Netspar Network for Studies on Pensions, Aging and Retirement

Arie Kapteyn

James P. Smith

Arthur van Soest

Life Satisfaction

Discussion Paper 01/2009 - 008

January, 2009

Life Satisfaction

Arie Kapteyn, James P. Smith, RAND & IZA
Arthur van Soest, Netspar, Tilburg University, RAND & IZA[1]

January 2009

JEL code: I31, J28, D31
Keywords: happiness, life satisfaction, Vignettes, Reporting bias

This research was supported by grants from the National Institute on Aging to RAND. We are grateful to Ed Diener, John Helliwell, an anonymous reviewer, workshop participants in Princeton and seminar participants in Tilburg for useful comments.

[1] Tilburg University, P.O. Box 90153, 5000 LE Tilburg, fax: +31-13-4663280; email avas@uvt.nl

Abstract

We analyze the determinants of global life satisfaction in two countries (The Netherlands and the U.S.), by using both self-reports and responses to a battery of vignette questions. We find global life satisfaction of happiness is well-described by four domains: job or daily activities, social contacts and family, health, and income. Among the four domains, social contacts and family have the highest impact on global life satisfaction, followed by job and daily activities and health. Income has the lowest impact.

As in other work, we find that American response styles differ from the Dutch in that Americans are more likely to use the extremes of the scale (either very satisfied or very dissatisfied) than the Dutch, who are more inclined to stay in the middle of the scale.

Although for both Americans and the Dutch, income is the least important determinant of global life satisfaction, it is more important in the U.S. than in The Netherlands. Indeed life satisfaction varies substantially more with income in the U.S. than in The Netherlands.

1. Introduction

Economists have discovered happiness (or rediscovered) or at least research on subjective well-being and its economic correlates (see, e.g., Layard, 2005, or Clark, Frijters and Shields, 2008). The rapidly growing research has touched on several important themes. These have included the so-called Easterlin paradox whereby average happiness remains relatively constant over time in spite of large increases in income per capita (Easterlin, 1974, 1995; see also the chapter by Graham, Chattopadhyay and Picon in this volume). In contrast, within country cross-sectional and panel data almost always show that rising incomes 'buy' additional satisfaction, although the magnitude of the within country cross-sectional effect of income on satisfaction is under dispute (Blanchflower and Oswald, 2004, Di Tella et al, 2007 and Stevenson and Wolfers, 2008). Resolving this paradox, which is often interpreted as a fundamental challenge to the conventional economic theory of utility maximization, has generated a substantial amount of subsequent research attempting to reconcile the finding of a zero correlation between income and life satisfaction in aggregate time series evidence with the positive correlation in cross-section micro-estimates within a given country.

This reconciliation has included adding relative incomes (of others or of oneself in the past) in the utility function (Van de Stadt et al. 1985, Clark et al, 2008) or a sometimes rapid process of adaptation to new circumstances (Di Tella et al, 2003) often labeled the 'hedonic treadmill'(Di Tella et al, 2007). A recent contrary view is provided by Deaton (2008) who documents that if one considers a much wider range of countries arrayed by their level of economic development, the normally positive association of income with subjective life satisfaction reappears. His work also leads to the conclusion that the effect of income on life satisfaction according to cross-country regressions is if anything higher in the high income

countries than in the low income countries. Stevenson and Wolfers (2008) revisit much of the earlier evidence and look at new data to reach similar conclusions.

A considerable amount of research has focused on cross-country differences in subjective well-being, in particular comparing Europe and the U.S. where the US appears to rank lower in satisfaction than many European countries with lower per capita incomes (Alesina et al, 2004, Di Tella et al, 2003, and Blanchflower and Oswald, 2004). For instance, Europeans apparently exhibit a stronger distaste for inequality than do Americans that may be partly explained by a perception of greater mobility in the US (Alesina et al, 2004). Blanchflower and Oswald (2004) study trends in well-being over time in the UK and the US and find that reported levels of well-being have been dropping over time in the US while they have been flat in the UK, despite the fact that in both countries average incomes have grown substantially over the last couple of decades.

A fundamental problem in international comparisons, cross-sectional, and time series analyses of subjective well-being is that one has to assume that somehow response scales are the same across countries, across time, and across groups of respondents within a country. This critical and largely untested assumption becomes even more tenuous if question phrasings change or differ across surveys, as is often the case (see Stevenson and Wolfers, 2008). Here we address these problems head on. In view of the specific interest of economists in the relation between life satisfaction and income, we focus on the role of income.

The population distribution of satisfaction in a country will depend on levels and distribution of incomes. Residents of alternative countries can however differ in the way they translate any given level of income into a subjective level of satisfaction. Moreover, residents of countries may differ in the subjective thresholds that they use in demarcating satisfaction into

discrete categories such as very satisfied or not satisfied. Income distributions, the translation from income to income satisfaction and the demarcation thresholds can all affect differences observed within and between countries in their distribution of stated level of satisfaction. These distinct factors are often confused in the existing literature on life satisfaction and happiness. In our research, we have created unique data sources in two countries—the United States and the Netherlands—and developed a statistical methodology that allows us to separate out these distinct factors.

This paper is divided into seven parts. Section 1 describes the data sources that we developed and will rely on in this analysis. The second section summarizes responses of Dutch and American respondents to questions about their own life satisfaction in several key domains of their lives while the third section describes the types of vignettes we developed and the responses to those vignettes by our Dutch and American respondents. In the next section, we summarize the vignette methodology that serves as the basis of our analysis and then sketch our statistical model that corrects for response scale differences across countries. Section 5 presents our main empirical results and their implications for interpreting observed differences in life satisfaction in the two countries. In section 6, simulations based on our estimated models are used to ascertain what Dutch distributions of life satisfaction would be if the Dutch had American parameters and thresholds rather than their own. The final section highlights our main conclusions.

1.1 Data Sources

Our analysis in this paper is based on information obtained from two Internet surveys, which we conducted in the Netherlands and the United States. For The Netherlands, we used CentERpanel, administered by CentERdata affiliated with Tilburg University. CentERpanel

includes about 2,250 households who have agreed to respond to questions every weekend over the Internet. The Dutch sample is not restricted to households with their own Internet access. Respondents are recruited by telephone. If they agree to participate and do not have Internet access, they are provided with Internet access (and if necessary, a set-top box). Thus, CentERpanel is representative of the adult Dutch population except the institutionalized. Sampling weights provided by CentERdata and based upon comparing with a much larger survey of Statistics Netherlands are used to correct for unit non-response and attrition. The sample used for estimation has 2,244 respondents who participated in an interview with questions on life satisfaction (self-assessments as well as vignettes) in 2006. From multiple waves collected in the past, CentERpanel has a rich set of variables on demographic, health, and economic characteristics of respondents. In our analysis we use the most recent measurement of these background variables, reported a few months before our life satisfaction survey. The Internet infrastructure makes CentERpanel an extremely valuable tool to conduct experiments, with possibilities for randomization of content. Production lags are very short, with about one month between module design and data delivery.

Our Internet survey for the United States is the RAND American Life Panel (ALP). This panel was initially recruited from respondents age 40 plus in the Monthly Survey (MS) of Michigan's Survey Research Center but has been subsequently supplemented with younger respondents.[2] Similar background information was collected for these respondents as was available for Dutch respondents. The American sample used for estimation consists of 1,113

[2]The MS, the leading consumer sentiments survey, produces the widely used Index of Consumer Attitudes. MS respondents are asked if they have Internet access and, if yes, if they are willing to participate in Internet surveys. Those who agree are added to our household panel to be interviewed regularly over the Internet. As with the CentERpanel, respondents who do not have Internet access are provided with a set top box (an MSN Web TV) that allows them to browse the Internet and send and receive email.

respondents interviewed during 2006-2007. The American data are weighted to match Current Population Survey demographic distributions in age, education, and gender.

2 Global Life Satisfaction by domain

Respondents are asked to rate themselves on a five point scale. They do so in the following specific life domains: income, family and social relations, job, and health. They are also asked a global question on their own life satisfaction. The scale that is used is the same for all domains: (*very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very dissatisfied*). The exact self-assessment questions of life satisfaction are contained in Table 1.

Table 1 summarizes responses obtained from the Dutch and American samples for the four domains of income, social contacts and family life, job or other daily activities, and health. The last panel in Table 1 presents responses to the question regarding global life satisfaction. For all domains and for global satisfaction, the distributions in the US and the Netherlands are significantly different (see the tests reported at the bottom of each panel).

Before turning to between country differences, it is useful to first highlight some differences across the domains. Both the Dutch and Americans appear to be less satisfied with their incomes than with the other domains. The health domain is next, at least in terms of most negative responses, followed by job and daily activities, with respondents in both countries most satisfied with their lives in the family and social contacts domain. Differences among the Dutch in the three domains besides income are relatively small with sharper distinctions present in these three domains among Americans. Finally, a much smaller proportion of both Dutch and American respondents appear to be dissatisfied with their lives when answering a global life satisfaction question than their answers in each of the four specific life domains would indicate. This appears to be due to relatively modest correlations in dissatisfaction across domains, so that

dissatisfaction in one domain may be compensated by satisfaction in a different domain. In fact, across the two samples, less than one percent of all respondents are not satisfied or very dissatisfied in all four domains and among those virtually all (94%) report to also be not satisfied or very dissatisfied with their life in general.

Turning to between country differences in life satisfaction, consider first how satisfied respondents in the two countries are with the total household income. As we have analyzed in more detail elsewhere[3], Americans are much less satisfied with their incomes than the Dutch are in spite of the fact that on average their incomes are considerably higher. Sixty-four percent of Dutch respondents say that they are either satisfied or very satisfied with their total household income. The comparable fraction for Americans is 46%- eighteen percentage points lower than the Dutch. Similarly, a much larger fraction of Americans respond that they are either not satisfied or very dissatisfied- a third of Americans compared to thirteen percent amongst the Dutch.

The Dutch are also more satisfied with their jobs than Americans are but these differences are smaller than those in the income domain. Fourteen percent of Americans are either not satisfied or very dissatisfied with their jobs compared to four percent of the Dutch- a differential of ten percentage points which is half as large as the differential in the income satisfaction domain. At the top of the scale, more than four in every five Dutch respondents are at least satisfied with their jobs as are more than two-thirds of Americans. In both countries, respondents are much more satisfied with their job and other daily activities than they are with their incomes.

[3] We analyze answers to questions on income satisfaction in depth in Kapteyn, Smith, and Van Soest (2008).

There are actually more Americans very satisfied as well as very dissatisfied with the social aspects of their lives compared to the Dutch. Relatively few respondents in either country register displeasure (not satisfied or very dissatisfied) with this domain, although once again Americans are more likely to go to the bottom of the scale compared to the Dutch (9% compared to 2%). This avoidance of extremes is a common feature of Dutch responses to subjective scale questions and is similar to what we have found in prior work (see for example Kapteyn, Smith and VanSoest, 2007). This tendency may well have its origins in the Dutch culture. According to Wikipedia, "The Dutch typically see their countrymen as sober, practical and down-to-earth people. Any form of ostentation is likely to be criticized, and straightforwardness is generally appreciated."

The final specific domain on which we asked about life satisfaction was health. Based on objective health measures, the Dutch are a healthier population than the Americans (Kapteyn, Smith, and Van Soest, 2007). In this case, that objective difference is reflected in their subjective answers to their satisfaction with their health. Nineteen percent of Americans are either not satisfied or very dissatisfied with their health compared to 8% of the Dutch.

The final panel in Table 1 displays the distribution of answers to questions evaluating own global life satisfaction (satisfaction with life – SWL). Eighty-eight percent of the Dutch are either satisfied or very satisfied compared to seventy-eight percent of the Americans. Similarly, while most respondents in both countries appear to relatively satisfied with their lives, 6.4% of the Americans say that they are at a minimum not satisfied compared to only 1.4% of the Dutch.

Table 2 examines the relationship between responses to the global life satisfaction question to the level of satisfaction within the four specific life domains. It does so by listing coefficient estimates (and the associated 'z' values) from an ordered probit model of global life

satisfaction. The initial set of regressors in the first two columns are responses to life satisfaction in the four specific life domains each indexed on a scale of one to five. Main effects are estimated Dutch coefficients while the US interactions test for differences between Americans and Dutch.[4]

As expected, these results show that satisfaction with life is positively associated with satisfaction within each of the four domains. As indicated by the estimated coefficients within each domain, income satisfaction received by far the lowest weight in global satisfaction with the health domain in second to last place. The highest weight is in the family and social relations domain.[5] While there is not much evidence of statistically significant differences between the two countries in the translation from satisfaction within a domain into global life satisfaction, there appears to be less weight in the US assigned to the health domain. Remember that the coding goes from very satisfied to very dissatisfied, so the negative sign on the US dummy means that US respondents are happier, keeping satisfaction in each domain constant. That result however is not statistically significant.

The second model in Table 2 adds a number of standard demographics to this model including age, marital status, education, gender, and income and once again allows all estimated effects to differ between the Dutch and Americans. All in all, the evidence for the need for demographics or interactions of these with the US dummy is very weak. A test of the null that the effects of the demographics are equal to zero does not lead to rejection. Thus, it seems a model with just the domain specific satisfaction variables is sufficient.

---

[4] To keep the specification parsimonious and following van Praag and Ferrer-i-Carbonel (2008), we include domain satisfactions as cardinal variables. Using dummies gives qualitatively similar results.

[5] Van Praag and Ferrer-i-Carbonell (2008, Chapter 4) perform similar regressions using panel data on Germany and the UK. They find a much larger role for satisfaction with the financial situation than we do for satisfaction with income. For the UK, they also find that social contacts are the most important factor; they do not have satisfaction with social contacts in the German data.

There is a slightly different way of interpreting this outcome. If we state as a null hypothesis that global life satisfaction is a function of just the four domains we consider here, then the test would not reject that null. Of course the power of that test will depend on how much the possibly omitted satisfaction dimensions are correlated with the demographics included in Table 2.

3.1 Description of Vignettes

In addition to their ratings of their own life satisfaction, respondents were given a set of vignettes cover the four life domains- income, family relations, job, and health. These domains were chosen because the current literature has documented them as key determinants of overall life satisfaction (see Easterlin, 2006). All vignettes were given with either a female or male name, which was randomized across respondents. Within each domain, vignettes were presented in random order to eliminate any possibility of order effects whereby the initial vignette presented could affect the ranking of subsequent vignettes.[6] Comparing rank ordering of vignette evaluations across respondents shows that different respondents tend to order vignettes in the same way. The scale that is used is the same for all domains: (*very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very satisfied*).

The vignette questions in the income domain specify an income for the vignette person that is selected randomly with the values being equal to either the median income in the Netherlands or the United States or a value that is half, twice, or four times the median income in each country. The six family relation vignettes vary conditions in the vignette person's family or friends life, including whether the vignette person is married or single, has children, has many or

[6] In earlier work on health and work disability vignettes, we found some effects of the order of vignettes within each domain on the vignette evaluations. It might also be useful to randomize the order in which self-assessments by domain are presented but we did not do this.

few friends and the nature of the relationship with these important others. The work vignettes focus on whether the vignette person is working or retired, the amount of hours worked the security of the job, and how stressful the job is. Finally, the four health domain vignettes vary conditions around the vignette person's ability to engage in (light, vigorous) exercise, and possible problems with sleep, anxiety, and depression.

In addition to these domain specific life satisfaction questions, respondents are also given a subset of ten possible vignettes on global life satisfaction. These global life satisfaction vignettes succinctly describe the vignette person in a single vignette across the four sub-domains mentioned above- family relations, work, income, and health, combining the descriptions given in the domain specific vignettes. This global vignette approach has the advantage of moving directly to an overall measure of life satisfaction and for that reason we will use them in the analysis in this paper. The domain specific vignettes are not analyzed in the current paper; the satisfaction with income vignettes are analyzed in Kapteyn et al. (2008).

The specific scenarios described in the global satisfaction vignettes are listed in the Appendix. Table 3 presents a summary of the type of variation present in the ten global vignettes. The six ways that the global vignettes can vary are gender, age, family, income, work, and health. Gender of the vignette person is randomly assigned to respondents through the feminine or masculine name of the vignette person. An exact age is always mentioned in each vignette and these ages are listed in the first column of Table 3.

With one exception (vignette 3 where income is always 'modest'), income is also randomly assigned in the vignette with up to four values assigned- (ranging from half the median, median, twice the median, and four times the median). In most scenarios, only three of the four possible income values are used.

The overall situation described in the family and social relations dimension relates to spouse, children, and/or the presence of close friends. In four of the vignettes, the social situation can reasonably be described as good, in one vignette as moderate, and in the other five it is problematic in some important way. The aspects of the work environment that are mentioned in the vignettes are stress, hours, control, security, and retirement status. In five vignettes, the vignette person has already retired. Finally, four of the vignettes describe a person in good health while the remainder point to some type of health problem ranging from moderate to serious.

3.2 Responses to Vignette questions on global life satisfaction

Table 4 lists the distribution of responses obtained in both countries for each global life satisfaction vignette. We divide these vignettes into three groups based on the age of the vignette person- those who have retired, those of the young, and those for middle age vignette persons. To eliminate the impact of income variation within the vignette, our comparisons in Table 4 apply to the highest income value mentioned in the vignette. The numbers in the rows at the top of the columns match the numbers of vignette in Appendix A. The rows below describe the main attributes of the vignette person in terms of their age, family relations, health, and job situation.

Let's first examine the panel for the five global vignettes for retired persons. In addition to a limited age variation within this group, the only meaningful variation in these vignette descriptions concerns the quality of family life and the health of the respondent. It is useful to start with the best case of the possible scenarios- vignette 8 where both the social life and health of the vignette respondent is good. In this case, vignette evaluations of both Dutch and American respondents mirror this best case description as three quarters of respondents say that one should be very satisfied and less than one percent respond as either not or very dissatisfied. When the

life situation appears to be very good as it certainly is in this case, a larger fraction of Americans answer very satisfied- 77% compared to 71% for the Dutch.

Health status has a major impact in these evaluations of retired vignette persons. Even with vignette number 7 where the vignette person suffers only from arthritis but has many friends, there is a very sharp reduction in the percent of respondents who would be satisfied. Less than one in five of Dutch and American respondents apply the label 'very satisfied' compared to at least 70% for the healthy vignette 8. As was the case with vignette 7, somewhat more Americans say 'very satisfied' with vignette 8 compared to the Dutch.

When the health situation of the vignette person gets even worse as in vignettes 3, 6, and 9, large fractions of Dutch and American respondents choose 'not satisfied' or 'very dissatisfied'. But this negative reaction to poor health appears to be more dramatic in the American sample. Using vignette 3 to illustrate the point, 59% of Americans respond as not dissatisfied or very dissatisfied compared to only 26% of the Dutch. This result seems in contrast to Table 2 where we found that life satisfaction depends to a lesser extent on health satisfaction in the US than in the Netherlands. Perhaps the result is misleading since we do not have an orthogonal design – variation in health is correlated with other variations in the vignette characteristics – and we need regressions to control for other differences in the vignette characteristics (see Table 5 below). Vignettes among the retired suggest the importance of health for life satisfaction among the retired population.

Turn next to the two vignettes pertaining to the young. Age and even health are basically the same in these two vignettes so that they differ only how they describe the family and work domains. Compared to vignette 5 where the social situation is good and the job situation is secure, the person in vignette 4 has no friends and does not feel secure about his/her job. Those

two problems take their toll on the evaluation of global life satisfaction. There is a tenfold reduction in the percent of respondents who are very satisfied and the percent of respondents who are not satisfied or very dissatisfied increases from about one percent to over forty percent. Even among the young, where income and income growth is quite central to their lives, a negative situation in terms of either job security or friendships reduces overall life satisfaction a great deal, at least that is what our respondents believe.

The final subset of global life satisfaction vignettes describes the middle aged. For them also, the age spread in the vignettes is quite limited so that the principal variation across alternative vignettes relates to job, social relations, and health. In some respects, vignette 10 may be the most interesting. In this case, the work situation is good and all vignettes are either assigned an income twice the median or four times the median. The answers summarized in Table 4 are for the four times median income in each country so that the economic situation of the household is very good in the economic/work domains. However, both the family life and health are not good and as a consequence many respondents evaluate the life situation described quite negatively. More than half of both Dutch and American respondents are not satisfied or are very dissatisfied with this vignette person's life.

Evaluating the independent impact of social relations, health, and job is difficult in the global vignettes as all three dimensions vary simultaneously across these vignettes. Table 5 presents an alternative approach for doing so, by showing the results of a regression of the rating of the vignettes on their main characteristics. The characterization chosen here is a little different from the one shown in Tables 3 and 4. Since there are only ten vignettes we have to choose a parsimonious set of indicators to characterize the vignettes, in order to be able to identify their effect. The differences in comparison with Table 3 are as follows. We characterize family and

social relations by means of two variables: whether the vignette person is married (vignettes 1, 5, 6, 9, 10) or not and whether the relation with family members or friends is good (vignettes 1, 2, 3, 5, 7, 8) or not. Work is characterized by whether one works long hours (50 per week or more, vignettes 2 and 4) or not and whether the vignette person has control over his or her job or job security (vignettes 2, 5, and 10) or not. Health, finally, is coded as good (vignettes 2, 4, 5, 8) or not.

Table 5 presents the results of an ordered probit, a regression, and a probit where we have combined dissatisfied and very dissatisfied into one category and the other three answers into another category. The latter has been included to investigate whether patterns are different at lower end of the scale. As one can see, qualitatively the results of the regression and the ordered probit are very similar. We will only discuss the regression, as one can immediately interpret the size of the effects, whereas that is more complicated in the ordered probit case. At the end we will comment on cases where the probit deviates from the regression and the ordered probit.

As before, the top panel presents the results for the Dutch sample, while the bottom panel presents the effects of interactions with a US dummy. The US dummy is negative, but insignificant. Thus the evidence for a uniform scale shift is limited. The difference in response patterns are more subtle than just a uniform shift in the scale used in both countries.

One observes that in the U.S. there is some evidence that a higher age is assumed to be associated with lower life satisfaction (remember once again that the scale runs from 1 for very satisfied to 5 for very dissatisfied). For instance, in the U.S. an age difference of 40 years would be associated with a deterioration in life satisfaction of .40 on the five-point scale. At first sight, being married is not considered to be a good thing, particularly in the US. Notice however that this has to be looked at in combination with the quality of the relation with family and friends.

Since good relations have an enormous effect of life satisfaction, being in a good marriage is a major source of happiness. Both countries attach a positive value to retirement, but the Americans more so than the Dutch.

Working long hours is believed to have a negative effect on life satisfaction, but more so in the U.S. than in The Netherlands, suggesting that working conditions are more attractive in the Netherlands than in the US. Having control over one's job or having a secure job is valued more positively in the U.S. than in The Netherlands. Good health is worth about half a point in The Netherlands and about three tenth of a point in the US, but the difference between the two countries is only marginally significant. The larger effect in the Netherlands is in line with Table 2 and shows that the differences in Table 4 are indeed somewhat misleading – they disappear when other vignette characteristics are controlled for.

Income has a positive effect on life satisfaction in both countries, although the effect is not all that large. This is in line with the results presented in Table 2, which showed that of the four domains, income is the least important one for global life satisfaction. The one exception is the dummy for modest income, which is quite large, but since this dummy uniquely identifies vignette 3, its coefficient really identifies the evaluation of that vignette as a whole, not just the income level. We note that income is thought to be more important in the US than in The Netherlands.

The binary probit which explains the probability of rating a vignette as dissatisfied or very dissatisfied shows some minor deviations from the results in the other two columns. Both in The Netherlands and the U.S., increasing age is associated with a greater probability of being rated dissatisfied or very dissatisfied. We also note that having a secure job or a job over which one exerts control reduces the chances of being rated dissatisfied or very dissatisfied.

4.1. The Theory of Vignettes

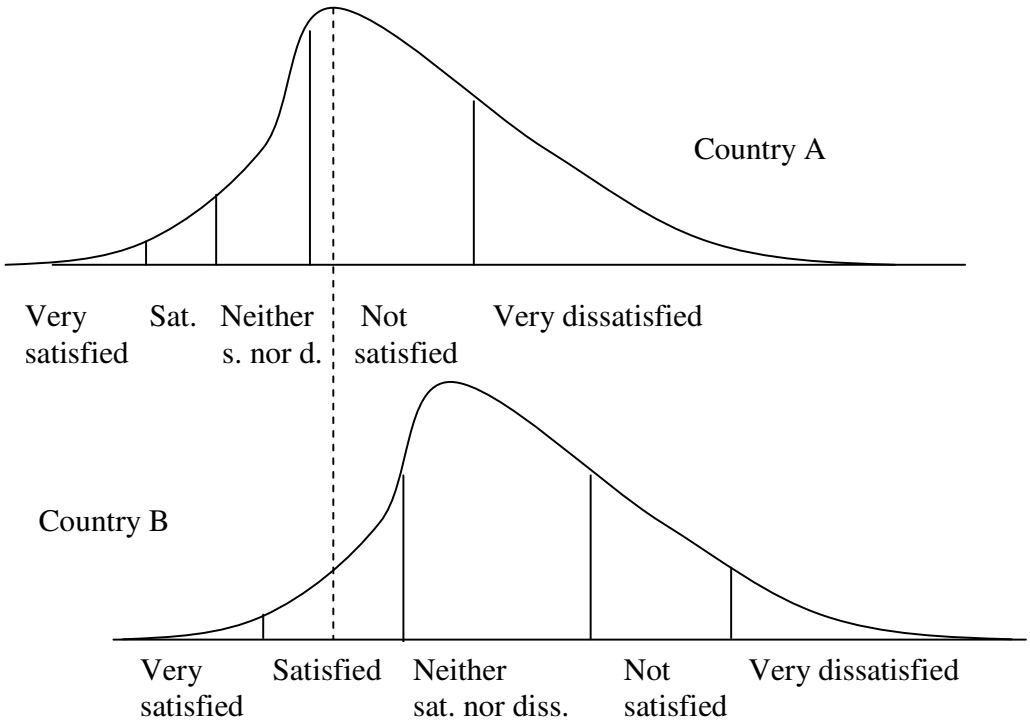
In this section, we provide an intuitive description of the use of vignettes for identifying response scale differences and then sketch our statistical approach. The basic idea is illustrated in Figure 1, which presents the distribution of life satisfaction or happiness in two hypothetical countries. The density of the continuous happiness variable in country A is to the left of that in country B, implying that on average, people in country A are happier than in country B. The people in the two countries, however, use very different response scales if asked to report their happiness on a five-point scale (very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very dissatisfied).

In the example in the figure, people in country B attach much more positive labels to given points on the life satisfaction scale than do people in country A. Someone in country A with the life satisfaction indicated by the dashed line would report to be not satisfied, while a person in country B with the same actual satisfaction would report to be satisfied. The frequency distribution of the self-reports in the two countries would suggest that people in country B are more satisfied than those in country A—the opposite of the true distribution. Correcting for the differences in the response scales (DIF, "differential item functioning," in the terminology of King et al., 2004) is essential to compare the actual health distributions in the two countries.

Vignettes can be used to do the correction. A vignette question describes the satisfaction of a hypothetical person and then asks the respondent to evaluate the satisfaction of that person on the same five-point scale that was used for the self-report of their satisfaction. Since the vignette descriptions are the same in the two countries, the vignette persons in the two countries have the same actual life satisfaction or happiness. For example, respondents can be asked to

evaluate the life satisfaction of a person whose satisfaction is given by the dashed line. In country B, this will be evaluated as "satisfied." In country A, the evaluation would be "not satisfied." Since the actual level of satisfaction is the same in the two countries, the difference in the country evaluations must be due to DIF.

Figure 1. Comparing self-reported happiness across two countries in case of DIF



Vignette evaluations thus help to identify differences between the response scales. Using the scales in one of the two countries as the benchmark, the distribution of evaluations in the other country can be adjusted by evaluating them on the benchmark scale. The corrected distribution of the evaluations can then be compared to that in the benchmark country—they are now on the same scale. In the example in the figure, this will lead to the correct conclusion that people in country A are more satisfied than those in country B, on average. The underlying

assumption is *response consistency:* a given respondent uses the same scale for self-reports and the vignette evaluations.

We will apply the vignette approach to life satisfaction, using vignettes not only to obtain international comparisons corrected for DIF, but also for comparisons of different groups within a given country.

4.2 Econometric Model

We will apply the vignette approach to life satisfaction, using vignettes not only to obtain international comparisons corrected for DIF, but also for comparisons of different groups within a given country. Our model explains respondents' self-reports on satisfaction by themselves as well as their reports on satisfaction of hypothetical vignette persons. Self-reports are modeled as a function of respondent characteristics X_i (including demographics, a country dummy and interactions of all demographics with that dummy) and an error term ε_i by the following ordered response equation:

$$Y_i^* = X_i\beta + \varepsilon_i;\ \ \varepsilon_i \sim N(0,\sigma^2),\ \varepsilon_i \text{ independent of } X_i \tag{1.1}$$

$$Y_i = j \text{ if } \tau_i^{j-1} < Y_i^* \le \tau_i^j,\ \ j = 1,\ldots 5 \tag{1.2}$$

The thresholds τ_j^i between the categories are given by

$$\begin{aligned} &\tau_i^0 = -\infty,\ \tau_i^5 = \infty,\ \tau_i^1 = \gamma^1 X_i + u_i,\ \tau_i^j = \tau_i^{j-1} + \exp(\gamma^j X_i),\ j = 2,3,4 \\ &u_i \sim N(0,\sigma_u^2),\ u_i \text{ independent of } X_i \text{ and the other error terms in the model} \end{aligned} \tag{1.3}$$

Since X_i includes a country dummy and interactions of all demographics with that dummy, this specification allows for completely different ways in which the response scales vary with demographics in the two countries. The various cut-off points can also vary in different ways, which seems useful because of the observed tendency of the Dutch to avoid extremes, suggesting

that a Dutch respondent will have a lower first cut-off point but a higher last cut-off point than a similar US respondent.

As noted before, the fact that different respondents *i* use different response scales τ_i^j is called "differential item functioning" (DIF). The term u_i introduces an unobserved individual effect in the response scale. It will imply that reported evaluations of different vignettes (see (1.5) below) are positively correlated with each other and with self-reports (conditional on X_i), since some respondents will tend to use high thresholds and others will use low thresholds in all their reports. Since such positive correlations are observed in the data, incorporating u_i helps to improve the model's ability to predict the observed outcomes (the model fit).

Define a benchmark respondent with characteristics $X_i = X(B)$. The DIF adjustment involves comparing Y_i^* to thresholds τ_B^j rather than τ_i^j, where τ_B^j is obtained in the same way as τ_i^j but using $X(B)$ instead of X_i. A respondent's reported satisfaction is computed using a benchmark scale instead of a respondent's own scale. This does not give an adjusted score for each individual (since Y_i^* is not observed) but it can be used to simulate adjusted *distributions* of Y_i for the whole population or conditional upon some of the characteristics in X_i.

Using self-reports on own life satisfaction only, parameters β and γ^1 are not separately identified, only the difference between β and γ^1. For example, consider country dummies: people in two different countries can have systematically different life satisfactions, but if the scales on which they report their life satisfaction can also differ across countries, then self-reports are not enough to identify the satisfaction difference between the countries. The vignettes will be used to identify β and γ^1 separately.

The evaluations Y_{li} of vignettes *l=1,…,L=10* are modeled using similar ordered response equations:

(1.4) $$Y_{li}^{*} = \theta_l + \kappa I_{li} + \varepsilon_{li}$$

(1.5) $$Y_{li} = j \text{ if } \tau_i^{j-1} < Y_{li}^{*} \le \tau_i^{j},\ j = 1,...5$$

(1.6) $$\varepsilon_{li} \sim N(0, \sigma^2), \text{ independent of each other, of } \varepsilon_{ri} \text{ and of } X_i$$

Thus we include a dummy for each of the 10 vignettes and allow the evaluations to depend on the log of the income assigned to the vignette (I_{li}), which is randomized across respondents. The unobserved vignette evaluations Y_{li}^{*} do not depend on respondent characteristics X_i (the assumption of *vignette equivalence).* The actually reported evaluations Y_{li} do depend on X_i, but only through the thresholds. The maintained assumption here is that of *"response consistency"*, meaning that the thresholds τ_i^j are the same for self-reports and the vignettes.

With these assumptions, it is clear how vignette evaluations can separately identify β and γ $(=\gamma^1,...,\gamma^5)$: From the vignette evaluations alone, γ, θ, $\theta_1,...\theta_{10}$ and κ can be identified (up to the usual normalization of scale and location). From self-reports, β can then be identified in addition. Thus the vignettes can be used to solve the identification problem due to DIF.

The two assumptions *vignette equivalence* and *response consistency* are crucial for solving the identification problem. Vignette equivalence may be problematic if life satisfaction is multidimensional and the weights are different in the two countries. The fact that in Table 2, the interactions between domain satisfactions and the US dummy are jointly insignificant suggests that this is not a serious problem in our case. *Response consistency* may be violated if, for example, people make systematic mistakes in evaluating vignette persons but are much better

able to evaluate their own satisfaction. *Response consistency* can be tested if an objective measure is available and such tests have typically supported the use of vignettes (see King et al., 2004, on vision, and Van Soest et al., 2007, on drinking behavior), but an objective measure of satisfaction with life seems hard to give.

5.1 Empirical Results

This section highlights our main empirical findings. We discuss our main parameter estimates determining overall satisfaction with life and assess the consequences of different threshold parameters in both countries.

The model presented above was estimated using the self-evaluations and vignettes in the Dutch CentERpanel and the RAND American Life Panel. The equations for global life satisfaction and for the response thresholds include a complete set of interactions with a country dummy for the United States. We also estimated the simpler model that does not allow for DIF. This amounts to a standard ordered probit for self-assessed satisfaction.

5.2 Model of Global Life Satisfaction

Table 6 lists parameter estimates for two models explaining global life satisfaction, where the scale is from good to bad (1: very satisfied, …, 5: very dissatisfied). All regressors in these models (except the country dummy) are measured in deviations of their country specific means, which makes it easier to interpret the constant term and most importantly the implications of the US dummy. Demographic regressors include dummy variables for whether the respondent is female, married, age brackets 40-50, 51-64, 65+ (the left out group is under 40 years old). Education is separated into three groups- low, medium or high with the low education group the left out category. Income is measured as log-equivalized family income where income is

adjusted by the logarithm of family size. Log-family size is also a separate regressor, in part to test for the adequacy of this choice of functional form for the equivalence scale. Finally, a dummy variable is included indicating whether the respondent is working.

For reasons outlined above, our preferred model is the model with DIF (adjusting for threshold differences). It is listed in the first two columns of Table 6. In the Dutch sample, there are no significant differences in satisfaction with life by gender or age. Higher income makes the Dutch more satisfied with their life. Conditional on income, higher education also makes the Dutch more satisfied. Since education is typically associated with higher income, this most likely reflects the fact that education is a reasonable proxy for permanent income of respondents. Finally, conditional on the equivalized income, married Dutch respondents and those with larger families are more satisfied with their lives. One interpretation of this finding is that marriage and family are on average a source of well-being for these households. Dutch respondents who work are more satisfied with their lives than those who do not.

Turn next to our estimates of the differences in parameters between the two countries which implicitly set the US parameters. Since regressors are measured in deviations from within-country means, the coefficient on the US dummy gives the difference between the average US person and the average Dutch person, whose characteristics are different (see also Section 6 for the consequences of these differences in the demographics). This coefficient is positive but insignificant, suggesting that the average Dutch and US respondents have similar satisfaction with life, according to both the model with and the model without DIF. Similar to the Dutch, there are no gender differences in life satisfaction among the Americans but the estimated age patterns indicate that life satisfaction among Americans increases with age and that retired Americans are particularly satisfied with their lives. There is no differential impact of work

among Americans. Here the contrast with some of the results in Table 5 is of interest. When evaluating vignettes, Americans seemed to think that getting older would reduce life satisfaction. Yet when it comes to their own satisfaction, getting older is a good thing. None of these effects are strongly significant, but they still would seem to cast some doubt on the assumption that Americans are able to evaluate the vignette persons in the same way as they evaluate themselves (response consistency) or that respondents of different ages evaluate the same vignette differently (vignette equivalence).

To explore this further we have included interactions between the respondents' own age (coded as dummies as in Table 6) and the age of the vignette person in the regressions reported in Table 5. These interactions turn out to be totally insignificant, as they should be under the assumption of vignette equivalence. This implies that there is no evidence that respondents make systematic errors in evaluating vignettes describing persons of different ages than their own age.

The most important variable for comparing the two countries is income. The impact of income in improving life satisfaction is much more pronounced in the US than in The Netherlands (more than four times larger in the US in the model with DIF).[7] Since we estimate no intercept difference between the countries and the data are all demeaned within countries, the Dutch and Americans are about equally satisfied at their country specific mean incomes. But Americans become more satisfied with life at high incomes levels and much less satisfied than the Dutch at low incomes levels.

Another important question is how the corrections for threshold differences within and across countries affect our interpretation of the determinants of life satisfaction. This question is

[7] The coefficients of log income in the US (0.504 in the model with DIF, 0.425 in the model without DIF) seem rather large compared to the coefficients reported in the chapter by Layard, Mayraz and Nickell in this volume (between 0.33 and 0.58 on a ten-point SWB scale, where we use a five-point scale). The coefficients for the Netherlands are much smaller than what Layard et al. find for European countries.

addressed by comparing the parameter estimates in the model without DIF to the model with DIF. Several estimated effects seem rather similar between the two models. We note however that for the Dutch the estimated effects of education and working are larger in the model with DIF than in the model without. Considering the interactions with the US dummy, the effect of income on life satisfaction in the US turns out to be more pronounced when we correct for DIF.

These differences between the models with and without DIF are of course directly related to the estimated equations for the thresholds in the model with DIF. For instance, consider the effect of log income interacted with the US dummy. The negative coefficient for this variable in the first threshold equation means that the first threshold shifts to the left when log income increases in the US. As a result of that, a response is less likely to lie to the left of that threshold. Since this effect of log income on the first threshold explains part of the existing negative correlation between income and life satisfaction, incorporating the effect on the threshold reduces the negative effect of log income in the US on self rated global satisfaction. This explains the difference of the income effects in the US on life satisfaction in the models with and without DIF. One should note however, that all thresholds play a role, not only the first one. Disentangling the effect of the threshold shifts may be a complicated matter. We prefer therefore to investigate the importance of threshold differences between countries and between demographic groups within countries by a series of simulations.

6. Model Simulations

A transparent way of understanding the implications of our approach is to simulate the distribution of life satisfaction in the two countries for different parameter values. Essentially we first simulate the Dutch distribution of self-reported life satisfaction and then replace various sets of parameters by the corresponding American values. Table 8 presents the results of these

simulations by four age groups—those less than 40, 40-50 years old, 50-64 years old, and at least 65 years old. The first row for each age group summarizes the distribution of satisfaction with income for the Dutch using their own parameters. The second row replaces Dutch thresholds by American thresholds (cf. Table 7).The third row simulates the Dutch distribution if we replace the parameters in the Dutch satisfaction equation (i.e. Table 6 with DIF) by the American parameters. The fourth row replaces all Dutch parameters by American parameters. The fifth row simulates distributions for the American sample using American parameters. Table 9 lists similar simulations by income quartile instead of age.

For each age group in Table 8, the first row approximately reproduces the distribution of self-reports in the Dutch sample, while the fifth row does the same for the US sample. Comparing the first two rows in each panel shows that the Dutch self reports would become more spread out when Dutch respondents would evaluate their satisfaction with life using US thresholds. Both the percentage very satisfied and the percentage dissatisfied go up. This corresponds to the notion that the Dutch tend to avoid extremes; giving them the US thresholds makes them more likely to report the two extreme categories. Comparing rows 1 and 2 with row 5 then shows that correcting for response scale differences does not make the distribution of life satisfaction in the Netherlands and the US more similar in all respects. For example, for all age groups combined (final panel), we find that after the correction a much larger fraction in the Dutch sample are very satisfied with their life than in the US sample. The fraction not satisfied/ dissatisfied or worse increases somewhat in the Dutch sample and comes somewhat closer to the US fraction, but remains substantially smaller. Both before and after correction for response scale differences, the Dutch population as a whole is more satisfied with their lives than the Americans. This does not apply to the oldest age group, however: Americans of 65 years and

older are somewhat more satisfied with their lives than their Dutch counterparts, on average, irrespective of whether we give them the same scales or not.

Rows 3 and 4 in each panel can be used to show how much of the remaining differences (keeping response scales constant across countries) is due to differences in observed characteristics, generalizing the traditional Oaxaca-Blinder decomposition to a non-linear model (cf., e.g., Yun, 2004). In particular, comparing rows 4 and 5 shows the differences explained by differences in background characteristics between the two countries, using US evaluation standards (both in the self-assessment equation and for the thresholds). The results show that, although the differences are modest, the characteristics make the Dutch in all age groups more satisfied with their lives than the Americans. The most important characteristic driving this is partnership status: having a partner has a strong positive effect on satisfaction with life, and the fraction with partner is much higher in the Netherlands than in the US (78% versus 64%).

On the other hand, comparing rows 2 and 3 in each panel of Table 8 shows that giving the Dutch the US parameters for the self-assessment (but keeping the Dutch thresholds) also brings about substantial shifts, where for younger ages the imposition of US parameters on Dutch respondents leads to lower simulated satisfaction, while for higher ages it leads to more satisfaction. This is a direct reflection of the results in Table 7, which show rather strong interaction effects with the US dummy for the age brackets 40-50 and 51-64.

Next, let's turn our attention to Table 9, which does the same thing as Table 8 but for income quartiles instead of age groups. The effect of assigning US thresholds again leads to more dispersion in the responses (row 2 compared to row 1). For the highest income quartile, comparing row 2 and row 5 shows that the US respondents are better off. This was not clear

from the first row, due to the reluctance of the Dutch to classify themselves as dissatisfied or very dissatisfied.

Assigning US self assessment parameters to the Dutch confirms the stronger effect of income on life satisfaction in the US than in The Netherlands (row 3 versus row 2). We see that with the US self assessment parameters, Dutch respondents with low incomes would be considerably less satisfied. Conversely with high incomes they would be more satisfied. When the Dutch are assigned both the US self assessment parameters and the US thresholds, then the satisfaction distribution more closely resembles that of the US (rows 4 and 5), and again show that the differences in background characteristics somewhat favor the Dutch, mainly in the third income quartile.

7. Conclusions

We have analyzed the determinants of global life satisfaction, by using both self-reports and responses to a battery of vignette questions. Although more work needs to be done, some preliminary conclusions can be drawn.

It appears that the four domains job or daily activities, social contact and family, health, and income provide a fairly complete description of global life satisfaction in both countries. Among the four domains, social contacts and family have the highest impact on global life satisfaction, followed by job and daily activities and health. Income has the lowest impact.

As in other work, we find that American response styles differ from the Dutch in that Americans are more likely to use the extremes of the scale (either very satisfied or very dissatisfied) than the Dutch, who are more inclined to stay in the middle of the scale.

Although for both Americans and the Dutch, income is the least important determinant of global life satisfaction, it is more important in the U.S. than in The Netherlands. Indeed life satisfaction varies substantially more with income in the U.S. than in The Netherlands.

There are some intriguing differences between the way respondents judge vignette persons and what turns out to influence their own satisfaction. Respondents in both The Netherlands and the U.S. appear to think that marriage does not contribute to life satisfaction when they judge vignettes. Yet their own satisfaction is positively influenced by being married. Similarly, respondents believe that other things being equal, older persons should be less satisfied. Yet their own satisfaction goes up with age.

The estimates of an econometric model are used to calculate counterfactual distributions of life satisfaction. Correcting for differences in response scales leads to some shifts though the shifts are not very large. For most age and income groups, the conclusion that the Dutch are more satisfied with their lives than the Americans remains valid. For the oldest age group (65+) and highest income group, however, the vignette corrections lead to different conclusions: giving Dutch respondents the American scales shows that these groups are somewhat less satisfied than their US counterparts. This was not clear from the distributions using own country's scales, mainly because of the Dutch reluctance to evaluate their satisfaction as dissatisfied or very dissatisfied.

Vignettes have been shown to bring objective and subjective measurements of health (in particular, vision) or drinking behavior closer in line with each other. An objective measure for life satisfaction seems hard to give, so that other ways of validation need to be considered, perhaps by looking at actual behaviors that are correlated with life satisfaction. This is one of the directions of future research.

References

Alesina, Alberto. Rafael Di Tella and Robert MacCulloch. 2004. "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics,* 88(9-10), 2009-2042.

Blanchflower, David G. and Oswald, Andrew J. 2004. 'Well-Being Over Time in Britain and the USA." *Journal of Public Economics,* 88(7-8), 1359-1386.

Clark, Andrew E., Paul Frijters, and Michael Shields. 2008. "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature*. 46(1), 95-144.

Deaton, Angus. 2008. "Income, Aging, Health and Well-Being Around the World: Evidence from the Gallup World Poll." *Journal of Economic Perspectives*. 22(2), 53-72.

Di Tella, Rafael, MacCulloch, Robert J and Blanchflower, David G. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics*. 85(4), 809-827.

Di Tella, Rafael, John Haisken–DeNew, and Robert MacCulloch, 2007, "Happiness adaptation to income and to status in an individual panel," processed, October

Easterlin Richard A. (1974), "Does Economic Growth Improve the Human Lot? Some Empirical Evidence", in: R. David and M. Reder (eds.) Nations and Households in Economic Growth: Essays in honor of Moses Abramowitz, New York, Academic Press, 89-125.

Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization*, 27(1), 35-48.

Easterlin, Richard A. 2006. "Life Cycle Happiness and its Sources: Intersections of Psychology, Economics, and Dempography." *Journal of Economic Psychology*, 27, 463, 482.

Kapteyn, Arie., Smith, James. P. & van Soest, Arthur. 2007. "Vignettes and self-reports of work disability in the U.S. and the Netherlands." *American Economic Review,* 97(1), 461-473.

Kapteyn, Arie., Smith, James. P. & van Soest, Arthur. 2008. "Are Americans Really Less Happy With Their Incomes?", RAND Working Paper, WR-591.

King, Gary; Murray, Christopher; Salomon, Joshua and Tandon, Ajay. 2004. "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research." *American Political Science Review,* 98(1), 567-583.

Layard, Richard. 2005. *Happiness: Lessons from a New Science*. Penguin Books: London.

Rao, Jon N. K., and Alastair J. Scott. 1984. "On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data." *Annals of Statistics,* 12: 46-60.

Stevenson, Betsey and Wolfers, Justin. 2008, "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," Working Paper, Wharton School, University of Pennsylvania, prepared for *Brookings Papers on Economic Activity*, Spring 2008.

Van Praag, Bernard M.S., and Ada Ferrer-i-Carbonell (2008), *Happiness Quantified – A Satisfaction Calculus Approach*, Oxford University Press, Oxford.

Van de Stadt, Huib, Arie Kapteyn & Sara van de Geer. 1985. "The Relativity of Utility: Evidence from Panel Data." *The Review of Economics and Statistics,* 67, 179-187.

Van Soest, Arthur., Liam Delaney, Liam, Harmon, Colm, Arie Kapteyn Arie., Smith, James P. 2007. "Validating the Use of Vignettes for Subjective Threshold Scales." RAND Labor and Population working paper WR-501.

Yun, Myeong-Su. 2004. "Decomposing Differences in the First Moment." *Economics Letters*, 82(2), 275-280.

Table 1
Self Reports on Satisfaction with Domains of Life

Self report: How satisfied are you with the total income of your household?	Country	
	NL	US
Very satisfied	9.9	6.5
Satisfied	53.6	39.4
Not satisfied or dissatisfied	23.6	21.5
Not satisfied	10.3	27.4
Very dissatisfied	2.7	5.2

Test for independence: F(3.64, 12207.95) = 20.3117; p-value = 0.0000

Self report: How satisfied are you with your job or other daily activities?		
	NL	US
Very satisfied	19.4	16.3
Satisfied	61.7	52.2
Not satisfied or dissatisfied	14.7	17.5
Not satisfied	3.4	12.1
Very dissatisfied	0.8	2.0

Test for independence: F(3.36, 11231.88) = 11.4447; p-value = 0.0000

Self report: How satisfied are you with your social contacts and family life?		
	NL	US
Very satisfied	23.0	27.1
Satisfied	62.8	48.2
Not satisfied or dissatisfied	11.7	15.7
Not satisfied	1.9	8.5
Very dissatisfied	0.6	0.5

Test for independence: F(3.58, 11978.15) = 13.9798; p-value = 0.0000

Self report: How satisfied are you with your health?		
	NL	US
Very satisfied	15.4	16.1
Satisfied	61.6	46.7
Not satisfied or dissatisfied	14.5	17.3
Not satisfied	7.0	16.5
Very dissatisfied	1.4	3.5

Test for independence: F(3.82, 12791.55) = 13.3638; p-value = 0.0000

Self report: How satisfied are you with your life in general?		
	NL	US
Very satisfied	19.3	20.1
Satisfied	68.2	58.0
Not satisfied or dissatisfied	10.9	15.4
Not satisfied	1.3	5.3
Very dissatisfied	0.3	1.1

Test for independence: F(3.17, 10610.55) = 9.2306; p-value = 0.0000

Note: all frequencies are weighted with sampling weights. Tests are Pearson chi-squared tests for independence converted into F-statistics, accounting for the weighting (see Rao and Scott, 1984).

Table 2
Ordered Probits for Global Life Satisfaction Against Satisfaction with Specific Domains

	Coef.	z	Coef.	z
Income Domain	.225	6.85	.220	6.53
Relations Domain	.721	12.22	.708	11.85
Job Domain	.625	12.27	.626	12.22
Health Domain	.486	11.02	.497	11.10
US Income Domain	.052	0.85	.031	0.49
US Relations Domain	.087	1.11	.101	1.26
US Job Domain	-.020	0.26	-.016	0.20
US Health Domain	-.131	1.98	-.143	2.12
Dummy US Domain	-.145	0.82	-.037	0.68
Married			-.128	1.38
Age 40-50			-.023	0.31
ln family size			-.024	0.35
Age 51-64			-.093	1.29
Age 65+			.023	0.22
Ed med			-.128	1.95
Ed high			-.117	1.77
Working			.116	1.73
ln eq income			.009	0.61
Dummy US			-.255	0.25
US female			.084	0.80
US married			-.088	0.59
US ln family size			.226	1.54
US age 40-50			.108	0.74
US age 51-64			.198	1.32
US age 65+			-.048	0.22
US ed med			-.038	0.25
US ed high			-.059	0.40
US working			-.077	0.60
US ln eq income			-.001	0.02

Table 3
Variation in Global Life Satisfaction Vignettes

	Age	Family	Income	Work	Health
1	42	good	median +	stressful	some pain
2	50	moderate	half median +	ok with long hours	good
3	65	bad	modest	retired	heart problems
4	25	no friends	half median +	no control or security	good
5	25	good	half-median to twice median	no control but secure	good
6	57	bad	median +	retired	pain
7	75	good	half median to twice median	retired	arthritis
8	62	good	half median +	retired	good
9	70	bad	half median +	retired	moderate
10	50	bad	twice median +	good	bad

Table 4

Global Vignettes for the Retired

	Vignette Number									
	3		**6**		**7**		**8**		**9**	
Family	*Bad*		*Bad*		*Good*		*Good*		*Bad*	
Health	*Heart bad*		*Pain*		*Arthritis*		*Good*		*Moderate*	
Age	*65*		*57*		*75*		*62*		*70*	
	NL	US	NL	US	NL	US	NL	US	NL	US
Very satisfied	0.1	1.0	1.6	0.0	13.3	18.1	70.7	77.4	1.4	5.4
Satisfied	25.0	12.7	19.6	10.5	67.4	69.2	26.1	18.8	7.5	4.4
Not satisfied or dissatisfied	48.9	27.1	48.6	40.6	16.3	9.0	3.2	3.0	39.5	19.6
Not satisfied	24.7	51.7	28.6	48.1	3.0	3.8	0.0	0.8	49.0	63.0
Very dissatisfied	1.3	7.6	1.6	0.8	0.0	0.0	0.0	0.0	2.7	7.6

Global Vignettes for the Young **Global Vignettes for the Middle Aged**

	Vignette Number									
	4		**5**		**1**		**2**		**10**	
Family	*no friends*		*good*		*good*		*moderate*		*bad*	
Work	*no control*		*no control*		*stressful*		*ok-long hours*		*good*	
	no security		*good security*							
Health	*good*		*good*		*some pain*		*good*		*bad*	
Age	*25*		*25*		*42*		*50*		*50*	
	NL	US	NL	US	NL	US	NL	US	NL	US
Very satisfied	1.9	3.2	23.9	33.1	16.9	15.4	19.8	23.0	1.5	1.1
Satisfied	14.9	14.9	66.8	58.8	64.3	65.0	56.5	50.0	11.7	8.5
Not satisfied or dissatisfied	39.8	37.2	8.0	7.4	18.8	22.5	19.8	21.6	36.6	36.0
Not satisfied	41.6	41.5	1.3	0.0	2.2	6.0	2.8	5.4	45.3	43.4
Very dissatisfied	1.9	3.2	0.0	0.3	0.5	0.0	1.1	0.0	5.0	11.1

All vignettes evaluated at highest income level in the vignette

Table 5: Effect of vignette descriptions on evaluation

	(1) Regression	(2) Ordered Probit	(3) Probit (very) dissatisfied
Age	0.003	-0.001	0.026
	(1.29)	(0.39)	(5.21)**
Married	0.171	0.263	0.259
	(3.97)**	(4.12)**	(2.17)*
Good relations	-1.211	-1.636	-1.964
	(39.11)**	(32.91)**	(24.24)**
Retired	-0.153	-0.091	-0.765
	(2.11)*	(0.83)	(4.40)**
Modest income	0.947	1.318	1.566
	(20.23)**	(18.77)**	(12.27)**
Half median income	0.218	0.334	0.406
	(8.22)**	(8.21)**	(6.28)**
Twice median income	-0.145	-0.234	-0.153
	(6.17)**	(6.43)**	(2.46)*
Four times median income	-0.279	-0.448	-0.294
	(9.35)**	(9.59)**	(4.14)**
Health good	-0.477	-0.955	-0.244
	(9.43)**	(11.47)**	(1.43)
job secure or under control	0.134	0.290	-0.094
	(3.66)**	(5.02)**	(0.88)
at least 50 hrs	0.693	1.215	0.958
	(12.91)**	(14.38)**	(6.09)**
Dummy US	-0.160	-0.216	-0.164
	(1.00)	(0.89)	(0.45)
US Age	0.007	0.010	0.009
	(2.41)*	(2.17)*	(1.22)
US Married	0.227	0.354	0.510
	(3.53)**	(3.63)**	(2.90)**
US Good relations	-0.218	-0.321	-0.111
	(4.65)**	(4.46)**	(0.95)
US Retired	-0.269	-0.400	-0.342
	(2.45)*	(2.38)*	(1.31)
US Modest income	0.652	0.925	0.885
	(8.75)**	(8.26)**	(4.52)**
US Half median income	0.100	0.153	0.020
	(2.37)*	(2.36)*	(0.20)
US Twice median income	-0.054	-0.093	-0.172
	(1.52)	(1.66)	(1.89)
US Four times median income	-0.058	-0.109	-0.158
	(1.27)	(1.49)	(1.48)
US Health good	0.172	0.200	0.418
	(2.25)*	(1.57)	(1.69)
US job secure or under control	-0.200	-0.284	-0.399
	(3.62)**	(3.25)**	(2.58)**
US at least 50 hrs	-0.078	-0.062	-0.235
	(0.97)	(0.49)	(1.02)
Constant	3.265		-1.168
	(30.70)**		(4.87)**
Observations	12051	12051	12051
R-squared	0.55		

Notes: Robust t statistics in parentheses; * significant at 5% level; **: significant at 1% level
Model (3): dependent variable: 1 if very dissatisfied or dissatisfied; 0 otherwise

Table 6

Self Assessment of Global Satisfaction

	Model with DIF		Model without DIF	
	β	s.e.	β	s.e.
Constant	1.005**	0.14	0.899**	0.12
Female	-0.081	0.06	-0.052	0.05
Married	-0.431**	0.11	-0.361**	0.10
ln family size	-0.205*	0.10	-0.205*	0.09
Age 40-50	0.043	0.09	0.119	0.08
Age 51-64	0.066	0.09	0.054	0.07
Age 65+	-0.141	0.11	-0.169+	0.09
Ed med	-0.018	0.08	-0.054	0.07
Ed high	-0.237**	0.08	-0.163*	0.07
Working	-0.160*	0.07	-0.034	0.06
ln eq income	-0.089*	0.04	-0.098**	0.04
Interactions with dummy US				
Dummy	0.217	0.24	0.270	0.20
Female	0.008	0.10	-0.055	0.83
Married	0.095	0.15	0.015	0.13
ln family size	-0.211	0.15	-0.078	0.13
Age 40-50	-0.274+	0.15	0.138	0.13
Age 51-64	-0.156	0.14	0.033	0.13
Age 65+	-0.346+	0.20	-0.397**	0.16
Ed med	-0.107	0.13	-0.018	0.11
Ed high	-0.026	0.14	-0.027	0.11
Working	-0.008	0.12	-0.127	0.10
ln eq income	-0.415**	0.08	-0.327**	0.06

** indicates significance at 1% level, * indicates significance at the 5% level, and + indicates significance at the 10% level.

Table 7
Thresholds of Estimated Equation for Global Life Satisfaction

	Threshold 1		ln (Threshold 2 – Threshold 1)		ln (Threshold 3 – Threshold 2)		ln (Threshold 4 – Threshold 3)	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.
Constant	0.00	0.00	0.53	0.17	0.81*	0.20	-0.34	0.36
Female	-0.04	0.05	0.02	0.02	0.06	0.04	-0.00	0.04
Married	-0.04	0.07	0.01	0.04	0.11	0.07	0.15*	0.08
ln family size	-0.00	0.07	0.04	0.04	-0.12*	0.06	-0.01	0.07
Age 40-50	-0.07	0.07	-0.03	0.03	0.02	0.05	0.06	0.06
Age 51-64	0.03	0.06	-0.03	0.03	0.07	0.05	-0.10	0.06
Age 65+	0.08	0.08	-0.06	0.05	0.16*	0.07	-0.16+	0.08
Ed med	0.06	0.06	0.02	0.03	0.01	0.04	-0.01	0.06
Ed high	-0.10+	0.06	0.06*	0.03	0.02	0.04	0.00	0.06
Working	-0.21*	0.06	0.09*	0.03	0.03	0.04	-0.05	0.05
ln eq income	0.02	0.04	0.01	0.02	-0.07*	0.02	0.08*	0.04
Interactions with dummy US								
Dummy	0.08	0.18	-0.31	0.34	-0.84+	0.43	-0.13	0.55
Female	0.08	0.07	-0.03	0.04	-0.10+	0.06	0.03	0.06
Married	-0.03	0.10	0.06	0.06	-0.02	0.09	-0.30*	0.10
ln family size	-0.17	0.11	-0.03	0.06	0.09	0.09	0.13	0.10
Age 40-50	0.23*	0.11	-0.06	0.06	-0.00	0.08	-0.05	0.09
Age 51-64	0.19+	0.10	-0.05	0.06	-0.15+	0.09	0.13	0.09
Age 65+	0.04	0.15	-0.02	0.08	-0.09	0.12	0.17	0.12
Ed med	-0.09	0.11	-0.01	0.06	-0.04	0.08	0.04	0.09
Ed high	-0.03	0.11	0.02	0.06	-0.09	0.08	0.08	0.09
Working	0.15	0.09	-0.06	0.05	-0.04	0.07	0.07	0.07
ln eq income	-0.16*	0.05	0.02	0.03	0.07	0.04	0.01	0.05

*Indicates significance at the 5% level; + indicates significance at the 10% level. N = 2244 for NL and 1093 for US.

Table 8

Simulations from Model with DIF: Percent Distribution of Global Satisfaction by Age Group

	Very Satisfied	Satisfied	Not Satisfied/ Dissatisfied	Dissatisfied	Very Dissatisfied
Age group younger than 40					
Dutch sample using own parameters	21.3	66.2	11.2	1.3	0.0
Dutch using US threshold parameters	24.2	59.8	12.9	3.0	0.0
Dutch using US self-assessment parameters	17.5	64.8	14.8	2.7	0.2
Dutch using all US parameters	19.4	59.0	16.1	5.1	0.4
US sample using US parameters	17.3	56.9	18.2	7.1	0.5
Age group 40-50					
Dutch sample using own parameters	19.1	66.0	13.2	1.8	0.0
Dutch using US threshold parameters	28.1	55.8	13.2	2.9	0.0
Dutch using US self-assessment parameters	10.8	63.1	21.0	4.8	0.3
Dutch using all US parameters	16.8	55.2	20.5	7.0	0.5
US sample using US parameters	17.1	55.0	20.3	7.2	0.4
Age group 50-64					
Dutch sample using own parameters	20.0	64.0	14.3	1.6	0.0
Dutch using US threshold parameters	27.5	55.9	13.0	3.6	0.1
Dutch using US self-assessment parameters	14.5	61.9	19.9	3.5	0.2
Dutch using all US parameters	20.1	55.6	17.5	6.5	0.4
US sample using US parameters	19.6	54.0	18.0	7.8	0.5
Age group 65 and older					
Dutch sample using own parameters	26.5	58.9	13.4	1.1	0.0
Dutch using US threshold parameters	27.6	55.3	14.3	2.7	0.0
Dutch using US self-assessment parameters	35.5	53.6	10.1	0.8	0.0
Dutch using all US parameters	36.3	51.3	10.5	1.9	0.1
US sample using US parameters	32.2	52.1	13.0	2.6	0.1
All age groups					
Dutch sample using own parameters	21.5	64.1	12.9	1.5	0.0
Dutch using US threshold parameters	26.7	56.9	13.3	3.1	0.0
Dutch using US self-assessment parameters	18.7	61.4	16.7	3.0	0.2
Dutch using all US parameters	22.4	55.6	16.4	5.3	0.3
US sample using US parameters	19.6	55.1	17.9	6.6	0.4

N=2244 for NL and 1093 for US

Table 9

Simulations from Model with DIF: Percent Distribution of Global Satisfaction by Income Group

	Very Satisfied	Satisfied	Not Satisfied/ Dissatisfied	Dissatisfied	Very Dissatisfied
Lowest Income Quartile					
Dutch sample using own parameters	20.9	65.4	12.4	1.3	0.0
Dutch using US threshold parameters	28.5	56.4	12.4	2.7	0.1
Dutch using US self-assessment parameters	11.3	59.4	22.9	5.8	0.6
Dutch using all US parameters	14.9	53.7	21.4	8.9	1.1
US sample using US parameters	14.2	51.8	22.8	10.3	0.9
Second Income Quartile					
Dutch sample using own parameters	22.3	64.4	12.0	1.3	0.0
Dutch using US threshold parameters	27.6	57.0	12.6	2.7	0.0
Dutch using US self-assessment parameters	17.6	64.0	16.0	2.3	0.0
Dutch using all US parameters	21.7	57.5	16.3	4.5	0.1
US sample using US parameters	22.7	58.4	14.9	3.9	0.1
Third Income Quartile					
Dutch sample using own parameters	22.1	63.0	13.3	1.6	0.0
Dutch using US threshold parameters	26.3	56.8	13.6	3.2	0.0
Dutch using US self-assessment parameters	22.1	61.2	14.5	2.1	0.0
Dutch using all US parameters	25.7	55.8	14.5	4.0	0.1
US sample using US parameters	22.7	56.2	15.9	5.0	0.1
Highest Income Quartile					
Dutch sample using own parameters	20.6	63.5	14.1	1.8	0.0
Dutch using US threshold parameters	24.0	57.5	14.6	3.8	0.0
Dutch using US self-assessment parameters	24.8	60.9	12.5	1.7	0.0
Dutch using all US parameters	28.0	55.7	12.9	3.4	0.0
US sample using US parameters	26.7	57.4	12.6	3.2	0.0
All Income groups					
Dutch sample using own parameters	21.5	64.1	12.9	1.5	0.0
Dutch using US threshold parameters	26.7	56.9	13.3	3.1	0.0
Dutch using US self-assessment parameters	18.7	61.4	16.7	3.0	0.2
Dutch using all US parameters	22.4	55.6	16.4	5.3	0.3
US sample using US parameters	19.6	55.1	17.9	6.6	0.4

N=2244 for NL and 1093 for US; Income is equivalized income (per capita). Quartiles are country specific.

APPENDIX - Life Satisfaction Vignettes

Global Life Satisfaction

Global 1: (*Name*) is 42 years old, happily married, with two children who are doing well at school and generally get on well with their parents. His/her family income is about xxx (median, twice median, four times median). He/she likes his work although some days it is somewhat stressful. (*Name*) suffers from rather serious back pain that keeps him/her awake at night about once a week, but has no other serious health problems.

Global 2: (*Name*) is 50 years and divorced. He/she has one daughter of 22 with whom he gets on well, although he/she sees her only once a year. *(Name)* works about 60 hours per week, and feels he/she has a very secure job over which he has a lot of control. He/she makes about xxx (half the median, median, twice the median, four times the median) per year. He/she has no serious health problems.

Global 3: (*Name*) is 65 years old. His/her wife/husband died 3 years ago and he/she still spends most of his time thinking about his/her and the good times they had together. He/she has four children and 10 grandchildren who visit him/her regularly. (*Name)* has a small pension and receives social security; he/she can make ends meet but has no money for extras such as expensive gifts to his/her grandchildren. He/she has had to stop working recently due to heart problems. He/she gets tired easily and can, for example, not walk more than one block without taking a pause. Otherwise, he/she has no serious health conditions.

Global 4: (*Name*) is 25 years old and single. He/she does not have many friends. He/she works about 50 hours a week and makes xxx (half the median, median, twice the median). He/she feels he has little control over his/her job and worries about loosing it. He/she has no health problems, but feels a little stressed sometimes. He/she does not exercise.

Global 5: (*Name*) is 25 years old and recently married, no children. He/she works about 35 hours per week and makes xxx (half the median, median, twice the median). He/she works out regularly and on vacations he/she makes long hikes in the mountains with his/her husband. He/she job is satisfying, though a bit dull sometimes. He/she feels she does not have a lot of control over his/her job, but it is a very secure job.

Global 6: (*Name*) is 57 years old and recently married his/her second wife/husband. He/she has two children from his/her first marriage, but has little contact with them. He/she draws DI, because he/she has serious back pains. He/she often has trouble sleeping. His/her DI benefits are xxx (half the median, median, twice the median).

Global 7: (*Name*) is 75 years old and a widow. His/her pension benefits are xxx (half the median, median, twice the median). He/she owns the house he/she lives in and has a large circle of friends. He/she plays bridge twice a week and goes on vacation regularly with some friends. Lately he/she has been suffering from arthritis, which makes work in the house and garden painful.

Global 8: (*Name*) is 62 and has been retired for five years. He/she quit his job as soon as he could. He/she has never regretted his/her decision to retire. His/her pension is xxx (median, twice median; four times the median) He/she is physically very active and makes long bicycle trips in Southern Europe. He/she is single, but usually makes the trips with friends his/her age.

Global 9: (*Name*) is 70 and has been retired for five years. His/her pension is xxx (median, twice median; four times the median). He/she still misses the contacts with his colleagues and would have liked to keep working part time. He/she and his wife/her husband take a few small vacations every year. For the rest they each lead their own lives and don't do many things together. They have two children but rarely see them. He/she is overweight and gets tired when walking more than a few blocks. He/she has been a smoker all his life.

Global 10: (*Name*) is 50 and does not exercise. He/she cannot climb stairs or do other physical activities because he/she is obese. He/she has pain in his/her knees, elbows, wrists and fingers, and the pain is present almost all the time. He/she has an executive job in a big firm and feels that he/she has a lot of control over his job. He/she makes xxx (twice the median, four times the median). He/she has been married for a long time, but he/she and his/her wife spend very little time together.

Life Satisfaction

Arie Kapteyn, James P. Smith, RAND & IZA
Arthur van Soest, Netspar, Tilburg University, RAND & IZA[1]

January 2009

JEL code: I31, J28, D31
Keywords: happiness, life satisfaction, Vignettes, Reporting bias

This research was supported by grants from the National Institute on Aging to RAND. We are grateful to Ed Diener, John Helliwell, an anonymous reviewer, workshop participants in Princeton and seminar participants in Tilburg for useful comments.

[1] Tilburg University, P.O. Box 90153, 5000 LE Tilburg, fax: +31-13-4663280; email avas@uvt.nl

Abstract

We analyze the determinants of global life satisfaction in two countries (The Netherlands and the U.S.), by using both self-reports and responses to a battery of vignette questions. We find global life satisfaction of happiness is well-described by four domains: job or daily activities, social contacts and family, health, and income. Among the four domains, social contacts and family have the highest impact on global life satisfaction, followed by job and daily activities and health. Income has the lowest impact.

As in other work, we find that American response styles differ from the Dutch in that Americans are more likely to use the extremes of the scale (either very satisfied or very dissatisfied) than the Dutch, who are more inclined to stay in the middle of the scale.

Although for both Americans and the Dutch, income is the least important determinant of global life satisfaction, it is more important in the U.S. than in The Netherlands. Indeed life satisfaction varies substantially more with income in the U.S. than in The Netherlands.

1. Introduction

Economists have discovered happiness (or rediscovered) or at least research on subjective well-being and its economic correlates (see, e.g., Van Praag, Frijters and Ferrer-i-Carbonell, 2003, Layard, 2005, or Clark, Frijters and Shields, 2008). The rapidly growing research has touched on several important themes. These have included the so-called Easterlin paradox whereby average happiness remains relatively constant over time in spite of large increases in income per capita (Easterlin, 1974, 1995; see also the chapter by Graham, Chattopadhyay and Picon in this volume). In contrast, within country cross-sectional and panel data almost always show that rising incomes 'buy' additional satisfaction, although the magnitude of the within country cross-sectional effect of income on satisfaction is under dispute (Blanchflower and Oswald, 2004, Di Tella et al, 2007 and Stevenson and Wolfers, 2008). Resolving this paradox, which is often interpreted as a fundamental challenge to the conventional economic theory of utility maximization, has generated a substantial amount of subsequent research attempting to reconcile the finding of a zero correlation between income and life satisfaction in aggregate time series evidence with the positive correlation in cross-section micro-estimates within a given country.

This reconciliation has included adding relative incomes (of others or of oneself in the past) in the utility function (Van de Stadt et al. 1985, Clark et al, 2008) or a sometimes rapid process of adaptation to new circumstances (Di Tella et al, 2003) often labeled the 'hedonic treadmill'(Di Tella et al, 2007). A recent contrary view is provided by Deaton (2008) who documents that if one considers a much wider range of countries arrayed by their level of economic development, the normally positive association of income with subjective life satisfaction reappears. His work also leads to the conclusion that the effect of income on life

satisfaction according to cross-country regressions is if anything higher in the high income countries than in the low income countries. Stevenson and Wolfers (2008) revisit much of the earlier evidence and look at new data to reach similar conclusions.

A considerable amount of research has focused on cross-country differences in subjective well-being, in particular comparing Europe and the U.S. where the US appears to rank lower in satisfaction than many European countries with lower per capita incomes (Alesina et al, 2004, Di Tella et al, 2003, and Blanchflower and Oswald, 2004). For instance, Europeans apparently exhibit a stronger distaste for inequality than do Americans that may be partly explained by a perception of greater mobility in the US (Alesina et al, 2004). Blanchflower and Oswald (2004) study trends in well-being over time in the UK and the US and find that reported levels of well-being have been dropping over time in the US while they have been flat in the UK, despite the fact that in both countries average incomes have grown substantially over the last couple of decades.

A fundamental problem in international comparisons, cross-sectional, and time series analyses of subjective well-being is that one has to assume that somehow response scales are the same across countries, across time, and across groups of respondents within a country. This critical and largely untested assumption becomes even more tenuous if question phrasings change or differ across surveys, as is often the case (see Stevenson and Wolfers, 2008). Here we address these problems head on. In view of the specific interest of economists in the relation between life satisfaction and income, we focus on the role of income.

The population distribution of satisfaction in a country will depend on levels and distribution of incomes. Residents of alternative countries can however differ in the way they translate any given level of income into a subjective level of satisfaction. Moreover, residents of

countries may differ in the subjective thresholds that they use in demarcating satisfaction into discrete categories such as very satisfied or not satisfied. Income distributions, the translation from income to income satisfaction and the demarcation thresholds can all affect differences observed within and between countries in their distribution of stated level of satisfaction. These distinct factors are often confused in the existing literature on life satisfaction and happiness. In our research, we have created unique data sources in two countries—the United States and the Netherlands—and developed a statistical methodology that allows us to separate out these distinct factors.

This paper is divided into seven parts. Section 1 describes the data sources that we developed and will rely on in this analysis. The second section summarizes responses of Dutch and American respondents to questions about their own life satisfaction in several key domains of their lives while the third section describes the types of vignettes we developed and the responses to those vignettes by our Dutch and American respondents. In the next section, we summarize the vignette methodology that serves as the basis of our analysis and then sketch our statistical model that corrects for response scale differences across countries. Section 5 presents our main empirical results and their implications for interpreting observed differences in life satisfaction in the two countries. In section 6, simulations based on our estimated models are used to ascertain what Dutch distributions of life satisfaction would be if the Dutch had American parameters and thresholds rather than their own. The final section highlights our main conclusions.

1.1 Data Sources

Our analysis in this paper is based on information obtained from two Internet surveys, which we conducted in the Netherlands and the United States. For The Netherlands, we used

CentERpanel, administered by CentERdata affiliated with Tilburg University. CentERpanel includes about 2,250 households who have agreed to respond to questions every weekend over the Internet. The Dutch sample is not restricted to households with their own Internet access. Respondents are recruited by telephone. If they agree to participate and do not have Internet access, they are provided with Internet access (and if necessary, a set-top box). Thus, CentERpanel is representative of the adult Dutch population except the institutionalized. Sampling weights provided by CentERdata and based upon comparing with a much larger survey of Statistics Netherlands are used to correct for unit non-response and attrition. The sample used for estimation has 2,244 respondents who participated in an interview with questions on life satisfaction (self-assessments as well as vignettes) in 2006. From multiple waves collected in the past, CentERpanel has a rich set of variables on demographic, health, and economic characteristics of respondents. In our analysis we use the most recent measurement of these background variables, reported a few months before our life satisfaction survey. The Internet infrastructure makes CentERpanel an extremely valuable tool to conduct experiments, with possibilities for randomization of content. Production lags are very short, with about one month between module design and data delivery.

Our Internet survey for the United States is the RAND American Life Panel (ALP). This panel was initially recruited from respondents age 40 plus in the Monthly Survey (MS) of Michigan's Survey Research Center but has been subsequently supplemented with younger respondents.[2] Similar background information was collected for these respondents as was

[2]The MS, the leading consumer sentiments survey, produces the widely used Index of Consumer Attitudes. MS respondents are asked if they have Internet access and, if yes, if they are willing to participate in Internet surveys. Those who agree are added to our household panel to be interviewed regularly over the Internet. As with the CentERpanel, respondents who do not have Internet access are provided with a set top box (an MSN Web TV) that allows them to browse the Internet and send and receive email.

available for Dutch respondents. The American sample used for estimation consists of 1,113 respondents interviewed during 2006-2007. The American data are weighted to match Current Population Survey demographic distributions in age, education, and gender.

2 Global Life Satisfaction by domain

Respondents are asked to rate themselves on a five point scale. They do so in the following specific life domains: income, family and social relations, job, and health. They are also asked a global question on their own life satisfaction. The scale that is used is the same for all domains: (*very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very dissatisfied*). The exact self-assessment questions of life satisfaction are contained in Table 1.

Table 1 summarizes responses obtained from the Dutch and American samples for the four domains of income, social contacts and family life, job or other daily activities, and health. The last panel in Table 1 presents responses to the question regarding global life satisfaction. For all domains and for global satisfaction, the distributions in the US and the Netherlands are significantly different (see the tests reported at the bottom of each panel).

Before turning to between country differences, it is useful to first highlight some differences across the domains. Both the Dutch and Americans appear to be less satisfied with their incomes than with the other domains. The health domain is next, at least in terms of most negative responses, followed by job and daily activities, with respondents in both countries most satisfied with their lives in the family and social contacts domain. Differences among the Dutch in the three domains besides income are relatively small with sharper distinctions present in these three domains among Americans. Finally, a much smaller proportion of both Dutch and American respondents appear to be dissatisfied with their lives when answering a global life satisfaction question than their answers in each of the four specific life domains would indicate.

This appears to be due to relatively modest correlations in dissatisfaction across domains, so that dissatisfaction in one domain may be compensated by satisfaction in a different domain. In fact, across the two samples, less than one percent of all respondents are not satisfied or very dissatisfied in all four domains and among those virtually all (94%) report to also be not satisfied or very dissatisfied with their life in general.

Turning to between country differences in life satisfaction, consider first how satisfied respondents in the two countries are with the total household income. As we have analyzed in more detail elsewhere[3], Americans are much less satisfied with their incomes than the Dutch are in spite of the fact that on average their incomes are considerably higher. Sixty-four percent of Dutch respondents say that they are either satisfied or very satisfied with their total household income. The comparable fraction for Americans is 46%- eighteen percentage points lower than the Dutch. Similarly, a much larger fraction of Americans respond that they are either not satisfied or very dissatisfied- a third of Americans compared to thirteen percent amongst the Dutch.

The Dutch are also more satisfied with their jobs than Americans are but these differences are smaller than those in the income domain. Fourteen percent of Americans are either not satisfied or very dissatisfied with their jobs compared to four percent of the Dutch- a differential of ten percentage points which is half as large as the differential in the income satisfaction domain. At the top of the scale, more than four in every five Dutch respondents are at least satisfied with their jobs as are more than two-thirds of Americans. In both countries, respondents are much more satisfied with their job and other daily activities than they are with their incomes.

[3] We analyze answers to questions on income satisfaction in depth in Kapteyn, Smith, and Van Soest (2008).

There are actually more Americans very satisfied as well as very dissatisfied with the social aspects of their lives compared to the Dutch. Relatively few respondents in either country register displeasure (not satisfied or very dissatisfied) with this domain, although once again Americans are more likely to go to the bottom of the scale compared to the Dutch (9% compared to 2%). This avoidance of extremes is a common feature of Dutch responses to subjective scale questions and is similar to what we have found in prior work (see for example Kapteyn, Smith and VanSoest, 2007). This tendency may well have its origins in the Dutch culture. According to Wikipedia, "The Dutch typically see their countrymen as sober, practical and down-to-earth people. Any form of ostentation is likely to be criticized, and straightforwardness is generally appreciated."

The final specific domain on which we asked about life satisfaction was health. Based on objective health measures, the Dutch are a healthier population than the Americans (Kapteyn, Smith, and Van Soest, 2007). In this case, that objective difference is reflected in their subjective answers to their satisfaction with their health. Nineteen percent of Americans are either not satisfied or very dissatisfied with their health compared to 8% of the Dutch.

The final panel in Table 1 displays the distribution of answers to questions evaluating own global life satisfaction (satisfaction with life – SWL). Eighty-eight percent of the Dutch are either satisfied or very satisfied compared to seventy-eight percent of the Americans. Similarly, while most respondents in both countries appear to relatively satisfied with their lives, 6.4% of the Americans say that they are at a minimum not satisfied compared to only 1.4% of the Dutch.

Using models similar to those developed by Ferrer-i-Carbonell and van Praag (2002), Van Praag et al. (2003), and Van Praag and Ferrer-i-Carbonel (2008), Table 2 examines the relationship between responses to the global life satisfaction question to the level of satisfaction

within the four specific life domains. It does so by listing coefficient estimates (and the associated 'z' values) from an ordered probit model of global life satisfaction. The initial set of regressors in the first two columns are responses to life satisfaction in the four specific life domains each indexed on a scale of one to five. Main effects are estimated Dutch coefficients while the US interactions test for differences between Americans and Dutch.[4]

As expected, these results show that satisfaction with life is positively associated with satisfaction within each of the four domains. As indicated by the estimated coefficients within each domain, income satisfaction received by far the lowest weight in global satisfaction with the health domain in second to last place. The highest weight is in the family and social relations domain.[5] While there is not much evidence of statistically significant differences between the two countries in the translation from satisfaction within a domain into global life satisfaction, there appears to be less weight in the US assigned to the health domain. Remember that the coding goes from very satisfied to very dissatisfied, so the negative sign on the US dummy means that US respondents are happier, keeping satisfaction in each domain constant. That result however is not statistically significant.

The second model in Table 2 adds a number of standard demographics to this model including age, marital status, education, gender, and income and once again allows all estimated effects to differ between the Dutch and Americans. All in all, the evidence for the need for demographics or interactions of these with the US dummy is very weak. A test of the null that

[4] To keep the specification parsimonious and following van Praag and Ferrer-i-Carbonell (2008), we include domain satisfactions as cardinal variables. Using dummies gives qualitatively similar results.

[5] Van Praag and Ferrer-i-Carbonell (2008, Chapter 4) perform similar regressions using panel data on Germany and the UK. They find a much larger role for satisfaction with the financial situation than we do for satisfaction with income. For the UK, they also find that social contacts are the most important factor; they do not have satisfaction with social contacts in the German data.

the effects of the demographics are equal to zero does not lead to rejection. Thus, it seems a model with just the domain specific satisfaction variables is sufficient.

There is a slightly different way of interpreting this outcome. If we state as a null hypothesis that global life satisfaction is a function of just the four domains we consider here, then the test would not reject that null. Of course the power of that test will depend on how much the possibly omitted satisfaction dimensions are correlated with the demographics included in Table 2.

3.1 Description of Vignettes

In addition to their ratings of their own life satisfaction, respondents were given a set of vignettes cover the four life domains- income, family relations, job, and health. These domains were chosen because the current literature has documented them as key determinants of overall life satisfaction (see Easterlin, 2006). All vignettes were given with either a female or male name, which was randomized across respondents. Within each domain, vignettes were presented in random order to eliminate any possibility of order effects whereby the initial vignette presented could affect the ranking of subsequent vignettes.[6] Comparing rank ordering of vignette evaluations across respondents shows that different respondents tend to order vignettes in the same way. The scale that is used is the same for all domains: (*very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very satisfied*).

The vignette questions in the income domain specify an income for the vignette person that is selected randomly with the values being equal to either the median income in the Netherlands or the United States or a value that is half, twice, or four times the median income in

[6] In earlier work on health and work disability vignettes, we found some effects of the order of vignettes within each domain on the vignette evaluations. It might also be useful to randomize the order in which self-assessments by domain are presented but we did not do this.

each country. The six family relation vignettes vary conditions in the vignette person's family or friends life, including whether the vignette person is married or single, has children, has many or few friends and the nature of the relationship with these important others. The work vignettes focus on whether the vignette person is working or retired, the amount of hours worked the security of the job, and how stressful the job is. Finally, the four health domain vignettes vary conditions around the vignette person's ability to engage in (light, vigorous) exercise, and possible problems with sleep, anxiety, and depression.

In addition to these domain specific life satisfaction questions, respondents are also given a subset of ten possible vignettes on global life satisfaction. These global life satisfaction vignettes succinctly describe the vignette person in a single vignette across the four sub-domains mentioned above- family relations, work, income, and health, combining the descriptions given in the domain specific vignettes. This global vignette approach has the advantage of moving directly to an overall measure of life satisfaction and for that reason we will use them in the analysis in this paper. The domain specific vignettes are not analyzed in the current paper; the satisfaction with income vignettes are analyzed in Kapteyn et al. (2008).

The specific scenarios described in the global satisfaction vignettes are listed in the Appendix. Table 3 presents a summary of the type of variation present in the ten global vignettes. The six ways that the global vignettes can vary are gender, age, family, income, work, and health. Gender of the vignette person is randomly assigned to respondents through the feminine or masculine name of the vignette person. An exact age is always mentioned in each vignette and these ages are listed in the first column of Table 3.

With one exception (vignette 3 where income is always 'modest'), income is also randomly assigned in the vignette with up to four values assigned- (ranging from half the

median, median, twice the median, and four times the median). In most scenarios, only three of the four possible income values are used.

The overall situation described in the family and social relations dimension relates to spouse, children, and/or the presence of close friends. In four of the vignettes, the social situation can reasonably be described as good, in one vignette as moderate, and in the other five it is problematic in some important way. The aspects of the work environment that are mentioned in the vignettes are stress, hours, control, security, and retirement status. In five vignettes, the vignette person has already retired. Finally, four of the vignettes describe a person in good health while the remainder point to some type of health problem ranging from moderate to serious.

3.2 Responses to Vignette questions on global life satisfaction

Table 4 lists the distribution of responses obtained in both countries for each global life satisfaction vignette. We divide these vignettes into three groups based on the age of the vignette person- those who have retired, those of the young, and those for middle age vignette persons. To eliminate the impact of income variation within the vignette, our comparisons in Table 4 apply to the highest income value mentioned in the vignette. The numbers in the rows at the top of the columns match the numbers of vignette in Appendix A. The rows below describe the main attributes of the vignette person in terms of their age, family relations, health, and job situation.

Let's first examine the panel for the five global vignettes for retired persons. In addition to a limited age variation within this group, the only meaningful variation in these vignette descriptions concerns the quality of family life and the health of the respondent. It is useful to start with the best case of the possible scenarios- vignette 8 where both the social life and health of the vignette respondent is good. In this case, vignette evaluations of both Dutch and American

respondents mirror this best case description as three quarters of respondents say that one should be very satisfied and less than one percent respond as either not or very dissatisfied. When the life situation appears to be very good as it certainly is in this case, a larger fraction of Americans answer very satisfied- 77% compared to 71% for the Dutch.

Health status has a major impact in these evaluations of retired vignette persons. Even with vignette number 7 where the vignette person suffers only from arthritis but has many friends, there is a very sharp reduction in the percent of respondents who would be satisfied. Less than one in five of Dutch and American respondents apply the label 'very satisfied' compared to at least 70% for the healthy vignette 8. As was the case with vignette 7, somewhat more Americans say 'very satisfied' with vignette 8 compared to the Dutch.

When the health situation of the vignette person gets even worse as in vignettes 3, 6, and 9, large fractions of Dutch and American respondents choose 'not satisfied' or 'very dissatisfied'. But this negative reaction to poor health appears to be more dramatic in the American sample. Using vignette 3 to illustrate the point, 59% of Americans respond as not dissatisfied or very dissatisfied compared to only 26% of the Dutch. This result seems in contrast to Table 2 where we found that life satisfaction depends to a lesser extent on health satisfaction in the US than in the Netherlands. Perhaps the result is misleading since we do not have an orthogonal design – variation in health is correlated with other variations in the vignette characteristics – and we need regressions to control for other differences in the vignette characteristics (see Table 5 below). Vignettes among the retired suggest the importance of health for life satisfaction among the retired population.

Turn next to the two vignettes pertaining to the young. Age and even health are basically the same in these two vignettes so that they differ only how they describe the family and work

domains. Compared to vignette 5 where the social situation is good and the job situation is secure, the person in vignette 4 has no friends and does not feel secure about his/her job. Those two problems take their toll on the evaluation of global life satisfaction. There is a tenfold reduction in the percent of respondents who are very satisfied and the percent of respondents who are not satisfied or very dissatisfied increases from about one percent to over forty percent. Even among the young, where income and income growth is quite central to their lives, a negative situation in terms of either job security or friendships reduces overall life satisfaction a great deal, at least that is what our respondents believe.

The final subset of global life satisfaction vignettes describes the middle aged. For them also, the age spread in the vignettes is quite limited so that the principal variation across alternative vignettes relates to job, social relations, and health. In some respects, vignette 10 may be the most interesting. In this case, the work situation is good and all vignettes are either assigned an income twice the median or four times the median. The answers summarized in Table 4 are for the four times median income in each country so that the economic situation of the household is very good in the economic/work domains. However, both the family life and health are not good and as a consequence many respondents evaluate the life situation described quite negatively. More than half of both Dutch and American respondents are not satisfied or are very dissatisfied with this vignette person's life.

Evaluating the independent impact of social relations, health, and job is difficult in the global vignettes as all three dimensions vary simultaneously across these vignettes. Table 5 presents an alternative approach for doing so, by showing the results of a regression of the rating of the vignettes on their main characteristics. The characterization chosen here is a little different from the one shown in Tables 3 and 4. Since there are only ten vignettes we have to choose a

parsimonious set of indicators to characterize the vignettes, in order to be able to identify their effect. The differences in comparison with Table 3 are as follows. We characterize family and social relations by means of two variables: whether the vignette person is married (vignettes 1, 5, 6, 9, 10) or not and whether the relation with family members or friends is good (vignettes 1, 2, 3, 5, 7, 8) or not. Work is characterized by whether one works long hours (50 per week or more, vignettes 2 and 4) or not and whether the vignette person has control over his or her job or job security (vignettes 2, 5, and 10) or not. Health, finally, is coded as good (vignettes 2, 4, 5, 8) or not.

Table 5 presents the results of an ordered probit, a regression, and a probit where we have combined dissatisfied and very dissatisfied into one category and the other three answers into another category. The latter has been included to investigate whether patterns are different at lower end of the scale. As one can see, qualitatively the results of the regression and the ordered probit are very similar. We will only discuss the regression, as one can immediately interpret the size of the effects, whereas that is more complicated in the ordered probit case. At the end we will comment on cases where the probit deviates from the regression and the ordered probit.

As before, the top panel presents the results for the Dutch sample, while the bottom panel presents the effects of interactions with a US dummy. The US dummy is negative, but insignificant. Thus the evidence for a uniform scale shift is limited. The difference in response patterns are more subtle than just a uniform shift in the scale used in both countries.

One observes that in the U.S. there is some evidence that a higher age is assumed to be associated with lower life satisfaction (remember once again that the scale runs from 1 for very satisfied to 5 for very dissatisfied). For instance, in the U.S. an age difference of 40 years would be associated with a deterioration in life satisfaction of .40 on the five-point scale. At first sight,

being married is not considered to be a good thing, particularly in the US. Notice however that this has to be looked at in combination with the quality of the relation with family and friends. Since good relations have an enormous effect of life satisfaction, being in a good marriage is a major source of happiness. Both countries attach a positive value to retirement, but the Americans more so than the Dutch.

Working long hours is believed to have a negative effect on life satisfaction, but more so in the U.S. than in The Netherlands, suggesting that working conditions are more attractive in the Netherlands than in the US. Having control over one's job or having a secure job is valued more positively in the U.S. than in The Netherlands. Good health is worth about half a point in The Netherlands and about three tenth of a point in the US, but the difference between the two countries is only marginally significant. The larger effect in the Netherlands is in line with Table 2 and shows that the differences in Table 4 are indeed somewhat misleading – they disappear when other vignette characteristics are controlled for.

Income has a positive effect on life satisfaction in both countries, although the effect is not all that large. This is in line with the results presented in Table 2, which showed that of the four domains, income is the least important one for global life satisfaction. The one exception is the dummy for modest income, which is quite large, but since this dummy uniquely identifies vignette 3, its coefficient really identifies the evaluation of that vignette as a whole, not just the income level. We note that income is thought to be more important in the US than in The Netherlands.

The binary probit which explains the probability of rating a vignette as dissatisfied or very dissatisfied shows some minor deviations from the results in the other two columns. Both in The Netherlands and the U.S., increasing age is associated with a greater probability of being

rated dissatisfied or very dissatisfied. We also note that having a secure job or a job over which one exerts control reduces the chances of being rated dissatisfied or very dissatisfied.

4.1. The Theory of Vignettes

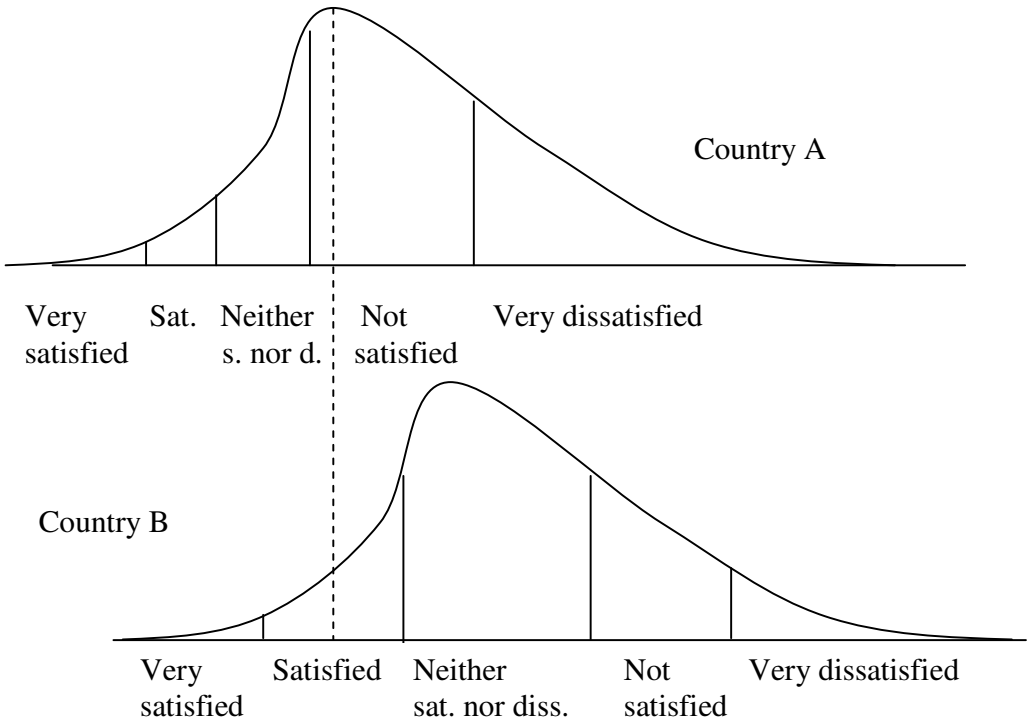
In this section, we provide an intuitive description of the use of vignettes for identifying response scale differences and then sketch our statistical approach.[7] The basic idea is illustrated in Figure 1, which presents the distribution of life satisfaction or happiness in two hypothetical countries. The density of the continuous happiness variable in country A is to the left of that in country B, implying that on average, people in country A are happier than in country B. The people in the two countries, however, use very different response scales if asked to report their happiness on a five-point scale (very satisfied, satisfied, not satisfied or dissatisfied, not satisfied, and very dissatisfied).

In the example in the figure, people in country B attach much more positive labels to given points on the life satisfaction scale than do people in country A. Someone in country A with the life satisfaction indicated by the dashed line would report to be not satisfied, while a person in country B with the same actual satisfaction would report to be satisfied. The frequency distribution of the self-reports in the two countries would suggest that people in country B are more satisfied than those in country A—the opposite of the true distribution. Correcting for the differences in the response scales (DIF, "differential item functioning," in the terminology of King et al., 2004) is essential to compare the actual health distributions in the two countries.

[7] Vignettes have been used earlier in economic research by Van Beek, Koopmans and van Praag (1997), who analyze employer preferences by presenting hypothetical descriptions of job applicants.

Vignettes can be used to do the correction. A vignette question describes the satisfaction of a hypothetical person and then asks the respondent to evaluate the satisfaction of that person on the same five-point scale that was used for the self-report of their satisfaction. Since the vignette descriptions are the same in the two countries, the vignette persons in the two countries have the same actual life satisfaction or happiness. For example, respondents can be asked to evaluate the life satisfaction of a person whose satisfaction is given by the dashed line. In country B, this will be evaluated as "satisfied." In country A, the evaluation would be "not satisfied." Since the actual level of satisfaction is the same in the two countries, the difference in the country evaluations must be due to DIF.

Figure 1. Comparing self-reported happiness across two countries in case of DIF



Vignette evaluations thus help to identify differences between the response scales. Using the scales in one of the two countries as the benchmark, the distribution of evaluations in the other country can be adjusted by evaluating them on the benchmark scale. The corrected distribution of the evaluations can then be compared to that in the benchmark country—they are now on the same scale. In the example in the figure, this will lead to the correct conclusion that people in country A are more satisfied than those in country B, on average. The underlying assumption is *response consistency:* a given respondent uses the same scale for self-reports and the vignette evaluations.

We will apply the vignette approach to life satisfaction, using vignettes not only to obtain international comparisons corrected for DIF, but also for comparisons of different groups within a given country.

4.2 Econometric Model

We will apply the vignette approach to life satisfaction, using vignettes not only to obtain international comparisons corrected for DIF, but also for comparisons of different groups within a given country. Our model explains respondents' self-reports on satisfaction by themselves as well as their reports on satisfaction of hypothetical vignette persons. Self-reports are modeled as a function of respondent characteristics X_i (including demographics, a country dummy and interactions of all demographics with that dummy) and an error term ε_i by the following ordered response equation:

(1.1) $$Y_i^* = X_i\beta + \varepsilon_i;\ \ \varepsilon_i \sim N(0,\sigma^2),\ \varepsilon_i \text{ independent of } X_i$$

(1.2) $$Y_i = j \text{ if } \tau_i^{j-1} < Y_i^* \le \tau_i^j,\ \ j = 1,...5$$

The thresholds τ_j^i between the categories are given by

$$\tau_i^0 = -\infty,\ \tau_i^5 = \infty,\ \tau_i^1 = \gamma^1 X_i + u_i,\ \tau_i^j = \tau_i^{j-1} + \exp(\gamma^j X_i),\ j = 2,3,4$$

(1.3)

$$u_i \sim N(0, \sigma_u^2),\ u_i \text{ independent of } X_i \text{ and the other error terms in the model}$$

Since X_i includes a country dummy and interactions of all demographics with that dummy, this specification allows for completely different ways in which the response scales vary with demographics in the two countries. The various cut-off points can also vary in different ways, which seems useful because of the observed tendency of the Dutch to avoid extremes, suggesting that a Dutch respondent will have a lower first cut-off point but a higher last cut-off point than a similar US respondent.

As noted before, the fact that different respondents *i* use different response scales τ_i^j is called "differential item functioning" (DIF). The term u_i introduces an unobserved individual effect in the response scale. It will imply that reported evaluations of different vignettes (see (1.5) below) are positively correlated with each other and with self-reports (conditional on X_i), since some respondents will tend to use high thresholds and others will use low thresholds in all their reports. Since such positive correlations are observed in the data, incorporating u_i helps to improve the model's ability to predict the observed outcomes (the model fit).

Define a benchmark respondent with characteristics $X_i = X(B)$. The DIF adjustment involves comparing Y_i^* to thresholds τ_B^j rather than τ_i^j, where τ_B^j is obtained in the same way as τ_i^j but using $X(B)$ instead of X_i. A respondent's reported satisfaction is computed using a benchmark scale instead of a respondent's own scale. This does not give an adjusted score for each individual (since Y_i^* is not observed) but it can be used to simulate adjusted *distributions* of Y_i for the whole population or conditional upon some of the characteristics in X_i.

Using self-reports on own life satisfaction only, parameters β and γ^1 are not separately identified, only the difference between β and γ^1. For example, consider country dummies: people in two different countries can have systematically different life satisfactions, but if the scales on which they report their life satisfaction can also differ across countries, then self-reports are not enough to identify the satisfaction difference between the countries. The vignettes will be used to identify β and γ^1 separately.

The evaluations Y_{li} of vignettes $l=1,\ldots,L=10$ are modeled using similar ordered response equations:

(1.4) $$Y_{li}^* = \theta_l + \kappa I_{li} + \varepsilon_{li}$$

(1.5) $$Y_{li} = j \text{ if } \tau_i^{j-1} < Y_{li}^* \leq \tau_i^j,\ j = 1,\ldots 5$$

(1.6) $$\varepsilon_{li} \sim N(0, \sigma^2), \text{ independent of each other, of } \varepsilon_{ri} \text{ and of } X_i$$

Thus we include a dummy for each of the 10 vignettes and allow the evaluations to depend on the log of the income assigned to the vignette (I_{li}), which is randomized across respondents. The unobserved vignette evaluations Y_{li}^* do not depend on respondent characteristics X_i (the assumption of *vignette equivalence).* The actually reported evaluations Y_{li} do depend on X_i, but only through the thresholds. The maintained assumption here is that of *"response consistency"*, meaning that the thresholds τ_i^j are the same for self-reports and the vignettes.

With these assumptions, it is clear how vignette evaluations can separately identify β and γ $(=\gamma^1,\ldots,\gamma^5)$: From the vignette evaluations alone, γ, θ, $\theta_1,\ldots\theta_{10}$ and κ can be identified (up to the usual normalization of scale and location). From self-reports, β can then be identified in addition. Thus the vignettes can be used to solve the identification problem due to DIF.

The two assumptions *vignette equivalence* and *response consistency* are crucial for solving the identification problem. Vignette equivalence may be problematic if life satisfaction is multidimensional and the weights are different in the two countries. The fact that in Table 2, the interactions between domain satisfactions and the US dummy are jointly insignificant suggests that this is not a serious problem in our case. *Response consistency* may be violated if, for example, people make systematic mistakes in evaluating vignette persons but are much better able to evaluate their own satisfaction. *Response consistency* can be tested if an objective measure is available and such tests have typically supported the use of vignettes (see King et al., 2004, on vision, and Van Soest et al., 2007, on drinking behavior), but an objective measure of satisfaction with life seems hard to give.

5.1 Empirical Results

This section highlights our main empirical findings. We discuss our main parameter estimates determining overall satisfaction with life and assess the consequences of different threshold parameters in both countries.

The model presented above was estimated using the self-evaluations and vignettes in the Dutch CentERpanel and the RAND American Life Panel. The equations for global life satisfaction and for the response thresholds include a complete set of interactions with a country dummy for the United States. We also estimated the simpler model that does not allow for DIF. This amounts to a standard ordered probit for self-assessed satisfaction.

5.2 Model of Global Life Satisfaction

Table 6 lists parameter estimates for two models explaining global life satisfaction, where the scale is from good to bad (1: very satisfied, …, 5: very dissatisfied). All regressors in these

models (except the country dummy) are measured in deviations of their country specific means, which makes it easier to interpret the constant term and most importantly the implications of the US dummy. Demographic regressors include dummy variables for whether the respondent is female, married, age brackets 40-50, 51-64, 65+ (the left out group is under 40 years old). Education is separated into three groups- low, medium or high with the low education group the left out category. Income is measured as log-equivalized family income where income is adjusted by the logarithm of family size. Log-family size is also a separate regressor, in part to test for the adequacy of this choice of functional form for the equivalence scale. Finally, a dummy variable is included indicating whether the respondent is working.

For reasons outlined above, our preferred model is the model with DIF (adjusting for threshold differences). It is listed in the first two columns of Table 6. In the Dutch sample, there are no significant differences in satisfaction with life by gender or age. Higher income makes the Dutch more satisfied with their life. Conditional on income, higher education also makes the Dutch more satisfied. Since education is typically associated with higher income, this most likely reflects the fact that education is a reasonable proxy for permanent income of respondents. Finally, conditional on the equivalized income, married Dutch respondents and those with larger families are more satisfied with their lives. One interpretation of this finding is that marriage and family are on average a source of well-being for these households. Dutch respondents who work are more satisfied with their lives than those who do not.

Turn next to our estimates of the differences in parameters between the two countries which implicitly set the US parameters. Since regressors are measured in deviations from within-country means, the coefficient on the US dummy gives the difference between the average US person and the average Dutch person, whose characteristics are different (see also

Section 6 for the consequences of these differences in the demographics). This coefficient is positive but insignificant, suggesting that the average Dutch and US respondents have similar satisfaction with life, according to both the model with and the model without DIF. Similar to the Dutch, there are no gender differences in life satisfaction among the Americans but the estimated age patterns indicate that life satisfaction among Americans increases with age and that retired Americans are particularly satisfied with their lives. There is no differential impact of work among Americans. Here the contrast with some of the results in Table 5 is of interest. When evaluating vignettes, Americans seemed to think that getting older would reduce life satisfaction. Yet when it comes to their own satisfaction, getting older is a good thing. None of these effects are strongly significant, but they still would seem to cast some doubt on the assumption that Americans are able to evaluate the vignette persons in the same way as they evaluate themselves (response consistency) or that respondents of different ages evaluate the same vignette differently (vignette equivalence).

To explore this further we have included interactions between the respondents' own age (coded as dummies as in Table 6) and the age of the vignette person in the regressions reported in Table 5. These interactions turn out to be totally insignificant, as they should be under the assumption of vignette equivalence. This implies that there is no evidence that respondents make systematic errors in evaluating vignettes describing persons of different ages than their own age.

The most important variable for comparing the two countries is income. The impact of income in improving life satisfaction is much more pronounced in the US than in The Netherlands (more than four times larger in the US in the model with DIF).[8] Since we estimate

[8] The coefficients of log income in the US (0.504 in the model with DIF, 0.425 in the model without DIF) seem rather large compared to the coefficients reported in the chapter by Layard, Mayraz and Nickell in this volume (between 0.33 and 0.58 on a ten-point SWB scale, where we use a five-point scale). The coefficients for the Netherlands are much smaller than what Layard et al. find for European countries.

no intercept difference between the countries and the data are all demeaned within countries, the Dutch and Americans are about equally satisfied at their country specific mean incomes. But Americans become more satisfied with life at high incomes levels and much less satisfied than the Dutch at low incomes levels.

Another important question is how the corrections for threshold differences within and across countries affect our interpretation of the determinants of life satisfaction. This question is addressed by comparing the parameter estimates in the model without DIF to the model with DIF. Several estimated effects seem rather similar between the two models. We note however that for the Dutch the estimated effects of education and working are larger in the model with DIF than in the model without. Considering the interactions with the US dummy, the effect of income on life satisfaction in the US turns out to be more pronounced when we correct for DIF.

These differences between the models with and without DIF are of course directly related to the estimated equations for the thresholds in the model with DIF. For instance, consider the effect of log income interacted with the US dummy. The negative coefficient for this variable in the first threshold equation means that the first threshold shifts to the left when log income increases in the US. As a result of that, a response is less likely to lie to the left of that threshold. Since this effect of log income on the first threshold explains part of the existing negative correlation between income and life satisfaction, incorporating the effect on the threshold reduces the negative effect of log income in the US on self rated global satisfaction. This explains the difference of the income effects in the US on life satisfaction in the models with and without DIF. One should note however, that all thresholds play a role, not only the first one. Disentangling the effect of the threshold shifts may be a complicated matter. We prefer therefore

to investigate the importance of threshold differences between countries and between demographic groups within countries by a series of simulations.

6. Model Simulations

A transparent way of understanding the implications of our approach is to simulate the distribution of life satisfaction in the two countries for different parameter values. Essentially we first simulate the Dutch distribution of self-reported life satisfaction and then replace various sets of parameters by the corresponding American values. Table 8 presents the results of these simulations by four age groups—those less than 40, 40-50 years old, 50-64 years old, and at least 65 years old. The first row for each age group summarizes the distribution of satisfaction with income for the Dutch using their own parameters. The second row replaces Dutch thresholds by American thresholds (cf. Table 7).The third row simulates the Dutch distribution if we replace the parameters in the Dutch satisfaction equation (i.e. Table 6 with DIF) by the American parameters. The fourth row replaces all Dutch parameters by American parameters. The fifth row simulates distributions for the American sample using American parameters. Table 9 lists similar simulations by income quartile instead of age.

For each age group in Table 8, the first row approximately reproduces the distribution of self-reports in the Dutch sample, while the fifth row does the same for the US sample. Comparing the first two rows in each panel shows that the Dutch self reports would become more spread out when Dutch respondents would evaluate their satisfaction with life using US thresholds. Both the percentage very satisfied and the percentage dissatisfied go up. This corresponds to the notion that the Dutch tend to avoid extremes; giving them the US thresholds makes them more likely to report the two extreme categories. Comparing rows 1 and 2 with row 5 then shows that correcting for response scale differences does not make the distribution of life

satisfaction in the Netherlands and the US more similar in all respects. For example, for all age groups combined (final panel), we find that after the correction a much larger fraction in the Dutch sample are very satisfied with their life than in the US sample. The fraction not satisfied/ dissatisfied or worse increases somewhat in the Dutch sample and comes somewhat closer to the US fraction, but remains substantially smaller. Both before and after correction for response scale differences, the Dutch population as a whole is more satisfied with their lives than the Americans. This does not apply to the oldest age group, however: Americans of 65 years and older are somewhat more satisfied with their lives than their Dutch counterparts, on average, irrespective of whether we give them the same scales or not.

Rows 3 and 4 in each panel can be used to show how much of the remaining differences (keeping response scales constant across countries) is due to differences in observed characteristics, generalizing the traditional Oaxaca-Blinder decomposition to a non-linear model (cf., e.g., Yun, 2004). In particular, comparing rows 4 and 5 shows the differences explained by differences in background characteristics between the two countries, using US evaluation standards (both in the self-assessment equation and for the thresholds). The results show that, although the differences are modest, the characteristics make the Dutch in all age groups more satisfied with their lives than the Americans. The most important characteristic driving this is partnership status: having a partner has a strong positive effect on satisfaction with life, and the fraction with partner is much higher in the Netherlands than in the US (78% versus 64%).

On the other hand, comparing rows 2 and 3 in each panel of Table 8 shows that giving the Dutch the US parameters for the self-assessment (but keeping the Dutch thresholds) also brings about substantial shifts, where for younger ages the imposition of US parameters on Dutch respondents leads to lower simulated satisfaction, while for higher ages it leads to more

satisfaction. This is a direct reflection of the results in Table 7, which show rather strong interaction effects with the US dummy for the age brackets 40-50 and 51-64.

Next, let's turn our attention to Table 9, which does the same thing as Table 8 but for income quartiles instead of age groups. The effect of assigning US thresholds again leads to more dispersion in the responses (row 2 compared to row 1). For the highest income quartile, comparing row 2 and row 5 shows that the US respondents are better off. This was not clear from the first row, due to the reluctance of the Dutch to classify themselves as dissatisfied or very dissatisfied.

Assigning US self assessment parameters to the Dutch confirms the stronger effect of income on life satisfaction in the US than in The Netherlands (row 3 versus row 2). We see that with the US self assessment parameters, Dutch respondents with low incomes would be considerably less satisfied. Conversely with high incomes they would be more satisfied. When the Dutch are assigned both the US self assessment parameters and the US thresholds, then the satisfaction distribution more closely resembles that of the US (rows 4 and 5), and again show that the differences in background characteristics somewhat favor the Dutch, mainly in the third income quartile.

7. Conclusions

We have analyzed the determinants of global life satisfaction, by using both self-reports and responses to a battery of vignette questions. Although more work needs to be done, some preliminary conclusions can be drawn.

It appears that the four domains job or daily activities, social contact and family, health, and income provide a fairly complete description of global life satisfaction in both countries.

Among the four domains, social contacts and family have the highest impact on global life satisfaction, followed by job and daily activities and health. Income has the lowest impact.

As in other work, we find that American response styles differ from the Dutch in that Americans are more likely to use the extremes of the scale (either very satisfied or very dissatisfied) than the Dutch, who are more inclined to stay in the middle of the scale.

Although for both Americans and the Dutch, income is the least important determinant of global life satisfaction, it is more important in the U.S. than in The Netherlands. Indeed life satisfaction varies substantially more with income in the U.S. than in The Netherlands.

There are some intriguing differences between the way respondents judge vignette persons and what turns out to influence their own satisfaction. Respondents in both The Netherlands and the U.S. appear to think that marriage does not contribute to life satisfaction when they judge vignettes. Yet their own satisfaction is positively influenced by being married. Similarly, respondents believe that other things being equal, older persons should be less satisfied. Yet their own satisfaction goes up with age.

The estimates of an econometric model are used to calculate counterfactual distributions of life satisfaction. Correcting for differences in response scales leads to some shifts though the shifts are not very large. For most age and income groups, the conclusion that the Dutch are more satisfied with their lives than the Americans remains valid. For the oldest age group (65+) and highest income group, however, the vignette corrections lead to different conclusions: giving Dutch respondents the American scales shows that these groups are somewhat less satisfied than their US counterparts. This was not clear from the distributions using own country's scales, mainly because of the Dutch reluctance to evaluate their satisfaction as dissatisfied or very dissatisfied.

Vignettes have been shown to bring objective and subjective measurements of health (in particular, vision) or drinking behavior closer in line with each other. An objective measure for life satisfaction seems hard to give, so that other ways of validation need to be considered, perhaps by looking at actual behaviors that are correlated with life satisfaction. This is one of the directions of future research.

References

Alesina, Alberto. Rafael Di Tella and Robert MacCulloch. 2004. "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics,* 88(9-10), 2009-2042.

Blanchflower, David G. and Oswald, Andrew J. 2004. 'Well-Being Over Time in Britain and the USA." *Journal of Public Economics,* 88(7-8), 1359-1386.

Clark, Andrew E., Paul Frijters, and Michael Shields. 2008. "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature*. 46(1), 95-144.

Deaton, Angus. 2008. "Income, Aging, Health and Well-Being Around the World: Evidence from the Gallup World Poll." *Journal of Economic Perspectives*. 22(2), 53-72.

Di Tella, Rafael, MacCulloch, Robert J and Blanchflower, David G. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics*. 85(4), 809-827.

Di Tella, Rafael, John Haisken–DeNew, and Robert MacCulloch, 2007, "Happiness adaptation to income and to status in an individual panel," processed, October

Easterlin Richard A. (1974), "Does Economic Growth Improve the Human Lot? Some Empirical Evidence", in: R. David and M. Reder (eds.) Nations and Households in Economic Growth: Essays in honor of Moses Abramowitz, New York, Academic Press, 89-125.

Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization*, 27(1), 35-48.

Easterlin, Richard A. 2006. "Life Cycle Happiness and its Sources: Intersections of Psychology, Economics, and Dempography." *Journal of Economic Psychology*, 27, 463, 482.

Ferrer-i-Carbonell, Ada, and Bernard M.S. van Praag (2002), "The subjective costs of health losses due to chronic diseases. An Alternative model for monetary appraisal." *Health Economics*, 11, 709-722.

Kapteyn, Arie., Smith, James. P. & van Soest, Arthur. 2007. "Vignettes and self-reports of work disability in the U.S. and the Netherlands." *American Economic Review,* 97(1), 461-473.

Kapteyn, Arie., Smith, James. P. & van Soest, Arthur. 2008. "Are Americans Really Less Happy With Their Incomes?", RAND Working Paper, WR-591.

King, Gary; Murray, Christopher; Salomon, Joshua and Tandon, Ajay. 2004. "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research." *American Political Science Review,* 98(1), 567-583.

Layard, Richard. 2005. *Happiness: Lessons from a New Science*. Penguin Books: London.

Rao, Jon N. K., and Alastair J. Scott. 1984. "On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data." *Annals of Statistics,* 12: 46-60.

Stevenson, Betsey and Wolfers, Justin. 2008, "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," Working Paper, Wharton School, University of Pennsylvania, prepared for Brookings Papers on Economic Activity, Spring 2008.

Van Beek, Krijn W. H., Carl C. Koopmans and Bernard M.S. van Praag (1997), "Shopping at the labour market: A real tale of fiction." European Economic Review, Elsevier, 41(2), 295-317.

Van Praag, Bernard M.S., and Ada Ferrer-i-Carbonell (2008), Happiness Quantified – A *Satisfaction Calculus Approach*, Oxford University Press, Oxford.

Van Praag, Bernard M.S., Paul Frijters and Ada Ferrer-i-Carbonell (2003), "The Anatomy of Subjective Well-being." *Journal of Economic Behavior and Organization*, 51, 29-49.

Van de Stadt, Huib, Arie Kapteyn & Sara van de Geer. 1985. "The Relativity of Utility: Evidence from Panel Data." *The Review of Economics and Statistics,* 67, 179-187.

Van Soest, Arthur., Liam Delaney, Liam, Harmon, Colm, Arie Kapteyn Arie., Smith, James P. 2007. "Validating the Use of Vignettes for Subjective Threshold Scales." RAND Labor and Population working paper WR-501.

Yun, Myeong-Su. 2004. "Decomposing Differences in the First Moment." *Economics Letters*, 82(2), 275-280.

Table 1
Self Reports on Satisfaction with Domains of Life

Self report: How satisfied are you with the total income of your household?	Country	
	NL	US
Very satisfied	9.9	6.5
Satisfied	53.6	39.4
Not satisfied or dissatisfied	23.6	21.5
Not satisfied	10.3	27.4
Very dissatisfied	2.7	5.2

Test for independence: F(3.64, 12207.95) = 20.3117; p-value = 0.0000

Self report: How satisfied are you with your job or other daily activities?		
	NL	US
Very satisfied	19.4	16.3
Satisfied	61.7	52.2
Not satisfied or dissatisfied	14.7	17.5
Not satisfied	3.4	12.1
Very dissatisfied	0.8	2.0

Test for independence: F(3.36, 11231.88) = 11.4447; p-value = 0.0000

Self report: How satisfied are you with your social contacts and family life?		
	NL	US
Very satisfied	23.0	27.1
Satisfied	62.8	48.2
Not satisfied or dissatisfied	11.7	15.7
Not satisfied	1.9	8.5
Very dissatisfied	0.6	0.5

Test for independence: F(3.58, 11978.15) = 13.9798; p-value = 0.0000

Self report: How satisfied are you with your health?		
	NL	US
Very satisfied	15.4	16.1
Satisfied	61.6	46.7
Not satisfied or dissatisfied	14.5	17.3
Not satisfied	7.0	16.5
Very dissatisfied	1.4	3.5

Test for independence: F(3.82, 12791.55) = 13.3638; p-value = 0.0000

Self report: How satisfied are you with your life in general?		
	NL	US
Very satisfied	19.3	20.1
Satisfied	68.2	58.0
Not satisfied or dissatisfied	10.9	15.4
Not satisfied	1.3	5.3
Very dissatisfied	0.3	1.1

Test for independence: F(3.17, 10610.55) = 9.2306; p-value = 0.0000

Note: all frequencies are weighted with sampling weights. Tests are Pearson chi-squared tests for independence converted into F-statistics, accounting for the weighting (see Rao and Scott, 1984).

Table 2
Ordered Probits for Global Life Satisfaction Against Satisfaction with Specific Domains

	Coef.	z	Coef.	z
Income Domain	.225	6.85	.220	6.53
Relations Domain	.721	12.22	.708	11.85
Job Domain	.625	12.27	.626	12.22
Health Domain	.486	11.02	.497	11.10
US Income Domain	.052	0.85	.031	0.49
US Relations Domain	.087	1.11	.101	1.26
US Job Domain	-.020	0.26	-.016	0.20
US Health Domain	-.131	1.98	-.143	2.12
Dummy US Domain	-.145	0.82	-.037	0.68
Married			-.128	1.38
Age 40-50			-.023	0.31
ln family size			-.024	0.35
Age 51-64			-.093	1.29
Age 65+			.023	0.22
Ed med			-.128	1.95
Ed high			-.117	1.77
Working			.116	1.73
ln eq income			.009	0.61
Dummy US			-.255	0.25
US female			.084	0.80
US married			-.088	0.59
US ln family size			.226	1.54
US age 40-50			.108	0.74
US age 51-64			.198	1.32
US age 65+			-.048	0.22
US ed med			-.038	0.25
US ed high			-.059	0.40
US working			-.077	0.60
US ln eq income			-.001	0.02

Table 3
Variation in Global Life Satisfaction Vignettes

	Age	Family	Income	Work	Health
1	42	good	median +	stressful	some pain
2	50	moderate	half median +	ok with long hours	good
3	65	bad	modest	retired	heart problems
4	25	no friends	half median +	no control or security	good
5	25	good	half-median to twice median	no control but secure	good
6	57	bad	median +	retired	pain
7	75	good	half median to twice median	retired	arthritis
8	62	good	half median +	retired	good
9	70	bad	half median +	retired	moderate
10	50	bad	twice median +	good	bad

Table 4

Global Vignettes for the Retired

	Vignette Number									
	3		**6**		**7**		**8**		**9**	
Family	*Bad*		*Bad*		*Good*		*Good*		*Bad*	
Health	*Heart bad*		*Pain*		*Arthritis*		*Good*		*Moderate*	
Age	*65*		*57*		*75*		*62*		*70*	
	NL	US	NL	US	NL	US	NL	US	NL	US
Very satisfied	0.1	1.0	1.6	0.0	13.3	18.1	70.7	77.4	1.4	5.4
Satisfied	25.0	12.7	19.6	10.5	67.4	69.2	26.1	18.8	7.5	4.4
Not satisfied or dissatisfied	48.9	27.1	48.6	40.6	16.3	9.0	3.2	3.0	39.5	19.6
Not satisfied	24.7	51.7	28.6	48.1	3.0	3.8	0.0	0.8	49.0	63.0
Very dissatisfied	1.3	7.6	1.6	0.8	0.0	0.0	0.0	0.0	2.7	7.6

Global Vignettes for the Young **Global Vignettes for the Middle Aged**

	Vignette Number									
	4		**5**		**1**		**2**		**10**	
Family	*no friends*		*good*		*good*		*moderate*		*bad*	
Work	*no control*		*no control*		*stressful*		*ok-long hours*		*good*	
	no security		*good security*							
Health	*good*		*good*		*some pain*		*good*		*bad*	
Age	*25*		*25*		*42*		*50*		*50*	
	NL	US	NL	US	NL	US	NL	US	NL	US
Very satisfied	1.9	3.2	23.9	33.1	16.9	15.4	19.8	23.0	1.5	1.1
Satisfied	14.9	14.9	66.8	58.8	64.3	65.0	56.5	50.0	11.7	8.5
Not satisfied or dissatisfied	39.8	37.2	8.0	7.4	18.8	22.5	19.8	21.6	36.6	36.0
Not satisfied	41.6	41.5	1.3	0.0	2.2	6.0	2.8	5.4	45.3	43.4
Very dissatisfied	1.9	3.2	0.0	0.3	0.5	0.0	1.1	0.0	5.0	11.1

All vignettes evaluated at highest income level in the vignette

Table 5: Effect of vignette descriptions on evaluation

	(1) Regression	(2) Ordered Probit	(3) Probit (very) dissatisfied
Age	0.003	-0.001	0.026
	(1.29)	(0.39)	(5.21)**
Married	0.171	0.263	0.259
	(3.97)**	(4.12)**	(2.17)*
Good relations	-1.211	-1.636	-1.964
	(39.11)**	(32.91)**	(24.24)**
Retired	-0.153	-0.091	-0.765
	(2.11)*	(0.83)	(4.40)**
Modest income	0.947	1.318	1.566
	(20.23)**	(18.77)**	(12.27)**
Half median income	0.218	0.334	0.406
	(8.22)**	(8.21)**	(6.28)**
Twice median income	-0.145	-0.234	-0.153
	(6.17)**	(6.43)**	(2.46)*
Four times median income	-0.279	-0.448	-0.294
	(9.35)**	(9.59)**	(4.14)**
Health good	-0.477	-0.955	-0.244
	(9.43)**	(11.47)**	(1.43)
job secure or under control	0.134	0.290	-0.094
	(3.66)**	(5.02)**	(0.88)
at least 50 hrs	0.693	1.215	0.958
	(12.91)**	(14.38)**	(6.09)**
Dummy US	-0.160	-0.216	-0.164
	(1.00)	(0.89)	(0.45)
US Age	0.007	0.010	0.009
	(2.41)*	(2.17)*	(1.22)
US Married	0.227	0.354	0.510
	(3.53)**	(3.63)**	(2.90)**
US Good relations	-0.218	-0.321	-0.111
	(4.65)**	(4.46)**	(0.95)
US Retired	-0.269	-0.400	-0.342
	(2.45)*	(2.38)*	(1.31)
US Modest income	0.652	0.925	0.885
	(8.75)**	(8.26)**	(4.52)**
US Half median income	0.100	0.153	0.020
	(2.37)*	(2.36)*	(0.20)
US Twice median income	-0.054	-0.093	-0.172
	(1.52)	(1.66)	(1.89)
US Four times median income	-0.058	-0.109	-0.158
	(1.27)	(1.49)	(1.48)
US Health good	0.172	0.200	0.418
	(2.25)*	(1.57)	(1.69)
US job secure or under control	-0.200	-0.284	-0.399
	(3.62)**	(3.25)**	(2.58)**
US at least 50 hrs	-0.078	-0.062	-0.235
	(0.97)	(0.49)	(1.02)
Constant	3.265		-1.168
	(30.70)**		(4.87)**
Observations	12051	12051	12051
R-squared	0.55		

Notes: Robust t statistics in parentheses; * significant at 5% level; **: significant at 1% level
Model (3): dependent variable: 1 if very dissatisfied or dissatisfied; 0 otherwise

Table 6

Self Assessment of Global Satisfaction

	Model with DIF		Model without DIF	
	β	s.e.	β	s.e.
Constant	1.005**	0.14	0.899**	0.12
Female	-0.081	0.06	-0.052	0.05
Married	-0.431**	0.11	-0.361**	0.10
ln family size	-0.205*	0.10	-0.205*	0.09
Age 40-50	0.043	0.09	0.119	0.08
Age 51-64	0.066	0.09	0.054	0.07
Age 65+	-0.141	0.11	-0.169+	0.09
Ed med	-0.018	0.08	-0.054	0.07
Ed high	-0.237**	0.08	-0.163*	0.07
Working	-0.160*	0.07	-0.034	0.06
ln eq income	-0.089*	0.04	-0.098**	0.04
Interactions with dummy US				
Dummy	0.217	0.24	0.270	0.20
Female	0.008	0.10	-0.055	0.83
Married	0.095	0.15	0.015	0.13
ln family size	-0.211	0.15	-0.078	0.13
Age 40-50	-0.274+	0.15	0.138	0.13
Age 51-64	-0.156	0.14	0.033	0.13
Age 65+	-0.346+	0.20	-0.397**	0.16
Ed med	-0.107	0.13	-0.018	0.11
Ed high	-0.026	0.14	-0.027	0.11
Working	-0.008	0.12	-0.127	0.10
ln eq income	-0.415**	0.08	-0.327**	0.06

** indicates significance at 1% level, * indicates significance at the 5% level, and + indicates significance at the 10% level.

Table 7
Thresholds of Estimated Equation for Global Life Satisfaction

	Threshold 1		ln (Threshold 2 – Threshold 1)		ln (Threshold 3 – Threshold 2)		ln (Threshold 4 – Threshold 3)	
	β	s.e.	β	s.e.	β	s.e.	β	s.e.
Constant	0.00	0.00	0.53	0.17	0.81*	0.20	-0.34	0.36
Female	-0.04	0.05	0.02	0.02	0.06	0.04	-0.00	0.04
Married	-0.04	0.07	0.01	0.04	0.11	0.07	0.15*	0.08
ln family size	-0.00	0.07	0.04	0.04	-0.12*	0.06	-0.01	0.07
Age 40-50	-0.07	0.07	-0.03	0.03	0.02	0.05	0.06	0.06
Age 51-64	0.03	0.06	-0.03	0.03	0.07	0.05	-0.10	0.06
Age 65+	0.08	0.08	-0.06	0.05	0.16*	0.07	-0.16+	0.08
Ed med	0.06	0.06	0.02	0.03	0.01	0.04	-0.01	0.06
Ed high	-0.10+	0.06	0.06*	0.03	0.02	0.04	0.00	0.06
Working	-0.21*	0.06	0.09*	0.03	0.03	0.04	-0.05	0.05
ln eq income	0.02	0.04	0.01	0.02	-0.07*	0.02	0.08*	0.04
Interactions with dummy US								
Dummy	0.08	0.18	-0.31	0.34	-0.84+	0.43	-0.13	0.55
Female	0.08	0.07	-0.03	0.04	-0.10+	0.06	0.03	0.06
Married	-0.03	0.10	0.06	0.06	-0.02	0.09	-0.30*	0.10
ln family size	-0.17	0.11	-0.03	0.06	0.09	0.09	0.13	0.10
Age 40-50	0.23*	0.11	-0.06	0.06	-0.00	0.08	-0.05	0.09
Age 51-64	0.19+	0.10	-0.05	0.06	-0.15+	0.09	0.13	0.09
Age 65+	0.04	0.15	-0.02	0.08	-0.09	0.12	0.17	0.12
Ed med	-0.09	0.11	-0.01	0.06	-0.04	0.08	0.04	0.09
Ed high	-0.03	0.11	0.02	0.06	-0.09	0.08	0.08	0.09
Working	0.15	0.09	-0.06	0.05	-0.04	0.07	0.07	0.07
ln eq income	-0.16*	0.05	0.02	0.03	0.07	0.04	0.01	0.05

*Indicates significance at the 5% level; + indicates significance at the 10% level. N = 2244 for NL and 1093 for US.

Table 8

Simulations from Model with DIF: Percent Distribution of Global Satisfaction by Age Group

	Very Satisfied	Satisfied	Not Satisfied/ Dissatisfied	Dissatisfied	Very Dissatisfied
Age group younger than 40					
Dutch sample using own parameters	21.3	66.2	11.2	1.3	0.0
Dutch using US threshold parameters	24.2	59.8	12.9	3.0	0.0
Dutch using US self-assessment parameters	17.5	64.8	14.8	2.7	0.2
Dutch using all US parameters	19.4	59.0	16.1	5.1	0.4
US sample using US parameters	17.3	56.9	18.2	7.1	0.5
Age group 40-50					
Dutch sample using own parameters	19.1	66.0	13.2	1.8	0.0
Dutch using US threshold parameters	28.1	55.8	13.2	2.9	0.0
Dutch using US self-assessment parameters	10.8	63.1	21.0	4.8	0.3
Dutch using all US parameters	16.8	55.2	20.5	7.0	0.5
US sample using US parameters	17.1	55.0	20.3	7.2	0.4
Age group 50-64					
Dutch sample using own parameters	20.0	64.0	14.3	1.6	0.0
Dutch using US threshold parameters	27.5	55.9	13.0	3.6	0.1
Dutch using US self-assessment parameters	14.5	61.9	19.9	3.5	0.2
Dutch using all US parameters	20.1	55.6	17.5	6.5	0.4
US sample using US parameters	19.6	54.0	18.0	7.8	0.5
Age group 65 and older					
Dutch sample using own parameters	26.5	58.9	13.4	1.1	0.0
Dutch using US threshold parameters	27.6	55.3	14.3	2.7	0.0
Dutch using US self-assessment parameters	35.5	53.6	10.1	0.8	0.0
Dutch using all US parameters	36.3	51.3	10.5	1.9	0.1
US sample using US parameters	32.2	52.1	13.0	2.6	0.1
All age groups					
Dutch sample using own parameters	21.5	64.1	12.9	1.5	0.0
Dutch using US threshold parameters	26.7	56.9	13.3	3.1	0.0
Dutch using US self-assessment parameters	18.7	61.4	16.7	3.0	0.2
Dutch using all US parameters	22.4	55.6	16.4	5.3	0.3
US sample using US parameters	19.6	55.1	17.9	6.6	0.4

N=2244 for NL and 1093 for US

Table 9

Simulations from Model with DIF: Percent Distribution of Global Satisfaction by Income Group

	Very Satisfied	Satisfied	Not Satisfied/ Dissatisfied	Dissatisfied	Very Dissatisfied
Lowest Income Quartile					
Dutch sample using own parameters	20.9	65.4	12.4	1.3	0.0
Dutch using US threshold parameters	28.5	56.4	12.4	2.7	0.1
Dutch using US self-assessment parameters	11.3	59.4	22.9	5.8	0.6
Dutch using all US parameters	14.9	53.7	21.4	8.9	1.1
US sample using US parameters	14.2	51.8	22.8	10.3	0.9
Second Income Quartile					
Dutch sample using own parameters	22.3	64.4	12.0	1.3	0.0
Dutch using US threshold parameters	27.6	57.0	12.6	2.7	0.0
Dutch using US self-assessment parameters	17.6	64.0	16.0	2.3	0.0
Dutch using all US parameters	21.7	57.5	16.3	4.5	0.1
US sample using US parameters	22.7	58.4	14.9	3.9	0.1
Third Income Quartile					
Dutch sample using own parameters	22.1	63.0	13.3	1.6	0.0
Dutch using US threshold parameters	26.3	56.8	13.6	3.2	0.0
Dutch using US self-assessment parameters	22.1	61.2	14.5	2.1	0.0
Dutch using all US parameters	25.7	55.8	14.5	4.0	0.1
US sample using US parameters	22.7	56.2	15.9	5.0	0.1
Highest Income Quartile					
Dutch sample using own parameters	20.6	63.5	14.1	1.8	0.0
Dutch using US threshold parameters	24.0	57.5	14.6	3.8	0.0
Dutch using US self-assessment parameters	24.8	60.9	12.5	1.7	0.0
Dutch using all US parameters	28.0	55.7	12.9	3.4	0.0
US sample using US parameters	26.7	57.4	12.6	3.2	0.0
All Income groups					
Dutch sample using own parameters	21.5	64.1	12.9	1.5	0.0
Dutch using US threshold parameters	26.7	56.9	13.3	3.1	0.0
Dutch using US self-assessment parameters	18.7	61.4	16.7	3.0	0.2
Dutch using all US parameters	22.4	55.6	16.4	5.3	0.3
US sample using US parameters	19.6	55.1	17.9	6.6	0.4

N=2244 for NL and 1093 for US; Income is equivalized income (per capita). Quartiles are country specific.

APPENDIX - Life Satisfaction Vignettes

Global Life Satisfaction

Global 1: (*Name*) is 42 years old, happily married, with two children who are doing well at school and generally get on well with their parents. His/her family income is about xxx (median, twice median, four times median). He/she likes his work although some days it is somewhat stressful. (*Name*) suffers from rather serious back pain that keeps him/her awake at night about once a week, but has no other serious health problems.

Global 2: (*Name*) is 50 years and divorced. He/she has one daughter of 22 with whom he gets on well, although he/she sees her only once a year. *(Name)* works about 60 hours per week, and feels he/she has a very secure job over which he has a lot of control. He/she makes about xxx (half the median, median, twice the median, four times the median) per year. He/she has no serious health problems.

Global 3: (*Name*) is 65 years old. His/her wife/husband died 3 years ago and he/she still spends most of his time thinking about his/her and the good times they had together. He/she has four children and 10 grandchildren who visit him/her regularly. (*Name)* has a small pension and receives social security; he/she can make ends meet but has no money for extras such as expensive gifts to his/her grandchildren. He/she has had to stop working recently due to heart problems. He/she gets tired easily and can, for example, not walk more than one block without taking a pause. Otherwise, he/she has no serious health conditions.

Global 4: (*Name*) is 25 years old and single. He/she does not have many friends. He/she works about 50 hours a week and makes xxx (half the median, median, twice the median). He/she feels he has little control over his/her job and worries about loosing it. He/she has no health problems, but feels a little stressed sometimes. He/she does not exercise.

Global 5: (*Name*) is 25 years old and recently married, no children. He/she works about 35 hours per week and makes xxx (half the median, median, twice the median). He/she works out regularly and on vacations he/she makes long hikes in the mountains with his/her husband. He/she job is satisfying, though a bit dull sometimes. He/she feels she does not have a lot of control over his/her job, but it is a very secure job.

Global 6: (*Name*) is 57 years old and recently married his/her second wife/husband. He/she has two children from his/her first marriage, but has little contact with them. He/she draws DI, because he/she has serious back pains. He/she often has trouble sleeping. His/her DI benefits are xxx (half the median, median, twice the median).

Global 7: (*Name*) is 75 years old and a widow. His/her pension benefits are xxx (half the median, median, twice the median). He/she owns the house he/she lives in and has a large circle of friends. He/she plays bridge twice a week and goes on vacation regularly with some friends. Lately he/she has been suffering from arthritis, which makes work in the house and garden painful.

Global 8: (*Name*) is 62 and has been retired for five years. He/she quit his job as soon as he could. He/she has never regretted his/her decision to retire. His/her pension is xxx (median, twice median; four times the median) He/she is physically very active and makes long bicycle trips in Southern Europe. He/she is single, but usually makes the trips with friends his/her age.

Global 9: (*Name*) is 70 and has been retired for five years. His/her pension is xxx (median, twice median; four times the median). He/she still misses the contacts with his colleagues and would have liked to keep working part time. He/she and his wife/her husband take a few small vacations every year. For the rest they each lead their own lives and don't do many things together. They have two children but rarely see them. He/she is overweight and gets tired when walking more than a few blocks. He/she has been a smoker all his life.

Global 10: (*Name*) is 50 and does not exercise. He/she cannot climb stairs or do other physical activities because he/she is obese. He/she has pain in his/her knees, elbows, wrists and fingers, and the pain is present almost all the time. He/she has an executive job in a big firm and feels that he/she has a lot of control over his job. He/she makes xxx (twice the median, four times the median). He/she has been married for a long time, but he/she and his/her wife spend very little time together.