

von Reumont, Louisa; Schöb, Ronnie; Hetschko, Clemens

Conference Paper

Embedding Effects in the OECD Better Life Index

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Political Economy II, No. E20-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: von Reumont, Louisa; Schöb, Ronnie; Hetschko, Clemens (2017) : Embedding Effects in the OECD Better Life Index, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2017: Alternative Geld- und Finanzarchitekturen - Session: Political Economy II, No. E20-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/168133>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Embedding Effects in the OECD Better Life Index: Evidence from an Experiment

Clemens Hetschko

Freie Universität Berlin, CESifo

Louisa von Reumont*

Freie Universität Berlin

Ronnie Schöb

Freie Universität Berlin, ifo Dresden, CESifo

This version: November 10, 2016

Abstract

The OECD's aim is to foster 'better policies for better lives'. To this end, the organization invites people to evaluate a variety of quality of life indicators according to their individual preferences, using the web-based Better Life Index. These indicators are embedded in 11 broader dimensions of well-being. Our experiment shows that this procedure yields strong embedding effects. This finding casts serious doubts on the reliability of the Better Life Index as the procedure fails to elicit the citizens' true preferences. The results thus challenge the usefulness of the Better Life Index in providing meaningful policy recommendations.

JEL Classification Codes: B41, C42, I31

Keywords: beyond GDP, welfare measurement, OECD Better Life Index, embedding effect

**Corresponding author.* Louisa von Reumont. Address. Freie Universität Berlin, School of Business & Economics, Chair of International Public Economics, Boltzmannstraße 20, 14195 Berlin, Germany, E-mail: Louisa.Reumont@fu-berlin.de.

Acknowledgements. We are very grateful to Stephen Lumenta for programming and implementing the experimental web-application as well as to Malte Preuss, Tobias Wolf, Katarina Bjelajac Witzheller, Daniel Nachtigall, Andreas Knabe, Marcel Thum, Robert Schwager, Nadine Riedel, Robert Fenge, Kerstin Schneider and Alfons Weichenrieder for their support in recruiting participants for the experiment. For very helpful comments, we thank participants of workshops at the Max Planck Institute for Tax Law and Public Finance (Munich, 2016) and the Institute for Labor Law and Industrial Relations in the EU (Trier, 2016).

1. Introduction

The welfare of nations is difficult to assess. For decades, the gross domestic product (GDP) has been the preferred measure to compare and evaluate national wealth. Yet right from the beginning, GDP has been acknowledged to be an inadequate measure of well-being (cf. Coyle 2014). Several commissions of experts have thus thought about alternatives to GDP. One of the most famous is the Stiglitz-Sen-Fitoussi Commission (see Stiglitz et al. 2009), which comprised a group of renowned scientists, including five Nobel laureates in economics. The commission suggests a broadly conceived multidimensional statistical system. Measures “centred on people’s well-being” (Stiglitz et al. 2009, p.12) as well as measures reflecting environmental sustainability are to supplement GDP. Such a coherent and explicit multidimensional database, including both objective and subjective data on wellbeing, could allow policy-makers to identify, assess and improve what matters to their citizens. The commission did not recommend aggregating these indicators into a social welfare function. By contrast, the OECD Better Life Initiative offers a concrete aggregation procedure with its Better Life Index, introduced in 2011 (BLI; OECD 2011).

Following the Stiglitz-Sen-Fitoussi Commission, the OECD considers 11 *dimensions* of life – housing, income, jobs, community, education, environment, civic engagement, health, life satisfaction, safety and work-life balance – to be universal and thus essential and relevant to all societies (Boarini and D’Ercole 2013). The dimensions summarize a set of 24 more detailed and measurable indicators, both objective and subjective by nature. The interactive web-based survey allows for the individual valuation of these dimensions as users are asked to rate the dimensions on a scale from 0 (not important) to 5 (very important) each. These values are transformed into relative weights that sum up to 100% and thus produce individualized social welfare functions (OECD 2011, p. 32). By directly asking people to judge the importance of different determinants of social welfare, the OECD aims to establish the BLI as a means to directly involve citizens in the public discourse (Boarini and D’Ercole 2013, Durand 2015). Aggregating the individuals’ assessments yields a social welfare function which should serve as a basis for designing ‘better policies for better lives’ (OECD 2015, p.17) – the principal goal of the OECD.¹ This procedure has already gained considerable media attention and political support. For instance, striving for a better understanding of its citizens’ preferences, the Austrian government employed the BLI to assess and investigate the priorities of the different

¹ The BLI also enables users to see and compare how different OECD countries rank on the various dimensions rated given their personal weighting of the 11 dimensions (OECD 2011).

dimensions through an extensive survey (WIFO 2012).

The BLI involves users rating the 11 aggregated well-being dimensions in lieu of the detailed underlying indicators. If people were fully aware of their social preferences and had the necessary capabilities to perform complex weighting, rating one aggregated set of dimensions would lead to the same weights as rating another set of dimensions aggregated from the same set of indicators. However, there is strong evidence that the valuation of a particular item might vary substantially depending on whether it is valued on its own or as a part of a larger, more inclusive item. This phenomenon has been labelled as *embedding effect* (Kahneman and Knetsch 1992a) in the contingent valuation method (CVM), where it severely biases subjects' willingness-to-pay (WTP) assessments for non-market environmental goods. Likewise, the valuation of different well-being dimensions which the BLI asks for may be sensitive to how the dimensions embed the underlying indicators. If users of the BLI are only asked to rate overarching dimensions instead of the underlying indicators, it remains unclear whether and to what extent the users make the necessary weighting of all the underlying indicators. In consequence, the relative weights might differ depending on whether a dimension is weighted as a whole or whether the indicators themselves are rated individually. Should the specific embedding indeed matter, the true indicator preferences of citizens cannot be inferred.

Evidence on the reliability of the OECD Better Life Index is missing so far. Previous studies which discuss the BLI mainly focus on problems arising when the index is used for country comparisons (see Mizobuchi 2014, Kasparian and Rolland 2012). Our study, by contrast, investigates the more fundamental concern that the BLI is prone to embedding effects and thus might not measure social welfare in a meaningful way. Our experiment assigns randomly selected subgroups of participants different versions of a replication of the BLI varying the embedding of the indicators that form a dimension. We find that the specific embedding of the same sets of indicators strongly impacts the relative weights assigned to these indicators and thus substantial embedding effects. Weights seem to be given on the fly and/or depend on individual preconceived notions rather than on a profound valuation of the specific indicators that are to be rated. By implication, the BLI does not yield reliable results and thus cannot reveal citizens' actual preferences.

We proceed as follows. Section 2 introduces the embedding phenomenon in more detail and describes its potential importance for the BLI. We describe the experiment, data and empirical methodology in Section 3. Section 4 presents the results and Section 5 concludes.

2. The embedding effect

Societies are confronted with complex assessments of environmental amenities. How much worth can be assigned to a certain landmark? How should one value an old-growth forest? The contingent valuation method is frequently employed for the purpose of environmental valuation. Survey respondents either state their willingness-to-pay or their willingness-to-accept for the item in question, through which the value of the (intangible) good is determined (Hanemann 1994, Carson and Hanemann 2005, Kling et al. 2012). The reliability of the method has been heavily disputed. Subjects “are often not responding out of stable well-defined preferences, but are essentially inventing their answers on the fly, in a way which makes resulting data useless for serious analysis” (Hausman 2012, p. 43).

For our purpose of evaluating the BLI, we expect one line of critique of the CVM to be of particular importance, the embedding phenomenon. It implies that the valuation of an item varies substantially “(...) depending on whether the good is assessed on its own or embedded as part of a more inclusive package” (Kahneman and Knetsch 1992a, p. 58). This effect can show up in a weak and a strong form, respectively called regular and perfect embedding (Loomis 1993, Svedsäter 2000, McDaniels et al. 2003).² *Regular embedding* occurs if the value assigned to a single good is higher when evaluated on its own compared to the value assigned to it when embedded within a larger, more inclusive good (Kahneman and Knetsch 1992b; Svedsäter 2000). *Perfect embedding* occurs if the valuation of a single good is similar to the value assigned to the more inclusive good, containing the item in question. Evidently, the stronger notion of *perfect embedding* implies *regular embedding* (Kahneman and Knetsch 1992b).

Kahneman and Knetsch (1992a) detect the embedding phenomenon in a survey collecting subjects’ willingness-to-pay for changes in the provision of environmental services in the Greater Vancouver region. A first group is asked about their WTP for the improvement of environmental services as overarching category, their WTP for the subcategory improvement of disaster preparedness and for the concrete measure of improving rescue equipment and personnel. A second group indicates the WTP only for the latter two items. A third group values solely the last item. The average WTP of this most specific item varies substantially depending on whether it is evaluated on its own (\$122.64), jointly with preparedness for disasters (\$74.65) or with both preparedness for disasters and environmental services (\$14.12). As the level of

² In their original study Kahneman and Knetsch (1992a) only define the term “embedding effect”. In their reply (Kahneman and Knetsch 1992b) to the critique by Smith (1992), they introduce the distinction between *regular embedding* and its stronger notion, *perfect embedding*.

embedding alters the values assigned to the same item, it becomes completely unclear which embedding structure reliably reveals respondents' actual preferences, if any. Diamond and Hausman (1994) therefore conclude that because of the existence of the embedding effect, contingent valuation should not be used for cost-benefit analysis.

As previously mentioned, we analyse in how far the BLI might also suffer from embedding effects. Both BLI and CVM ask respondents to reveal their preferences by considering a trade-off when assigning values to the different items in question. In the assessment of an environmental good, subjects have to assign a particular monetary value given their limited budget. For the BLI the values assigned to each dimension are converted into relative weights that reflect the individuals' marginal rate of substitution between two social objectives measured by the respective indicators. Scarcity implies that an increase of the relative weight of one particular objective (i.e. the aggregated individual preferences for an objective) should raise the amount of scarce resources invested in achieving this particular objective and reduce the resources invested in all of the other objectives. Whether or not a policy that shifts resources from one dimension to another dimension enhances social welfare thus crucially hinges on the accurate and consistent measurement of the relative weights. Embedding effects may distort these relative weights and thus make it impossible to provide reliable information for 'better policies for better lives'.

3. Experiment and hypotheses

In order to investigate the possible existence of the embedding effect, we set up a web application very similar to the BLI web application, which allows us to test the reliability of the BLI. Using this web application we conducted an experiment throughout Germany.

3.1 Recruitment of participants

From the 18th of January to the 5th of February 2016, we recruited participants for our experiment across Germany, namely at the universities of Bochum, Dresden, Frankfurt, Göttingen, Magdeburg, Rostock, Wuppertal as well as FU Berlin and TU Berlin.³ In introductory undergraduate courses we distributed information flyers that indicate that, for research purposes, it is of interest to know "what makes life worth living" (see Appendix 1). Each flyer shows a randomly created six-digit identification number (in the following: 'ticket')

³ Our sample therefore matches the age group that uses the OECD weighting tool the most as the vast majority of worldwide and also of German users are between 15 and 35 years old (OECD 2014).

and a web address. The ticket ensures anonymous participation and randomly assigns participants to one of the treatment and control groups (see below). 2,730 tickets were distributed, 538 students participated in the experiment, i.e. the response rate is 19.7%. The web address leads participants to the landing page of our replication of the BLI website, which we name ‘Replicated BLI (R-BLI)’ in the following. It was accessible from the 18th of January to the 12th of February 2016. The landing page introduces the BLI and its usage (see Appendix 2 to compare the OECD-BLI website and the R-BLI landing page). At first, participants enter their ticket number.⁴ Then, they provide information about their year of birth, gender, the size of their home town (less than 20k, 20k-100k, 100k-500k, 500k-1M, more than 1M) and their major subject of study (economics, business administration, mathematics, engineering, natural sciences, law, medicine, languages, social sciences, applied arts, teaching, non-enrolled). If they finally agree to participate, they are redirected to the R-BLI weighting tool.

3.2 Weighting of well-being dimensions

Our weighting tool invites participants to rate the dimensions of the BLI on a scale from 0 to 5. As with the original BLI the valuation bar is set to an equal valuation of all dimensions at 1 at the beginning, but can be moved over the whole scale. The ratings assigned to each dimension are converted into relative weights, which, identical to the BLI weighting mechanism, add up to 100% (i.e. the relative weight is the rating of one dimension divided by the sum of the ratings assigned to all of the dimensions). To make participants aware of the trade-off resulting from the conversion into relative weights, a bar chart is included in the R-BLI weighting tool. Each bar reflects the relative weight assigned to the respective dimension and adjusts immediately in response to changes of the weights (see Appendix 2). The visualization of the trade-off resulting from the relative weighting is important in order to replicate the BLI, where the trade-off is visualized through the dynamic adjustment of petals.

The information provided about the R-BLI is in principle the same as provided on the BLI website. Informative mouse-over pop-ups, summarizing the indicators of the dimensions to be rated, show up once participants move their cursor over the title of each dimension. In contrast to the BLI website, the pop-up also appears on the R-BLI page when moving the cursor over the weighting bars. Thus, the underlying indicators are even more visible using our weighting tool than using the original tool. At any stage of the rating process, participants can thus easily see which indicators are embedded in the respective dimension. We measured the time subjects

⁴ As a consequence, the ticket number becomes invalid to avoid multiple use.

spent to complete the ratings. Respondents who took less than 45 seconds are excluded from the sample as it is rather unlikely that they reflected seriously upon the importance of all the dimensions. As it turns out, we would obtain the same qualitative results without this restriction.

Each dimension title on the BLI website is linked to a subpage providing detailed information on the underlying indicators. Our website also includes such links to pdf files giving the same information as the BLI subpages (see Supplementary Material). It was measured whether users accessed this additional material or not. When the rating is completed, respondents finally indicate whether they are familiar with the original BLI.

3.3 Main treatment

Participants are assigned different versions of the R-BLI weighting tool, which only varies with respect to the *Jobs* dimension. The BLI version of *Jobs* embeds four indicators, namely long-term unemployment rate, employment rate, personal earnings and job security. *Jobs* is thus the dimension with the highest number of indicators, providing us with many opportunities to redefine dimensions by regrouping indicators.

A first version of the R-BLI weighting tool replicates the original BLI weighting tool. It includes eleven dimensions in total, one of them is *Jobs*, which covers the four indicators listed above. Note, however, that the OECD's mouse-over pop-up differs slightly from the underlying indicators. It contains the two indicators personal earnings and job security but subsume the other two indicators under the heading 'unemployment'. People whose tickets assign them to this first version of our weighting tool constitute a first control group (C1). A first treatment group (T1) consists of participants who rate twelve dimensions, since the *Jobs* dimension is split up into two distinct dimensions. Personal earnings and job security are individual job-specific indicators, reflecting an important part of the job quality employed people enjoy themselves (e.g. Clark 2001). We subsume these two indicators in the new dimension *Job Quality (JQ)*. The long-term unemployment rate and the employment rate, by contrast, refer to the overall labor market situation. They constitute the new dimension *Labor Market (LM)*. The information in the mouse-over pop-ups changes accordingly. The *JQ* dimension prompts the pop-up information 'personal earnings' and 'job security', whereas *Labor Market* prompts 'unemployment'.

If people are fully aware of the weights attributed to the four indicators, the sum of the relative weights (*RW*) the treated assign to the two new dimensions should not differ from the controls' relative weighting of *Jobs*. In this case, we would not observe any embedding effect.

If embedding matters, then the sum of the relative weights in T1 should exceed the relative weighting of *Jobs* in C1. We therefore postulate the following hypothesis:

$$\text{Hypothesis 1a: } RW_{C1}(Jobs) < RW_{T1}(JQ) + RW_{T1}(LM).$$

Hypothesis 1a assumes regular embedding. To investigate the stronger notion of perfect embedding, we go one step further. In our experimental design, perfect embedding implies that the maximum of the relative weights T1 participants assign to *JQ* or to *LM* exactly equals the control groups' relative weight of *Jobs*. Accordingly, we formulate

$$\text{Hypothesis 2: } RW_{C1}(Jobs) = \max(RW_{T1}(JQ); RW_{T1}(LM)).$$

We could have split the dimension *Jobs* into four dimensions representing one indicator each. This, however, would not fit into the concept of dimensions as overarching categories (e.g. environment, education), thus leading to a discontinuity within the weighting tool. Participants' attention might be directed especially to these single-indicator dimensions, which we wanted to rule out as alternative explanation for embedding effects.

3.4 Embedding or framing?

The first treatment varies from the first control group with respect to the dimension titles describing working life. Hence, the treatment frames the same information differently. It is therefore in principle possible that a confirmation of the first hypothesis reflects a framing effect, as described by Tversky and Kahneman (1981). For instance, if the term 'job quality' were perceived more positively than 'job security' and 'personal earnings', the *Job Quality* dimension could receive more attention and would thus be given a higher weight.

To investigate whether framing effects matter in our experiment, we introduce a second control group, C2, who has to rate the same 11 dimensions as C1 (including *Jobs*), but the mouse-over pop-ups are framed differently. C2 participants see the pop-ups 'job quality' and 'labor market', whereas C1 participants see 'unemployment', 'job security' and 'personal earnings'. The underlying indicators of *Job Quality* are personal earnings and job security, while the underlying indicators of *Labor Market* are long term unemployment rate and employment rate. If framing plays no role in our context, the relative weights assigned to *Jobs* should not differ between the two control groups, i.e.

$$\text{Hypothesis 3: } RW_{C1}(Jobs) = RW_{C2}(Jobs).$$

holds. Compared to C2, the first treatment T1 changes the embedding as *Job Quality* and *Labor Market* become dimensions, but it does not change the labelling of aspects of working life. Hence, we expect regular embedding also shows up when we compare C2 and T1:

$$\text{Hypothesis 1b: } RW_{C2}(Jobs) < RW_{T1}(JQ) + RW_{T1}(LM).$$

If our analysis supports either Hypothesis 1a or Hypothesis 1b but not both, it will indicate the existence of a framing effect.

3.4 The role of preconceived individual notions

People may have preconceived individual notions of a dimension title based on previous experience and pre-determined attitudes. This might strongly influence the perception of the dimension and thus its rating. If these notions are very strong, the ratings will be immune to the information provided by the OECD about what is actually rated, i.e. the underlying indicators. Any difference in the *Jobs* weight(s) from C1 and C2 and T1 would indirectly point in this direction. Once new dimension titles are used, the ratings change irrespectively of the fact the underlying indicators remain the same. Individual notions could thus induce embedding effects. To shed more light on this conjecture, we test whether dimensions that are labelled the same, and should thus recall the same individual notion, get the same relative weights although they carry distinct sets of indicators. We therefore introduce two further treatment groups, T2 and T3. As in the two control groups, both rate *Jobs* as one dimension of the original list of eleven dimensions, but the pop-ups showing up once people move the mouse over *Jobs* as well as the underlying information is fairly different. T2 only covers the job-specific aspects, namely personal earnings and job security, whereas T3 only gets to see the labor market indicators long-term unemployment rate and unemployment rate. Subjects who take the information provided into account probably value the *Jobs* dimension differently, depending on the importance they assign to the underlying items. If people only translate their preconceived notions of the term ‘jobs’ into the relative weights and do not care about any information provided, the following hypothesis holds:

$$\text{Hypothesis 4: } RW_{T2}(Jobs) = RW_{T3}(Jobs) = RW_{C1}(Jobs).$$

3.5 Summary of experimental groups and descriptive statistics

Table 1 sums up the specifications of the R-BLI weighting tool assigned to the different control and treatment groups as well as information on the five experimental groups gathered by the landing page. The randomization process produced fairly similar groups given that the group

sizes are not tremendous. The groups only differ on a statistically significant level in a very few characteristics. Comparing many characteristics across five groups inflates the likelihood of type I errors. It becomes very likely to find at least one significant difference even if by pure coincidence. We apply the Bonferroni correction to tackle this so-called multiple-comparison problem.⁵ The only significant difference emerging is that ‘economics as major’ is more likely in group T3 than in group T2. To be on the safe side, we consider in our regression analyses the characteristics described in Table 1 as control variables.

Table 1: Overview of the experimental groups

	Control group C1	Control group C2	Treatment group T1	Treatment group T2	Treatment group T3
Number of dimensions	11	11	12	11	11
Label of <i>Jobs</i>	<i>Jobs</i>	<i>Jobs</i>	<i>Job Quality; Labor Market</i>	<i>Jobs</i>	<i>Jobs</i>
Mouse-over pop-up*	Personal earnings, Job security; Unemployment	Job quality; Labor market	Personal earnings, Job security; Unemployment	Personal earnings; Job security	Long-term unemployment rate, Employment rate
Number of observations	105	119	101	96	101
Female (share)	44%	39%	56%	45%	46%
Age (in years)	22.18 (4.64)	21.71 (2.85)	22.27 (5.51)	22.23 (3.50)	22.34 (3.02)
Knowledge of the BLI (share)	21%	16%	27%	25%	20%
Time spent weighting (in minutes, median)	1:42	1:39	1:46	1:43	1:44
Accessed information (share)	26%	26%	19%	27%	27%
<i>Size of home town (shares)</i>					
20,000 or less	30%	29%	29%	18%	25%
20,000 – 100,000	22%	17%	20%	23%	17%
100,000 – 500,000	14%	20%	19%	19%	23%
500,000 – 1,000,000	10%	9%	8%	9%	13%
1,000,000 or more	25%	24%	25%	31%	23%
<i>Major (shares)</i>					
Economics	17%	18%	24%	8%	23%
Business	33%	34%	28%	39%	33%
Administration					
Mathematics	13%	13%	17%	16%	15%
Languages	8%	12%	7%	6%	11%
Arts	10%	6%	7%	14%	5%
Other	18%	17%	18%	18%	14%

*Note: Standard deviation in parentheses. *Most mouse-over pop ups display indicators. Unemployment and labor market each represent the two indicators long-term unemployment rate and employment rate. Job Quality consists of the two indicators personal earnings and job security.*

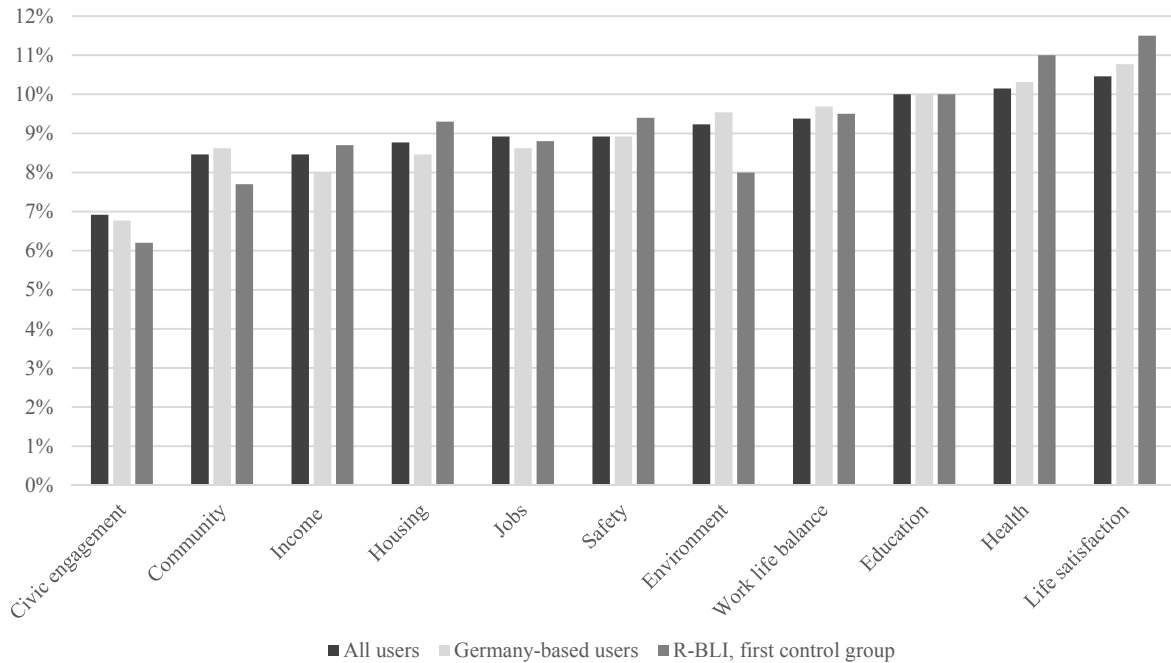
⁵ The basic idea is that the type I errors of the number of dependent tests (e.g. the tests per group comparison) add up at the maximum. Thus, the 10% level of significance must be divided by the number of tests to obtain the test specific equivalent level of 10% significance, i.e. in our case the p -value of each test must be lower than 0.1/15.

4. Experimental results

4.1 Mean analysis

We present the results of our replication of the original BLI in Figure 1, i.e. the ratings of C1. It allows us to cautiously compare the results obtained from our weighting tool to those published by the OECD to assess how successful our replication is. The relative weights of the 11 dimensions range from 6.2% (civic engagement) to 11.5% (life satisfaction). This matches the data from the worldwide ranking and the country specific ranking of Germany. There the lowest rankings are somewhat above 6% and the highest ranking somewhat above 10%. In addition, the order of these rankings is very similar to that of our first control group. In all three cases, civic engagement is ranked the least, whereas the sequence of the highest four ranks is life satisfaction, followed by health, education and work-life balance (OECD 2014). Thus, we find no evidence that participating in our experiment and using a replication of the BLI weighting tool affects the rating process in a way different from the original. Obviously, our participants have been able to provide meaningful indications. Thus, we consider the following results to generalize to the use of the original BLI.

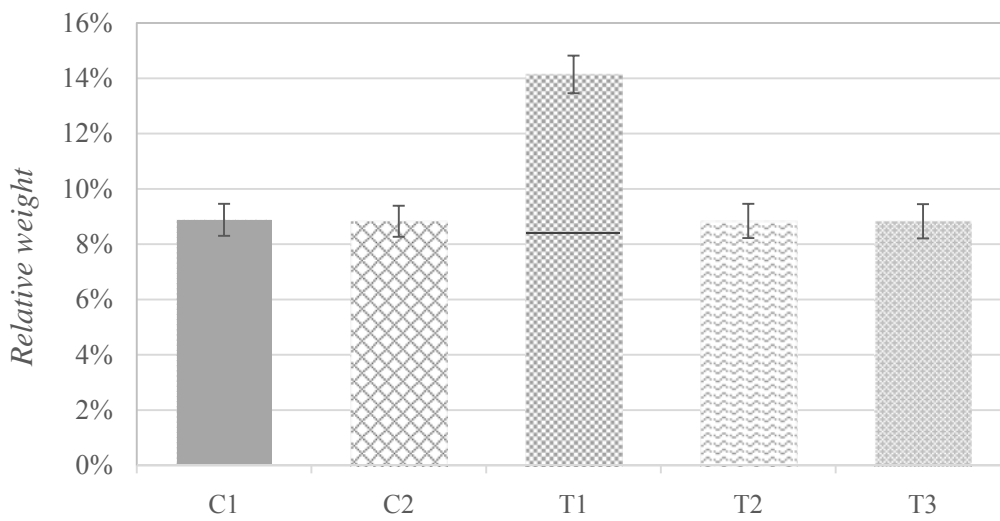
Figure 1: Dimension ratings across different user groups



Note: The bars displays average relative weights of dimensions according to ratings of all users (left bars) and Germany-based users (central bars) of the original BLI (source: OECD 2014) as well as of our control group 1 (right bars).

We start testing our hypotheses by comparing the mean relative weights assigned to the *Jobs* dimension and its treatment 1 replacements *Labor Market* and *Job Quality*. To test for statistical significance, we always use the parametric t-test and the non-parametric Mann-Whitney-U test. The t-test is based on the assumption of normality, which is confirmed by the Shapiro-Wilk test.⁶ We also assess the variance of the relative weights across all of the groups. Only when comparing T1 and C1 is variance homogeneity not supported. Therefore, the Welch's adjustment for the t-test is employed.

Figure 2: Dimension ratings across different user groups



Note: For C1 and C2 as well as T2 and T3 the bars display average relative weights of the *Jobs* dimension. The treatment group 1 bar illustrates the sum of average relative weights of the dimensions *Job Quality* (below the black line) and *Labor Market* (above the black line).

C1 as well as T1 rate the same embedded indicators, made transparent by mouse-over pop ups. As Figure 2 illustrates, the treated assign a substantially and significantly ($p \approx 0.0000$) higher relative weight to *Labor Market* and *Job Quality* in sum (0.141) than the controls attribute to *Jobs* (0.088). This result strongly supports Hypothesis 1a, implying at least regular embedding. As group T1's relative weight of *Job Quality* (0.086) does not differ significantly from group C1's relative weight of *Jobs*, perfect embedding cannot be ruled out, in line with Hypothesis 2.

Our results would reflect a framing effect rather than an embedding effect if treated subjects perceived the labels 'job quality' and 'labor market' differently from 'personal earnings', 'job security' or 'unemployment' and therefore assigned higher weights to the new dimensions than the controls C1 assign to *Jobs*. However, C2 that can see 'job quality' and 'labor market' as

⁶ Since each group sample is considered to be a large sample ($n > 30$), one can rest on the central limit theorem which predicts a normal sampling distribution.

mouse-over pop-ups does not rate *Jobs* higher than the first control group based on the originally framed mouse-over pop ups. Hence, the framing Hypothesis 3 is rejected. In addition, all of the effects emerging from the comparison of groups T1 and C1 also appear when comparing T1 and C2, supporting Hypotheses 1b. The same new labels of the indicators receive a higher relative weight when rated as dimensions than when embedded in *Jobs*, suggesting the expected embedding effect.

Figure 2 also allows us to investigate the role of preconceived notions of dimension titles as an explanation for our results. Subjects may update their perception of a dimension either only partly or even not at all by the information they receive from the BLI's website about underlying indicators. T3 and T4 vary the set of information provided by our R-BLI website. While C1 receives the full set of information of the BLI, T2 and T3 have to rate distinct subsets, which add up to the full set (see Subsection 3.5). As Figure 2 illustrates, the relative weights of the *Jobs* dimension do not vary across groups C1, T2 and T3. Hypothesis 4 cannot be rejected. The information provided by the website is largely ignored and subsets of indicators receive the same value as the complete set. This result is striking as nevertheless a fraction of 26%-28% of people in groups C1, T2 and T3 have accessed the detailed information about the underlying indicators (see Table 1).

4.2 Regression analysis

To consider demographic and further information gathered by the landing page in the identification of embedding effects (see Table 1), we conduct ordinary least square (OLS) regression analyses. The conditioning variables will help us to better approach the true treatment effects compared to the mean analysis if potential differences in group compositions matter for our results. In a further step, the additional data are used to identify subgroups that drive the results in particular. The relative weight a participant i assigns to the *Jobs* dimension $RW_i(Jobs)$ is our dependent variable. For T1, this variable adds up the relative weights of *Labor Market* and *Job Quality*. Binary independent variables indicate the three treatment groups and the second control group (T1, T2, T3 / C2), while C1 serves as reference category. We include various individual characteristics, namely gender (*FEM*), age (*AGE*), the size of home town (vector *SIZE* of binary variables indicating different classes) and people's major (vector *MAJOR* of binary variables indicating different fields of study). In addition, variables related to the weighting exercise are controlled for: previous knowledge of the BLI (*KNOW*), time spent weighting (binary variable *LONG* = 1 if above-median length of 1:42 minutes) and having accessed extra information while weighting (*INFO*). Altogether, this yields the following

econometric model with α as the relative weight of *Jobs* of the reference group and ε as individual error term:

$$\begin{aligned} RW_i(Jobs) = & \alpha + \beta_1 T_{1,i} + \beta_2 T_{2,i} + \beta_3 T_{3,i} + \gamma C_1 \\ & + \delta FEM_i + \phi AGE_i + \lambda SIZE_i + \mu MAJOR_i \\ & + \theta KNOWS_i + \rho TIME_i + \sigma INFO_i + \tau LONG_i + \varepsilon_i \end{aligned}$$

Table 2 summarizes the results of the regression analysis. The explanatory variables are added stepwise (I: group variables, II: individual characteristics, III: weighting variables). The results closely resemble those of the previous mean analysis in Section 4.1. Once *Jobs* is split into *Labor Market* and *Job Quality*, the sum of the relative weights of these two dimensions exceeds the relative weight of *Jobs* for the first control group, implying a regular embedding effect (columns I–III.1). In addition, the *Jobs* weights of groups C2 and C1 do not vary significantly. At the same time, a post-estimation test reveals the *Jobs* weight of C2 to be significantly lower than the sum of the relative weights of *Labor Market* and *Job Quality* of T1. This again confirms that it is not the particular framing of the terms ‘labor market’ and ‘job quality’ that drives the treatment effect.

Subjects who took some time to weight the dimensions and/or accessed the extra information about the underlying indicators might be less prone to the regular embedding effect as they set their wits to the indicators. As Table 3 reveals, however, the effect of T1 shows up robustly across these participants as well as across a variety of further subsamples. It seems to vary neither in size nor in statistical significance. Post-estimation tests confirm that the regular embedding effect also emerges across subsamples when comparing T1 to C2 where a potential framing effect is ruled out.

Our regression analyses presented in Table 2 also provide evidence for perfect embedding. For the estimation underlying column III.2, we have redefined the dependent variable for T1 as the relative weight assigned to *Job Quality* only. The coefficient does not differ significantly from zero and thus implies that *Job Quality* does not receive a weight by the treated that varies from the controls’ rating of *Jobs*. Again, a subdimension gets the same value as the whole dimension. We repeat this procedure for *Labor Market* (column III.3), which is assigned a significantly lower weight by group T1 compared to that the *Jobs* dimension receives from group C1. These results also hold across all of the subsamples.

Table 2: Results of multiple regression analyses (OLS)

	I	II	III.1	III.2	III.3
<i>Experimental groups (ref. Control group 1)</i>					
Control group 2	0.000 (0.004)	0.000 (0.004)	0.000 (0.004)	0.001 (0.004)	-0.000 (0.004)
Treatment group 1	0.053*** (0.005)	0.053*** (0.005)	0.054*** (0.005)	-0.002 (0.004)	-0.032*** (0.004)
Treatment group 2	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)	-0.005 (0.004)
Treatment group 3	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)
Female		-0.002 (0.003)	-0.001 (0.003)	-0.003 (0.003)	-0.001 (0.003)
Age (ref. 21 years)		0.000 (0.000)	0.000 (0.000)	0.001** (0.000)	-0.000 (0.000)
<i>Size of home town (ref. 100,000 - 500,000)</i>					
below 20,000		-0.003 (0.004)	-0.002 (0.004)	-0.001 (0.004)	-0.005 (0.004)
20,000-100,000		-0.001 (0.005)	-0.001 (0.005)	0.001 (0.004)	-0.003 (0.004)
500,000-1,000,000		0.003 (0.005)	0.004 (0.005)	0.004 (0.004)	0.001 (0.005)
1,000,000+		-0.005 (0.004)	-0.005 (0.004)	-0.004 (0.004)	-0.006 (0.004)
<i>Major (ref. Economics)</i>					
Business Administration		0.002 (0.005)	0.002 (0.005)	0.002 (0.004)	0.001 (0.004)
Mathematics		-0.003 (0.006)	-0.004 (0.006)	-0.002 (0.005)	-0.005 (0.005)
Languages		-0.007 (0.005)	-0.007 (0.006)	-0.007 (0.005)	-0.009 (0.005)
Arts		0.003 (0.006)	0.003 (0.006)	0.000 (0.005)	0.002 (0.005)
Other		-0.000 (0.005)	-0.000 (0.006)	-0.001 (0.005)	-0.001 (0.005)
Knows the OECD Better Life Index			-0.000 (0.004)	-0.000 (0.003)	-0.002 (0.003)
Above-median time spent weighting			-0.005 (0.003)	-0.004 (0.003)	-0.004 (0.003)
Accessed information			0.004 (0.004)	0.004 (0.004)	0.006 (0.003)
Constant	0.088*** (0.003)	0.081*** (0.010)	0.081*** (0.010)	0.071*** (0.012)	0.099*** (0.011)
Observations	522	522	522	522	522
R ²	0.313	0.325	0.329	0.037	0.161

*Note: The table presents OLS estimates of the relative weight of the Jobs dimension (T1: the sum of the relative weights of the Labor Market dimension and the Job Quality dimension). Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.*

At the close of this section, we return to the first three columns of Table 2 to shed light on the role of preconceived notions. The coefficients of the second and third treatment show that the two distinct subsets of the *Jobs* dimension receive the same weights as C1 attributes to the

whole dimension. Again, we ascertain that people do not update their own perception of what the term *Jobs* means to them by the specific indicators to be weighted. At least in this case, one would expect the results to change for people who accessed the extra information or took a relatively long time weighting, resulting in lower weights for the subsets of indicators compared to the full set. Table 3 reveals a converse pattern, if any. Subjects who forwent accessing additional information or rated relatively quickly tend to assign slightly lower weights to the *Jobs* dimension if it represents long-term unemployment rate and employment rate only (3rd treatment) than if it covers all of the indicators. The same applies to men and people coming from relatively large cities.

Table 3: Subgroup analyses, dependent on individual and weighting characteristics

	<i>initial sample</i>	<i>female</i>	<i>male</i>	<i>age below 21 years</i>	<i>age above 21 years</i>	<i>small town</i>	<i>large town</i>	<i>major econ./bus adm. or business</i>	<i>major not econ./bus. adm.</i>
<i>Experimental groups (ref. Control group 1)</i>									
Control group 2	0.000 (0.004)	0.003 (0.006)	-0.001 (0.006)	0.003 (0.005)	-0.003 (0.007)	-0.004 (0.007)	0.002 (0.005)	0.002 (0.006)	-0.004 (0.006)
Treatment group 1	0.054*** (0.005)	0.059*** (0.006)	0.047*** (0.007)	0.056*** (0.006)	0.049*** (0.007)	0.056*** (0.007)	0.051*** (0.006)	0.051*** (0.007)	0.054*** (0.007)
Treatment group 2	-0.005 (0.004)	-0.009 (0.006)	-0.001 (0.006)	-0.005 (0.006)	-0.007 (0.007)	0.002 (0.007)	-0.012** (0.006)	-0.005 (0.007)	-0.005 (0.006)
Treatment group 3	-0.004 (0.004)	0.007 (0.006)	-0.012* (0.006)	-0.004 (0.006)	-0.007 (0.007)	0.002 (0.007)	-0.010* (0.005)	-0.003 (0.006)	-0.009 (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*** (0.010)	0.083*** (0.015)	0.080*** (0.013)	0.093*** (0.009)	0.094*** (0.008)	0.054*** (0.021)	0.089*** (0.010)	0.070*** (0.012)	0.097*** (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

	<i>initial sample</i>	<i>knows BLI</i>	<i>does not know BLI</i>	<i>short time spent</i>	<i>long time spent</i>	<i>read extra info</i>	<i>did not read extra info</i>
<i>Experimental groups (ref. Control group 1)</i>							
Control group 2	0.000 (0.004)	0.009 (0.010)	-0.003 (0.005)	-0.006 (0.006)	0.003 (0.006)	0.013 (0.008)	-0.005 (0.005)
Treatment group 1	0.054*** (0.005)	0.064*** (0.010)	0.050*** (0.005)	0.050*** (0.007)	0.056*** (0.006)	0.049*** (0.011)	0.054*** (0.005)
Treatment group 2	-0.005 (0.004)	0.003 (0.008)	-0.009* (0.005)	-0.007 (0.006)	-0.006 (0.007)	-0.001 (0.010)	-0.006 (0.005)
Treatment group 3	-0.004 (0.004)	0.003 (0.011)	-0.006 (0.005)	-0.015** (0.007)	0.004 (0.006)	0.014 (0.010)	-0.011** (0.005)
Individual characteristics	yes	yes	yes	yes	yes	yes	yes
Weighting characteristics	yes	yes	yes	yes	yes	yes	yes
Constant	0.081*** (0.010)	0.096*** (0.016)	0.080*** (0.013)	0.082*** (0.012)	0.072*** (0.017)	0.062*** (0.023)	0.087*** (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

*Note: The table presents OLS estimates of the relative weight of the Jobs dimension (treatment 1: the sum of the relative weights of the Labor Market dimension and the Job Quality dimension) across various subgroups. Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.*

5. Discussion

Our experimental results show strong embedding effects in the BLI and identify a likely reason for this. Individual preconceived notions of the dimensions seem to decisively affect subjects' ratings, whereas the underlying indicators are largely ignored. Further reasons might play a part in explaining our findings as well. The embedding effect could also originate from a tendency to assign average relative weights to all dimensions, since people might not spend enough time and effort during the weighting process. This refers to two problems that generally serve as explanations for embedding effects, namely that people complete questionnaires on the fly and that they are not informed enough about the implications of their ratings. However, we consider these two reasons less important in our context as subgroups of participants who spent a relatively long time weighting and those who accessed the additional information provided by the weighting tool show the same embedding effects as the whole samples. Furthermore, the first treatment group assigns the lowest weight to *Labor Market*, which is far from the average weight. Thus, at least a large group of subjects not only assign average weights on the fly, but also think their preferences through while weighting. In sum, preconceived notions that are recalled by the dimension titles seem to be the main source of embedding effects in the BLI and providing more information about the indicators will probably not solve the problem.

One might think of one simple way to avoid embedding effects in the BLI: People should directly weight the 24 indicators that currently underlie the dimensions. The embedding phenomenon, however, may persist. For instance, personal earnings as one of the indicators can include fixed wage rate plus performance bonus, end-of-the-year bonus, fringe benefits, and many more aspects. Indicators always embed subcategories such that the weighting process remains prone to embedding effects. An extensive list of very detailed indicators might circumvent embedding effects, but would probably overwhelm participants. Even in this case, embedding effects cannot be ruled out as long as not all indicators that at least some participants consider as relevant are included.

Our results yield various implications. Most importantly, the results cast serious doubts that the BLI will provide reliable measures of the citizens' true preferences. As the *Jobs* dimension receives a much lower weight in the BLI than two similar dimension names containing the same information, we cannot infer the true relative weight work-related aspects of life should have in the social welfare function. There is no reason why this result should only concern the *Jobs* dimension. In fact, it will certainly apply to all of the other dimensions as well. Hence, the weights assigned to the dimensions should not be interpreted in any respect, and any

comparison of dimension weights, such as the statement ‘education is more important to people than income’, is not based on reliable data or analysis. Obviously, this makes the BLI invalid for at least two types of implications. First, it does not meet its target of providing societies with a credible order of aspects of quality of life which aggregates citizens’ preferences. Second, it does not allow us to assess a policy that redistributes scarce resources from one aspect of life to another.


Often, measurement problems of surveys can be solved by only relying on variation over time. Though we cannot interpret the weight of *Jobs* or that of any other dimension, an increase of a certain weight from one year to the next might imply that this specific aspect got more important to the participants. However, in the specific case of the BLI, the data provided are not quite informative as long as we do not know what people actually associate with the specific dimension. Which policy will take account of a growing importance of, for example, the *Jobs* dimension? A policy that leads to higher wages at the cost of higher unemployment or just the opposite measure? A dimension title is only a label that respondents ultimately interpret for themselves. Thus, policy makers are left with no better understanding of what exactly people consider worthwhile or which policy implications such preferences could have. The sad lesson from our experiment is therefore that the OECD Better Life Index does not accelerate the ongoing search for a valid measure of social welfare.

References

- Boarini, Romina and Marco Mira D'Ercole (2013): "Going beyond GDP: An OECD Perspective", *Fiscal Studies* 34(3), pp. 289–314.
- Carson, Richard T. and W. Michael Hanemann (2005): "Chapter 17 Contingent Valuation", *Handbook of Environmental Economics* 2(5), pp. 821–936.
- Clark, Andrew E. (2001): "What really matters in a job? Hedonic measurement using quit data", *Labour Economics* 8(2), pp. 223–242.
- Coyle, Diane (2014): *GDP. A Brief but Affectionate History*, Princeton University Press: Princeton and Oxford.
- Diamond, Peter A. and Jerry A. Hausman (1994): "Contingent Valuation: Is Some Number Better than No Number?", *Journal of Economic Perspectives* 8(4), pp. 45–64.
- Durand, Martine (2015): "The OECD better life initiative: How's life? And the measurement of well-being", *Review of Income and Wealth* 61(1), pp. 4–17.
- Hanemann, W. Michael (1994): "Valuing the Environment Through Contingent Valuation", *Journal of Economic Perspectives* 8(4), pp. 19–43.
- Hausman, Jerry (2012): "Contingent Valuation: From Dubious to Hopeless", *Journal of Economic Perspectives* 26(4), pp. 43–56.
- Kahneman, Daniel and Jack L. Knetsch (1992a): "Valuing Public-Goods - the Purchase of Moral Satisfaction", *Journal of Environmental Economics and Management* 22(1), pp. 57–70.
- Kahneman, Daniel and Jack L. Knetsch (1992b): "Contingent Valuation and the Value of Public Goods: Reply", *Journal of Environmental Economics and Management* 22(1), pp. 90–94.
- Kasparian, Jérôme and Antoine Rolland (2012): "OECD's 'Better Life Index': can any country be well ranked?", *Journal of Applied Statistics* 39(10), pp. 2223–2230.
- Kling, Catherine L., Daniel J. Phaneuf, and Jinhua Zhao (2012): "From Exxon to BP: Has Some Number Become Better than No Number?", *Journal of Economic Perspectives* 26(4), pp. 3–26.
- Loomis, John (1993): "Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection", *Journal of Environmental Economics and Management* 24, pp. 45–55.
- McDaniels, Timothy L., Robin Gregory, Joseph Arvai, and Ratana Chuenpagdee (2003): "Decision structuring to alleviate embedding in environmental valuation", *Ecological Economics* 46(1), pp. 33–46.
- Mizobuchi, Hideyuki (2014): "Measuring World Better Life Frontier: A Composite Indicator for OECD Better Life Index", *Social Indicators Research* 118(3), pp. 987–1007.
- OECD (2011): *How's Life? Measuring Well-Being*, OECD Publishing: Paris.
- OECD (2014): *OECD Better Life Index Country Reports*, OECD Publishing: Paris.
- OECD (2015): *How's Life? 2015: Measuring Well-Being*, OECD Publishing: Paris.
- Smith, V. Kerry (1992): "Arbitrary Values, Good Causes, and Premature Verdicts", *Journal of Environmental Economics and Management* 22(1), pp. 71–89.
- Stiglitz, Joseph E., Amartya K. Sen, and Jean-Paul Fitoussi (2009): *Report by the Commission on the Measurement of Economic Performance and Social Progress*, Unpublished Work .

-
- Svedsäter, Henrik (2000): "Contingent valuation of global environmental resources: Test of perfect and regular embedding", *Journal of Economic Psychology* 21, pp. 605–623.
- Tversky, Amos and Daniel Kahneman (1981): "The framing of decisions and the psychology of choice", *Science* 211(4481), pp. 453–458.
- WIFO (2012): *Mehr als Wachstum Messung von Wohlstand und Lebensqualität in ausgewählten OECD-Ländern Ausgangslage und Aufgabenstellung*, mimeo.

Appendix 1 Recruitment flyer


 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 Number"

This ticket is valid only once. Your answers will remain completely anonymous. 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@fu-berlin.de
 Chair of International Public Economics: LS-Schoeb@fu-berlin.de

Appendix 2: Webpages of OECD BLI and R-BLI

Landing Page of the OECD BLI

OECD Better Life Index

Index Responses Countries Topics FAQ

Create Your Better Life Index

What is your recipe for a better life – a good education, clean air, nice home, money?
See how your country measures up on the topics important to you. [Help](#)

Start with all topics rated equally or set your own preferences [here](#).

How's life?

There is more to life than the cold numbers of GDP and economic statistics – This Index allows you to compare well-being across countries, based on 11 topics the OECD has identified as essential, in the areas of material living conditions and quality of life.

[Download executive summary](#)
[Download the index data](#)
[Learn more about the Better Life Initiative](#)

Mapping well-being

What matters most to people around the world?

More than 100,000 users of the Better Life Index around the world have shared their views on what makes for a better life. [Explore the map](#)

Find out which topics are most important to people around the world.

Create Your Better Life Index

Rate the topics according to their importance to you:

- Housing
- Income
- Jobs
- Community
- Education
- Environment
- Civic Engagement
- Health
- Life Satisfaction
- Safety
- Work-Life Balance

[Gender differences](#)
[Compare with others](#)
[Start your index](#)

Source: www.oecdbetterlifindex.org

Landing Page of the R-BLI

Better Life Index

There is more to life than the cold numbers of GDP and economic statistics. This Index allows you to compare societal well-being based on various topics identified as essential.

Participate

Ticket N°

Year of birth

Please choose

Gender

☐ Male

☐ Female

Size of hometown

Please choose

Major of study

Please choose

[Participate](#)

R-BLI experimental page (default setting)

Create your Better Life Index!

Rate the topics according to their importance to you.

[Click here to obtain further information about the specific topics](#)



Housing

Income

Jobs

Community

Education

Environment

Civic Engagement

Health

Life Satisfaction

Safety

Work-Life-Balance

Housing

Income

Jobs

Community

Education

Environment

Civic Engagement

Health

Life Satisfaction

Safety

Work-Life-Balance

I am familiar with the OECD Better Life Index

Please choose

Participate

R-BLI experimental page with exemplified weighting

Create your Better Life Index!

Rate the topics according to their importance to you.

[Click here to obtain further information about the specific topics](#)



Housing

Income

Jobs

Community

Education

Environment

Civic Engagement

Health

Life Satisfaction

Safety

Work-Life-Balance

Housing

Income

Jobs

Community

Education

Environment

Civic Engagement

Health

Life Satisfaction

Safety

Work-Life-Balance

I am familiar with the OECD Better Life Index

Please choose

Participate

Supplementary Material

Information provided on the R-BLI web-application

The information provided below was extracted from the OECD BLI web-application and used in the experimental set-up for further informational purposes. Dimensions are in bold letters, underlying indicators are in italics.

1 Information on the dimension jobs, C1

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

2 Information on the dimension jobs, C2

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

‘Job quality’

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing

earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

‘Labour market’

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

3 Information on the dimensions *job quality* and *labour market*, T1

Job quality

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

Labour market

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

4 Information on the dimension jobs, T2

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Average earnings

The wages and other monetary benefits that come with employment are an important aspect of job quality. Earnings represent the main source of income for most households. Analysing earnings may also suggest how fairly work is remunerated. In the OECD on average, people earn USD 36,118 per year, but average earnings differ significantly across OECD countries.

Job security

Another essential factor of employment quality is job security. Workers facing a high risk of job loss are more vulnerable, especially in countries with smaller social safety nets.

5 Information on the dimension jobs, T3

Jobs

Work has obvious economic benefits, but having a job also helps individuals stay connected with society, build self-esteem, and develop skills and competencies. Societies with high levels of employment are also richer, more politically stable and healthier.

Employment rate

Across the OECD, about 65% of the working-age population aged 15 to 64 has a paid job. Employment rates are generally higher for individuals with a higher level of education.

Long-term unemployment rate

Unemployed persons are defined as those who are currently not working but are willing to do so and actively searching for work. Long-term unemployment can have a large negative effect on feelings of well-being and self-worth, and result in a loss of skills, further reducing employability. Such effects can last a long time, even after a return to work.

A3.6 Information on the other dimensions for all experimental groups

Housing

Living in satisfactory housing conditions is one of the most important aspects of people's lives. Housing is essential to meet basic needs, such as shelter, but it is not just a question of four walls and a roof. Housing should offer a place to sleep and rest where people feel safe and have privacy and personal space; somewhere they can raise a family. All of these elements help make a house a home. And of course there is the question whether people can afford adequate housing.

Housing expenditure

Housing costs take up a large share of the household budget and represent the largest single expenditure for many individuals and families, by the time you add up elements such as rent,

gas, electricity, water, furniture and repairs. In the OECD, households on average spend around 18% of their gross adjusted disposable income on keeping a roof over their heads.

Rooms per person and dwellings with basic facilities

When looking at housing, it is important to examine living conditions, such as the average number of rooms shared per person and whether dwellings have access to basic facilities.

The number of rooms in a dwelling, divided by the number of persons living there, indicates whether residents are living in crowded conditions. Overcrowded housing may have a negative impact on physical and mental health, relations with others, and children's development.

Income

While money may not buy happiness, it is an important means to achieving higher living standards and thus greater well-being. Higher economic wealth may also improve access to quality education, health care and housing.

Household net-adjusted disposable income

Household net-adjusted disposable income is the amount of money that a household earns, or gains, each year after taxes and transfers. It represents the money available to a household for spending on goods or services.

Household financial wealth

Household financial wealth is the total value of a household's financial worth, or the sum of their overall financial assets minus liabilities. Financial wealth takes into account: savings, monetary gold, currency and deposits, stocks, securities and loans.

Community

Humans are social creatures. The frequency of our contact with others and the quality of our personal relationships are thus crucial determinants of our well-being. Studies show that time spent with friends is associated with a higher average level of positive feelings and a lower average level of negative feelings than time spent in other ways.

Social support network

A strong social network, or community, can provide emotional support during both good and bad times as well as access to jobs, services and other material opportunities. Across the OECD, 88% of people believe that they know someone they could rely on in time of need.

Education

Education plays a key role in providing individuals with the knowledge, skills and competences needed to participate effectively in society and in the economy. In addition, education may improve people's lives in such areas as health, civic participation, political interest and happiness. Studies show that educated individuals live longer, participate more actively in politics and in the community where they live, commit fewer crimes and rely less on social assistance.

Years in education

In a fast-changing knowledge economy, education is about learning skills for life. But how many years of school, college, or training will future generations expect to have?

Educational attainment

Having a good education greatly improves the likelihood of finding a job and earning enough money. Highly-educated individuals are less affected by unemployment trends, typically because educational attainment makes an individual more attractive in the workforce. Lifetime earnings also increase with each level of education attained.

Environment

The quality of our local living environment has a direct impact on our health and well-being. An unspoiled environment is a source of satisfaction, improves mental well-being, allows people to recover from the stress of everyday life and to perform physical activity. Having access to green spaces for example, is an essential part of quality of life. Also, our economies rely not only on healthy and productive workers but also on natural resources such as water, timber, fisheries, plants and crops. Protecting our environment and natural resources therefore remains a long-term priority for both our generation and those to come.

Air pollution

Outdoor air pollution is one important environmental issue that directly affects the quality of people's lives. Despite national and international interventions and decreases in major pollutant emissions, globally the health impacts of urban air pollution continue to worsen, with air pollution set to become the top environmental cause of premature mortality by 2050.

Water quality

Access to clean water is fundamental to human well-being. Managing water to meet that need is a major – and growing – challenge in many parts of the world. Many people are suffering from inadequate quantity and quality of water. Despite significant progress in OECD countries in reducing water pollution, from fixed sources such as industrial and municipal wastewater treatment plants, diffuse pollution from agriculture and urban run-offs remains a challenge, and improvements in freshwater quality are not always easy to discern.

Civic engagement

Trust in government is essential for social cohesion and well-being. Today, more than ever, citizens demand greater transparency from their governments. Information on the who, why and how of decision making is essential to hold government to account, maintain confidence in public institutions and support a level playing field for business. Greater transparency is not only key to upholding integrity in the public sector; it also contributes to better governance. Indeed, openness and transparency can ultimately improve public services by minimising the risk of fraud, corruption and mismanagement of public funds.

Voter turnout

High voter turnout is a measure of citizens' participation in the political process. Voter turnout is defined as the percentage of the registered population that voted during an election. High voter turnout is desirable in a democracy because it increases the chance that the political system reflects the will of a large number of individuals, and that the government enjoys a high degree of legitimacy.

Consultation on rule-making

Transparency in communication and open access to regulations promotes government accountability, a business-friendly environment and public trust in government institutions.

Indeed, a country's regulations contain much information about how a society is organised, the rules of the game and the political decisions taken. If citizens can readily access and understand regulations, it is more likely that they will participate in the legislative process and comply with the rules.

Health

Good health is one of the most important things to people and also brings many other benefits, including enhanced access to education and the job market, an increase in productivity and wealth, reduced health care costs, good social relations, and of course, a longer life.

Life expectancy

Life expectancy is the most widely used measure of health, although it only takes into account the length of people's life and not their quality of life. There have been remarkable gains in life expectancy over the past 50 years in OECD countries.

Self-reported health

Most OECD countries conduct regular health surveys which allow respondents to report on different aspects of their health. The commonly-asked question, "How is your health?" is one way of collecting data on self-perceived health status. Despite the subjective nature of this question, the answers received have been found to be a good predictor of people's future health care use.

Life satisfaction

Measuring feelings can be very subjective, but is nonetheless a useful complement to more objective data when comparing quality of life across countries. Subjective data can provide a personal evaluation of an individual's health, education, income, personal fulfilment and social conditions. Surveys, in particular, are used to measure life satisfaction and happiness.

Life satisfaction

Life satisfaction measures how people evaluate their life as a whole rather than their current feelings.

Safety

Personal security is a core element for the well-being of individuals, and includes the risks of people being physically assaulted or falling victim to other types of crime. Crime may lead to loss of life and property, as well as physical pain, post-traumatic stress and anxiety. One of the biggest impacts of crime on people's well-being appears to be through the feeling of vulnerability that it causes.

Assault rate

According to recent data, 3.9% of people in OECD countries say they have been assaulted or mugged over the past 12 months.

Homicide rate

Homicide rates (the number of murders per 100,000 inhabitants) only represent the most extreme form of contact crime and thus do not provide information about more typical safety conditions. They are however a more reliable measure of a country's safety level because, unlike other crimes, murders are usually always reported to the police.

Work-life Balance

Finding a suitable balance between work and daily living is a challenge that all workers face. Families are particularly affected. The ability to successfully combine work, family commitments and personal life is important for the well-being of all members in a household. Governments can help to address the issue by encouraging supportive and flexible working practices, making it easier for parents to strike a better balance between work and home life.

Employees working long hours

An important aspect of work-life balance is the amount of time a person spends at work. Evidence suggests that long work hours may impair personal health, jeopardise safety and increase stress.

Time devoted to leisure and personal care

Furthermore, the more people work, the less time they have to spend on other activities, such as personal care or leisure. The amount and quality of leisure time is important for people's overall well-being, and can bring additional physical and mental health benefits.