Volume Title: Measuring the Subjective Well-Being of Nations: National Accounts of Time Use and Well-Being

Volume Author/Editor: Alan B. Krueger, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-45456-8

Volume URL: http://www.nber.org/books/krue08-1

Conference Date: December 7-8, 2007

Publication Date: October 2009

Title: Thoughts on "National Time Accounting: The Currency of Life"

Author: Erik Hurst

URL: http://www.nber.org/chapters/c5060

8

Thoughts on "National Time Accounting: The Currency of Life"

Erik Hurst

8.1 Introduction

In their article "National Time Accounting: The Currency of Life," Krueger et al. (see chapter 1 of this volume) propose an alternate way of computing individual well-being. The foundation of the new measure of well-being is the construction of a U-index (where the "U" stands for "unpleasantness"). The U-index is formed by surveying households about their enjoyment of the activities in which they participated during the prior day. For example, suppose last night the survey respondent had dinner with their spouse. Today, the respondent would be asked to assess the feelings they were experiencing during the previous night's dinner. The measurement of feelings occurs along a variety of dimensions (happiness, sadness, pain, stress levels, etc.). The measure of intensity of the feeling occurs along a 6-point scale (with six being the most intense feeling along the respective dimension). For an individual, an activity is deemed unpleasant if the negative feelings (sadness, pain, stress, etc.) experienced while engaging in the activity are more intense than the positive feelings (happiness) experienced while engaging in the activity. That is, for each individual-specific activity, the U-index is either one (negative emotions dominate) or zero (negative emotions do not dominate). The overall U-index for an activity in the population is simply the average U-index across all people performing the activity in the survey. That is, the average U-index for an activity globally takes on values between 0 and 100 percent. A global activity level U-index of one hundred means that 100 percent of the people engaging in that activity had a U-index of one.

Erik Hurst is the V. Duane Rath Professor of Economics and Neubauer Family Faculty Fellow at the Booth School of Business, University of Chicago, and a research associate of the National Bureau of Economic Research.

The goal of the U-index method is to compute activity level U-indices. With these U-indices in hand, researchers can use existing time diaries (such as the American Time Use Survey [ATUS] conducted by the Bureau of Labor Statistics [BLS]) to compute a measure of the average quality of an individual's day. In essence, they propose computing a measure that converts time units (like minutes for a given activity during a day) into unpleasantness units (using the U-index). The fraction of time people spend in relatively unpleasant activities could then be used as an alternate measure of well-being for individuals. This measure can be tracked over time (at the national level or group-specific levels) to ask whether individuals within a country are spending more time in pleasant activities today than they did at some time in the past. Likewise, the measure can be used to assess whether the well-being of one group (i.e., the lower educated) is converging or diverging from the well-being of another group (i.e., the higher educated).

Overall, I think this research design has merit. I think it would be good to create a time series of the U-index and see whether it adds any additional information in terms of computing trends in well-being aside from our existing traditional sources (wages, time allocation, overall GDP and inflation statistics, other happiness measures, etc.). I think the goal should be to assess whether changes in the U-index provide additional information about well-being above and beyond changes in other readily available series. The only way to know the answer to that question is to develop the U-index and monitor its properties over time. I applaud the authors for starting that process.

My comments will be structured around three points. My first set of comments (in section 8.2) expands upon the themes outlined in the previous paragraph. In that section, I ask what is it that we hope to capture about changing well-being using the U-index that would not already be captured by the changes in existing well-being measures. Also in that section, I talk about other conceptual issues pertaining to the measurement of well-being. In section 8.3, I take a more philosophical turn and ask how the U-index is designed to measure activities where extreme negative emotions are desired (such as many forms of art). In particular, I will ask how to interpret the U-index if people seek out (and are willing to pay) for stress or sadness (such as movies or television shows that market themselves as thrillers or dramas). In section 8.4, I offer a series of comments pertaining to the implementation of the U-index. In doing so, I address many potential selection issues inherent in the construction of the U-index. Some of these can be addressed empirically. All of them, however, need to be thought about before implementing the U-index for larger purposes (such as measuring changing aggregate well-being). The final section concludes.

8.2 Some Thoughts on Motivation

Before getting into specific comments with respect to how the U-index is constructed, I wanted to comment on some bigger issues. First, I want to think about what it is that the U-index is trying to measure. Second, I want to comment on what the U-index is intending to measure that is not measured by more general "happiness" surveys.

8.2.1 Why is the U-Index Necessary?

In Nordhaus' comment on the Krueger et al. chapter (see chapter 5 of this volume), he laid out a model of individual optimization that shows how, under certain assumptions, individual well-being evolves over time. One conclusion from his work is that if individuals have utility over consumption commodities (defined in a Beckerian sense) and also receive some process flow from spending time in a given activity, individual utility evolves at exactly the same rate as market wages, assuming that the productivity used to augment the production of each commodity (including market labor) grows at the same rate. This is seen in Nordhaus's equation 16.

In other words, according to the Nordhaus model, if market productivity and each component of nonmarket productivity grow at the same rate, the *change* in well-being for an individual over time is perfectly measured by the change in their return to working. The return to working is just the wage less any disutility from work. This is exactly analogous to using full income to measure individual well-being (where full income is just the wage multiplied by the time endowment). In these full-income models, there is no disutility to working. As a result, given a constant per period time endowment, the growth in full income is just the growth in the wage. The reason that the growth in the wage is the relevant measure for the change in well-being in these models is that consumers equate the marginal return across different activities. If individuals always have the opportunity to work, the marginal return to any activity should be set equal to the marginal return to working.

Therefore, the relevant question at hand is, when is the growth in the wage rate not the appropriate measure of well-being? The Nordhaus model shows a few instances. First, if there is changing disutility to work, changing well-being will not be appropriately measured by changes in the wage. Likewise, if there are different growth rates in productivity between market and nonmarket sectors, the growth in the wage rate will not perfectly capture the growth in well-being (see Nordhaus equation [16]). Lastly, if individuals are not on their labor supply curve, the wage may not truly represent their marginal value of time. This is not in the Nordhaus model but will fall out of any model of labor supply in which there are frictions in the labor market that cause labor supply to be "lumpy" (i.e., we can only work forty hours per week, twenty hours per week, or nothing).

Moreover, as emphasized in the Krueger et al. chapter, there are other things that individuals care about that are not embedded in models similar to the one put forth by Nordhaus. For example, we may care about national security (safety) or environmental quality. These extranalities are out of an individual's control, yet they enter into their utility function. So, changes in these extranalities will also affect the growth rate of individual well-being.

So, if that is the case, the goal of the U-index is very specific. The U-index (or any other subjective measure of well-being) hopes to capture either (a) changes in nonmarket technologies over time, (b) the changing ability of the wage to measure the marginal valuation of an individual's time (either due to changing disutility of work or because of changing constraints in the labor market), or (c) the changing nature of extranalities over time. These are fine goals to have for developing alternate measures of well-being. It would be valuable to understand how important these omissions are in terms of their effect on changes in well-being relative to changes in the wage. I share Nordhaus' belief that as this project moves forward, it would be useful to highlight what the U-index is intended to capture beyond our standard methods of measuring well-being (such as the full-income method). The authors currently do some of that. I just think it is important for them to continue doing so forcefully.

8.2.2 The U-Index versus Traditional Happiness Measures

This brings me to a related point. Not only do the authors have to convince people that there is value added in measuring changing well-being by a broader measure than just changes in the wage; they also have to convince people that their proposed U-index actually has the potential to add value relative to other existing broader measures of well-being. As noted in the Blanchflower comment on chapter 1 (see chapter 7 of this volume), there are many existing happiness or life satisfaction surveys collected within the United States and many other countries. Most of these surveys ask people some variant of the question, "On a whole, how satisfied are you with the life you lead?" As shown in the Blanchflower comment, the cross-sectional patterns of these existing happiness surveys are very similar in most instances to the cross-sectional patterns of the U-index.

One natural question to ask is, what is the gain of developing the U-index, given that we already have well-developed existing questions on happiness or life satisfaction? The U-index is an innovation on existing well-being measures in two directions. First, it measures life satisfaction at the individual activity level. Second, the U-index measures life satisfaction for a particular day (yesterday), where the short-term emotional memories of that activity are still fresh in respondents' minds.

In terms of measuring the components of an individual's changing well-being not captured by changes in the wage, how valuable is the activity level

data? It depends on the component of well-being that is not being measured. Take, for example, large societal externalities like pollution, fear of terrorism, economic uncertainty, or the quality of our children's play (all of which were emphasized in the Krueger et al. piece as a rationalization for the U-index). The effects of these large societal externalities on well-being are likely not to be activity specific. For example, if I am more uncertain about terrorism, it is not likely to manifest itself only when I fly. I will sometimes be thinking about it when I am eating dinner, working at my job, or watching television. Similar stories can be told about the externalities from clean air or happy children. These types of externalities likely affect an individual's general emotional experience as opposed to activity-specific emotional experiences. If we think that the primary mismeasurement of changing well-being as proxied by the changing wage is that it does not account for large unmeasured societal externalities, it is not certain that activity-based measures of affect are better than general affect measures (like the existing measures of life satisfaction).

The activity-specific measures are likely to be very informative, however, if there are changing technological advances in the production of the experience by activity. For example, if we are truly happier now watching television because the quality of television sets has increased so dramatically (holding price constant), the U-index will likely be able to isolate this activity-specific trend.

In summary, the true innovation of the activity-specific U-indices, in terms of measuring previously unmeasured well-being, is that it can capture activity-specific advances in technology. The externalities can be measured by more general (nonactivity-specific) affect measures. A discussion by the authors of how the U-index could improve upon the unmeasured components of well-being with respect to existing measures of life satisfaction would be very useful. I think that such reasons do exist, so it should be easy for the authors to do.[1]

8.2.3 Relative Preferences and Adaptation

The last comment I wish to address in this section is how we would expect the U-index to evolve over time if people have relative (or adaptive) preferences. There is ample evidence (many by these authors) that convince me that relative well-being enters directly into utility functions. Such preferences explain in part why happiness measures tend not to trend upward over time,

1. One question that I have is whether there is any new information in how the U-index evolves over time relative to traditional happiness measures, but a time series on the U-index will allow us to answer this question. Again, as Blanchflower's comment has already indicated (see chapter 7 of this volume), the cross-sectional patterns in the U-index, for the most part, are very similar to the cross-sectional patterns in traditional happiness surveys. But in terms of changing well-being, we care about the changes in the U-index relative to the changes in traditional happiness surveys.

despite the huge increases in real incomes within an economy. I am not sure whether such preferences matter at all for the construction and measurement of the U-index. However, the existence of such preferences is certainly important for the interpretation of trends in the U-index. I would have liked a little more discussion about what the authors think with respect to the implementation (and the value of the implementation) of the U-index in a world where individuals care about relatives rather than absolutes. In such a world, would the U-index even measure changes in well-being resulting from changes in extranalities like pollution or terrorism? If everyone eventually gets used to the pollution and terrorism, would the negative effects on well-being actually show up in the U-index? I am not sure of the answers; I just thought that it should be addressed somewhere (especially given the previous work of some of the authors).

8.3 The Importance of Television Watching

One important output of the National Time Accounting system developed in chapter 1 was to measure changing well-being for men and women since 1965. In figures 1.9 through 1.11, Krueger et al. use their activity level U-index to show that over the last forty years, women experienced a smaller decline in unhappiness than did men. In this subsection, I discuss the importance of the U-index for television watching for making these conclusions.

In my 2007 *Quarterly Journal of Economics* paper with Mark Aguiar (hence referred to as AH2007), we documented the major trends in time use for men and women within the United States between 1965 and 2003.[2] To do this, we harmonized the five major nationally representative time-use studies conducted in the United States during this time period. The major trends can be summarized as follows. First, for men, total time spent in total market work declined substantially (by over ten hours per week). This number also includes ancillary work activities like commuting to work and taking breaks while at work. Men also increased the time they allocated to nonmarket work (by roughly five hours per week). Leisure time for men (time spent with friends, watching television, exercising, going to the movies, etc.) increased by roughly five hours per week during this time.

For women, there was a slightly smaller increase in leisure time (by about three hours per week). The increase in leisure was facilitated by women dramatically decreasing the time they allocated to nonmarket production (by about ten hours per week), while simultaneously increasing the time they spent in market work. The majority of the decline in nonmarket production was due to a decrease in food preparation and cleanup. Like us, Krueger

2. The paper is referenced in Krueger et al. (see chapter 1 of this volume).

et al. find that the increase in leisure (or decline in unhappiness) was greater for men. Also like us, they find that most of the increase took place prior to 1985.

The relevant question that the U-index can shed light upon is, "How much happier did men and women become over the last forty years?" To interpret the Krueger et al. results, we should note that AH2007 shows that almost the entire increase in leisure can be explained by an increase in television watching. This is similar to the Krueger et al. findings (which should not be surprising, given that they are using the same underlying data as AH2007). Over the last forty years, women have substituted housework, such as food prep, essentially for television watching. According to the U-index results (table 1.8), food prep and television watching are roughly similar in terms of unpleasantness (19.0 versus 18.1). Men, on the other hand, substituted market work for television watching. According to the U-index, this was a huge gain in well-being, given that market work is reported as being much more unpleasant compared to television watching (26.9 versus 18.0).

Reading the results of Krueger et al., I was struck by how individuals report feeling while they are watching television. Television is reported as being one of the more unpleasant leisure activities. In table 8.1, I summarize the time individuals spend on various leisure activities (from the 2003 to 2005 American Time Use Surveys) and the corresponding U-index for that activity as reported in Krueger et al.[3]

As seen in this table, individuals find many other leisure activities to be more enjoyable than watching television. For example, some of the most enjoyable activities, according to the U-index, are listening to music, engaging in sports or exercise, participating in religious activities, and relaxing or general leisure activities. However, households allocate very little time to these activities. Yet, individuals spend an abundance of time watching television, which is on par with washing dishes and cooking in terms of reported U-index.

Does this fact violate individual-revealed preference? It does if we take the U-index seriously. Take, as an example, watching television versus listening to music. In terms of cognitive resources needed to engage in the activity, both are similar. For example, one can just as easily passively watch television as they can passively listen to the radio. Additionally, the necessary start-up costs are probably lower for listening to music. A nice music system is equally as expensive (if not less expensive) than a nice television system. If people like music so much more than watching television (which they do according to the U-index), why are they watching so much television? Why

3. The sample used is similar to the sample used in Aguiar and Hurst (2007). Basically, the sample consists of twenty-one- to sixty-five-year-olds who were nonretired and nonstudents. The only difference is that Aguair and Hurst (2007) only looked at data from the 2003 ATUS, as opposed to aggregating together the ATUS from 2003 to 2005.

Table 8.1 Hours per week spent in activity versus U-index: By leisure activity

Leisure activity	Hours per week (2003)	U-index
Television watching	16.4	18.1
Hobbies	0.2	13.4
Socializing	7.0	13.5
Sports and exercise	2.2	7.4
Religion	2.0	6.4
Listening to music	0.2	0.0

Notes: This table shows the hours per week spent in the activity according to the ATUS from 2003 to 2005. The sample for the ATUS is the same used in Aguiar and Hurst 2007, which is basically all individuals between the ages of twenty-one and sixty-five who were nonstudents and who were nonretired. The U-index numbers came from the Krueger et al. chapter.

do people not shut off the television and turn on the radio? The failure of people to do so implies one of two things: either individuals are persistently irrational (and keep watching television despite their relatively low enjoyment as compared to listening to music), or the U-index is not capturing what it intends to capture.

Before addressing that latter question, I want to address one other question first. Particularly, does it matter how we view television watching for understanding changing well-being over time? Because of the fact that increased television watching has been one of the most dominant trends in how we allocate our time over the last forty years, we would expect the classification of how we view television watching to be critical to assessing changing well-being over time. Table 8.2 confirms this fact. In table 8.2, I first restate the change in the U-index for men and women between 1965 and 2003 as in Krueger et al. (figure 1.9). These are found in row 1 of panels A (men) and B (women). In row 2 of each panel, I assign television watching the same U-index as listening to music.[4] This latter assumption is extreme, but it serves an illustrative purpose. It says that if people choose to watch television, they must like television watching more than listening to music (at least on average). If we implicitly assume that the listening-to-music U-index is correct, the U-index for television watching has got to be at least as low as the U-index for listening to music. The results in row 2 of Table 8.2 show that the decline in the U-index over the last forty years is much greater for both men and women if we change the evaluation of television watching in a way that would be consistent with revealed preference. In other words, a simple change in the U-index to make the measure consistent with revealed preference only for television watching would dra-

4. I am indebted to Alan Krueger for running these hypotheticals using the actual data that underlies figure 1.9. I did nothing more than ask about the hypothetical; all the work was done by Alan.

Table 8.2 **Sensitivity of Krueger et al. change in U-index over time to the treatment of television watching**

	U-index (1965)	U-index (2003)	Difference (%)
	Panel A: men		
Men (original)	20.9	19.6	–6.2
Men (adjusted)	18.8	16.6	–11.7
	Panel B: women		
Women (original)	19.4	19.2	–1.0
Women (adjusted)	17.9	16.7	–6.2

Notes: This table explores the sensitivity of the change in the U-index over time for men and women to the U-index attributed to television watching. Row 1 of panels A and B show the original time trend in the U-index reported in Krueger et al., figure 1.9. Row 2 of panels A and B show the recomputed U-index in both years, assigning television watching the same U-index as listening to music. Given that listening to music is reported as being much more pleasant than watching television, and given the fact that television watching increased dramatically over this time period, the adjusted series shows much greater declines in the U-index than the original series.

matically change the conclusions about changing well-being over the last forty years.[5]

Such a discussion brings me to the more substantive (and philosophic) question of why people watch television and how the U-index would respond to those reasons. The premise of the U-index is to measure the intensity of positive and negative emotions during an experience. If the negative emotions were more pronounced than the positive emotions, the U-index takes a value of one for the individual during the activity (zero otherwise). The average U-index for an activity is the average of individual U-indices for individuals participating in the activity. The positive and negative emotions measured when computing the U-index include whether the individual felt happy, sad, pain, and stressed.

Often the goal of participating in various forms of art (such as movies, television, or music) is to experience extreme human emotions (positive or negative). Think of recent Oscar winning movies. Upon leaving *Million Dollar Baby* or *Schindler's List,* I felt really, really sad. If you asked me how I felt when watching *Million Dollar Baby,* I would have provided an extreme report on the sad scale. Was I happy during the movie? Absolutely not. Did I expect this going in to the movie? Without a doubt. (I am an avid reader of movie and television reviews before I view them.) The reason I went is to experience the human emotion, knowing that it would be extreme (and in the process, I may learn something about myself or human nature more

5. Again, this is done for illustrative purposes. It is also likely that television watching is measured correctly and that listening to music is measured incorrectly. In that situation, the results in row 1 of panels A and B of table 8.1 are measured correctly (given that listening to music is such a small portion of individual time).

generally). Much of television is also like this. I am an avid Miami Dolphins football fan. If you ask me how I felt (retrospectively) while I was watching a game, my answer would definitely depend on whether they won. I still watch the games every week, but I am honestly more sad than happy when I am watching the Dolphins, depending on how they are playing. This is verifiable within my family—I have been much sadder watching games this year, given the Dolphin's 1-15 win/loss record. This is common for most sports fans when following their team. They know one of the teams is going to lose while they are watching the game, yet they still knowingly watch.

More generally, there is a large industry within television that caters to extreme emotions. Movies on the Lifetime channel are often very depressing (yet garner sizeable ratings). A large fraction of Oprah's episodes are based on topics designed to illicit extreme negative emotions (elderly depression, violence against women, racism). Yet millions of people tune in daily. I never feel happy when I am watching television shows like *24, The Sopranos, Oz,* or *The Wire* (for example), but I am often very stressed or sad when watching them (less so with *The Sopranos,* more so with *Oz*). That is by design—*24* bills itself as a thriller, and that is exactly what I am seeking out when I watch the show. Often, dramas and thrillers are designed to deliver extreme negative emotions like sadness and stress.

The question that I think the authors need to think much harder about is, how does the U-index deal with such art forms where the design of the experience is to seek out negative emotions? This could be one reason that television watching has such a high U-index relative to other leisure activities. If part of the experience of watching television programs (or participating in art forms more generally) is to experience the full range of human emotions (both good and bad), am I really worse off when I watch shows that induce me to experience negative emotions? This seems very unlikely to me. People seek out such negative emotions (and advertisers regularly market those negative emotions). As previously seen, how we interpret an individual's well-being from watching television is critical to understanding the trends in well-being over time (given the large increase in television watching). Regardless, the chapter needs to at least acknowledge the U-indices' problem with television watching (or dealing with art forms more generally). The results, as currently presented, appear to be a strong violation of revealed preference. If people do not like television (especially compared to similar leisure activities), why do they watch so much of it?

8.4 The Potential Importance of Selection

As noted in the prior section, one of the drawbacks of the U-index seems to be its handling of experiences that are designed to elicit extreme negative emotions (like movies, music, or television). In this section, I set out three other issues pertaining to the U-index that the authors need to think about

more explicitly in their framing of the U-index. My sense is that these issues can be dealt with empirically. They just need to be acknowledged. Additionally, these issues will only be relevant when we try to predict an activity's enjoyment out of sample. If the goal is to measure individual well-being, eliciting affect measures for each activity for a given individual will not be subject to the selection issues I describe next.

8.4.1 Selection Issue One: More on Revealed Preference

Suppose, for simplicity, that the sample for our survey is only comprised of three people. Suppose further that persons 1, 2, and 3 are exactly the same in all observable dimensions (income, family size, similar distance to job, etc.) except for that the three individuals allocate their time differently to the categories found in table 8.3, panel A. We will assume that for all other time-use categories, the individuals allocate identical amounts of time. However, for food prep, walking, and watching television, the three individuals spend different amounts of time in these activities. For example, person

Table 8.3 **Hypothetical example of the importance of selection when computing the U-index**

	Person		
	1	2	3
A Time allocation for three hypothetical individuals (in hrs./week)			
Time spent on food prep/cleanup	2	0	0
Time spent walking	0	2	0
Time spent watching TV	2	2	4
B Assumed "true" U-index for the three individuals			
Time spent on food prep/cleanup	8	17	14
Time spent walking	18	8	16
Time spent watching TV	9	10	8
C Assumed "measured" U-index for the three individuals			
Time spent on food prep/cleanup	8	—	—
Time spent walking	—	8	—
Time spent watching TV	9	10	8

Notes: This table provides a simple example to show the importance of revealed preference and selection when computing the U-index. In the first panel, I provide the allocation of time for three hypothetical individuals in three activities: food preparation and clean up, walking, and watching television. In my hypothetical example, person 1 does not do any waking, person 2 does not do any food prep, and person 3 does neither walking nor food prep. In the second panel, I make up a corresponding U-index that could be consistent with the data in panel A. For example, the reason that person 1 does not walk is that for them, walking is a very unpleasant activity. In the third panel, I show the U-index that would be measured using the methodology used by Krueger et al. Notice, given that Krueger et al. only measure the U-index for activities that an individual performs, they would not measure the U-index for walking for persons 1 and 3, nor would they measure the U-index for food preparation for persons 2 and 3. In panel C, a dash indicates the unmeasured U-index for activities that were not performed during the week.

2 spends two hours per week on food prep, zero hours per week walking, and two hours per week watching television. Given this information, what could we conclude? Given the patterns of time use, we may conclude that person 1 does not like walking relative to preparing meals or watching television. Likewise, person 3 may like watching television more than walking and preparing meals.

Suppose we further survey these households to construct a U-index (in a manner similar to Krueger et al.). Suppose the following would be the *true* (as opposed to the measured) unpleasantness that each household feels while performing each activity. Again, these numbers are just for illustrative purposes. I made my fictional U-index on a zero to one hundred scale for each individual for illustrative purposes (with zero being least unpleasant and one hundred being most unpleasant). This makes it easier to make my point. Table 8.3, panel B shows the true U-index for each activity for each individual. For example, person 1 gets eight units of unpleasantness from food prep, eighteen units of unpleasantness from walking, and nine units of unpleasantness from watching television. Given these affect measures for each activity, it is not surprising that person 1 does not engage in any walking.

What would be the average U-index for each activity if we averaged the U-indices across our three people? According to these figures, the average U-index would be thirteen, fourteen, and nine for meal preparation, walking, and watching television, respectively. In this fictional world, time spent watching television is the most enjoyable activity on average. As a result, it is not surprising that all three individuals allocated positive amounts of time to television watching.

However, given that the U-index as measured by Krueger et al. only records the affect for activities that individuals actually engage in, the actual data that I would have at my disposable to compute my fictional U-index would be as shown in table 8.3, panel C. The measured U-index only includes the U-index for activities in which the person chose to participate. This is exactly analogous to the data available to Krueger et al., who only observe an individual's affect for activities where the individuals allocate positive amounts of time. In my example, given the way we measure the data, the measured U-indices will differ dramatically than the actual U-indices. Specifically, the measured U-indices found that food prep, walking, and watching television have measured U-indices of eight, eight, and nine, respectively. This is a direct violation of the underlying uncensored data shown in the previous table. The reason for this is that the only people observed walking and preparing meals are the individuals who really enjoy those activities. If we projected these values out of sample to someone who did not have a dishwasher and was forced to wash dishes, the utility we would get from washing dishes would be dramatically overstated (the unpleasantness would be understated).

In practice, how could such selection bias the results? The way that activ-

ity level affect is measured is through probabilistic sampling of a person's day. The more they like an activity, the more they will engage in an activity. The more they engage in an activity, the more likely that activity is going to be sampled. As a result, the affect measures will tend to be biased toward sampling activities that people like. Applying the affect measure for such activities to all others will likely overstate the utility (understate the unpleasantness) of the activity.

Second, given that changes in technology change the costs and benefits of engaging in certain activities, the selection I previously alluded to can be more or less pronounced when comparing activities across time (as the authors do). For example, as technological advances have occurred for food preparation (microwaves, take-out food, dishwashers, etc.) during the last forty years, the cost of reducing time inputs into food preparation has fallen. If food preparation is relatively unpleasant, we should see less people engaging in food preparation today. Those that do engage in food preparation should be those that relatively enjoy food preparation (because they chose not to purchase cheap market substitutes). Even if individuals' tastes have not changed over the last forty years with respect to the unpleasantness of preparing meals, the measured U-index would likely decline for food prep.

There is some evidence that the degree of selection has changed substantially over time. The fraction of households who engaged in some sort of food preparation in 1965 was 65 percent. In 2003, only 55 percent of households engaged in some sort of food preparation. Similar patterns are found among all home production time-use categories (and for market work for men). During the last forty years, people have seemed to substitute away from unpleasant activities. Conversely, the fraction of households who watch television on a given day has increased by 10 percentage points over this time period. Households should be substituting toward relatively more enjoyable activities and away from less pleasant activities.

Overall, the choice to spend time on an activity is related to how much one enjoys that activity relative to other activities. The U-index is based on enjoyment measures only for people who allocate time to a given activity. The more time that they allocate to an activity, the more likely it is that their enjoyment will be a component of the U-index. Given this, the U-index will be biased downward (enjoyment will be biased upward) relative to people's underlying preferences. This creates a problem with trying to project the U-index out of sample.

One solution to this problem is to ask people about their expected enjoyment if they were doing alternate activities. However, this method would be subject to the filtering issues associated with recalling distant memories that the authors are trying to avoid. A second solution is to just sample everyone's entire day with respect to measuring their affect (and do no projection out of sample).

In summary, I would like the authors to discuss this selection issue in their work and think about ways to address this issue when predicting out

of sample. Otherwise, if the selection issue cannot be overcome (perhaps because there are no instruments to deal with the selection), it severely limits the usefulness of the U-index relative to other existing measures of subjective well-being.

8.4.2 Selection Issue Two: Individual Fixed Effects

The authors already recognize the potential selection issue arising from individual fixed effects. However, when doing different analyses (like measuring trends in happiness over time), they do not account for such selection. The relevant issue for this type of selection is that different types of people do different types of activities. Also, different types of people have differing underlying levels of happiness. If the choice of activities is correlated with the underlying level of happiness, the U-index will be confounding individual fixed effects with the activity's latent enjoyment level.

The best way to deal with this issue is to remove individual fixed effects when computing the U-index. The authors do this in their chapter. They are able to do this, given that they have multiple observations of affect for a given individual. While they did not emphasize this directly, the level of affect changes dramatically once conditioning on individual fixed effects. For example, the data suggest that low educated individuals watch much more television than high educated individuals. Additionally, low educated individuals do much less exercising than high educated individuals. The happiness literature (as well as the U-index) suggests that low educated individuals are much less happy than high educated individuals. If this is the case, we would speculate that television watching has too high a U-index (because it is more intensively consumed by low happiness individuals) and exercise has too low a U-index (because it is more intensively consumed by high happiness individuals).

Comparing tables 1.8 and 1.9 confirms my predictions. Without controlling for individual fixed effects, exercise and television watching look very much different with respect to their unpleasantness (7.4 and 18.1, respectively—a gap of 10.7 on the U-scale). However, after controlling for individual fixed effects, the two activities look much more similar to each other. Specifically, the respective U-index for exercise and television watching are now 11.9 and 15.7 (a gap of only 3.8 on the U-scale). In fact, the gap in the U-index between television watching and almost all other leisure activities is relatively small once controlling for individual fixed effects. However, the U-index for housework and food prep did not change much after controlling for individual fixed effects. This is not surprising, given that there is a much smaller education and income gradient with respect to time spent in nonmarket work within the population.

So in summary, in all future work, I recommend that the authors only work with the fixed effect version of their U-index. Also, I would encourage them to highlight this issue in future iterations of their work. If this

research was implemented in different settings, those implementing their methodology should be encouraged to take multiple observations of affect for the same individual so individual fixed effects can be removed from their analysis.

8.4.3 Selection Issue Three: Time-of-Day Effects

One thing I would have liked to see is a control for time-of-day effects when comprising the U-index. If one's U-index changes throughout the day (regardless of activities) and some activities are more intensively consumed at certain times of the day (like television watching), the U-index for certain activities could be contaminated by time-of-day effects.

For example, the authors show that individuals are more likely to report being tired at the end of the day (figure 1.3). The end of the day is when people are most likely to watch television. So, if we classify television watching as a less enjoyable activity, is it because people find television more unpleasant than other activities, or is it because people watch television at the end of the day—when all activities are more unpleasant? To make policy prescriptions about moving individuals across different activities, we would want to know the true unpleasantness of the activity.

In summary, I would like to see the authors pull out time-of-day effects when computing their U-index measure for different activities. Again, I would also like them to caution other researchers who are trying to implement their research design that such time-of-day effects can be important.

8.5 Conclusions

Overall, this is a very ambitious and worthwhile project. The main short-term goal is to assess the value added by measuring changes in the U-index to assess changing well-being relative to changes in other existing measures of well-being (wages, GDP, happiness indices, etc.). If a large-scale data collection effort is to be created to measure the U-index, we need to understand the value added so as to start to think about the appropriate cost-benefit analysis. Only time will tell if the U-index adds substantive value to our understanding of the evolution of societal (or individual) well-being. The work of Krueger et al. provides a necessary first step in this evaluation process.

References

Aguiar, M., and E. Hurst. 2007. Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics* 122 (3): 969–1006.

U.S. Department of Labor, Bureau of Labor Statistics. 2003–2005. *2003–2005 American Time Use surveys.* Washington, D.C.: Government Printing Office.