Clemens Hetschko, Louisa von Reumont & Ronnie Schöb

Pitfalls for the Construction of a Welfare Indicator: An Experimental Analysis of the Better Life Index

University Alliance of Sustainability

Spring Campus

Berlin, 28 March 2017

Welfare measurement

- empirical tools to assess a nation's shape and progress
  - Are people doing well? Are they better off than ten years ago?
  - comparisons of countries / over time
  - policy implications
- GDP, HDI, lengthy lists of capabilities, subjective well-being, …
- when the ultimate list is finally agreed we face the weighting issue, i.e. clarifying how the many indicators translate into overall welfare

Weighting issue – an example

- Human Development Index (HDI)
  - life expectancy, years of education, income - weighted by a very specific formula

  → e.g. one more year of education would increase the German HDI twice as much as one more year of life expectancy
- the formula decides about normative trade-offs
- general problem of all approaches

Resolving the weighting issue: Better Life Index

- ‘top-down part’: 24 indicators of quality of life, chosen by OECD based on conclusions of the Stiglitz-Sen-Fitoussi commission
- ‘bottom-up part’: people weight the indicators in the course of a web-based survey
- tool gains enormous media attention
- tool plays an important role in the ongoing scientific debate about welfare measurement
- OECD seems undecided how far to push the tool and how to deal with the results, at least it reports the results

Weighting process

- 11 dimensions, to be rated from 1 to 5
- relative weight = dimension weight over all weights
- dimensions embed indicators that can be measured
- access to additional information about indicators

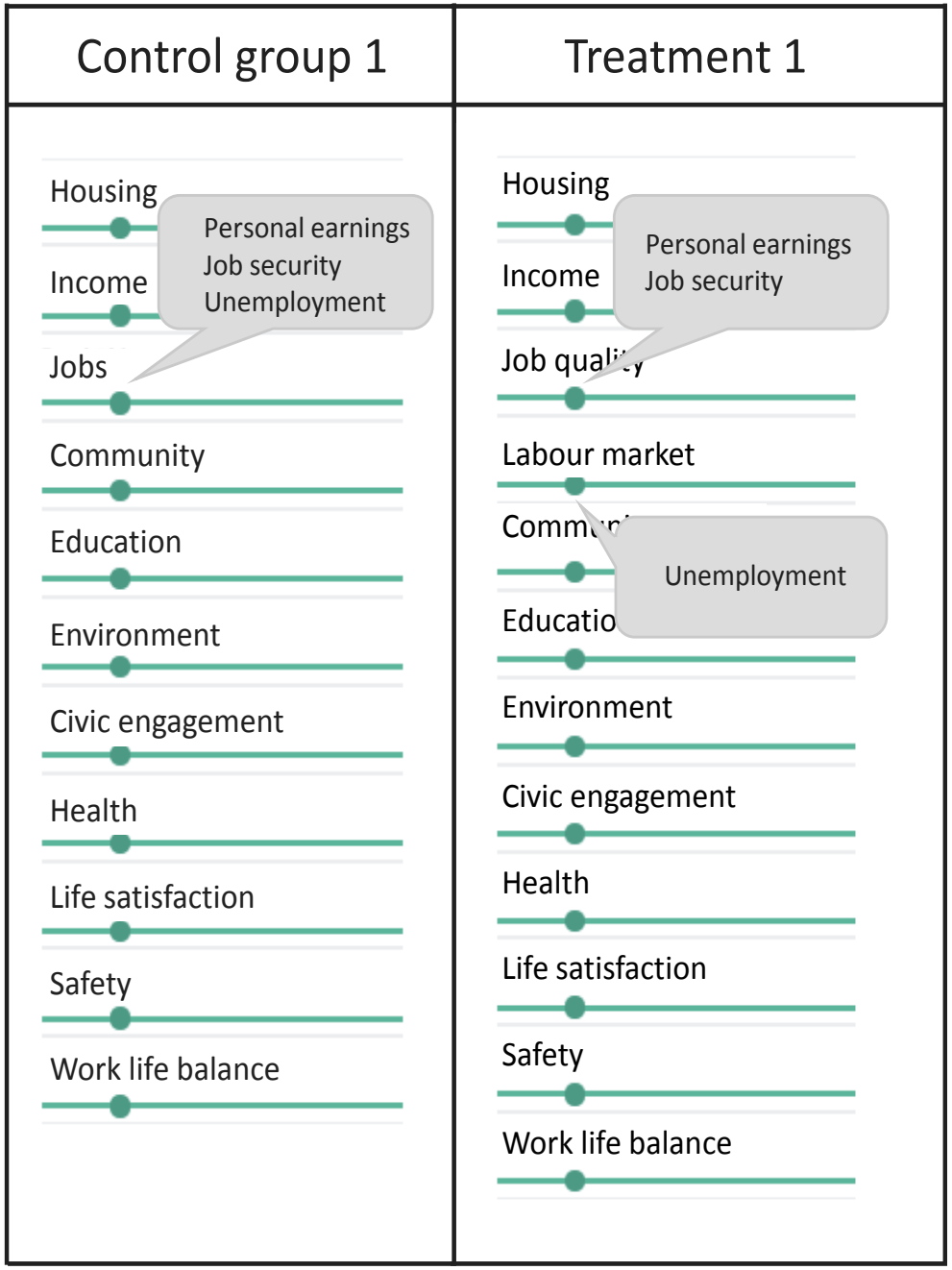
The embedding phenomenon

- detected in surveys where people indicate their willingness to pay for some public project
- people indicate different willingness to pay for a project depending on whether it is presented on its own or as part of a larger category

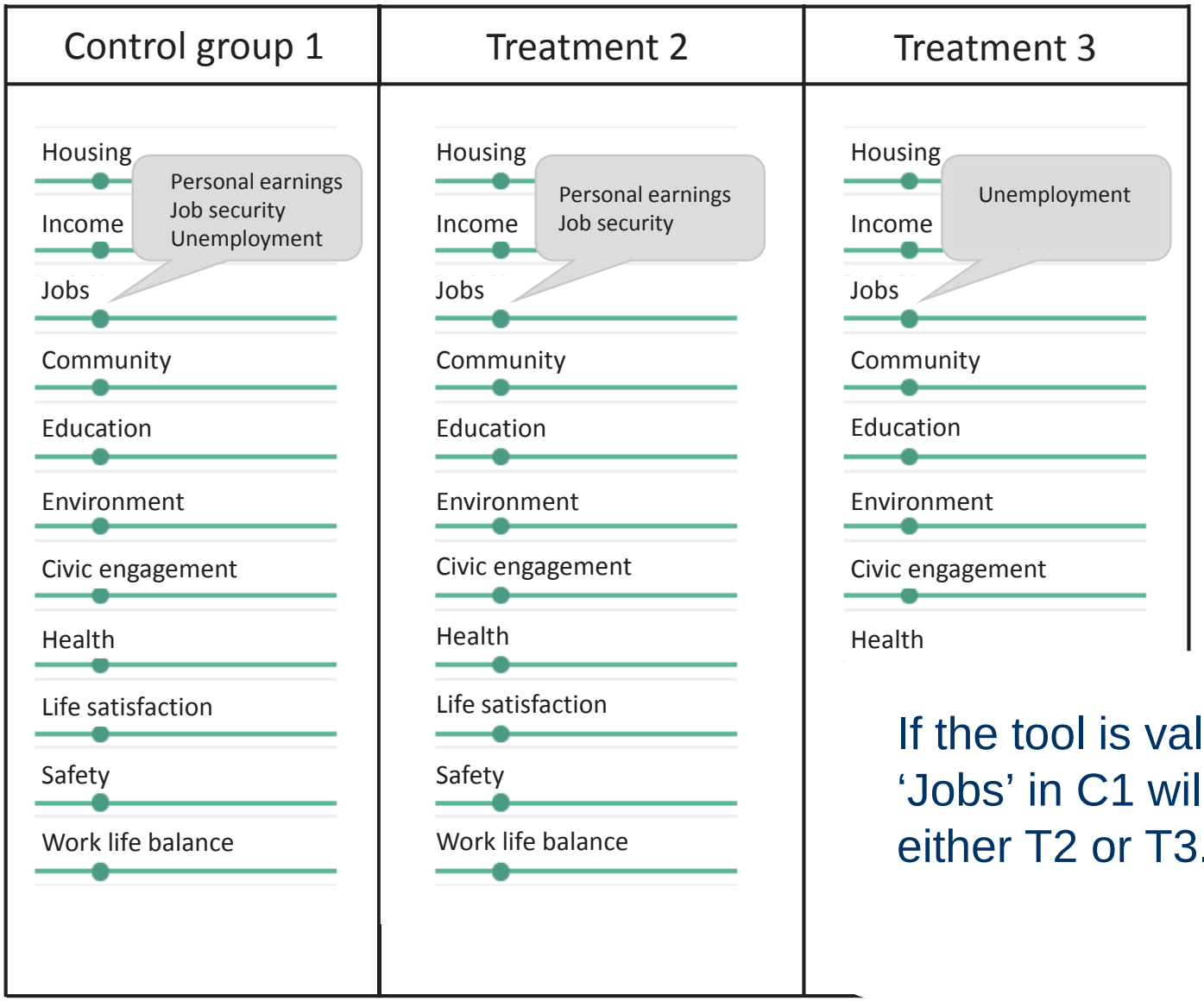
  → true willingness to pay remains unclear
- analogy: specific embedding of indicators in dimensions could affect subjects' ratings of the indicators

Idea of our experiment

- we vary the ‘Jobs’ dimension to test for embedding effects
- reminder: Jobs embeds earnings, job security and unemployment



If the BLI is valid, the weight of 'Jobs' in C1 will equal the sum of the weights of 'Labor Market' and 'Job Quality' in T1.



If the tool is valid, the weight of 'Jobs' in C1 will exceed that of either T2 or T3.

The experiment

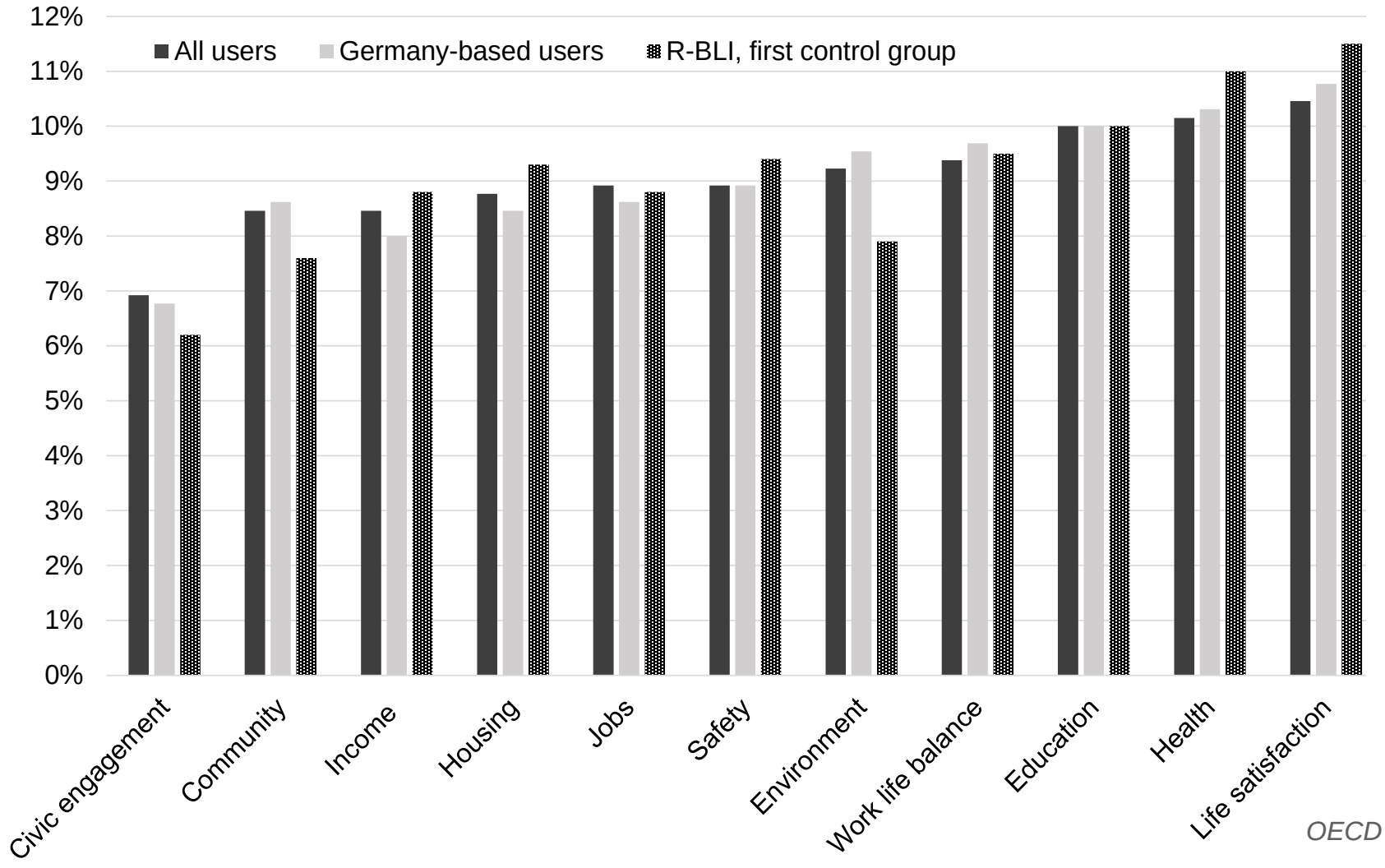
- based on an replication of the OECD's weighting tool ('RBLI')
- RBLI website was accessible from 18/01/16 to 12/02/16, using a ticket (six digit number)
- 2,370 flyers with the **web address** and a ticket were distributed in undergraduate lectures across Germany
  - universities: Rostock, Berlin (TU, FU), Magdeburg, Göttingen, Bochum, Wuppertal, Dresden, Frankfurt
  - response rate of 19.7% (number of observations: 538)
  - tickets assigned participants randomly to control group / treatment groups

Descriptive statistics

Female (share)	46%	
Age (in years)	22.13	(SD = 4.00)
Knowledge of the OECD BLI (share)	21%	
Time spent weighting (in minutes, median)	1:42	
Accessed information (share)	25%	
Size of home town (shares)		
20,000 or less	26%	
20,000 – 100,000	20%	
100,000 – 500,000	19%	
500,000 – 1,000,000	10%	
1,000,000 or more	25%	
Major (shares)		
Economics	18%	
Business Administration	33%	
Mathematics	15%	
Languages	9%	
Arts	8%	
Other	17%	

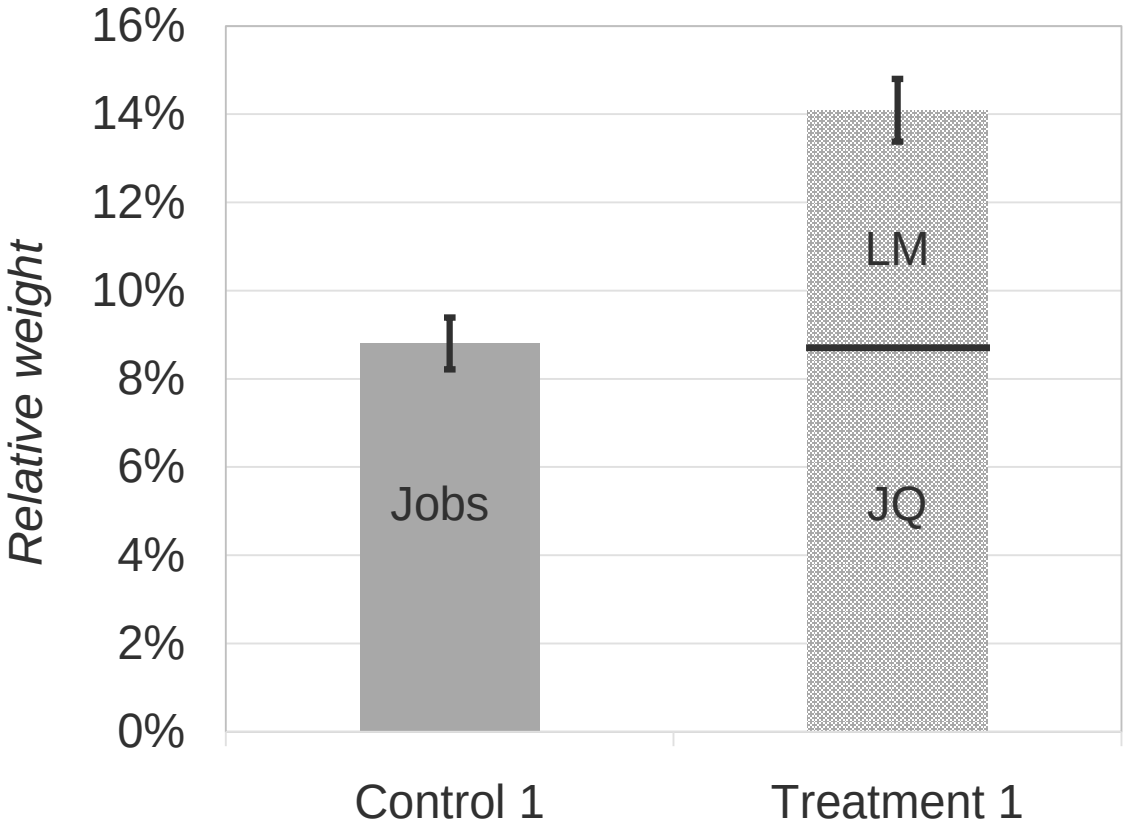
- based on 522 obs.
- drop outs:
  - ✓ 1 invalid
  - ✓ 15: time < 0:45 Min.

Overiew weighting results Control group 1



OECD 2015

Control 1 vs. Treatment 1



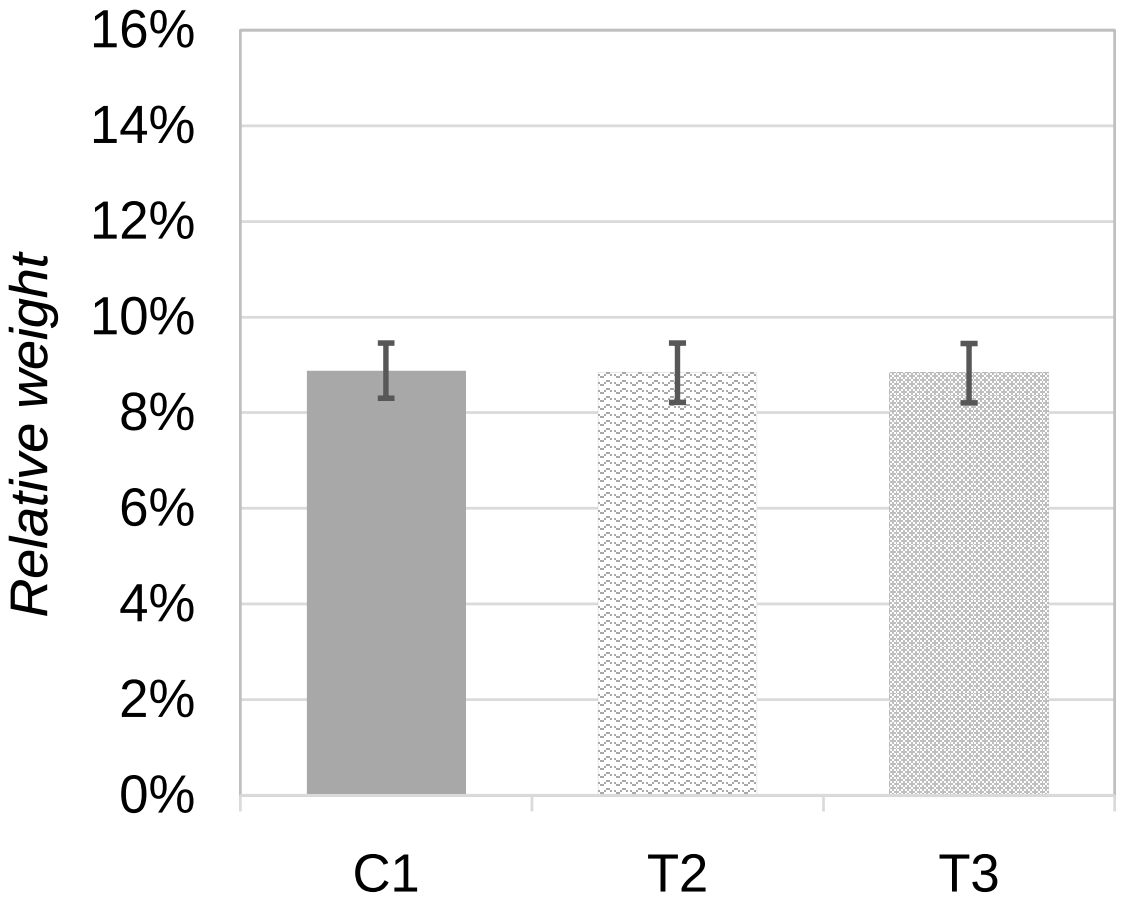
- $\Delta = 0.053$ ($p < 0.0001$)

→ significant embedding effects

- perfect embedding

Whiskers denote 95% confidence intervals.

Control 1 vs. Treatments 2, 3



Finding:

Withdrawing indicators does not affect the Jobs weight at all!

Whiskers denote 95% confidence intervals.

Further analyses

- tests do not imply framing effects to drive C1 vs T1
- regression analyses accounting for socio-demographic characteristics yield the same results
- subgroup tests imply that people who spent a long time weighting / accessed the extra information show the same results

Implications

- strong embedding effects undermine OECD Better Life Index
- possible reasons
  - people answer ‘on the fly’, may tend to apply ‘1/n heuristic’
  - preconceived notions of the dimension titles affect the ratings much more than the embedded indicators

→ Better Life Index no solution to weighting issue

→ results may extend to other survey-based approaches

Freie Universität Berlin

Thank you for your attention!

Clemens.Hetschko@fu-berlin.de

Freie Universität Berlin

Was macht das Leben lebenswert?

Geld? Familie? Karriere? Umwelt?
Im Rahmen einer sehr (!) kurzen Umfrage kannst Du festlegen, was für Dich im Leben am wichtigsten ist und gleichzeitig einen Beitrag zur Beantwortung aktueller Forschungsfragen leisten.

Besuche dazu die Internetseite:

http://goo.gl/hAR8Nv

und gib diese Ticket-Nummer an:

b6c9e2

Dieses Ticket ist nur einmal gültig. Deine Angaben sind völlig anonym. Die Umfrage ist Teil eines Forschungsprojekts des Lehrstuhl Schöb.

Vielen Dank für Deine Unterstützung!

Kontakt:
Louisa.Reumont@fu-berlin.de
Ls-Schoeb@wiwiss.fu-berlin.de

Freie Universität Berlin

What makes life worth living?

Is it Money? Is it Family? A Career? The environment?

This very short (!) survey allows you to determine what is most important for you in your life.

Please visit the website

https://goo.gl/hAR8Nv

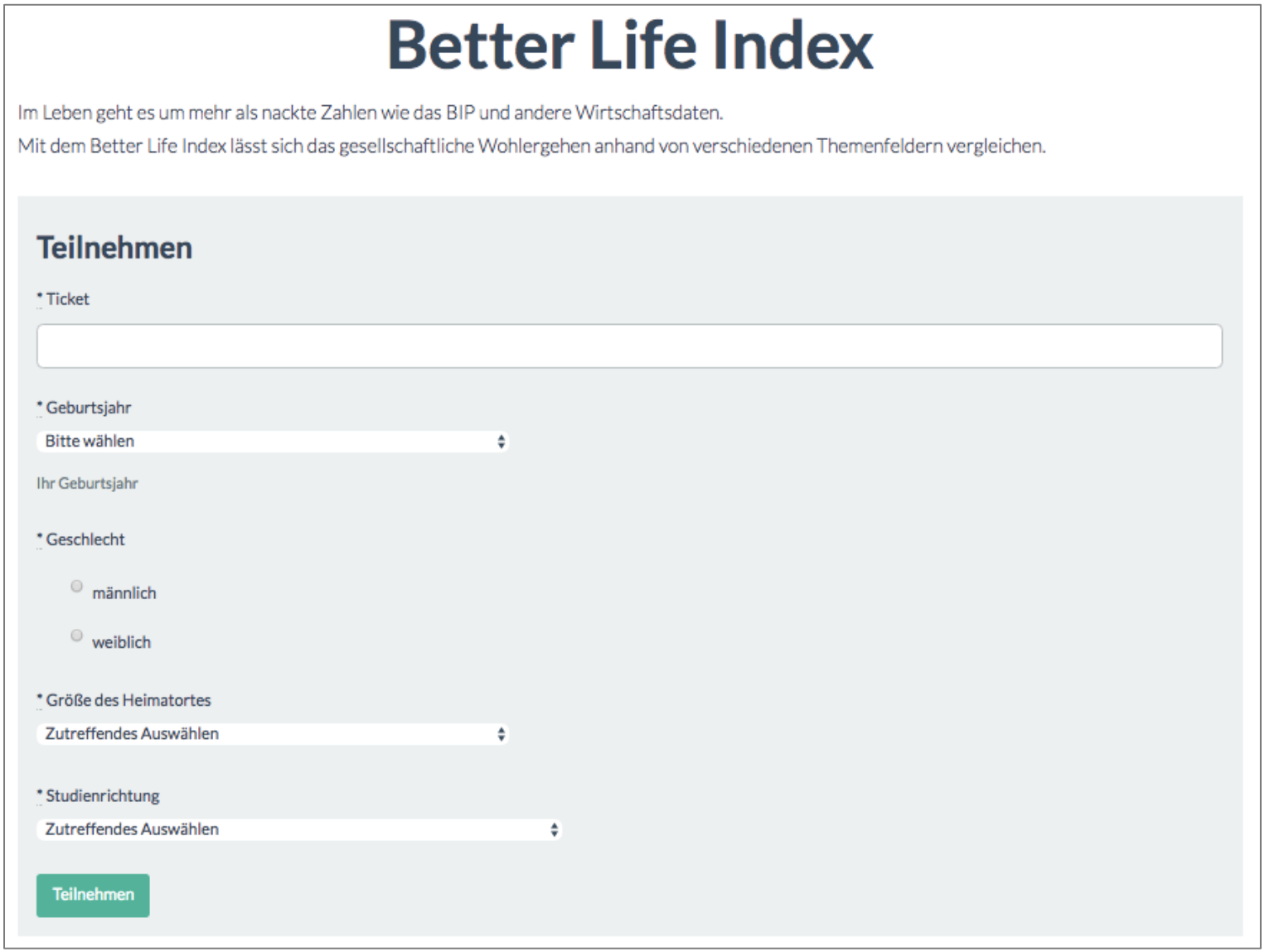
and use the following number as identification:

"Ticket Nummer"

This ticket is valid only once. Your answers will remain completely anonymus. This survey is part of a research project of the Chair of International Public Economics.

Thank you very much for your support!

Contact:
Louisa Reumont: Louisa.Reumont@fu-berlin.de
Chair of International Public Economics: LS-Schoeb@fu-berlin.de



Better Life Index

Im Leben geht es um mehr als nackte Zahlen wie das BIP und andere Wirtschaftsdaten.

Mit dem Better Life Index lässt sich das gesellschaftliche Wohlergehen anhand von verschiedenen Themenfeldern vergleichen.

Teilnehmen

* Ticket

* Geburtsjahr

Bitte wählen

Ihr Geburtsjahr

* Geschlecht

männlich

weiblich

* Größe des Heimatortes

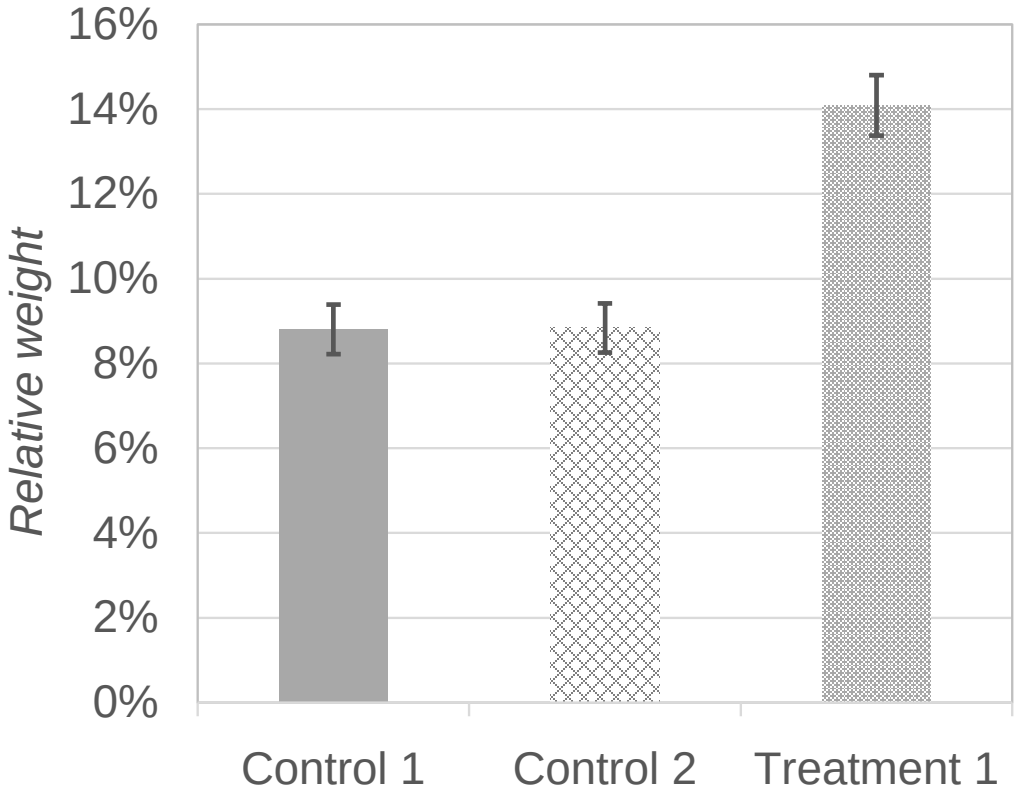
Zutreffendes Auswählen

* Studienrichtung

Zutreffendes Auswählen

Teilnehmen

Control 1 vs Control 2 vs Treatment 1



- no difference C2 / C1
- difference C2/T1 the same as C1/T1

→ $\Delta = 0.053$ ($p < 0.0001$)

Whiskers denote 95% confidence intervals.

Regression analyses

$$RW_i(Jobs) = \alpha + \beta_1 T_{1,i} + \beta_2 T_{2,i} + \beta_3 T_{3,i} + \gamma C_2$$

$$+\delta FEM_i + \phi AGE_i + SIZE_i'\lambda + MAJOR_i'\mu_i$$

$$+\theta KNOWS_i + \sigma INFO_i + \tau LONG_i + \varepsilon_i$$

Regression analyses

	I	II	III
Experimental groups (ref. control group 1)			
Control group 2	0.000	0.000	0.000
	(0.004)	(0.004)	(0.004)
Treatment group 1	0.053***	0.053***	0.054***
	(0.005)	(0.005)	(0.005)
Treatment group 2	-0.005	-0.005	-0.005
	(0.004)	(0.004)	(0.004)
Treatment group 3	-0.004	-0.004	-0.004
	(0.004)	(0.004)	(0.004)
Individual characteristics (gender, age, size of home town, major)		Yes	Yes
Weighting characteristics (knows BLI, time spent weighting, accessed extra information)			Yes
Constant	0.088***	0.081***	0.081***
	(0.003)	(0.010)	(0.010)
Observations	522	522	522
R^2	0.313	0.325	0.329

*Dependent variable: relative weight of Jobs, estimation: OLS, robust standard errors in parentheses, ***p < 0.01, **p < 0.05, *p < 0.1*

Subgroup analyses

	initial sample	*female*	*male*	*age below 21 years*	*age above 21 years*	*small town*	*large town*	*major econ./bus adm. or business*	*major not econ./bus. adm.*
Experimental groups (ref. Control group 1)									
Control group 2	0.000	0.003	-0.001	0.003	-0.003	-0.004	0.002	0.002	-0.004
	(0.004)	(0.006)	(0.006)	(0.005)	(0.007)	(0.007)	(0.005)	(0.006)	(0.006)
Treatment group 1	0.054***	0.059***	0.047***	0.056***	0.049***	0.056***	0.051***	0.051***	0.054***
	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
Treatment group 2	-0.005	-0.009	-0.001	-0.005	-0.007	0.002	-0.012**	-0.005	-0.005
	(0.004)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)
Treatment group 3	-0.004	0.007	-0.012*	-0.004	-0.007	0.002	-0.010*	-0.003	-0.009
	(0.004)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.005)	(0.006)	(0.006)
Individual characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Weighting characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.081***	0.083***	0.080***	0.093***	0.094***	0.054***	0.089***	0.070***	0.097***
	(0.010)	(0.015)	(0.013)	(0.009)	(0.008)	(0.021)	(0.010)	(0.012)	(0.012)
Observations	522	239	283	301	221	239	283	269	253
R^2	0.329	0.436	0.266	0.366	0.311	0.336	0.350	0.317	0.374

Robust standard errors in parentheses.

Subgroup analyses

	initial sample	*knows BLI*	*does not know BLI*	*short time spent*	*long time spent*	*read extra info*	*did not read extra info*
Experimental groups (ref. Control group 1)							
Control group 2	0.000	0.009	-0.003	-0.006	0.003	0.013	-0.005
	(0.004)	(0.010)	(0.005)	(0.006)	(0.006)	(0.008)	(0.005)
Treatment group 1	0.054***	0.064***	0.050***	0.050***	0.056***	0.049***	0.054***
	(0.005)	(0.010)	(0.005)	(0.007)	(0.006)	(0.011)	(0.005)
Treatment group 2	-0.005	0.003	-0.009*	-0.007	-0.006	-0.001	-0.006
	(0.004)	(0.008)	(0.005)	(0.006)	(0.007)	(0.010)	(0.005)
Treatment group 3	-0.004	0.003	-0.006	-0.015**	0.004	0.014	-0.011**
	(0.004)	(0.011)	(0.005)	(0.007)	(0.006)	(0.010)	(0.005)
Individual characteristics	yes	yes	yes	yes	yes	yes	yes
Weighting characteristics	yes	yes	yes	yes	yes	yes	yes
Constant	0.081***	0.096***	0.080***	0.082***	0.072***	0.062***	0.087***
	(0.010)	(0.016)	(0.013)	(0.012)	(0.017)	(0.023)	(0.011)
Observations	522	112	410	257	265	130	392
R^2	0.329	0.505	0.299	0.385	0.331	0.246	0.382

Robust standard errors in parentheses.