OPTIMIZING STANDARDIZED SURVEY QUESTIONS FOR MEASURING POLITICAL SOLIDARITIES AND RELATED CONCEPTS IN ONLINE SURVEYS

Achim Goerres & Jan Karem Höhne
University of Duisburg-Essen (Germany)

General Online Research Conference
Berlin (Germany) – September 7 to 9, 2022

POLITSOLID

- Questions with rating scales are frequently used in attitude measurement
    - *American National Election Study (ANES)*
    - *European Social Survey (ESS)*

- Design of rating scales can impact answer behavior
    - *Affecting answer distributions*
    - *Inducing systematic measurement errors*
    - *Reducing measurement quality (e.g., validity)*

- Ratings scales may have implications for response effort
    - *Inflating response times*
    - *Affecting answer clicks*

POLITSOLID

- Rating scale characteristics that impact answer behavior and response effort:
  - ***length (number of scale points),***
  - ***verbalization (fully or end verbalized),***
  - ***non-substantive options ("don't know"),***
  - *polarity (unipolar or bipolar),*
  - *numeric labels (with or without numbers),*
  - *direction (decremental or incremental),*
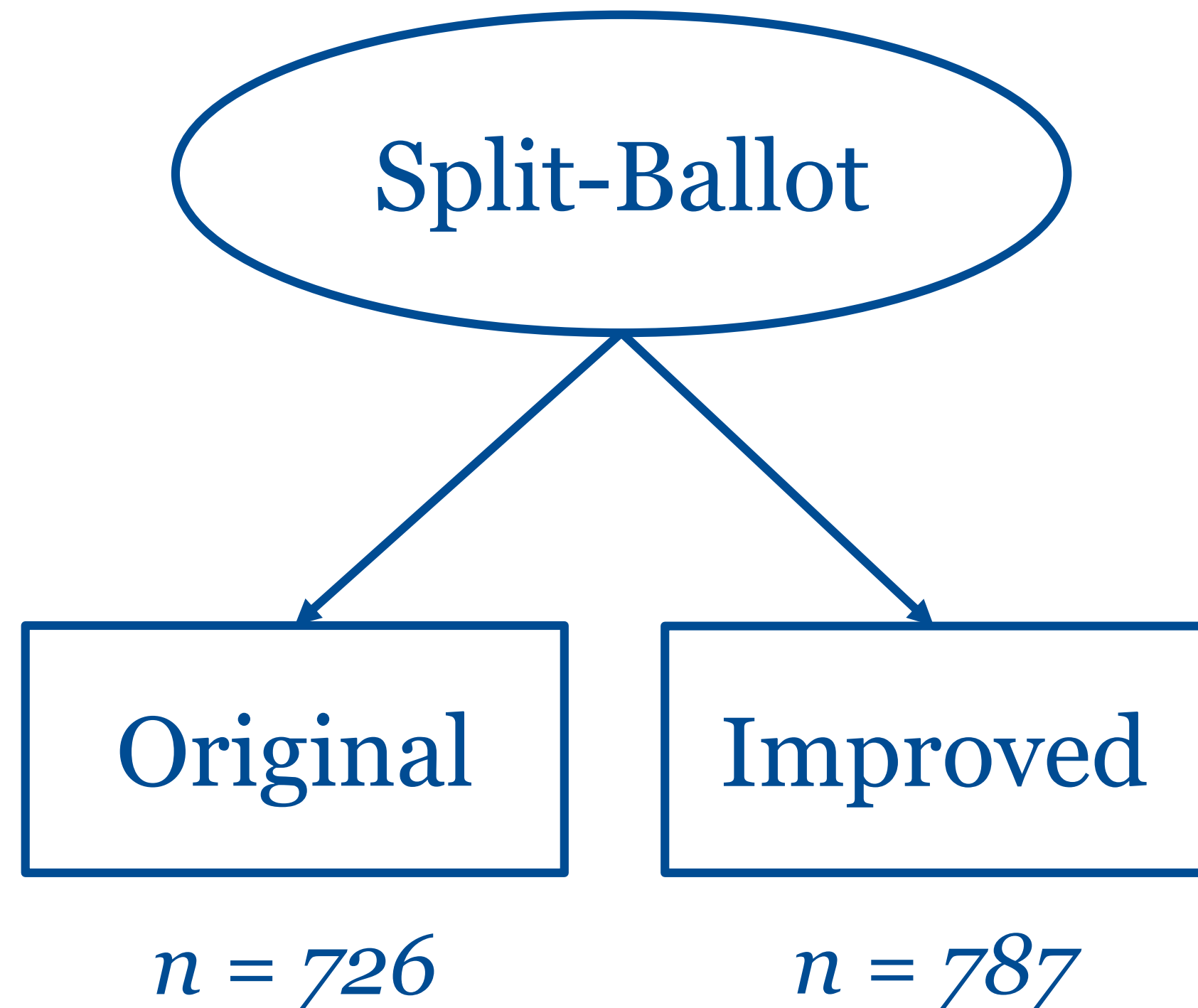 - *alignment (horizontal or vertical).*

Bold indicates characteristics addressed in this study.

- ## Research scope:
    - *Optimizing rating scales of established questions on political solidarity and related concepts*

- ## Research questions (RQs):
    - *Do methodologically improved survey questions, compared to original ones, decrease response effort in terms of response times?*
    - *Do methodologically improved survey questions, compared to original ones, increase data quality in terms of criterion validity?*

Improved: Five-point, end labeled scales without non-substantive options.

POLITSOLID

POLITSOLID

Split-Ballot

Original

Improved

n = 726

n = 787

- Pre-registered study via OSF:
  - *Link: bit.ly/3APag76_OSF_Goerres*
- A total of 16 target questions
  - *Three on redistribution, five on governmental scope, three on social trust, and five on welfare chauvinism*
  - *Taken from social surveys (e.g., ESS)*
- Scale characteristics
  - *Original: Mixed*
  - *Improved: Five-point, end labeled without non-substantive options*
- Vertical alignment
- Optimized survey layout
  - *No horizontal scrolling*

The experiment was conducted in the Forsa Omninet Panel (Germany) in July and August 2021

POLITSOLID

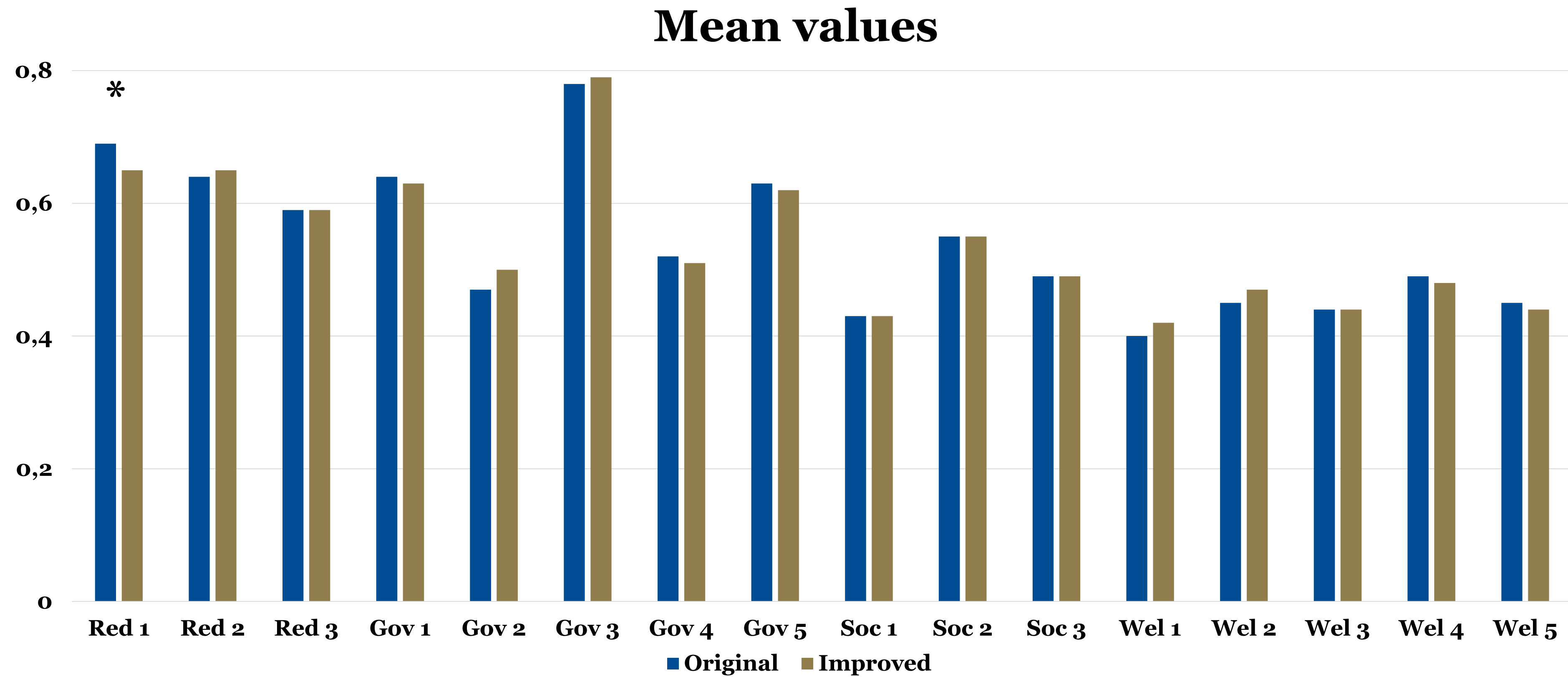|	
Sample size:	*N = 1,513*
Gender:	*49% female*
Age (in years):	*Mean = 52*
Education:	*33% lower secondary school* *27% intermediate secondary school* *40% at least college preparatory secondary school*

Chi-square tests revealed no significant differences between the two experimental groups regarding gender, age, and education.

- Comparing answer distributions
  - *Mean comparisons (t-tests)*
  - *Standardized scales: 0 to 1*

- Comparing response effort
  - *Response times in milliseconds: Embedded Client Side Paradata (Schlosser & Höhne 2018)*
  - *Median comparison (U-tests)*

- Comparing criterion validity
  - *OLS regressions with unstandardized coefficients*
  - *Strength of associations between target and criterion questions*
  - *Criterion questions were significantly correlated with target questions*

POLITSOLID

POLITSOLID

Mean values
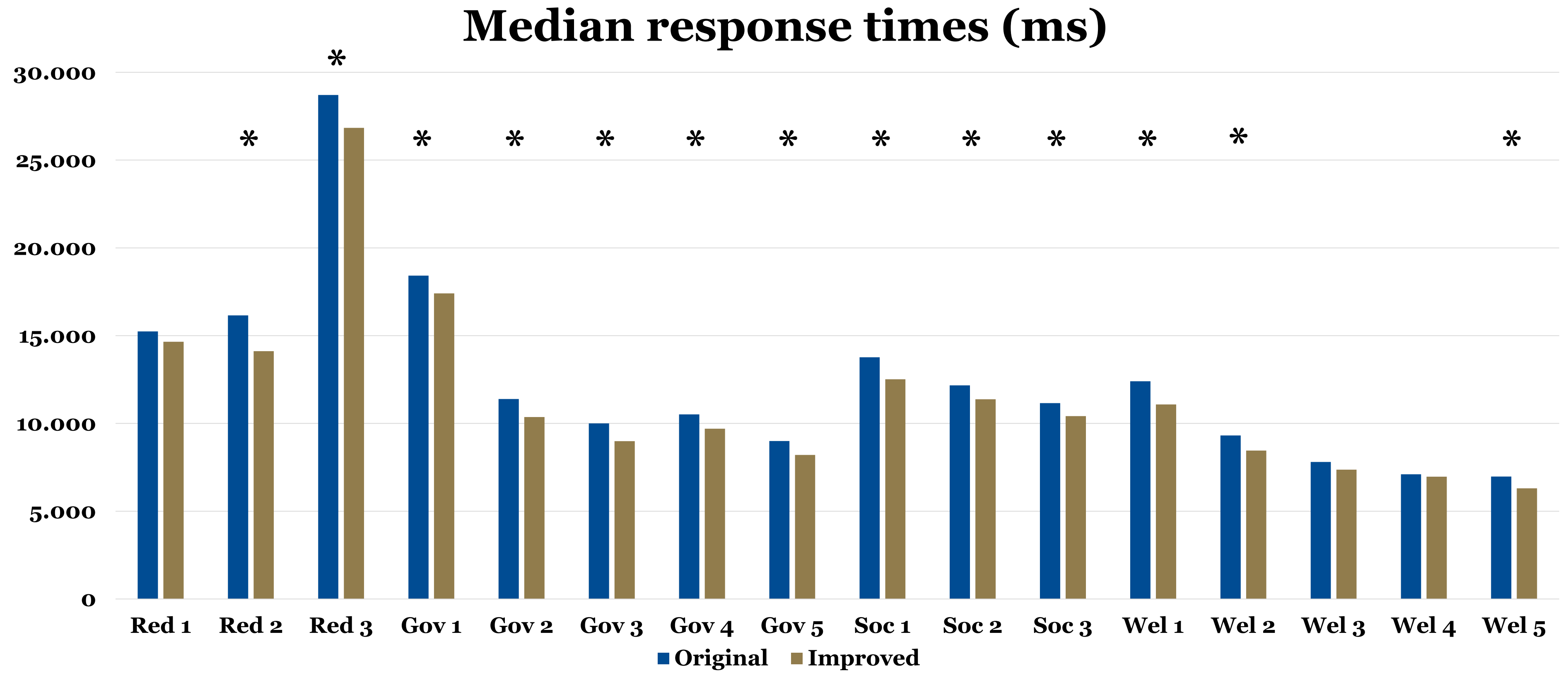
*p < 0.05. t-test. Abbreviations: red = redistribution, gov = governmental scope, soc = social trust, wel = welfare chauvinism.

Median response times (ms)
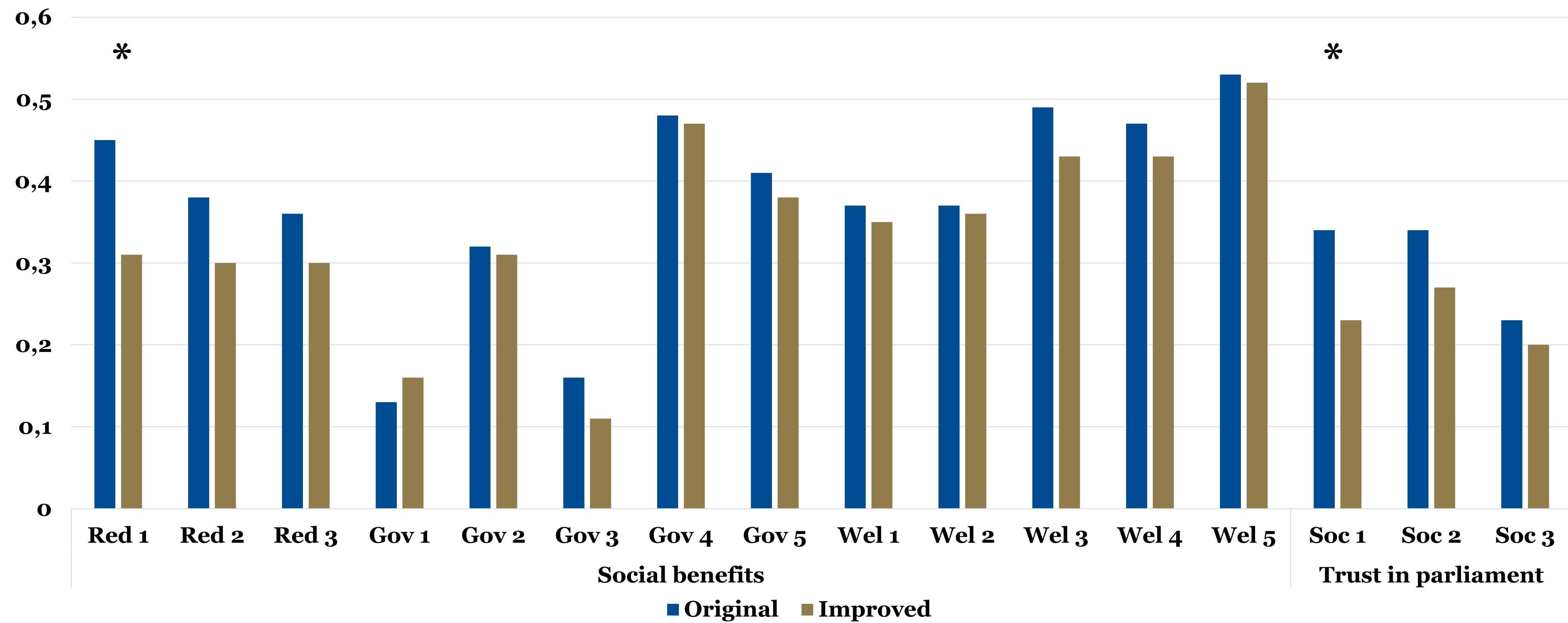
p < 0.05. U-test. Abbreviations: red = redistribution, gov = governmental scope, soc = social trust, wel = welfare chauvinism.

POLITSOLID

9

POLITSOLID

Unstandardized coefficients
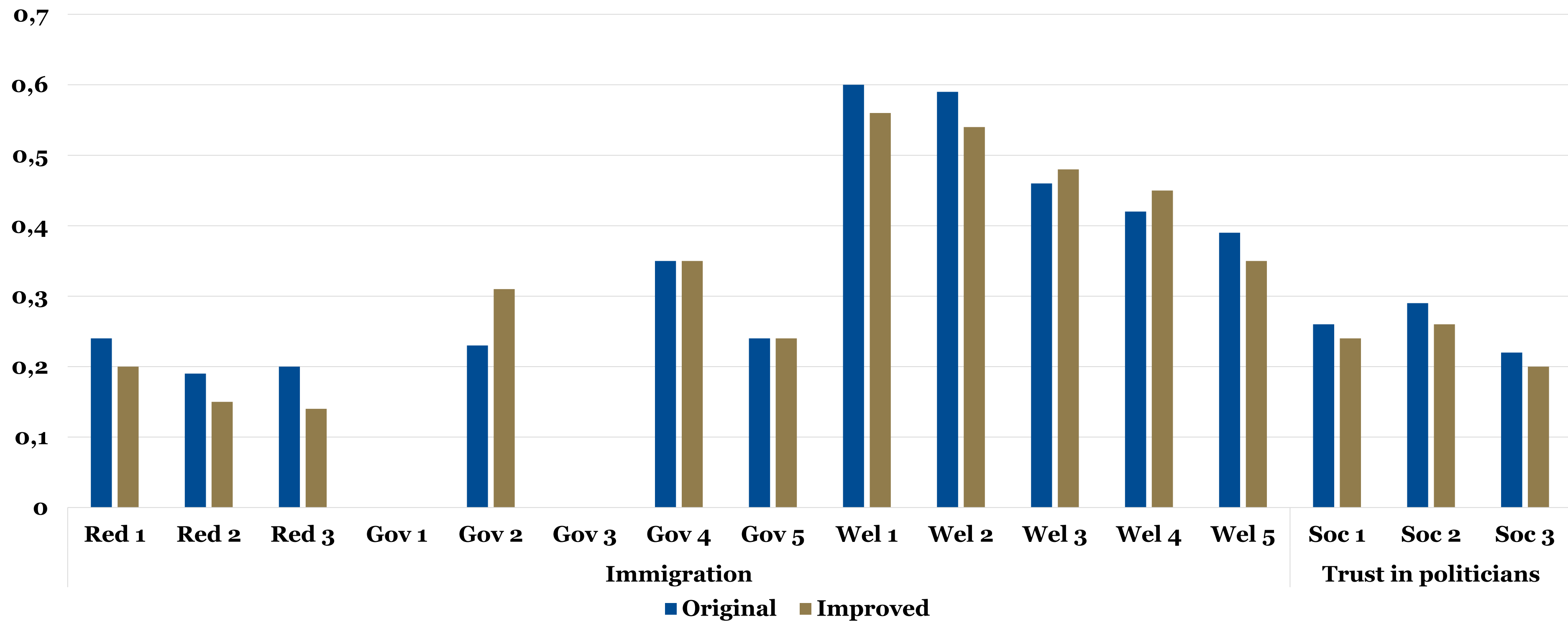
*p < 0.05. OLS regression. Abbreviations: red = redistribution, gov = governmental scope, soc = social trust, wel = welfare chauvinism.

POLITSOLID

Unstandardized coefficients

*p < 0.05. OLS regression. Abbreviations: red = redistribution, gov = governmental scope, soc = social trust, wel = welfare chauvinism. gov 1 and 3 did not correlate significantly with the criterion question on immigrants in the full sample and thus we do not report regression coefficients.

11

POLITSOLID

- **Established measures of political solidarities seem to be robust**
  - *No rating scale effects on answer distributions*
  - *No reduced data quality in terms of criterion validity*
  - *Exploring further data quality indicators*

- **Response effort improvement is desirable**
  - *Established measures result in higher response times*
  - *Reducing burden on respondents and potentially saving costs*
  - *Exploring further response effort indicators*

- **Future research may consider other scale characteristics**
  - *Polarity, numeric labels, direction, alignment etc.*

- **We recommend five-point, end labeled scales without non-substantive options**

This research is funded by the European Research Council (ERC)

POLITSOLID

13

Many thanks for your attention!

jan.hoehne@uni-due.de

www.jkhoehne.eu

POLITSOLID