Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR WIRTSCHAFTS- UND
SOZIALWISSENSCHAFTEN

For “better” or “worse”: a direct approach to elicit preference rankings from life-satisfaction data

Sonja Köke
Grischa Perino

WiSo-HH Working Paper Series
Working Paper No. 43
August 2017

WiSo-HH Working Paper Series
Working Paper No. 43
August 2017

For “better” or “worse”: a direct approach to elicit preference rankings from life-satisfaction data

Sonja Köke, Universität Hamburg
Grischa Perino, Universität Hamburg

ISSN 2196-8128

Font used: „TheSans UHH“ / LucasFonts

Kontakt:
WiSo-Forschungslabor
Von-Melle-Park 5
20146 Hamburg
E-Mail: experiments@wiso.uni-hamburg.de
Web: http://www.wiso.uni-hamburg.de/forschung/forschungslabor/home/

For "better" or "worse": a direct approach to elicit preference rankings from life-satisfaction data

Sonja Köke[1,2], Grischa Perino[1,*]

Affiliations:

[1]Department of Socioeconomics, University of Hamburg, Welckerstr. 8, 20354 Hamburg, Germany.

[2]Department of Economics, Christian-Albrechts-University Kiel, Wilhelm-Seelig-Platz 1, 24118 Kiel

*Correspondence to: grischa.perino@wiso.uni-hamburg.de

Abstract

We provide empirical evidence that intra-personal changes in life-satisfaction are a biased measure of preference rankings and, as a remedy, propose the *ranking measure* where subjects state whether their life has become "better" or "worse". Three representative datasets reveal: intra-personal changes in satisfaction levels are dominated by noise and less well explained by socio-economic variables than the ranking measure; the deviation between the two measures is systematic and adaptation (e.g. to income and unemployment) is only observed for changes in satisfaction levels but not for rankings, indicating that adaptation is driven by the elicitation method rather than changes in preferences.

Keywords: Life satisfaction, preference rankings, adaptation

JEL codes: I31, C83

1. Introduction

The definition of social welfare has occupied philosophers and social scientists for centuries if not millennia. On the measurement side, differences in schools of thought are compounded with limited availability of data and challenges in eliciting the specific object of interest. Purely monetary measures such as the highly influential but strongly disputed Gross Domestic Product compete with "objectively" measurable indicators (Human Development Index, Jones and Klenow 2016 or the capabilities approach, Sen 1985) and subjective measures of wellbeing such as happiness or life satisfaction (Frey and Stutzer 2002, Benjamin et al. 2014b).

A prominent example that objective and subjective welfare measures might produce vastly different assessments over time is the Easterlin Paradox. While income and subjective well-being tend to be positively correlated in the short run, average life satisfaction (LS) has been almost constant in industrialized countries over the past decades despite a substantial increase in per capita GDP (Easterlin 1974, Easterlin et al. 2010, Stevenson and Wolfers 2008, 2013). Several explanations for this discrepancy between measures have been proposed. A leading one is that individuals get used to, i.e. adapt, to new (especially higher) levels of income (Frederick and Loewenstein 1999, Diener et al. 2006, Loewenstein and Ubel 2008, diTella et al. 2010). The literature so far has not addressed the question of whether adaptation of reported levels of LS reflect changes in preferences or are (merely) a re-scaling of the physiological and psychological sensitivity to external stimuli. While preference rankings over states of the world can plausibly include a very large number of strictly separate ranks, our bodies and minds face limitations in both range and step size when attributing satisfaction levels to states of the world (Rayo and Becker 2007a,b). The latter extends – in an even stronger form - to the answer scales used in surveys to elicit levels of LS where participants are periodically asked to state their LS on a discrete scale from "completely dissatisfied" to "completely satisfied", spanning three to

eleven categories.[1] We label this measure of LS the level measure (LM). There is evidence that responses correspond well both to neurological measures and other indicators of well-being such as frequency of genuine smiles, spousal fights and suicide (Berridge and Kringelbach 2011, Camerer et al. 2005, Fehr and Rangel 2011). It hence plausibly captures how intensely satisfaction is perceived as a sensation.

Differences in reported levels of life satisfaction (LS) across individuals or time have been used to identify the socio-economic drivers of LS[2] and strong prescriptions have been proposed on the normative premise that life satisfaction should be maximized. Take the response to the Easterlin Paradox. It has led some to conclude that fostering economic growth should no longer be among governments' objectives. On the contrary, income should be more heavily taxed due to the externality a higher income imposes on other citizen (Layard 2006). The claim being that rising income levels don't make us more satisfied at the aggregate societal level but keep us captured in the "hedonic treadmill" (Diener et al. 2006).

If one regards the objective of policymaking to be the maximization of perceived satisfaction levels in the population, then the above might be effective. This position is held by representatives of what Fleurbaey (2009) calls "hedonic welfarism". However, these policies might fail at maximizing utility based on preference rankings in the tradition of liberal welfare economics[3], which is concerned with "obtaining what one wants" instead of "being satisfied" (Fleurbaey 2009). As Frederick and Loewenstein (1999) point out, the two concepts are likely to differ in many situations. When using

[1] Surveys including such a question are e.g. World Value Survey, German SOEP, UK Understanding Society, Eurobarometer, Latinobarometer, US General Social Survey, Happy Planet Index, World Happiness Index and the OECD Better Life Index.
[2] Frey and Stutzer 2002, Di Tella et al. 2003, Di Tella and MacCulloch 2006, Luttmer 2005, Frey et al. 2004, Ferrer-i-Carbonell 2005, Oswald 1997, Stevenson and Wolfrers 2013, Dolan et al. 2008.
[3] A preference ranking is an ordinal ranking of situations in which higher ranked situations are preferred over lower ranked ones.

satisfaction data to inform policy making it is thus important to know how these two concepts differ and how they can best be measured. If they differ, the recommendations will as well and researchers will need to be clear about the concept they are using.

We show that changes in LS are unreliable indicators of individuals' preference rankings over changes in states of the world as evaluated ex-post[4]. We identify conceptual reasons for them to differ and provide evidence from representative surveys indicating that basic requirements for identifying preference rankings are not met by reported levels of LS. This challenges interpretations of LS data as valid indicators of preference rankings (as formulated e.g. by Clark et al. 2008, MacKerron 2012, Frey et al. 2004, Welsch and Kühling 2009, Layard et al. 2008, Oswald and Wu 2010, Decancq et al. 2015).[5] However, a reliable subjective measure of preference rankings is desirable (Fleurbaey 2009).[6] In contrast to decision utility obtained from revealed preferences, life satisfaction refers to states of the world rather than bundles of goods and therefore captures a much broader set of policy relevant aspects such as the distribution of income and environmental quality.

We therefore develop and introduce the ranking measure (RM), which asks subjects to directly compare the present situation to a past situation, i.e. two situations they have actually experienced[7]. The RM avoids the conceptual shortcomings of the LM for measuring preference rankings. Most importantly, it does not need to cardinalize an ordinal concept when collecting

[4] Welfare economics traditionally is concerned with decision utility, i.e. ex-ante preferences, this difference in interpretation is necessary due to the elicitation method of stated as compared to revealed preferences and that LS is not elicited on hypothetical but rather on actually experienced situations.

[5] We focus on *life satisfaction* as a *cognitive* measure of subjective well-being instead of *happiness* or others that are more *affective* since preferences also are a cognitive concept.

[6] „Indeed, in spite of the aspiration treadmill, a clever use of satisfaction data, or the use of new questionnaires that would enable the respondents to express their ordinal preferences more directly than through the prism of a satisfaction level, may provide valuable information about people's preferences and values relative to the various dimensions of life…"

[7] Focusing on experienced situations distinguishes the subjective well-being approaches from stated preference approaches, which ask questions about hypothetical situations.

the data and to then make it ordinal again when analyzing it as is the case for the LM, making it prone to errors in the conversion process. Instead, it directly asks for the preference ranking. Comparing the preference rankings implied by LM and RM for three representative surveys with a total of well over a hundred thousand observations confirms that changes in LM are only weakly correlated with RM. Moreover, deviations between the two measures are systematically correlated with socio-demographic variables indicating that results based on the LM are biased representations of preference rankings. The ranking measure correlates more strongly and intuitively with changes in socio-economic aspects, while the LM has an inherent tendency towards adaptation both conceptually and empirically, even more so in case of a trend.

We extend previous work on the relationship between subjective well-being measures and preference rankings in several dimensions:

First, we are interested in the comparability of satisfaction measures within subjects across time, evaluating validity of changes over time. Other studies have analyzed validity of LS at one point in time, comparing it with measures of revealed or hypothetical choices and anticipated subjective well-being (e.g. Benjamin et al. 2012, 2014a, 2014b, Perez-Truglia 2015). The inter-temporal comparability is most important when using life satisfaction data from representative panel surveys and hence for much of the empirical work on life satisfaction (Fleurbaey and Schwandt 2016). It also seems the more natural direction of comparison, as compared to inter-personal comparison, when trying to identify preference rankings, which by definition are about intra-personal comparisons.

Second, we combine conceptual reasoning with empirical evidence drawing on three representative surveys spanning two countries, almost two decades and well beyond 100k observations. Fleurbaey (2009) has formulated a need for a stronger connection of LS measures and welfare theory. While some of

our concerns have been raised before, no suggestion has been made as to how to test them.

Third, we propose a new elicitation method, the ranking measure, fixing several of the conceptual issues arising when comparing the level measure across time to elicit preference rankings over states of the world.

Last, we provide evidence that the adaptation results that are characteristic of the established level measure might not primarily be driven by changes in preferences but rather represent adaptation in the reporting function.

The remainder of this paper is organized as follows: The next section relates the established level measure of life satisfaction to requirements for eliciting preference rankings over states of the world when using panel surveys. Section 3 presents evidence from a field experiment with a representative sample testing for some of the key assumptions necessary for the level measure to represent preference rankings. The ranking measure of life satisfaction is introduced in Section 4 and Section 5 compares the two measures based on three representative datasets. The last section concludes.

2. Preference rankings and the level measure of life satisfaction

We now analyze under which conditions LS measures are able to capture individual preference rankings[8]. Assume that individuals have complete preferences over all relevant pairs of situations. Let A and B be two such situations, where each is a vector including individual characteristics like number of children, size of the house, education but also aggregates such as GDP, its distribution, inflation, pollution or combinations of both such as own

[8] Note that we don't normatively argue that preferences *should* be the objective of policies nor that they are the only possible interpretation of life satisfaction. We positively *identify the conditions* under which the level measure of life satisfaction is able to represent the underlying preferences of respondents.

position in the income distribution of a peer group etc. In short: they represent the state of the world as relevant to the individual.

A preference ranking of an individual states for any pair of situations which one is preferred or whether it is indifferent between them. Continuous and rational preference rankings can be represented by utility functions. Utility function $u_i(.)$ attributes a higher value to A than to B if and only if individual i strictly prefers situation A over situation B. Indifference is represented by assigning the same utility level to both situations. Utility is an ordinal concept. A utility level holds no information in itself but only gains meaning by comparison to another level of the same utility function.

For a LS measure to capture preferences, a higher reported value for A than for B has to imply that the respondent strictly prefers situation A over situation B.[9] Formally, $LM_{i,t}(A_{i,t}) = f_{i,t}(u_{i,t}(A_{i,t}))$ is the reported satisfaction level of individual i for situation A at time t and $f_{i,t}$ a monotonically increasing reporting function mapping preferences onto answer categories in the survey. We introduce both a utility function and a reporting function to distinguish between differences in the preference order (represented by a different $u_{i,t}(.)$) and differences in how a given preference order is reported in a survey (represented by a different $f_{i,t}(.)$) that might occur across individuals or time. Layard et al. (2008) and Fleurbaey and Schwandt (2016) e.g. use a reporting function but allow variation only across individuals not across time.

LM questions slightly differ across surveys, but they all ask for the current level of LS and the answer scale provides a fixed number of categories represented by labels and numbers, e.g. the British Household Panel Survey (BHPS) asks "[...] how dissatisfied or satisfied are you with your life

[9] In contrast to the concept of (ex-ante) decision utility that is relevant for choices and relies on expectations about the outcome of choices frequently used in welfare economics, life satisfaction is more an ex-post evaluation of the current situation after decisions have been made. Thus, our utility function captures the latter of the two concepts.

overall?" and the answer categories range from 1 "Not satisfied at all" to 7 "Completely satisfied".[10]

Since interpersonal comparison is not possible with an ordinal concept, the only way to elicit information on preference rankings from LM values is from within-subject comparisons over time. To infer the underlying preferences over situations *A* and *B* from the observed LM, both the utility function and the reporting function need to remain unchanged over the two periods.[11]

In principle, both functions might change from one period to the next. For the moment, let's assume that the utility function remains unchanged for at least two consecutive measurements and focus on adjustments of the reporting function. Let's also assume that respondents try to report their preferences, i.e. that monotonicity of $f_{i,t}$ at any point in time is given. Two main reasons for adjusting the reporting function are: first, the answer scale has no intuitive meaning. It is not apparently clear what the condition is for reporting to be in the fifth out of seven satisfaction categories. Respondents need to come up with an ad hoc reporting function making it hard to remember the function until the next wave of the survey. They might hence use a different one each time. As long as there is no systematic change which distorts the representation of the underlying preferences this simply introduces noise to the data and in principle can be fixed by using large data sets (Bertrand and Mullianathan 2001, Frey and Stutzer 2002).

The second and more problematic reason for changes in the reporting function is that the scales of the preference concept and of the answer scale don't match. While the answer scale has only a few categories, e.g. 7 in the BHPS, a complete mapping of preference relations over all possible states of the

[10] One problem with this type of question is that it does not give a clear time horizon for evaluation. It would be better to ask for satisfaction with the current situation in order to capture preferences for the current situation.

[11] If respondents don't intend to report their preferences, the reporting function would not meet the monotonicity requirement and would also lead to a misrepresentation of preferences.

world requires as many categories as there are states that can be ranked in a strict sense. For most respondents this will be a much higher number than there are categories on the answer scale.

Respondents therefore face a trade-off when choosing a reporting function. They might either choose a steep one to report small to medium changes relevant in the present or they opt for a flat one to capture the "big picture". However, both can be achieved when using a steep reporting function but adjusting it to recent experiences. Such an adjustment is common for sensory perceptions[12] and in line with neurological limits to perceive sensations. A moving reference base for neurologically perceived LS allows an individual to capture the direction of changes in the short run even with a limited set of discrete levels of perception available (Rayo and Becker 2007a,b). With such an adjustment, the adaptation of the reporting function is not random but systematically depends on past and expected changes, which introduces biases that cannot be rectified by large numbers of observations (Bertrand and Mullianathan 2001). In the literature, adaptation of LS to new situations is often interpreted as a "getting used to", i.e. a change in $u_{i,t}(.)$ (e.g. DiTella et al. 2010). However, in general one cannot tell whether the observed pattern is due to a preference change or a reporting function shift.[13] Thus, the nature of the level measure of LS does not allow to reliably infer preference rankings

[12] When judging sensory magnitudes with an objective cardinal scale, humans build an ad-hoc reference frame that depends on the range of stimuli they are exposed to when reporting them on a categorical scale. *"Sensory magnitudes are selected for this review of biases in judgement because the stimuli can be measured on a physical scale. Judgements of the quality of life or the likeableness of people lack a precise measure of the stimuli. Thus the biases are more difficult to specify exactly."* (Poulton 1979).

[13] The possibility that reporting functions change over time has been ignored not only in much of the economics literature on life satisfaction but also in psychology. See e.g. Karney and Coombs (2000) where current subjective assessments of marital satisfaction are taken to be the accurate measure when compared with recollections ten years later although the recollection question asked neither implies that participants should use the past utility nor the past reporting function. Hence, in line with the *Reassess* treatment reported in Section 3, there is no reason to expect that recollection answers match former current assessments. One might therefore question, whether the study really identifies a memory bias or merely an adjustment in either utility or reporting function.

from changes in its level even if participants attempt to report their preferences.

The identification problem is stronger when many improvements (or deteriorations) are experienced or expected in a row, because then the limited answer scale is particularly restricting. While in dimensions such as income improvements might be expected to happen repeatedly, in others such as marital status this seems less plausible. A respondent might therefore choose a flatter reporting function or a different adaptation rule for the income dimension than for the marriage dimension such that the relative impact of both dimensions on LS cannot be compared to infer preferences. Hence, estimates of marginal rates of substitution are distorted (Benjamin et al. 2014a). For example the individual might adapt the reporting function to income changes but not to marriage since only with income the bounds of the answer scale are limiting factors. This would then distort the conclusions on the relative preferences for income and marriage and is in line with observations that income coefficients in LS regressions are often found to be relatively small (see e.g. Kopmann and Rehdanz 2013).

As a result, one might draw misleading conclusions on people's preferences. The Easterlin Paradox (Easterlin 1974, Easterlin et al. 2010) for example might not imply that people are indifferent to higher levels of (average) income in the long run. LS data might simply not reveal the preference for increasing income due to an adapting reporting function. While the problem of preference misrepresentation has been mentioned in the literature before, solutions are still rare (Bertrand and Mullianathan 2001, Frey and Stutzer 2002, Benjamin et al. 2014a, Fleurbaey 2009, Frederick and Loewenstein 1999, Loewenstein and Ubel 2008).

We want to stress that our analysis only applies to the preference relation interpretation of LS. The LM might be a good representation of the mental state, i.e. the intensity of sensation in a given situation. Intensity of sensation

might thus well have adapted to higher levels of income. But this does not mean that respondents are indifferent to increasing income. Our aim is to show why these two interpretations of LS will systematically differ. We abstain from making a normative judgment as to which of the concepts should be policy relevant.

3. Experimental evidence on the level measure

In this section we test how well participants remember past answers and the reporting function which we identified to be preconditions for measuring preference rankings. To this end, we commissioned a survey of a representative sample of the German population. The survey was conducted in two waves about three months apart by the market research company Lightspeed between June and October 2015. In wave 1 we asked a total of 2,300 participants for their age and gender, the LM (LM1) and whether they would be willing to participate in a later wave on a similar topic. This creates a situation similar to those faced by participants in large annual panels like the BHPS or GESIS. All participants agreed to be re-contacted. A total of 1,600 participants participated in the second wave equally and randomly distributed over four treatments.

The treatments differed in the type and order of LS questions in wave 2 (Table 1). In treatments *RecallExAnte* and *RecallExPost* participants were asked to recall the answer given to LM1 three months ago (LM1recall) and state whether they felt sure or unsure about the accuracy of their answer. The option "I cannot remember at all" was also available. The two treatments differed only with respect to the ordering of questions. In *RecallExAnte* LM1recall was asked first and hence participants could use it to anchor their response to the level measure (LM2) and the ranking measure (RM, see next section). This was not possible in *RecallExPost* where LM1recall was elicited

last and participants were prevented from changing their responses to previous questions.

In treatment *Reassess* participants had to give a retrospective assessment of LS at the point in time they had completed wave 1 of the survey (LM1reassess). In contrast to LM1recall there is no objectively correct answer to LM1reassess. LM1reassess might differ from LM1 because of imperfect recollection of the situation three months ago or because either the reporting function or preferences might have changed. Answering LM1reassess requires participants to be able and willing to apply the current reporting function to the situation three month ago. A precondition is that they perceive their recollection of that situation to be sufficient to do that. All questions as well as the summary statistics can be found in Appendix A1.

Table 1. Order of questions asked in Lightspeed survey. 400 participants per treatment.

Treatment	Wave 1	Wave 2			
Control	LM1		LM2	RM	
RecallExAnte	LM1	LM1recall	LM2	RM	
RecallExPost	LM1		LM2	RM	LM1recall
Reassess	LM1	LM1reassess	LM2	RM	

Treatments *RecallExAnte* and *RecallExPost* check how well participants remember answers to the level measure over three months (the typical gap between two waves of panel surveys is one year). Interpreting changes in the level measure over time as preference rankings requires that participants use their previous answer as a reference point and correctly report improvements or deteriorations accordingly. This requires that they either remember the answer given in the previous wave of the survey or that they correctly

remember both the situation they were in and the reporting function used in order to reconstruct their previous response.

Over both treatments 43.3 percent of participants stated that they could not at all recollect their answer to LM1 or gave no number (see Table 2). Out of the participants that answered LM1 and took part in the wave 2 only 20 percent correctly recalled this answer three months later. In *RecallExAnte* the rate of accurate answers was higher for those participants that stated a lower confidence (32.8 vs. 37.8 percent).[14] Thus, confidence here is a poor predictor of ability. The frequency of accurate answers is better than a random guess but the amount of noise introduced by imperfect recollection is substantial. Next, we test whether the noise dominates the signal, i.e. the information on the underlying preferences the respondent tries to transmit.

Table 2. Self-assessed and real accuracy of recollection of LM1.

LM1recall and accuracy conditional on answering LM1.

			LM1recall	
Treatment	**LM1 (#)**	**Confidence level**	**Total (#)**	**Accurate (#)**
RecallExAnte	397	Precise	67	22
		Guess	188	71
		Cannot recall/no answer	119/23	-
RecallExPost	399	Precise	27	13
		Guess	169	53
		Cannot recall/no answer	192/11	-

The direction of the observed change in LS (sign(LM2 – LM1)) only measures the preference ranking accurately if the recollection error

[14] Over both treatments combined the accuracy rate of the confident is 37.2 percent vs. 34.7 percent of the less confident.

(LM1recall – LM1) does not change the "signal" that participants try to give on their preferences (LM2 – LM1recall). In *RecallExPost*[15], the recollection error results in a misrepresentation of the preference order for 102 out of 196 participants (52%) that answered all three questions. Thus, the preference rankings implied by (LM2 – LM1) and (LM2 - LM1recall) differ. Hence using the reference point that is commonly observable in surveys (LM1) results in a different preference ranking than using the one that participants have in their minds (LM1recall). A further 203 participants that answered LM1 and LM2 were unable to recall LM1. For the latter the consistency rate cannot be determined but is likely to be lower than for those that did recall LM1. Pooled over all participants in the treatment, the percentage of measurements consistent with the signal is hence in between 48 percent and the random rate of a third.[16] Consistency rates in standard surveys will be lower due to the four times longer gap between consecutive waves. Hence, even when abstracting from the possibility that reporting and utility functions can change over time, within-subject comparisons of the level measure seem ill fitted to produce reliable information on preference rankings.

Asking participants to reassess the situation three months ago rather than asking them to recall their prior assessment, increases response rates significantly. Instead of 64 and 49 percent (of those answering LM1) in treatments *RecallExAnte* and *RecallExPost*, respectively, 98.25 percent (393 out of 400) answered the LM1reassess question in treatment *Reassess* including four participants that did not answer LM1. The difference to LM1 is comparable to the answers given to LM1recall in terms of mean deviation and number of exact matches. Only the variance of (LM1 - LM1reassess) is significantly larger than that of (LM1 - LM1recall) ($p < 0.001$, variance ratio

[15] We analyze *RecallExPost* since in *RecallExAnte* the answer to LMrecall might influence the later answer to LM2 and thus distort the observed change in LS. On the other hand, we thereby have to accept that LMrecall might be influenced by the earlier answer to LM2.
[16] The upper bound for *RecallExAnte* is 45.9 percent.

test). There is no reason to expect participants (even those with perfect memories) giving the same answers to LM1reassess and LM1 since both the reporting function and preferences might have changed.

Respondents struggle to remember the category selected in a LM question three month ago but have no trouble in reporting a reassessment of the past situation using a new ad hoc reporting function. This implies that many cannot remember the previously used reporting function as otherwise they could reconstruct their past answer. The experimental test indicates that changes in the LM are a poor indicator of preference rankings even with only three months in between.

4. The ranking measure

Asking individuals directly to rank the current situation against the situation from the previous period prevents an uncontrollable change in reporting function from one wave to the next and the memory issues detected in the previous section. An example of such a question from the BHPS:

"Would you say that you are more satisfied with life, less satisfied or feel about the same as you did a year ago?" where the answer options were *"More satisfied"*, *"Less satisfied"*, *"About the same"* and *"Don't know"*.

Such RM data can be described by $RM_{i,t}(A_{i,t},B_{i,t-1}) = r_{i,t}(u_{i,t}(A_{i,t}) - u_{i,t}(B_{i,t-1}))$. The individual compares today's situation $A_{i,t}$ with last period's situation $B_{i,t-1}$ represented by the utility difference. This utility difference (preference ranking) is transformed into an answer on the answer scale according to the reporting function $r_{i,t}(.)$. The RM captures preferences in period t over the two situations, i.e. it gives a retrospective evaluation of the experienced change in life.

In contrast to the LM, the two situations are compared using the same reporting and utility functions. This could not be guaranteed for the LM,

causing the identification problem. The RM scale is not bound in terms of levels. It is possible to observe arbitrarily long sequences of improvements or deteriorations. The RM also uses more intuitive categories. It seems easier to agree on what *"improvement"* means than on what is a "*4 out of 7"*. This increases comparability across individuals. It especially avoids transforming an ordinal concept into a cardinal answer in order to then evaluate it again in an ordinal manner by comparing changes of the LM, thereby avoiding mistakes in this transformation procedure.

Under which conditions is the RM reliably able to capture preference rankings? First, if there is no bias in memory about last period's situation, the preference ranking of the two situations is weakly correct. The reporting function might be coarser than actual preferences, such that a very small improvement might still be reported as indifference. The RM requires individuals to remember the past situation (without bias). However, people might be reluctant reporting that things have become worse or glorify the past, therefore biasing the RM up or down. While the RM is not immune against memory biases (Hoffrage et al. 2000, Karney and Coombs 2000, Levine and Safer 2002, Morewedge et al. 2005), it requires memories of one's life only while the LM on top of that critically relies on remembering an ad hoc reporting function or the answer to a specific survey questions for an entire year.[17]

Secondly, and this condition is equivalent to the monotonicity requirement of $f_{i,t}(.)$ for the LM: Participants need to intend to report their preference rankings and not another concept. Thus, $r_{i,t}(.)$ needs to preserve the sign of the change in utility. We think that the BHPS question failed to clearly ask for preferences instead of sensation intensity and can be improved in that respect.

[17] Note that studies of memory bias are not immune against misspecification either (see footnote 15).

Therefore we commissioned the following question in the GESIS panel and also used this formulation in the Lightspeed experiment:

"In what follows we would like to know how you evaluate the development of your life over the past 12 months. In your opinion, has your life overall improved or deteriorated over the past 12 months? Please evaluate the changes from today's perspective."

Direct measurements of changes have been found to outperform computed changes based on repeatedly reported levels in subjective assessments of health (Gunasekara et al. 2012) and preference reversals (Bateman et al., 2007).

The RM thus eliminates the uncontrollable change in utility and reporting functions, reduces the memory issues, is more intuitive, avoids the bounded scale, is ordinal and promises therefore to be better suited to measure preference rankings as compared to the LM.

5. Comparison of measures: empirical evidence

We have argued above that with the LM preference rankings cannot be reliably elicited and that the RM is a more reliable measure of preference rankings. The empirical difference between the two measures sheds light on the identification problem of the LM.

When reduced to the relevant information for a preference ranking, i.e. whether an individual perceives the present situation as "better" or "worse" than or is "indifferent" to the situation a year ago, the two measures are comparable. This is achieved by using the sign of the change in the level measure $signdLM_{i,t} = \text{sign}(dLM_{i,t})$ with $dLM_{i,t} = LM_{i,t} - LM_{i,t-1}$ and the ranking measure $RM_{i,t}$.

In order to compare RM and signdLM, we use their difference, which we call *deviation,* revealing when the information on preferences provided by the RM

deviates from the information provided by the LM: *deviation = RM - signdLM*. If *deviation = 0*, both measures report the same preference ranking. If, however, |*deviation*| = *2* they give opposite rankings. If |*deviation*| = *1*, they weakly disagree.

We draw on data from BHPS (University of Essex 2010), GESIS (GESIS 2017) and Lightspeed surveys which contain the LM and the RM. While the BHPS has ten years of observations, the GESIS Panel has so far only three waves of RM but features a more precise RM question for detecting preferences (as discussed in the previous section). The RM question in the GESIS panel and the complete Lightspeed survey were commissioned by us. Data descriptions and summary statistics can be found in Appendix A.1.

The BHPS panel is representative for Great Britain and starting from 2001 also for the whole of the UK. As far as we are aware, the data of the RM in the BHPS has never been used so far in any publication. For 117,244 observations both signdLM and RM are available. The GESIS panel is representative for Germany. While the LM was included in the GESIS panel starting 2014, the RM was added on our behalf in 2015. There are 9,553 valid observations for the years 2015 to 2017 in total for which both LM and RM are available. The Lightspeed dataset is representative of Germany and contains 794 observations from the *Control* and *RecallExPost* treatments where LM2 and RM are not affected by the experimental setup.

Table 3 reports the descriptive data and test results for signdLM, RM and *deviation*. Across all data sets, comparing the means, we find that signdLM tends more towards the negative than RM, confirmed by a significant positive mean *deviation*. Still, the share of participants reporting no noteworthy change in LS is higher for the RM than the LM which is consistent with substantially more noise in LM answers. Across all datasets, the preference rankings elicited from the two measures at least weakly disagree in more than half of all cases. Correlation of both measures is around .2 and tests of

asymptotic symmetry and marginal homogeneity are rejected. The two measures are thus significantly different in terms of means and distributions.

Table 3. Comparison of RM and signdLM. Standard error in brackets. Lightspeed based on treatments *Control* and *RecallExPost*. Difference from zero is reported at * p<0.05, **p<0.01, ***p<0.001 significance levels.

	BHPS	**GESIS**	**Lightspeed**
Country	UK	Germany	Germany
Years covered	1996 - 2000 2002 - 2008	2014 - 2017	2015
Observations	117,244	9,553	794
Mean signdLM	-.15*** (.002)	.006 (.008)	.006 (.028)
Mean RM	.11*** (.002)	.279*** (.010)	.144*** (.022)
Mean deviation	.046*** (.002)	.273*** (.010)	.137*** (.033)
signdLM = 0 (%)	46.4	30.732.1	37.2
RM = 0 (%)	57.9	38.739.2	59.4
signdLM < 0 (%)	27.6	42.233.6	31.1
RM < 0 (%)	15.6	16.4	13.1
Mean deviation	.046*** (.002)	.272835*** (.0102)	.137*** (.033)
Weak disagreement (%)	47.1	42.73	48.7
Agreement (%)	45.5	41.1	41.8
Corr(signdLM,RM)	.21	.178	.17
t-test signdLM = RM (p-value)	.000	.000	.000
Marginal homogeneity (Stuart Maxwell, p-value)	.000	.000	.000

Some of the inconsistency between signdLM and RM is due to noise in signdLM that is caused by inaccurate but unbiased recollection of the LM reported in the previous wave. Treatment *RecallExAnte* in the Lighspeed dataset allows to quantify this by comparing the inconsistency rates between signdLM and RM (55.6%) and between LM2 – LM1recall and RM (31.8%), thus 57% (= 31.8 / 55.6) of inconsistency between signdLM and RM is due to inaccurate recollection. In the first case participants had to remember their

answer for three months, in the second only for a few seconds. Systematic adjustments in the reporting function cannot be eliminated by this procedure.

Next we test whether the difference between the measures is systematic. Tables 4a and 4b contain results for the BHPS and Table 5 for the GESIS data set. The tables report regressions of deviation, signdLM and RM depending on changes in explanatory variables ($dX_{i,t}$) such as income, employment status and marital status, on time-invariant personal variables (X_i) such as gender, on year dummies (X_t) and allowing for the possibility of unobserved personal fixed effects (u_i):

$$\begin{aligned} signdLM_{i,t} &= sign\left(LM_{i,t} - LM_{i,t-1}\right) \\ &= \beta_1\, dX_{i,t} + \beta_2\, X_i + \beta_3\, X_t + u_i + e_{i,t} \end{aligned}$$

$$RM_{i,t} = \hat{\beta}_1\, dX_{i,t} + \hat{\beta}_2\, X_i + \hat{\beta}_3\, X_t + \hat{u}_i + \hat{e}_{i,t}$$

$$deviation = RM_{i,t} - signdLM_{i,t} = \tilde{\beta}_1\, dX_{i,t} + \tilde{\beta}_2\, X_i + \tilde{\beta}_3\, X_t + \tilde{u}_i + \tilde{e}_{i,t}$$

For BHPS, two specifications are presented. Table 4a contains self-assessed changes in financial situation (improved, stayed the same, deteriorated), Table 4b contains change in household income. For GESIS, the change in financial situation is reported in Table 5.[18] We have checked for multicollinearity by means of the variance inflation factor, which for all variables in all specifications is between 1 and 2 except for age and age^2, which by definition are correlated with each other.

We report the results of pooled OLS regressions. According to the Breusch-Pagan test, unobserved personal fixed effects are present for the RM and a Hausman test suggests using a fixed effects regression which we report in addition to pooled OLS. The results are quite similar though. This fixed effects specification takes into account personal fixed effects in *changes* over

[18] The version with income is presented in the appendix (Table A.6).

time. Person fixed effects in *levels* are already accounted for in the pooled OLS regressions since we use first differences (changes over time) as dependent variable and for most independent variables. Since the dependent variables have either three or five categories, an ordered logistic regressions is reported in the appendix as robustness check (Table A.7). Specifications using the natural logarithm of income are also reported in the appendix (Table A.9). This does not improve the significance of the income variable but also doesn't change the other coefficients.

Regressions (1) in Tables 4a and 4b confirm that the *deviation* between the preference relations elicited by LM and RM is systematically correlated with key socio-demographic variables. Age, gender, changes in the financial situation, changes in employment status and changes in marital status are significant drivers of *deviation* in the BHPS. The difference between the LM and the RM is hence systematically correlated with changes in key socio-economics variables and therefore at least one of them provides biased estimates of preference rankings over these attributes.[19] The R^2 is consistently about ten times higher in the RM specification than for signdLM, consistent with the RM being less noisy.

Next, we test for adaptation in the reporting function, focusing again on Tables 4a and 4b.[20] When comparing the coefficients of changes and lagged changes in socio-economic variables, a clear pattern emerges. In the BHPS, an improvement in financial situation, finding a job, leaving the labor market (not due to unemployment but e.g. into retirement) and getting married all have positive immediate impacts on both measures, a positive impact on next year's RM but a negative impact on next year's LM. The same pattern but with reversed signs holds for becoming unemployed.

[19] This also holds for the Lightspeed dataset (see Table A.3).
[20] We also report the results of Tables 4a and 4b excluding the lags in Tables A.4a and A.4b.

The LM therefore has a clear tendency towards findings that people get used to whatever happens to them. The key question is whether this adaptation reflects changes in preferences ('once you got to know it, more money isn't that great and unemployment not that bad after all'), which has been a dominant interpretation in the adaptation literature, or merely an adjustment of the reporting function.

The RM exhibits the opposite pattern. Some of the benefits of an improvement in the financial situation and some of the downsides of unemployment only seem to occur to (or hit) people with some delay. By construction of the RM, this cannot be attributed to changes in the reporting function. At least for some of the variables plausible explanations for the reinforcing impact of lags exist. After losing a job, people might first be hopeful that they find a new one soon. One year later, this hope might have died. Note that both changes in the financial situation and whether a new job has been found in the meantime are controlled for.

For the GESIS data in Table 5, this adaptation pattern is only observed for the change in financial situation.[21] The results for GESIS using income instead of financial situation can be found in Table A.6.

[21] This seems to be due to the fact that there are only three years of observations. When the BHPS is only analyzed for three periods (instead of ten), the picture gets very similar (see appendix Tables A.5a and A.5b).

Table 4a. BHPS: Regressions on key socio-economic variables (financial situation). Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) **deviation OLS**	(2) **signdLM OLS**	(3) **signRM OLS**	(4) **signRM FE**
female	0.0171**	0.00332	0.0205***	0
	(0.00615)	(0.00336)	(0.00567)	
age	-0.0209***	0.00466***	-0.0162***	-0.0292
	(0.00109)	(0.000649)	(0.000971)	(0.0326)
age^2	0.000157***	-0.0000456***	0.000111***	-0.00000131
	(0.0000106)	(0.00000642)	(0.00000942)	(0.0000231)
change_fin_situation	**0.105*****	**0.0844*****	**0.189*****	**0.163*****
	(0.00474)	(0.00399)	(0.00386)	(0.00349)
L.change_fin_situation	**0.105*****	**-0.0593*****	**0.0458*****	**0.0179*****
	(0.00469)	(0.00400)	(0.00357)	(0.00350)
stayed unemployed	-0.120***	-0.00902	-0.129***	-0.0907**
	(0.0332)	(0.0245)	(0.0277)	(0.0326)
stayed out of labor market	-0.0428***	0.00382	-0.0390***	0.0359**
	(0.00883)	(0.00522)	(0.00792)	(0.0121)
got employed	**0.0548****	**0.0439****	**0.0987*****	**0.125*****
	(0.0175)	(0.0148)	(0.0136)	(0.0139)
lag got employed	**0.0198**	**0.00470**	**0.0245***	**0.0415*****
	(0.0151)	(0.0129)	(0.0115)	(0.0115)
got unemployed	**-0.0922*****	**-0.0896*****	**-0.182*****	**-0.114*****
	(0.0250)	(0.0212)	(0.0199)	(0.0186)
lag got unemployed	**-0.154*****	**0.0501***	**-0.104*****	**-0.0635****
	(0.0272)	(0.0233)	(0.0207)	(0.0199)
exited labor market	**0.0726*****	**0.0432****	**0.116*****	**0.138*****
	(0.0166)	(0.0141)	(0.0132)	(0.0127)
lag exited labor market	**0.109*****	**-0.0403****	**0.0688*****	**0.0398****
	(0.0169)	(0.0142)	(0.0126)	(0.0126)
got married	**0.190*****	**0.0535****	**0.244*****	**0.113*****
	(0.0230)	(0.0194)	(0.0173)	(0.0180)
lag got married	**0.192*****	**-0.0328**	**0.159*****	**0.150*****
	(0.0234)	(0.0197)	(0.0180)	(0.0177)
marriage ended	**-0.141*****	**-0.0213**	**-0.163*****	**-0.252*****
	(0.0310)	(0.0284)	(0.0277)	(0.0233)
lag marriage ended	**-0.0469**	**0.0895*****	**0.0426**	**-0.0251**
	(0.0307)	(0.0264)	(0.0228)	(0.0219)
stayed married	-0.00202	-0.00939*	-0.0114	-0.174***
	(0.00705)	(0.00400)	(0.00645)	(0.0136)
Year dummies	yes	yes	yes	yes
Constant	0.233***	-0.117***	0.116***	1.099
	(0.0306)	(0.0197)	(0.0271)	(1.368)
Observations	83533	83533	83533	83533
R^2	0.052	0.012	0.109	0.052
Adjusted R^2	0.051	0.011	0.109	-0.220
AIC	207001.0	183091.7	151818.8	114298.3
BIC	207243.7	183334.4	152061.5	114531.6

Table 4b. BHPS: Regressions on key socio-economic variables (income). Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects (not reported). Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1)	(2)	(3)	(4)
	deviation OLS	**signdLM OLS**	**signRM OLS**	**signRM FE**
female	0.0146*	0.00300	0.0176**	0
	(0.00633)	(0.00336)	(0.00591)	(.)
age	-0.0252***	0.00407***	-0.0211***	-0.0126
	(0.00110)	(0.000642)	(0.000989)	(0.0332)
age^2	0.000191***	-0.0000402***	0.000151***	0.0000175
	(0.0000108)	(0.00000637)	(0.00000961)	(0.0000233)
d_hh_inc_month	**0.00177**	**0.00159**	**0.00336***	**0.00151**
	(0.00192)	(0.00181)	(0.00142)	(0.00137)
L.d_hh_inc_month	**0.00262**	**-0.00237**	**0.000245**	**-0.00131**
	(0.00194)	(0.00163)	(0.00144)	(0.00137)
stayed unemployed	-0.183***	-0.0230	-0.206***	-0.166***
	(0.0333)	(0.0247)	(0.0285)	(0.0328)
stayed out of labor market	-0.0822***	-0.00102	-0.0833***	-0.00271
	(0.00895)	(0.00515)	(0.00811)	(0.0122)
got employed	**0.0440***	**0.0695***	**0.114***	**0.140***
	(0.0174)	(0.0148)	(0.0139)	(0.0141)
lag got employed	**0.0418****	**0.00164**	**0.0434***	**0.0538***
	(0.0153)	(0.0129)	(0.0120)	(0.0116)
got unemployed	**-0.165***	**-0.125***	**-0.290***	**-0.210***
	(0.0250)	(0.0213)	(0.0206)	(0.0187)
lag got unemployed	**-0.186***	**0.0755****	**-0.111***	**-0.0610****
	(0.0273)	(0.0233)	(0.0212)	(0.0201)
exited labor market	**0.0203**	**0.0158**	**0.0361****	**0.0700***
	(0.0166)	(0.0141)	(0.0133)	(0.0128)
lag exited labor market	**0.0799***	**-0.0327***	**0.0472***	**0.0291***
	(0.0169)	(0.0141)	(0.0127)	(0.0127)
got married	**0.200***	**0.0526****	**0.252***	**0.117***
	(0.0231)	(0.0195)	(0.0176)	(0.0182)
lag got married	**0.200***	**-0.0366**	**0.163***	**0.152***
	(0.0235)	(0.0197)	(0.0182)	(0.0179)
marriage ended	**-0.185***	**-0.0398**	**-0.225***	**-0.306***
	(0.0311)	(0.0285)	(0.0282)	(0.0234)
lag marriage ended	**-0.0867****	**0.0967***	**0.0100**	**-0.0452***
	(0.0311)	(0.0263)	(0.0233)	(0.0220)
stayed married	-0.00838	-0.00995*	-0.0183**	-0.181***
	(0.00723)	(0.00399)	(0.00670)	(0.0138)
Year dummies	yes	yes	yes	yes
Constant	0.809***	-0.0493**	0.760***	0.766
	(0.0265)	(0.0166)	(0.0234)	(1.390)
Observations	84112	84112	84112	84112
R^2	0.036	0.005	0.064	0.020
Adjusted R^2	0.035	0.004	0.064	-0.259
AIC	209865.3	184980.1	157105.5	118102.5
BIC	210108.1	185222.9	157348.3	118336.0

Table 5. GESIS: Regressions on key socio-economic variables. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects (not reported). Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) **deviation OLS**	(2) **signdLM OLS**	(3) **signRM OLS**	(4) **signRM FE**
female	-0.00509	-0.00523	-0.0103	-
	(0.0247)	(0.0191)	(0.0195)	(.)
age	-0.0252**	0.00475	-0.0205**	0.0460
	(0.00821)	(0.00648)	(0.00630)	(0.0704)
age^2	0.000167*	-0.0000501	0.000117	-0.000501
	(0.0000847)	(0.0000664)	(0.0000650)	(0.000664)
change_fin_situation	**0.232*****	**0.119*****	**0.351*****	**0.304*****
	(0.0163)	(0.0144)	(0.0123)	(0.0194)
L.change_fin_situation	**0.152*****	**-0.0937*****	**0.0584*****	**0.0247**
	(0.0161)	(0.0140)	(0.0112)	(0.0180)
stayed unemployed	0.165	-0.274	-0.109	-0.0599
	(0.212)	(0.161)	(0.131)	(0.289)
stayed out of labor market	-0.0748	0.0112	-0.0635	0.0263
	(0.0425)	(0.0340)	(0.0336)	(0.129)
got employed	**-0.0336**	**-0.0159**	**-0.0496**	**-0.0710**
	(0.0738)	(0.0603)	(0.0503)	(0.104)
lag got employed	**0.0536**	**0.0225**	**0.0761***	**0.0969**
	(0.0621)	(0.0526)	(0.0388)	(0.0685)
got unemployed	**-0.0671**	**0.0510**	**-0.0161**	**-0.00631**
	(0.138)	(0.116)	(0.0861)	(0.161)
lag got unemployed	**0.0812**	**-0.0521**	**0.0292**	**-0.146**
	(0.177)	(0.137)	(0.105)	(0.164)
exited labor market	**0.0498**	**0.0653**	**0.115***	**0.0966**
	(0.0673)	(0.0548)	(0.0470)	(0.0773)
lag exited labor market	**0.0815**	**0.0251**	**0.107***	**0.0878**
	(0.0665)	(0.0559)	(0.0470)	(0.0757)
stayed married	0.0364	-0.0332	0.00315	-0.345
	(0.0294)	(0.0228)	(0.0238)	(0.181)
got married	**-0.00950**	**0.0972**	**0.0877**	**-0.0392**
	(0.113)	(0.100)	(0.0748)	(0.135)
lag got married	**-0.0174**	**0.0785**	**0.0611**	**0.0527**
	(0.0952)	(0.0836)	(0.0694)	(0.101)
marriage ended	**-0.0725**	**0.0743**	**0.00175**	**-0.128**
	(0.132)	(0.103)	(0.0881)	(0.159)
lag marriage ended	**-0.0332**	**0.0957**	**0.0625**	**0.0652**
	(0.104)	(0.0914)	(0.0745)	(0.115)
Year dummies	yes	yes	yea	yes
Constant	-0.108	-0.195	-0.303	-1.509
	(0.201)	(0.163)	(0.159)	(1.841)
Observations	4864	4864	4864	4864
R^2	0.107	0.022	0.257	0.127
Adjusted R^2	0.103	0.018	0.254	-0.998
AIC	13106.7	11680.3	9219.6	4058.3
BIC	13236.5	11810.0	9349.4	4175.2

Table 6. Deviation between measures on direction of change and trends in financial situation (BHPS). Description of variables given in Appendix A.1. Results from pooled OLS regression including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) deviation	
female	0.0182**	(0.00594)
age	-0.0206***	(0.00105)
age^2	0.000155***	(0.0000104)
financial deterioration	-0.0438***	(0.00907)
Lag financial deterioration	-0.102***	(0.00968)
financial improvement	0.134***	(0.00957)
Lag financial improvement	0.0783***	(0.00903)
negative trend	**-0.0462****	**(0.0157)**
positive trend	**0.0442****	**(0.0146)**
stayed unemployed	-0.116***	(0.0320)
stayed out of labor market	-0.0409***	(0.00851)
got employed	0.0489**	(0.0169)
got unemployed	-0.0986***	(0.0227)
exited labor market	0.0668***	(0.0159)
lag got employed	0.0209	(0.0147)
lag got unemployed	-0.150***	(0.0253)
lag exited labor market	0.111***	(0.0162)
got married	0.189***	(0.0227)
marriage ended	-0.150***	(0.0297)
stayed married	-0.00147	(0.00678)
lag got married	0.192***	(0.0229)
lag marriage ended	-0.0453	(0.0284)
Year dummies	yes	
Constant	0.624***	(0.0265)
Observations	83533	
R^2	0.053	
Adjusted R^2	0.052	
AIC	206936.4	
BIC	207216.4	

Is it possible to identify what is driving this difference in the impact of lagged changes in life circumstances? One of the motivations for adjusting the reporting function of the LM is to preserve the ability to report future changes on a bound answer scale. This is especially relevant if there is a clear trend in the variable to be assessed. Changes in the financial situation are the only

candidate among the variables available here that can exhibit a trend over several periods.

Table 6 provides some additional insights into the adaptation pattern that is observed for the LM but not the RM. It reports how deviation is affected by trends in the financial situation. For this purpose, changes in financial situation have been categorized into improvements, no change and deteriorations that enter both directly and lagged (similar to the previous tables). Moreover, the dummy variables 'positive trend' and 'negative trend' have been created. They equal one if the financial situation has improved (deteriorated) for two successive periods. Coefficients for both are highly significant and lend support to the hypothesis that reporting functions of the LM are adjusted more if there is an experienced trend in the assessed variable.[22]

The main regression results remain when excluding observations where LM is either 1 or 7 in the BHPS (Table A.10) or the lag of LM is either 1 or 7 (Table A.11), i.e. when participants have reached the end of the answer scale and thus have no further opportunity of indicating improvement or deterioration respectively without adjustments in the reporting function. This indicates that adaptation of reporting functions starts before the bounds of the answer scale are reached.

Our results provide a wide range of supporting evidence for adaptation of the reporting function and for systematic distortion in the representation of preference rankings when using the LM.[23]

[22] Results are similar but not significant when using leads instead of lags for creating the trends, which can be interpreted as proxies for expected trends (see Table A.8).
[23] For evidence that the level measure is unreliable in eliciting estimates of the marginal rate of substitution between different attributes that are complementary to those presented here, see Benjamin et al. (2014a).

6. Conclusion and Discussion

Several conceptual concerns raise doubt that the established level-measure of life satisfaction is suitable to elicit preference rankings over states of the world. The adjustment of preferences and, more importantly, of the way they are mapped onto the discrete and bound answer scale via ad hoc reporting functions reduce the comparability of answers given by survey panel members in subsequent years. Empirical evidence from three representative surveys covering two countries, almost two decades and well over a hundred thousand observations give credence to these concerns. An alternative elicitation method, the ranking measure of life satisfaction is introduced and compared with the level measure.

The key empirical findings are that survey participants are largely unable to remember answers given to life satisfaction questions three months ago but are able to reassess the situation they experienced three months ago using a new reporting function. The noise introduced by imprecise recollection of answers given three months ago reverts the preference ranking elicited by the level measure in more than half of all cases. The ranking measure raises consistency rates by more than 50 percent compared to the level measure in a controlled experiment. In all three panel surveys the change in the level measure and the ranking measure are only weakly correlated and systematically differ in their means and distributions. These differences are systematically correlated with key socio-economic variables such as age, changes in the financial situation, employment and marital status. The level measure but not the ranking measure systematically features adaptation to changes in life's circumstances and especially so when there is a trend in the underlying socio-economic variable.

Combining the conceptual concerns and the empirical evidence strongly suggests that the level measure of life satisfaction produces biased estimates of preference rankings. In particular the widely reported adaptation to income

changes and other aspects of life seems to be mainly driven by adjustments in the reporting function rather than by changes in the underlying preferences. This challenges at least some interpretations of key results from the life satisfaction literature such as the Easterlin Paradox. Levels of life satisfaction have remained mostly flat in developed countries despite substantial increases in average income over the last decades. This might be an adequate description of the satisfaction sensation in the population, but the results presented here call into question that this implies an indifference towards increases in per capita income in the long run. Depending on which concept policy makers decide to target, recommendations will be very different. According to ex-post preferences, increases in income do matter, also in the long run, whereas for long-run satisfaction intensities this seems not to be the case.

The reference to the Easterlin Paradox raises another issue. The level measure allows to compare satisfaction levels at very different points in time and to plot charts in terms of cardinal levels. However, given that the level measure is a poor proxy for preference rankings even for two subsequent years, it is important to exercise great care in interpreting these charts, i.e. not to take them as representations of preferences. The ranking measure does not allow drawing the same types of charts. This draws attention to the fact that the ordinal concept of preference rankings requires a direct comparison of two states of the world. Hence, if states five or ten years apart are to be compared, corresponding questions need to be included in the survey. However, the issue of imperfect and potentially biased recollection of past states of the world might then be a more serious issue. The level measure obscures problems with comparability (even for subsequent periods) by using an absolute (cardinal) scale for a relative (ordinal) concept. While such comparisons might be enlightening for some research questions and policy issues, their use and interpretation requires careful judgment by the analyst as not to be taken as

indicators of preference rankings. The latter are more reliably captured by the ranking measure of life satisfaction.

Knowing how people evaluate changes they have experienced is important for policy evaluation since this is both an essential input for many conventional welfare measures and, more practically, it is valuable information for policymakers that want to be reelected. The ranking measure is well suited to provide such insights.

Future research could compare the level and the ranking measure of life satisfaction in more detail shedding light on the differences in drivers, developments over time, and especially marginal rates of substitution for public goods (e.g. as in Welsch and Kühling 2009, Frey et al. 2004, Kopmann and Rehdanz 2013). This will further identify where the two measures and interpretations of life satisfaction differ and where a clear political decision for one of the target concepts needs to be made.

References

Bateman, I., Day, B., Loomes, G., & Sugden, R. (2007). Can ranking techniques elicit robust values? *Journal of Risk and Uncertainty*, 34(1), 49-66.

Benjamin, D. J., & Heffetz, O., Kimball, M.S., & Rees-Jones, A. (2012). What do you think would make you happier? What do you think you would choose? *American Economic Review*, *102*(5), 2083-2110.

Benjamin, D. J., Heffetz, O., Kimball, M. S., & Rees-Jones, A. (2014a). Can marginal rates of substitution be inferred from happiness data? Evidence from residency choices. *American Economic Review*, *104*(11), 3498-3528.

Benjamin, D. J., Kimball, M. S., Heffetz, O., & Szembrot, N. (2014b). Beyond happiness and satisfaction: Toward well-being indices based on stated preference. *American Economic Review*, *104*(9), 2698-2735.

Berridge, K. C., & Kringelbach, M. L. (2011). Building a neuroscience of pleasure and well-being. *Psychology of Well-Being: Theory, Research and Practice*, *1*(1), 3.

Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, *91*(2), 67-72.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, *43*(1), 9-64.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, *46*(1), 95-144.

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: revising the adaptation theory of well-being. *American Psychologist*, *61*(4), 305.

Di Tella, R., Haisken-De New, J., & MacCulloch, R. (2010). Happiness adaptation to income and to status in an individual panel. *Journal of Economic Behavior & Organization*, *76*(3), 834-852.

Di Tella, R., & MacCulloch, R. (2006). Some uses of happiness data in economics. *Journal of Economic Perspectives*, *20*(1), 25-46.

Di Tella, R., MacCulloch, R. J., & Oswald, A. J. (2003). The macroeconomics of happiness. *Review of Economics and Statistics*, *85*(4), 809-827.

Decancq, K., Fleurbaey, M., Schokkaert, E. (2015). Happiness, equivalent incomes, and respect for individual preferences, Economica 82: 1082-1106.

Dolan, P., Peasgood, T., & White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, *29*(1), 94-122.

Easterlin, R. (1974). Does economic growth improve the human lot? Some empirical evidence. In P. David & M. Reder (Eds.), *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz* (pp. 98-125). New York and London: Academic Press.

Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. *Proceedings of the National Academy of Sciences*, *107*(52), 22463-22468.

Fehr, E., & Rangel, A. (2011). Neuroeconomic foundations of economic choice—recent advances. *Journal of Economic Perspectives*, *25*(4), 3-30.

Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. *Journal of Public Economics*, *89*(5), 997-1019.

Fleurbaey, M. (2009). Beyond GDP: The quest for a measure of social welfare. *Journal of Economic Literature*, 47(4), 1029-1075.

Fleurbaey, M. & Schwandt, H. (2016). Do People Seek to Maximize Their Subjective Well-Being – And Fail?. Paper presented at the Annual Meeting of the American Economic Association, Chicago, 6th Jan. 2017, https://www.aeaweb.org/conference/2017/preliminary/paper/zYFd2Rdd [17th August 2017].

Frederick, S., & Loewenstein, G. (1999). Hedonic Adaptation. in Kahneman, Diener, and Schwarz (eds.) Well-being: The foundations of hedonic psychology, New York: Russell Sage Foundation, 302-29.

Frey, B. S., Luechinger, S., & Stutzer, A. (2004). Valuing public goods: The life satisfaction approach. *CESifo Working Paper*, (No. 1158).

Frey, B. S., & Stutzer, A. (2002). What can economists learn from happiness research? *Journal of Economic Literature*, *40*(2), 402-435.

GESIS (2017). GESIS Panel - Standard Edition. GESIS Data Archive, Cologne. ZA5665 Data file Version 19.0.0, doi:10.4232/1.12743

Gunasekara, F. I., Carter, K., & Blakely, T. (2012). Comparing self-rated health and self-assessed change in health in a longitudinal survey: Which is more valid?. *Social Science & Medicine*, *74*(7), 1117-1124.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 566.

Jones, Charles I. and Peter J. Klenow. (2016). "Beyond GDP? Welfare across Countries and Time." *American Economic Review*, 106(9): 2426-57.

Karney, B. R., & Coombs, R. H. (2000). Memory bias in long-term close relationships: Consistency or improvement?. *Personality and Social Psychology Bulletin*, 26(8), 959-970.

Kopmann, A., & Rehdanz, K. (2013). A human well-being approach for assessing the value of natural land areas. *Ecological Economics*, *93*, 20-33.

Layard, R. (2006). Happiness and public policy: A challenge to the profession. *Economic Journal*, *116*(510), C24-C33.

Layard R., G. Mayraz, S. Nickell (2008). The marginal utility of income, *Journal of Public Economics*, 92, 1846 -1857.

Levine, L. J., & Safer, M. A. (2002). Sources of bias in memory for emotions. *Current Directions in Psychological Science*, 11(5), 169-173.

Loewenstein, G., & Ubel, P. A. (2008). Hedonic Adaptation and the Role of decision and Experience Utility in Public Policy. *Journal of Public Economics*, 92(8–9): 1795–1810.

Luttmer, Erzo F. P. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *Quarterly Journal of Economics*, *120*(3), 963-1002.

MacKerron, G. (2012). Happiness economics from 35 000 feet. *Journal of Economic Surveys*, *26*(4), 705-735.

Morewedge, C. K., Gilbert, D. T., & Wilson, T. D. (2005). The least likely of times: How remembering the past biases forecasts of the future. *Psychological Science*, 16(8), 626-630.

Oswald, A. J. (1997). Happiness and economic performance. *Economic Journal*, *107*(445), 1815-1831.

Oswald, A. J., & Wu, S. (2010). Objective confirmation of subjective measures of human well-being: Evidence from the USA, *Science*, 327(5965), 576-579.

Perez-Truglia, R. (2015). A Samuelsonian validation test for happiness data. *Journal of Economic Psychology*, 49, 74-83.

Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin*, *86*(4), 777-803.

Rayo, L., & Becker, G. S. (2007a). Evolutionary efficiency and happiness. *Journal of Political Economy*, *115*(2), 302-337.

Rayo, L., & Becker, G. S. (2007b). Habits, peers, and happiness: An evolutionary perspective. *American Economic Review*, *97*(2), 487-491.

Sen, A. (1985). *Commodities and Capabilities*. North-Holland, Amsterdam.

Stevenson, B., & Wolfers, J. (2008). Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. *Brookings Papers on Economic Activity*, 39(1), 1-102.

Stevenson, B., & Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation? *American Economic Review*, *103*(3), 598-604.

University of Essex. Institute for Social and Economic Research (2010). *British Household Panel Survey: Waves 1-18, 1991-2009*. [data collection]. 7th Edition. UK Data Service. SN: 5151, http://doi.org/10.5255/UKDA-SN-5151-1

Welsch, H., & Kühling, J. (2009). Using happiness data for environmental valuation: issues and applications. *Journal of Economic Surveys*, *23*(2), 385-406.

Appendix (For Online Publication)

A.1 Description of variables

Here we describe the explanatory variables used in the regressions. The following variables[24] are available and summarized in Table A.1 and Table A.2:

Table A.1 Summary Statistics of BHPS variables

	(1) count	mean	sd	min	max
signdLM	117244	-.0148408	.732213	-1	1
signRM	117244	.1101122	.6397049	-1	1
difference	117244	.1249531	.8655879	-2	2
female	115413	.5476419	.4977272	0	1
age	115404	46.91606	18.12847	17	100
change_fin_situation	116830	2.041051	.6940744	1	3
hh_inc_month	117244	2.676986	2.134119	-.009	86.70329
d_hh_inc_month	117244	.0820844	1.704277	-65.97318	86.54298
fin_trend3	116449	2.041383	.461874	1	3
inc_trend3	104883	2.351201	.639902	1	3
fin_trend3_lead	85219	2.031789	.4571693	1	3
inc_trend3_lead	104654	2.344583	.6417428	1	3
d_employment_status	117203	2.100305	1.3621	1	6
d_marital_status_legal	116846	3.386466	.6048385	1	4

Table A.2 Summary Statistics of GESIS variables

	(1) count	mean	sd	min	max
signdLM	9553	.0059667	.8239416	-1	1
signRM	9553	.2788653	.7277624	-1	1
difference	9553	.2728986	.9995578	-2	2
female	9553	.5174291	.4997223	0	1
age	9523	49.893	14.11716	20	74
change_fin_situation	9465	3.191337	.8915453	1	5
hh_income	7812	3.185625	1.656383	.45	7
d_hh_inc	6656	.0282527	1.011679	-6.55	6.55
d_employment_status	8604	1.900628	1.454252	1	6
d_marital_status_legal	8849	3.337326	.600452	1	4

[24] The codebooks are available under http://www.gesis.org/unser-angebot/daten-erheben/gesis-panel/gesis-panel-data-usage/ and https://www.iser.essex.ac.uk/bhps/documentation/volb where we use the "individual-level data for respondents" in waves F to R except K.

d_hh_income_month is the change in monthly household income from past to current year in 1,000 EUR for Gesis and 1,000 GBP for BHPS. In GESIS net monthly household income is computed from changes in reported categories of income ranges of which we took the middle value to describe the category. The question was: "`If you take a look at the total income from all members of the household: how high is the monthly average household income today? I.e. the sum of all incomes including pensions and social benefits? Please use the list below"'[25] There are 14 answer categories ranging from "`700 Euro and less"' to "`6000 Euro and more"'. In BHPS, the household monthly income is computed from household income in different categories (labor income, capital income...) that are asked separately. The answer scales for the subcategories are open.

Change_fin_situation is a subjective evaluation of the change in financial situation. In GESIS, the question is: "In your opinion, has your Life improved or deteriorated in the following domains? Financial situation:"'[26]. There are five answer options: "Considerably improved", "Slightly improved", "Stayed the same", "Slightly deteriorated", "Considerably deteriorated" and "Don't know".

In BHPS, the question was "`Would you say that you yourself are better off or worse off financially than you were a year ago?"' and the answer options were "`Better off"', "`Worse off"', "`About the same"' and "`Don't know"'. In both GESIS and BHPS, the question about current employment status contains categories like employed part time, employed full time, unemployed,

[25] The German original version is: "Wenn man nun die Einkünfte aller Mitglieder Ihres Haushalts zusammen nimmt: Wie hoch ist das durchschnittliche monatliche Nettoeinkommen aller Haushaltsmitglieder also die Summe aller Einkünfte einschließlich aller Bezüge und Sozialleistungen insgesamt? Benutzen Sie bitte wieder die Liste".
[26] German original: "Hat sich Ihr Leben Ihrer Meinung nach in den folgenden Bereichen in den letzten 12 Monaten verschlechtert oder verbessert? Finanzielle Situation:"

student, old-age pensioner, unable to work etc. We computed 6 categories for the change in employment situation d_employment_status depending on current and previous answer:

- *got unemployed*: switched to "`unemployed"' from any other category,
- *got employed*: switched to "`employed"' or "`part time employed"' from any other category,
- *exit labor market*: switched from "`emplpoyed"' or "`unemployed"' or "`part time employed"' to any but those categories,
- *stayed employed*: was in any of the categories "`emplpoyed"' or "`part time employed"',
- *stayed unemployed*: stayed in "`unemployed"',
- *stayed out of labor market*: was in anything but "`emplpoyed"' or "`unemployed"' or "`part time employed"'.

In both, GESIS and BHPS, marital status contains different categories like married, divorced, never married, widowed, and separated. We computed the 4 following categories of the change in marital status d_marital_status_legal:

- *got married*: switch to married from any other category,
- *marriage ended*: switch to divorced from any other category,
- *stayed married*: stayed in category married,
- *stayed not married*: stayed in any group or switch between these groups: divorced, widowed, separated.

To compute the trend in financial situation or in income for BHPS (in Gesis there are not enough waves in the panel yet) we computed fin_trend3 and inc_trend3. They both are categorical variables with three categories computed for both change_fin_situation and d_hh_income_month:

- *Negative trend* if there was improvement in t and t-1
- *No trend* if there was anything else than positive or negative trend
- *Positive trend* if there was deterioration in t and t-1

For the expected trend inc_trend3_lead and fin_trend3_lead we used t+1 instead of t-1.
We also include gender and age.

Lightspeed data:

LM1 and LM2:
We would like to ask you how you evaluate your life.
Everything taken together, how satisfied are you with your CURRENT life situation?

10 categories from "completely dissatisfied" to "completely satisfied", with "don't know" option.

> *Im Folgenden würden wir gerne von Ihnen wissen, wie Sie Ihr Leben bewerten.*
> *Alles in allem betrachtet, wie zufrieden sind Sie mit Ihrer* ***GEGENWÄRTIGEN*** *Lebenssituation?*
>
> *Ganz und gar unzufrieden – Ganz und gar zufrieden, weiß nicht*

LM1recall:

How well can you remember the answer you gave during the first questionnaire 3 months ago concerning the satisfaction with your situation back then?

I can remember exactly, it was:
I am not sure but I think it was:
I cannot remember.

> *Wie gut können Sie sich daran erinnern, welche Antwort Sie in der ersten Befragung vor 3 Monaten bezüglich Ihrer Zufriedenheit mit Ihrer damaligen Lebenssituation gegeben haben?*
>
> *Ich kann mich genau erinnern, es war:*
> *Ich bin mir nicht sicher, aber ich denke es war:*
> *Ich kann mich nicht erinnern.*

LM1reassess:

We would like to ask you how you evaluate your life AT THE TIME OF THE FIRST QUESTIONNAIRE (beginning of August 2015).
Everything taken together, how satisfied were you with your life situation 3 months ago?

> *Im Folgenden würden wir gerne von Ihnen wissen, wie Sie Ihr Leben* ***ZUM ZEITPUNKT DER ERSTEN BEFRAGUNG*** *(Anfang Juli 2015) bewerten.*
> *Alles in allem betrachtet, wie zufrieden waren Sie mit Ihrer Lebenssituation vor 3 Monaten?*

RM:

We would like to ask you how you evaluate the CHANGES in your life since the first questionnaire (beginning of August 2015).

In your opinion, has your life as a whole since then improved or become worse?

Please evaluate any changes from today's perspective.

Considerably deteriorated	*Slightly deteriorated*	*Stayed the same*	*Slightly imporved*	*Considerably imporved*	*Don't know*

Im Folgenden würden wir gerne von Ihnen wissen, wie Sie die ***VERÄNDERUNGEN*** *Ihres Lebens seit der ersten Befragung (Anfang Juli 2015) bewerten.*

Hat sich Ihr Leben Ihrer Meinung nach in dieser Zeit insgesamt verbessert oder verschlechtert?

Bewerten Sie die Veränderungen aus heutiger Sicht.

Deutlich verschlechtert	*Leicht verschlechtert*	*Gleich geblieben*	*Leicht verbessert*	*Deutlich verbessert*	*Weiß nicht*

A.2 Additional regression results

Table A.3 Lightspeed: Ordered logit regressions on changes in perceived changes in socio-economic circumstances (Treatments: *Control* and *RecallExPost*). Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) difference	(2) signdLM	(3) RM
Age	-0.0161**	0.0120*	-0.0111
	(0.00551)	(0.00528)	(0.00653)
Female	0.106	-0.228	-0.204
	(0.135)	(0.134)	(0.156)
ch_job	0.372**	0.161	1.151***
	(0.126)	(0.141)	(0.200)
ch_money	0.309*	0.210	1.059***
	(0.129)	(0.127)	(0.177)
ch_leisure	0.380*	-0.160	0.565**
	(0.167)	(0.141)	(0.208)
ch_relationship	0.108	0.253	0.803***
	(0.150)	(0.145)	(0.236)
ch_health	0.176	0.155	0.689***
	(0.152)	(0.148)	(0.204)
ch_family	0.130	0.220	0.735**
	(0.150)	(0.147)	(0.227)
ch_friends	-0.326	0.0218	-0.598**
	(0.178)	(0.154)	(0.232)
ch_neighbour	0.0202	0.0658	0.182
	(0.160)	(0.159)	(0.249)
cut1, Constant	-4.289***	-0.357	-3.028***
	(0.354)	(0.261)	(0.363)
cut2, Constant	-1.818***	1.254***	0.973**
	(0.281)	(0.264)	(0.346)
cut3, Constant	0.0761		
	(0.274)		
cut4, Constant	2.139***		
	(0.303)		
Observations	794	795	796
r2_p	0.0313	0.0168	0.245
p	5.98e-09	0.00311	1.26e-30

Table A.4a. BHPS: Regressions on key socio-economic variables (financial situation) without lags. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, **p<0.01, ***p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.0171**	0.00406	0.0212***	0
	(0.00554)	(0.00284)	(0.00510)	(.)
age	-0.0232***	0.00622***	-0.0170***	-0.0106
	(0.000929)	(0.000530)	(0.000832)	(0.0233)
age^2	0.000171***	-0.0000555***	0.000115***	0.00000305
	(0.00000925)	(0.00000532)	(0.00000823)	(0.0000181)
change_fin_situation	0.132***	0.0689***	0.201***	0.162***
	(0.00402)	(0.00309)	(0.00341)	(0.00292)
stayed unemployed	-0.204***	0.0353	-0.169***	-0.0996***
	(0.0263)	(0.0189)	(0.0209)	(0.0232)
stayed out of labor market	-0.0448***	0.00543	-0.0393***	0.0227**
	(0.00752)	(0.00430)	(0.00673)	(0.00876)
got employed	-0.00485	0.0612***	0.0563***	0.0821***
	(0.0132)	(0.0113)	(0.0100)	(0.00964)
got unemployed	-0.0861***	-0.0960***	-0.182***	-0.114***
	(0.0206)	(0.0174)	(0.0164)	(0.0148)
exited labor market	0.0469***	0.0472***	0.0941***	0.110***
	(0.0139)	(0.0116)	(0.0111)	(0.0103)
got married	0.205***	0.0513**	0.257***	0.122***
	(0.0193)	(0.0162)	(0.0146)	(0.0147)
marriage ended	-0.110***	-0.0228	-0.133***	-0.196***
	(0.0266)	(0.0242)	(0.0235)	(0.0187)
stayed married	0.0163**	-0.0209***	-0.00459	-0.116***
	(0.00621)	(0.00333)	(0.00571)	(0.0101)
Year dummies	yes	yes	yes	yes
Constant	0.507***	-0.290***	0.218***	0.327
	(0.0250)	(0.0166)	(0.0216)	(0.948)
Observations	114568	114568	114568	114568
R^2	0.046	0.008	0.107	0.049
Adjusted R^2	0.045	0.008	0.107	-0.175
AIC	286652.3	252794.3	209947.1	163799.3
BIC	286864.6	253006.5	210159.3	164001.9

Table A.4b. BHPS: Regressions on key socio-economic variables (income) without lags. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, **p<0.01, ***p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.0154**	0.00323	0.0186***	0
	(0.00568)	(0.00283)	(0.00533)	(.)
age	-0.0262***	0.00454***	-0.0216***	-0.00448
	(0.000940)	(0.000524)	(0.000855)	(0.0237)
age^2	0.000194***	-0.0000415***	0.000153***	0.0000305
	(0.00000937)	(0.00000527)	(0.00000845)	(0.0000183)
d_hh_inc_month	-0.000520	0.00310*	0.00258*	0.00232*
	(0.00156)	(0.00153)	(0.00103)	(0.00103)
stayed unemployed	-0.250***	0.0129	-0.237***	-0.172***
	(0.0265)	(0.0190)	(0.0215)	(0.0235)
stayed out of labor market	-0.0705***	-0.00797	-0.0784***	-0.0157
	(0.00761)	(0.00424)	(0.00692)	(0.00886)
got employed	0.0145	0.0695***	0.0840***	0.102***
	(0.0133)	(0.0113)	(0.0104)	(0.00978)
got unemployed	-0.163***	-0.132***	-0.295***	-0.209***
	(0.0205)	(0.0174)	(0.0170)	(0.0149)
exited labor market	-0.00545	0.0207	0.0152	0.0461***
	(0.0138)	(0.0116)	(0.0111)	(0.0104)
got married	0.207***	0.0524**	0.260***	0.123***
	(0.0193)	(0.0163)	(0.0148)	(0.0149)
marriage ended	-0.153***	-0.0401	-0.193***	-0.243***
	(0.0265)	(0.0241)	(0.0240)	(0.0189)
stayed married	0.0121	-0.0223***	-0.0102	-0.124***
	(0.00636)	(0.00332)	(0.00594)	(0.0103)
Year dummies	yes	yes	yes	yes
Constant	0.877***	-0.0953***	0.781***	0.388
	(0.0226)	(0.0142)	(0.0198)	(0.964)
Observations	114980	114980	114980	114980
R^2	0.035	0.004	0.064	0.018
Adjusted R^2	0.035	0.004	0.064	-0.214
AIC	288942.9	254197.4	216222.7	168243.2
BIC	289155.3	254409.7	216435.0	168445.9

Table A.5a. BHPS: Regressions on key socio-economic variables (financial situation) including only the three first years of the panel, making it comparable to the current time horizon of GESIS. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.00461	0.00857	0.0132	0
	(0.0126)	(0.00940)	(0.0104)	(.)
age	-0.0219***	0.00483**	-0.0171***	-0.0198
	(0.00232)	(0.00174)	(0.00183)	(0.0264)
age^2	0.000175***	-0.0000539**	0.000121***	-0.000124
	(0.0000229)	(0.0000174)	(0.0000177)	(0.000264)
change_fin_situation	**0.101*****	**0.0821*****	**0.184*****	**0.162*****
	(0.00998)	(0.00862)	(0.00799)	(0.0118)
L.change_fin_situation	**0.100*****	**-0.0591*****	**0.0411*****	**0.00365**
	(0.00963)	(0.00842)	(0.00750)	(0.0120)
stayed unemployed	-0.114	-0.000554	-0.115*	-0.177
	(0.0663)	(0.0577)	(0.0550)	(0.116)
stayed out of labor market	-0.0216	-0.00405	-0.0257	0.0146
	(0.0189)	(0.0143)	(0.0151)	(0.0684)
got employed	**0.0255**	**0.0917****	**0.117*****	**0.160****
	(0.0363)	(0.0325)	(0.0275)	(0.0560)
lag got employed	**0.0311**	**0.0190**	**0.0501***	**0.0248**
	(0.0316)	(0.0280)	(0.0240)	(0.0397)
got unemployed	**-0.0618**	**-0.0774**	**-0.139****	**-0.0947**
	(0.0547)	(0.0483)	(0.0463)	(0.0699)
lag got unemployed	**-0.0645**	**0.0118**	**-0.0527**	**-0.0619**
	(0.0572)	(0.0517)	(0.0439)	(0.0666)
exited labor market	**0.103****	**0.0278**	**0.131*****	**0.170*****
	(0.0359)	(0.0318)	(0.0289)	(0.0456)
lag exited labor market	**0.142*****	**-0.0633**	**0.0783****	**0.0225**
	(0.0378)	(0.0328)	(0.0271)	(0.0454)
got married	**0.237*****	**0.0352**	**0.272*****	**0.107**
	(0.0520)	(0.0449)	(0.0392)	(0.0694)
lag got married	**0.119***	**0.0570**	**0.176*****	**0.210*****
	(0.0511)	(0.0417)	(0.0380)	(0.0629)
marriage ended	**-0.156****	**-0.00368**	**-0.159****	**-0.300****
	(0.0593)	(0.0565)	(0.0544)	(0.0949)
lag marriage ended	**-0.0802**	**0.117***	**0.0370**	**-0.0799**
	(0.0621)	(0.0556)	(0.0477)	(0.0728)
stayed married	-0.0372*	0.00653	-0.0307*	-0.289**
	(0.0148)	(0.0112)	(0.0121)	(0.0992)
Year dummies	Yes (3 years)	Yes (3 years)	Yes (3 years)	Yes (3 years)
Constant	0.273***	-0.112*	0.161**	1.156
	(0.0620)	(0.0463)	(0.0510)	(0.653)
Observations	16426	16426	16426	16426
R^2	0.050	0.017	0.112	0.046
Adjusted R^2	0.049	0.016	0.111	-1.231
AIC	40450.0	36069.9	29755.0	12099.0
BIC	40604.2	36224.0	29909.1	12237.7

Table A.5b. BHPS: Regressions on key socio-economic variables (income) including only the three first years of the panel, making it comparable to the current time horizon of GESIS. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * $p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.00271	0.00892	0.0116	0
	(0.0128)	(0.00940)	(0.0107)	(.)
age	-0.0261***	0.00423*	-0.0219***	-0.0197
	(0.00232)	(0.00173)	(0.00186)	(0.0267)
age^2	0.000207***	-0.0000477**	0.000160***	-0.000136
	(0.0000230)	(0.0000173)	(0.0000180)	(0.000268)
d_hh_inc_month	0.00252	0.00174	0.00426	0.00206
	(0.00394)	(0.00377)	(0.00287)	(0.00380)
L.d_hh_inc_month	0.000442	0.000271	0.000713	-0.00325
	(0.00401)	(0.00347)	(0.00290)	(0.00466)
stayed unemployed	-0.187**	-0.0107	-0.198***	-0.216
	(0.0659)	(0.0585)	(0.0563)	(0.116)
stayed out of labor market	-0.0645***	-0.0113	-0.0758***	-0.0478
	(0.0189)	(0.0141)	(0.0153)	(0.0690)
got employed	0.00764	0.121***	0.129***	0.180**
	(0.0364)	(0.0327)	(0.0285)	(0.0566)
lag got employed	0.0548	0.0121	0.0668**	0.0393
	(0.0321)	(0.0282)	(0.0249)	(0.0403)
got unemployed	-0.138*	-0.105*	-0.243***	-0.181*
	(0.0545)	(0.0489)	(0.0480)	(0.0705)
lag got unemployed	-0.0920	0.0306	-0.0614	-0.0518
	(0.0583)	(0.0523)	(0.0458)	(0.0675)
exited labor market	0.0468	0.00380	0.0506	0.103*
	(0.0358)	(0.0316)	(0.0293)	(0.0458)
lag exited labor market	0.112**	-0.0514	0.0604*	0.0223
	(0.0379)	(0.0326)	(0.0273)	(0.0457)
got married	0.253***	0.0345	0.288***	0.100
	(0.0522)	(0.0452)	(0.0400)	(0.0699)
lag got married	0.116*	0.0510	0.167***	0.180**
	(0.0513)	(0.0419)	(0.0389)	(0.0637)
marriage ended	-0.195**	-0.0297	-0.225***	-0.373***
	(0.0601)	(0.0565)	(0.0570)	(0.0960)
lag marriage ended	-0.123	0.123*	0.0000408	-0.122
	(0.0630)	(0.0553)	(0.0492)	(0.0736)
stayed married	-0.0351*	0.00698	-0.0281*	-0.280**
	(0.0150)	(0.0111)	(0.0124)	(0.100)
Year dummies	Yes (3 years)	Yes (3 years)	Yes (3 years)	Yes (3 years)
Constant	0.824***	-0.0509	0.773***	1.543*
	(0.0528)	(0.0391)	(0.0426)	(0.658)
Observations	16506	16506	16506	16506
R^2	0.034	0.010	0.067	0.015
Adjusted R^2	0.033	0.009	0.066	-1.299
AIC	40931.4	36359.6	30748.7	12757.2
BIC	41085.6	36513.8	30902.9	12896.0

Table A.6. GESIS: Regressions on key socio-economic variables (income). Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, **p<0.01, ***p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	-0.0374	0.00256	-0.0349	0
	(0.0318)	(0.0236)	(0.0266)	(.)
age	-0.0280*	0.00232	-0.0257**	-0.0796
	(0.0114)	(0.00879)	(0.00939)	(0.102)
age^2	0.000188	-0.0000316	0.000156	0.000643
	(0.000116)	(0.0000890)	(0.0000960)	(0.000951)
d_hh_inc	**0.0318**	**0.00866**	**0.0405****	**0.0403***
	(0.0199)	(0.0170)	(0.0140)	(0.0186)
L.d_hh_inc	**0.0215**	**0.00218**	**0.0236**	**0.0154**
	(0.0173)	(0.0147)	(0.0133)	(0.0170)
stayed unemployed	-0.106	-0.516**	-0.623**	-0.440
	(0.277)	(0.177)	(0.214)	(0.430)
stayed out of labor market	-0.179**	0.0368	-0.142**	-0.0884
	(0.0567)	(0.0428)	(0.0461)	(0.179)
got employed	**0.0480**	**-0.0862**	**-0.0382**	**-0.212**
	(0.0960)	(0.0784)	(0.0726)	(0.144)
lag got employed	**0.0713**	**0.0276**	**0.0989**	**0.0864**
	(0.0777)	(0.0667)	(0.0548)	(0.0935)
got unemployed	**-0.343**	**0.0343**	**-0.309***	**-0.454***
	(0.185)	(0.149)	(0.150)	(0.217)
lag got unemployed	**-0.0303**	**0.0523**	**0.0221**	**-0.0401**
	(0.222)	(0.179)	(0.159)	(0.241)
exited labor market	**-0.0492**	**0.0432**	**-0.00595**	**-0.0195**
	(0.0859)	(0.0666)	(0.0644)	(0.106)
lag exited labor market	**0.116**	**-0.00752**	**0.108**	**0.0726**
	(0.0845)	(0.0699)	(0.0614)	(0.110)
got married	**0.00865**	**0.0611**	**0.0698**	**-0.185**
	(0.149)	(0.125)	(0.108)	(0.207)
lag got married	**0.0222**	**0.0794**	**0.102**	**0.0267**
	(0.114)	(0.0995)	(0.0846)	(0.134)
marriage ended	**0.0473**	**0.0365**	**0.0838**	**0.0593**
	(0.166)	(0.128)	(0.120)	(0.225)
lag marriage ended	**-0.00834**	**0.0705**	**0.0621**	**0.158**
	(0.137)	(0.108)	(0.104)	(0.165)
stayed married	-0.0319	-0.0120	-0.0440	-0.301
	(0.0410)	(0.0323)	(0.0335)	(0.254)
Year dummies	yes	yes	yes	yes
Constant	1.308***	-0.0595	1.248***	2.827
	(0.268)	(0.210)	(0.220)	(2.703)
Observations	3442	3442	3442	3442
R^2	0.032	0.005	0.064	0.018
Adjusted R^2	0.027	-0.001	0.059	-1.535
AIC	9486.4	8260.7	7301.9	2899.1
BIC	9609.3	8383.6	7424.8	3009.7

Table A.7. BHPS: Regressions on key socio-economic variables (financial situation). Variable description in Appendix A.1. Results from **ordered logistic regressions** including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) deviation	(2) signdLM	(3) signRM
female	0.0366**	0.00695	0.0758***
	(0.0132)	(0.0133)	(0.0142)
age	-0.0469***	0.0123***	-0.0586***
	(0.00236)	(0.00237)	(0.00253)
age^2	0.000354***	-0.000119***	0.000410***
	(0.0000233)	(0.0000235)	(0.0000248)
change_fin_situation	**0.226*****	**0.222*****	**0.655*****
	(0.0100)	(0.0101)	(0.0113)
L.change_fin_situation	**0.230*****	**-0.156*****	**0.154*****
	(0.0101)	(0.0101)	(0.0110)
stayed unemployed	-0.271***	-0.0340	-0.435***
	(0.0728)	(0.0742)	(0.0782)
stayed out of labor market	-0.0821***	0.00896	-0.121***
	(0.0190)	(0.0191)	(0.0204)
got employed	**0.113****	**0.115****	**0.378*****
	(0.0380)	(0.0384)	(0.0417)
lag got employed	**0.0417**	**0.0120**	**0.0910***
	(0.0329)	(0.0330)	(0.0358)
got unemployed	**-0.198*****	**-0.264*****	**-0.648*****
	(0.0514)	(0.0527)	(0.0570)
lag got unemployed	**-0.356*****	**0.147***	**-0.353*****
	(0.0573)	(0.0584)	(0.0626)
exited labor market	0.166***	0.111**	0.419***
	(0.0356)	(0.0360)	(0.0397)
lag exited labor market	0.246***	-0.103**	0.234***
	(0.0365)	(0.0365)	(0.0395)
got married	**0.427*****	**0.139****	**0.861*****
	(0.0509)	(0.0507)	(0.0566)
lag got married	**0.434*****	**-0.0792**	**0.547*****
	(0.0515)	(0.0512)	(0.0562)
marriage ended	**-0.308*****	**-0.0668**	**-0.609*****
	(0.0664)	(0.0694)	(0.0763)
lag marriage ended	**-0.112**	**0.249*****	**0.155***
	(0.0643)	(0.0654)	(0.0699)
stayed married	-0.0000336	-0.0234	-0.0439**
	(0.0151)	(0.0152)	(0.0163)
Year dummies	yes	yes	yes
Cut1, Constant	-3.975***	-0.720***	-1.971***
	(0.0695)	(0.0668)	(0.0721)
Cut 2,Constant	-1.592***	1.339***	1.044***
	(0.0666)	(0.0669)	(0.0720)
Cut3, Constant	0.505***		
	(0.0664)		
Cut4, Constant	2.889***		
	(0.0680)		
Observations	83533	83533	83533

Percent correctly predicted		0.414	0.520
Log likelihood	-103336.9	-87923.9	-74544.3
chi2	4253.6	1004.5	10233.9
chi2type	LR	LR	LR
df_m	25	25	25
p	0	6.47e-196	0

Table A.8. Deviation between measures on direction of change and trends in financial situation using leads instead of lags to capture expected trends (BHPS). Description of variables given in Appendix A.1. Results from pooled OLS regression including year fixed effects. Standard errors in parentheses. * p<0.05, **p<0.01, ***p<0.001.

	(1) deviation	
female	0.0186**	(0.00700)
age	-0.0208***	(0.00126)
age^2	0.000160***	(0.0000125)
financial deterioration	-0.0491***	(0.0114)
Lag financial deterioration	-0.0985***	(0.00930)
Lag financial improvement	0.0977***	(0.00879)
financial improvement	0.132***	(0.0105)
negative trend (lead)	**-0.0219**	**(0.0152)**
positive trend (lead)	**0.0329***	**(0.0136)**
stayed unemployed	-0.0891*	(0.0388)
stayed out of labor market	-0.0412***	(0.0100)
got employed	0.0369	(0.0199)
got unemployed	-0.0893**	(0.0275)
exited labor market	0.0562**	(0.0187)
lag got employed	0.0191	(0.0174)
lag got unemployed	-0.138***	(0.0297)
lag exited labor market	0.114***	(0.0191)
got married	0.212***	(0.0264)
marriage ended	-0.162***	(0.0340)
stayed married	0.00147	(0.00797)
lag got married	0.173***	(0.0269)
lag marriage ended	-0.0389	(0.0335)
Year dummies	yes	
Constant	0.621***	(0.0312)
Observations	60133	
R^2	0.049	
Adjusted R^2	0.049	
AIC	148716.0	
BIC	148968.1	

Table A.9. BHPS: Regressions on key socio-economic variables (log of income). Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, **p<0.01, ***p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.0156*	0.00232	0.0179**	0
	(0.00637)	(0.00338)	(0.00594)	(.)
age	-0.0252***	0.00403***	-0.0212***	-0.0133
	(0.00111)	(0.000648)	(0.000997)	(0.0331)
age^2	0.000191***	-0.0000402***	0.000151***	0.0000181
	(0.0000109)	(0.00000642)	(0.00000968)	(0.0000235)
d_log_hh_inc_month	**0.00728**	**0.00867**	**0.0159*****	**0.0173*****
	(0.00545)	(0.00493)	(0.00431)	(0.00398)
L.d_log_hh_inc_month	**0.0126***	**-0.00974***	**0.00283**	**0.00302**
	(0.00552)	(0.00490)	(0.00424)	(0.00400)
stayed unemployed	-0.173***	-0.0282	-0.201***	-0.165***
	(0.0339)	(0.0251)	(0.0290)	(0.0336)
stayed out of labor market	-0.0827***	-0.0000113	-0.0827***	-0.00344
	(0.00902)	(0.00517)	(0.00817)	(0.0123)
got employed	**0.0389***	**0.0694*****	**0.108*****	**0.135*****
	(0.0178)	(0.0151)	(0.0141)	(0.0143)
lag got employed	**0.0360***	**0.00525**	**0.0412*****	**0.0483*****
	(0.0155)	(0.0131)	(0.0121)	(0.0118)
got unemployed	**-0.164*****	**-0.124*****	**-0.289*****	**-0.205*****
	(0.0254)	(0.0217)	(0.0211)	(0.0191)
lag got unemployed	**-0.194*****	**0.0785*****	**-0.115*****	**-0.0621****
	(0.0280)	(0.0238)	(0.0216)	(0.0205)
exited labor market	**0.0210**	**0.0173**	**0.0383****	**0.0757*****
	(0.0168)	(0.0142)	(0.0135)	(0.0129)
lag exited labor market	**0.0835*****	**-0.0373****	**0.0462*****	**0.0295***
	(0.0171)	(0.0143)	(0.0129)	(0.0129)
got married	**0.200*****	**0.0527****	**0.253*****	**0.115*****
	(0.0233)	(0.0195)	(0.0177)	(0.0183)
lag got married	**0.198*****	**-0.0365**	**0.162*****	**0.153*****
	(0.0237)	(0.0199)	(0.0184)	(0.0181)
marriage ended	**-0.186*****	**-0.0389**	**-0.225*****	**-0.301*****
	(0.0317)	(0.0289)	(0.0284)	(0.0238)
lag marriage ended	**-0.0823****	**0.0970*****	**0.0147**	**-0.0397**
	(0.0316)	(0.0266)	(0.0236)	(0.0224)
stayed married	-0.00820	-0.00946*	-0.0177**	-0.186***
	(0.00729)	(0.00404)	(0.00674)	(0.0140)
Year dummies	yes	yes	yes	yes
Constant	0.808***	-0.0483**	0.760***	0.795
	(0.0267)	(0.0168)	(0.0236)	(1.391)
Observations	82908	82908	82908	82908
R^2	0.036	0.005	0.064	0.020
Adjusted R^2	0.035	0.005	0.064	-0.262
AIC	206721.6	182117.6	154677.8	116036.7
BIC	206964.0	182360.0	154920.3	116269.8

Table A.10. BHPS: Regressions on key socio-economic variables (financial situation) excluding observations where the current LM is either 1 or 7. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.0222**	-0.00645	0.0158**	0
	(0.00703)	(0.00421)	(0.00596)	(.)
age	-0.0243***	0.0102***	-0.0141***	-0.0131
	(0.00129)	(0.000833)	(0.00103)	(0.0374)
age^2	0.000200***	-0.000115***	0.0000844***	-0.0000260
	(0.0000129)	(0.00000849)	(0.0000100)	(0.0000271)
change_fin_situation	**0.115*****	**0.0746*****	**0.189*****	**0.170*****
	(0.00513)	(0.00425)	(0.00411)	(0.00385)
L.change_fin_situation	**0.111*****	**-0.0736*****	**0.0376*****	**0.0161*****
	(0.00507)	(0.00426)	(0.00379)	(0.00385)
stayed unemployed	-0.0968*	-0.0519	-0.149***	-0.0863*
	(0.0377)	(0.0281)	(0.0289)	(0.0369)
stayed out of labor market	-0.0362***	-0.00166	-0.0379***	0.0428**
	(0.0103)	(0.00660)	(0.00830)	(0.0137)
got employed	**0.0605****	**0.0303**	**0.0908*****	**0.124*****
	(0.0191)	(0.0160)	(0.0147)	(0.0156)
lag got employed	**0.0218**	**0.00142**	**0.0232**	**0.0405****
	(0.0163)	(0.0137)	(0.0124)	(0.0127)
got unemployed	**-0.0978*****	**-0.0913*****	**-0.189*****	**-0.124*****
	(0.0274)	(0.0228)	(0.0212)	(0.0206)
lag got unemployed	**-0.169*****	**0.0756****	**-0.0933*****	**-0.0542***
	(0.0302)	(0.0253)	(0.0223)	(0.0224)
exited labor market	**0.0946*****	**0.0278**	**0.122*****	**0.153*****
	(0.0187)	(0.0155)	(0.0145)	(0.0144)
lag exited labor market	**0.117*****	**-0.0414****	**0.0754*****	**0.0424****
	(0.0191)	(0.0157)	(0.0137)	(0.0144)
got married	**0.205*****	**0.0206**	**0.226*****	**0.104*****
	(0.0255)	(0.0209)	(0.0190)	(0.0201)
lag got married	**0.192*****	**-0.0437***	**0.149*****	**0.148*****
	(0.0257)	(0.0213)	(0.0194)	(0.0199)
marriage ended	**-0.139*****	**-0.0115**	**-0.151*****	**-0.235*****
	(0.0338)	(0.0304)	(0.0298)	(0.0260)
lag marriage ended	**-0.0494**	**0.0971*****	**0.0477**	**-0.00475**
	(0.0341)	(0.0290)	(0.0247)	(0.0246)
stayed married	0.00380	-0.0292***	-0.0254***	-0.185***
	(0.00808)	(0.00497)	(0.00681)	(0.0152)
Year dummies	yes	yes	yes	yes
Constant	0.276***	-0.197***	0.0785**	0.422
	(0.0350)	(0.0236)	(0.0287)	(1.517)
Observations	71449	71449	71449	71449
R^2	0.051	0.016	0.110	0.055
Adjusted R^2	0.050	0.015	0.110	-0.252
AIC	180607.9	157794.0	131606.7	99764.0
BIC	180846.5	158032.6	131845.3	99993.4

Table A.11. BHPS: Regressions on key socio-economic variables (financial situation) excluding observations where the lag of LM is either 1 or 7. Variable description in Appendix A.1. Results from pooled OLS and fixed effects regressions including year fixed effects. Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001.

	(1) deviation OLS	(2) signdLM OLS	(3) signRM OLS	(4) signRM FE
female	0.00286	0.0159***	0.0187**	0
	(0.00652)	(0.00420)	(0.00604)	(.)
age	-0.0143***	-0.000441	-0.0147***	-0.0248
	(0.00117)	(0.000841)	(0.00105)	(0.0362)
age^2	0.0000730***	0.0000183*	0.0000914***	-0.0000126
	(0.0000116)	(0.00000862)	(0.0000102)	(0.0000274)
change_fin_situation	**0.0956*****	**0.0975*****	**0.193*****	**0.171*****
	(0.00494)	(0.00428)	(0.00414)	(0.00386)
L.change_fin_situation	**0.0866*****	**-0.0475*****	**0.0391*****	**0.0172*****
	(0.00484)	(0.00426)	(0.00382)	(0.00387)
stayed unemployed	-0.161***	0.0125	-0.148***	-0.0817*
	(0.0359)	(0.0288)	(0.0289)	(0.0375)
stayed out of labor market	-0.0489***	0.00945	-0.0395***	0.0454***
	(0.00925)	(0.00655)	(0.00845)	(0.0138)
got employed	**0.0396***	**0.0587*****	**0.0983*****	**0.128*****
	(0.0185)	(0.0161)	(0.0148)	(0.0157)
lag got employed	**0.00345**	**0.0144**	**0.0179**	**0.0368****
	(0.0157)	(0.0138)	(0.0124)	(0.0128)
got unemployed	**-0.100*****	**-0.100*****	**-0.201*****	**-0.126*****
	(0.0257)	(0.0231)	(0.0215)	(0.0208)
lag got unemployed	**-0.139*****	**0.0428**	**-0.0957*****	**-0.0660****
	(0.0287)	(0.0253)	(0.0225)	(0.0225)
exited labor market	**0.0710*****	**0.0543*****	**0.125*****	**0.164*****
	(0.0178)	(0.0155)	(0.0146)	(0.0143)
lag exited labor market	**0.0922*****	**-0.0278**	**0.0644*****	**0.0386****
	(0.0182)	(0.0158)	(0.0141)	(0.0146)
got married	**0.150*****	**0.0823*****	**0.232*****	**0.109*****
	(0.0240)	(0.0207)	(0.0190)	(0.0201)
lag got married	**0.170*****	**-0.0194**	**0.151*****	**0.155*****
	(0.0250)	(0.0217)	(0.0199)	(0.0201)
marriage ended	**-0.132*****	**-0.00616**	**-0.138*****	**-0.222*****
	(0.0330)	(0.0310)	(0.0303)	(0.0264)
lag marriage ended	**-0.0529**	**0.0974*****	**0.0445**	**-0.0167**
	(0.0322)	(0.0286)	(0.0249)	(0.0248)
stayed married	-0.0299***	0.00964	-0.0203**	-0.182***
	(0.00749)	(0.00495)	(0.00692)	(0.0153)
Year dummies	yes	yes	yes	yes
Constant	0.122***	-0.0335	0.0888**	0.870
	(0.0328)	(0.0238)	(0.0292)	(1.464)
Observations	70759	70759	70759	70759
R^2	0.057	0.014	0.111	0.055
Adjusted R^2	0.057	0.014	0.111	-0.254
AIC	172994.9	156368.7	131099.4	98790.1
BIC	173233.3	156607.1	131337.8	99019.3