

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/74895>

Please be advised that this information was generated on 2019-03-21 and may be subject to change.

NiCE Working Paper 09-107

February 2009

Measuring values for cross-cultural research

Robbert Maseland

André van Hoorn

Nijmegen Center for Economics (NiCE)

Institute for Management Research

Radboud University Nijmegen

P.O. Box 9108, 6500 HK Nijmegen, The Netherlands

<http://www.ru.nl/nice/workingpapers>

Abstract

This paper investigates the empirical relevance of the recent critique that values surveys, as they are, suffer from the problem of measuring marginal preferences rather than values. By surveying items from cross-cultural surveys by Hofstede, Inglehart and GLOBE, we show that the marginal preferences problem is pervasive, even in cases where improved survey items are used. As an alternative to survey items about the importance respondents attach to various objectives, we therefore propose the use of so-called happiness or experienced utility functions as method to assess more directly what people across cultures care about and how much. Several recent contributions show the feasibility of such methods. We conclude that value measures based on cultural variation in the structure of happiness appear viable and superior substitutes for survey-based value indicators.

Robbert Maseland, Radboud University Nijmegen, Department of Political Science, PO Box 9108, 6500 HK, Nijmegen, the Netherlands, Tel: +31 243 611 564, Fax: +31 243 612 379, E-mail: R.Maseland@fm.ru.nl.

André van Hoorn (corresponding author), Radboud University Nijmegen, Department of Economics, PO Box 9108, 6500 HK, Nijmegen, the Netherlands, Tel: +31 243 612 344, Fax: +31 243 612 379, E-mail: A.vanHoorn@fm.ru.nl.

1. Introduction

Critique of values surveys and quantitative measures of culture is probably as old as the indicators themselves. Typically, critical assessments of cross-cultural indicators have focused on methodology. Hofstede (2001: 73) reviews critiques on the survey-based approach to measuring value preferences in general and his framework in particular. Recently, Maseland and van Hoorn (2008) have added a new theoretical dimension to this debate. Employing textbook microeconomics to explain the negative relation between values and practices reported by the GLOBE values survey (House, Hanges, Javidan, Dorfman, & Gupta, 2004), they find that these results indicate that GLOBE is not eliciting values (the importance attached to an objective in general) but marginal preferences (the importance attached to an objective on top of the current level of satiation).

An open question is whether this problem is specific to the GLOBE survey or is more widespread. Since the GLOBE study follows the same basic methodology as other values surveys, it is likely that these other surveys are also prone to mistaking marginal preferences for values. This paper investigates whether this is indeed the case. In order to do so, we survey two of the most widely used values surveys, namely the World Values Survey (WVS) coordinated by Ronald Inglehart (e.g. Inglehart, 1990) and the work of Geert Hofstede (Hofstede, 2001). The value items we take concern a range of objectives, and data is available at the individual level, at the societal level or both. The empirical tests show that values surveys scores tend to correlate negatively with practices. This indicates that the failure of values surveys to elicit “true” values is widespread, and that there is much evidence consistent with the marginal-preferences interpretation of values surveys scores.

A second open question is what we can do about this. Maseland and van Hoorn (2008) are short on solutions to the marginal-preferences problem and the format for questions they propose remains rather unconvincing, as we show on the basis of empirical results. The paper subsequently introduces an alternative approach to measuring cultural differences by looking at cross-cultural differences in the impact of various factors on people’s happiness (cf. Di Tella and MacCulloch, 2006: 39-43).

Accordingly, the paper contributes to the body of literature on cross-cultural comparisons in two ways. Firstly, it points out serious problems with one of the more popular measurements of cultural differences, which has important implications for the type of research in international business that relies on cross-cultural comparative research. Secondly, it discusses a solution to the marginal-preferences problem in the form of so-called

happiness or experienced utility functions. Allowing the exact shape of these functions to differ between cultures, appears a promising alternative strategy to elicit the value differences that are of relevance to researchers seeking to do cross-cultural comparative research.

The next section reviews the theoretical argument of Maseland and van Hoorn (2008), sketching the implications of the distinction between values and marginal preferences. In the section that follows, we present evidence from several sources, supporting the view that values surveys elicit marginal preferences rather than values. A fourth section considers different value indicators that alleviate the marginal-preferences problem, proposing cross-cultural heterogeneity in happiness functions as an alternative to the survey-based method of measuring values. The final section offers some summarizing thoughts.

2. Theoretical framework

Economic theory assumes that individuals have an infinite range of objectives. Each separate want is limited; the desire for any objective declines with the level of satiation of that objective. Alfred Marshall (1920 [1890]) labeled this tendency of human nature the *law of diminishing utility*.¹ Graphically, this principle can be captured by a convex utility function (Figure 1), indicating that the extra utility gained by consuming an additional unit of a particular good falls with the consumption of the good. Thus, at point B in the figure, the value attached to more of the good is lower than at point A, because the desire for the good is satisfied to a larger extent already.

[Insert Figure 1 about here]

For the interpretation of values surveys, it is important to note that the value attached to extra satiation of an objective tells us nothing about underlying tastes for objectives in general. For each objective, no matter how important, a point exists at which satiating it any further no longer takes precedence over other objectives. When more and more police are added, for instance, even the most dedicated law-and-order enthusiast eventually comes to a

¹ Laboratory experiments by Horowitz, List and McConnell (2007) provide strong support for the principle of diminishing marginal value. Cross-sectional evidence from the literature on the economics of happiness (e.g. Frey & Stutzer, 2002) shows that self-reported happiness or experienced utility increases with income but at a diminishing rate.

point where (s)he cares little about more police on the street. To conclude from this that this person attaches low value to law and order is clearly wrong; it is simply that the desire for law and order has been satisfied to a great extent already, whereas other desires have not (cf. Maseland and van Hooft, 2008).

The question raised by Maseland and van Hooft (2008) is whether values surveys succeed in eliciting values, or tend to capture marginal preferences instead. In so far as values surveys elicit marginal preferences, they are flawed instruments for measuring cultural differences. In order to answer this question, we have to look into the relation between values surveys scores and practices. Values are thought to correlate positively with practices (or, more general, levels of “consumption”), since it is reasonable to assume that people by and large act upon their values (Hofstede, 2001: 11). The marginal preference for an objective, in contrast, declines with satiation of an objective, implying a negative correlation with practices. This results in two rival hypotheses:

Hypothesis 1a. If questionnaire items predominantly measure values, response scores and practice or consumption measures will be positively correlated.

Hypothesis 1b. If questionnaire items predominantly measure marginal preferences, response scores and practice or consumption measures will be negatively correlated.

The GLOBE project (House et al., 2004), including values items alongside practice measures, offered a first test of these hypotheses. On seven of its nine dimensions of culture, it reported a negative relation between values items and practice measures, indicating that GLOBE elicited marginal preferences rather than values (Maseland & van Hooft, 2008; see also Hofstede, 2006; Javidan, House, Dorfman, Hanges & Sully de Luque, 2006).² In the next section, we test whether this verdict extends beyond the GLOBE study to values surveys in general.

² The dimensions Assertiveness ($r = -0.26$), Institutional Collectivism (-0.61), Future Orientation (-0.41), Humane Orientation (-0.32), Performance Orientation (-0.28), Power Distance (-0.43), and Uncertainty Avoidance (-0.62) all show statistically significant negative correlations ($p < 0.05$; $n = 61$). In-Group Collectivism shows a positive but insignificant correlation, while Gender Egalitarianism is the only dimension with a positive correlation ($r = 0.32$; $n = 61$; $p < 0.05$) that is significant at usual levels (House et al., 2004: 736).

3. Empirical evidence

In order to establish the empirical relevance of the marginal preferences-critique, we have sampled a number of items from various values surveys for which practice or consumption measures are available. This latter requirement necessarily limits the range of items for which we can test the marginal preferences hypothesis; in so far as explicit measures of practices are not included in values surveys, consumption levels of rather abstract issues like collectivism are difficult to obtain.

The World Values Survey (WVS) coordinated by Ronald Inglehart provides the richest source for our purpose, though it is perhaps not that frequently used in international business research. The publicly available World Values Survey dataset (European Values Study Group & World Values Survey Association, 2006) contains scores on specific items, which can be more readily linked to practices than overarching dimensions. Also, individual-level data is available, which can be aggregated to the country level so that we can test whether values surveys elicit marginal preferences or, as generally assumed, values at both levels. In addition, we look at the work of Hofstede (2001), which is more widely used in international business studies. Hofstede's data itself does not include direct information on practices other than limited general background data. In order to assess the relation between values and practices, we have to rely on other sources for practice measures, such as the GLOBE study.

3.1 Inglehart's World Values Survey

Not unlike the GLOBE project, the WVS is a large-scale data-gathering effort by researchers from around the world and nowadays coordinated by Ronald Inglehart. The WVS has grown out of the European Values Survey (EVS), which in the period 1981-1984 surveyed respondents from a number of European countries. Since then, the WVS has evolved in a global project, which, in different waves, spans some 80% of world population. Individual responses for all waves (1: 1981-1984, 2: 1989-1993, 3: 1994-1999 (WVS only), and 4: 1999-2004), almost 270,000 individuals in total, are now available in a single dataset (European Values Study Group & World Values Survey Association, 2006). This dataset covers 84 country regions, and items involve many aspects of respondents' beliefs, values and attitudes. In addition, the combined dataset contains details on socio-demographics such

as health status, age, scale of income, marital status, sex, number of children, and employment status. The websites of the WVS and the EVS, <http://www.worldvaluessurvey.org> and <http://www.europeanvalues.nl> respectively, give more information.

The first EVS-WVS item we focus upon refers to the attitude toward women, measured by responses to the question:

For each of the following statements I read out, can you tell me how much you agree with each. Do you?: 1 - agree strongly; 2 - agree; 3 - disagree; 4 - disagree strongly
“Both the husband and wife should contribute to household income.”

If this item elicits values, we would expect a positive relation between income and agreement with this statement. Households in which both partners, out of conviction, contribute to income are generally likely to earn more. If the item elicits marginal preferences, we would expect agreement with the statement to fall with income. A high income takes away the necessity and desire for more income, so that people are less inclined to agree that both partners should contribute. Table 1 shows the estimated relation between a respondent's income and agreement with the above statement. It clearly reveals that income is negatively correlated with the view that both the husband and wife should contribute to household income. The correlation between income scale (1-10) and mean values scores is statistically highly significant ($r = 0.96$; $p < 0.001$). Apparently, individuals enjoying higher incomes are less inclined to believe that contribution to household income should be made by both partners.

[Insert Table 1 about here]

To further assess what this item is measuring, we aggregate the data to the country level. Combining the resulting average country scores with data on gross domestic product (GDP) per capita from The Conference Board & Groningen Growth and Development Centre (2008), allows us to establish a similar pattern between income and attitude towards women

as at the individual level.³ Figure 2 depicts the resulting relation between average view on contributions to household income and national income. Clearly, higher income is again associated with disagreement with the statement that both partners should contribute to household income.

[Insert Figure 2 about here]

These results are in support of Hypothesis 1b (although for reasons of coding, the statistical correlation is actually positive). The item asking whether both husband and wife should contribute to household income seems to elicit marginal preferences for income.

Similar conclusions can be drawn about other items in the WVS. For example, the WVS asks about the importance of pay in a job:

Here are some more aspects of a job that people say are important. Please look at them and tell me which ones you personally think are important in a job? Good pay. 0 "Not mentioned" 1 "Mentioned"

If this item elicits values, one would expect a positive relation between scores on this item and actual pay: people attaching a lot of importance to pay in a job are less likely to settle for a low-paying job. If the item elicits marginal preferences primarily, one would expect high-wage earners to attach less importance to pay relative to other objectives, because their desire for material rewards has been satisfied already.

[Insert Table 2 about here]

Table 2 gives the results for this item. We find a statistically highly significant negative correlation between income (1-10) and the importance attached to good pay in a job ($r = -0.80$; $p < 0.01$). Apparently, people enjoying high incomes care less about pay than poorer people do. Figure 3 shows that the same applies to the relation at the societal level: individuals in richer countries care significantly less about pay than those in poorer countries

³ We refer to the joint website of The Conference Board and Groningen Growth and Development Centre website, <http://www.conference-board.org/economics/database.cfm>, for details about this dataset and a downloadable Excel file.

do. This confirms our earlier result, further suggesting that the WVS elicits marginal preferences rather than values (and in this case the statistical correlation is also negative).

[Insert Figure 3 about here]

3.2 Hofstede

We now move towards another popular survey-based measure of cultural values, namely the Hofstede indices for cultural differences. Hofstede's indices are based on answers given in IBM's international employee attitude survey program, for which between 1967 and 1972 approximately 116,000 questionnaires have been obtained from 72 countries (Hofstede, 2001). Through data-reduction techniques, Hofstede was able to identify four values dimensions: Uncertainty Avoidance (UA), Individualism-Collectivism (IND), Power Distance (PD), and Masculinity-Femininity (MAS). A replication of the survey in China resulted in the same four dimensions, adding a fifth called Long-Term Orientation (LTO; Chinese Culture Connection, 1987). In contrast to the World Values Survey, for the Hofstede indices only country aggregate scores on these overarching value dimensions are available; individual responses to single items cannot be retrieved. This implies that we can test Hypotheses 1a and 1b on the level of societies only. Again the approach is to relate scores on values dimensions to scores on corresponding practices dimensions. Hofstede himself, however, has not studied practices. Hence, for the scores on the practices dimensions we must turn to other sources. We rely on the GLOBE study, whose dimensions partly overlap with those identified by Hofstede. House et al. (2004: 138-141) establish six cases in which GLOBE's practice measures match Hofstede's value dimensions (see also Leung, Bhagat, Buchan, Erez & Gibson, 2005: 366): (1) Uncertainty Avoidance & UA; (2) Power Distance & PD; (3) Assertiveness Orientation & MAS; (4) Gender Egalitarianism & MAS; (5) Institutional Collectivism & IND; and (6) In-Group Collectivism dimension & IND.

House et al. (2004: 140) have actually examined the correlation between values and practices for all these combinations. Their findings are as follows. In two cases (1 and 6) the statistical evidence supports hypothesis 1b, implying that Hofstede's cultural framework tends to elicit marginal preferences rather than values. There is a statistically significant negative correlation between Hofstede's UA values and GLOBE's UA practices ($r = -0.61$; $n = 41$; $p < 0.01$) and between Hofstede's scores on Individualism-Collectivism and GLOBE's In-Group Collectivism practices ($r = -0.82$; $n = 41$; $p < 0.01$). For two further values-practices

combinations (2 and 3) the correlations found are consistent with hypothesis 1a. Hofstede's PD index correlates positively and statistically significantly with GLOBE's PD practices ($r = 0.61$; $n = 40$; $p < 0.01$) as does Hofstede's Masculinity-Femininity index and GLOBE's Assertiveness scale ($r = 0.42$; $n = 41$; $p < 0.01$). Finally, in the remaining two matches (4 and 5), Gender Egalitarianism & Masculinity and Institutional Collectivism & Individualism-Collectivism, no significant statistical correlation between GLOBE's practice measures and Hofstede's value dimensions was found.

In order to go beyond value dimensions and say anything about less broadly constructed cultural indicators in Hofstede's survey, we have to rely on occasional information from Hofstede's own research. One of the few value indicators for which he provides data at a less-aggregated level concerns "work goal importance" (Hofstede, 2001: 56-58). This measure is compiled from questions asking respondents, for example, how important the opportunity for high earnings or to have good fringe benefits is in their job (Hofstede, 2001: 467-468). As with the closely related WVS item about the importance of good pay in a job, we would expect this indicator to correlate positively with income (or other levels of consumption) if it measures values, but negatively in so far as it elicits marginal preferences. Unfortunately, Hofstede (2001) does not relate "work goal importance" to income or levels of consumption but he does report its relation to average level of education (on the basis of 38 occupational groups). Since, on average, one's level of education is related to various aspects of one's job, not least having good fringe benefits or a high income we can further test hypotheses 1a and 1b. Hofstede (2001: 57) reports a strong correlation ($r = 0.83$; $n = 38$; $p < 0.001$) between work goal importance and the mean formal level of education in years (see Figure 4), which he attributes to acquiescence ("yes-man-ship"). In particular, he argues that it is acquiescence that leads people with lower levels of education (and which are in a certain occupation) to rank work goals more often as of utmost importance (score 1) or very important (score 2). Under the reasonable assumption that level of education relates positively to important features of one's job, the negative association between work goal importance and education Hofstede is also consistent with hypothesis 1b, however. The work-goal importance items could simply be eliciting marginal preferences rather than values—just like the similar WVS item on importance of pay in a job. Acquiescence may thus play a role, but need not be invoked to account for the finding that work goals such as the opportunity for high earnings or to have good fringe benefits are more important for those with lower levels of education.

[Insert Figure 4 about here]

All in all, the results on Hofstede's data suggest that the Hofstede's framework does not unambiguously measure values. In some cases, they clearly reflect marginal preferences. In general, the picture is simply unclear, suggesting a mix of values and marginal preferences within a single dimension. Hofstede's survey perhaps fares better than GLOBE or the WVS, but it is not immune for the marginal preferences problem.

3.3 Discussion: the validity of values surveys

The results of our analysis testify to the pervasiveness of the marginal preferences-problem in values measured using multi-country surveys. Maseland and van Hoon's (2008) marginal-preferences critique of values surveys appears not to be a rare incident or measurement artifact of the GLOBE study. On the contrary, our results show that the marginal-preferences problem is widespread, and extends beyond the values measured by the GLOBE project (House et al. 2004; Javidan et al. 2006). That is not to say that all values surveys elicit only marginal preferences; but marginal preferences do seem to at least play a role in a diverse set of items and dimensions. What is more, the sheer prevalence of the problem implies that the validity of value indicators cannot be assessed on basis of correlations with values scores from other surveys (convergent or discriminant validity), as is commonly done in the literature. Whereas the face validity of various value items may be high, all are likely to succumb to the marginal-preferences problem, leaving the need to establish what is being elicited by any of the survey questions. The development of new cultural indicators, not suffering the marginal-preferences problem, would be useful. We discuss and compare alternative methods for measuring value preferences and cultural differences in the next section.

4. Measurement alternatives

Alternative approaches to measuring value indicators to be used in cross-cultural comparative research generally need to meet two requirements: (1) measure values as they are commonly understood (and not, for example, marginal preferences); and (2) retain the important advantages of survey-based measures—ease of measurement and pervasiveness being among

the foremost—as much as possible. This latter requirement rules out cross-cultural experiments (Camerer & Fehr, 2004; Leung et al., 2005; Boyd, 2008).⁴

The most candid way of going about the task of meeting these two requirements is to stick to the survey-based approach. New questionnaires consisting of items better able to measure values can be developed and administered—Maseland and van Hoorn (2008) present a format for such questions. The crucial element in their proposal is that people are asked for their general attitudes towards ideal situations, so as to induce respondents not to take in their present context when answering questions. We have identified one item in the WVS that meets this requirement rather well, and, like one of the items we have discussed already, attempts to elicit respondents' attitudes towards gender egalitarianism in marriage⁵. This allows us to investigate whether this improved formulation helps in ruling out marginal preferences. The item is formulated as:

Here is a list of things which some people think make for a successful marriage. Please tell me, for each one, whether you think it is very important, rather important or not very important for a successful marriage? An adequate income 1 “Very” - 2 “Rather” - 3 “Not very”

In spite of the supposedly superior formulation of this item, our analysis indicates that this does not solve the problem. If we relate respondents' scores to income (Table 3), we again find a negative relation between the income people earn and the importance they attach to income ($r = 0.978$; $p < 0.001$). The same applies at the nation level, where higher per-capita income implies lower importance (Figure 5).

⁴ A further limitation of using experimental measures of cultural differences is the potential mixing up of the explanans (the thing that explains) with the explanandum (the thing that requires explaining). In general, we would like to account for and even predict cross-cultural variation in experimental behavior on the basis of clear differences in culture. Reversing this order of analysis operationalises culture as a residual factor, i.e. as observed differences in experimental behavior that cannot be otherwise explained.

⁵ The 1994 and 2008 Values Survey Modules developed by Geert Hofstede and collaborators (see <http://www.geerthofstede.nl>) explicitly ask respondents to disregard their present circumstances when answering the module's questions, thus also coming close to this superior format.

[Insert Table 3 about here]

It appears that the tendency to elicit marginal preferences rather than values is robust for improvements in the formulation of the question. This suggests that improvements may do some good, but cannot eliminate the marginal preferences problem. We have to look for alternative methods.

[Insert Figure 5 about here]

4.1 Cross-cultural differences in the structure of happiness

An alternative measurement method that, in principle, meets both the two general criteria and appears unlikely to succumb to the marginal-preferences problems exists in the form of so-called “happiness functions”. Self-reported happiness has been studied by psychologists for some five decades and is receiving increasing attention from economists. It can be seen as an empirical proxy for Kahneman’s (1999; Kahneman, Wakker & Sarin, 1997) concept of experienced utility (cf. Alesina, Di Tella & MacCulloch, 2004; Di Tella, MacCulloch & Oswald, 2001; Kahneman & Krueger, 2006).⁶ In psychology, happiness is less colloquially referred to as subjective well-being (SWB), where a straightforward definition of the latter is “a broad category of phenomena that includes people’s emotional responses, domain satisfactions, and global judgments of life satisfaction” (Diener, Suh, Lucas & Smith, 1999: 277). Typical measures of SWB simply ask respondents: “*All things considered, how satisfied are you with your life as a whole these days?: 1, Dissatisfied – 10, Satisfied*” or “*Taking all things together, would you say you are: 1, Very happy; 2, Quite happy; 3, Not very happy; 4, Not at all happy*”. These are the basic questions, asked by, for example, the World Values Survey (European Values Study Group & World Values Survey Association, 2006). Items can further be formulated to target certain domains, for example, job satisfaction or financial satisfaction. In addition, more elaborate indicators that use multiple items to

⁶ Kahneman defines experienced utility as “the hedonic quality” of an outcome and contrasts it with “decision utility”, defined as the “weight of an outcome in a decision” and the traditional utility concept in economics (Kahneman et al., 1997: 375).

measure the same underlying construct exist (e.g. Kahneman & Krueger, 2006) and generally these are more reliable and valid (Krueger & Schkade, 2008).

Economists have used happiness indicators, amongst others, to measure the costs of environmental externalities (van Praag & Baarsma, 2005), to assess the well-being effects of policy actions such as the implementation of excise taxes (Gruber & Mullainathan, 2005), and to test and add insights to existing economic theories (Di Tella & MacCulloch, 2006; Di Tella et al., 2001; cf. Frey & Stutzer, 2002). Most happiness research in economics, however, focuses on the economic factors underlying variation in levels of SWB. The typical approach of such studies is to regress happiness scores, sometimes aggregated at the country level, on a broad range of possible determinants. The happiness function thus estimated reveals something about the structure of happiness, the factors and conditions of life that add or detract from SWB, and how much people value certain aspects of their life in terms of these aspects' effect on happiness. Di Tella et al.'s (2001) paper on the happiness effects of inflation and aggregate unemployment nicely illustrates this. Controlling for individual factors such as employment status, income, gender, and marital status, known to affect happiness, this study finds that people are bothered both by inflation and by aggregate unemployment, as is assumed by standard models in monetary economics. Specifically, a rise in the unemployment rate of one percentage point lowers reported satisfaction with life by 0.0233 on a 1-4 scale, while a one percentage-point increase in inflation lowers average life satisfaction by 0.014 points.⁷ This compares with a negative SWB effect of -0.33 for individual joblessness (relative to being full-time employed) and -0.18 for being divorced (relative to being single).

To make happiness functions thus estimated a useful tool for measuring value differences across groups and societies subsequently requires only a small step. Building on the basic idea that data on self-reported happiness or experienced utility provides a mean to assess directly what people care about and how much, the challenge is simply to allow for systematic heterogeneity in the structure of happiness. To the extent that the coefficients for a range of determinants of happiness vary across groups, they provide a clear indicator of value

⁷ Frey and Stutzer (2002) provide an overview of key economic determinants of self-reported happiness. Diener et al. (1999) and Lyubomirsky, King and Diener (2005) survey other significant causes and correlates of SWB. Happiness data is scarcely used in international business research, but see Hui, Au and Fock (2004) and Smith Speck and Roy (2008) for notable exceptions.

differences between these groups. Examples of research that have used heterogeneous happiness functions to assess differences in value preferences include Clark (2003), Alesina et al. (2004), and Lelkes (2006).⁸

Clark (2003) uses data on self-reported happiness to analyze the psychic costs of unemployment. He reports that unemployment lowers happiness, but that the size of the effect depends on social norms. In particular, unemployment lowers an individual's happiness to a much smaller extent if one's peers are unemployed as well and the stigma the unemployed suffer is weaker.

Alesina et al. (2004) examine the effect of inequality on happiness in both Europe and the United States. Their results show that higher inequality is generally associated with lower happiness (with several other factors such as income controlled for), but that there are differences in the way inequality affects happiness. Most notably inequality has a larger effect on happiness in Europe than in the U.S. Furthermore, in Europe poorer individuals and left-wingers are hurt more by inequality whereas in the U.S. inequality hurts the rich. And at the same time, both left-wing and poorer Europeans care more about inequality than left-wing and poorer Americans.

Lelkes (2006) relates heterogeneity in the structure of happiness to religiosity. Her findings show that the effect of economic variables including income on happiness is smaller among the religious than among the non-religious. Apparently, religious people value income less than non-believers do.

4.2 Discussion and possible objections

Whilst it seems possible to use happiness functions for constructing useful substitutes for survey-based measures of value differences, this approach of course is not without limitations either. Firstly, and most importantly, heterogeneity in the structure of happiness captures only part of the cultural differences between groups and societies that matter for behavior and

⁸ Di Tella and MacCulloch (2006: 39-43) review a few applications of heterogeneous happiness functions in economics. Some work, mainly in psychology, has more explicitly assessed cross-cultural variation in the predictors of SWB (e.g. Oishi, Diener, Lucas & Suh, 1999; Diener, Gohm, Suh & Oishi, 2000; Hui et al., 2004; Gelade, Dobson & Auer, 2008). In particular, happiness and (domain) satisfaction are structured differently in individualistic countries than in countries deemed collectivistic.

practices. However, it captures exactly that part for which surveys are not appropriate: values. In measuring aspects like trust, beliefs, practices (as in the GLOBE project), or norms, which are also part of culture, marginal preferences are not a problem and surveys can still be used.

A second line of critique applies more to the method and idea of using happiness data and estimating culture-specific happiness function more generally. For example, there is widespread, and to some extent rightful, doubt whether measures of SWB are meaningful in the sense that they convey important and useful information on individuals' well-being. In formal terms, extensive checks show that happiness data have a reasonable degree of reliability (Krueger & Schkade, 2008). Concerning the validity of SWB data, a large literature has assessed the relation between measures of happiness and related constructs purporting to measure similar or opposing constructs. This work shows that measured happiness correlates with, for example, the frequency of genuine, so-called Duchenne, smiling, SWB reports by friends, family and spouses, and the ability to recollect pleasant memories. Furthermore, happiness scores have some neurophysiologic correlates such as patterns of brain activation and cardiovascular activity. Di Tella and MacCulloch (2006), Diener et al. (1999), Frey and Stutzer (2002), and Kahneman and Krueger (2006) discuss these and other validity assessments in somewhat more detail.

Beyond appraisals of convergent and/or discriminant validity, self-reported happiness is also found to predict outcomes in several domains of life. Higher levels of happiness appear to foster longevity, diminish suicide risk and general health risk, and increase the chance of surviving coronary heart disease and other conditions, amongst others (Lyubomirsky et al., 2005). All in all, most of the work judging the reliability and validity of happiness data reports that these measures are indeed meaningful and carry useful and important information.

Another potential problem in the use of differences in the structure of happiness to construct value measures is that cultural groups need to be pre-defined. When estimating the effects of unemployment on happiness in Germany and the United States, for example, it is not *a priori* clear that the group of unemployed people shares the same values as the rest of society. On the other hand, it is also possible to let the data speak for itself. Important value differences can be reflected in heterogeneous happiness functions, for which the relevant group is defined endogenously. Clark, Etilé, Postel-Vinay, Senik and Van der Straeten (2005) take this approach, classifying individuals on the basis of the way in which they transform income into financial satisfaction. Using data from the European Community Household Panel and applying a latent class technique, they establish that individuals can be classified in

four groups on the basis of the effect of income on satisfaction with one's financial situation. Furthermore, the probability that a respondent belongs to a particular group is found to be a function of individual characteristics, such as education and age, and country of residence as well.

Finally, a limitation of the happiness approach is that it requires a minimum amount of individual observations per culture. Practically, if a group, for whatever reason, contains too few people it is not possible to estimate a happiness function, the coefficients of which can be compared with those of other groups. That being said, the wide coverage of happiness data implies that it can be used almost as extensively as values surveys instruments. The WVS appears to offer a rich source, as does the Gallup organization's "World Poll" conducted annually since 2005 and covering over 140 countries (see <http://www.gallup.com>), although results from the latter are not publicly available.

5. Summarizing thoughts

On the face of it, asking an individual how much he or she values, work relative to leisure, or freedom relative to equality, tells us something about the person's values. In practice, however, questionnaire items like these tend to elicit marginal preferences rather than values. Scores reflect the degree to which a respondent has already satisfied an objective rather than the underlying taste for this objective. Respondents are more likely to stress uncertainty avoidance as something important in a world full of uncertainty, than in a society in which every move is tightly regulated. This tells us nothing about respondents' underlying values; it is merely a reflection of how society appears to them.

In this paper, we have assessed to what extent this marginal-preferences problem affects various values surveys instruments often used in research. Scrutinizing data taken from the World Values Survey and Hofstede's cultural indices, we show that what they claim are values, generally correspond negatively with practices. This indicates that indeed values surveys tend to pick up marginal preferences rather than values.

Outside the survey-based approach to measuring variation in culture, so-called happiness functions appear a viable method of measuring values and how they differ between groups and societies. Psychologists have been measuring happiness for several decades, and researchers (economists, psychologists) increasingly use data on self-reported happiness to assess value preferences more directly. Exactly how much people care about certain aspects of their life, either at the individual or at the macro level, can be ascertained by simply

looking at the impact the different factors have on individuals' level of happiness. By allowing for group heterogeneity in happiness functions thus estimated, it is possible to uncover systematic variation in the value attached to a wide range of factors. Pioneering work that has taken this approach to deal with cultural variation, finds, amongst others, that Europeans are hurt more by inequality than are U.S. citizens, and that religious individuals care less about income than non-religious individuals do. Although happiness functions have not been much applied, and the method is not without limitations, the method can be used to construct substitutes for survey-based value indicators.

Overall, cross-cultural researchers seem to have some work cut out for them.

References

- Alesina, A., Di Tella, R., & MacCulloch, R. 2004. Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics*, 88 (9-10): 2009-2042.
- Boyd, R. 2008. cross-cultural experiments. In: Durlauf, S. N., & Blume, L. E., Eds. *The New Palgrave Dictionary of Economics*. Second edition. The New Palgrave Dictionary of Economics Online. Palgrave Macmillan.
- Camerer, C. F., & Fehr, E. 2004. Measuring social norms and preferences using experimental games: A guide for social scientists. In: Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H., eds. *Foundations of Human Sociality. Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press: 55-95.
- Chinese Culture Connection. 1987. Chinese Values and the Search for Culture-Free Dimensions of Culture. *Journal of Cross-Cultural Psychology*, 18 (2): 143-164.
- Clark, A. E. 2003. Unemployment as a social norm: Psychological evidence from panel data. *Journal of Labor Economics*, 21 (2): 323-351.
- Clark, A. E., Etilé, F., Postel-Vinay, F., Senik, C., & Van der Straeten, K. 2005. Heterogeneity in Reported Well-Being: Evidence from Twelve European Countries. *Economic Journal*, 115 (502): C118-C132.
- Diener, E., Gohm, C. L., Suh, E., & Oishi, S. 2000. Similarity of the Relations between Marital Status and Subjective Well-Being Across Cultures. *Journal of Cross-Cultural Psychology*, 31 (4): 419-436.

- Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. 1999. Subjective Well-Being: Three Decades of Progress. *Psychological Bulletin*, 125 (2): 276-302.
- Di Tella, R., & MacCulloch, R. 2006. Some Uses of Happiness Data. *Journal of Economic Perspectives*, 20 (1): 25-46.
- Di Tella, R., MacCulloch, R., & Oswald, A. 2001. Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, 91 (1): 335-341.
- European Values Study Group & World Values Survey Association. 2006. *European and World Values Surveys four-wave integrated data file, 1981-2004*. Available from www.jdsurvey.net.
- Frey, B. S., & Stutzer, A. 2002. What Can Economists Learn from Happiness Research? *Journal of Economic Literature*, 40 (2): 402-435.
- Gelade, G. A., Dobson, P., & Auer, K. 2008. Individualism, Masculinity, and the Sources of Organizational Commitment. *Journal of Cross-Cultural Psychology*, 39 (5): 599-617.
- Gruber, J. H., & Mullainathan, S. 2005. Do Cigarette Taxes Make Smokers Happier. *Advances in Economic Analysis & Policy*, 5 (1): Article 4.
- Hofstede, G. 2001. *Culture's Consequences* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Hofstede, G. 2006. What did GLOBE really measure? Researchers' minds versus respondents' minds. *Journal of International Business Studies*, 37 (6): 882-896.
- Horowitz, J., List, J., & McConnell, K. E. 2007. A Test of Diminishing Marginal Value. *Economica*, 74 (November): 650-663.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. 2004. *Culture, Leadership and Organizations: The Globe Study of 62 Societies*. Thousand Oaks, CA: Sage Publications.
- Hui, M. K., Au, K., & Fock, H. 2004. Empowerment effects across cultures. *Journal of International Business Studies*, 35 (1): 46-60.
- Inglehart, R. 1990. *Culture Shift in Advanced Industrial Society*. Princeton, NJ: Princeton University Press.
- Javidan, M., House, R. J., Dorfman, P. W., Hanges, P. J., & Sully de Luque, M. 2006. Conceptualizing and measuring cultures and their consequences: a comparative review of GLOBE's and Hofstede's approaches. *Journal of International Business Studies*, 37 (6): 897-914.

- Kahneman, D. 1999. Objective Happiness. In: Kahneman, D., Diener, E., & Schwarz, N., eds. *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation: 3-25.
- Kahneman, D., & Krueger, A. B. 2006. Developments in the Measurement of Subjective Well-Being. *Journal of Economic Perspectives*, 20 (1): 3-24.
- Kahneman, D., Wakker, P., & Sarin, R. 1997. Back to Bentham? Explorations of Experienced Utility. *Quarterly Journal of Economics*. 112 (2): 375-405.
- Krueger, A. B., & Schkade, D. A. 2008. The reliability of subjective well-being measures. *Journal of Public Economics*, 92 (8-9): 1833-1845.
- Lelkes, O. 2006. Tasting freedom: Happiness, religion and economic transition. *Journal of Economic Behavior & Organization*, 59 (2): 173-194.
- Leung, K., Bhagat, R. S., Buchan, N. R., Erez, M., & Gibson, C. B. 2005. Culture and international business: recent advances and their implications for future research. *Journal of International Business Studies*, 36 (4): 357-378.
- Lyubomirsky, S., King, L., & Diener, E. 2005. The Benefits of Frequent Positive Affect: Does Happiness Lead to Success? *Psychological Bulletin*, 131 (6): 803-855.
- Marshall, A. 1920 [1890]. *Principles of Economics: An Introductory Volume*. (8th ed.). The Macmillan Press: London and Basingstoke.
- Maseland, R., & Hoorn, A. van. 2008. Explaining the negative correlation between values and practices: A note on the Hofstede-GLOBE debate. *Journal of International Business Studies*, advance online publication September 25. doi:10.1057/jibs.2008.68.
- Oishi, S., Diener, E., Lucas, R. E., & Suh, E. M. 1999. Cross-Cultural Variations in Predictors of Life Satisfaction: Perspectives From Needs and Values. *Personality and Social Psychology Bulletin*, 25 (8): 980-990.
- Praag, B. M. S. van, & Baarsma, B. 2005. Using Happiness Surveys to Value Intangibles: The Case of Airport Noise. *Economic Journal*, 115 (500): 224-246.
- Rokeach, M. 1973. *The Nature of Human Values*. New York: The Free Press.
- Smith Speck, S. K. & Roy, A. 2008. The interrelationships between television viewing, values and perceived well-being: A global perspective. *Journal of International Business Studies*, 39 (7): 1197-1219.
- The Conference Board & Groningen Growth and Development Centre. 2008. *Total Economy Database*. January 2008.

Table 1. Husband and wife should both contribute to income.

Scale of incomes	Mean	N	Distribution of responses by answer category [%]			
			1 Agree strongly	2 Agree	3 Disagree	4 Disagree strongly
Lower step	1.75 0.77	18,215	42.5	42.8	12.2	2.6
Second step	1.79 0.76	24,732	38.8	45.9	12.9	2.4
Third step	1.81 0.75	27,850	37.1	47.2	13.4	2.4
Fourth step	1.81 0.75	28,485	37.5	46.6	13.7	2.3
Fifth step	1.81 0.76	26,076	38.0	45.3	14.3	2.4
Sixth step	1.83 0.76	19,990	36.6	45.9	15.2	2.4
Seventh step	1.87 0.78	15,572	35.2	45.7	16.4	2.7
Eighth step	1.86 0.77	11,633	35.5	45.2	16.9	2.3
Ninth step	1.91 0.80	7,687	33.5	44.4	19.1	3.0
Tenth step	1.95 0.81	7,306	32.0	44.3	20.6	3.1

Notes: Question asked in waves 2, 3 and 4 of both the WVS and the EVS. Standard deviation in parentheses. “Don’t know” answers and respondents with otherwise missing responses are dropped. The sample covers 81 countries. The chi-square test rejects the null-hypothesis that the distribution of the responses across the four answer categories are the same for all 10 income scales ($p < 0.001$). Syntax for all the empirical analyses we conducted is available on request.

Table 2. Income scale and the importance of pay in a job.

Scale of incomes	Mean	N
Lower step	0.80 0.40	19,573
Second step	0.81 0.39	26,872
Third step	0.82 0.38	30,580
Fourth step	0.82 0.39	31,081
Fifth step	0.82 0.39	28,633
Sixth step	0.80 0.40	23,038
Seventh step	0.80 0.40	18,721
Eighth step	0.79 0.41	14,505
Ninth step	0.78 0.42	9,621
Tenth step	0.76 0.43	8,837

Notes: See Table 1. Question asked in all waves. The sample covers 80 countries. The chi-square test rejects the null-hypothesis that the distribution of the responses across the two answer categories are the same for all 10 income scales ($p < 0.001$).

Table 3. Importance of adequate income for successful marriage by income scale.

Scale of incomes	Mean	N	Distribution of responses by answer category [%]		
			1 Very	2 Rather	3 Not very
Lower step	1.67 0.69	8,137	46.2	40.7	13.1
Second step	1.67 0.67	10,570	44.0	44.8	11.1
Third step	1.69 0.67	13,301	42.5	45.8	11.7
Fourth step	1.69 0.67	13,628	42.4	45.8	11.9
Fifth step	1.70 0.67	13,320	41.4	46.8	11.8
Sixth step	1.76 0.67	11,047	38.0	48.6	13.5
Seventh step	1.77 0.67	9,430	36.5	50.2	13.2
Eighth step	1.78 0.66	6,997	35.8	50.8	13.3
Ninth step	1.81 0.68	5,047	33.9	50.7	15.4
Tenth step	1.83 0.66	4,885	32.0	53.0	15.0

Notes: See Table 1. Question asked in wave 1, 2 and 4 of the EVS and wave 2 of the WVS. The sample covers 46 countries. The chi-square test rejects the null-hypothesis that the distribution of the responses across the four answer categories are the same for all 10 income scales ($p < 0.001$).

Figure 1. A utility function with diminishing marginal utility.

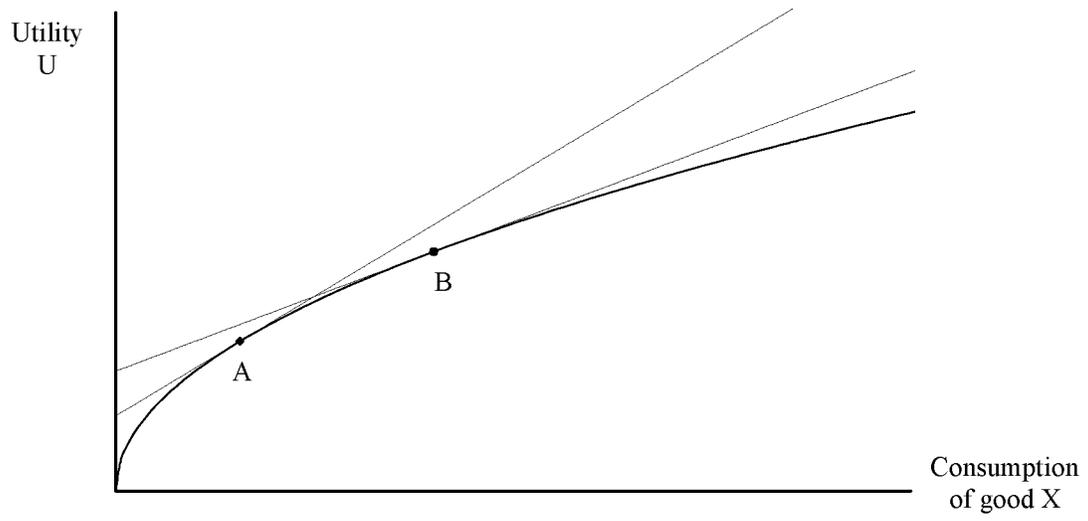
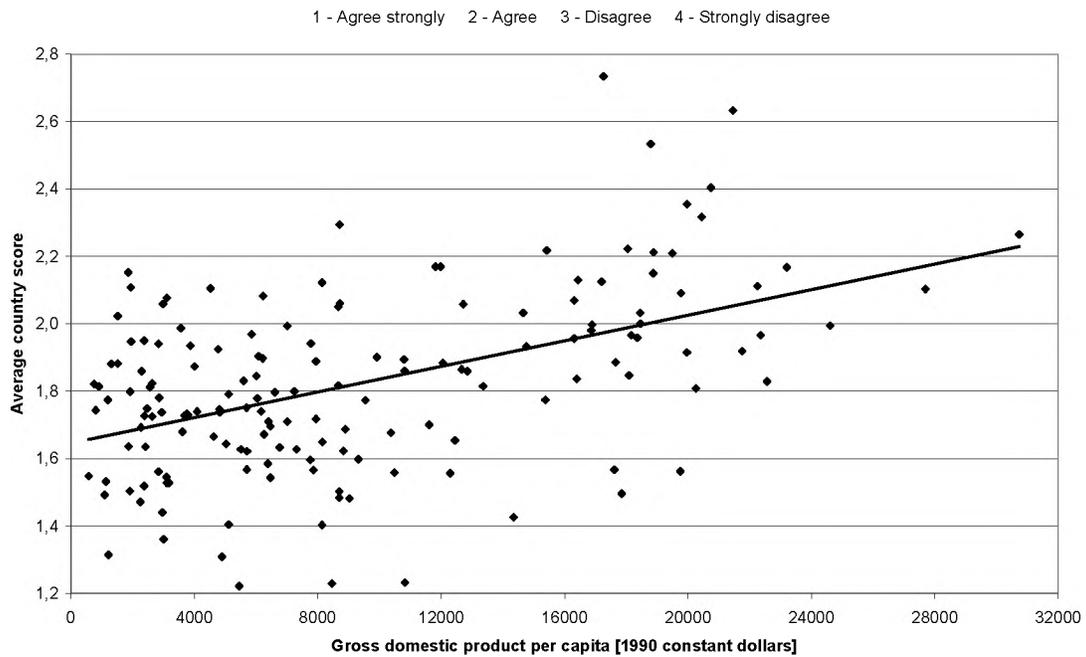
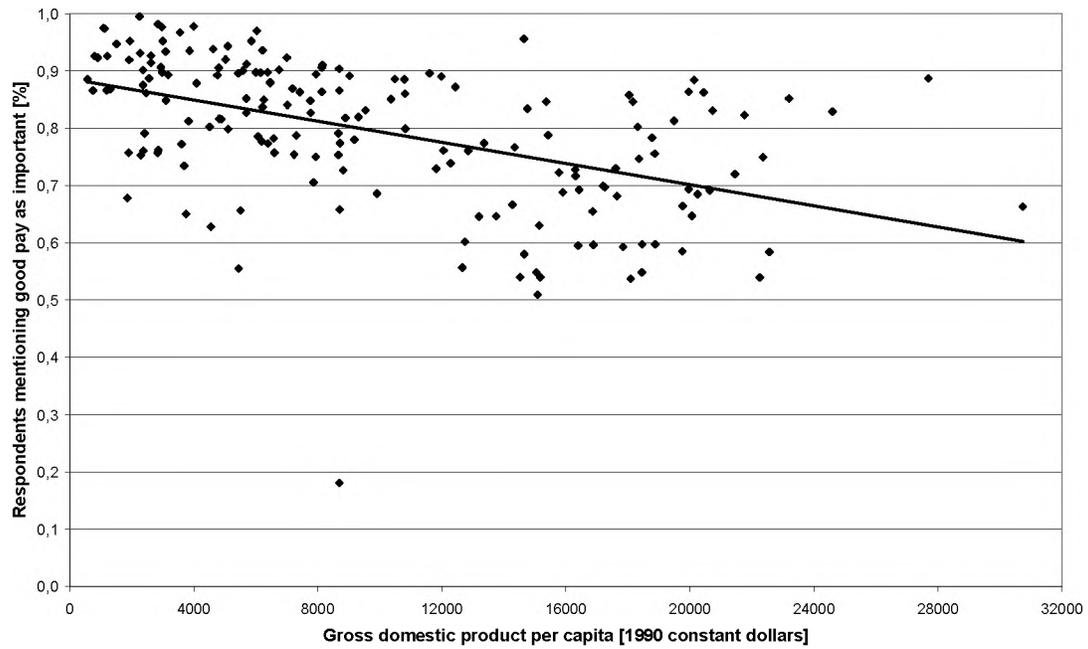


Figure 2. Husband and wife should both contribute to income and per-capita income.



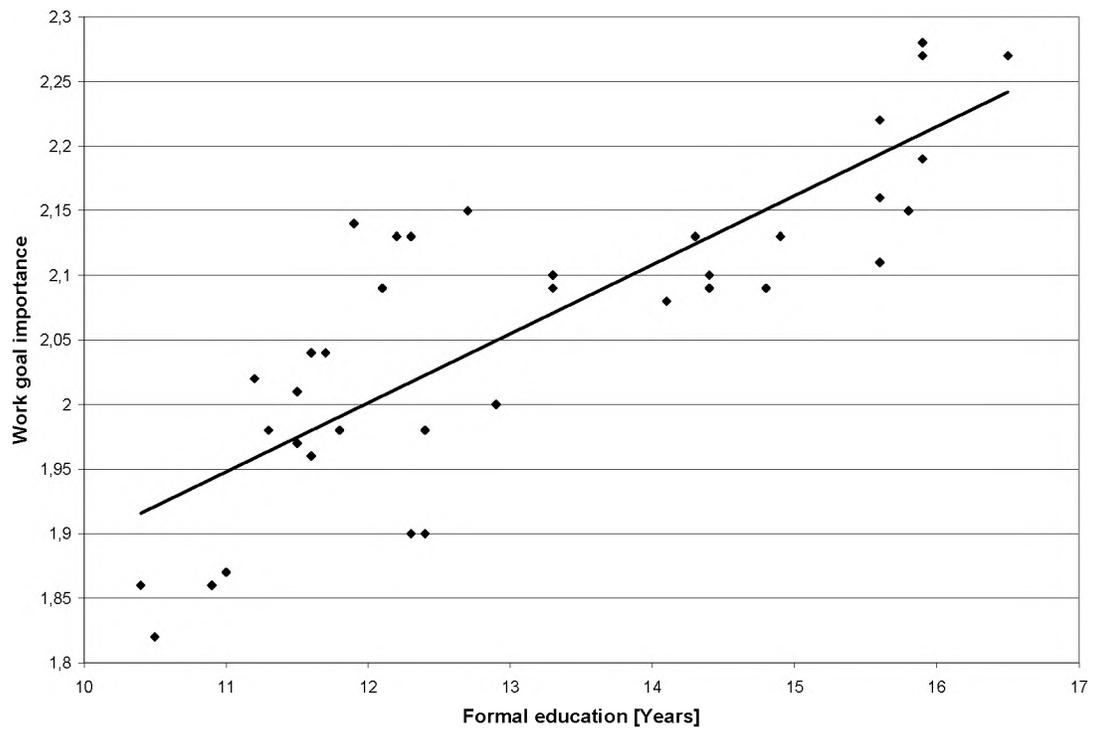
Notes: See Table 1. Aggregated country scores are based on weighted individual answers. $r = 0.50$ ($p < 0.001$). The sample consists of 79 countries and 158 year observations.

Figure 3. Good pay important in a job and per-capita income.



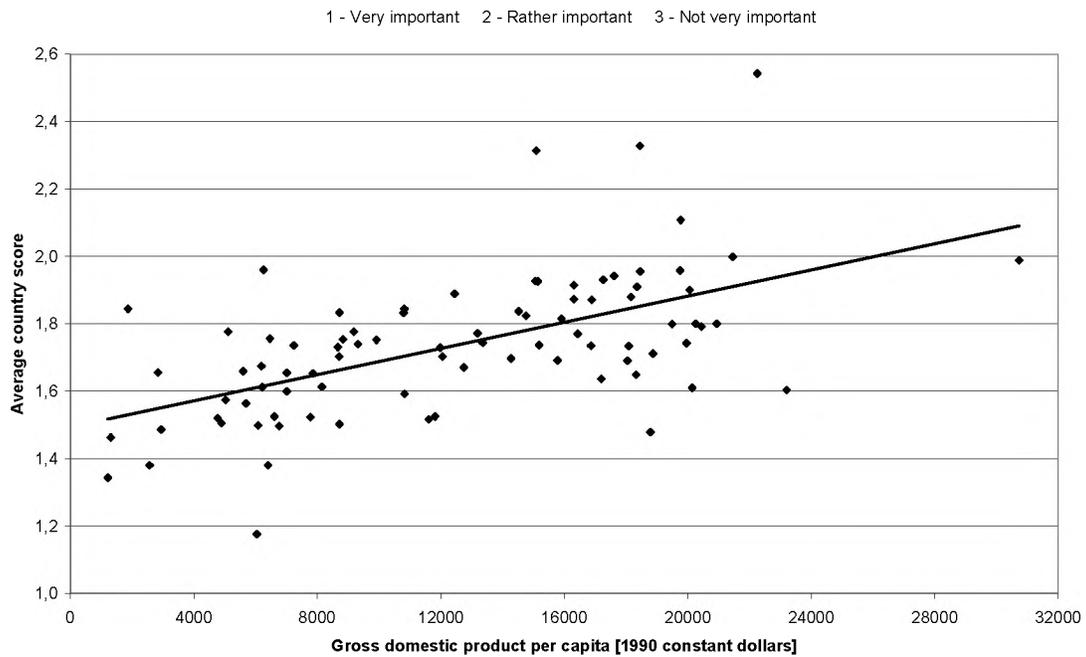
Notes: See Table 2. Aggregated country scores are based on weighted individual answers. $r = -0.49$ ($p < 0.001$). The low outlier (0.18) is South Korea (in 1990). Excluding the outlier gives: $r = -0.53$ ($p < 0.001$). The sample consists of 78 countries and 177 year observations.

Figure 4. Hofstede's work goal importance and formal years of education.



Notes: Data from Hofstede (2001: 57). $r = 0.83$ ($n = 38$; $p < 0.001$).

Figure 5. Importance of adequate income for successful marriage and per-capita income.



Notes: See Table 3. Aggregated country scores are based on weighted individual answers. $r = 0.57$ ($p < 0.001$). The sample consists of 45 countries and 85 year observations.