INCOMPLETE DRAFT

National Time Accounting: The Currency of Life

Alan B. Krueger
Princeton University & NBER

Daniel Kahneman
Princeton University

David Schkade
University of California, San Diego

Norbert Schwarz
University of Michigan

Arthur A. Stone
Stony Brook University

First Draft: November 19, 2007

*We thank the National Institute of Aging, the Hewlett Foundation and Princeton University for generous financial support. We thank Leandro Carvalho, Marie Connolly, Molly Fifer, David Kamin, Amy Krilla and Doug Mills for excellent research assistance, and Ed Freeland, Jack Ludwig, John McNee and Rajesh Srinivasan for survey assistance. We are grateful to colleagues too numerous to thank individually for their constructive comments and criticisms, but they have improved our collective U-index.

National Time Accounting: The Currency of Life

Abstract

This monograph proposes a new approach for measuring features of society's subjective well-being, based on time allocation and affective experience. We call this approach *National Time Accounting* (*NTA*). National Time Accounting is a set of methods for measuring, comparing and analyzing how people spend and experience their time -- across countries, over historical time, or between groups of people within a country at a given time. The approach is based on evaluated time use, or the flow of emotional experience during daily activities. After reviewing evidence on the validity of subjective well-being measures, we present and evaluate diary-based survey techniques designed to measure individuals' emotional experiences and time use. We illustrate NTA with: (1) a new cross-sectional survey on time use and emotional experience for a representative sample of 4,000 Americans; (2) historical data on the amount of time devoted to various activities in the United States since 1965; and (3) a comparison of time use and well-being in the United States and France. In our applications, we focus mainly on the U-index, a measure of the percentage of time that people spend in an unpleasant state, defined as an instance in which the most intense emotion is a negative one. The U-index helps to overcome some of the limitations of interpersonal comparisons of subjective well-being. National Time Accounting strikes us as a fertile area for future research because of advances in subjective measurement and because time use data are now regularly collected in many countries.

Time is the coin of your life. It is the only coin you have, and only you can determine how it will be spent. Be careful lest you let other people spend it for you.

- Carl Sandburg

1. Introduction

The development of the National Income and Product Accounts (NIPA) was arguably the foremost contribution of economics in the last century, and the National Bureau of Economic Research's role in developing the accounts remains an unparalleled achievement. Nearly every country tracks its national income today, and limiting fluctuations in national income is a goal of public policy around the world. The National Accounts have been used to estimate bottlenecks in the economy, to forecast business growth, and to inform government budgeting.[1] As then Treasury Secretary Robert Rubin said, "the development of the GDP measure by the Department of Commerce is a powerful reminder of the important things that government can and does do to make the private economy stronger and our individual lives better."[2]

Yet GDP, national income, consumption, investment and other components of the National Accounts have long been viewed as partial measures of society's well-being -- by economists and non-economists alike. For one thing, the National Accounts miss "near-market" activities, such as home production (e.g., unpaid cleaning, cooking and child care), which produce services that could be purchased on the market. Perhaps more significantly, the National Accounts do not value social activities, such as interactions between friends or husbands and wives, which have an important effect on subjective well-being. Because economic activity is measured by prices (marginal valuations) in the National Accounts, the accounts also miss consumer surplus from market transactions. Diamonds are counted as more valuable than water, for example, yet one could question whether diamonds contribute more to society's well-being. Other limitations of the National Accounts that have long been recognized are: externalities are improperly accounted for; prices are distorted in imperfectly competitive markets; and the particular distribution of income in a country influences prices and marginal valuations. While attempts have been made to adjust the National Accounts for some of these limitations – such as by valuing some forms of non-market activity – these efforts are unlikely to go very far in overcoming these problems.

Many of these sentiments were alluded to by Robert Kennedy in his "On Gross National Product" speech at the University of Kansas on March 18, 1968:

[1] In one important early application, Fogel (2001; p. 213) describes how Simon Kuznets and Robert Nathan "used national income accounting together with a crude form of linear programming to measure the potential for increased [military] production and the sources from which it would come and to identify the materials that were binding constraints on expansion" prior to the U.S. entry in World War II.

[2] Quoted from "GDP: One of the Great Inventions of the 20th Century," *Survey of Current Business*, January 2000.

> Too much and for too long, we seemed to have surrendered personal excellence and community values in the mere accumulation of material things. Our Gross National Product, now, is over $800 billion dollars a year, but that Gross National Product - if we judge the United States of America by that - that Gross National Product counts air pollution and cigarette advertising, and ambulances to clear our highways of carnage. It counts special locks for our doors and the jails for the people who break them. It counts the destruction of the redwood and the loss of our natural wonder in chaotic sprawl. … And the television programs which glorify violence in order to sell toys to our children. Yet the gross national product does not allow for the health of our children, the quality of their education or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages, the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage, neither our wisdom nor our learning, neither our compassion nor our devotion to our country, it measures everything in short, except that which makes life worthwhile.[3]

The problem is not so much with the National Accounts themselves as with the fact that policy makers and the public often lose sight of their limitations, or misinterpret national income as the sole object of policy and primary measure of well-being.

In this volume, we propose an alternative way of measuring society's well-being, based on time use and affective (emotional) experience. We call our approach *National Time Accounting* (*NTA*). National Time Accounting is a set of methods for measuring, comparing and analyzing the way people spend their time, across countries, over historical time, or between groups of people within a country at a given time. Currently, time use is tracked according to the amount of time spent in various activities -- such as traveling, watching television, and working for pay -- but the evaluation and grouping of those activities is decided by external researchers and coders. Determining whether people are spending their time in more or less enjoyable ways than they were a generation ago is either impossible or subject to researchers' judgments of what constitutes enjoyable leisure activities and arduous work. In addition to the obvious problem that researchers may not view time use in the same way as the general public, other problems with this approach are that: 1) many people derive some pleasure from non-leisure activities; 2) not all leisure activities are equally enjoyable to the average person; 3) the nature of some activities change over time; 4) people have heterogeneous emotional experiences during the same activities; and 5) emotional responses during activities are not unidimensional. The methods we propose provide a means for evaluating different uses of time based on the population's own evaluations of their emotional experiences, what we call *evaluated time use*, which can be used to develop a system of national time accounts.

---

[3]Transcription available from: www.jfklibrary.org/Historical+Resources/Archives/Reference+Desk/Speeches/RFK/RFKSpeech68Mar18UKansas.htm

We view NTA as a compliment to the National Income Accounts, not a substitute. Like the National Income Accounts, NTA is also incomplete, providing a partial measure of society's well-being. National time accounting misses people's general sense of satisfaction or fulfillment with their lives as a whole, apart from moment to moment feelings.[4] Still, we will argue that evaluated time use provides a valuable indicator of society's well-being, and the fact that our measure is connected to time allocation has analytical and policy advantages that are not available from other measures of subjective well-being, such as overall life satisfaction.

There have been some attempts at NTA in the past, primarily by time-use researchers. Our approach builds on Juster's (1985; p. 333) seminal observation that "an important ingredient in the production and distribution of well-being is the set of satisfactions generated by activities themselves." To assess the satisfactions generated by activities, Juster asked respondents to rate on a scale from 0 to 10 how much they generally enjoy a given type of activity, such as their job or taking care of their children. Later research found that such general enjoyment ratings can deviate in important and theoretically meaningful ways from episodic ratings that pertain to specific instances of the activity (Schwarz, Kahneman, & Xu, in press). To overcome this problem, we utilize a time diary method more closely connected to the recalled emotional experiences of a day's actual events. Gershuny and Halpin (1996) and Robinson and Godbey (1997), who analyzed a well-being measure (extent of enjoyment) and time use collected together in a time diary, are closer forerunners to our approach. Our project is distinguished from past efforts in that we approach NTA from more of a psychological well-being and Ecological Momentary Assessment (EMA) perspective. For example, our measure of emotional experience is multidimensional, reflecting different core affective dimensions. We also developed an easily interpretable and defensible metric of subjective well-being, which combines the data on affective experience and time use to measure the proportion of time spent in an unpleasant state. And we use cluster analysis to determine which groups of activities are associated with similar emotional experiences to facilitate the tracking of time use with historical and cross-country data. Lastly, our survey methods attempt to have respondents re-instantiate their day before answering affect questions, to make their actual emotional experiences at the time more vivid and readily accessible for recall.

Past calls for National Time Accounting have largely foundered. It is instructive to ask why these efforts were not more influential in academic circles and why government statistical agencies have not implemented them. One possible explanation is that it is difficult to collect time diary information along with affective experience in a representative population sample. To this end, we developed a telephone survey, called the Princeton Affect and Time Survey (PATS), patterned on the Bureau of Labor Statistics' (BLS's) American Time Use Survey (ATUS), that is practical and easily adaptable for use in on-going official time use surveys. Another possible explanation is that evidence on the validity of subjective well-being measures has progressed greatly in the last decade. While subjective data cannot be independently verified, a range of findings presented in Section 3 suggests that self-reports of subjective experience indeed

[4] For surveys of economics research using the more conventional measures of life satisfaction, see Frey and Stutzer (2002) and Layard (2005).

have signal. The earlier efforts may have been ahead of their time and taken less seriously than they should have because such evidence was not yet available. Finally, it is difficult to track down documentation on the precise methods used in past diary *cum* well-being surveys. To facilitate replication and extensions, we have posted our main data sets, questionnaires and background documents on the web at www.krueger.princeton.edu/Subjective.htm.

The remainder of this paper is organized as follows. Section 2 provides a conceptual framework for using evaluated time use in National Time Accounting and discusses perspectives on well-being in economics and psychology. Section 3 provides evidence on the link between self-reports of subjective well-being and objective outcomes, such as health and neurological activity. Section 4 introduces the evaluated time use measures that we have developed and provides some evidence on their reliability and validity. Section 5 uses the PATS data to describe time use and affective experience across groups of individuals and activities. Section 6 provides a method for grouping activities into categories based on the emotional experiences that they are associated with. To illustrate the utility of our techniques, Section 7 describes long-term historical trends in the desirability of time-use and Section 8 provides a cross-country comparison. Section 9 concludes by considering some knotty unresolved issues and by pointing to some opportunities for NTA in the future.

2. Conceptual Issues

2.1. Economics of Time Use, Goods and Utility

In a standard economic model, households receive utility from their consumption of leisure and goods. People choose to work because of the income and hence consumption of goods that work makes possible. Available time and the wage rate are the constraints that people face. The national income and product accounts only value market output (or, equivalently, paid inputs and profits). Some attempts have been made to value non-market time using the wage rate as the shadow price of leisure. Becker (1965) argued that households combine resources (e.g., food) and time to produce output (e.g., meals), just like firms. Thus, in Becker's model cooking only affects utility through the subsequent enjoyment of eating. Pollak and Wachter (1975) expand this framework to allow home production activities to affect utility through their direct effect on utility during the activities themselves and through the consumption of the output produced during the activities.

Dow and Juster (1985) and Juster, Courant and Dow (1985) emphasize the notion of "process benefits," or the flow of utility that accrues during particular activities, such as work and consumption.[5] Juster, Courant and Dow illustrate this idea in a Robinson

[5] They define process benefits as the "direct subjective consequences from engaging in some activities to the exclusion of others. … For instance, how much an individual likes or dislikes the activity 'painting one's house,' in conjunction with the amount of time one spends in painting the house, is an important

Crusoe economy. Robinson can divide his time among three distinct activities, working in the market, cooking, and eating. He is constrained by the amount of food or clothing he can obtain through work, the amount of meals he can cook in a given period of time, and 24 hours in a day.[6] With the assumption that process benefits from activities are separable, utility can be written as:

(1) $U = V_w(t_w,x_c) + V_c(t_c,x_c,x_f) + V_e(t_e,x_c,x_m)$,

where V_w, V_c, and V_e are the process benefits derived during work, cooking and eating, respectively, x_c is the quantity of clothing, x_f is the quantity of food, x_m is the amount of meals cooked, and t is the amount of time devoted to each activity. Juster, Courant and Dow (p. 128) make the critical but sensible assumption "that the process benefit obtained from each activity is independent of the time and goods devoted to other activities." They defend this assumption by noting that "any stocks produced by activity i are permitted to affect the process benefits from other activities."[7]

The data that we collect are divided into episodes of varying length, not activities, so it is more natural to model the time devoted to episodes and the average process benefit during those episodes. Consider someone who spends her first t_1 hours of the day working, her next t_2 hours preparing meals, her next t_3 hours eating the meals prepared earlier, and her final t_4 hours working again. (Of course, this could easily be extended to allow for more episodes and other activities.) Under the assumption of separability, the utility function can be written as:

(2) $$U_i = \int_0^{t_1} v_1(t, X_c)dt + \int_0^{t_2} v_2(t, X_c, X_f)dt + \int_0^{t_3} v_3(t, X_c, X_m)dt + \int_0^{t_4} v_4(t, X_c)dt$$

Taking means of the flow utilities over the relevant intervals gives:

(3) $$U_i = t_1\overline{v_1}(t_1, X_c) + t_2\overline{v_2}(t_2, X_c, X_f) + t_3\overline{v_3}(t_3, X_c, X_m) + t_4\overline{v_4}(t_4, X_c)$$

It follows that a person's total utility can be obtained from the duration weighted sum of average process benefits during the time the individual is engaged in each episode. There is no need to collect additional information on resources or prices to summarize the person's well-being. Notice also that equation (3) does not require utility maximization. Even if the individual allocates his or her time suboptimally, if the mean process benefit can be estimated it is possible to estimate his or her well-being.

In this framework, which loosely guides our empirical work, the average well-being among N members of society, W, is $W = \sum U_i/N$. If one wants to put a dollar value on W, in principle it is possible to estimate the monetary price that people are willing to pay

---

determinant of well-being independent of how satisfied one feels about having a freshly painted house." The idea of process benefits is closely related to Kahneman's notion of "experienced utility".

[6] We ignore sleep to simplify the exposition.

[7] An exception might be exercise. A period of exercising may raise someone's mood during the rest of the day. We return to this below.

on the margin to increase their process benefit in some activity by one unit, and use the inverse of this figure as a numeraire. For example, the way workers trade off pay for a more or less pleasant job can give an estimate of the marginal willingness to pay to improve time spent in a pleasant state. Alternatively, the amount that people are willing to spend on various types of vacations can be related to the flow of utility they receive during those vacations to place a monetary value on additional utility. Although it is possible, under the assumption of rational decision making, to place a dollar value on W in this framework, we shy away from this step and focus instead on providing credible estimates of W.

Of course, measuring the flow of utility or emotions during various activities is no easy task, and some scholars doubt its feasibility entirely. Juster (1985) attempts to measure process benefits by using responses to the following question: "Now I'm going to read a list of certain activities that you may participate in. Think about a scale, from 10 to zero. If you enjoy doing an activity a great deal, rank it as a '10'; if you dislike doing it a great deal, rank it as a '0'; if you don't care about it one way or the other, rank it in the middle as '5'. … Keep in mind that we're interested in whether you like doing something, not whether you think it is important to do." The activities included: cleaning the house, cooking, doing repairs, taking care of your child(ren), your job, grocery shopping, etc. For activity j, the enjoyment score multiplied by the amount of time devoted to activity j is assumed to equal the process benefit, Vj.

There are several important limitations to Juster's type of enjoyment data, which we describe as a "general activity judgment" measure, because it focuses on a general response to a domain of life, not specific events that actually occurred. First, respondents are likely to develop a theory of how much they should enjoy a activity in order to construct an answer to the question. Second, respondents' may be sensitive to the interviewers' reactions to their answers. For example, someone may be concerned that they will be viewed as a bad parent or worker if they respond that they do not like taking care of their children or their job. Third, people are unlikely to correctly aggregate their experiences over the many times that they engaged in a particular activity in providing a general activity judgment. Other research (e.g., Kahneman, Wakker and Sarin, 1997) has found that individuals ignore the duration of events and instead place excessive weight on the end and peak of the experience when answering general evaluative recall questions. Fourth, and related, individuals are likely to exercise selection bias in choosing from the best or worst moments of past incidents of the specified activities. Results presented below cast some doubt on the validity of general activity judgments. Fifth, it is unclear if individuals utilize the enjoyment scales in an interpersonally comparable way.

Nonetheless, as a description of time-use and well-being, the process benefit approach has many advantages. Most importantly, the output of home production does not have to be observed or evaluated. A major goal of our work therefore has been to develop more informative measures of the flow of emotional experience during specific moments of the day.

2.2. The Psychology of Well-Being

Contemporary psychology recognizes a variety of informative SWB measures. Our view of the structure of subjective well-being concentrates on two qualitatively distinct constituents that both contribute to SWB. The first component pertains to how people experience their lives moment to moment as reflected in the positive and negative feelings that accompany their daily activities. We refer to this component as "experienced happiness," or the average of a dimension of subjective experience reported in real time over an extended period. The second component pertains to how people evaluate their lives. It is typically assessed with measures of life-satisfaction, like "Taking all things together, how satisfied would you say you are with your life as a whole these days?" There are many ways in which these components of SWB can be measured, but we view them as reflecting overlapping but distinct aspects of people´s lives.

Much of the variance of both experienced happiness and life satisfaction is explained by variation in personal disposition that probably has a significant genetic component (Diener & Lucas, 1999; Lykken, 1999). We focus here on two other determinants: the general circumstances of people´s lives (marital status, age, income), and the specifics of how they spend their time.

Evaluating one´s life-as-a-whole poses a difficult judgment task (see Schwarz and Strack, 1999, for a process model). Like other hard judgments, the evaluation of one´s life is accomplished by consulting heuristics - the answers to related questions that come more readily to mind (Kahneman, 2003). Experimental demonstrations of priming and context effects provide evidence for the role of such heuristics in reports of life satisfaction (Schwarz and Strack, 1999). Two heuristic questions that are used are: `How fortunate am I?´ and `How good do I feel?´ The first involves a comparison of the individual´s circumstances to conventional or personal standards, while the second calls attention to recent affective experience. Research indicates, for example, that reported life satisfaction is higher on sunny than on rainy days, consistent with the influence of the weather on their temporary moods. If individuals are first asked explicitly about the weather, however, they become aware that their current feelings may only reflect a temporary influence, which eliminates the effect of weather on reported life satisfaction (Schwarz and Clore, 1983).

In addition to personal effects, affective experience is determined by the immediate context and varies accordingly during the day; most people are happier sharing lunch with friends than driving alone in heavy traffic. Russell (1980) provides a theory of core affect, in which emotions are described along two dimensions. One dimension ranges from pleasure to displeasure, and the other from highly activated to deactivated. Happiness, for example, is an activated, pleasurable state. We define an individual´s experienced happiness on a given day by the average value of this dimension of affective experience for that day. Experienced happiness, so defined, is influenced by the individual´s allocation of time: a longer lunch and a shorter commute make for a better day. A person´s use of time, in turn, reflects his or her circumstances and choices. Favorable life circumstances are more strongly correlated with activation than with experienced happiness.

A classic puzzle in SWB research involves the limited long-term hedonic effects of outcomes that are greatly desired in anticipation and evoke intense emotions when they occur (Brickman, Coates and Janoff-Bulman, 1978). In a recent study using longitudinal data, Oswald and Powdthavee (2005) find that average life satisfaction drops after the onset of a moderate disability but fully recovers to the pre-disability level after two years.[8] This process is known as adaptation or habituation. Oswald and Powdthavee find that adaptation takes place but is incomplete for severe disabilities. Life events, such as marriage and bereavement have substantial short-run effects on happiness and life satisfaction, but these effects are mainly temporary (e.g., Clark et al., 2003). Findings like these invite the idea of a potent process of hedonic adaptation that eventually returns people to a set point determined by their personality (see Diener, Lucas & Scollon, 2006; Headey & Wearing, 1989).

Kahneman and Krueger (2006) conclude that adaptation to both income and to marital status is at least as complete for measures of experienced happiness as for life satisfaction. This conclusion is also consistent with Riis et al. (2003), who used experience sampling methods to assess the feelings of end-stage renal dialysis patients and a matched comparison group. They found no significant differences in average mood throughout the day between the dialysis patients and the controls.

A focus on time use and activities suggests to factors in addition to hedonic adaptation for understanding the stability of SWB. First, although personality surely matters, the claim that an individual´s experienced happiness must return to a set-point that is independent of local circumstances is probably false. For someone who enjoys socializing much more than commuting, a permanent reallocation of time from one of these activities to the other can be expected to have a permanent effect on happiness (Lyubomirsky, Sheldon & Schkade, 2005). Second, one must recognize that there are substantial substitution possibilities when it comes to activities. People who suffer injuries, for example, can substitute games like chess or checkers for competitive sports in their leisure time. These substitution possibilities are probably not anticipated. Thus, the largely unanticipated opportunity to substitute activities could attenuate the actual loss or gain in SWB associated with major changes in life circumstances, relative to anticipations.

A final observation is that the withdrawal of attention is another mechanism of adaptation to life changes. Attention is normally associated with novelty. Thus, the newly disabled, lottery winner or newlywed are almost continuously aware of their state. But as the new state loses its novelty it ceases to be the exclusive focus of attention, and other aspects of life again evoke their varying hedonic responses. Research indicates that paraplegics are in a fairly good mood more than half the time as soon as one month after their crippling accident. Intuitive affective forecasts will miss this process of attentional adaptation,

[8] Smith et al. (2005) find that the onset of a new disability causes a greater drop in life satisfaction for those in the bottom half of the wealth distribution than for those in the top half, suggesting an important buffering effect of wealth, although low-wealth individuals still recovered some of their pre-disability well-being.

unless they are corrected by specific personal knowledge (Ubel, Loewenstein, Schwarz, and Smith, 2005).

2.3 The U-Index: A Misery Index of Sorts

Two challenges for developing a measure of the process benefit of an activity are that the scale of measurement is unclear, and different people are likely to interpret the same scale differently. Indeed, modern utility theory in economics dispenses with the concept of cardinal utility in favor of preference orderings.

Survey researchers try to anchor response categories to words that have a common and clear meaning across respondents, but there is no guarantee that respondents use the scales comparably. Despite the apparent signal in subjective well-being data (documented in the next section), one could legitimately question whether one should give a cardinal interpretation to the numeric values attached to individuals' responses about their life satisfaction or emotional states because of the potential for personal use of scales. This risk is probably exacerbated when it comes to comparisons across countries and cultures.

We propose an index, called the U-index (for "unpleasant" or "undesirable"), designed to address both challenges.[9] The U-index measures the proportion of time an individual spends in an unpleasant state. This statistic has the virtue of being immediately understandable, and has other desirable properties as well. Most importantly, the U-index is an ordinal measure *at the level of feelings*.

The first step in computing the U-index is to determine whether an episode is unpleasant or pleasant. There are many possible ways to classify an episode as unpleasant or pleasant. The data collected with Experience Sampling Methods (ESM) or the Day Reconstruction Method (DRM) include descriptions of an individual's emotional state during each episode in terms of intensity ratings on several dimensions of feelings, some of which are positive (e.g., "Happy", "Enjoy myself", "Friendly") and some of which are negative (e.g., "Depressed", "Angry", "Frustrated"). We classify an episode as unpleasant if the most intense feeling reported for that episode is a negative one -- that is, if the maximum rating on any of the negative affect dimensions is strictly greater than the maximum of rating of the positive affect dimensions.[10] Notice that this definition relies purely on an *ordinal ranking* of the feelings within each episode. It does not matter if respondent A uses the 2 to 4 portion of the 0 to 6 intensity scale and Respondent B uses the full range. As long as they employ the same personal interpretation of the scale to

[9] The remainder of this section borrows heavily and unabashedly from Kahneman and Krueger (2006).

[10] Our approach bears some resemblance to a procedure proposed by Diener, Sandvik and Pavot (1991), which categorized moments as unpleasant if the average rating of positive emotions was less than the average rating of negative emotions. Unlike the U-index, however, averaging ratings of feelings requires a cardinal metric. Notice also that because the correlations between negative emotions tend to be low, their procedure will categorize fewer moments as unpleasant than the U-index.

report the intensity of their positive and negative emotions, the determination of which emotion was strongest is unaffected.[11]

From a psychological perspective, the U-index has some desirable attributes. First, the predominant emotional state for the majority of people during most of the time is positive, so any episode when a negative feeling is the most intense emotion is a significant occurrence. It is not necessary to have more than one salient negative emotion for an episode to be unpleasant. Second, the selection of a negative feeling as more intense than all positive ones is likely to be a mindful and deliberate choice: the maximal rating is salient, especially when it is negative, because negative feelings are relatively rare. Third, because at a given moment of time, the correlation of the intensity among various positive emotions across episodes is higher than the correlation among negative emotions, one dominant negative emotion probably colors an entire episode and it is potentially misleading to average negative emotions.

Of course, the dichotomous categorization of moments or episodes as unpleasant or pleasant obscures some information about the intensity of positive and negative emotions, just as a dichotomous definition of poverty misses the depths of material deprivation for those who are below the poverty line. However, we see the ordinal definition of unpleasant episodes as a significant advantage. In addition to reducing interpersonal differences in the use of scales, the question of how to numerically scale subjective responses is no longer an issue with our dichotomous measure.

Once we have categorized episodes as unpleasant or pleasant, we define the U-index as the fraction of an individual's waking time that is spent in an unpleasant state. The U-index can be computed for each individual (what proportion of the time is this person in an unpleasant emotional state?), and averaged over a sample of individuals. The same index can also be used to describe situations (what proportion of the time that people spend commuting is experienced as unpleasant?). Notice that because the U-index is aggregated based on time, it takes on useful cardinal properties. Like the poverty rate, for example, one could compute that the U-index is X percent lower for one group than another, or has fallen by Y percent over time

3. Is There Useful Signal In What People Report About Their Subjective Experiences?

Economists often treat self-reported data with a high degree of suspicion, especially when those data pertain to subjective internal states, such as well-being or health. Is there any useful signal in what people tell us about their subjective experiences? To answer this question, we first discuss how social scientists assess the validity of self-reports of behavior and subsequently develop a strategy for assessing the validity of self-reports of

[11] Formally, let f( ) be any monotonically increasing function. If P is the maximum intensity of the positive emotions and N is the maximum intensity of the negative emotions, than f(P) > f(N) regardless of the monotonic transformation.

subjective experiences before we turn to relevant empirical findings. Following the review of the evidence, we identify some limiting conditions and highlight that self-reports of affect are most meaningful when they pertain to recent specific episodes in a person's life, a fact that we exploit later in the design of the Day Reconstruction Method and the Princeton Affect and Time-use Survey.

3.1 Rationale

Many surveys ask respondents to report on their behavior. The validity of such reports can be assessed by comparing them with external records at the individual or aggregate level. For example, banking records can be used to evaluate the validity of self-reported expenditures at the individual level (e.g., Blair & Burton, 1987), and national sales figures can be used to assess the validity of purchase reports in representative sample surveys at the aggregate level (e.g., Sudman & Wansink, 2002). Neither of these strategies is feasible for assessing the validity of self-reported feelings, like moods, emotions, worries or pain. Feelings are subjective experiences and the final arbiter is the person who experiences them. The same holds for other subjective evaluations, like reports of life-satisfaction, which pertain to individuals' subjective assessments of the quality of their lives. The subjective nature of feelings and evaluations precludes direct validation against objective records. It is also expected that comparisons of subjective and objective reports will not be identical, because people interpret the objective world in idiosyncratic ways.

Nevertheless, one can gauge the validity of these reports in other, less direct ways. To begin with, one can assess interpersonal agreement: Do "close others" perceive the person in ways that are compatible with the person's self-reports? While interpersonal agreement is comforting, it is less than compelling and subject to numerous biasing factors. As a more informative alternative, one can relate self-reports of subjective experience to objective outcomes with the expectation that there should be at least a modest correspondence. If reports of positive affect are associated with increased longevity, for example, they obviously capture *something* real – yet it remains unclear whether that something is indeed positive affect or some other variable correlated with its expression (the so-called "third variable" explanation). Perhaps people who present themselves in a positive light when answering questions also follow other strategies of social interaction that reduce daily friction and benefit health? Such ambiguities are attenuated when studies that do not rely on self-reports for the assessment of affect show similar results. Finally, interpretative ambiguities are further attenuated when experimental results, based on random assignment, support the naturalistic observation, e.g., when induced positive affect also has beneficial health consequences. Such supporting results will typically be more limited in scope due to ethical constraints on the experimental induction of affect (especially negative affective states such as stress or anger) and the more limited time frame of experimental studies.

Our discussion follows the latter strategy. Specifically, we review illustrative findings from longitudinal studies that show that self-reported affect predicts some important objective outcomes in life. Paralleling these naturalistic observations, a growing number of experimental studies documents compatible effects of induced affect, based on random assignment of participants to positive or negative "affect induction" conditions. For example, positive affect can be induced by giving subjects a cookie or placing a dime in a spot where they can find it. Other approaches to inducing affect include placing subjects in a situation where they overhear a compliment or insult, showing subjects a funny vs. sad movie, asking subjects to recall a happy vs. sad event, and giving subjects a task that is easy or impossible to perform; see Schwarz and Strack (1999).

3.2 Affect and Objective Outcomes: Social Life

In a comprehensive review of cross-sectional and longitudinal studies, Lyubomirsky, King, and Diener (2005) observed that a preponderance of positive over negative affect predicts numerous beneficial outcomes, from the quality of one's social life and work life to longevity and the quality of one's health. Here, we focus on studies that are particularly informative with regard to the validity of affective self-reports, namely studies in which (i) the person's affect was assessed through self-reports several months or years prior to the observed outcome, (ii) the outcome itself is objective (e.g., longevity or health status rather than subjective satisfaction with one's health), and (iii) studies in which the affect assessment is *not* based on self-reports show compatible effects.

3.2.1 Finding a Spouse

Most people would prefer to be married to a partner who is happy and satisfied rather than depressed and dissatisfied. Consistent with this preference, several longitudinal studies show that people who report in sample surveys that they are happy (Marks & Fleming, 1999) or satisfied with their lives (Lucas, Clark, Georgellis, & Diener, 2003; Spanier & Fuerstenberg, 1982) are indeed more likely marry in the following years. For example, Marks and Fleming (1999) observed in a 15-year longitudinal study with a representative sample of young Australians that those who were 1 standard deviation above the mean of happiness reports were 1.5 times more likely to marry in the ensuing years; those 2 standard deviations above the mean were twice as likely to marry.

This relationship can also be observed with measures of affect that do *not* rely on self-report. For example, Harker and Keltner (2001) coded the affect expressed in women's college yearbook photographs, following the well-established procedures of Ekman's facial action coding system (Ekman & Rosenberg, 1997). They observed that women who expressed genuine positive affect (in form of a Duchenne smile) at age 21 were more likely to be married by age 27 and less likely to remain single through middle adulthood. Of course, people may report being happy because they anticipate being married in the next year, but the long lag in the in Ekman and Rosenberg study is harder to reconcile with reverse causality.

3.2.2 Helping Others

Several studies show that self-reported daily mood is associated with the likelihood of helping others. For example, Lucas (2001) observed that students who reported a preponderance of positive mood in their daily diaries also reported spending more time helping others than did those with less positive moods. Similarly, Csikszentmihalyi, Patton, and Lucas (1997) found that self-reported helping behavior increased with the percentage of time spent in a good mood among school-age youths.

Numerous experimental studies, with random assignment to different affect induction conditions, support the link between positive mood and prosocial behavior. People in induced positive moods are more likely to help others by donating money (Cunningham, Steinberg, & Grev, 1980), blood (O'Malley & Andrews, 1983), and time (Berkowitz, 1987) to worthy causes. Receiving a cookie or finding a dime is sufficient to elicit increased prosocial behavior (Isen & Levin, 1972).

3.2.3 Income

Several studies show a positive relationship between self-reported positive affect at a given time and later income. Diener, Nickerson, Lucas, and Sandvik (2002) observed that self-reported cheerfulness at college entry predicted income 16 years later, controlling for numerous other variables, including parents' income. For example, the most cheerful offspring of well-off parents earned $25,000 more per year than the least cheerful offspring. Similarly, Marks and Fleming (1999) observed in their Australian panel study of young adults that respondents' self-reported happiness in one wave predicted the size of the pay raises they had received by the time of the next interview, two years later. Finally, Russian respondents who reported high happiness in 1995 enjoyed higher incomes in 2000 and were less likely to have experienced unemployment in the meantime (Graham, Eggers, & Sukhtankar, 2006).

3.3 Affect and Objective Outcomes: Health

Numerous longitudinal studies show that happy people have a better chance to live a long and healthy life (for reviews see Lyubomirsky et al., 2005, and Howell, Kern, & Lyubomirsky, 2007). This observation holds for mortality in general as well as for specific health outcomes; moreover, it is supported by studies that relied on affect measures other than self-report.

3.3.1 Mortality

Based on data of the Berlin Aging Study, Maier and Smith (1999) reported that a preponderance of self-reported positive over negative affect (assessed with the PANAS) predicted mortality in a sample of 513 older adults three to six years later. Studies with clinical samples reinforce this observation. For example, Devins and colleagues (1990) observed that end-stage renal patients who reported overall happiness were more likely to survive over a 4 year period than were their less happy peers. Similarly, Levy and colleagues (1988) found that women who reported more joy in life were more likely to survive a recurrence of breast cancer over a 7 year period. Studies based on personality tests that assess enduring affective predisposition replicate this conclusion (see Lyubomirsky et al., 2005, for a review).

Complementary support for the observed relationship between positive affect and mortality comes from studies that asked the interviewer to rate the respondent's affective state. In one study (Zuckerman, Kasl, & Ostfeld, 1984), healthy as well as unhealthy respondents who were rated as happier enjoyed lower mortality than their peers over a 2 year period; Palmore (1999) replicated this observation over a more impressive period of 15 years. Finally, in a study that attracted broad attention, Snowdon and his colleagues (Danner, Snowdon, & Friesen, 2001; Snowdon, 2001) analyzed autobiographical essays that young catholic nuns of the American School Sisters of Notre Dame had written in 1930, when most were in their early twenties. Coding the essays for emotional content, they discovered that positive affect expressed in these early essays was highly predictive of mortality by the time the writers were 80 to 90 years old. On average, nuns whose essays placed them in the top quartile of positive affect in the sample lived 10 years longer than nuns whose essays placed them in the bottom quartile. Given that all nuns lived under highly comparable conditions in terms of daily routines, diet and health care, this finding provides particularly compelling evidence for the repeatedly observed relationship between positive affect and longevity.

3.3.2 Physiological Associations

Several conceptual models in the fields of health psychology and behavioral medicine posit a central role for positive and negative affect in the translation of the psychosocial environment into physiological states and, subsequently, health outcomes, such as those mentioned above. Empirical demonstrations of affect-physiology associations are a compelling source of validation for affect. We present representative findings in two physiological systems – the immune system and the endocrine system – because of their close linkage with health outcomes.

3.3.3 Immune Response

Alterations in immune system functioning – either above or below normative levels – can result in greater susceptibility to invading organisms and neoplastic diseases, and to autoimmune conditions. Therefore, many studies have examined how psychosocial factors and affect are related to various compartments of the immune system.

Several longitudinal studies observed that the frequency of self-reported hassles and uplifts and their accompanying affect is predictive of immune response. In one daily study, Evans and colleagues (1993) related participants' daily reports of life-events and mood over a two week period to markers of immune function in daily saliva samples. They observed a higher secretion of immunoglobulin A on days that were characterized by many positive and few negative events. Stone and colleagues showed their daily studies of events, mood, and symptoms that the impact of daily events on the secretory immune system was mediated through changes in negative and positive affect associated with daily events (Stone, et al., 1987; Stone, et al., 1996). A similar line of work by Vitaliano and colleagues (1998) monitored natural killer (NK) cell activity in cancer survivors. They found that participants who reported more uplifts than hassles (and presumably decreased levels of negative affect based on prior work [Stone, 1987]) in daily life showed higher NK cell activity 18 months later, an indicator of enhanced immune function.

Moving to more major events, a classic extensive line of work by Kiecolt-Glaser and colleagues demonstrated that naturalistic situations such as students taking exams or martially distressed individuals discussing their marital situation results in declines in immune functioning (for example, Kiecolt-Glaser et al, 1988). Changes in the immune system have been shown by the same investigators to have health consequences, such as in the resolution of experimentally induced wounds.

A particularly interesting series of studies by Cohen and colleagues demonstrated that people's level of affect is associated with their susceptibility to an experimentally induced viral infection and this is strongly supportive of the role of affect in physiology. In particular, recent evidence has indicated that pro-inflammatory cytokines are associated with positive affect (Doyle, Gentile, & Cohen, 2006) when measured on a daily basis.

Beneficial immune function effects of positive affect were also observed in experimental studies, based on random assignment to different affect induction conditions. For example, watching a humorous video clip has been found to increase NK cell activity and several other immune function markers (Berk et al., 2001), including salivary immunoglobulin A (Dillon, Minchoff, & Baker, 1985) and salivary lysozyme (sLys) concentration (Perera, Sabin, Nelson, & Lowe, 1998). Induction of stressful situations has also produced changes in immune function. For example, Stone (Stone et al., 1993) exposed participants to challenging mental tasks and they subsequently had lower responsiveness of t-cells stimulated with standard antigens compared to participants who were not exposed. A recent review article by Marsland, Pressman, & Cohen (2007) concludes that positive affect is associated with up-regulation of the immune system.

3.3.4 Hormones

Many bodily functions are regulated by the actions of hormones, which are biological active substances secreted by various organ systems. One hormone that has been of particular interest to psychosocial researchers is cortisol, a product of the hypothamalic-pituatary-adrenal (HPA) system. Cortisol is often called the "stress hormone". It affects aspects of metabolism in general, but of special interest for this discussion is its impact of the immune system and its anti-inflammatory role.

Observational and experimental studies have confirmed that cortisol levels are responsive to changes in affect and to experiences that are closely linked with affect changes. In an impressive line of research, Kirschbaum and colleagues (Kirschbaum, Pirke, & Hellhammer, 1993) showed that a laboratory manipulation involving stressful student presentations quickly increased levels of cortisol; such changes could at least temporarily suppress the immune system. Supporting the experimental work, there is evidence from naturalistic studies that sampled respondents affect and cortisol repeatedly throughout the day. Those studies showed that momentary negative affect is associated with higher levels of cortisol and positive affect with lower levels of cortisol (relative to when affect levels were at the opposite level) (Smyth et., 1998). Furthermore, both state (momentary) and trait measurement of affect is associated in the same manner with cortisol levels (Polk, et al., 2005).

3.3.5 Other Systems

Levels of positive and negative affect have also been associated with and shown to affect other physiological systems and we mention some of them here. Positive affect has been shown to increase performance on cognitive tasks and this could be associated with brain dopamine levels (Ashbym Isen, & Turken, 1999). Relatedly, measures of brain activity have been associated with affective levels (Wheeler, Davidson, & Tomarken, 1993). Some aspects of cardiovascular function and affect have been studied. Shapiro and colleagues (Shapiro, Jamner, & Goldstein, 1997) used daily monitoring of affect and blood pressure to show that specific mood states such as anger were associated with increased levels of blood pressure.

3.3.6 Neuro Activity

Findings from neuroscience research also lend some support for the view that subjective reports are related to individuals' emotional states. By way of background, note that there is strong clinical and experimental evidence that the left prefrontal cortex of the brain is associated with the processing of approach and pleasure, whereas the corresponding area in the right hemisphere is active in the processing of avoidance and aversive stimuli. In particular, the left prefrontal cortex is more active when individuals are exposed to pleasant images or asked to think happy thoughts, while the right prefrontal cortex is more active when individuals are shown unpleasant pictures and asked to think sad thoughts. A study using several measures of psychological well-being reported a statistically significant correlation of 0.30 between survey evidence on life satisfaction and the left-right difference in brain activation (Urry et al., 2004).

In a striking demonstration of the validity of subjective reports, Coghill and colleagues compared subjects' self-reported pain levels to functional magnetic resonance imaging (fmri) while applying a *standardized* pain stimulus to 17 subjects. The pain stimulus consisted of hot presses against the lower leg. They found that individuals reporting higher levels of pain to the thermal pain stimulus produced greater activation of various cortical regions of the brain, some of which corresponded with the stimulated limb, than individuals who reported lower pain ratings to the same stimulus (see figure 3.1; Coghill, McHaffie, & Yen, 2003). The strong implication of this work is that variation in self-reports to standard stimuli are not simply a function of interpersonal differences in scale usage, but reflect, at least in part, differential neural processes associated with the perception of pain. They concluded (p. 8358), “By identifying objective neural correlates of subjective differences, these findings validate the utility of introspection and subjective reporting as a means of communicating a first-person experience.”

3.4 Assessing Subjective Experiences

As our review indicates, there is systematic signal in people’s self-reports of their affective experiences. Nevertheless, self-reports of affect are subject to systematic methodological biases, which depend on the assessment method used. Next, we summarize what has been learned (for reviews see Robinson & Clore, 2002; Schwarz, 2007).

When people report on their *current* feelings, the feelings themselves are accessible to introspection, allowing for more accurate reports on the basis of experiential information. But affective experiences are fleeting and not available to introspection once the feeling dissipated. Accordingly, the opportunity to assess emotion reports based on experiential information is limited to methods of momentary data capture (Stone, Shiffman, Atienza, & Nebeling, 2007), like experience sampling (Stone, Shiffman, & DeVries, 1999), which we address in more detail in section 4. Once the feeling dissipated, the affective experiences need to be reconstructed on the basis of other information. When the report pertains to a specific *recent episode*, people can draw on episodic memory, retrieving specific moments and details of the recent past. Such reports can often recover the actual experience with some accuracy, as indicated by their convergence with concurrent reports (e.g., Kahneman et al, 2004; Stone, et al, 2006). The Day Reconstruction Method, presented in section 4, takes advantage of this observation.

In contrast, *global* reports of past feelings are based on semantic knowledge. When asked how they “usually” feel during a particular activity, people draw on their general beliefs about the activity and its attributes to arrive at a report. The actual experience does not figure prominently in these global reports because the experience itself is no longer accessible to introspection and episodic reconstruction is not used to answer a global question. Finally, the same semantic knowledge serves as a basis for *predicting* future feelings, for which episodic information is not available to begin with (Schwarz, Kahneman, & Xu, in press). These hedonic predictions, in turn, often serve as a basis for behavioral *choice* (March, 1978).

These processes result in a systematic pattern of convergences and divergences in affect reports. First, concurrent reports and retrospective reports pertaining to specific recent episodes usually show good convergence, provided that the episode is sufficiently recent to allow detailed reinstantiation in episodic memory. Second, retrospective global reports of past feelings and predictions of future feelings also show good convergence, given that both are based on the same semantic inputs. Hence, global memories are likely to “confirm” predictions. Third, choices are based on predicted hedonic consequences, and are therefore usually consistent with predictions and global memories. However, fourth, global retrospective reports as well predictions and choices will often diverge from concurrent and episodic reports, given that the different types of reports are based on different inputs. As a result, a person's expectations and global memories go hand in hand, but often fail to reflect what the person actually experienced moment to moment (for a review see Schwarz et al., in press).

These observations have important implications for the assessment of affective experience in time-use studies. They highlight that global reports of how much one usually enjoys a given activity are a fallible indicator of people's actual affective experience in situ. Such global reports were used in Juster and colleagues' pioneering studies (e.g., Juster & Stafford, 1985). Our work builds on Juster's (1985) conceptual approach while heeding the lessons learned from recent psychological research by employing measures of affective experience that pertain to specific episodes of the preceding day. Next, we turn to the development of these measures.

4. Methods for Collecting Evaluated-Time-Use Data: From EMA to DRM to PATS

The Experience Sampling Method (ESM) and Ecological Momentary Assessment (EMA) were developed to collect information on people's reported feelings in *real time* in natural settings during selected moments of the day (Csikszentmihalyi, 1990; Stone and Shiffman, 1994). Participants in real-time studies carry a handheld computer that prompts them several times during the course of the day (or days) to answer a set of questions immediately.[12] Participants are typically shown several menus, on which they indicate their physical location, the activities in which they were engaged just before they were prompted, and the people with whom they were interacting. They also report their current subjective experience by indicating the extent to which they feel the presence or absence of various feelings, such as angry, happy, tired, and impatient. Momentary real-time surveys are often viewed as the gold standard for collecting data on affective experience because it minimizes effects of judgment and of memory.

Survey Techniques for Collecting Data on Evaluated Time Use

Experience Sample Method (ESM) and Ecological Momentary Assessment (EMA). ESM and EMA are techniques for collecting data on time use and emotional experiences in real time. Respondents carry a computer device and indicate features of their activity and the feelings prior to being signaled by the device.

Day Reconstruction Method (DRM). DRM is a paper-and-pencil questionnaire that first collects time diary information from individuals for the preceding day. The diaries can list personal details, as they are not collected. Then, for each indicated episode, individuals indicate the nature of the activity, who was present, and the extent to which various emotions were present or absent.

Princeton Affect and Time Survey (PATS). PATS is a telephone survey patterned after the American Time Use Survey. After individuals report the activities of the preceding

[12] Other survey technologies can also be used for EMA, such as paper diaries and cell phones.

day (who with, what doing, where, when started and ended), three 15-minute intervals are randomly sampled and respondents are asked the extent to which various emotions were present or absent during that time.

So far, however, real-time data collection has proved prohibitively expensive and burdensome to administer to large, representative samples. An alternative to ESM which relies on a short recall period is the Day Reconstruction Method (DRM), which is described in Kahneman et al. (2004). The DRM combines elements of experience sampling and time diaries, and is designed specifically to facilitate accurate emotional recall.[13] Respondents, who participated in the survey in a central location, were provided with four packets containing separate questionnaires, and were asked to answer them in sequence. The first packet had standard questions on life, health and work satisfaction and demographics. Satisfaction questions were asked first so that answers were not contaminated by the other questions and diary that followed. Second, respondents filled out a time diary summarizing episodes that occurred in the preceding day. The third packet asked respondents to describe each episode of the day by indicating: when the episode began and ended; what they were doing (by selecting activities from a provided list); where they were; and with whom they were interacting. To ascertain how they felt during each episode in regards to selected affective dimensions, respondents were also asked to report the intensity of their feelings along 12 categories on a scale from 0 ("Not at all") to 6 ("Very Much"). The affective categories were specified by descriptors, mostly adjectives, such as happy, worried/anxious, and angry/hostile. The anchor, "Not at all," is intended to be a natural zero point that has a common meaning across respondents for these descriptors. The final packet contained personality and work questions. Subjects were paid $75 for filling out the DRM questionnaire, which usually took 45 minutes to 75 minutes to complete.

The emotions that respondents were asked to rate for each episode in the DRM were selected in part to represent points along the Russell (1980) affect circumplex. This distinguishes the DRM from the small number of past diary studies that included a question on how much individuals enjoyed (or liked/disliked) the activity they were doing. Russell models emotions as consisting of two core dimensions, *pleasantness* (pleasant vs. unpleasant) and *activation* (aroused vs. unexcited), with emotions positioned on a circle in this space. We interpret the duration-weighted average of the reported affect intensities as the average flow of "process benefits" or experienced well-being during the interval.

An early version of the Day Reconstruction Method was applied to a sample of 909 working women in Dallas and Austin, which we refer to as the Texas DRM (Kahneman, et al., 2004).[14] Another DRM survey was conducted of 810 women in Columbus, Ohio

[13] Robinson and Godbey (1997), Gershuny and Halpin (1996) and Michelson (2005) have used data collected from related survey techniques.
[14] The sample consisted of 535 respondents who were recruited through random selection from the driver's license list plus a screen for employment and age 18-60, and another 374 workers in three occupations:

and 820 women in Rennes, France in the Spring of 2005.[15] A major goal of the Texas DRM study was to determine whether, despite its reliance on memory, the DRM reproduces results found in ESM. We looked in particular for features of experience captured by ESM and DRM that deviate from people's lay intuitions. If DRM reproduces these patterns we can conclude that it captures respondents' actual experiences during the preceding day rather than their general intuitions about what their experiences "must have been like." One comparison along these lines is shown in Figure 4.1, which shows hourly mean ratings of "tired" in the DRM and from an independent study that used experience sampling. Whereas people's intuitions might hold that tiredness rises monotonically throughout the day, ESM studies show that tiredness reaches a minimum around noon. The DRM data replicate this V-shaped pattern, and the results obtain with ESM and DRM methods are remarkably similar. Moreover, this V-shaped pattern of tiredness was found in four subsequent DRM studies.

Other results of the Texas DRM conformed reasonably well to basic results frequently observed in Experience Sampling, despite differences in the sample demographics.[16] For example, the incidence of negative emotions is relatively rare in DRM -- "angry/hostile" was rated above zero only 23 percent of the time, while feeling "happy" was rated above zero 95 percent of the time. The same pattern is found in ESM studies. The correlations among the emotions, particularly the positive ones, were quite high across episodes -- around 0.7 for positive emotions and 0.4 for negative emotions. This pattern also replicates ESM findings. For example, the correlation of happy and "enjoying myself" across episodes is 0.73 in the DRM and 0.80 for a specialized sample of arthritis patients who participated in an ESM study.[17] Unfortunately, we are not aware of an real-time data capture study that collected sufficiently comparable data to compare activity ratings in the two methods.

Though not definitive, these findings suggest that DRM provides a reasonable approximation to the results of the more demanding ESM.

We also compared the DRM to a set of general activity judgment questions that closely replicated Juster (1985). Specifically, we asked the following questions to 252 women in Texas who were recruited in the same fashion as the Texas DRM sample:

Juster-Like Question. We would like to learn how likable or dislikable various activities are. Below we list a number of different things that you may often likely to do in your life. For each one, please circle the response that indicates how much you like or dislike it: (if one does not apply to you, you may skip it)

nurses, telemarketers and teachers. Because most results were similar for both subsamples, we present results for the full sample.

[15]Sampled individuals were identified by random-digit dialing.

[16] See Kahneman, et al. (2004) for further examples of non-intuitive patterns obtained with both methods.

[17] This correlation was computed using a sample of 84 arthritis patients who were prompted to report their feelings on a 0 to 100 visual analog scale three to twelve times a day, over an entire week.

	Dislike a great deal										Like a great deal
commuting to work	-5	-4	-3	-2	-1	0	1	2	3	4	5
working in your main job	-5	-4	-3	-2	-1	0	1	2	3	4	5
having lunch on a workday	-5	-4	-3	-2	-1	0	1	2	3	4	5
socializing at work	-5	-4	-3	-2	-1	0	1	2	3	4	5
commuting to home from work	-5	-4	-3	-2	-1	0	1	2	3	4	5
socializing with friends	-5	-4	-3	-2	-1	0	1	2	3	4	5
talking on the phone at home	-5	-4	-3	-2	-1	0	1	2	3	4	5
taking care of your children	-5	-4	-3	-2	-1	0	1	2	3	4	5
doing housework	-5	-4	-3	-2	-1	0	1	2	3	4	5
cooking/preparing food	-5	-4	-3	-2	-1	0	1	2	3	4	5
having dinner on a workday	-5	-4	-3	-2	-1	0	1	2	3	4	5
relaxing at home	-5	-4	-3	-2	-1	0	1	2	3	4	5
watching TV	-5	-4	-3	-2	-1	0	1	2	3	4	5

We then used just the adjective "enjoy" on a 0-6 scale from the Texas DRM to compute the average reported enjoyment while women engaged in these 13 activities according to the diary study. Table 4.1 compares the ranking of activities from the two approaches. The correlation between the ranks is 0.69. With small samples and some possible differential selection as to who participated in the activities on the diary day, the results should be read cautiously. Still, the results are quite similar to Juster (1985). The original Juster survey found that work and child care ranked particularly highly in terms of enjoyment, while our replication survey finds a similar result, especially for child care. Child care is reported as more enjoyable when asked about as an activity than in the diary-based study.[18] Work is ranked 8th in the Juster-like survey, perhaps not as highly as in the original because of our focus on women, but still higher than in the DRM. Interestingly, socializing after work is ranked much more highly in the DRM than in the general activity question. The contrast between these results, together with the contrast between the DRM and the original Juster rankings of activities, highlights the importance of collecting event-based data. Asking people to respond about how they feel about activities in general tends to provide a different ranking than when their actual experiences are used to guide their reported feelings during those activities.[19]

[18] Robinson and Godbey (1997) found a similar result comparing his diary based study to Juster's ranking.
[19] Gershuny and Halpin (1996) also cast doubt on the utility of general activity judgments. They analyzed data from a survey of British married couples in 1986 that asked a set of general questions about enjoyment with various activities. Respondents also maintained a diary for 5 days in which they reported their main activity during 30 minute intervals and, for each interval, how much they enjoyed their main activity, on a scale of 1 (very much) to 5 (not at all). Looking across subjects for a given activity, the proportion of the variation in the diary-derived enjoyment scale explained by the corresponding general activity enjoyment response was low, only 11 percent for supervising kids and 10 percent for cooking. Thus, people did a poor job predicting their own reported emotional experiences with a general activity enjoyment question.

4.1. PATS: A Phone Survey Version of DRM

The DRM is also burdensome and difficult to implement in a national sample. We designed the Princeton Affect and Time Survey to collect data from respondents over the phone more expeditiously. A related goal was to develop a module that could be added to the U.S.'s main time-use survey, the ATUS. The PATS survey works as follows. We started with the BLS ATUS questionnaire and eliminated a small number of questions that were not relevant. Respondents were first asked to describe each episode (defined as an interval of time in which the respondent was engaged in a specified activity; the average respondent reported 17.8 episodes) of the preceding day, using the ATUS protocols. Information about the activity individuals engaged in -- what they were doing, where they were, and who was with them -- was collected for each episode.

After the entire day was described in this manner, three episodes were randomly selected in proportion to duration and without replacement.[20] For these episodes, respondents were asked a 5-minute module of questions, covering the extent to which they experienced six different feelings (pain, happy, tired, stressed, sad, and interested) during each episode on a scale from 0 to 6. They were instructed that a 0 meant they did not experience the feeling at all at the time and a 6 meant the feeling was very strong. Specifically, respondents were asked to report their feelings during a randomly selected 15-minute interval of the sampled episodes. They were also reminded of what activity they said they were doing at that time in the diary part of the questionnaire. The order in which the feelings were presented was randomly assigned across respondents from six different permutations. The sampled episodes were ordered chronologically in the module. We also collected information on whether the individual was interacting with someone during sampled episodes.

The adjectives used in the PATS only partially overlap with those used in our DRM studies for a few reasons. First, we asked a smaller number of adjectives to save respondent time. Second, we avoided using compound adjectives, which we thought could be confusing to respondents over the phone. Third, the Gallup Organization conducted a set of 25 cognitive interviews with respondents to check their understanding of the affect questions and to make sure the questions made sense during most non-sleeping activities. These interviews helped us narrow down the set of emotions asked about.

The survey was administered by the Gallup Organization on our behalf in a random digit dial telephone survey of U.S. residents from May to August of 2006. Interviews were conducted in English and Spanish. A total of 3,982 people completed the survey, for a response rate of 37 percent. Weights were developed by Gallup to make the sample representative of the general population in terms of geographic region, gender, age and race. The weights were based on counts from the Current Population Survey (CPS).

---

[20] More specifically, the BLAISE computer program divided the day into 15-minute intervals and randomly selected three 15-minute intervals. If any of those intervals was in the same episode, additional 15-minute intervals were selected that were in other episodes so an episode was only included at most once.

Sixty-one percent of the unweighted respondents were women, a majority were white (88 percent), 90 percent had a high school education or higher, and 40 percent had household income less than $40,000 per year. The average age was 51.4 years. Re-weighting the sample to represent the population resulted in some significant distributional changes. Most notably, compared with the unweighted sample, the weighted sample had fewer women (53 percent), higher income (36 percent below $40,000), and a lower average age (45.2 years). Unless otherwise noted, we apply sample weights in all of the statistics we report based on PATS.

4.2 Evaluating PATS

We will use the PATS to illustrate NTA, so it is important to evaluate its properties in comparison to other time-use data sets and in comparison to results for affective experience captured in ESM and DRM.

Figure 4.2 shows that the allocation of time across activities (weighting individuals by sample weights) from the PATS closely matches that in the ATUS for the same months of 2004 and 2005. The correlation between time spent in these activities from the two surveys is an impressive 0.99. This high concordance suggests that the weighted sample is representative of the population, at least in terms of time use.

In Figure 4.3 we show the distribution of responses to the questions about feeling happy and tired over episodes in the PATS and Texas DRM. These adjectives were selected because they display different patterns – strongly skewed to the left for happy and slightly skewed to the left for tired except for a prominent mode at 0. It is reassuring that the distributions are very similar in both methods. Moreover, the incidence of reports of negative emotions was rare in PATS as was found in DRM and ESM.

We can also compare correlations between feelings across episodes in PATS to those in DRM and ESM. The correlation between feeling happy and feeling tired, for example, is -0.13 for women in the PATS and -0.21 in the Texas DRM survey of women and -0.34 in a Columbus, OH DRM survey of women. The correlation between feeling happy and stressed is -0.29 across women's episodes in PATS, and -0.44 in the Columbus DRM. The correlation between pain and happiness across episodes in the PATS is -0.10, while the corresponding correlation across moments in ESM data is -0.20 for the sample of arthritis patients mentioned above. These results suggest that the correlation between pairs of reported emotions in the PATS is a little weaker than the corresponding correlations in ESM and DRM, but they point in the same direction and are qualitatively similar.

With only three sampled episodes per interview, it is probably more difficult for respondents to reproduce their precise pattern of tiredness over the day. Still, the correspondence between the diurnal pattern of tiredness in PATS and DRM and ESM is reasonable (see Figure 4.4). The pattern displayed by the PATS data is much less V-shaped than was the case in the other surveys, but the increasing pattern of tiredness in the afternoon and evening is clearly evident. The correlation between the average rating of tiredness each hour in PATS and DRM is 0.87, and between PATS and ESM is 0.86. Moreover, the PATS data show similar age interactions to what we found earlier; namely, a sharper decline in tiredness in the morning for younger respondents.

The correlation between reported life satisfaction and net affect across people was also similar in PATS and the Texas DRM. In the (random sample component of the) Texas DRM, the correlation between life satisfaction and net affect is 0.44 and in the PATS it is 0.35. Because net affect can be computed for only three episodes per person in the PATS, however, one would expect the 0.35 correlation to be biased downward. To make a fairer comparison, we randomly selected three episodes per person from the DRM. In this more comparable sample, the correlation fell to 0.39, quite close to the 0.35 computed with PATS. Krueger and Schkade (2007) provide estimates of the reliability of life satisfaction and net affect. Using their estimates to adjust for attenuation bias, the correlation between life satisfaction and net affect would rise from 0.44 to around 0.70. This figure suggests that interpersonal variations in average net affect over many days reflects about half of the variability in life satisfaction.

Table 4.2 considers how the average rating of happy compares across common activities in the PATS and the random sample of the Texas DRM, both on a 0-6 scale.[21] The Pearson correlation between the two measures is 0.78, and the rank-order correlation is 0.74. Childcare is the largest outlier, with a one-half point lower rating in the DRM. Television is another outlier, with the DRM exceeding the PATS. In these respects, the PATS ranking of activities are intermediate between the rankings in the Juster-like survey and the DRM. It is possible that in the PATS respondents reflect more on the activity in general than the particular episode. Another possibility is that differences in the sample populations between PATS and the DRM account for the discrepancies.

Table 4.3 summarizes results on how the order of emotions affected reported intensity of feelings in PATS. As mentioned, we randomly assigned respondents to one of six different orderings for the affect questions. Once an order was selected, the same order was used for each of the three sampled 15-minute intervals. The order effect for each of the emotions is statistically significant at the 0.025 level, and usually much lower. As a general rule, when positive emotions were asked about early on, their ratings tended to be higher, and when negative emotions were asked about early their ratings tended to be lower. If happy was asked first, for example, its mean response was 4.35, compared with 3.99 when it was asked last; when pain was asked first its mean response was 0.89, compared with 1.08 when it was asked last. Interestingly, the adjective “interested” behaved like a positive emotion in this regard. Table 4.1 combines results for the first, second and third episode that was inquired about. Surprisingly, when we disaggregated the order effects were not notably stronger for the first of the three episodes. We expected to find stronger order effects for the first episode, as the order was known to respondents by the second and third episode. One interpretation of these results is that the first emotion provides an anchor for the subsequent ones. Respondents are typically in a positive mood before the affect questions are asked (judging from the high frequency

[21] Attempts were made to make the activities as comparable as possible.

of positive affect), and the response to the first emotion question is anchored relative to this positive feeling. Because the order in which emotions were presented was randomly assigned to respondents in PATS, our results should not be biased by order effects in any event.

It is also worth noting that the particular ordering used did not have a significant effect on the level of the U-index (p-value = 0.37 for joint F-test of constant U-index). Thus, a salutary feature of the U-index is that it is apparently robust to order effects, because the anchoring that produces the order effects does not substantially alter the ordinal ranking of emotional ratings.

Lastly, we can examine how the weather relates to the PATS affect and satisfaction data. Table 4.4 summarizes results from Connolly (2007), who merged daily weather data from the National Climate Data Center to the PATS survey. Specifically, she merged data on the mean temperature and amount of rainfall on the interview day and diary day (which is the day prior to the interview day), as well as the normal temperature and rainfall for the season and geographic area. Because temperature is highly correlated on adjacent days, it was not possible to estimate separate effects of the temperature on the interview and diary day. Rainfall, however, varies considerably from day to day. Women's reports of their life satisfaction and affect were more sensitive to the weather than men's, so we focus on results for women here. As in Schwarz and Clore's survey, Connolly found that life satisfaction was lower in the PATS if women were interviewed on rainy days. Life satisfaction was also lower in areas with higher normal precipitation levels and temperature. Temperature on the interview day was unrelated to life satisfaction, but a higher temperature on the diary day was associated with lower net affect. Since PATS was conducted in the late spring and summer, one might expect hotter days to be associated with lower net affect. Rain on the interview day was insignificantly related to net affect, while a small amount of rain on the diary day was associated with lower net affect. These results suggest that the weather influences reported net affect in the PATS data in a plausible way that is consistent with the true effect of the weather on people's moods, while the weather on the interview day is unrelated to net affect reported for the preceding day, as one would hope.

* Note: In the next draft, we will have results from an ongoing study in which 200 subjects are participating in an EMA study for 3 days and responding to the PATS survey for the overlapping 3 days. This will provide the strongest validation data yet available for both the time use and affect questions.

5. Well-Being Across Groups and Activities

5.1. Differences in Well-Being Between Groups

We use the PATS to compare affective experience across groups of individuals and frequent uses of time. Table 5.1 reports the average U-index for several demographic groups, and some of those results are highlighted here. (The Appendix Table presents results of the effect of demographic and other variables on the U-index in a multiple regression framework.) The U-index is 2 points higher for men than women (p-value < 0.10). The U-index is higher for Blacks and Hispanics than for Whites. The U-index falls with household income and education. Those in households with income below $30,000 per year spend almost 50 percent more time in an unpleasant state than do people with income above $100,000 per year (22.5 percent versus 15.7 percent). The data indicate a mild inverse-U shape pattern in unpleasant moments with age for women. These patterns are often found in life satisfaction data and in our earlier DRM studies.

Married men and women have the same U-index, 17.4 percent. The U-index for never married men and cohabiting men is also around 17 percent. The U-index is notably higher for unmarried women and divorced men. The former result is a contrast to our previous DRM studies, which found that married and unmarried women exhibited a similar U-index. Interestingly, the U-index is around 23 percent for all groups of unmarried women, divorced, widowed, cohabiting and never married. In a regression, the married-unmarried gap is not accounted for by controlling for demographic variables or activities. Controlling for differences in household income, however, accounts for more than half of the marriage gap in the U-index for women.

5.2. Activities

Table 5.2 reports the U-index and mean of five reported emotions during various primary activities. The order of activities is ranked by the U-index. The U-index is relatively low during discretionary activities, including religion/prayer, sports and exercise, relaxing and leisure, and socializing. Watching television is rated in the middle of the activities shown, as are food preparation and volunteering. The highest U-index activities include housework, working for pay, household management, receiving medical care, education, and caring for adults. This pattern is quite plausible, although it deviates in some important respects from the Juster-like general activity results.

Some of the ratings of the specific emotions are also worth discussing. The intensity of both pain and happiness are high during episodes of sports and exercise, especially for men. This pattern, which is not surprising, may result from elevated endorphins during exercise. The low rating of “interested” during education-related activities might be related to the high dropout rate of college-age students in the U.S. Telephone calls seem to evoke a high level of diverse emotions, with above-average ratings of happy, stressed, sad, and interest. Medical care is rated as an especially painful activity, particularly by women. The emotional experience of watching television appears quite close to the overall average emotional experience during the day, except for stress, which is below average.

A salutary feature of the PATS is that the same individual reports on multiple episodes of the day. As a result, individual fixed-effects (means) can be removed when studying differences in activities. Table 5.2b reports the U-index and affective ratings during the various activities after removing individual fixed-effects. In essence, this analysis compares the emotional ratings of the same individual as he or she moves from one activity to another. In general, the activities are ranked similarly with or without fixed-effects removed. The correlation between the U-index across activities in Table 5.2a and 5.2b is 0.93. The biggest movement occurs for medical care and personal care, both of which become less unpleasant when person-effects are removed, indicating that the people who tend to engage in these activities have a higher-than-average U-index during other episodes of the day. Since people tend to seek medical care when they are in pain or ill, this finding is quite plausible.

Another feature of the PATS is that affect can be modeled before, during and after participating in a specific activity. Figure 5.1 illustrates this point by showing the average rating of the emotion "happy" in relation to the occurrence of an episode involving sports or exercise. Specifically, we regressed the happiness rating on the number of minutes before or after an episode involving exercise with an interaction to allow for a different slope before and after exercise, for the subset of people who exercised on the interview day. The model was estimated both with and without person fixed-effects. Time 0 corresponds to the period of exercise. Especially in the model that removes person fixed-effects, an inverse-V pattern is evident: Happiness rises as a period of exercise approaches and then decays afterwards. With more observations, a less constraining model could be estimated.

5.3. Interaction Partners

The presence of others during an episode affects the pleasantness of the experience. Table 5.3 presents the U-index for men and women, disaggregated by who else was present during the episode. The tabulations do not control for other features of the episode, but the pattern is generally similar when we control for the activity engaged in during the episode as well as person fixed-effects. For simplicity, we present the unadjusted results here.

When people are alone, the U-index is higher than when they interact with others. The identity of the "others" matters, however. For men, the U-index is lower when friends and relatives are present. Spending time with co-workers is associated with a higher U-index for both men and women, primarily because work has a high incidence of negative emotions, particularly stress. Spending time with the boss makes the experience of work notably more unpleasant. The pattern for men and women is similar, except for the striking elevation in the U-index for women when it comes to spending time with one's parents or children.[22] These differences are partly explained by the different mix of activities that men and women engage in when they are with their parents and children.

[22] The ranking in Table 5.3 for women is exactly the same as was found for interaction partners in the Texas DRM, except parents were not separately identified in the DRM.

For example, men spend relatively more of their time with children watching television and traveling than do women, while women spend relatively more of their time with children engaged in childcare and doing chores. Even holding activities constant, however, there are sizable differences in the U-index between men and women when they are in the company of their parents or children.

5.4. *Day of Week*

Table 5.4 reports the U-index by day of the week (i.e., the diary day). A test of a constant U-index across days is rejected at the 0.01 level. Not surprisingly, weekend days are associated with less unpleasant feelings than weekdays, although the U-index is slightly lower on Fridays than on Saturdays. (For many people, apparently the weekend starts on Friday.) The U-index is lowest on Sundays and slightly higher on Mondays than on Tuesdays through Thursdays. Almost half of the weekend-weekday difference in the U-index can be accounted for by the different mix of activities that take place on the weekend. The empirical support for the song "rainy days and Mondays always get me down" thus far is limited, as a statistical test does not find the U-index on Monday to be significantly higher than on other weekdays (t=1.41), and the evidence on rain on the diary day cited in Table 4.4 was mixed as well.

5.5. *Goods and Time Use*

In the standard economic model, people consume goods to increase their utility. Time use data are notably lacking in information on goods consumption. Instead, it can be hoped that the activity description reflects the goods consumed during an episode or that no goods are involved. In many situations, however, this is likely to be inadequate. For example, food must be involved during episodes of eating, but we lack information on the quantity or quality of food. Dinners at McDonalds' or the French Laundry are obviously not equivalent experiences, yet these events are lumped together in the time use data. When computed at the episode level, the U-index potentially reflects features of the episode such as consumption of goods that are not captured elsewhere in the data. Unobserved features of activities, including goods consumption, surely account for some of the variability in emotional responses across respondents engaged in a given activity.

The largest expenditure item for most people is their housing. Wong (2007) merged data on housing values and other housing characteristics to the Columbus DRM to explore the effect of housing consumption on subjective well-being. She finds that respondents who live in larger or more expensive homes do not report higher net affect while they are at home (either absolutely or in comparison to time spent away from home). This conclusion holds for both women with and without children living at home. She also finds that reported joy from one's house and home is unrelated to the market value of the home but is positively related to the market value of the homes in the neighborhood.

To illustrate the effect of the consumption of goods on the affective experience of time use, in the PATS we collected information on the size of the television set being viewed during episodes of watching television. Because television absorbs such a large proportion of people's time, this seemed a particularly worthwhile activity to focus on. Specifically, we asked respondents whether the television screen they were watching was greater than or smaller than 25 inches. (If we were to redo the survey today, we might ask about flat screen versus not-flat screen.) We regressed each of the reported emotions during television watching on an indicator for the size of the television set, education, household income and the mean affect rating during other episodes of the day. The results indicated some emotional benefit from watching a larger television: stress was lower (t=-2.7) and net affect was higher (t=2.0) if a larger television was being watched. Although we would not make too much out of this result, it does suggest the utility of collecting information on the nature of the goods involved during participation in certain activities.

Clearly more could be done in connecting goods to the quality of experiences. For example, the nature of kitchen equipment could be related to affect during episodes involving cooking, and the make and model of cars could be related to affect during episodes of travel. Note, however, that goods only affect people's hedonic experience when they attend to them. For example, Schwarz, et al (in press) explored how the quality of the car driven (as indexed by the car's Bluebook value) affects the driver's emotional experience. They found that drivers feel better driving luxury cars than economy cars – but only during episodes that are car-focused; that is, in the 2 percent of episodes that the drivers categorized as "driving for fun." In the other 98 percent of driving episodes, like commuting to work or shopping, the type of car driven was unrelated to drivers' emotional experiences. In short, the car only made a difference when the car was on the driver's mind. However, drivers are not aware of this contingency and drivers of luxury cars reported that they "generally" feel much better while driving than drivers of economy cars. Such discrepancies between global and episodic reports of enjoyment highlight that global reports of one's "usual" experience are based on general beliefs about the type of activity, which are often at odds with actual experience as captured by episodic assessments.

5.6. Decomposing Group Differences: The Case of Age and Income

5.6.1 Age

Past research finds that older individuals report fewer negative emotional experiences and greater emotional control than younger individuals (e.g., Gross, et al. 1997). Consistent with this result, we find that the U-index is lower for those age 65 and older than for the younger population. The younger group works more and spends more time taking care of children, activities associated with stress (see Figure 5.2). How much of the difference in the U-index between young and old is accounted for by differences in their activities? Here we provide an example of how the difference in well-being between groups can be attributed to differences in time allocated across activities and differences in affect derived from a given set of situations and a residual.

To simplify the analysis, we focus on the gap in the U-index between people age 25-64 and those 65 and over. We also confine our attention to weekdays, when differences in activities are more pronounced. Table 5.5 summarizes our results. The U-index is 20.4 percent for the younger group and 16.1 percent for the older group, a gap of 4.3 points (p= 0.007). If we compute the U-index using each groups actual time allocation and the average activity ratings for the combined sample, so the entire difference is due to differences in time allocated across activities, the gap is predicted to be 2.5 points.[23] Thus, 58 percent (= 100 x 2.5/4.3) of the difference in the U-index between young and working-age is solely a result differences in their activities. The remaining 1.8 point gap is a result of differences in emotional responses to the same set of activities or an interaction between differences in ratings and differences in time allocation.

A further indication that choice of activity plays a role here comes from comparing the weekend and weekdays. On the weekend, the U-index falls to 16.8 percent for the younger group, not very different from the U-index for the older group during the week.[24]

Table 5.5. Decomposition of U-index for 25-64 year olds and those 65 and over

Group	**Actual**	**Predicted**	**Unexplained by Activities**
25-64 year olds	20.4%	20.0%	0.4%
65 +	16.1	17.5	-1.4
Difference	4.3	2.5	1.8

Notes: Table gives actual episode-level U-index and the predicted U-index using the overall sample's average U-index at the activity-level. Seventy-two harmonized activities are used.

5.6.1 Income

To be added. Bottom line: Activities do not explain much of the gap in U-index for top and bottom half of income distribution. Even episodes of TV watching have a lower U-

[23] This is mostly a result of the difference in working hours. During weekdays the younger group spent 24 percent of its awake time at work compared with just 2.6 percent for the older group. The U-index is 9 points higher during work-related episodes. So 9% x (.24-.026) = 1.9 points of the 2.5 points is due to the difference in time spent at work.
[24] The U-index also falls for the older group, but by a smaller amount, to 13.4 percent. Perhaps the elderly are more cheerful on the weekend because they interact with more cheerful younger people on those days.

index for the rich. But there is a bigger issue about reverse causality here. The rich may be rich because they have personalities that enable them to make a lot of money.

6. Identifying Affectively Similar Activities[25]

Summarizing time use data at the activity level can be unwieldy. The ATUS, for example, has hundreds of detailed activity codes. To make the analysis tractable, it is necessary to group activities into common categories. But classifying activities requires judgments of what activities are similar. Should gardening and lawn care be classified with leisure activities or with home production activities, for example? Researchers may have a different view of the enjoyment derived from such activities than the general public. (See Aguiar and Hurst, 2007 and Ramey and Francis, 2006 for alternative results in which researchers classified time use into broad categories, such as leisure, home production and market work.)

Rather than externally assign activities to groups, we propose an alternative approach: Use the average of the emotional ratings that respondents reported during each activity to assign activities with similar emotional experiences to the same group. Specifically, we use K-means cluster analysis to identify K groups of activities associated with similar emotional experiences. Cluster analysis is a family of techniques for assigning observations to groups (clusters) in a way that minimizes the discrepancies within groups and maximizes discrepancies between groups. For a single outcome measure (e.g., happy), the K-means cluster technique minimizes the within cluster variance, which also has the feature of maximizing the between cluster variance in means. The interpretation is more complicated with more than one outcome measure, but the intuition is the same. The algorithm for the Stata cluster procedure used here minimizes the sum of squared Euclidean distances of the emotions associated with the activities from their cluster means (which is equivalent to maximizing between group differences as well due to a multivariate extension of the Pythagorean identity from ANOVA).

We illustrate this approach using ratings of pain, happy, tired, stressed, sad, and interested to cluster activities. Activities form the unit of observation. For each activity, we computed the weighted average of each of those six emotional responses. Activities in the PATS were originally coded with the same system that the Census Bureau uses for ATUS. Because we will use the groups to make historical comparisons in Section 7, we converted the ATUS activity codes to 72 "harmonized" codes used in the American Heritage Time Use Studies (AHTUS).[26] These harmonized codes are activity codes that can be compared over time in a consistent way. We set K to equal 6, mainly because 6 is

[25] This section and the next one borrow heavily from Krueger (2007).

[26] The concordance was from the Center for Time Use Research (www.timeuse.org/athus/documentation). The concordance contains 92 activities, 14 of which could not be coded in the ATUS. We combined child care regardless of the child's age. We omitted sleeping and napping and a small number of infrequent activities that were not covered by PATS, resulting in 72 harmonized activities.

a tractable number of categories and because it is not very different from the number of categories that researchers have used in the past. It would be possible to explore the sensitivity of the results to other values of K, or to select K on the basis of a goodness of fit test.

Two additional features of the analysis are worth noting. First, the activities were weighted by their relative frequencies.[27] Thus, the resulting clusters can be thought of as minimizing the weighted sum of within-group variances. Second, because cluster analysis is an iterative procedure that can be sensitive to the starting point, we executed the cluster command 35 times using random starting points and selected the estimates with the highest Calinski and Harabasz pseudo-F statistic, defined as:

$$F = \frac{\text{trace(B)/(g - 1)}}{\text{trace(W)/(n - g)}} \;,$$

where B is the between-cluster sum of squares and cross-products matrix, W is the within-cluster sum of squares and cross-products matrix, g is the number of groups and n is the sample size.

Table 6.1 reports the optimal cluster assignments for the most common activities and the average ratings for each of the six emotions. In addition, the table reports net affect, the positive emotion (happy) less the average of the negative ones (sad, pain, stressed). Many of the cluster assignments make intuitive sense. Paid work performed at home and away from home, for example, are both in cluster 6, as is helping someone with homework. Home production activities, including cleaning and putting away dishes, are mostly assigned together in cluster 5. There are some unexpected results, however. For example, time on a second job is classified in cluster 2 while other paid work is in cluster 6. Unfortunately, we did not collect occupation or industry for secondary jobs. Compared with surveyed episodes during the main job, people on a second job were much less likely to work with co-workers and were more likely to work alone or with their spouse.

In addition to tracking and organizing time use, another application of the classification of activities that result from this exercise would be for non-market NIPAs. In particular, a question often arises in valuing non-market activities whether an activity should be valued at the wage rate, at the market wage for hiring someone to perform a task, or at some other price. Another issue concerns whether particular activities such as schooling are primarily consumption activities or investment activities. One answer to this question is that activities that are as stressful and uninteresting as someone's main job should be valued at the same wage as the main job. Likewise, activities that are as enjoyable as socializing should be treated as leisure. The cluster analysis provides a means for identifying activities that are associated with similar emotional experiences. For

[27] Because Stata does not have a weight option with cluster, we created a new data set in which each activity could be represented multiple times, in proportion to its relative frequency.

example, time spent in school does not appear to be a consumption activity in our data, and time spent taking care of teenagers appears as taxing as one's main job.

Table 6.2 reports the mean of the emotions and net affect for each cluster of activities. The lowest rated cluster in terms of net affect is cluster 1, which includes receiving medical care, purchasing medical services, seeking government services and doing homework. Cluster 2 involves tasks like writing and using a computer. The most enjoyable and interesting activities are in cluster 3, including religious activities, exercise, attending parties, listening to music, playing with children and recreation. Cluster 4 is a mixture of activities, such as watching television, relaxing, cooking and gardening, that are close to average in terms of affect ratings. Cluster 5, which includes domestic activities such as doing laundry, ironing, caring for adults, and cleaning, is slightly above cluster 6 (work) in terms of net affect but well below it in terms of interest.
If we were to assign value laden terms to describe the clusters, we could think of cluster 1 as unpleasant personal maintenance, cluster 2 as moderately enjoyable tasks, cluster 3 as engaging leisure and spiritual activities, cluster 4 as neutral downtime and cooking, cluster 5 as mundane chores, and cluster 6 as work-like activities.

One caveat to bear in mind is that average affect ratings are conditional on engaging in the activity for a given length of time. People probably sort into the activities that they engage in based, in part, on how much utility they derive from them. If the cluster analysis is redone using residuals of the six emotions after removing person effects, however, 83 percent of activities (weighted by frequency) remain in the same cluster as in the original assignment that did not remove person effects. Thus, the cluster analysis seems to provide a reasonably robust and plausible set of groups of activities that can be used to compare time use in over time or between countries.

7. Comparing Time Use Over Time in Groups of Activities and Generally

We propose three techniques for tracking time use over time: (1) following groups of activities defined in Section 6, (2) computing an overall U-index based on the U-index associated with various activities at a point in time; and (3) computing the U-index at the episode level. To illustrate the first two techniques, we used data from a project originally of the Yale University Program on Non-Market Accounts, known as the American Heritage Time Use Studies (AHTUS). The AHTUS consists of five time-use surveys conducted from 1965-66 through 2003. The disparate activity codes were harmonized to a common set of 72 main activities (plus missing/unclassified). In addition, we merged the harmonized activity codes to the 2005 ATUS and include it as well. The underlying sources of the harmonized data are described in the following box. Unfortunately, it is not possible to compute the episode-level U-index over time as PATS like data are not available in earlier years so we just illustrate the technique, but we hope that data will be available in the future for episode-level analyses.

Historical Time Use Surveys

● 1965-1966: Original source is Multinational Comparative Time-Budget Research Project conducted by the University of Michigan's Survey Research Center. N=1,968.

● 1975-1976: Original source is American's Use of Time: Time Use in Economic and Social Accounts, conducted by the University of Michigan's Survey Research Center. N=5,869.

● 1985: Original source is American's Use of Time, conducted by the University of Michigan's Survey Research Center. N=2,308.

● 1992-1994: Original source is National Human Activity Pattern Survey, conducted by the University of Maryland's Survey Research Center. N=5,964.

● 2003: Original source is ATUS, conducted by Census Bureau for Bureau of Labor Statistics. N=15,999.

● 2005: Original source is ATUS, conducted by Census Bureau for Bureau of Labor Statistics. N=10,112.

Sample weights were used for all estimates using the AHTUS data sets. Because we lack affect ratings during sleep, we focus on the waking day.[28] One issue that we can only partially address is that the data sets use different methods and sampling frames. For example, the 1965-66 survey sampled people from households in which someone was employed in a nonagricultural industry, and only covered certain months of the year. The samples were restricted to those from age 19 to 64 to have a consistent age range. The average age was fairly similar in the data sets, ranging from 38.4 in 1985 to 40.6 in 2003.

7.1. Tracking Groups of Activities

Tables 7.1A and 7.1B present the average proportion women's and men's awake time spent in the harmonized activities, respectively. A motivation of the cluster analysis was to classify these activities into affectively similar categories so that changes in time use could be tracked in a more manageable set of categories.

Specifically, for each person we first computed the average percent of the awake day spent in each of the six clusters described above. We next averaged over every individual in the sample.[29] Table 7.2A summarizes the results for men and women combined. The picture that emerges is one of stability for clusters 1 (unpleasant personal maintenance), 2 (moderately enjoyable tasks) and 6 (work-like activities). Time spent on cluster 4 (neutral downtime) is up while cluster 3 (engaging leisure) and cluster 5 (mundane

[28] Sleep rose from 7.95 hours in 1965-66 to 8.5 hours in 2005, or by 2.3 percentage points on a 24 hour day.
[29] Because a small number of activities (accounting for less than 3 percent of awake time each year) were not assigned to clusters in the PATS, they are omitted here. The percentages were renormalized to sum to 100 percent accordingly.

chores) are down. Overall, these figures suggest that affectively neutral downtime activities like watching television have gained at the expense of mundane chores and engaging leisure activities over the last 40 years.

Tables 7.2B and 7.2C report separate results for men and women, respectively. For men, the share of the day devoted to cluster 6 (work-like activities) has declined by 6 percentage points since 1965-66, while the share devoted to cluster 4 (neutral downtime) has increased by 8.5 points. Women, not surprisingly, have increased time in cluster 6 activities by 5 percentage points because of higher labor force participation, while time spent on mundane chores fell even more, by almost 7 points. The amount of time women spend in cluster 3 (engaging leisure) fell by roughly the same amount (3 points) as their time devoted to cluster 4 (neutral downtime) increased. These shifts, on balance, do not suggest significant improvements in affective experience for women over this entire 40 year time span.

7.2. Activity-based U-Index

In addition to classifying and tracking time use in categories, it is useful to summarize time allocation in a single welfare measure. The U-index can be used for this purpose. As before, the U-index measures the percent of moments spent in an unpleasant state during each activity, where an unpleasant state is defined as one where a negative emotion (sad, stress, or pain) strictly dominates the positive emotions (happy in this case).

Specifically, we first computed the U-index for each harmonized activity using the 2006 PATS data for a pooled sample of men and women. For example, the U-index during paid work was 27%, during exercise was 8%, and during television viewing was 18%. We next computed the weighted average U-index where the weights were the percent of awake time the average person spent in each activity. Formally, the weighted average U-index, denoted $\overline{\overline{U}}_t$, each year is:

$$\overline{\overline{U}}_t = \frac{\sum_i w_{it} (\sum_j p_{ijt} \overline{U}_j)}{\sum_i w_{it}} ,$$

where w_{it} is the sample weight for individual i, p_{ijt} is the proportion of time individual i spent in activity j in year t, and $\overline{U}_j$ is the U-index for activity j from the PATS.

Panel A of Table 7.3 reports the results. The activity-based U-Index shows very little trend over the last 40 years for men and women combined or for women as a group. For men, however, there has been a shift away from activities associated with unpleasant feelings. To put the estimates in context, note that the difference between the activity-based U-index on weekends and weekdays is about 3 percentage points.[30] Thus, the one

[30] With episode-level data, the weekend-weekday difference is about twice as large.

point drop in the U-index from 1965-66 to 2005 is about one third of the difference in unpleasant feelings associated with activities during the week and those on the weekend.

Although the U-index is highly correlated across activities for men and women, there are some notable differences in a small number of activities. Women, for example, find supervising/helping with homework and voluntary acts less unpleasant than do men. Thus, we computed the U-index separately for men and women. We then assigned the gender-specific U-index for each activity to each observation in the historical sample, and computed the activity-level U-index separately for men and women. Panel B of Table 7.3 and Figure 7.1 display the results, combining 2003 and 2005 for presentation. The results are generally consistent with those in Panel A, though they are a noisier. The gender-specific weighted U-index displays no trend for women and has trended downward for men over the last 40 years, indicating an improvement in daily experience.

Table 7.4 presents regressions to control for possible changes in the age and education composition of the samples, as well as the survey day and month. The unit of observation for the regressions is an individual. The dependent variable is the duration-weighted U-index for each person's activities on the survey day, or $\sum_j p_{ijt} \overline{U}_j$ where $\overline{U}_j$ is the U-index for activity j for men and women combined. The regression-adjusted estimates reveal a similar pattern: very little shift toward or away from unpleasant activities, on net, for women, but about a one percentage point shift away from activities associated with unpleasant feelings for men since the mid 1960s.

7.2.1 Dispersion in Activity-Level U-Index

The activity-level U-index masks some important trends across people and groups. The standard deviation of the activity-level U-index was calculated across people each year (see Figure 7.2). This measure of dispersion has grown by about 15 percent over the 40 year period. Thus, the spread in time people spend in activities according to their frequency of unpleasant moments is increasing over time.

Additionally, the U-index has declined by more for men with a high school degree or less schooling than it has for men with a college degree or higher (see Figure 7.3). This result is consistent with Aguiar and Hurst's (2007) finding that leisure time increased more for the less educated than highly educated, partially offsetting the rise in income associated with additional schooling.

7.3. Episode-Level U-Index

Table 5.1 provides what we refer to as episode-level estimates of the U-index for various groups. These are tabulations of the proportion of time spent in an unpleasant state where the episode is the unit of observation. The calculations do not require information on activities. If the nature of activities changes over time, the episode-level U-index will reflect this change. The episode-level U-index will also reflect the presence of others during the episode and other features of the episode. Moreover, if the U-index is calculated at the episode level, it allows for the fact that some people may respond emotionally to the same activity in different ways. Because activity and other measured features of episodes account for a small proportion of variability in affect – for example, controlling for 71 activity dummies only account 6 percent of the variability in reported happiness across episodes -- tracking changes over time in the episode-level U-index can be more informative than tracking how changes in activities are likely to affect well-being.

Unfortunately, an episode-level U-index – either for a representative national sample or for selected groups -- can only be calculated for 2006 because the PATS data set is cross-sectional. Nevertheless, the PATS data provide proof of the applicability of the idea and a baseline against which future measurements can be compared. If the affect questions are added to subsequent time-use surveys, such as ATUS, than the episode-level U-index can be computed at regular intervals in the future.

8. International Comparison[31]

In addition to comparing subjective well-being over time, social scientists and policymakers have long been interested in comparing SWB across countries. This interest partly stems from a desire to rank countries based on SWB. Additionally, cross-country data have been used to study the effect of various public policies, economic conditions and institutions (e.g., Blanchflower and Oswald, Alesina, Frey). The most common measures of SWB in these studies are reports of overall life satisfaction or happiness, which reflect global evaluations of one's life relative to some standard. In this section, we compare SWB in two "representative" cities, one in France and the other in the United States, and ask whether the standard measure of life satisfaction and the DRM yield the same conclusion concerning relative well-being. Specifically, we designed a survey to compare overall life satisfaction, time use, and recalled affective experience during episodes of the day for random samples of women in Rennes in France and Columbus, Ohio in the United States. These cities were selected because they represent "middle America" and "middle France". We also present results using time allocation derived from national samples in the United States and France to extend our analysis beyond these two cities. This comparison illustrates national time accounting in a cross-national context.

To preview the main results, based on the standard life satisfaction question, we find that Americans report higher levels of life satisfaction. Yet based on the DRM we find that the French spend their days in a more positive mood, on average. Moreover, the national time-use data indicate that the French spend relatively more of their time engaged in activities that tend to yield more pleasure than do Americans. Our results suggest that considerable caution is required in comparing standard life satisfaction data across populations with different cultures. In particular, the Americans seem to be more emphatic when reporting their well-being. The U-index apparently overcomes this inclination.

[31] This section is based on work that we did together with Claude Fischler.

8.1. Study Design

The sample consists of 810 women in Columbus, Ohio and 820 women in Rennes, France. They were invited to participate based on random-digit dialing in the Spring of 2005. Respondents were paid approximately $75 for their participation in both countries. The age range spanned 18 to 68, and all participants spoke the country's dominant language at home. The Columbus sample was older (median age of 44 versus 39), more likely to be employed (75 percent versus 67 percent) and better educated (average of 15.2 years of schooling years versus 14.0) than the Rennes sample. In addition, the Rennes sample was more likely to be currently enrolled in school (16 percent versus 10 percent). The differences in demographic characteristics partly reflect different circumstances in the countries (e.g., the employment rate is 8 percentage points higher than in the U.S. than in France, and average education is 0.9 years higher in the U.S.), and partly reflect idiosyncrasies of our two cities and sample. Because we compare SWB measured with different methods for the *same* samples, our results should reflect differences in the methods, not demographic differences between the samples.

Essentially the same protocols as those used in the Texas DRM were followed. Groups of participants were invited for a weekday evening to a central location, where they completed a series of questionnaires contained in separate packets. The first packet included general satisfaction and demographic questions. The life satisfaction question was taken from the World Values Survey. The second packet asked respondents to construct a diary of the previous day as a series of episodes, noting the content and the beginning and ending time of each.[32] The average number of episodes described was 13.2 in Columbus and 14.5 in Rennes.

In the third packet, respondents completed a form for *each* of the episodes they had previously listed. The form included a list of 22 activities and 8 interaction partners, with an instruction to mark all that apply. Respondents who had checked multiple activities were requested to indicate the one that "seemed the most important to you at the time" (we call it *focal*). Unless specifically noted, all analyses below refer to focal activities. The form also requested ratings of 10 emotions that were experienced at the time on a scale from 0 (Not at all) to 6 (Very Strongly). We focus on the following emotions: 'Happy', 'Tense/stressed', 'Depressed/blue', and 'Irritated/angry'. The questionnaire was translated back and forth between French and English to ensure common meanings, and some questions were modified and deleted as a result of this procedure.

The data were re-weighted by day of week to be representative of a random day. Weekdays received 5/7th of the weight and Saturday and Sunday received 1/7th of the

[32] About 300 participants in each country were recruited for Mondays to describe a weekend day. Half of them were instructed to describe the preceding Saturday and half the preceding Sunday. Data were not collected pertaining to Fridays.

weight in the weighted samples. Additional details of the procedures and all questionnaires are available on line.[33]

8.2. Life Satisfaction

Table 8.1 contains tabulations of reported life satisfaction in the two cities. As in most populations, reports of being very unsatisfied are rare. The American women, however, are twice as likely to say they are very satisfied with their lives as are the French women (26 percent versus 13 percent). Furthermore, assigning a number from 1 to 4 indicating life satisfaction, a common practice, also indicates that the Americans are more satisfied, on average, and the difference is statistically significant at the .05 level.

On further inspection, however, Table 8.1 provides less clear cut evidence that the Americans' responses exhibit higher life satisfaction. American respondents are over represented in both extremes, in both the very satisfied and the unsatisfied categories. If the top two categories on the satisfaction scale (very satisfied and satisfied) are combined, the French actually indicate higher life satisfaction: 83 percent versus 77 percent. Thus, it is unclear from these data whether the French are less satisfied or less prone to use the extreme ends of the scales. The propensity to express oneself in extremes can be influenced by cultural and social expectations. Cultural and social norms may discourage French women from reporting themselves as very satisfied compared with Americans.

8.3. Comparing SWB with the U-Index

The U-index is less susceptible to a tendency for the Americans to be more emphatic than the French as long as both apply their interpretation of the scales consistently to positive and negative emotions. To take an extreme example, suppose the French only use the 1-5 portion of the 0-6 scale, while the Americans utilize the full scale. Provided that the French use the 1-5 range consistently for reporting positive and negative emotions – i.e., an emotion reported as a 5 is always experienced more intensively than an emotion reported as a 4 – then, apart from integer concerns, the U-index is unaffected by this differential use of scales. As commonly applied, the standard life satisfaction measure is not robust to such reporting differences across people.

The first row of Table 8.2 reports the average episode-level U-index for the two samples. In this case, the U-index for an episode is defined as equal to 1 if the maximum rating of tense/stressed, depressed/blue, or irritated/angry strictly exceed the rating of happy, and 0 if not. The U-index was weighted by the proportion of each person's waking day spent in an episode to derive an overall estimate. In contrast to reported life satisfaction, the U-index is 2.8 percentage points lower in the French sample (16%) than in the American

[33]http://homepage.mac.com/WebObjects/FileSharing.woa/wa/default?user=dschkade&fpath=WB%20structure%20supplemental&templatefn=FileSharing1.html

sample (18.8%). Thus, the French appear to spend less of their time engaged in unpleasant activities (i.e., activities in which the dominant feeling is a negative one) than do the Americans in our samples.

We explored whether the lower U-index for the French is a result of any single negative emotion, or combinations of them. The lower U-index for the French appears to be a fairly robust result. If we required that at least two negative feelings were rated more strongly than happy, for example, the U-index was still 2.8 points lower in France than in the U.S. (10.1% versus 7.4%) And if we dropped any one of the negative emotions and compared the remaining two to happy, the U-index was lower in France than in the U.S. in each case. These results suggest that the lower U-index in France is not due to the rating of any particular negative emotion in our study.

The other rows of Table 8.2 provide comparisons of the episode-level U-index for various subpopulations. The general pattern is sensible. For example, the U-index in both countries is considerably lower on weekends than on weekdays. The French-American gap is largest for non-students, employed people, low-income people and during the week. Interestingly, in both countries -- but especially in the U.S. -- the U-index of the unemployed is much higher during the week than it is during weekends. This pattern suggests that observing others go to work during the week worsens the mood of the unemployed during weekdays.

There is greater inequality in the U-index across people in the American sample than in the French sample. Figure 8.1 displays the average U-index by quintile of the individual-level U-index distribution in each country. The average woman in Columbus in the top quintile of the distribution spent 57.5 percent of her time in an unpleasant state, while her counterpart in Rennes spent 49.0 percent of her time in an unpleasant state. Regression analysis indicated that the gap in the upper tail is only partially accounted for by independent variables such as the log of household income, a quadratic in age, school enrollment and day of week. Controlling for these variables reduced the U.S.-French gap in the upper quintile from 8.5 points to 5.3 points.

Another issue concerns vacations. In our sample, the French report taking 21 more vacation days than the Americans. We were not able to interview people if they were away from home, so we did not sample most vacation days. Accounting for vacations would almost certainly lower the U-index in France relative to the U.S. as vacation days are likely to have a lower U-index than non vacation days. The following back-of-the-envelope calculation suggests, however, that this is not a large bias. The 21 day difference in vacations amounts to only 5.8 percent of the year. If the U-index is 10 points lower on vacation days than non-vacation days, which is almost double the difference on weekdays and weekends, then the French U-index would be an additional 0.58 percentage points lower than the American U-index.

8.4. Counterfactual Cross-Country Comparisons: Activity Level Analysis

Table 8.3 presents the U-index for 21 activities and the proportion of the day the average person devoted to each activity based on the DRM. (These activities are different from those in some of our other DRMs because of translation issues.) If more than one activity was engaged in at a time, we selected the activity that was indicated by respondents as being most important at the time. Activities such as working, commuting, and childcare have a high U-index, and activities such as walking, making love, and exercising have a low U-index, similar to our earlier findings.

Both the pattern of time allocation and the U-index for each activity are similar in the two countries, with correlations of 0.93 and 0.85, respectively. The most notable exceptions to this pattern are that the Americans find childcare substantially more unpleasant than do the French, and the French spend less time engaged in childcare and more time eating. The latter is explained mainly by the fact that Americans are much less likely to indicate eating as their main activity when they engage in multiple activities that include eating. It is also worth noting that the French women in our sample are slightly less likely to have children living at home (56% versus 60%).

The data in Table 8.3 can be used to perform counterfactual calculations. Specifically, we can use the time allocation across activities for one country to weight the U-index for the other country and thus create a "synthetic" U-index. To be more precise, define the synthetic U-index using country j's time allocation ($\overline{\mathrm{H}}_{\mathrm{i}}^{\mathrm{j}}$) and country k's U-index ($\overline{\mathrm{U}}_{\mathrm{i}}^{\mathrm{k}}$) for activities denoted i as $U_{j,k} = \sum_i \overline{\mathrm{H}}_{\mathrm{i}}^{\mathrm{j}} \overline{\mathrm{U}}_{\mathrm{i}}^{\mathrm{k}}$. The "synthetic" U-index indicates how the average French woman, say, would feel if she experienced her activities in the same way as the average American woman. Table 8.4 reports the synthetic U-indexes for each country.[34]

The results indicate that if the French and American women's allocation of time is weighted by either the average American woman's rating of activities or the average French woman's rating of activities, the average French woman is predicted to have a lower synthetic U-index than the average American woman. But only about one third to 40 percent of the between-country difference in the U-index comes about because of differences in time allocation. Moreover, with small samples to compute time allocation, the difference in the synthetic U-index is not statistically significant regardless of which country's activity ratings are used.

We can calculate the synthetic U-indexes using larger samples of time allocation data from national time-use surveys, however. This provides a check on whether our results for Rennes and Columbus can be extended to the countries as a whole, and yields more precise estimates. Specifically, we analyzed national time use data on American women from the 2003-04 ATUS and on French women from the 1998-99 *Enquête Emploi du Temps* survey by INSEE. We restrict both samples to women age 18-60. Although the

[34] Notice that when the same country's time allocation and activity-level U-indexes are used the synthetic U-index is slightly different from the episode-level U-indexes reported in the first row of Table 8.2. This discrepancy arises because there is a weak correlation between time allocation and the U-index at the individual level.

French data are from an earlier time period, they are the most recent national data publicly available, and time allocation does not change very rapidly over time within countries. Because the activity categories in national time-use data are not harmonized, we collapsed the activities in these surveys into 6 broad categories: work; compulsory activities; active leisure; passive leisure; eating; and other. The U-index for these categories was computed from the DRM for Rennes and Columbus for the same activities.

Results are reported in Table 8.5. The national time allocations are generally similar to what we found for Rennes and Columbus. In particular, using national data the French women spend less time working, less time participating in passive leisure (e.g., watching TV), and more time participating in active leisure (e.g., exercise and reading) and eating than do the American women. As was found before, the French allocation of time produces a slightly higher synthetic U-index regardless of whether the American or the French U-index is used to rate each activity. Using either U-index to rate the activities, the French allocation of time produces about a 1 percentage point lower synthetic U-index. With the larger national time-use samples, the differences are statistically significant at the 0.10 level, although they are similar in magnitude to the differences reported in Table 8.4.

9. Conclusion

National Time Accounting provides a method for tracking time allocation and assessing whether people are experiencing their daily lives in more or less enjoyable ways. This paper demonstrates how NTA can be used to compare groups of individuals, countries and eras. Many economists argue that a decline in the amount of time spent working has been a major source of improvement in Americans' daily lives over the last century (Fogel, 1999). Shifts in time use among non-work activities also affect the experience of daily life. If non-work time increases in the next century as much as it did in the last century, it will be even more important to understand the experience of non-work time. Tracking the U-index over time, either at the episode level or at the activity level, provides a means for measuring whether daily life is becoming more or less pleasant, and of understanding why. To facilitate NTA in the future, we think that adding a module on affective experience to on-going time use surveys, such as ATUS, should be a priority.

Like the NIPAs, NTA is a descriptive technique, not prescriptive. The method of NTA does not lead to immediate policy recommendations. For example, the fact that spending time socializing may be more enjoyable than working for pay for the average person does not lead to the immediate recommendation that people should socialize more and work less. Paid work is obviously required to afford a certain life style. A similar limitation applies to the NIPAs: Although national income would be increased if all workers trained for higher paying professions, there are psychic and monetary costs that must be taken into account before making such a policy recommendation. To draw policy conclusions, we would recommend using the PATS or related instruments to measure outcomes of policy relevant experiments, such as the Moving to Opportunities public housing experiment.

Existing time use data sets provide several opportunities for additional applications of NTA. One possibility is to use the harmonized international time use data sets to compare how people in different countries devote time to various activities and to evaluate the activities by their average emotional experience according to the PATS. The clusters of activities identified in Section 6 would seem particularly appropriate for comparing time use across countries. Another possibility is to use existing time use data for the U.S. to study the effect of aging on the allocation of time across activities by following cohorts as they age. Again, the clusters of similar activities identified in Section 6 could facilitate the analysis.

Several extensions, unresolved issues and research issues concerning NTA should also be noted. First, although we based the emotions that we surveyed partly on the Russell circumplex and partly on practicality, the precise set of emotions could be tailored for the particular application at hand. For example, studies related to health and aging might focus on feelings of aches, pain, weariness, fatigue and disorientation. In addition, PATS might be adapted to measure people's sense of purpose about their daily routines. For example, people could be asked whether they considered their use of time during sampled episodes to be meaningful or a waste of time. If additional emotions are included, the robustness of the U-index to the set of surveyed emotions can be further explored, although some features of experience (e.g., meaningfulness) would seem to represent separate subjective components of well-being.

Another issue concerns the context of time use. That is, the precise situations that people are engaged in during their daily activities. Available time-use surveys collect only coarse information on the nature of activities. The fact that activity dummies account for such a small share of the variability in affective experience suggests that important features of activities are not measured by time-use surveys. Thus, tracking the change in activities over time weighted by the activity-level U-index (or some other activity-level measure of emotional experience) is susceptible to missing important changes in people's affective experiences because a great deal of what generates emotional experience occurs within a given set of measured activities.

A related issue is that the nature of some activities changes over time. For example, the experience of television viewing is likely to be quite different today than 40 years ago, when there were few channels, television sets were black and white, and Tivo was not available to skip over commercials. While changes in the nature of activities present a problem for all studies that track time-use over historical time, the problems are particular apparent for NTA. In some respects, the problem is akin to changes in product quality in the consumer price index. The prospect of tracking affective experience at the episode-level in the future, however, provides a way to avoid problems caused by changes in the nature of activities because it would not depend on the *a priori* assignment of activities. In addition, a time-series of episode-level data on affective experience would enable research into the changing hedonic nature of activities.

Finally, it is unclear how to fully integrate sleep and health into NTA. To some extent, both factors are reflected in our measures of affect. For example, people who are in poor health experience more pain during their daily lives (Krueger and Stone, 2007). And a bad night sleep is associated with a bad mood and greater tiredness throughout the day (Kahneman, et al., 2004). In other words, sleep and health both affect the process benefit of various uses of time. But if people learn to sleep half as much without lowering their average emotional experience during waking moments, our current summary measures would not credit an improvement in well-being. In addition, health surely has a direct effect on well-being independent of any effect on momentary emotional experience.

While these limitations of NTA are important, they are not insurmountable. We suspect that many of the current limitations of NTA are amenable to research, just as research helped to overcome some of the problems posed by changes in product quality in the NIPAs. Moreover, the choices that people make regarding their allocation of time, particularly labor supply, have long been subject to economic analysis. Research on the allocation and experience of non-work time is less developed, but no less important for economics and policy. Evaluated time use also strikes us as a fertile area for research because most determinants of subjective well-being are not well captured by data on market transactions, and this will be even more so in the future as people live longer and spend a smaller share of their lives engaged in market work and home production.

References (incomplete)

Aguiar, Mark, and Erik Hurst. 2007. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades." *Quarterly Journal of Economics* 122, no. 3 (August): 969-1006.

Ashby, F.G., Isen, A.M., & Turken, A.U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106,* 529-550.

Becker, Gary. 1965. "A Theory of the Allocation of Time." *Economic Journal* 75: 493-517.

Berk, L.S., Felten, D.L., Tan, S.A., Bittman, B.B., & Westengard, J. (2001). Modulation of neuroimmune parameters during the eustress of humor-associated mirthful laughter. *Alternative Therapies, 7*, 62-76.

Berkowitz, L. (1987). Mood, self-awareness, and willingness to help. *Journal of Personality and Social Psychology, 52,* 721–729.

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, *14*, 280 - 288.

Coghill, R.C., J.G. McHaffie, and Y. Yen. 2003. Neural correlates of interindividual differences in the subjective experience of pain. Proceedings of the National Academy of Sciences, 100, 8538-8542.

Cohen, S., Doyle, W. J., Turner, R. B., Alper, C. M., & Skoner, D. P. (2003). Emotional style and susceptibility to the common cold. *Psychosomatic Medicine, 65,* 652–657.

Connolly, Marie. 2007. "Some Like It Mild and Not Too Wet: The Influence of Weather on Subjective Well-Being," in *It's About Time: Three Essays on Time Use, Weather and Well-Being*, Dissertation, Princeton University.

Csikszentmihalyi, M., Patton, J. D., & Lucas, M. (1997). Le bonheur, l'experience optimale et les valeurs spirituelles: Une etude empirique aupres d'adolescents [Happiness, the optimal experience, and spiritual values: An empirical study of adolescents]. *Revue Quebecoise de Psychologie, 18,* 167–190.

Cunningham, M. R., Steinberg, J., & Grev, R. (1980). Wanting to and having to help: Separate motivations for positive mood and guiltinduced helping. *Journal of Personality and Social Psychology, 38,* 181–192.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology, 80,* 804–813.

Devins, G. M., Mann, J., Mandin, H. P., & Leonard, C. (1990). Psychosocial predictors of survival in end-stage renal disease. *Journal of Nervous and Mental Disease, 178,* 127–133.

Diener, E., Nickerson, C., Lucas, R. E., & Sandvik, E. (2002). Dispositional affect and job outcomes. *Social Indicators Research, 59,* 229–259.

Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E. Diener, & N. Schwarz (eds.), *Well-being:The foundations of hedonic psychology* (pp. 213-229). New York: Russel-Sage.

Dillon, K. M., Minchoff, B., & Baker, K. H. (1985). Positive emotional states and enhancement of the immune system. *International Journal of Psychiatry in Medicine, 15,* 13–18.

Dow, Gregory K. and F. Thomas Juster. 1985. "Goods, Time, and Well-being: The Joint Dependence Problem," in *Time, Goods, and Well-Being*, ed. F. Thomas Juster and Frank P. Stafford. Survey Research Center, Institute for Social Research, University, pp. 397-413.

Doyle, W. J., Gentile, D. A., & Cohen, S. (2006). Emotional style, nasal cytokines, and illness expression after experimental rhinovirus exposure. Brain, Behavior, and Immunity, 20, 175-181.

Ekman, P., & Rosenberg, E. L. (eds.) (1997). *What the face reveals.* New York: Oxford University Press.

Evans P, Bristow M, Hucklebridge F, Clow A, Walters N. (1993). The relationship between secretory immunity, mood and life-events. *British Journal of Clinical Psychology,32,* 227-236.

Fogel, Robert W. 1999. “Catching up with the economy,” *American Economic Review* 89, pp.1-21.

Fogel, Robert W. 2001. “Simon S. Kuznets” *Biographical Memoirs*, Vol 79, Washington D.C. : The National Academy Press.

Frey, Bruno and Alois Stutzer. 2002. Happiness & Economics, Princeton: Princeton University Press.

Gershuny, Jonathan and Brendan Halpin. 1996. “Time Use, Quality of Life and Process Benefits” in *In Pursuit of the Quality of Life*, Avner Offer (ed.), Oxford: Clarendon Press.

Graham, C., Eggers, A., & Sukhtankar, S. (2006). Does happiness pay? An exploration based on panel data from Russia. *Journal of Economic Behaviour and Organization, 55,* 319-342.

Gronau, Reuben. 1977. “Leisure, Home Production and Work: The Theory of the Allocation of Time Revisited.” *Journal of Political Economy* 85, no. 6: 1099-1123.

Gross, JJ, LL Carstensen, M. Pasupathi, J. Tasi, C.G. Skorpen and A.Y. Hsu. 1997. “Emotion and aging: experience, expression and control,” *Psychol Aging*, Dec. 12(4): 590-599.

Harker, L., & Keltner, D. (2001). Expressions of positive emotions in women’s college yearbook pictures and their relationship to personality and life outcomes across adulthood. *Journal of Personality and Social Psychology, 80,* 112–124.

Howell, R.T., Kern, M.L. & Lyubomirsky, S. (2007). Health benefits: Meta-analytically determining the impact of well-being on objective health indicators. Health Psychology Review, 1, 83-136.

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21,* 384–388.

Juster, F. Thomas. 1985. “Preferences for Work and Leisure,” in *Time, Goods, and Well-being*. F.T. Juster and F.P. Stafford, eds., Ann Arbor, MI: Institute for Social Research, University of Michigan, pp. 333-51.

Juster, F. T., & Stafford, F. P. (eds.) (1985). *Time, goods, and well-being*. Ann Arbor: Institute for Social Research.

Kahneman, Daniel, Alan Krueger, David Schkade, Norbert Schwarz, Arthur Stone. 2004. “A survey method for characterizing daily life experience: the Day Reconstruction Method.” *Science* 306: 1776-80.

Kahneman, Daniel, and Alan Krueger. 2006. “Developments in the Measurement of Subjective Well-Being.” *Journal of Economic Perspectives* 20, no. 21: pp. 3-24.

Kahneman, D., Wakker, P.P., & Sarin, R. (1997). "Back to Bentham?". *Quarterly Journal of Economics,* **112**, 375-405.

Kiecolt-Glaser, J., Fisher, L.D., Ogrocki, P., et al., (1988). Martial quality, marital disruption, and immune function. *Psychosomatic Medicine, 49*, 13-34.

Kirschbaum, C., Pirke, K.M. & Hellhammer, D.H. (1993). The “Trier Social Stress Test“ – A tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology, 28, 76-81.

Krueger, Alan. 2007. “Are we having more fun yet? Categorizing and evaluating changes in time allocation.” Forthcoming in *Brookings Papers on Economic Activity*, 2007.

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

Levy, S. M., Lee, J., Bagley, C., & Lippman, M. (1988). Survival hazard analysis in first recurrent breast cancer patients: Seven-year follow-up. *Psychosomatic Medicine, 50,* 520–528.

Lucas, R. E. (2001). Pleasant affect and sociability: Towards a comprehensive model of extraverted feelings and behaviors. *Dissertation Abstracts International, 61* (10-B), 5610. (UMI No. AAI9990068)

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2003). Reexamining adaptation and the set point model of happiness: Reactions to changes in marital status. *Journal of Personality and Social Psychology, 84,* 527–539.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131,* 803-855.

Maier, H., & Smith, J. (1999). Psychological predictors of mortality in old age. *Journal of Gerontology, 54B,* 44–54.

March, J. (1978). Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics*, 9, 587-608.

Marks, G. N., & Fleming, N. (1999). Influences and consequences of well-being among Australian young people: 1980–1995. *Social Indicators Research, 46,* 301–323.

Marsland, A. L., Pressman, S., & Cohen, S. (2007). Positive affect and immune function. In R. Ader (Ed.), Psychoneuroimmunology (4th ed., Vol. 2, pp. 761-779). San Diego, CA: Elsevier.

Michelson, William. 2005. *Time Use: Expanding Explanation in the Social Sciences*. Boulder, CO: Paradigm Publishers.

O’Malley, M. N., & Andrews, L. (1983). The effect of mood and incentives on helping: Are there some things money can’t buy? *Motivation and Emotion, 7,* 179–189.

Palmore, E. B. (1969). Predicting longevity: A follow-up controlling for age. *Gerontologist, 9,* 247–250.

Perera, S., Sabin, E., Nelson, P., & Lowe, D. (1998). Increases in salivary lysozyme and IgA concentrations and secretory rates independent of salivary flow rates following viewing of humorous videotape. *International Journal of Behavioral Medicine, 5,* 118-128.

Polk, D. E., Cohen, S., Doyle, W. J., Skoner, D. P., & Kirschbaum, C. (2005). State and trait affect as predictors of salivary cortisol in healthy adults. Psychoneuroendocrinology, 30, 261-272.
Ramey, Valerie, and Neville Francis. 2006. "A Century of Work and Leisure," NBER working paper 12264.

Robinson, John, and Geoffrey Godbey. 1997. *Time for Life: The Surprising Ways Americans Use Their Time*. University Park, PA: Pennsylvania State University Press.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*, 934-960.

Russell, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39: 1161-1178.

Schwarz, N. (2007). Retrospective and concurrent self-reports: The rationale for real-time data capture. In A. A. Stone, S. S. Shiffman, A. Atienza, & L. Nebeling (eds.), *The science of real-time data capture: Self-reports in health research* (pp. 11-26). New York: Oxford University Press.

Schwarz, N., Kahneman, D., & Xu, J. (in press). Global and episodic reports of hedonic experience. In R. Belli, D. Alwin, & F. Stafford (eds.), *Using calendar and diary methods in life events research* (pp. 00-00). Newbury Park, CA: Sage.

Shapiro, D., Jamner, L.D., & Goldstein, I.B. (1997). Daily mood state and ambulatory blood pressure. Psychophysiology, 34, 399-405.

Smyth, J. Ockenfels, M.C., Porter, L., Kirschbaum, C., Hellhammer, D.C., & Stone, A.A. (1998). Stressors and mood measured on a momentary basis are associated with salivary cortisol secretion. Psychoneuroendocrinology, 23, 353-370.

Snowdon, D. (2001). *Aging with grace.* New York: Bentam

Spanier, G. B., & Furstenberg, F. F. (1982). Remarriage after divorce: A longitudinal analysis of well-being. *Journal of Marriage and the Family, 44,* 709–720.

Stone, A.A. Event content in a daily survey differentially predicts mood. (1987). *Journal of Personality and Social Psychology, 52*, 56-58.

Stone, A.A., Marco, C., Cruise, C.E., Cox, D.S., & Neale, J.M. (1996). Are stress-induced immunological changes mediated by mood? A closer look at how both desirable and undesirable daily events influence sIgA antibody. *International Journal of Behavioral Medicine, 3,* 1-13.

Stone, A.A., Neale, J.M., Cox, D.S., Napoli, A., Valdimarsdottir, H., & Kennedy-Moore, E. (1994). Daily events are associated with a secretory immune response to an oral antigen in men. *Health Psychology, 13*, 440-446.

Stone, A.A., Cox, D.S., Valdimarsdottir, H., Jandorf, L., & Neale, J.M. (1987). Evidence that secretory IgA antibody is associated with daily mood. *Journal of Personality and Social Psychology, 52*, 988-993.

Stone, A. A., Schwartz, J. E., Schwarz, N., Schkade, D., Krueger, A., & Kahneman, D. (2006). A population approach to the study of emotion: Diurnal rhythms of a working day examined with the Day Reconstruction Method (DRM). *Emotion, 6,* 139-149.

Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). In D. Kahneman, E. Diener, & N. Schwarz (eds.), *Well-being: The foundations of hedonic psychology* (pp. 26-39). New York: Russell-Sage.

Sudman, S., & Wansink, B. (2002). *Consumer panels*. Chicago: American Marketing Association.

Ubel, P.A., Loewenstein, G., Schwarz, N., & Smith, D. (2005). Misimagining the unimaginable: The disability paradox and health care decision making. *Health Psychology, 24 (4, Suppl),* S57-S62.

Vitaliano, P. P., Scanlan, J. M., Ochs, H. D., Syrjala, K., Siegler, I. C., & Snyder, E. A. (1998). Psychosocial stress moderates the relationship of cancer history with natural killer cell activity. *Annals of Behavioral Medicine, 20,* 199–208.

Wheeler, R.E., Davidson, R.J. & Tomarken, A.J. (1993). Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. Psychophysiology, 30, 82-89.

Wong, Grace. 2007. “Be it Ever So Humble: Understanding Housing Using Subjective Well-being Data,” Working Paper, Wharton School, University of Pennsylvania.

Zuckerman, D. M., Kasl, S. V., & Ostfeld, A. M. (1984). Psychosocial predictors of mortality among the elderly poor. The role of religion, well-being, and social contacts. *American Journal of Epidemiology, 119,* 410–423.

Table 4.1: Rank of activities in terms of average enjoyment from DRM and general activity enjoyment question similar to Juster (1985)

Activity	DRM (Enjoy)	Juster Enjoy/ Dislike
Child care	9	2
Commuting from work	12	11
Commuting to work	13	13
Cooking	8	9
Dinner	3	3
Housework	10	12
Lunch	4	4
Phone at home	7	10
Relaxing	2	1
Socializing after work	1	7
Socializing at work	6	5
Watching TV	5	6
Working	11	8

Table 4.2. Comparison of PATS and DRM Average Happiness Rating (0-6) by Activity

Activity	PATS	DRM	Difference
Housework	3.77	4.10	-0.33
Commuting	3.80	3.84	-0.04
Working	3.82	3.74	0.08
Watching TV	3.91	4.32	-0.41
Computer	4.06	3.94	0.12
Shopping	4.11	4.00	0.11
Preparing food	4.25	4.27	-0.02
On the phone	4.47	4.00	0.47
Relaxing	4.49	4.55	-0.06
Eating	4.57	4.43	0.14
Childcare	4.59	4.06	0.53
Socializing	4.74	4.48	0.26
Prayer/worship	4.97	4.56	0.41
Exercising	5.09	4.77	0.32
Unwtd. Average	4.37	4.23	0.15

Notes: PATS sample is men and women combined.
DRM sample is random component of Texas survey.

Table 4.3. Average Response by Order of Affect Questions in PATS Sample

	Average					
	Happy	**Tired**	**Stressed**	**Sad**	**Interested**	**Pain**
First Question	4.35	2.31	1.37	0.71	4.34	0.89
Second	4.22	2.62	1.41	0.68	4.10	0.97
Third	4.19	2.67	1.62	0.69	3.90	0.98
Fourth	4.18	2.65	1.58	0.83	3.92	0.96
Fifth	3.88	2.67	1.49	0.70	4.10	1.03
Sixth	3.99	2.71	1.54	0.69	4.07	1.08
All	4.13	2.61	1.50	0.72	4.07	0.99

Notes: One of the following six different orderings was randomly selected for each respondent:
Order 1: Happy, Tired, Stressed, Sad, Interested, Pain
Order 2: Tired, Stressed, Sad, Interested, Pain, and Happy
Order 3: Stressed, Sad, Interested, Pain, Happy, and Tired
Order 4: Sad, Interested, Pain, Happy, Tired, and Stressed
Order 5: Interested, Pain, Happy, Tired, Stressed, and Sad
Order 6: Pain, Happy, Tired, Stressed, Sad, and Interested

Results are unweighted.

Table 4.4. Summary of Effects of Weather on Reported Well-Being in the PATS Survey

Variable	Life Satisfaction	Net Affect
Normal Rainfall	-	0
Rain on Interview Day	-	0
Rain on Diary Day	0	-/0
Normal Temperature	+	0
Temperature on Interview Day	0	NA
Temperature on Diary Day	NA	-

Notes: Connolly entered dummy variables for ranges of the rain and temperature variables in her regression analysis. A negative sign here indicates a negative and statistically significant effect of the climate measure, a positive sign indicates a positive and statistically significant effect of the climate measure, and NA indicates that the measure was not included in the particular analysis because of multicolinearity. Sample consists of women from PATS. The satisfaction regression also controlled for demographic variables (education, age, marital status, race and ethnicity). The net affect regression also controlled for activity dummies, month, day, state and demographic variables. See Tables 3.4, 3.12 and 3.16 of Connolly (2007) for the underlying estimates.

Table 5.1. U-Index for Various Demographic Groups, PATS Data

Sex	
Men	17.6%
Women	19.6
Race/Ethnicity	
White	17.5
Black	23.8
Hispanic	21.9
Household Income	
<$30,000	22.5
$30,000-$50,000	18.6
$50,000-$100,000	18.6
>$100,000	15.7
Education	
<High School	20.5
High School	21.3
Some College	19.6
College	15.6
Masters	16.6
Doctorate	11.3

	Men	**Women**
Age		
15-24	18.8%	18.9%
25-44	17.1	20.5
45-64	18.7	20.9
65+	15.6	16.1
Marital Status		
Married	17.4	17.4
Divorced/Separated	24.3	24.5
Widowed	20.2	22.3
Never Married	16.9	23.2
Cohabiting	17.3	23.3

Notes: U-index is proportion of time that rating of sad, stressed or pain exceeds happy.

Table 5.2A. U-Index and Selected Emotions by Activity

ATUS Activity Category	U-Index*	Happy	Stressed	Sad	Interested	Pain	No Epis
Religious	6.4%	4.97	0.90	0.66	5.09	0.61	
Sports and exercise	7.4%	5.08	0.84	0.25	4.92	1.20	
Eating and drinking	9.7%	4.57	1.11	0.52	4.03	0.80	
Relaxing and leisure	13.4%	4.34	1.08	0.70	4.55	0.91	
Socializing	13.5%	4.74	1.21	0.66	4.65	0.88	
Lawn and garden	14.2%	4.23	0.98	0.47	3.92	1.37	
Child care	15.6%	4.63	1.76	0.39	4.41	0.56	
Shopping	16.9%	4.11	1.42	0.45	4.04	0.85	
Volunteer	17.7%	4.22	1.40	0.61	4.86	0.57	
Watching TV	18.1%	3.91	1.17	0.82	3.97	0.94	
Food prep and clean-up	19.0%	4.02	1.58	0.62	3.62	1.07	
Travel	20.7%	4.05	1.69	0.59	3.46	0.81	
Telephone calls	23.5%	4.47	2.02	1.14	4.99	0.86	
Personal care	23.6%	4.02	1.83	0.91	3.32	1.30	
Housework	24.0%	3.55	1.46	0.61	3.16	1.02	
Working	26.9%	3.80	2.37	0.69	3.99	0.71	
Household management	27.9%	3.50	1.85	0.82	3.94	0.76	
Medical care	29.0%	3.64	2.50	0.75	4.06	1.66	
Education	32.3%	3.62	2.66	0.87	3.87	0.82	
Adult care	33.8%	3.54	1.89	1.46	3.63	1.34	
All	**18.6%**	**4.13**	**1.53**	**0.66**	**4.03**	**0.88**	**1**

Notes: U-index indicates the proportion of 15-minute intervals in which stressed, sad, or pain exceeded ha
Source: Authors calculations based on PATS.

Table 5.2B. U-Index and Selected Emotions by Activity after Removing Individual Fixed-Effects

ATUS Activity Category	U-Index*	Happy	Stressed	Sad	Interested	Pain		No. (Episo
Religious	8.3%	4.81	0.94	0.83	5.14	0.88		
Eating and drinking	10.7%	4.49	1.14	0.55	3.99	0.78		1
Sports and exercise	11.9%	4.89	1.22	0.48	4.87	1.48		
Socializing	13.0%	4.68	1.21	0.59	4.65	0.84		
Child care	13.6%	4.59	1.44	0.49	4.49	0.65		
Relaxing and leisure	15.1%	4.35	1.24	0.68	4.49	0.88		1
Watching TV	15.7%	4.00	1.16	0.71	4.01	0.77		1
Lawn and garden	16.7%	4.21	1.21	0.55	3.92	1.25		
Personal care	17.4%	4.07	1.47	0.60	3.20	0.96		
Food prep and clean-up	17.6%	4.02	1.42	0.51	3.39	0.92		
Shopping	18.0%	4.15	1.68	0.63	4.01	0.92		
Travel	19.8%	4.06	1.62	0.63	3.64	0.89		1
Telephone calls	20.4%	4.50	1.73	0.94	5.14	0.84		
Volunteering	20.7%	4.28	1.72	0.81	4.71	0.96		
Medical care	22.6%	3.76	2.20	0.83	4.52	1.22		
Housework	25.6%	3.56	1.57	0.68	3.11	1.08		
Household management	27.4%	3.70	1.68	0.78	4.00	0.76		
Education	28.7%	3.55	2.39	0.90	4.09	0.80		
Working	28.8%	3.83	2.34	0.78	4.09	0.89		1
Adult care	32.0%	3.50	1.79	1.15	3.37	1.23		
All	**18.6%**	**4.13**	**1.53**	**0.66**	**4.03**	**0.89**		**11**

Notes: U-index indicates the proportion of 15-minute intervals in which stressed, sad, or pain exceeded hap
Source: Authors calculations based on PATS.

Table 5.3 U-Index by Whom With

	Men	Women	p-value for difference
Alone	18.3	21.9	0.033
Spouse	15.8	15.3	0.808
Children	10.2	17.7	0.034
Parents	7.2	27.1	0.025
Friends	11.8	12.8	0.792
Co-Workers	25.9	27.5	0.615
Boss/Supervisor	46.9	30.5	0.522

Table 5.4. U-Index by Day of Week

Monday	21.7%
Tuesday	19.0
Wednesday	20.9
Thursday	20.1
Friday	16.8
Saturday	17.7
Sunday	13.7

Table 6.1. Clusters assigned based on six emotions, 2006 PATS

Activity	Cluster	Net Affect	Happy	Tired	Stress	Sad	Interested	Pain	No. of Episodes
personal medical care	1	0.21	2.34	3.69	2.21	1.06	2.70	3.10	24
financial/government services	1	0.32	2.87	3.19	3.40	1.86	3.34	1.92	20
homework	1	0.80	2.71	3.08	3.32	0.94	3.08	1.47	43
purchase medical services	1	2.08	3.67	2.77	2.51	0.74	4.08	1.63	80
writing by hand	2	2.79	3.46	1.97	0.96	0.52	3.69	0.53	34
purchase routine goods	2	3.08	4.03	2.29	1.46	0.52	3.96	0.88	218
other child care	2	3.08	3.93	2.43	1.32	0.48	3.79	0.73	30
use computer	2	3.24	3.99	2.17	1.16	0.55	4.52	0.55	240
second job, other paid work	2	3.40	4.39	2.49	1.42	0.66	4.48	0.90	67
other meals & snacks	2	3.61	4.47	2.42	1.15	0.58	3.91	0.83	971
walking	2	3.95	4.66	1.56	0.64	0.27	4.21	1.22	56
general voluntary acts	3	3.36	4.22	2.41	1.40	0.61	4.86	0.57	53
conversation, phone, texting	3	3.42	4.55	2.44	1.50	0.93	4.61	0.98	377
read books	3	3.49	4.36	2.35	0.94	0.83	4.81	0.87	474
receive or visit friends	3	3.79	4.71	2.71	1.25	0.59	4.77	0.90	187
read to/with, talk with children	3	3.92	4.73	2.61	1.45	0.39	4.72	0.58	35
travel related to consumption	3	4.04	5.02	2.87	1.86	0.51	4.23	0.55	18
other in-home social, games	3	4.08	4.77	2.23	1.04	0.25	4.92	0.78	121
pet care, walk dogs	3	4.14	4.91	2.89	1.06	0.49	4.51	0.75	104
worship and religious acts	3	4.24	4.97	1.70	0.90	0.66	5.09	0.61	151
sports & exercise	3	4.26	5.09	2.87	0.89	0.25	4.97	1.34	208
café, bar	3	4.39	5.00	2.24	0.88	0.29	4.59	0.66	255
general out-of-home leisure	3	4.39	4.91	1.91	0.46	0.38	4.49	0.69	29
purchase personal services	3	4.43	5.06	2.08	0.69	0.16	4.33	1.05	22
parties or receptions	3	4.72	5.24	2.04	0.88	0.29	5.00	0.38	90
hunting, fishing, boating, hiking	3	4.73	5.32	1.91	0.74	0.36	5.26	0.68	30
attend sporting event	3	4.74	5.24	1.73	0.78	0.04	4.97	0.69	21
play with children	3	4.81	5.41	2.49	0.74	0.21	4.69	0.86	40
listen to music (cd etc.)	3	4.81	5.33	1.56	0.38	0.35	5.06	0.84	22
watch television, video	4	2.94	3.91	2.94	1.17	0.82	3.97	0.94	1946
food preparation, cooking	4	3.14	4.25	2.65	1.63	0.60	3.91	1.11	452
relax, think, do nothing	4	3.25	4.40	2.77	1.31	0.80	3.96	1.34	313
gardening	4	3.34	4.26	2.79	0.92	0.43	3.88	1.41	306
set table, wash/put away dishes	5	2.28	3.32	2.81	1.45	0.68	2.76	0.93	145
laundry, ironing, clothing repair	5	2.46	3.33	2.28	1.11	0.61	2.73	0.94	187
adult care	5	2.56	3.90	2.56	1.72	1.19	3.82	1.10	87
Cleaning	5	2.63	3.72	2.85	1.61	0.62	3.54	1.05	327
other domestic work	5	2.63	3.76	2.59	1.85	0.66	3.87	0.90	368
travel related to leisure/other	5	3.00	4.02	2.73	1.66	0.57	3.43	0.79	1120
wash, dress, personal care	5	3.11	4.31	3.16	1.78	0.77	3.39	1.02	140
home repairs, maintain vehicle	6	2.22	3.50	2.76	1.97	0.85	3.95	1.03	89
paid work at home	6	2.35	3.47	2.66	2.01	0.63	4.00	0.71	207
regular schooling, education	6	2.42	3.77	3.73	2.69	0.89	4.01	0.48	70
main paid work (not at home)	6	2.55	3.83	2.72	2.44	0.69	3.98	0.71	1425
general care of older children	6	3.55	4.54	3.41	1.98	0.45	4.36	0.54	235

Table 6.2. Average of emotions by cluster

Cluster	Happy	Tired	Stressed	Sad	Interested	Pain	Net Affect
1	3.09	2.97	2.92	1.18	3.57	1.80	1.12
2	4.29	2.31	1.18	0.55	4.06	0.78	3.45
3	4.79	2.37	1.05	0.56	4.79	0.84	3.97
4	4.05	2.87	1.23	0.76	3.95	1.06	3.04
5	3.86	2.72	1.64	0.63	3.44	0.89	2.80
6	3.88	2.83	2.35	0.69	4.04	0.69	2.63

Note: averages are weighted by episode frequency and sample weights. All emotions are reported on a 0 to 6 scale. Based on July 5, 2007 cluster6_freqwgt_ctus_best.log

Table 7.1A. Percentage of women's days spent in each activity, 1965-66 to 2005

	Main Activity	1965-66	1975-76	1985	1992-94	2003	2005
1	general or other personal care	1.52%	0.20%	0.79%	0.32%	0.25%	0.09%
2	wash, dress, personal care	5.80%	4.90%	6.67%	5.84%	5.22%	4.96%
3	personal medical care	0.06%	0.11%	0.04%	0.06%	0.44%	0.64%
4	meals at work	0.74%	0.69%	0.72%	0.00%	0.05%	0.03%
5	other meals & snacks	7.09%	7.83%	7.32%	6.88%	5.27%	5.51%
6	main paid work (not at home)	14.32%	14.07%	15.83%	21.10%	19.51%	19.13%
7	paid work at home	0.62%	0.56%	1.36%	0.81%	1.36%	1.28%
8	second job, other paid work	0.14%	0.17%	0.26%	0.01%	0.64%	0.62%
9	work breaks	0.51%	0.34%	0.18%	0.06%	0.02%	0.02%
10	other time at workplace	0.23%	0.19%	0.16%	0.00%	0.00%	0.00%
11	time looking for work	0.00%	0.08%	0.08%	0.06%	0.18%	0.14%
12	regular schooling, education	0.19%	0.30%	0.33%	1.01%	0.61%	0.43%
13	homework	0.30%	0.42%	0.48%	0.77%	0.79%	0.70%
14	short course or training	0.21%	0.20%	0.28%	0.04%	0.06%	0.21%
15	other education or training	0.72%	0.03%	0.16%	0.09%	0.02%	0.02%
16	food preparation, cooking	7.46%	7.08%	5.77%	4.09%	3.74%	3.77%
17	set table, wash/put away dishes	3.71%	2.26%	1.87%	0.68%	1.23%	1.22%
18	cleaning	5.94%	5.76%	4.52%	4.79%	3.97%	4.58%
19	laundry, ironing, clothing repair	4.43%	2.45%	1.99%	1.58%	2.21%	2.37%
20	home repairs, maintain vehicle	0.30%	0.60%	0.40%	0.39%	0.32%	0.28%
21	other domestic work	1.58%	0.59%	1.49%	1.40%	1.26%	1.24%
22	purchase routine goods	1.90%	2.94%	3.10%	0.93%	3.35%	3.31%
23	purchase consumer durables	0.14%	0.12%	0.08%	2.60%	0.01%	0.02%
24	purchase personal services	0.27%	0.26%	0.16%	0.18%	0.26%	0.19%
25	purchase medical services	0.13%	0.25%	0.30%	0.37%	0.43%	0.33%
26	purchase repair, laundry services	0.33%	0.16%	0.10%	0.09%	0.12%	0.11%
27	financial/government services	0.06%	0.14%	0.20%	0.12%	0.09%	0.10%
28	purchase other services	1.52%	0.10%	0.19%	0.10%	0.06%	0.06%
29	general care of older children	3.47%	2.36%	2.23%	1.44%	2.60%	2.37%
30	medical care of children	0.09%	0.12%	0.07%	0.02%	0.16%	0.17%
31	play with children	0.32%	0.30%	0.41%	0.33%	0.87%	0.81%
32	supervise/help with homework	0.25%	0.13%	0.16%	0.18%	0.52%	0.45%
33	read to/with, talk with children	0.24%	0.36%	0.18%	0.06%	0.38%	0.43%
34	other child care	0.30%	0.57%	0.23%	0.43%	0.54%	0.53%
35	adult care	0.67%	1.10%	0.51%	0.51%	1.65%	1.35%
36	general voluntary acts	0.45%	0.29%	0.43%	0.05%	0.91%	0.78%
37	political and civic activity	0.09%	0.04%	0.01%	0.00%	0.02%	0.00%
38	worship and religious acts	0.95%	1.09%	0.84%	1.02%	0.98%	0.89%
39	general out-of-home leisure	0.16%	0.18%	0.16%	0.00%	0.19%	0.21%
40	attend sporting event	0.11%	0.26%	0.28%	0.31%	0.22%	0.16%
41	theater, concert, opera	0.02%	0.09%	0.06%	0.14%	0.11%	0.08%
42	museums, exhibitions	0.01%	0.04%	0.01%	0.06%	0.06%	0.05%
43	café, bar	0.11%	0.27%	0.49%	0.30%	1.63%	1.44%
44	parties or receptions	1.54%	0.55%	0.55%	0.69%	0.68%	0.61%

45	sports & exercise	0.34%	0.60%	0.98%	1.50%	0.90%	0.84%
46	walking	0.10%	0.13%	0.25%	0.00%	0.31%	0.26%
47	cycling	0.00%	0.03%	0.02%	0.00%	0.03%	0.02%
48	physical activity/sports with child	0.05%	0.13%	0.15%	0.10%	0.02%	0.04%
49	hunting, fishing, boating, hiking	0.08%	0.21%	0.25%	0.00%	0.08%	0.10%
50	gardening	0.27%	0.55%	0.36%	0.26%	0.82%	0.80%
51	pet care, walk dogs	0.13%	0.37%	0.57%	0.44%	0.60%	0.65%
52	receive or visit friends	4.97%	4.78%	2.94%	4.01%	4.62%	1.81%
53	other in-home social, games	0.46%	0.69%	0.71%	0.56%	0.58%	0.80%
54	artistic activity	0.07%	0.15%	0.11%	0.09%	0.02%	0.02%
55	crafts	1.24%	1.44%	0.76%	0.55%	0.11%	0.17%
56	hobbies	0.04%	0.04%	0.02%	0.03%	0.02%	0.03%
57	relax, think, do nothing	0.59%	1.16%	0.74%	1.81%	1.77%	1.69%
58	read books	3.02%	2.97%	2.68%	2.44%	1.96%	2.15%
59	listen to music (cd, etc.)	0.08%	0.20%	0.08%	0.04%	0.10%	0.07%
60	listen to radio	0.28%	0.19%	0.23%	0.11%	0.07%	0.11%
61	watch television, video	8.47%	12.74%	13.02%	14.87%	13.60%	14.68%
62	writing by hand	0.74%	0.23%	0.39%	0.72%	0.19%	0.15%
63	conversation, phone, texting	1.60%	2.20%	3.37%	1.42%	0.92%	3.45%
64	use computer	0.00%	0.00%	0.08%	0.26%	0.89%	1.00%
65	imputed travel	0.00%	0.05%	0.00%	0.00%	0.33%	0.03%
66	travel related to personal care	0.71%	0.96%	0.86%	1.76%	1.56%	0.97%
67	travel related to work	1.35%	1.37%	1.97%	2.26%	1.68%	1.66%
68	travel related to education	0.11%	0.13%	0.22%	0.23%	0.13%	0.11%
69	travel related to consumption	2.13%	2.06%	2.33%	2.22%	2.50%	1.26%
70	travel related to child care	0.55%	0.53%	0.53%	0.36%	0.77%	0.72%
71	travel related to voluteering/worship	0.39%	0.91%	0.67%	0.37%	0.27%	0.26%
72	travel related to leisure	1.89%	1.87%	2.04%	2.00%	1.71%	1.56%
73	missing/unclassified	1.34%	2.79%	2.18%	1.66%	0.47%	2.92%

Table 7.1B. Percentage of men's days spent in each activity, 1965-66 to 2005

	Main Activity	1965-66	1975-76	1985	1992-94	2003	2005
1	general or other personal care	0.93%	0.19%	0.74%	0.34%	0.25%	0.17%
2	wash, dress, personal care	4.60%	4.04%	4.93%	4.10%	3.67%	3.51%
3	personal medical care	0.06%	0.04%	0.02%	0.04%	0.31%	0.60%
4	meals at work	1.55%	1.18%	0.90%	0.00%	0.05%	0.06%
5	other meals & snacks	7.49%	8.42%	7.63%	7.13%	5.55%	5.93%
6	main paid work (not at home)	34.98%	30.28%	25.57%	29.27%	28.44%	27.41%
7	paid work at home	0.97%	1.76%	2.62%	1.23%	1.54%	1.89%
8	second job, other paid work	0.96%	0.71%	0.54%	0.06%	1.00%	0.96%
9	work breaks	1.16%	0.60%	0.27%	0.08%	0.03%	0.03%
10	other time at workplace	0.68%	0.40%	0.35%	0.00%	0.00%	0.00%
11	time looking for work	0.00%	0.16%	0.12%	0.10%	0.30%	0.15%
12	regular schooling, education	0.32%	0.67%	0.64%	1.23%	0.64%	0.50%
13	homework	0.73%	0.76%	0.93%	0.93%	0.68%	0.90%
14	short course or training	0.26%	0.25%	0.20%	0.03%	0.03%	0.09%
15	other education or training	0.29%	0.09%	0.12%	0.07%	0.04%	0.00%
16	food preparation, cooking	0.84%	1.03%	1.44%	1.52%	1.42%	1.42%
17	set table, wash/put away dishes	0.35%	0.22%	0.38%	0.14%	0.33%	0.30%
18	cleaning	0.94%	1.79%	2.13%	2.54%	1.88%	1.89%
19	laundry, ironing, clothing repair	0.11%	0.10%	0.26%	0.30%	0.42%	0.45%
20	home repairs, maintain vehicle	0.99%	1.75%	1.80%	1.64%	1.49%	1.47%
21	other domestic work	0.79%	0.72%	1.35%	1.13%	0.88%	0.84%
22	purchase routine goods	1.05%	1.31%	1.69%	0.44%	2.17%	1.95%
23	purchase consumer durables	0.18%	0.15%	0.10%	1.24%	0.03%	0.01%
24	purchase personal services	0.09%	0.05%	0.06%	0.04%	0.06%	0.06%
25	purchase medical services	0.17%	0.14%	0.19%	0.21%	0.24%	0.28%
26	purchase repair, laundry services	0.25%	0.13%	0.15%	0.18%	0.13%	0.11%
27	financial/government services	0.04%	0.13%	0.16%	0.10%	0.08%	0.07%
28	purchase other services	1.02%	0.11%	0.23%	0.10%	0.05%	0.04%
29	general care of older children	0.40%	0.48%	0.38%	0.25%	0.83%	0.84%
30	medical care of children	0.00%	0.02%	0.01%	0.00%	0.05%	0.01%
31	play with children	0.46%	0.17%	0.23%	0.20%	0.60%	0.54%
32	supervise/help with homework	0.08%	0.05%	0.04%	0.05%	0.23%	0.17%
33	read to/with, talk with children	0.06%	0.11%	0.08%	0.07%	0.12%	0.12%
34	other child care	0.11%	0.13%	0.06%	0.15%	0.25%	0.25%
35	adult care	0.47%	0.91%	0.54%	0.40%	1.22%	1.13%
36	general voluntary acts	0.21%	0.24%	0.26%	0.10%	0.72%	0.67%
37	political and civic activity	0.10%	0.02%	0.00%	0.03%	0.00%	0.05%
38	worship and religious acts	0.59%	0.76%	0.54%	0.65%	0.74%	0.57%
39	general out-of-home leisure	0.03%	0.08%	0.19%	0.00%	0.22%	0.17%
40	attend sporting event	0.14%	0.30%	0.28%	0.40%	0.26%	0.29%
41	theater, concert, opera	0.05%	0.08%	0.09%	0.06%	0.09%	0.16%
42	museums, exhibitions	0.02%	0.05%	0.03%	0.03%	0.06%	0.01%
43	café, bar	0.66%	0.48%	0.83%	0.78%	1.67%	1.65%
44	parties or receptions	1.40%	0.59%	0.61%	0.61%	0.62%	0.52%
45	sports & exercise	0.72%	1.24%	1.75%	2.21%	1.39%	1.36%

46	walking	0.16%	0.19%	0.26%	0.00%	0.23%	0.22%
47	cycling	0.00%	0.03%	0.03%	0.00%	0.05%	0.07%
48	physical activity/sports with child	0.04%	0.07%	0.10%	0.04%	0.04%	0.07%
49	hunting, fishing, boating, hiking	0.52%	0.63%	0.99%	0.00%	0.53%	0.50%
50	gardening	0.16%	0.38%	0.61%	0.33%	1.39%	1.64%
51	pet care, walk dogs	0.06%	0.34%	0.52%	0.40%	0.45%	0.47%
52	Receive or visit friends	3.29%	3.36%	2.50%	3.60%	3.86%	1.63%
53	other in-home social, games	0.54%	0.52%	0.51%	0.51%	1.00%	1.06%
54	Artistic activity	0.11%	0.05%	0.09%	0.03%	0.02%	0.00%
55	Crafts	0.01%	0.22%	0.03%	0.04%	0.18%	0.13%
56	hobbies	0.28%	0.32%	0.30%	0.04%	0.04%	0.06%
57	relax, think, do nothing	0.31%	1.21%	0.77%	1.74%	1.75%	1.93%
58	read books	3.46%	2.61%	2.42%	2.44%	1.55%	1.44%
59	listen to music (cd, etc.)	0.10%	0.42%	0.13%	0.08%	0.26%	0.32%
60	listen to radio	0.44%	0.28%	0.33%	0.24%	0.12%	0.13%
61	watch television, video	11.21%	12.77%	14.55%	16.41%	16.08%	17.25%
62	writing by hand	0.27%	0.12%	0.23%	0.60%	0.12%	0.11%
63	conversation, phone, texting	0.99%	1.53%	2.05%	0.73%	0.44%	2.69%
64	use computer	0.00%	0.00%	0.17%	0.58%	1.24%	1.25%
65	imputed travel	0.00%	0.04%	0.01%	0.00%	0.24%	0.03%
66	travel related to care	0.97%	1.48%	1.08%	1.83%	1.66%	1.09%
67	travel related to work	3.68%	3.19%	3.45%	3.35%	2.86%	2.69%
68	travel related to education	0.19%	0.27%	0.17%	0.22%	0.15%	0.09%
69	travel related to consumption	1.63%	1.41%	1.86%	1.59%	2.12%	0.95%
70	travel related to child care	0.28%	0.21%	0.23%	0.11%	0.32%	0.26%
71	travel related to voluteering/worship	0.37%	0.81%	0.62%	0.35%	0.24%	0.18%
72	travel related to other purposes	2.06%	1.97%	2.58%	2.35%	1.79%	1.71%
73	missing/unclassified	1.60%	2.67%	2.00%	2.23%	0.47%	2.47%

Table 7.2. Average percent of day by cluster, 1965-66 to 2005

Panel A: All

Cluster	1965-66	1974-75	1985	1992-94	2003	2005
1	4.2%	3.6%	3.9%	5.8%	4.4%	3.8%
2	10.7%	12.1%	11.8%	9.5%	11.1%	11.5%
3	19.8%	19.6%	19.0%	16.5%	18.3%	17.1%
4	16.3%	20.3%	20.1%	21.2%	20.6%	22.3%
5	17.6%	15.2%	16.3%	14.6%	14.0%	14.1%
6	31.4%	29.2%	28.9%	32.4%	31.6%	31.2%

Panel B: Men

Cluster	1965-66	1974-75	1985	1992-94	2003	2005
1	4.5%	4.0%	4.2%	5.0%	3.9%	3.6%
2	10.7%	11.5%	11.2%	9.4%	10.8%	11.1%
3	18.2%	17.5%	17.8%	15.5%	17.4%	16.1%
4	14.5%	17.3%	18.8%	20.7%	20.9%	23.0%
5	9.7%	10.2%	12.6%	11.4%	10.4%	10.2%
6	42.4%	39.5%	35.4%	38.0%	36.5%	36.0%

Panel C: Women

Cluster	1965-66	1974-75	1985	1992-94	2003	2005
1	4.0%	3.2%	3.6%	6.5%	4.9%	3.9%
2	10.7%	12.5%	12.3%	9.6%	11.3%	11.9%
3	21.2%	21.5%	20.2%	17.3%	19.2%	18.1%
4	17.9%	23.0%	21.3%	21.6%	20.2%	21.7%
5	24.7%	19.6%	19.6%	17.2%	17.5%	17.9%
6	21.5%	20.1%	23.0%	27.8%	26.9%	26.5%

Table 7.3. U-Index based on time in various activities each year

A. U-Index from Men and Women Combined

	1965-66	1975-76	1985	1992-94	2003	2005
All	20.1%	19.5%	19.5%	20.0%	19.3%	19.6%
Men	20.9	20.4	20.1	20.2	19.6	19.9
Women	19.4	18.7	19.0	19.8	19.2	19.4

B. Gender-Specific U-Indexes and Time Allocation

	1965-66	1975-76	1985	1992-94	2003	2005
Men	20.2	20.1	19.2	18.8	18.7	19.0
Women	20.8	19.4	20.0	21.0	20.1	20.4

Note: A small number of missing and unclassified activities were assigned the mean U-index each year.

Table 7.4. Regression Models for Activity-based U-Index
Dependent variable is the duration-weighted average U-index

	All		Men		Women	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Intercept	20.905	0.224	21.108	0.356	19.862	0.279
Year=1975-76	-0.518	0.074	-0.338	0.118	-0.689	0.094
Year=1985	-0.544	0.070	-0.731	0.111	-0.363	0.088
Year=1992-94	-0.031	0.071	-0.677	0.113	0.551	0.089
Year=2003	-0.682	0.070	-1.255	0.110	-0.130	0.090
Year=2005	-0.409	0.070	-0.950	0.109	0.110	0.089
Tuesday	-0.137	0.071	-0.122	0.113	-0.149	0.090
Wednesday	0.007	0.071	0.035	0.113	-0.023	0.090
Thursday	-0.194	0.071	-0.049	0.112	-0.325	0.090
Friday	-0.513	0.071	-0.553	0.112	-0.474	0.090
Saturday	-2.231	0.071	-2.599	0.113	-1.893	0.090
Sunday	-3.018	0.072	-3.431	0.113	-2.645	0.090
February	0.022	0.089	-0.128	0.140	0.158	0.113
March	0.203	0.092	-0.072	0.146	0.451	0.115
April	0.056	0.095	-0.179	0.149	0.243	0.121
May	-0.118	0.093	-0.272	0.146	0.004	0.117
June	-0.146	0.089	-0.302	0.142	-0.018	0.112
July	-0.406	0.111	-0.351	0.177	-0.470	0.139
August	-0.405	0.107	-0.473	0.171	-0.363	0.134
September	-0.018	0.096	-0.221	0.152	0.177	0.121
October	0.088	0.095	0.028	0.150	0.109	0.120
November	0.142	0.087	-0.031	0.140	0.313	0.109
December	0.102	0.089	0.082	0.140	0.092	0.113
Age	0.036	0.011	0.054	0.017	0.018	0.013
Age-Squared	-0.001	0.000	-0.001	0.000	0.000	0.000
Female	-0.921	0.038	---		---	
< HS	-0.048	0.059	-0.025	0.093	-0.113	0.074
Some College	0.438	0.052	0.511	0.084	0.329	0.066
College	0.152	0.056	0.103	0.087	0.142	0.072
> College	0.009	0.075	-0.006	0.112	-0.054	0.099
R-Square	0.104		0.115		0.084	
Sample Size	40,388		17,921		22,467	

Notes: Regressions are estimated by weighted least squares. Person weights have been normalized to sum to one in each sample. Weighted mean (and standard deviation) of the dependent variable is 19.7% (4.0) for all, 20.1% (4.3) for men and 19.3% (3.8) for women. All explanatory variables are dummy variables except age and age-squared. Base year is 1965-66.

Table 8.1. Distribution of reported life satisfaction in Columbus, OH and Rennes, France. Life satisfaction is based on the question, "Taking all things together, how satisfied are you with your life as a whole these days?" Sample size is 810 women for Columbus and 816 women for Rennes. Chi-square test of identical distributions rejects at $p < 0.001$

	U.S.	France
Not at all satisfied	1.6%	1.1%
Not very satisfied	21.4	16.1
Satisfied	51.0	70.0
Very Satisfied	26.1	12.9

Table 8.2. U-index for various groups in Columbus, OH and Rennes, France. U-index is computed as proportion of time in which the rating of the maximum of tense, blue and angry is strictly greater than the rating of happy. P-values are for test of country differences for each group.

Group	**U.S.**	**France**	**Difference**
All	0.188	0.160	0.028**
Enrollment Status			
Non-Student	0.181	0.144	0.037**
Student	0.243	0.229	0.014
Employment Status			
Employed	0.189	0.143	0.046***
Unemployed	0.219	0.190	0.029
Household Income			
Bottom Half	0.203	0.173	0.030*
Top Half	0.169	0.143	0.026
Day of Week			
Weekday	0.205	0.174	0.031*
Weekend	0.144	0.122	0.022

*** p < .05; ** p < .01; *** p < .001**

Table 8.3. The U-index and allocation of time across activities.

	U-index per activity		Percent of time	
Focal Activity	US	France	US	France
Walking	0.04	0.09	0.63%	1.69%
Making love	0.05	0.03	0.77%	0.98%
Exercise	0.06	0.03	0.88%	1.21%
Playing	0.07	0.02	1.47%	1.26%
Reading, non-work	0.09	0.07	2.97%	4.36%
Eating	0.10	0.09	5.22%	11.11%
Prayer	0.11	0.16	1.70%	0.25%
TV	0.12	0.14	7.07%	7.32%
Relaxing	0.13	0.13	2.88%	2.85%
Preparing food	0.14	0.13	2.92%	3.29%
Talking, non-work	0.14	0.12	9.35%	11.58%
Trooming	0.15	0.14	5.19%	4.76%
Other	0.16	0.13	8.54%	5.72%
Housework	0.18	0.23	5.91%	5.16%
Sleep	0.18	0.15	2.70%	2.32%
Other travel	0.20	0.20	3.23%	3.22%
Shop	0.22	0.20	4.86%	4.35%
Computer, non-work	0.23	0.22	2.52%	2.28%
Childcare	0.24	0.11	6.85%	4.50%
Commute	0.27	0.26	2.22%	1.68%
Work	0.29	0.26	22.10%	20.12%

Table 8.4. Synthetic U-index based on country's aggregate time allocation and country's U-index by activity. Calculations based on data in Table 3.

Country's	*Country's Time:*			
U-index:	U.S.	France	Difference	t-ratio
U.S.	0.189	0.177	0.012	1.02
French	0.169	0.159	0.010	0.90

Note: standard errors for t-ratios are derived from a bootstrap procedure that takes into account sampling variability in the U-index and in the time allocation.

Table 8.5. National Time-Use Data for U.S. and France and Synthetic U-Indexes

	Fraction of awake time spent in each activity					
	Work / Commute	Compulsory	Passive leisure	Active leisure	Eating	Other
US	24.6%	35.2%	24.8%	7.5%	6.6%	1.3%
France	21.8%	34.8%	18.1%	10.6%	14.3%	0.5%

	Average U-Index per activity					
US	0.29	0.19	0.15	0.10	0.10	0.15
France	0.26	0.17	0.14	0.09	0.09	0.13

Synthetic U-index based on country's aggregate time allocation from national time-use data and country's U-index by activity from DRM.

Country's U-index:	*Country's Time:* U.S.	France	Difference	t-ratio
U.S.	0.193	0.184	0.010	1.67
France	0.173	0.164	0.009	1.74

Notes: standard errors for t-ratios are derived from a bootstrap procedure that takes into account sampling variability in the U-index and in the time allocation. The work activity combines working and commuting, the compulsory activity combines shopping, housework, preparing food and grooming, passive leisure combines watching TV, non-work computer use, relaxing, and napping, activity leisure combines exercise, walking, making love, playing and talking.

Appendix Table
Linear Probability Multiple Regression Models for U-Index, Full Sample and By Sex

Explanatory Variable	Full Sample Coefficient	Full Sample t-ratio	Women Coefficient	Women t-ratio	Men Coefficient	Men t-ratio
Female	0.024	1.96	---		---	
Black	0.052	1.84	0.042	1.18	0.065	1.42
Hispanic	0.033	1.26	0.057	1.62	0.010	0.25
Log Income	-0.023	-2.31	-0.027	-2.15	-0.020	-1.21
< High School	-0.005	-0.21	0.006	0.17	-0.010	-0.27
Some College	-0.017	-0.96	0.006	0.27	-0.052	-1.98
College	-0.056	-3.27	-0.045	-1.99	-0.070	-2.76
College+	-0.045	-2.31	-0.020	-0.71	-0.082	-3.00
Age	0.003	1.67	0.009	3.11	0.000	-0.02
Age-squared	0.000	-1.95	0.000	-3.58	0.000	-0.07
Married	-0.017	-1.17	-0.051	-2.69	0.020	0.94
Tuesday	-0.012	-0.51	0.019	0.61	-0.043	-1.28
Wednesday	0.004	0.18	0.026	0.89	-0.022	-0.62
Thursday	0.005	0.22	0.035	1.13	-0.024	-0.68
Friday	-0.020	-0.86	0.000	0.00	-0.049	-1.42
Saturday	-0.009	-0.36	0.027	0.82	-0.055	-1.52
Sunday	-0.061	-2.62	-0.052	-1.79	-0.070	-1.87
June	-0.015	-0.92	-0.036	-1.66	0.010	0.41
July	-0.025	-1.67	-0.022	-1.01	-0.031	-1.50
August	0.046	2.32	0.030	1.16	0.065	2.14
No. of Episodes	9,989		6,136		3,853	

Notes: All regressions also control for 15 "who with" dummies, 5 dummies indicating the order in which affect questions were asked, and an intercept. Heteroskedasticity consistency standard errors that allow for within-person correlated errors were calculated. Data are from the PATS.

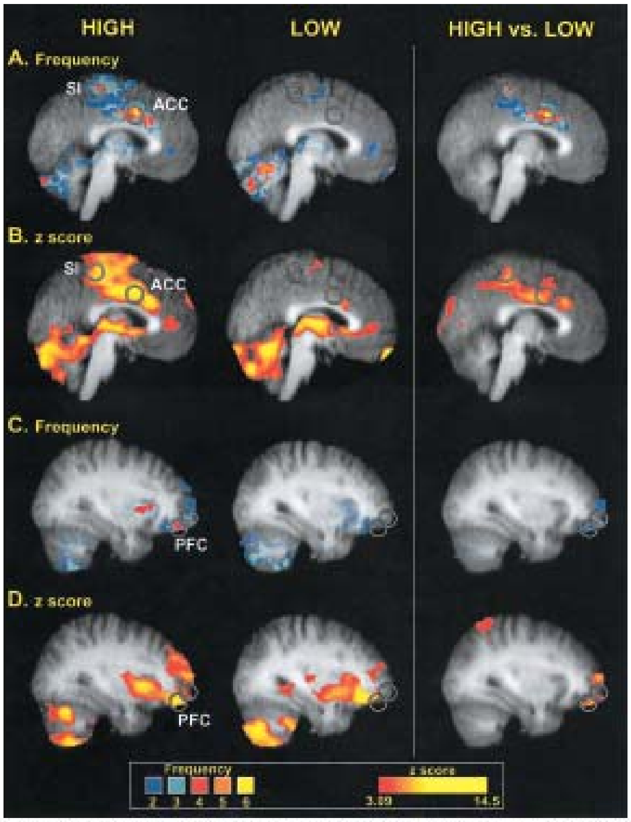
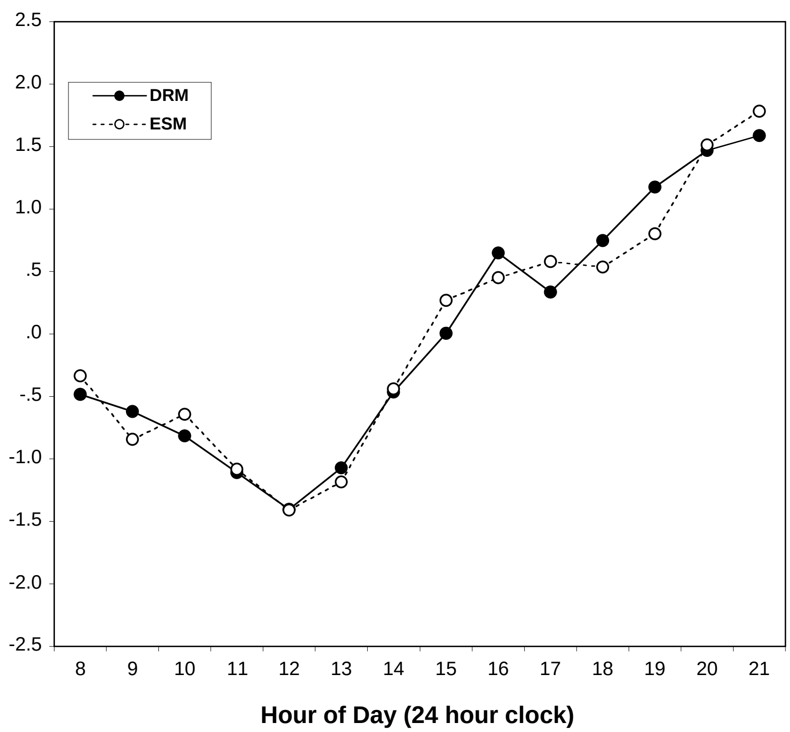


Figure 3.1. Brain regions displaying different frequencies of activation between high- and low-[pain rating] sensitivity subgroups. Circles are centered on regions where the peak differences between groups were located. Colors in *A* and *C* correspond to the number of individuals displaying statistically significant activation at a given voxel (frequency), whereas colors in *B* and *D* correspond to the *z*-score of the subgroup analysis. Slice locations in *A* and *B* are -2 mm from the midline, whereas slice locations in *B* and *C* are 32 mm from the midline (in standard stereotaxic space). Structural MRI data (gray) are averaged across all individuals involved in corresponding functional analysis. Reproduced from: Coghill, McHaffie, andYen (2003).

Figure 4.1. Comparison of pattern of tiredness over the day based on DRM and ESM samples. Points are standard scores computed across hourly averages within each sample.



Source: Kahneman, et al. 2004.

Figure 4.2. Comparison of pattern of tiredness over the day based on PATS, DRM and ESM samples. Points are standard scores computed across hourly averages within each sample.

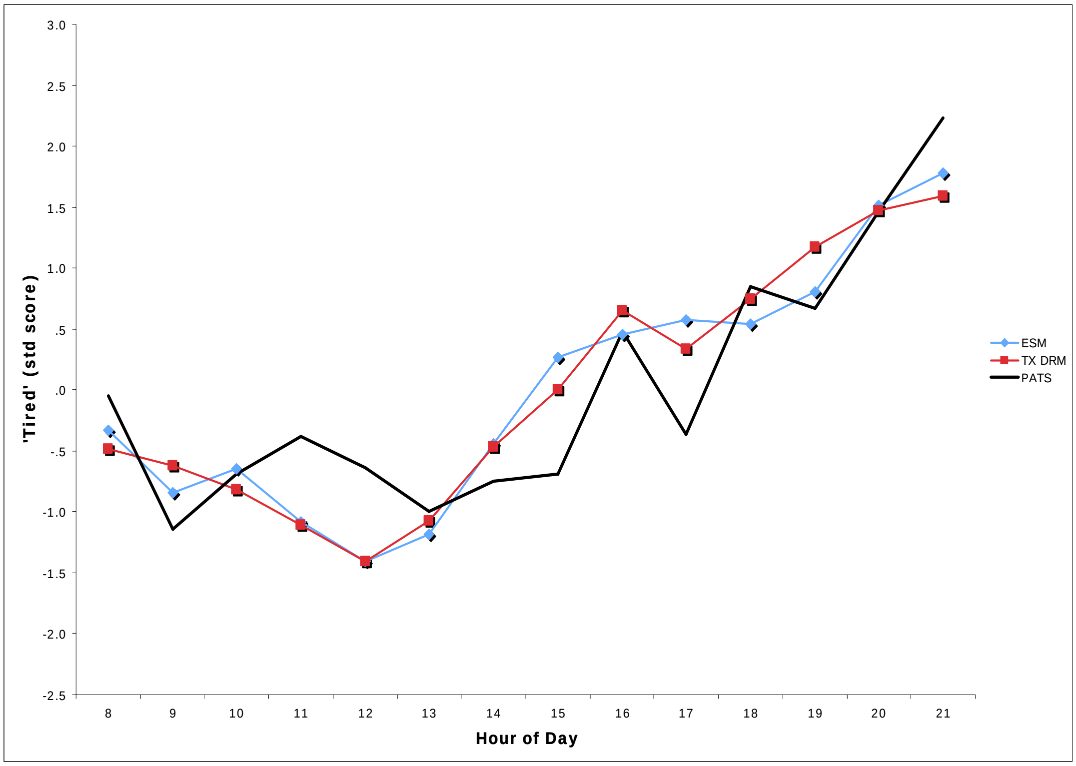
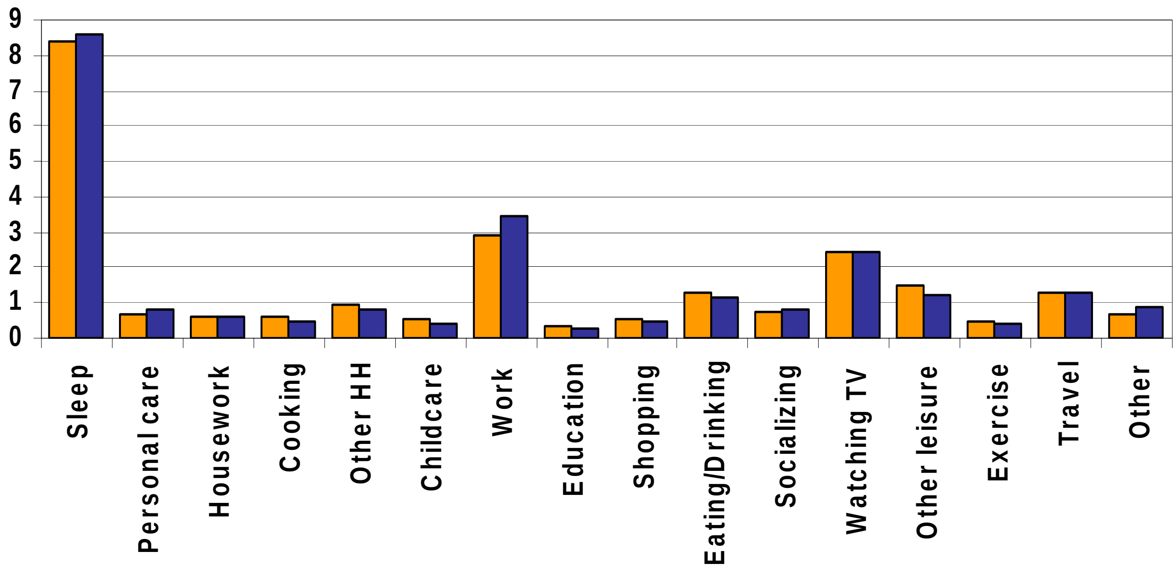
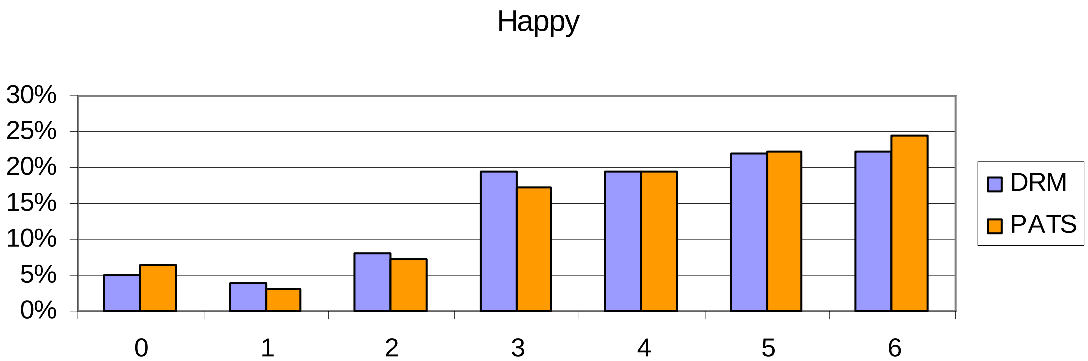


Figure 4.2. Average hours per major activity in PATS and ATUS



Notes: PATS shown in orange and ATUS in blue. PATS was conducted in May-August 2006 and ATUS is for May-August 2004-05.

Figure 4.3. Distribution of reported happiness and tired in PATS and DRM



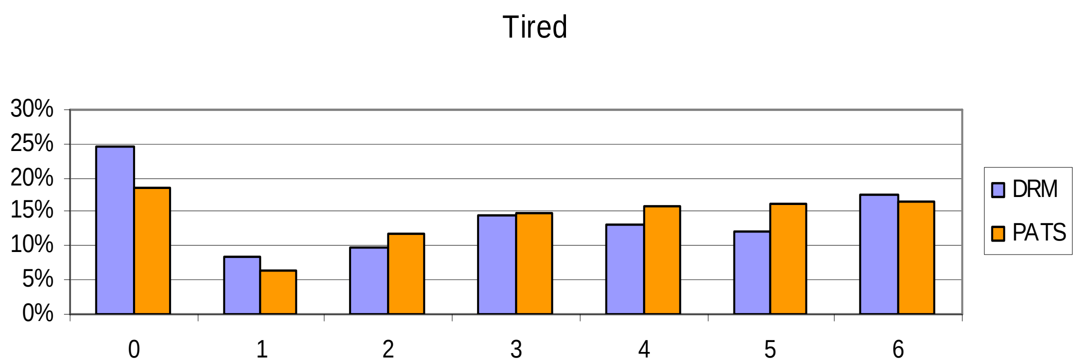


Figure 5.1. Happiness rating before and after exercise, results of a linear spline.

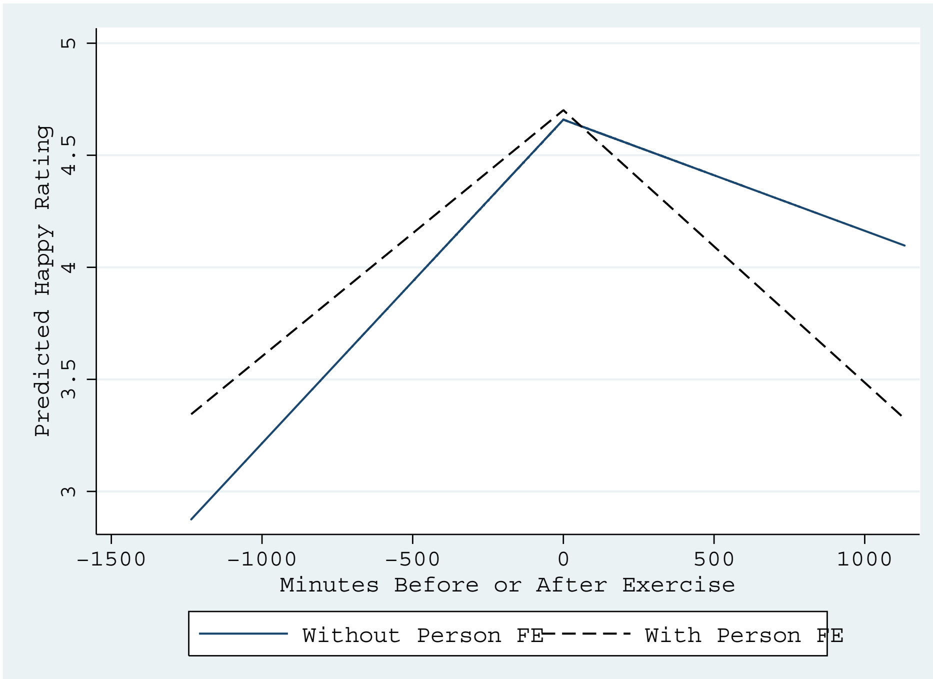


Figure 5.2. Time spent in various activities by age, 2005 ATUS

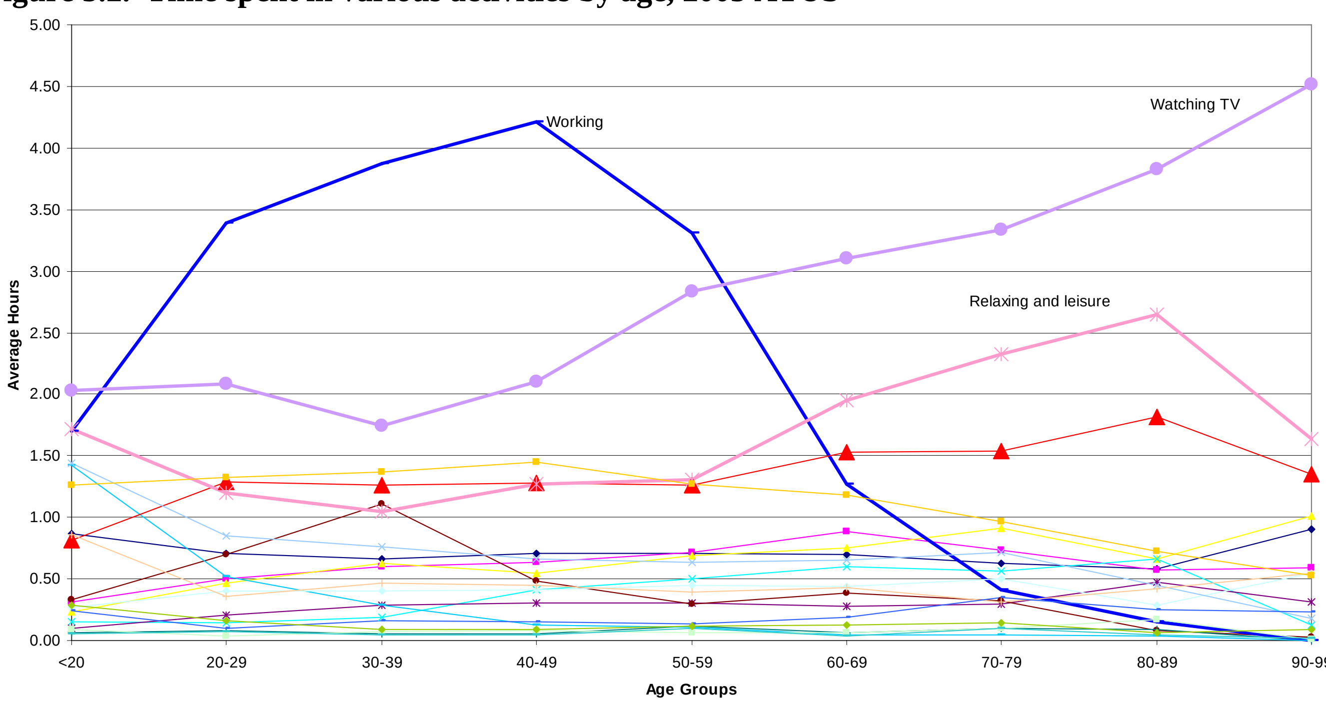


Figure 7.1. Activity-level U-index over time, using gender-specific U-indexes

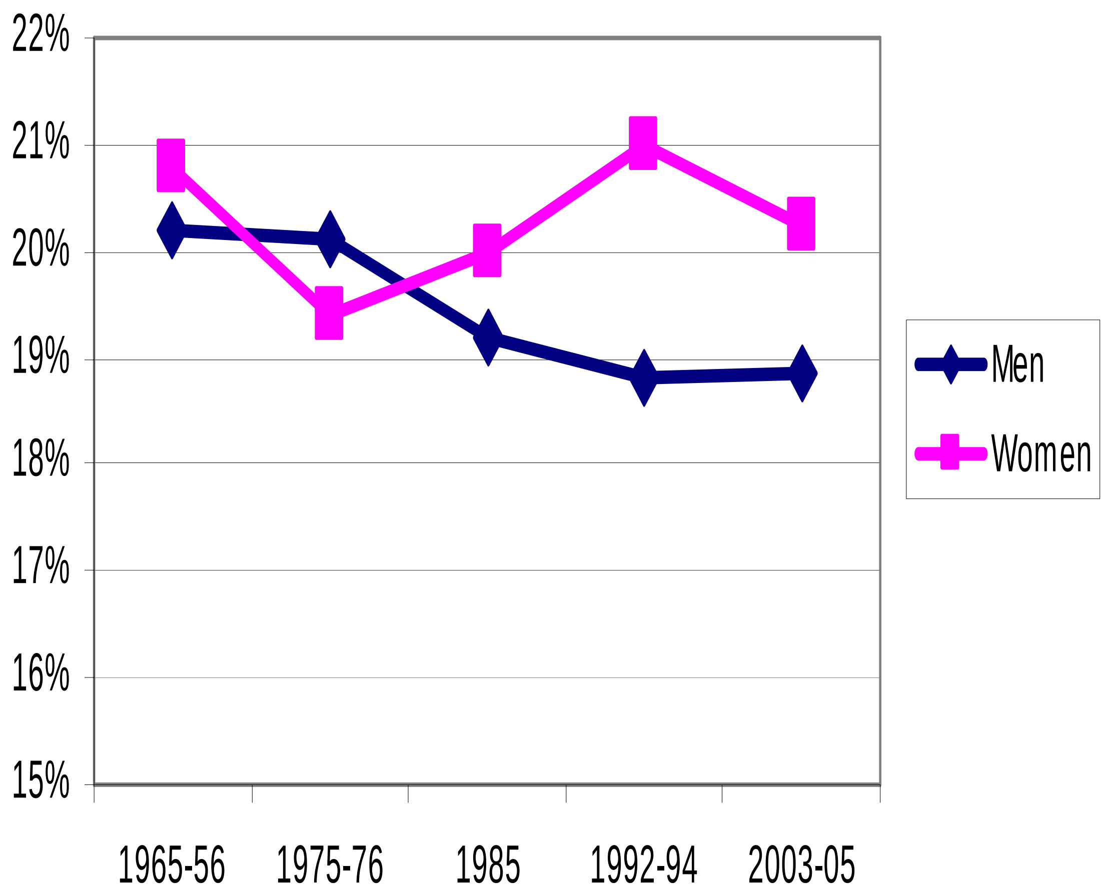


Figure 7.2. Dispersion of activity-level U-index across people, 1965-66 to 2003-05

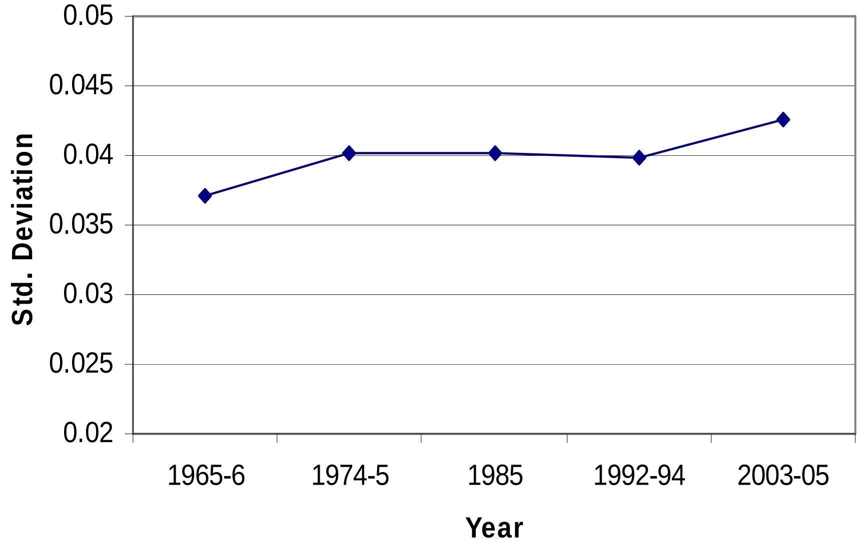


Figure 7.3. U-index for Men, by Education, 1965-66 to 2003-05

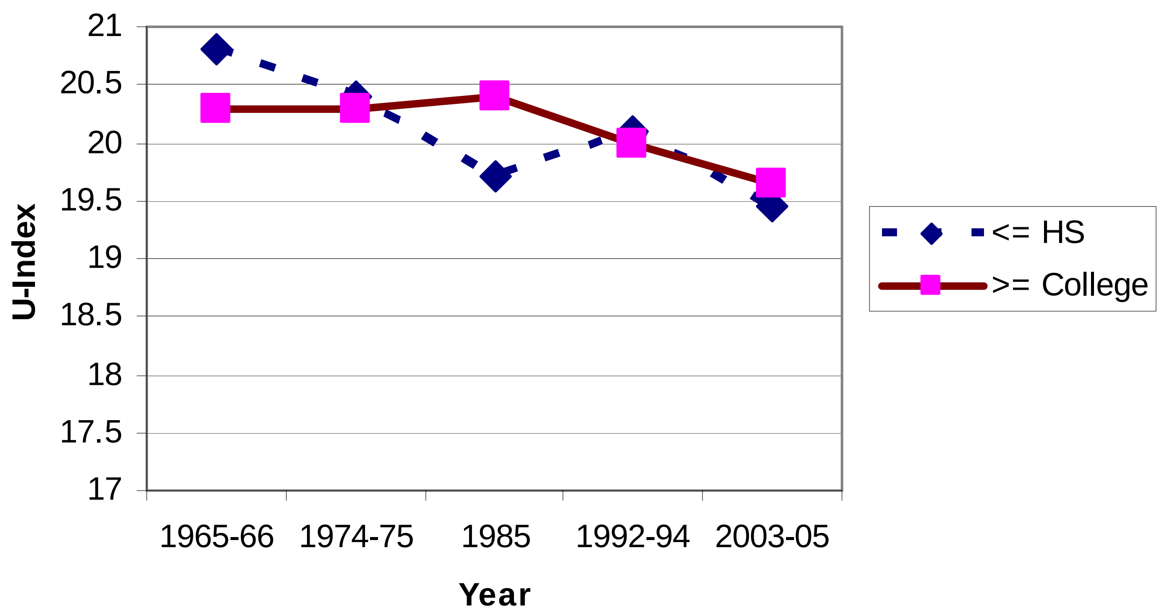


Figure 8.1. Average U-index by quintile of the U-index distribution in U.S. and France