Journal of Economic Behavior & Organization 81 (2012) 556–570

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Frame-of-reference bias in subjective welfare☆

Kathleen Beegle, Kristen Himelein, Martin Ravallion*

Development Research Group, World Bank, 1818 H Street NW, Washington, DC 20433, USA

ARTICLE INFO

Article history:
Received 23 October 2009
Received in revised form 28 July 2011
Accepted 29 July 2011
Available online 6 August 2011

JEL classification:
C81
D60
I32

Keywords:
Subjective welfare
Scales
Differential item functioning
Vignettes

ABSTRACT

The inferences drawn from the most widely used regression models of subjective welfare are subject to a "frame-of-reference bias," stemming from non-ignorable heterogeneity in subjective scales, such as what it means to be "rich" or "poor." To test for this bias, respondents in Tajikistan were asked to rank the economic status of theoretical vignette households, as well as their own. Respondents are found to hold diverse scales, but there is very little bias in either the economic gradient of subjective welfare or most other coefficients of interest. These results provide a foundation for standard survey methods and regression specifications for subjective welfare data.

© 2011 Published by Elsevier B.V.

1. Introduction

Subjective measures of welfare are widely used in psychological and social sciences, including economics.[1] Typically a survey respondent is asked to rate their "economic welfare," "satisfaction with life" or "happiness" on an ordinal scale.[2] A large literature has studied the covariates of answers to such subjective welfare questions.[3] The most common method is a regression (typically an ordered probit) of the survey responses on individual and household characteristics, including age, gender, income, education, employment status and household demographics. Such regressions have been used to assess the welfare effects of, *inter alia*, own income ("does money buy you happiness?"), employment ("does unemployment lower welfare at given income?") and relative position ("do people care about relative deprivation?").

☆ This paper would not have been possible without the work of the Tajikistan Living Standards Survey Team, including Diane Steele, Sasun Tsirunyan, Vladimir Kolchin, Farhod Khamidov, Oleksiy Ivaschenko and staff of the National Statistics Committee of Tajikistan (Goskomstat). We thank Gero Carletto, seminar participants at IFPRI and the AEA annual meetings, and the Journal's anonymous referees for very useful comments. All views are those of the authors and do not reflect the views of the World Bank or its member countries.

* Corresponding author.
E-mail address: mravallion@worldbank.org (M. Ravallion).

[1] A cross-country compendium of the questions asked and a summary of the answers can be found in Veenhoven et al. (1993). The literature in economics is reviewed by Di Tella and MacCulloch (2006). The psychological literature on subjective welfare is reviewed in Diener et al. (1999) and Furnham and Argyle (1998). Since 2000, a scholarly journal has been devoted to the scientific study of subjective welfare, namely the *Journal of Happiness Studies*.

[2] An alternative approach is to ask what level of income is needed to attain a given position on a ladder, such as not being "poor." This is the "Leyden method" devised by Van Praag (1968). While we do not use this type of data here, the same concerns about bias arise in the Leyden method.

[3] Examples include Van Praag (1968), van de Stadt et al. (1985), Clark and Oswald (1994, 1996), Kapteyn et al. (1998), Easterlin (1995), Oswald (1997), Winkelmann and Winkelmann (1998), Pradhan and Ravallion (2000), McBride (2001), Ravallion and Lokshin (2001, 2002, 2010) Senik (2004), Luttmer (2005), Ferrer-i-Carbonell (2005) and Fafchamps and Shipli (2009).

0167-2681/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.jebo.2011.07.020

Measurement problems clearly confound interpersonal comparisons of welfare using subjective data (as they do with objective data). A long standing concern about subjective welfare (and health) questions is that different people may use different criteria for scaling their welfare – that they have different ideas about what it means to be "rich" or "poor," or what it means to be "satisfied" or not with one's life.[4] Such latent heterogeneity in scales has often been seen as invalidating the use of subjective welfare questions for inferring utility.[5] This recognizes explicitly that latent heterogeneity in factors that are essentially irrelevant to welfare but influence responses to subjective welfare questions casts doubt on the implied interpersonal comparisons of welfare from a subjective measure. In these circumstances, heterogeneity in scales will translate into corresponding differences in subjective welfare at any given level of objective welfare or other relevant covariates. This will, of course, reduce the explanatory power of the regression models for subjective welfare. However, if such heterogeneity was purely random then it would not invalidate inferences from such regressions (at least for linear models). Thus it has been argued that, while inter-personal welfare comparisons are invalidated by heterogeneous scales, the regressions are likely to be robust to such heterogeneity.[6]

However, that claim is questionable. People will answer subjective questions in surveys relative to their personal *frame-of-reference*, which depends on the respondent's own knowledge and experience, and therefore is likely to vary systematically with the characteristics of that person, including objective measures of economic welfare. For example, it can be conjectured that people living in poor areas of a developing country tend to have a more limited knowledge of the full range of levels of living found in the society as a whole. Someone living in a poor, remote area who has only infrequently left the village and gone no further than the district town may rate her welfare higher than someone with the same real income who lives in a city and sees far greater affluence around her. Similarly, it can be conjectured that relatively well-off people are often unaware of how poor some people are, and may thus rate their own welfare lower on the scale.

When this effect is present and it impacts on the coefficients of interest, we shall say that there is a *frame-of-reference bias* (FORB). The potential for FORB raises concerns about the (enumerable) regression models found in past literature. Consider, for example, the many papers that have used subjective welfare regressions to test for reference-group effects, such as whether higher neighbors' income makes one feel worse off through perceptions of relative deprivation. It seems likely that the same reference group also influences the respondents' interpretation of scales used in subjective questions. The reference group acts as both the comparator in assessing relative position and a key element of the information set used by respondents when interpreting the scales.

To give a sharp illustration of the problem in the present context, suppose that three people are asked to rate their own welfare on a scale of 1, 2 and 3, with "1" the poorest. For the sake of the argument, let us also suppose that "wealth" is the only parameter for defining "welfare." Wealth is normalized to be in the $[0, 1]$ interval. The first person is relatively poor, the second has the overall modal wealth, denoted M, and the third is relatively rich. The frame-of-reference effect implies that the poor person is only aware of levels of wealth in the interval $[0, M]$ while the rich person is only aware of those in the interval $[M, 1]$. Furthermore, the poor person has a wealth somewhere near the middle of the $[0, M]$ interval while the rich person has wealth near the middle of $[M, 1]$. This suggests a potentially large downward bias in the regression coefficient of subjective welfare on wealth. Indeed, all three responses to the subjective welfare question may well be "2". This type of bias may also be present when using other concepts of well-being such as life satisfaction or happiness.

There are antecedents to the idea of frame-of-reference effects in the literature. It is a well-established idea that people assess their welfare relative to some "comparison group" such as neighbors or co-workers.[7] This argument has emphasized relativist welfare comparisons. It has also been argued that reference groups play an important role in expectations formation.[8] It is a small step from these ideas to the proposition that survey respondents answer questions with reference to their immediate experiences and that this may well be highly localized in some relevant social or geographic dimensions. In the specific context of subjective welfare measurement, Seidl (1994) argues that Van Praag's (1968) method of calibrating a utility function to subjective data confounds the underlying utility function with a "welfare evaluation function" whereby (for example) "respondents belonging to the middle income strata can evaluate the welfare of the middle income range more precisely" (Seidl, 1994, p. 1653).[9] The frame-of-reference effect can also be interpreted as a special case of what is termed "differential item functioning" (DIF) in the literature on educational testing. In this literature, DIF exists if students with equal latent ability have different probabilities of giving a correct answer.[10]

[4] While this paper focuses on heterogeneity in scales, there are other concerns. For example, Conti and Pudney (2008) find that minor re-designs in questions on satisfaction of life/work led to large changes in answers, particularly for women. Moreover, they conclude that these distortions in survey responses influence research findings with respect to understanding women's job satisfaction.

[5] For a recent discussion, see Di Tella and MacCulloch (2006).

[6] For example, Frey and Stutzer (2002) note the possibility of heterogeneity in the scales used in self-reported welfare questions but claim that this does not invalidate regression models for such data.

[7] See, for example, Frank (1997). Frey and Stutzer (2002) provide a useful overview of evidence related to comparison-group effects.

[8] This was argued by Hirshman (1973). For more recent discussions and evidence, see Ravallion and Lokshin (2000) and Senik (2004).

[9] See the response by Van Praag and Kapteyn (1994) to this critique.

[10] For an overview of the history and methods of addressing DIF, see Angoff (1993). King et al. (2004) note that they apply a slightly different meaning to the term DIF that what is meant when it has been traditionally applied to education testing. We also feel that DIF (as defined in education work) is not really the appropriate term for the concept we are studying. We interpret FORB as a special case of DIF, and in the same way that King et al. (2004) apply the term DIF in their work.

Table 1
Plausible rankings of vignettes households by households with different objective wealth levels when there is a frame-of-reference bias.

	Respondent's wealth		
	Poor	Middle	Rich
Poor vignette household	2	1	1
Middle vignette household	3	2	1
Rich vignette household	3	3	2

This paper explores the role of the frame-of-reference effect in influencing self-reported economic status and offers various tests for FORB, to see whether the standard regressions for subjective welfare found in the literature are contaminated by this effect and to assess the consequences of corrections for this bias to the types of results reported in the literature. Toward these aims, we adapt the vignette methodology that has been used in some recent studies of subjective data on health status, political efficacy and job satisfaction. King et al. (2004) and King and Wand (2007) designed vignettes to establish common points on the heterogeneous reference scales regarding political efficacy in China and Mexico. Kristensen and Johansson (2008) used vignettes in anchoring subjective scales for job satisfaction. Kapteyn et al. (2008) use vignettes to compare life satisfaction between respondents in the U.S. and the Netherlands. Bago d'Uva et al. (2008) used them for correcting self-assessed health data for reporting bias. Vignettes have also been used in testing the competence of doctors (Das et al., 2008). To our knowledge, this is the first work to use vignettes for anchoring self-reported economic status.

The paper provides four tests for FORB where our subjective welfare measure is self-reported economic status. In the first we simply test whether vignette responses are correlated with the regressors typically used in subjective welfare analysis, including objective welfare measures. If everyone has essentially the same idea of what it means to be "poor" or "rich" then we would not expect to find significant correlations between the vignette responses and the covariates used to explain subjective welfare. Under certain forms of FORB, we may find a negative economic gradient in the vignette responses. Consider again the three people in the example above and now suppose that these three people are the vignettes. A plausible set of responses to the vignette questions is given in Table 1, indicating that the wealthier the respondent, the lower (or at least not higher) the rating of each vignette.

So Test 1 provides a very direct test for FORB in subjective welfare regressions, under which we would look for a negative wealth gradient in each set of vignette responses as well as correlation with other household characteristics.

We then provide three further tests that can help quantify the extent of any bias due to heterogeneity in scales by providing a method of correcting subjective welfare regressions for the presence of individual-specific scales or standards. In Test 2, the vignette responses enter as control variables in standard regressions, to purge the error term of the heterogeneity in scales used, under the assumption that differences in vignette responses are solely attributable to differences in the personal scales used. In the third test, an alternative to Test 2, we use the re-scaling method proposed by King et al. (2004) in which subjective welfare measures are re-calibrated for consistently across respondents based on the vignette responses. In our fourth test, we jointly model the thresholds that define the personal scales and own subjective welfare, under the identifying assumption that common scales are used for own welfare and the vignettes. We employ the estimator for this type of data proposed by King et al. (2004).

We use data from Tajikistan. For the purpose of this paper, we specially designed and included a set of vignettes in the 2007 national household survey for Tajikistan. Respondents were asked to place themselves on a subjective welfare ladder with six rungs. Later in the questionnaire they were asked to place four vignette households on this ladder, and finally to (again) place their own household on the same ladder after ranking the vignettes. Unlike some past research using vignettes, the subjective welfare questions were asked both before and after the vignettes. The second (post-vignette) subjective reporting enables the re-scaling of subjective responses, since respondents place their household in reference to the vignettes themselves.

We begin with a description of our data in Section 2. Section 3 presents our results on FORB while Section 4 concludes.

2. Data

Tajikistan is one of the poorest and most isolated of the countries in the former Soviet Union, with a per capita income of 430 USD in 2007. Its mountainous location and deteriorating physical infrastructure make transportation difficult and leave certain parts of the country completely isolated during winter months. About one-third of households are located in the capital Dushanbe or other urban areas. Overall, 47 percent of the population lived below the country's poverty line (World Bank, 2008), and 21 percent in 2004 lived below the World Bank's international poverty line of $1.25 a day at 2005 purchasing power parity (estimated using PovcalNet at iresearch.worldbank.org/PovcalNet/jsp/index.jsp). From 2003 to 2007, there was considerable pro-poor growth in Tajikistan, with large income gains experienced in the poorest region of Gbao, and moderate gains in the wealthiest region of Dushanbe. By 2007, then, mean income levels were not widely different across the five regions. Gbao and Dushanbe had the same poverty rate. Despite lack of difference in mean incomes, poverty rates in 2007 (which capture the both income distribution and level) were notably higher in Sogd (69%) compared to other regions (43–49%). Thirty percent of the population resides in Sogd, while Sogd is home to 38 percent of the poor.

Table 2
Pre-vignette and post-vignette subjective welfare rankings.

Pre-vignette	Post-vignette						Total
	1 poorest	2	3	4	5	6 richest	
1 poorest	247	64	23	3	4	0	341
2	72	933	232	53	7	5	1302
3	34	242	1735	223	37	4	2275
4	4	33	112	592	34	1	776
5	2	0	2	11	51	1	67
6 richest	0	0	1	1	4	4	10
Total	359	1272	2105	883	137	15	4771

2.1. Survey data and instrument design

The 2007 Tajikistan Living Standards Measurement Survey (TLSMS) surveyed a random sample of 4860 households in September and November 2007. The sample is designed to be representative at the national level, urban and rural levels, and at the *oblast* (administrative region) level. Data were collected in two visits, with the subjective welfare modules being asked in the first visit. Summary statistics for the variables used can be found in Appendix 1.

In addition to the standard questions on household characteristics and expenditures (including imputed values of self-produced and consumed food) common to multi-topic household questionnaires, subjective welfare data were collected at two different points in the questionnaire. In the Subjective Poverty and Food Security Module, respondents were asked: "*Imagine a 6-step ladder where on the bottom, the first step, stand the poorest people, and the highest step, the sixth, stand the rich. On which step are you today?*" In a later section of the questionnaire, respondents were asked to place four vignettes of hypothetical households on a six-step ladder and then to place themselves on the same scale. The latter is the same question asked in the earlier portion of the questionnaire. The actual vignettes from the questionnaire (translated into English) are given in Appendix 2.

We developed the vignettes for this experiment in consultation with local counterparts. The vignettes were designed to capture representative snapshots of various levels of welfare in Tajikistan. Characteristics incorporated in the vignettes include land holdings, education, diet, clothing, and the ability to heat the home during the winter. The vignettes were developed in a clear, expected hierarchy with respect to welfare, with all aspects of socio-economic status increasing monotonically. This structure was used to minimize the effects of multi-dimensionality. Multi-dimensionality can lead to the perverse sequencing of the vignettes with respect to the intended ordering if respondents place different values on various characteristics contained in the vignettes with respect to overall welfare (King et al., 2004). This does not seem to be a major concern for our vignettes: of the 4860 households in the sample, only 89 have any instances of an "incorrect" ordering. The most common characteristic of respondents who perversely order the vignettes is a low level of education of the household head; the Working Paper version contains the full results from a probit estimation of the correlates of perverse ordering (Beegle et al., 2009). These 89 households are excluded from our analysis.

2.2. Pre and post vignettes

If responses are influenced by heterogeneous scales, it might be expected that subjective welfare responses will be affected by familiarity with the vignettes. By asking the vignette questions, the survey may focus the respondent to think about, and possibly revise, their own scale used to report their self-assessment welfare (Hopkins and King, 2010).

In Table 2 we compare the pre- and post-vignette responses to the household's self-assessment welfare. On average, respondents place themselves between steps 2 and 3 in both the pre-vignette and post-vignette question, though the mean was slightly higher post-vignette (2.75 vs. 2.80). Most respondents place themselves in the same position in the pre- and post-vignette self-reporting, although fully 25 percent change their position. Of those who change their position, 57 percent adjust upwards and 43 percent downwards. Generally respondents give similar responses; 82 percent of those changing their position move only one step up or down. Among the 4771 households, only 39 respondents switch dramatically from the "rich" category (steps 4–6) when asked before the vignettes to the "poor" (steps 1 and 2) category when asked after the vignettes. In the opposite direction, 72 report themselves in the "poor" categories pre-vignette and in the "rich" category post-vignette.

We explore whether specific types of households are more inclined to change pre- and post-vignette reports, but we do not observe any striking differences in such changes across different categories of households (Table 3). For example, comparing those classified as poor using objective measures (per capita households expenditure below the poverty line), the poor report a lower subjective welfare score than the non-poor. Both groups adjust their scores up slightly following the vignettes, but neither experiences a marked transformation. We observe similar patterns for other groups, namely urban/rural populations, households in which the head is employed or unemployed, male and female headed households, and households headed by persons older or younger than 65.

Table 3
Summary statistics pre and post-vignettes.

	Pre-vignette	Post-vignette
Overall mean	2.75 (0.87)	2.8 (0.92)
Household characteristics		
Poor	2.53	2.59
Non-poor	2.94	2.99
Rural	2.67	2.73
Urban	2.91	2.93
Unemployed head	2.61	2.66
Employed head	2.82	2.88
Female head	2.58	2.62
Male head	2.79	2.84
Pensioner head	2.78	2.63
Non-pensioner	2.59	2.83

Note: Standard deviation in parentheses.

Table 4
Comparison of pre-vignette subjective welfare with objective measure.

Subjective welfare rank	Expenditure per capita rank						Total
	1 poorest	2	3	4	5	6 richest	
1 poorest	84	120	108	26	3	0	341
2	109	461	586	137	8	1	1302
3	130	582	1184	345	29	5	2275
4	18	135	369	227	23	4	776
5	0	3	25	36	3	0	67
6 richest	0	1	3	5	1	0	10
Total	341	1302	2275	776	67	10	4771

Notes: Cramer's $V = 0.1484$; Chi-square = 525 (significant at 1%).

2.3. Subjective vs. objective economic welfare

To compare the two subjective measures against our objective measure based on expenditure per person, we define the sample sizes of six categories of expenditure per capita, based on the distribution of households in the subjective categories. If there are N respondents who place themselves on the lowest subjective step, the lowest N households in terms of expenditure per capita will make up the lowest category in the objective measure. Table 4 presents the results for the pre-vignette subjective rankings.

If the subjective measures are perfectly explained by the objective measure, all observations in the matrix would be along the shaded diagonal. Though the subjective measures are highly correlated with the objective measure, the matching is imperfect. Nearly half (43 percent) of those in the lowest objective classification place themselves on steps 3 or higher in the subjective measure. Of the richest decile of the population according to objective measures, 19 percent place themselves on the lowest two rungs of the subjective ladders, and less than half position themselves on the top three rungs. Of those households classified as extreme poor, living below the food poverty line (roughly 15 percent of the population), only 55 percent place themselves on the lowest two subjective rungs. Similarly, among poor households, only 45 percent place themselves on the lowest two subjective rungs.

The post-vignette rankings show a similar relationship to objective measures as for the pre-vignette rankings (Table 5). Twenty percent of those in the richest decile position themselves on the lowest two rungs. Among extreme poor households and poor households, 53 and 44 percent respectively place themselves on the bottom two rungs. It is also interesting to note that the pre-vignette rankings seem more consistent with our objective measure than the post-vignette rankings. The

Table 5
Comparison of post-vignette subjective welfare with objective measure.

Subjective welfare rank	Expenditure per capita rank						Total
	1 poorest	2	3	4	5	6 richest	
1 poorest	78	128	98	47	7	1	359
2	126	422	549	154	19	2	1272
3	126	560	998	362	53	6	2105
4	23	150	402	260	43	5	883
5	6	10	53	54	13	1	137
6 richest	0	2	5	6	2	0	15
Total	359	1272	2105	883	137	15	4771

Notes: Cramer's $V = 0.1355$; Chi-square = 438 (significant at $p < 0.0005$).

Table 6
Correlations with objective measures.

	Expenditure per capita (Tajik somoni)	Correlation with pre-vignette ranking	Correlation with post-vignette ranking
Overall mean	176 (162)	0.203	0.178
Household characteristics			
Poor	100	0.188	0.176
Non-poor	245	0.135	0.106
Rural	157	0.206	0.187
Urban	218	0.193	0.169
Unemployed head	161	0.176	0.176
Employed head	185	0.209	0.173
Female head	192	0.167	0.135
Male head	173	0.225	0.202
Pensioner head	157	0.151	0.157
Non-pensioner	180	0.209	0.179

Note: Standard deviation in parentheses.

Table 7
Placement of vignettes.

First vignette ranked on step	Expenditure per capita	Subjective position	*N*
1	173	2.64	3276
2	181	2.95	1352
3	195	3.12	140

Note: Three households which ranked the first vignette above step 4 are excluded.

Cramer's V statistic is higher in the pre-vignette question, indicating a stronger association between the rows and columns, and a better overall fit to the objective measure.

Comparing the correlation between the pre- and post-vignette rankings and the objective per capita expenditure measure, the correlations are higher for pre-vignette reports (Table 6, row 1). This is also true universally across various populations of interest. It is interesting to note that within some household categories we find lower average subjective welfare reports among the sub-group with higher per capita expenditure. This is true, for example, among female and male headed households (where female-headed households are on average richer by objective measures). Across all household categories the post-vignette scores are higher than those reported prior to the vignettes.

2.4. Placement of vignettes

The vignettes were designed such that the first vignette presented a scene of extreme poverty, the second vignette of improved conditions though still poor, the third of middle class and the fourth of relative affluence.

If the vignettes are an effective method of imposing a uniform scale, we would expect the placement of particularly the first vignette to be a good indicator of objective welfare. As the first vignette presents a picture of extreme poverty, most respondents would be expected to place this vignette on step 1. Households which position the first vignette on step 2 might be expected, on average, to be poorer than those who placed it on step 1, as they are able to conceive of a household situation that was even poorer than the first vignette. Similarly households which place the first vignette on step 3 we might assume them to be poorer than those who place the vignette on steps 1 or 2 as they could imagine even worse living conditions compared to that depicted in the first vignette. Following a similar logic, we would also expect the average subjective welfare score to be higher for those that locate the first vignette on step 1 as opposed to those placing it on steps 2 or 3.

The data, however, did not bear out this expectation (Table 7). Those households that place the first vignette on step 1 are, on average, poorer than those who place the first vignette on steps 2 or higher. Though there is no statistical significance in terms of objective measures between households that place the first vignette on step 1; for those who place it on steps 2 or higher, there is a statistically significant difference in terms of subjective welfare measures. This would indicate that those households who position the first vignette on steps 2 or higher perceive themselves as better off than their objective circumstances would indicate.

3. Results

Test 1: We begin by asking whether vignette responses are correlated with widely used covariates from the literature, including objective measures of economic welfare. We assume an ordered probit specification, which has become standard in the literature. The specification for the underlying continuous variable (generating the ordinal categorical responses) is as follows:

$$\mathrm{VW}_{ik} = \beta_k \ln \mathrm{PCE}_i + \pi_k X_i + \varepsilon_{ki} \quad (k = 1, 4;\ i, \ldots, N) \tag{1}$$

Table 8
Significant predictors of how households rank the four vignettes.

Vignette	1 (poorest)	2	3	4 (richest)
Significant covariates at the 10% level or better	Household size (+)	Special secondary schooling (+)	PCE (+)	PCE (+)
	Basic education (−)	Number of employed (+)	Uzbek (+)	Uzbeck (+)
	Services sector occupation (+)	Small holding (−)	Primary schooling (+)	Number of employed (+)
	Small holding (+)	Sogd (+)	General secondary (+)	Agriculture sector (−)
	Khatlon urban (−)	Khatlon (+)	Number of employed (+)	Small and medium holding (−)
	Gbao rural (−)	Gbao urban (−)	Small holding (−)	Sogd (−)
		Gbao rural (+)	Sogd rural (−)	Khatlon urban (+)
			Khatlon urban (+)	Khatlon rural (−)
			RRP rural (−)	RRP (−)
			Gbao (+)	Gbao rural (+)
Pseudo R^2	0.022	0.018	0.016	0.019

where VW_{ik} is a latent continuous variable for respondent i's assessment of vignette k's welfare, which generates a discrete response on the scale from 1 to 6, PCE is per capita expenditure, X is a vector of other household-level variables.

Table 8 summarizes the results. Full details are given in the working paper version; see Beegle et al. (2009). The pseudo R^2s' are low, at approximately 0.02. In general, vignette rankings are not consistently or significantly correlated with household characteristics. Geographic characteristics are more likely to be significant for the vignettes higher on the consumption scale (vignettes 3 and 4).

For vignettes 3 and 4 (but not 1 and 2), we find a positive and statistically significant relationship between log PCE and the vignette rankings. Richer households are more likely to give a high welfare ranking to the better-off households described by vignettes 3 and 4. This is not what one would expect with the frame-of-reference effect described in Section 1, which would suggest that poor people would tend to rank the relatively rich vignette higher than rich people. More suggestive of this FORB is our finding that smallholders (in terms of land) tend to rate the poorest vignette higher than do other households.

So the results of Test 1 show that there are only a few significant correlates of vignette responses among the types of regressors commonly used in subjective welfare regressions. But the evidence is mixed on FORB: How much do these effects bias the standard subjective welfare regressions found in the literature?

Test 2: In our second test we examine a standard subjective welfare regression, employing widely used covariates from the literature, with the difference that we also estimate specifications augmented with the vignettes. The augmented specification is as follows:

$$SW_i = \beta \ln PCE_i + \pi X_i + \gamma V_i + \varepsilon_i \quad (i = 1, \ldots, N) \tag{2}$$

where SW_i is a latent continuous variable for the subjective welfare of respondent i, which also generates a discrete response on the scale from 1 to 6 and V is a vector representing the vignette responses. The vector V translates the vignette responses into a series of dummy variables. This eliminates FORB under the assumption that inter-personal differences in vignette responses stem solely from differences in how the scales are interpreted. We refer to the estimated β as the economic gradient in subjective welfare.

With six possible steps for each of the four vignettes, there would theoretically be twenty dummy variables with one step omitted for each vignette. In practice, however, some steps are omitted due to an insufficient number of responses, leaving a total of fourteen vignette dummy variables to capture the complete set of responses observed in the data.

Table 9 presents the results of the ordered probit based on Eq. (2) using the pre-vignette self-reported welfare. We estimate four alternative specifications. In column 1 we find that logged PCE is significantly positively associated with SW. When vignettes are introduced in the second specification (column 2), the coefficient on PCE is basically unchanged. We do, however, find that the set of vignette dummy variables are jointly significant. Column 2 suggests that there is a frame-of-reference effect on SW, although comparing columns 1 and 2 the vignette effects are not sufficiently correlated with the household's PCE to generate more than negligible bias in the unconditional economic gradient in subjective welfare. That is, there is negligible FORB.

In column 3 of Table 9 we include non-income household characteristics and omit the vignettes. A number of these characteristics have significant effects on subjective welfare controlling for PCE. Female-headed households have lower SW. Households where the head has completed higher education, those where the head has a professional job (such as sales, service and public administration) and larger households generally have higher SW. We do not find consistent urban/rural patterns across areas. Households in urban Gbao report higher SW compared to the reference group in Dushanbe. In rural Khatlon and RRP as well as urban Khatlon, households have lower SW compared to households in Dushanbe. One might expect that households with migrants would have lower subjective welfare scores as they have a wider scope of knowledge and are therefore less likely to overstate their position. The coefficient is positive though not statistically significant.

After adding controls for other household characteristics, we find an increase in the economic gradient (the increase in the coefficient on PCE in columns 1 and 3 of Table 9). This is a statistically significant difference at the 10 percent level. The

Table 9
Pre-vignette self-assessed subjective welfare positions (ordered probit).

	(1) PCE		(2) PCE + vignettes		(3) PCE + household controls		(4) PCE + household controls + vignettes	
	Coef.	se	Coef.	se	Coef.	se	Coef.	se
Log expenditure per capita	0.630***	0.037	0.631***	0.037	0.775***	0.044	0.786***	0.045
Household demographics								
Female headed household				−0.128**	0.053	−0.126**	0.053	
Age of household head				−0.128**	0.053	−0.126**	0.053	
Log household size				0.467***	0.051	0.481***	0.051	
Number of older adults				0.039	0.043	0.024	0.043	
Any migrants				0.467***	0.051	0.481***	0.051	
Ethnicity (Reference: Tajik)								
Uzbek					0.018	0.046	0.017	0.046
Russian					−0.182	0.162	−0.190	0.169
Other					−0.602***	0.157	−0.624***	0.147
Education (Reference: no education)								
Primary					0.067	0.115	0.002	0.117
Basic					−0.127	0.108	−0.141	0.110
General secondary				0.001	0.001	0.105	−0.023	
Special secondary					0.115	0.112	0.080	0.113
Technical secondary				0.071	0.071	0.115	0.032	
Higher education					0.271***	0.113	0.249***	0.115
Graduate school					0.538	0.385	0.611*	0.324
Employment characteristics of household, head's occupation (Reference for occupation: unemployed)								
Number of employed					0.037**	0.017	0.035**	0.017
Agriculture/fishing/forestry				−0.004	−0.004	0.063	−0.020	
Manufacture/mining				−0.028	−0.028	0.094	−0.016	
Services					0.353***	0.129	0.294**	0.128
Construction					−0.079	0.079	−0.112	0.079
Public administration/education/health			0.174**	0.069	0.174**	0.069		
Sales and services					0.300***	0.064	0.302***	0.064
Other					−0.173	0.106	−0.174*	0.104
Agriculture (Reference for holdings: no land)								
Small holding					−0.086	0.058	−0.061	0.059
Medium holding					−0.134*	0.069	−0.119*	0.071
Large holding					−0.006	0.070	0.023	0.071
Geography (Reference: Dushanbe)								
Sogd urban					0.004	0.076	−0.051	0.076
Sogd rural					−0.110	0.079	−0.164**	0.080
Khatlon urban					−0.197**	0.096	−0.362***	0.097
Khatlon rural					−0.525***	0.074	−0.588***	0.074
RRP urban					−0.180*	0.109	−0.224**	0.110
RRP rural					−0.281***	0.075	−0.337***	0.076
Gbao urban					0.310***	0.107	0.375***	0.112
Gbao rural					−0.080	0.078	−0.094	0.079
Vignette 1 (positions 3–omitted)								
vign1==1			−0.258*	0.140			−0.287**	0.144
vign1==2			−0.022	0.133			−0.073	0.139
Vignette 2 (position 1 omitted)								
vign2=2			1.106***	0.223			1.178***	0.229
vign2=3			1.337***	0.230			1.450***	0.235
vign2=4			1.515***	0.261			1.682***	0.267
vign2=5			1.724***	0.574			1.716***	0.514
Vignette 3 (positions 1 and 2 omitted)								
vign3=3			−0.777***	0.257			−0.720**	0.287
vign3=4			−0.852***	0.263			−0.744**	0.292
vign3=5			−0.844***	0.270			−0.729**	0.298
vign3=6			−1.505***	0.401			−1.616***	0.457
Vignette 4 (positions 1–3 omitted)								
vign4=4			0.735***	0.245			0.593**	0.274
vign4=5			0.794***	0.247			0.555**	0.275
vign4=6			0.793***	0.251			0.547**	0.278
Number of observations	4,771		4,771		4,771		4,771	
Pseudo R^2	0.037		0.055		0.082		0.101	

* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Table 10
Log per capita expenditure beta coefficients.

	(1) PCE		(2) PCE + vignettes		(3) PCE + household controls		(4) PCE + household controls + vignettes	
	Coef.	se	Coef.	se	Coef.	se	Coef.	se
1: Pre-vignette self-assessed subjective welfare positions (ordered probit)								
Log expenditure per capita	0.630***	0.037	0.631***	0.037	0.775***	0.044	0.786***	0.045
2: Pre-vignette self-assessed subjective welfare positions (CHOPIT)								
Log expenditure per capita	0.5801***	0.0327	NA	NA	0.731***	0.0392	NA	NA
3: Post-vignette self-assessed subjective welfare positions (ordered probit)								
Log expenditure per capita	0.560***	0.037	0.573***	0.037	0.698***	0.043	0.724***	0.043
4: Post-vignette self-assessed subjective welfare positions (CHOPIT)								
Log expenditure per capita	0.5131***	0.0318	NA	NA	0.6504***	0.0378	NA	NA
5: Rescaled SW responses using post vignettes (ordered probit)								
Log expenditure per capita	0.466***	0.034	0.559***	0.035	0.593***	0.041	0.692***	0.042

Row 1 results are the log expenditure per capita results in Table 9.
*$p < 0.1$.
**$p < 0.05$.
***$p < 0.01$.

household characteristics in column 3 may capture several things, including frame-of-reference effects or other effects such as perceived vulnerability, permanent income, or security.

Introducing both the vignettes and the set of non-income household characteristics in column 4, we find that the economic gradient is basically unchanged from column 3. The set of vignette dummy variables remain statistically significant. Thus, while it appears that there is a frame-of-reference effect being picked up by the vignettes, it is still not influencing the economic gradient in subjective welfare. In addition, including the vignettes does not alter the coefficients on other household characteristics.

Table 10, row 3, presents the results for post-vignette subjective welfare outcomes regressed on PCE, and other controls; see Beegle et al. (2009) for the full set of results. Results are similar to Table 9. The magnitude of the coefficients tended to be lower than in the pre-vignette response. As before, we observed an increase in the economic gradient when we control for non-income household characteristics (comparing Table 10, row 3 columns 1–4). And, again, there was no significant change in this coefficient when we further include the vignettes. Likewise, the coefficients on other covariates remain largely unchanged. Our finding that FORB is negligible remains valid.

Test 3: A concern with Test 2 is that the vignette responses are strictly endogenous, given that they come from the same respondent at the same time and so could be jointly influenced by some latent characteristic. One response to this concern is to instead use the vignette responses to re-calibrate the subjective welfare responses. Our rescaling method follows that developed by King et al. (2004) and King and Wand (2007) in which only the relative position of the self-reported score in relation to the vignette rankings matters to the analysis. For example, all respondents who ranked themselves below vignette 1 would have a score of 1 in the re-scaled rankings, regardless of the actual values given to either the self-reported score or the ranking of the first vignette. Similarly, all those respondents who placed themselves at the same level as the first vignette would have a re-scaled ranking of 2, those between vignettes 1 and 2 would have a 3, and so on. Rescaling therefore gives nine possible values to the dependent variable.

Table 10, row 5, reports the coefficient on PCE for the rescaled post-vignette ranking; full results are given in Beegle et al. (2009). The rescaled regressions showed similar relationships with the household characteristics, including significant correlations with PCE, higher education, professional jobs, larger households and the geographic variables. The vignettes, in general, lost their significance.

So far we have assessed the impact of the vignettes on the economic gradient in subjective welfare under the assumption that this impact is constant across the income distribution. To allow for more flexibility in the specification we estimate non-parametric regression functions with linear controls ("partial linear regressions").[11] By comparing the non-parametric results with and without vignettes (as linear controls), we can test for FORB across the income distribution. We define two alternative binary outcome variables: SW being poor (steps 1 and 2) and SW being rich (steps 4–6). Since we have both pre- and post-vignette SW, we have 4 outcome variables (pre/post, poor/rich). Further, we assess the FORB from estimates with and without the non-income traits. Thus, we have a total of eight pairs of partial linear regressions (Figs. 1–4).

Fig. 1 shows the results for reporting oneself as poor and as rich, with and without vignettes as controls. The "poor" curve is downward sloping: households are less likely to place themselves on the lowest two steps as PCE increases. The opposite is true for those who place themselves on the upper rungs ("rich"). The economic gradient in subjective wellbeing is unchanged if we control for the household's scoring of vignettes. That is, the FORB appears to be minimal across the whole

[11] We used the PLREG program for STATA written by Lokshin (2006).

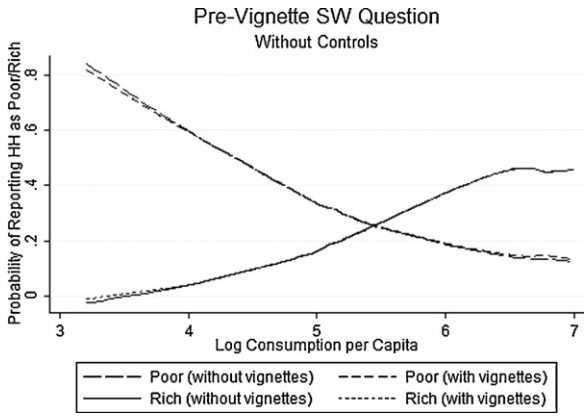

Fig. 1. Pre-vignettes subjective welfare as a function of log consumption per capita, without controls.

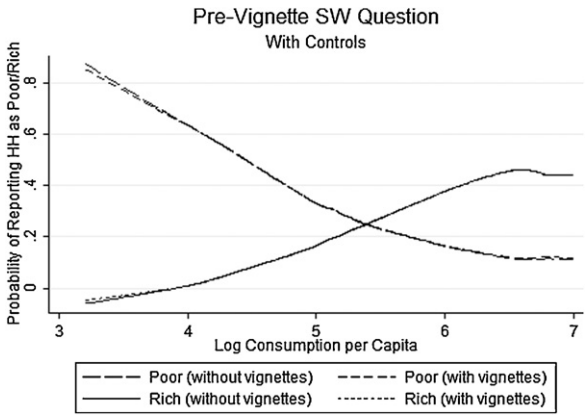

Fig. 2. Pre-vignettes subjective welfare as a function of log consumption per capita, with controls.

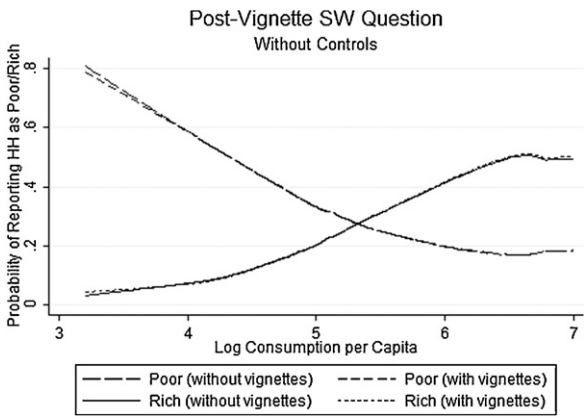

Fig. 3. Post-vignettes subjective welfare as a function of log consumption per capita, without controls.

income distribution. If there was FORB, we would expect the curves with the vignettes to be steeper than those without, assuming that the poor would overstate and the rich understate their subjective welfare, with attenuation toward the mean.

We find the same results when we include non-income controls (Fig. 2) and when we define the dependent variables using the post-vignette SW (Figs. 3 and 4).

As noted above, the economic gradient was lower for the post-vignette SW compared to the pre-vignette score (Table 10, rows 1 and 3). This difference in the economic gradient may also vary across the income distribution. Partial linear regressions of poor SW and rich SW are presented in Fig. 5, which compares the results based on pre- and post-vignette SW. Fig. 5 shows that households at all income levels are more likely to report being rich after scoring the vignettes compared to their prior

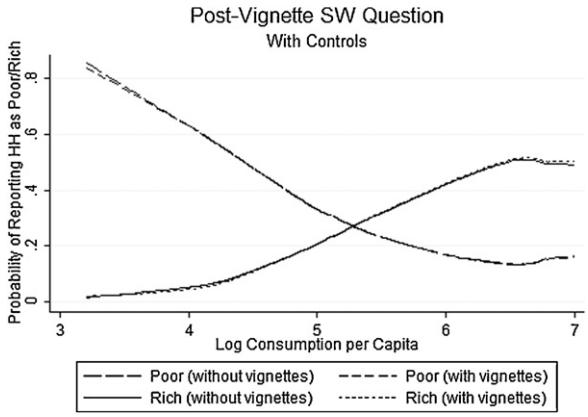

Fig. 4. Post-vignettes subjective welfare as a function of log consumption per capita, with controls.

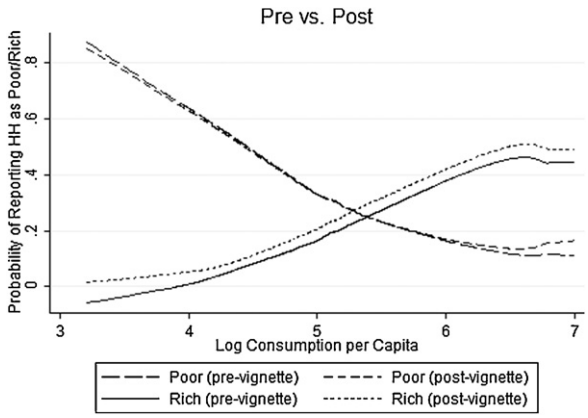

Fig. 5. Comparison of pre-vignettes and pre-vignettes subjective welfare as a function of log consumption per capita.

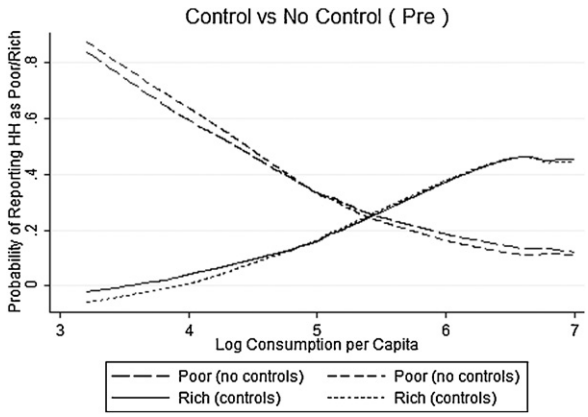

Fig. 6. Comparison of pre-vignettes subjective welfare with and without controls as a function of log consumption per capita.

response; that is, the "rich" curve is shifted up after the hearing the vignettes. There is little change in the probability of reporting being poor for the pre- versus post-vignette SW.

Non-income household characteristics have a modest impact the estimates of the economic gradient of subjective welfare. The economic gradient with respect to self-reporting as poor is slightly steeper when additional control variables are included (Fig. 6, "poor"). There is only a slight shift in the gradient with respect to self-reporting as rich after adding controls (Fig. 6, "rich"). These results are the same for the post-vignette SW (Fig. 7).

Test 4: In Test 1 we modeled vignette rankings as functions of socioeconomic characteristics of the respondent. By construction, these rankings differ due to the scales used. In Tests 2 and 3 we modeled subjective welfare allowing for the

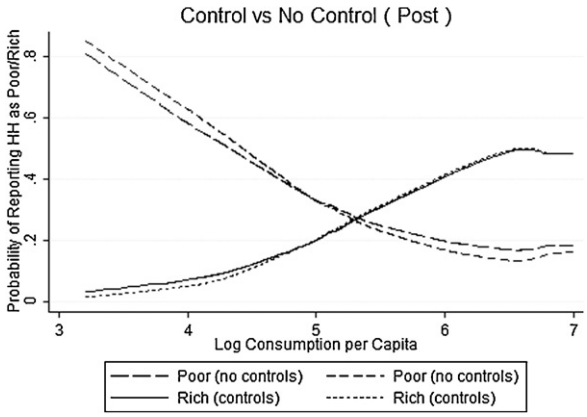

Fig. 7. Comparison of post-vignettes subjective welfare with and without controls as a function of log consumption per capita.

Table 11
Threshold equation results from CHOPIT estimation.

	Threshold 1		Threshold 2		Threshold 3		Threshold 4		Threshold 5	
	Coef.	se	Coef.	se	Coef.	se	Coef.	se	Coef.	se
Specification with only lnpce										
Log expenditure per capita	−0.0451*	0.0268	0.0279	0.0296	−0.045*	0.0234	−0.0251	0.0246	0.0775***	0.0285
Specification with full controls										
Log expenditure per capita	−0.0892***	0.0325	0.0765**	0.0361	−0.011	0.0293	−0.0001	0.0313	0.0584	0.0371
Household demographics										
Female headed household	0.0269	0.0417	−0.0169	0.0456	−0.0400	0.0382	−0.0235	0.0419	0.0421	0.0497
Age of household head	−0.0008	0.0015	0.0023	0.0017	−0.0006	0.0014	0.0001	0.0015	0.0015	0.0018
Log household size	−0.0723*	0.0395	0.0983**	0.0432	0.0081	0.0361	0.0586	0.0394	−0.0330	0.0475
Number of older adults	−0.0033	0.0343	−0.0320	0.0375	0.0254	0.0309	−0.0311	0.0338	−0.0390	0.0394
Number of migrants	−0.0053	0.0350	−0.0177	0.0384	−0.0268	0.0313	−0.0106	0.0346	−0.0160	0.0405
Ethnicity (Reference: Tajik)										
Uzbek	0.0017	0.0403	0.0365	0.0443	−0.0955***	0.0360	−0.0019	0.0407	0.0736	0.0482
Russian	0.1494	0.1238	−0.1280	0.1335	−0.0210	0.1187	0.1002	0.1297	−0.3310**	0.1437
Other	0.0090	0.1081	0.2053*	0.1175	−0.1404	0.1022	−0.0624	0.1067	0.0605	0.1312
Education (Reference: no education)										
Primary	−0.1503	0.0928	−0.1370	0.1010	0.0387	0.0875	0.0408	0.0994	0.1483	0.1126
Basic	−0.0288	0.0865	−0.0921	0.0944	0.0046	0.0822	−0.0612	0.0931	0.1035	0.1052
General secondary	−0.0834	0.0840	−0.0068	0.0920	0.0132	0.0795	−0.1109	0.0902	0.1109	0.102
Special secondary	−0.1737*	0.0909	0.0562	0.0998	−0.0081	0.0846	−0.1302	0.0953	0.1826*	0.1089
Technical secondary	−0.0724	0.0922	−0.0795	0.1010	0.0181	0.0864	−0.0883	0.0978	0.1364	0.1107
Higher education	−0.0687	0.0906	−0.0222	0.0996	−0.0304	0.0843	−0.0709	0.0952	0.1311	0.109
Graduate school	0.5711*	0.3331	−0.258	0.376	0.018	0.3309	0.0367	0.3573	−0.2320	0.3629
Employment characteristics of household, head's occupation (Reference: for occupation: unemployed)										
Number of employed	−0.0029	0.0146	−0.0215	0.0161	0.0058	0.013	−0.0262*	0.0143	0.0006	0.0170
Agriculture/fishing/forestry	−0.0496	0.0512	0.0455	0.0561	0.0609	0.0466	0.0128	0.0519	0.0224	0.0605
Manufacture/mining	0.0153	0.0792	0.0612	0.0875	−0.0174	0.0709	−0.0303	0.0775	0.0372	0.0932
Services	−0.4357***	0.1408	0.3557**	0.1584	0.0626	0.1110	0.0479	0.1155	0.0604	0.1424
Construction	−0.0699	0.0666	−0.0088	0.0727	0.0465	0.0605	0.0080	0.0660	−0.0070	0.0778
Public Administration/education/health	0.0090	0.0561	−0.0184	0.0619	0.0677	0.0496	0.0824	0.0543	−0.1240*	0.0634
Sales and services	−0.0214	0.0514	−0.0137	0.0566	0.0255	0.0452	0.0565	0.0497	−0.0030	0.0587
Other	0.0001	0.0867	0.1210	0.0951	0.0243	0.0811	−0.1680**	0.0835	−0.0890	0.1000
Agriculture (Reference for holdings: no land)										
Small Holding	0.0096	0.0468	0.0669	0.0510	0.0044	0.0428	0.0107	0.0473	−0.0370	0.0556
Medium Holding	−0.0348	0.0581	0.1248**	0.0635	0.0350	0.0530	−0.0467	0.0590	0.0171	0.0694
Large Holding	0.0094	0.0561	0.1012*	0.0614	−0.0677	0.0506	−0.0447	0.0562	−0.0290	0.0658
Geography (Reference: Dushanbe)										
Sogd urban	−0.1344*	0.0688	0.0167	0.0759	0.0935	0.0610	0.2829***	0.0671	0.1206	0.0827
Sogd rural	−0.0416	0.0659	−0.0093	0.0723	0.1481**	0.0604	0.2776***	0.0668	−0.1320*	0.0775
Khatlon urban	−0.1341	0.0830	−0.3095***	0.0879	−0.0982	0.0704	0.2625***	0.0778	−0.0800	0.0891
Khatlon rural	−0.0115	0.0620	−0.1018	0.0678	0.0145	0.0564	0.2910***	0.0628	−0.0910	0.0733
RRP urban	−0.0742	0.0904	0.0987	0.1001	0.0831	0.0816	0.0338	0.0860	0.0124	0.1066
RRP rural	−0.1099*	0.0611	0.0793	0.0674	0.0665	0.0554	0.2429***	0.0610	0.0303	0.0729
Gbao urban	0.1356	0.1060	−0.0335	0.1204	−0.0984	0.0927	−0.1138	0.0880	−0.0990	0.1107
Gbao rural	0.0050	0.0678	−0.1065	0.0742	0.0428	0.0610	0.0496	0.0651	−0.0840	0.0771

* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

heterogeneity in scales using the vignettes rankings. It is of interest to see if we can learn more about the heterogeneity in scales by modeling them jointly with own-welfare rankings.

King et al. (2004) propose an econometric method ("CHOPIT") for modeling the same type of data we use here in which there is an equation for both the latent response of interest and one for the vignettes, with the two equations estimated jointly. As above, there is a continuous latent variable for the respondents' self-rated welfare and an observed ordinal response, with thresholds that are explicit functions of a vector of explanatory variables. The extra information on the vignette responses provides the basis for identification, under the assumption that the thresholds for self-assessed welfare are determined identically as for these thresholds in the vignette responses. This is a natural assumption to make. Without the vignettes, identification would only be possible under questionable assumptions about the nonlinearity of the functional forms involved (as discussed in King et al., 2004). Thus we are able to model determinants of the thresholds separately to those of the latent continuous variable for subjective welfare.[12]

Table 10, rows 2 and 4, gives the coefficients of PCE implied by this estimator. The CHOPIT estimates for the determinants of subjective welfare were quiet similar to those we have already reported and discussed. What is new about Test 4 is the extra insight we get into the determinants of the thresholds. Table 11 presents CHOPIT's estimates for the threshold equations, corresponding to the "pre-vignettes" case. (Results were very similar for the post-vignettes case.)

We find a striking pattern in the effect of differences in PCE on the thresholds. The top threshold has a significantly positive expenditure gradient, such that richer households use a higher threshold in defining who is on the top rung of the ladder. By contrast, the gradient is negative for the lowest threshold. Between the two there is not much pattern, but then one cannot reject the null that there is no gradient in most cases. So there is a clear suggestion here that the relatively rich respondents tend to have a wider range of thresholds than the poor. There are very few other significant covariates of the thresholds.

4. Conclusions

A cloud of doubt has hung over subjective welfare regressions, arising from concerns about likely heterogeneity in the interpretation of the scales widely used to measure subjective welfare. This heterogeneity undoubtedly reduces the power of standard covariates in explaining perceived welfare. More worrying, however, is the possibility that the heterogeneity in scales is leading to biased inferences from studies of subjective welfare, including biases in identifying its economic gradient, as well as the effects of other variables such as education, employment status and relative economic position. Bias arises if these variables are correlated with the latent heterogeneity in scales. It can be conjectured that poorer people tend to have more limited horizons in life, due to more limited experiences with the extent of the disparities in levels of living in society as a whole; a poor person's idea of what it means to be "rich" may then be very different to that of a middle or upper income person with a very different frame of reference. This may be correlated with certain attributes of the household, such as if rural or more isolated households may overstate their welfare given that they have a limited experience or exposure to higher living standards with which to judge their own economic standing.

We have provided various tests for bias due to such heterogeneity in individual scales. The tests entailed adding vignettes of hypothetical households to a national household survey for Tajikistan in 2007. Respondents placed these vignettes on the same six-step ladder used to report their own subjective economic welfare. In our first test for this bias, we do find some significant covariates for vignette responses among the regressors commonly used to explain subjective welfare. However, the effects are neither very strong nor consistent across different vignettes.

In the second and third tests, we explore the extent of bias due to the frame of reference effect by embedding vignette rankings by households into standard regressions for subjective welfare. We do this in two ways: we include vignettes among the covariates and, to address endogeneity concerns with this approach, we also recalibrate the self-assessments of welfare to accord with the heterogeneity in scales revealed by the vignette responses. The striking finding of these further tests is that the bias is negligible with respect to the "income effect" on subjective welfare as well as other covariates. Based on this study, the concerns that past uses of subjective economic welfare data are compromised by systematic differences in the meaning given to the scales used appear to be unwarranted.

Our fourth test has involved estimating a model of the heterogeneity in scales, whereby the thresholds are modeled as functions of covariates, assuming that the structure of this heterogeneity is common between "own welfare" and the vignettes. This shows similar results on the factors influencing subjective welfare, but also throws new light on the heterogeneity in scales. In particular, we find that poorer households tend to have narrower range in the thresholds used in identifying where they lie and where the vignettes lie on the subjective welfare ladder.

A frame-of-reference effect on subjective welfare is evident in our findings; people with different socioeconomic backgrounds use systematically different scales in responding to questions on their welfare. However, our results do not suggest that this is an important source of bias in past efforts to model the objective determinants of subjective welfare.

[12] We implemented the CHOPIT analysis using the R statistical analysis program, using the programs 'anchors,' 'rgenoud,' and 'Zelig.' Further information and documentation on these packages is available at http://sekhon.berkeley.edu/rgenoud, http://wand.stanford.edu/anchors and http://gking.harvard.edu/zelig.

Appendix 1.

Table A1 Summary statistics (weighted).

Variable	Mean	sd
Log expenditure per capita (Tajik somoni)	5.0	0.54
Household Demographics		
Female-headed household	0.19	
Age of household head	50.6	13.74
Household size	6.3	2.80
Number of older adults (65+)	0.3	0.57
Number of migrants	0.3	0.68
Ethnicity		
Tajik	0.79	
Uzbek	0.18	
Russian	0.01	
Other	0.02	
Education of household head		
No education	0.04	
Primary (grades 1–4)	0.07	
Basic (grades 1–8)	0.12	
Secondary general (grades 9–10)	0.34	
Secondary special	0.13	
Secondary technical	0.11	
Higher education	0.19	
Graduate school/*aspirantura*	0.002	
Employment characteristics of household, head's occupation		
Number of employed	1.76	1.33
Not employed	0.37	
Agriculture, fishing and forestry	0.18	
Manufacture and mining	0.04	
Services (electricity, gas, hot water, etc.)	0.02	
Construction	0.07	
Public administration, education, health	0.13	
Sales and services	0.16	
Other	0.03	
Agriculture		
No land holdings	0.37	
Small holding (1–10 acres)	0.29	
Medium holding (11–20 acres)	0.15	
Large holding (21+ acres)	0.20	

Appendix 2. The vignettes in the TLSMS

Vignette 1: Family A can only afford to eat meat on very special occasions. During the winter months, they are able to partially heat only one room of their home. They cannot afford for children to complete their secondary education because the children must work to help support the family. When the children are able to attend school, they must go in old clothing and worn shoes. There is not enough warm clothing for the family during cold months. The family does not own any farmland, only their household vegetable plot.

Vignette 2: Family B can afford to eat meat only once or twice a week. During winter months, they can heat several rooms, but not the entire house. They cannot afford for all their children to complete secondary education. Their clothing is sufficiently warm, but they own only simple garments. In addition to their household vegetable plot, they own a small plot of poor quality farmland that is distant from their home.

Vignette 3: Family C can afford to eat meat everyday. During the winter months, generally they are able to keep their home warm. They can afford for all their children to complete secondary education. They have sufficient clothing to keep warm in the winter. Their everyday clothing is simple, but they also have some fancy items for special occasions. In addition to their household vegetable plot, they have a larger plot of good quality farmland, not too distant from their home.

Vignette 4: Family D can afford to eat whichever foods they would like, including sweets and imported food. During the winter months, they have no problems with heating and are able to keep their entire house warm. They can afford for all of their children to complete their education, and then to continue at a local university. They are able to afford a variety of fancy traditional clothes and also imported brand clothing. The family owns property, including a good car. The family also has a large farm and acts as landlord to others in their area.

References

Angoff, W.H., 1993. Perspectives on differential item functioning methodology. In: Holland, P., Wainer, H. (Eds.), Differential Item Functioning. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 3–24.

Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., O'Donnell, O., 2008. Does reporting heterogeneity bias the measurement of health disparities? Health Economics 17 (3), 351–375.
Beegle, K., Himelein, K., Ravallion, M., 2009. Frame-of-Reference Bias in Subjective Welfare Regressions. Policy Research Working Paper 4904. World Bank, Washington, DC.
Clark, A., Oswald, A., 1994. Unhappiness and unemployment. Economic Journal 104 (424), 648–659.
Clark, A., Oswald, A., 1996. Satisfaction and comparison income. Journal of Public Economics 61, 359–381.
Conti, G., Pudney, S., 2008. If You're Happy and You Know it, Clap Your Hands! Survey Design and the Analysis of Satisfaction. Institute for Social and Economic Research Working Paper 2008-39.
Das, J., Hammer, J., Leonard, K., 2008. The quality of medical advice in low-income countries. Journal of Economic Perspectives. 22 (2), 93–114.
Diener, E., Suh, E., Lucas, R., Smith, H., 1999. Subjective well-being: three decades of progress. Psychological Bulletin 125, 276–302.
Di Tella, R., MacCulloch, R., 2006. Some uses of happiness data in economics. Journal of Economic Perspectives. 20 (1), 25–46.
Easterlin, R., 1995. Will raising the incomes of all increase the happiness of all? Journal of Economic Behavior and Organization. 27, 35–47.
Ferrer-i-Carbonell, A., 2005. Income and well-being: an empirical analysis of the comparison income effect. Journal of Public Economics. 89, 997–1019.
Fafchamps, M., Shipli, F., 2009. Isolation and subjective welfare: evidence from South Asia. Economic Development and Cultural Change 57 (4), 641–683.
Frank, R.H., 1997. The frame of reference as a public good. Economic Journal 107, 1832–1847.
Frey, B., Stutzer, A., 2002. What can economists learn from happiness research? Journal of Economic Literature XL, 402–435.
Furnham, A., Argyle, M., 1998. The Psychology of Money. Routledge, London.
Hirshman, A., 1973. The changing tolerance for income inequality in the course of economic development, with a mathematical appendix by Michael Rothschild. Quarterly Journal of Economics 87, 544–566.
Hopkins, D., King, G., 2010. Improving anchoring vignettes designing surveys to correct for interpersonal incomparability. Public Opinion Quarterly, 1–22.
Kapteyn, A., Kooreman, P., Willemse, R., 1998. Some methodological issues in the implementation of subjective poverty definitions. Journal of Human Resources 23 (2), 222–242.
Kapteyn, A., Smith, J., Van Soest, A., 2008. Comparing Life Satisfaction. Working Paper WR-623, Rand Cooperation.
King, G., Wand, J., 2007. Comparing incomparable survey responses: evaluating and selecting anchoring vignettes. Political Analysis 15 (1), 46–66.
King, G., Murray, C., Salomon, J., Tandon, A., 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. American Political Science Review 98 (1), 191–207.
Kristensen, N., Johansson, E., 2008. New evidence on cross-country differences in job satisfaction using vignettes. Labor Economics 15 (1), 96–117.
Lokshin, M., 2006. Difference-based semiparametric estimation of partial linear regression models. STATA Journal 6 (3), 377–384.
Luttmer, E., 2005. Neighbors as negatives: relative earnings and well-being. Quarterly Journal of Economics 120 (3), 963–1002.
McBride, M., 2001. Relative income effects on subjective well-being in the cross-section. Journal of Economic Behavior and Organization 45, 251–278.
Oswald, A., 1997. Happiness and economic performance. Economic Journal 107, 1815–1831.
Pradhan, M., Ravallion, M., 2000. Measuring poverty using qualitative perceptions of consumption adequacy. Review of Economics and Statistics 82 (3), 462–471.
Ravallion, M., Lokshin, M., 2000. Who wants to redistribute? The tunnel effect in 1990s Russia. Journal of Public Economics 76 (1), 87–104.
Ravallion, M., Lokshin, M., 2001. Identifying welfare effects from subjective questions. Economica 68 (271), 335–357.
Ravallion, M., Lokshin, M., 2002. Self-rated economic welfare in Russia. European Economic Review 46 (8), 1453–1473.
Ravallion, M., Lokshin, M., 2010. Who cares about relative deprivation? Journal of Economic Behavior and Organization 73 (2), 171–185.
Seidl, C., 1994. How sensible is the Leyden individual welfare function of income? European Economic Review 38, 1633–1659.
Senik, C., 2004. When information dominates comparison: learning from Russian subjective panel data. Journal of Public Economics 88, 2099–2123.
van de Stadt, H., Kaptyn, A., van de Geer, S., 1985. The relativity of utility: evidence from panel data. Review of Economics and Statistics 67, 179–187.
Van Praag, B., 1968. Individual Welfare Functions and Consumer Behavior. North-Holland, Amsterdam.
Van Praag, B., Kapteyn, A., 1994. How sensible is the Leyden individual welfare function of income? A reply. European Economic Review 38 (9), 1817–1825.
Veenhoven, R., Ehrhardt, J., Sie Dhian Ho, M., de Vries, A., 1993. Happiness in Nations: Subjective Perceptions of Life in 56 Nations 1946–1992. Erasmus University, Rotterdam.
Winkelmann, L., Winkelmann, R., 1998. Why are the unemployed so unhappy: evidence from panel data. Economica 65, 1–15.
World Bank, 2008. Basic Information Document: 2007 Tajikistan Living Standards Measurement Study. Mimeo, World Bank, Washington, DC.