

Example 4.2

Happiness Surveys

Are you happy with your economic lot in life? Economists have periodically shown a keen interest in answering questions such as this. In the late 1800s, they joined with psychiatrists and psychologists to conduct experiments in which electrodes were placed on subjects' heads in an attempt to measure the pleasure they experienced while consuming various goods and services. These studies proved to be useless and were abandoned. About thirty years ago, there was a renewed interest in discovering how satisfied or happy people are with certain aspects of their lives, most commonly their income, health, work, and family. This time economists chose surveys as their tool of analysis, asking representative samples of people to record their satisfaction of these aspects on a scale ranging from very dissatisfied to highly satisfied. The motivation for these surveys and the earlier experiments was the same, that the standard objective measures of economic well being such as income and wealth are clearly incomplete and inadequate measures of people's overall utility or satisfaction. Economists have a natural interest in determining how consumption, income, health, and the like affect people's utility, because individuals' utilities are the basis of social welfare. This example considers the relationship between individual satisfaction and income.¹

The survey literature on the relationship between satisfaction and income is by now quite extensive, and it has generated some puzzling results. The economic theory of consumer behavior implies that satisfaction or utility rises with income. Surveys conducted within a country over time tend to find this result, although in most cases the relationship between income and satisfaction is not as strong as one would expect. In contrast, surveys comparing satisfaction with income across countries often find that satisfaction with income is not higher in the richer countries. The cross-country result became known as the Easterlin paradox after Richard Easterlin, who first documented it.

¹ Economists are teaming up once again with medical specialists to use state of the art brain imaging techniques to record how people react to various economic activities. This research agenda, known as neuroeconomics, is in its infancy.

The United States is commonly cited as a particularly striking example of the Easterlin paradox. Surveys comparing satisfaction with income in the U.S. versus European countries typically find that Americans are *less* satisfied with their incomes, even in those instances when they have much higher incomes, on average, than the Europeans they are being compared with. Americans appear to be a relatively grumpy folk, and there is no obvious reason why this should be.

Economists Arie Kapteyn, James P. Smith, and Arthur van Soest (KSvS) may have resolved the Easterlin paradox. In their 2008 paper, “Are Americans Really Less Happy With Their Incomes?”, they point out that the Easterlin paradox may simply be due to the design of the surveys. The respondents in the surveys are asked to place their satisfaction with their incomes, or anything else for that matter, into one of a discrete set of categories, such as from ‘highly satisfied’ to ‘very dissatisfied’.² These discrete responses can be quite misleading if they are used to compare how much satisfaction people in different countries receive from their incomes. This is an important point, because it also applies to survey comparisons of satisfaction between groups of people within a country differentiated by gender, or race, or religion. Indeed, it applies to any survey of subjective preferences across different groups of people that asks them to place their preferences into discrete categories.

SURVEY BIAS

To see the nature of the problem, consider the relationship between income and satisfaction. A sample of individuals is meant to be representative of an entire population, and over the entire population the relationship between satisfaction and income is essentially continuous, as Figure 1 illustrates. It pictures the frequency distribution of the level of satisfaction received by different people who have a given selected level of income. The level of satisfaction is on the horizontal axis and the proportion of people with each level of satisfaction is on the vertical axis. The pictured relationship is roughly bell-shaped; as one would expect: a few people will be ‘very dissatisfied’ with that income, a few others ‘highly satisfied’, with most people somewhere between ‘mildly dissatisfied’ to ‘mildly satisfied’ in the middle of the distribution.

² A. Kapteyn, J. P. Smith, and A. van Soest, “Are Americans Really Less Happy With Their Incomes?”, *Working Paper WR-591*, Labor and Population Group working paper series, RAND Corporation, June 2008.

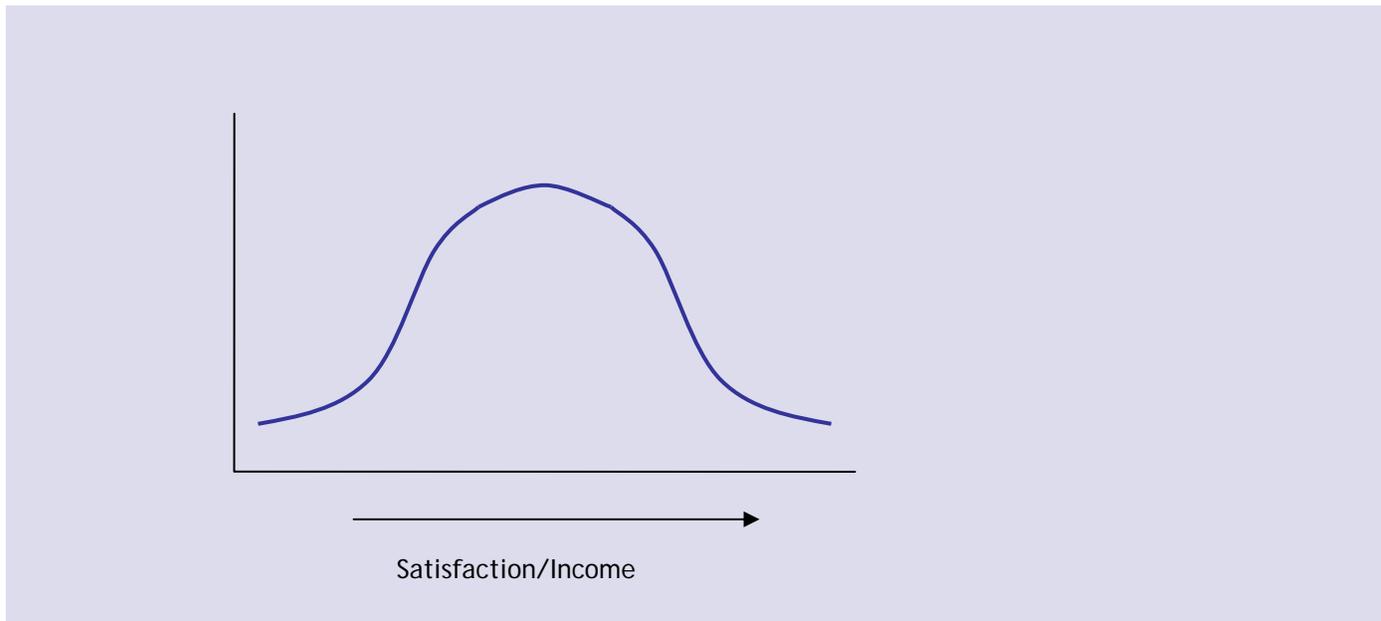


Figure 1: Frequency distribution of satisfaction for a given level of income

The surveys cannot capture the continuous distribution, however, because individuals can never say precisely how much satisfaction they receive from their income, or anything else. The best they can do is place their level of satisfaction within a broadly defined discrete category, which is what the surveys ask them to do. But even this forces the survey participants to implicitly choose thresholds or lines of demarcation between at least two categories, such as between ‘dissatisfied’ and ‘very dissatisfied’. If people in different countries choose different thresholds, on average, then comparing the survey results between the two countries can be seriously misleading. (The same point applies to different groups of people within a country.)

To see why, refer to Figure 2, which shows the continuous frequency distribution of satisfaction for all people in countries 1 and 2 who have a given selected amount of income. The distribution for country 1 is skewed more toward the right than that of country 2, indicating that people in country 1 are generally more satisfied with that given income. (Assume for the sake of discussion that this same skewed pattern holds at all income levels.) Suppose the people surveyed in each country who have that income are asked to indicate whether they are ‘satisfied’, ‘highly satisfied’, or ‘very dissatisfied’ with their income and the thresholds are as drawn. Note that the threshold lines in country 1 are drawn to the right of the lines in country 2. That is, when it comes to setting the thresholds for the purposes of responding to the survey question, the people in country 1 choose to be more negative in expressing their satisfaction with their incomes.

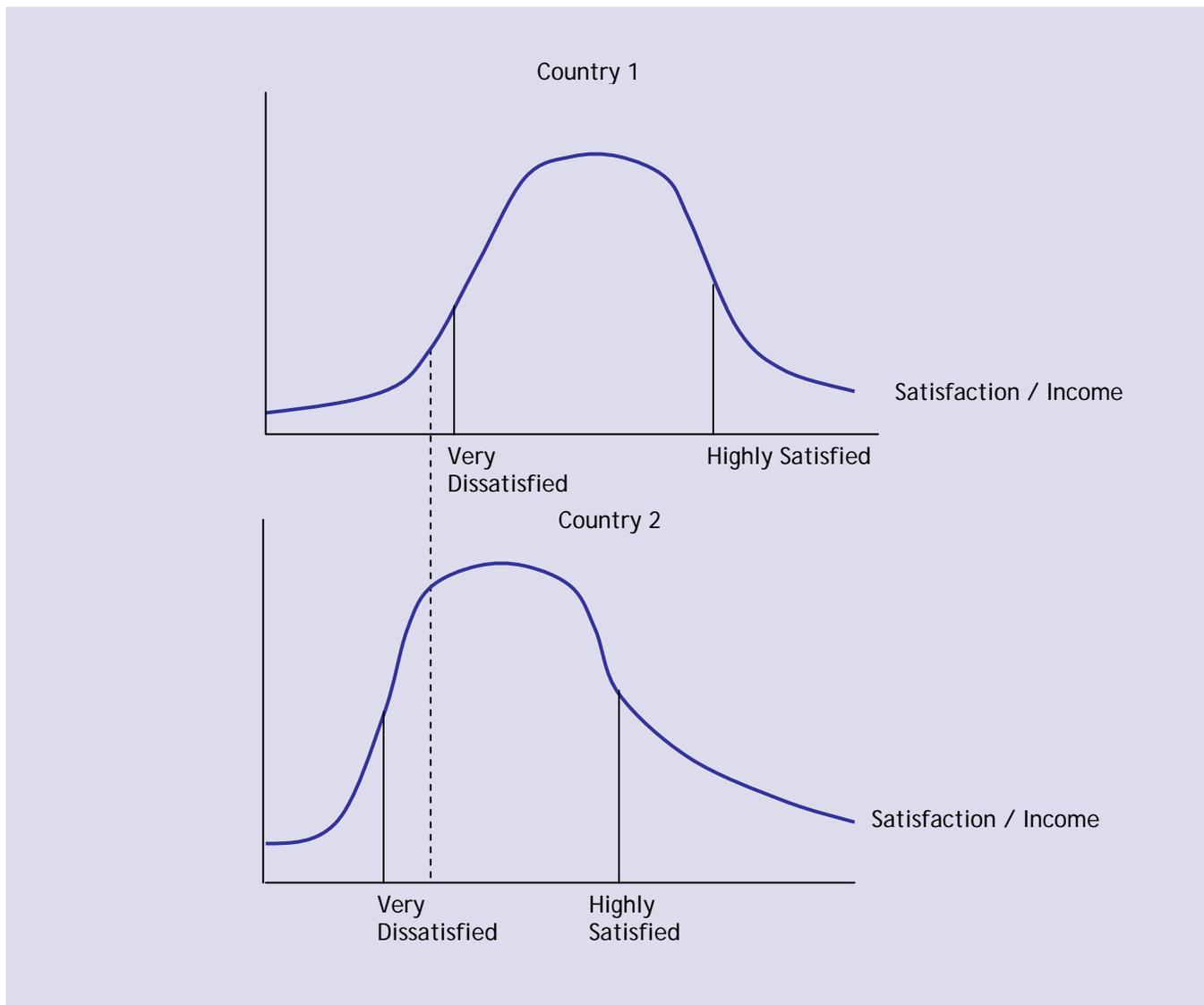


Figure 2: Continuous frequency distribution of satisfaction for all people with a given level of income in countries 1 and 2

Now consider a person in each country with the level of satisfaction indicated by the dotted line. The person in country 1 responds ‘very dissatisfied’ in the survey whereas the person in country 2 responds ‘satisfied’. With the thresholds drawn as indicated, many respondents in country 1 report lower levels of satisfaction than respondents in country 2 for each given level of income even though they have the same level of satisfaction with the income. As a result, the survey suggests that people in country 1 are less satisfied with their incomes, whereas the reverse is actually true.

But if large numbers of people in different countries with the same income and same amount of satisfaction record their satisfaction in different categories because they are using different thresholds for the discrete categories, then econometric estimates of the effects of income on satisfaction within each country are clearly not comparable across countries.

Unfortunately, different groups of people are quite likely, on average, to choose different discrete thresholds in preference surveys. As the example above shows, the bias introduced by the choice of different discrete thresholds can be so large that it reverses the true relationship between satisfaction and income between the groups. The choice of thresholds across different groups of people has to be standardized in some way to remove the bias.

■ Removing Survey Bias with Vignettes

KSvS propose standardizing the choice of thresholds by using what they refer to as anchoring vignettes. The idea is that, in addition to asking survey participants about their satisfaction with their own incomes, they are asked to respond to a constructed example or vignette. Each vignette consists of a hypothetical individual or family household with a particular income level. In their study, the vignettes have four possible income levels: the median income for the participant's country, half the median income, twice the median income, and four times the median income. Different participants receive different vignettes and are asked to indicate how satisfied they think the individual or family in the chosen vignette should be with their given income, using the same discrete categories as for their own incomes.

KSvS then make two assumptions. The first is that each participant chooses the same satisfaction thresholds for his or her own income as for the vignettes, which they refer to as response consistency. The second is that survey participants in two countries (or in two different groups within a country) who receive the same vignette interpret the vignette the same way, which they refer to as vignette equivalence. This follows because the individual or family in an identical vignette must have the same level of satisfaction. Therefore, if the two participants place the same vignette into different categories, this is entirely the result of selecting different thresholds for the discrete categories.

The econometric technique that makes use of the vignette responses is quite involved and need not concern us here. Suffice it to say that under the assumptions of response consistency and vignette equivalence, the responses to the vignettes by the participants in either country can be used to define the thresholds for the participants in both countries. This removes the potential bias that arises from the participants in different countries choosing different thresholds for the discrete categories, and permits an accurate, unbiased comparison of the estimated relationship between satisfaction with income and socioeconomic variables such as the income, age, education, and gender of the respondents in both countries. The technique does not work if either response consistency or vignette equivalence does not hold, and they may not. But they do appear to be reasonable assumptions.

■ The Americans and the Dutch

KSvS demonstrate the advantage of using vignettes based on surveys of satisfaction in the Netherlands and the United States. The surveys are conducted on the Internet. The Netherlands survey is the CentERPanel, which interviews 2,250 households who agree to respond to questions every weekend. The U.S. survey is the RAND American Life Panel, a survey of 1,113 people conducted in 2006-7. The vignettes use the median income of the Netherlands for the Dutch participants and the median income of the United States for the American participants. Participants in each survey were asked to record the satisfaction with their incomes in one of five categories: ‘very satisfied’, ‘satisfied’, ‘neither satisfied nor dissatisfied’, ‘not satisfied’, and ‘very dissatisfied’. The same five categories were used for the vignettes.

The American participants set their thresholds at much lower levels of satisfaction, despite having much higher median and mean incomes than the Dutch. For example, 64% of the Dutch participants said they were satisfied or very satisfied with their incomes, compared with 46% of the American participants. Conversely, 33% of the American participants reported being not satisfied or very dissatisfied with their incomes, compared with 13% of the Dutch participants.³ This is the same grumpy American pattern that has appeared in the satisfaction survey literature: Americans appear to need much higher incomes to achieve the same level of satisfaction as the Dutch.

The same general pattern obtains with the vignettes. For example, in vignettes with income at the median, 45% of the American respondents said that the vignette individual or family should be ‘not satisfied’ compared with 22% of the Dutch for the identical vignette; for half the median income, the percentages of ‘should not be satisfied’ were 47% for the American participants, and 27% for the Dutch participants.⁴ The U.S. participants were clearly more negative in their choice of thresholds for the discrete categories, suggesting the possibility of bias in comparing the American and Dutch survey responses.

KSvS discovered that this was indeed the case. Using the vignettes to standardize the choice of the discrete thresholds removes virtually all the differences in the distribution of income satisfaction in the two countries; indeed, the distribution of satisfaction is virtually identical. A second important result is that satisfaction with income rises much more sharply with income in the United States under the standardized thresholds. The estimated relationship between income and satisfaction with income using the vignette corrected thresholds is twice as steep in the U.S. as the estimated relationship on the uncorrected, raw data, and much steeper than estimates in the literature tend to be. People with high incomes are much more satisfied with their incomes, and people with low incomes are much less satisfied, under the vignette

³ *Ibid.*, Table 1, p. 24.

⁴ *Ibid.*, Table 2, p. 25.

standardized thresholds. As noted earlier, a strong positive relationship between income and satisfaction is consistent with the theory of consumer behavior. KSvS conclude that the Americans are at least as happy as the Dutch (although not necessarily happier—the Easterlin paradox may still apply to some extent since Americans have higher incomes). A third result of note is that the setting of thresholds varies with family size and education in the U.S., with larger families and higher educated individuals being more negative in setting the thresholds.⁵ Thresholds do vary across groups in the U.S. (not so much in the Netherlands, however).

The more general lesson to take from their study is this: Be wary of comparisons between different groups of people based on discrete surveys of satisfaction if no attention is paid to the possibility of different threshold choices across the groups. The use of vignettes is likely to become commonplace in these surveys to overcome the potential biases introduced by differences in setting the satisfaction thresholds.

⁵ The results are discussed in *Ibid.*, pp. 13-21 and presented in Tables 3-5, pp. 26-28.