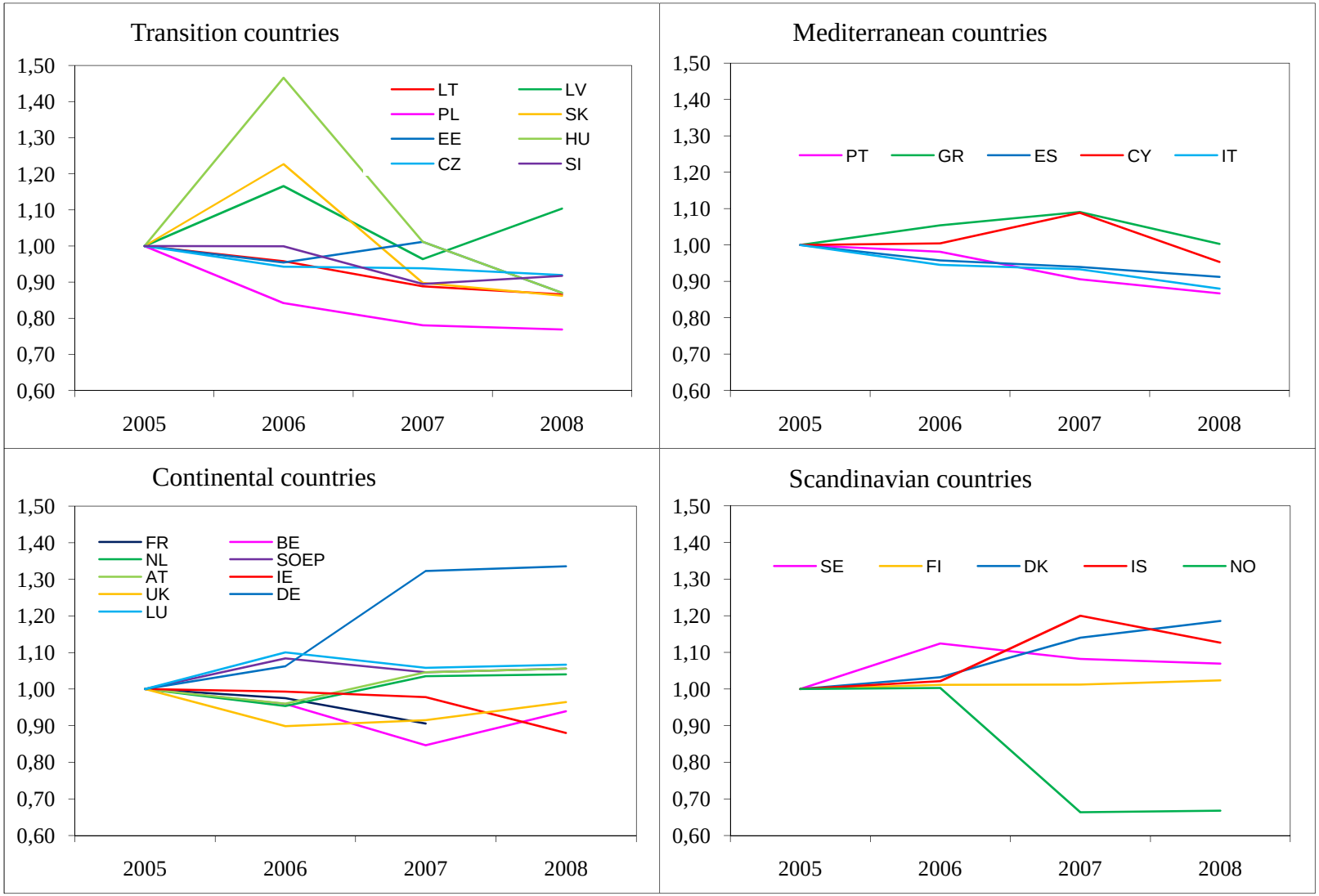
Poverty rate 2005-2008



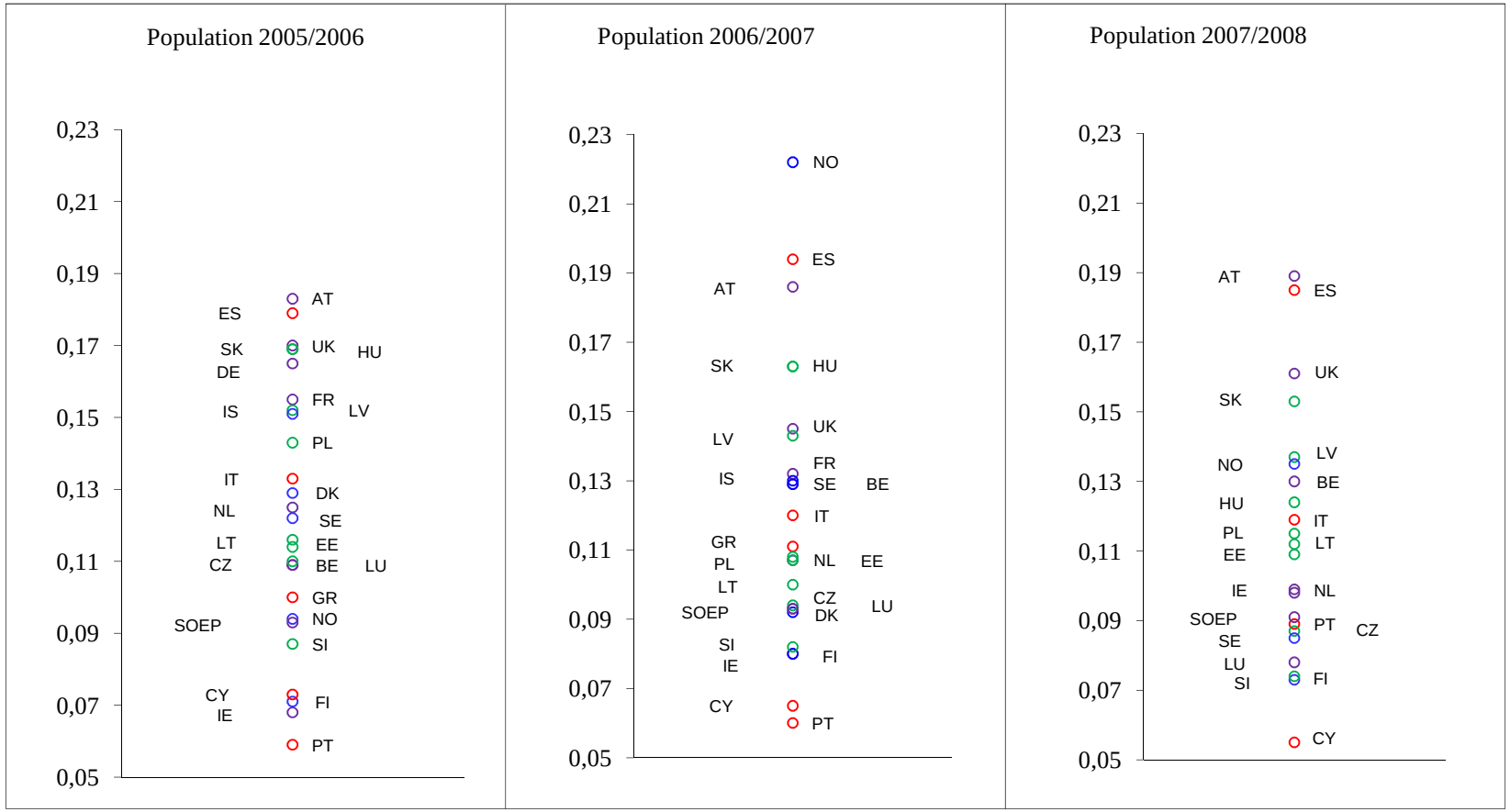
Development of median income



Income inequality – Mean Log Deviation



Shorrocks equalisation measure



Matrix-mobility: Normalised Average Jump