Utility and Happiness

by Miles Kimball and Robert Willis[1]

University of Michigan

March 3, 2006

Abstract: Psychologists have developed effective survey methods of measuring how happy people feel at a given time. The relationship between how happy a person feels and utility is an unresolved question. Existing work in Economics either ignores happiness data or assumes that felt happiness is more or less the same thing as flow utility. The approach we propose in this paper steers a middle course between the two polar views that "happiness is irrelevant to Economics" and the view that "happiness is a sufficient statistic for utility."

We argue that felt happiness is not the same thing as flow utility, but that it does have a systematic relationship to utility. In particular, we propose that happiness is the sum of two components: (1) *elation*--or short-run happiness--which depends on recent news about lifetime utility and (2) *baseline mood*--or long-run happiness--which is a subutility function much like health, entertainment, or nutrition. In principle, all of the usual techniques of price theory apply to baseline mood, but the application of those techniques is complicated by the fact that many people may not know the true household production function for baseline mood.

If this theory is on target, there are two reasons data on felt happiness is important for Economics. First, short-run happiness in response to news can give important information about preferences. Second, long-run happiness is important for economic welfare in the same way as other higher-order goods such as health, entertainment, or nutrition.

[1] We would like to acknowledge first and foremost the substantial contributions of Norbert Schwarz to this paper. Our discussions with him from the very first beginnings of this paper clarified many things for us, particularly about the empirical evidence on happiness measures. However, we need to make clear that there are important aspects of our theoretical position he would not agree with. In addition to Norbert Schwarz, we would like to thank George Akerlof, Toni Antonucci, Robert Barsky, Kerwin Charles, Fred Conrad, Mick Couper, Michael Elsby, Gwenith Fisher, Bruno Frey, Christopher House, Michael Hurd, Helen Levy, Charles Manski, Randolph Nesse, Fumio Ohtake, Antonio Rangel, Luis Rayo, Matthew Shapiro, Daniel Silverman, Alois Stutzer, Yoshiro Tsutsui, Janet Yellen and participants in seminars at Osaka University, the Stanford Institute for Theoretical Economics, University of Michigan, Harvard, and Brown for helpful discussions and comments on early versions of this material. They must also be absolved from complicity in any errors we perpetrate, small or large. We are grateful for support from National Institute on Aging grant P01 AG026571-01.

1. Introduction

On first impression, *"utility"* and *"happiness"* seem to refer to the same concept. However, over the last century, economists and psychologists respectively have developed technical meanings for the words *"utility"* and *"happiness"* that refer to logically distinct concepts.

The success of the Ordinalist Revolution of Lionel Robbins (1932) and of John Hicks and R. G. D. Allen (1934)—codified as "Revealed Preference" by Paul Samuelson (1938, 1947)[2]—has fixed the meaning of "*utility*" for more than a half-century of economists as a representation of an individual's preferences over alternatives. The practice of Economics has made this concept of utility immensely valuable in thousands of applications.

In the aftermath of the Cognitive Revolution, the success of Hedonic Psychology—exemplified in the volume edited by Daniel Kahneman, Ed Diener and Norbert Schwarz (1999)—has fixed the scientific meaning of "*happiness*" within Psychology as the overall goodness or badness of an individual's felt experience at any point in time. To be more explicit, operationally, psychologists define current happiness as how people answer questions such as "*On a scale from one to seven, where one is extremely unhappy and seven is extremely happy, how do you feel right now?*" This concept of happiness has attracted increasing interest among economists in recent years.

Throughout this paper, we follow the convention that the technical meaning of *"utility"* is determined by the tradition in Economics, while the technical meaning of *"happiness"* is determined by the tradition in Hedonic Psychology. Thus, *utility* is a reflection of people's choices; *happiness* is a reflection of people's feelings. Once one recognizes these two concepts as distinct, *discovering the nature of the empirical relationship between utility and happiness* stands out in sharp relief as one of the central questions at the frontier between Economics and Psychology.

In the existing literature attempting to link utility and happiness, the dominant explicit or implicit hypothesis is that current felt happiness is equal to flow utility.[3] We argue that the hypothesis that felt happiness equals flow utility is empirically untenable. Instead, to oversimplify our discussion below, we argue, in effect, that a large component of happiness is much more like the recent *change* or *innovation* in lifetime utility than it is like flow utility.

Of course, even unchanging, predictable circumstances can have an effect on happiness, so it is important to allow for another, longer-lasting component of happiness. We argue that this long-run component of happiness is not always aligned with utility, since people often knowingly and without regret make decisions that sacrifice a pleasant mental state day after day for the sake of some other goal.

---

[2] For more of the history of these developments, see also George Stigler (1950).

[3] Kahneman (1999), Gruber and Mullainathan (2002), Frey and Stutzer (2004b), and Layard (2005) are some of the most explicit in equating happiness and flow utility.

Thus, in our view, happiness is the sum of a transitory response to good and bad news and a long-run response of mood to circumstances that is distinct from utility. To be specific, we propose that happiness is the sum of two components: (1) *elation*--or short-run happiness--which depends on recent news about lifetime utility and (2) *baseline mood*--or long-run happiness. Baseline mood is a subutility function—or output of a household production function—much like health, entertainment, or nutrition. In other words, long-run happiness is a "valuable commodity," that cannot be purchased directly, though inputs to it can be.

Such a theory of happiness not only makes sense of existing happiness data, but provides a road map for future research. According to this theory, data on felt happiness can make two contributions to Economics. First, short-run happiness in response to news can give important information about preferences. Second, long-run happiness is important for economic welfare in the same way as other higher-order goods such as health, entertainment, or nutrition. Policy issues surrounding long-run happiness arise because of the value of producing and disseminating knowledge about the household production function for happiness and from any externalities in the causes or effects of long-run happiness.

Desmond Morris, at the outset of his wonderful little book *The Nature of Happiness,* writes:

> "The true nature of happiness is frequently misunderstood. It is often confused with contentment, satisfaction or peace of mind. The best way to explain the difference is to describe contentment as the mood when life is good, while happiness is the sensation we experience when life suddenly gets better. At the very moment when something wonderful happens to us, there is a surge of emotion, a sensation of intense pleasure, an explosion of sheer delight—and this is the moment when we are truly happy. Sadly, it does not last very long. Intense happiness is a transient, fleeting sensation. We may continue to feel good for quite a while, but the joyful elation is quickly lost."

Morris's description of "happiness" emphasizes what we call *elation*—the word Morris also uses to describe this type of happiness. The "contentment" he refers to is close to our concept of *baseline mood*, which unlike Morris, we also consider a fully legitimate component of happiness, since both the contentment when life is good and the joy when life gets better are likely to affect measured subjective well-being.

A word is in order about the length of this paper. We have learned from experience in talking to colleagues and others that because of the widely varying preconceptions almost everyone has on the subject of happiness, the perspective we propose on happiness is easy to misunderstand. Therefore, we make an effort to lay out the issues very carefully. Moreover, as we discuss below, there is an existing consensus among most psychologists and economists who are involved in studying happiness with which we disagree. It is incumbent upon us to make clear exactly why we disagree with the existing consensus, which requires a reexamination of all of the key types of evidence that are used to back up that consensus.

The remainder of the paper can be divided into two halves. The first half, Sections 2-5, is conceptual. In it we make the case for utility and happiness as logically and empirically distinct concepts. The second half, Sections 6-10, is mathematical. In it we lay out a specific model of the relationship between utility and happiness, along with interpretations, extensions and applications.

2. Distinguishing Between Utility and Happiness

A. The Need to Establish Clear Terminology. One of the difficulties we face in explaining our viewpoint is that the tradition of equating "happiness" to flow utility runs deep in the history of economic thought. Indeed, Jeremy Bentham's (1781) first definition of 'utility' made the equation of utility and happiness explicit:

> "By the principle of utility is meant that principle which approves or disapproves of every action whatsoever according to the tendency it appears to have to augment or diminish the happiness of the party whose interest is in question …."

The "Revealed Preference" definition of utility—to which we resolutely adhere—is closer to Bentham's second, more inclusive, definition of utility, in the immediately following paragraph:

> "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered: if that party be the community in general, then the happiness of the community: if a particular individual, then the happiness of that individual."

Another difficulty we face in distinguishing utility and happiness is that, while "Revealed Preference" guides economic research, a more naïve Marginalism has remained very common in economic teaching. For example, "Principles of Economics" courses often teach about diminishing marginal utility by engaging students' intuitions about how happy they would feel in consuming different consumption bundles.

Let us state clearly that, throughout this paper, when we discuss utility, we do so from the perspective of Paretian Welfare Economics. Whether explicitly or implicitly, welfare questions motivate a large share of economic research; an orientation toward welfare questions is particularly important in informing our assessment of utility in cases where people are liable to mistakes. As for the focus on Pareto optimality, in our view, the use of happiness data is *not* a Philosopher's Stone that magically solves the difficulties in comparing utility interpersonally, but happiness data—used judiciously—can give useful information about individual preferences.[4]

Any adequate theory of utility and happiness must explain why the meanings of happiness and utility seem so similar. The right nuances for explaining the semantic relationship between "happiness" and "utility" can be found in the first two definitions for "happy" in the *American Heritage Dictionary* (1976, Houghton Mifflin):

---

[4] One can then make the leap from individual preferences to statements about social welfare on more or less the same terms as one could in the absence of happiness data. To the extent that happiness data give the illusion of providing a cardinal utility function, it is an illusion similar to that provided by expected utility theory—where one may sometimes need to be reminded that a monotonic transformation $f(E(U))$ of the *overall* objective function $E(U)$ leaves preferences unaltered. Just as there is no *necessary* reason why the curvature of U in the expected utility representation $E(U)$ tells us how to aggregate preferences interpersonally, there is no *necessary* reason why whatever structure is revealed in preferences *as they relate to happiness data* tells how preferences must be aggregated. At a minimum, any debate about what happiness says about social welfare must take into account the existing literature on social welfare and social choice theory.

happy … 1. Characterized by luck or good fortune; prosperous. 2. Having or demonstrating pleasure or satisfaction; gratified.''

The second definition is the meaning of "happy" in Psychology. The first definition talks about prosperity, which seems closely linked to utility, but there is a hint of a stochastic element in the nature of happiness: "luck or good fortune." Our view of happiness emphasizes *recent* good luck by positing that an important component of happiness has to do with an individual's reaction to recent news about lifetime utility. Although the differences are important, news about lifetime utility and lifetime utility itself are linked tightly enough that it is not surprising to find a certain confusion between the two in the structure of the lay lexicon. In other words, if people feel happy whenever they receive good news about lifetime utility, it is not hard to see why they would sometimes use the word "happiness" to describe lifetime utility itself. Yet scientifically, we consider it crucial to have two distinct, clearly delineated concepts for revealed preference utility and happiness in the psychological sense of current feelings. Maintaining two distinct concepts—on an equal footing—in a situation where each has a certain tendency to subordinate or engulf the other, is one of the main contributions of this paper.

One way to think about the distinction between utility and happiness is that one's commitment to an Ordinalist, "revealed preference" definition of utility is confronted with an acid test when confronted with happiness data. There is a sense in which the most radical implications of the Ordinalist Revolution are apparent only in the light of data on experienced happiness.

Both felt happiness and choice-based utility are well-defined, observable concepts. Our aim is to determine the dynamic relationship between the standard psychological concept of current affect—felt happiness—and the standard economic concept of lifetime utility. Establishing *any systematic relationship* between happiness and utility would provide an important bridge between Psychology and Economics, allow psychological data and theory to be used in Economics in a way that is complementary to standard economic data and theory, and enable economists to bring to bear all the tools of economic theory toward understanding happiness.

B. Distinguishing Between Utility and Happiness as a Matter of Logic. In Psychology, the term "*subjective well-being*" refers to a multidimensional concept that includes evaluations of one's life-as-a-whole and of specific life-domains as well as the pleasantness of one's average experienced affect. Though the terminology has not been entirely standardized in the literature, *affect* is a useful term to refer to how happy a respondent currently feels, as opposed to judgments about his or her whole life. An attractive feature of affect measures is that the cognitive burden they place on respondents is modest in contrast to the extremely difficult cognitive task of forming a judgment about the quality of one's entire life. Throughout this paper, we use "current affect" and "happiness" interchangeably.

Economists have been slower than psychologists to focus on subjective well-being data. But a growing economic literature has made use of subjective well-being data. Richard Layard's (2005) book gives a good introduction to this literature and Bruno Frey and Alois Stutzer (2002)

give a partial survey.[5] This literature lays out many provocative findings, but with a few exceptions, the focus of this literature has been on the cross-sectional and trend properties of subjective well-being rather than on its detailed dynamic properties. Two key motivations for the use of subjective well-being data in Economics (shared in large measure by Hedonic Psychology itself) have been (i) the desire to study the welfare implications of non-traded goods[6] (something that is especially important for older people for whom market work is a less dominant part of their lives) and (ii) the desire to study welfare implications in contexts where preferences are potentially inconsistent and to diagnose optimization mistakes.[7]

Despite this growing literature, many economists are still very skeptical of the use of subjective well-being data,[8] in large part because the theoretical status of affect--"happiness"—within economic theory is unclear. A simple multiple-choice question illustrates this lack of clarity:

What is Happiness?

a. Flow utility?
b. The individual's overall objective function?
c. The part of the individual's objective function that abstracts from the desire to do one's duty?
d. The individual's objective function plus pleasure from memory?
e. None of the above?

To begin to answer this kind of question, it is important first to distinguish utility and happiness as a matter of logic. Then the relationship between utility and happiness will ultimately be an empirical matter. Using the shorthand "lifetime utility" to refer to an individual's overall objective function—*including things the individual cares about that occur after his or her death*—we can distinguish lifetime utility and current affect ("happiness") as follows:

- **Lifetime Utility** = The extent to which people get what they want, where what they want is indicated by their *choices*.
- **Current Affect** = How positive people's *feelings* are at a given time.

In thinking about lifetime utility, it is important to remember that people's choices clearly show that they value a wide range of goods that are not traded in markets or only partially traded in markets. Thus, the economic concept of lifetime utility is not limited to what are sometimes called "economic goods" but includes the value an individual places on *non-traded goods* such as respect, freedom, clean air, a vibrant community, being married to a particular person, and such *partially-traded goods* as time allocations--which are partially traded because people are paid for work time---and health and longevity--which are partially traded because people pay for health care.

C. Utility and Happiness as Empirically Distinct Candidates for a Welfare Measure.
Lifetime utility is the standard welfare measure in economics at the individual level. It is often

[5] Some of the recent empirical papers in economics using happiness data are John Helliwell (2002), David Blanchflower and Andrew Oswald (2004), Clark (1999), Rafael Di Tella, Alberto Alesino and Robert MacCulloch (2004), Di Tella and MacCulloch (1999), Di Tella, MacCulloch and Oswald (2001, 2003), and Wolfers (2003).
[6] See for example Frey, Simon Luechinger and Stutzer (2004) and Frey and Stutzer (2000, 2004a).
[7] See for example, Jonathan Gruber and Sendhil Mullainathan (2002) and Frey and Stutzer (2004b).
[8] See, for example, Daniel Hamermesh (forthcoming).

thought of as a discounted sum over time of "flow utility." As a counterpoint to this, Kahneman (1999), in a chapter that has been influential among psychologists who study well-being, has urged a discounted sum over time of affect (momentary experienced happiness) as the appropriate measure of overall individual welfare.[9] A *prima facie* case can be made for each of these views. Both subjective well-being and utility are based on trusting an individual's own judgment, but different judgments are trusted in each case: as a welfare measure, lifetime utility puts trust in an individual's (conscious and subconscious) judgments as reflected in choices, while the discounted sum of affect puts trust in an individual's (largely subconscious) judgments as expressed in feelings.

It would be very convenient if flow utility and affect were essentially equivalent; in that case the standard economic measure of individual welfare would match Kahneman's (1999) proposed measure of individual welfare. Unfortunately, things are not so easy. In brief, as we discuss below, affect seems to behave very differently from at least our traditional notions of the behavior of flow utility:

- The Easterlin Paradox: Flow utility trends upward while affect has no strong trend.
- Hedonic Adaptation: Flow utility is usually thought to respond permanently to permanent shocks, while affect seems to be very strongly mean-reverting.

Clearly, one could attempt to modify either one's ideas about utility or the mode of measuring happiness to try to bring flow utility and affect closer together. We argue that it is better to accept *utility* as determined by standard, best-practice economic methods of measurement, and *affect* as determined by standard, best-practice psychological methods of measurement, then pose as an *open-ended question* the nature of the *relationship* between these two concepts. We make a specific proposal for what this relationship might be, but we consider the question—thus posed— more important than our attempt at an answer.

3. Measuring Happiness

The logical distinction between happiness and utility becomes clearer the more closely one pays attention to the way each concept is measured. In this section we argue that psychologists can reliably measure happiness, in the carefully defined sense of how people feel at a given time. Of course, that leaves the question of what happiness *is*.

To say the same thing in a different way, some economists think happiness can't be measured well. *This is just not true.* Happiness (current affect) is one of the easiest of all subjective concepts to measure. What *is* true (that these economists are intuiting) is that once happiness is measured, we don't know what it means in terms of economic theory.

Psychologists have taken measurement issues in assessing emotions in general, and happiness in particular, very seriously. Randy Larsen and Barbara Fredrickson (1999) give a survey of research touching on this issue. Self-report measures of happiness and sadness (the most common type of measure) have been related to impressionistic observer ratings of happiness,

[9] Kahneman calls momentary affect "instant utility," but here, to avoid confusion, it is best to reserve the term "utility" for the concept of overall individual welfare in Economics.

highly-structured coding of facial expressions by trained observers, assessment of voice tone, electromyographic measurement of face muscle activation, measurement of skin conductance, heart-rate, blood pressure and respiration, electro-enchephalograms, positron emission tomography and functional magnetic resonance imaging of brain activity (where "…approach related positive emotions are associated with left anterior activation whereas withdrawal related negative emotions are associated with right anterior activation" Larsen and Fredrickson (1999), p. 53.) Self-report measures of happiness and sadness have also been shown to predict many types of cognition and behavior in the laboratory, including writing speed and performance speed on other tasks, judgments of probabilities, the output of free word association and word completion under time pressure, speed of judging positive and negative words versus nonwords, and the speed of the startle reflex after a loud sound. All of these experiments add up to consistent evidence that happiness is a measurable psychological state.[10]

Among self-report measures of happiness, the gold standard is experience sampling, in which people are signaled at random intervals to report their current happiness. Kahneman, Alan Krueger, David Schkade, Schwarz, and Arthur Stone (2004) argue that the day reconstruction method is a close second. Measuring happiness as part of a large-scale survey presents an extra issue in that the survey itself may represent a significant slice of a day. To avoid too much emphasis on the feeling states engendered by the interview process itself one can ask about happiness over a longer, but still relatively short span of time.[11] The Health and Retirement Study measures affect by the following series of questions:

> *"Now think about the past week and the feelings you have experienced. Please tell me if each of the following was true for you much of the time this past week:* [12]
>
> *a. Much of the time during the past week, you felt you were happy. (Would you say yes or no?)*
> *b. (Much of the time during the past week,) you felt sad. (Would you say yes or no?)*
> *c. (Much of the time during the past week,) you enjoyed life. (Would you say yes or no?)*
> *d. (Much of the time during the past week,) you felt depressed. (Would you say yes or no?)"*

---

[10] In general, self-report measures of emotions can be affected by social desirability and the semantic framing effects that arise cross-culturally, and lack of conscious awareness of emotions. For the most part, social desirability and semantic framing effects should be fairly constant over time within a given culture and can be dealt with empirically using fixed effects. The likelihood that people might lack conscious awareness of emotions is a subject of debate within Psychology. Some psychologists insist on conscious awareness as part of the definition of an emotion. (Larsen and Fredrickson, 1999 reports that "some would question whether an unperceived emotion is an emotion at all.") But even Tim Wilson (2002), in a book-length argument for the possibility of unconscious feelings, points out that "feelings differ from the rest of the adaptive unconscious in their potential to reach awareness" and allows that "It might even be the case that the default is for feelings to emerge into awareness, and that it takes special circumstances to prevent them from doing so." (See Wilson, 2002, p. 134.) It seems likely that the overall positive or negative aspect of feelings that we are focusing on under the label of "happiness" makes it into consciousness more reliably than the detailed reasons behind feelings or finer categorizations of emotions. Wilson (2002) goes on to discuss repression, inattention and "the obscuring of feelings by the smoke screen of people's conscious theories and confabulations." Repression and inattention seem unlikely to cause serious problems for the survey measurement of happiness. However, "the smoke screen of people's conscious theories" about happiness is a serious issue, which we address below.

[11] Michael Robinson and Gerald Clore (2002, p. 950) looked at evidence on happiness reports with different time frames. Their evidence led them to conclude that a few week's time is the longest interval for which one can get happiness reports that are not contaminated in an important way by people's theories of how they "should" feel.

[12] In the first wave respondents were instead asked "*Please tell me how often you have experienced the following feelings during the past week: all or almost all of the time, most of the time, some of the time or none or almost none of the time."*

Operationally, one can treat happiness as the latent variable behind these four yes/no questions. This series of questions on the Health and Retirement Study is a subset of the Center for Epidemiologic Studies Depression (CES-D) measure of depressive symptoms.[13] These questions illustrate what we mean when we say that the concept of happiness we are referring to is about current feelings. These questions ask about easily accessible feelings and memories of feelings in the past week. One indication of how readily respondents answer these questions is that the average amount of survey time required for all four questions put together is less than 35 seconds.

It is important to contrast current affect measures like those on the HRS with life satisfaction measures, such as those on the German Socioeconomic Panel—*"On a scale from 1 to 10, how satisfied are you with your life?"*—and "global happiness" questions, such as those on the World Values Survey:

"Taking all things together, would you say you are
1. *Very happy*
2. *Quite happy*
3. *Not very happy*
4. *Not at all happy*
9. *Don't Know [DO NOT READ OUT]"*

An extensive body of psychological research explores the cognitive processes underlying global judgments of happiness and life-satisfaction (for a review and process model see Schwarz and Strack, 1999). It converges on the following conclusions[15]:

1. Reported life-satisfaction does not reflect stable inner states of respondents. Instead, these judgments are formed on the spot and depend on which aspects of life happen to come to mind at the time of judgment, which gives rise to pronounced context effects. For example, when students are asked to report their overall life-satisfaction and their dating frequency, both correlate r = .1 when the life-satisfaction question is answered first, but r = .7 when the dating frequency question precedes the life-satisfaction question, thus bringing the domain of dating to mind (Strack, Martin, and Schwarz, 1988).
2. The use of comparison standards is similarly context dependent. People can evaluate their current lives relative to their expectations, their past situation, the situation of others, and so on, resulting in profoundly different judgments. For example, the mere presence of a handicapped other in the room is sufficient to increase global life-satisfaction (Strack, Schwarz, Chassein, Kern, and Wagner, 1990) and one's current life looks good or bad depending on which aspect of one's past was brought to mind (Strack, Schwarz, and Gschneidinger, 1985).
3. People can simplify the complex task of evaluating their life-as-a-whole by drawing on their current feelings as an indicator of their overall well-being. For example, survey respondents report higher life-satisfaction when called on sunny rather than rainy days—unless a preceding question about the weather

[13] See Steffick (2000) for a detailed description and assessment of the CES-D questions in the HRS. Besides omitting the other less relevant questions, we have reversed the order of the first two questions even after those omissions in order to give the version of the question that we would recommend for use on other surveys that do not have a more extensive CES-D battery of questions.

[15] We are particularly grateful to Norbert Schwarz for this summary of the psychological research on different subjective well-being measures.

makes them aware that their current mood may not provide diagnostic information about the overall conditions of their lives (Schwarz and Clore, 1983).[16]

This context-dependence of evaluative measures of well-being attenuates any meaningful relationship with objective circumstances of life and motivates approaches to the measurement of well-being based on people's momentary affective experience.

In comparison to global evaluations of one's life-as-a-whole, assessments of current affect pose more reasonable cognitive demands. As noted in point 3 above, experimental evidence suggests that survey responses to questions about overall life satisfaction or about global happiness with life rely heavily on the readily accessible internal information a respondent has about current affect (Schwarz and Clore, 1983).[17] Thus, how a respondent feels *right now* has a strong effect on answers to overall life-satisfaction and global happiness questions, whether we like it or not. We maintain that it is clearer to focus on current happiness directly, so that we know what we are getting, in a transparent way. Finally, to the extent that respondents are not using current affect as a shortcut to make an overall evaluation of life satisfaction or global happiness, there is a serious danger that they will report how happy or satisfied they think they *should* feel about their lives according to whatever folk theories they have about happiness and satisfaction.[18]

4. Measuring Utility

For economists, a discussion of measuring utility is only a reminder. Utility is defined by revealed preference—the information gleaned from the choices people make. Some of the accumulated wisdom from economic research is encoded in standard functional forms that are repeatedly applied and tested. The techniques of revealed preference can be applied to tradeoffs over seemingly incommensurable values, and apply even to situations involving choices over time.

For non-economists, one can say that the concept of utility relies on an individual's judgment of his or her priorities, as reflected especially in his or her actual choices when faced with a tradeoff, or, at a minimum, his or her choices in a hypothetical situation. Higher utility out of any two choices is defined by what the individual chooses (or would choose) when presented with those two choices. Thus, utility is a measure of the extent to which people get what they want, and differences in utility are predictors of behavior. This allows a deep connection between positive (descriptive) and normative (prescriptive) aspects of utility theory in Economics.

A. The Upward Trend in Utility. In view of the Easterlin (1974, 1995, 2003) Paradox of secularly nontrending happiness, an important application of the principle of revealed preference

[16] The relationship of such context-dependence to decision-making is an important research question. For example, Hirshleifer and Shumway (2003) indicates that sunny days have a detectable effect on stock-market trading.
[17] See also Schwarz (1996, 1999) and Schwarz and Bohner (2001).
[18] For example, consider the fact reported by Lucas, Clark, Georgellis and Diener (2004), that life satisfaction is permanently dragged down by an episode of unemployment. Even if the affective sting of past unemployment has long since faded away, asking for an overall evaluation of life satisfaction invites the respondent to evaluate the past as well as the present. It is not surprising that a past episode of unemployment permanently affects one's assessment of one's autobiography.

is to a hypothetical choice between the comprehensive consumption bundle (including all externalities, public goods and time use patterns) now and the comprehensive consumption bundle fifty years ago in the U.S. Real per capita income has grown dramatically over that period of time, which means that the total set of marketed consumption bundles that people can choose from has expanded. Higher real per capita income allows people more choices, out of which they typically elect to spend in ways they could not previously afford--rejecting the available option of continuing to spend in the same way they did at the lower income level. Average work hours have trended slightly downward. Moreover, as Easterbrook (2003) points out, a large set of goods not traded or only partially traded in the market have either stayed about the same or improved over the last fifty years. Among partially-traded goods, medical care and longevity have been improving dramatically, while household conveniences have reduced the time necessary for housework and increased the time available for genuine leisure even for many who do spend longer hours in market work. Equality between the sexes and races, while far from complete, is much better than two generations ago; the number of democratic nations is on the rise; and even the War on Terror, which at worst could involve the nuclear destruction of a large portion of Manhattan, is an improvement over the Cold War, which at worst could have destroyed human life from the face of the earth. Finally, many of the non-traded goods that worsened for a while after 1955, such as rates of crime, teenage pregnancy and drug abuse, have turned the corner and begun trending in a favorable direction for the last two decades. In short, although many problems remain, and are the focus of nightly news reports, we argue that it would be a bad deal to trade the problems we face today for the problems of yesteryear, implying that utility is higher than fifty years ago.

Of course there are those who look back at the past with nostalgia. In part, the increasing individual freedom that comes with higher per capita income may have some undesirable side effects such as a diminished sense of community, of the sort Robert Putnam (2000) describes in his book *Bowling Alone: The Collapse and Revival of American Community*. (Trends in the divorce rate and other aspects of family structure can be seen as part of the same phenomenon.) Similarly, the rise in per capita income may have increased the availability of illegal drugs and access to a wide variety of delicious foods and drinks that create important intra-psychic conflicts. Still, how many would really want to go back to the way it used to be if they saw clearly the way it *really* used to be? It is easy to forget the legitimate and the irrational fears engendered by the Cold War, the toll of racial and other injustices on those mistreated, the enforced conformity that went along with the greater sense of community in the past, and how effective the long-available drug of alcohol is at messing up the lives of those who are prone to intra-psychic conflict. And it is easy to take for granted boons such as word processing, the ability to watch any of a huge range of movies at home, the existence of J. K. Rowling's *Harry Potter* series, and inexpensive access by means of free internet access at the public library to a huge range of fascinating scientific findings that were not known fifty years ago, let alone available at the click of a mouse.

Besides ordinary selective memory that often leads us to forget former difficulties once they have been surmounted, there is another kind of bias that helps to fuel such nostalgia: despite recent trends in historiography toward telling the stories of those at the bottom of the social

ladder[19] as well as the stories of common men and women, our image of the past is still often dominated by the biographies of those near the top of the social ladder who were much better off than the average person in their time. Even when we assess the past by thinking of the experience of our own grandparents, they are far from a random sample of people in their time. The ancestors of a randomly chosen individual in the present are likely to be people who were more successful than average in number of descendants--and likely to be above average in the degree of success in their life experience more generally.[20] The experience of those who died young in the past or who never found mates is not remembered as well as the experience of those who did. The travails of those in the present who die young or who never find mates is more apparent.[21]

One reason it is important to hold imagined social rank constant when assessing the past versus the present is that most people care a lot about social rank.[22] From the standpoint of revealed preference, it is not difficult to observe people making choices that sacrifice other valuable things in order to attain higher social rank. As a consequence, it is not incredible that someone might choose to be a king or queen in a bygone era rather than a middle-class person today, even if the real value of the market consumption bundle of the middle-class person today is worth much more. Individuals' positional concerns are not irrational. In many domains of life, relative standing is more crucial to obtaining desired outcomes than absolute standing (for discussions see Frank, 1985, 1999; Hirsch, 1976; Sen, 1983). Social rank yields hard-to-measure but real benefits in terms of respect and favorable treatment by others. Because, in practice, social rank is so highly correlated with income, at least in the United States, secular comparisons are useful for distinguishing concern with income from concern with social rank.[23]

The point of this extended discussion is to argue that average lifetime utility for people of a given age is higher than it was fifty years ago, even after accounting for a wide range of tradeoffs going far beyond those that are in obvious monetary terms. Of course, the choice between the comprehensive consumption bundle of fifty years ago and the comprehensive consumption bundle now is a hypothetical choice. But every year millions of people make a choice that is similar in important respects by migrating from a poor home country to the U.S. and other rich countries. Many leave behind tight-knit communities in which they have high local social status

[19] Our choice of the "social ladder" metaphor is influenced by the form of a question on the HRS leave-behind survey asking people to mark their perceived social rank on a printed ladder, pioneered by Michael Marmot's Whitehall II studies of British civil servants. (See Marmot, 2004.)

[20] In the future, this bias could go the other way, since the ratio of family size for high-status parents to family size for low-status parents seems to be falling over time.

[21] Note that, in comparing the past to the present, it is important to abstract from people's preference for the familiar and status-quo bias more generally, which would have worked in favor of the actual experience in the past as much as in the present. A good way to abstract from the attraction of familiar idiosyncratic details of one's life is to imagine a choice between being (a) thrown into the life of a randomly chosen individual in the present and (b) being thrown into the life of a randomly chosen individual in the past. Unfortunately, there is no such helpful device to help in abstracting from the familiarity of one's entire era.

[22] For some direct evidence on the strength of preferences over social rank based on hypothetical choices, see Solnick and Hemenway (1998). For a discussion of social comparison by psychologists, see Sulls and Wills (1991).

[23] There are some instructive instances of social rank diverging from income rank even in the present-day U.S. Clergy and teachers (including professors) often have considerably higher social rank than their income rank. This relatively high social rank is important in making many people willing to sacrifice a significant amount of income to go into these fields.

for a foreign land where they will be at the bottom of the social status ladder and where they cannot even speak the language. Clearly, in making the enormous effort of migrating, with all the psychic costs of being uprooted from one's familiar cultural surroundings, they are choosing something that they value highly—the modern consumption bundle in the U.S. and other rich countries.[24]

B. Mistaken Choices. The greatest difficulties in measuring utility arise when people make mistaken choices or have inconsistent preferences.

Garden-variety mistakes based on a lack of knowledge of objective facts are the easiest to deal with. Consider the case of someone who chooses a particular car, thinking that it will get good gas mileage, but then discovers that it gets bad gas mileage. Once she learns this, the purchaser regrets the earlier decision and wishes that she had chosen a different car. In this case, in judging utility one needs to either use the choice the purchaser *would have made* if fully informed, *or* the choice the purchaser did make between ideas of cars with assumed characteristics.

Sometimes a mistake arises not from lack of basic facts, but from a failure of computation. For example, one of the authors has only very recently begun to adjust his book-buying habits for the high shadow-cost of available book-reading time that is generated by the large number of books already in his personal library, by thinking "What are the chances that on any future date this book will win out over all of the competition?" This is the kind of reasoning that does affect one's choices when the calculation has finally been made. Taking Paretian Welfare Economics as our touchstone, we consider it relatively uncontroversial to suggest that utility (strictly speaking, preferences) be measured according to what people would choose when not only well-informed in terms of raw information, but also when they are aware of relevant calculations and lines of reasoning.[25]

A third type of mistake is making mistakes about what one's subjective experience will be after a given choice. There is nothing disruptive of standard economic theory about the existence of experience goods, such as a new flavor of ice cream, for which preferences are known only after trying some of the good. Marketing strategies by firms selling experience goods vary from actively providing a free taste to forcing people to buy a substantial package based on guesswork. When free samples are not provided, it is easy to make mistakes due to not fully knowing one's own tastes, even if the physical properties of the product itself are well-known.

For an expensive durable good, an optimal decision of whether to purchase the good should involve considering the time-path of one's subjective experiences with the good. It is not hard to imagine someone changing her mind about buying an expensive car upon being shown evidence that after a year people report roughly the same experience when driving an expensive car as

[24] Note that, at $9600, the per capita GDP of Mexico—an important source of migration to the U.S.—is not far below the U.S. real per capita GDP in 1955.

[25] There is a practical problem of distinguishing between the force of a calculation or line of reasoning itself and the desire to agree with the person urging that line of reasoning. In principle there are ways to deal with that problem. For example, in presenting a hypothetical choice, it is important to even-handedly present correct calculations and lines of reasoning that favor both the pro and con side of a decision. Also, to minimize social pressure, it may be possible to present calculations and lines of reasoning by a prepared text or an interactive computer setup. Making sure the agent is able to make the decision with as much anonymity as possible may also be helpful.

when driving a much cheaper car—say because driving is one of those activities that becomes reasonably automatic and so fades into the background of awareness—pushed out by thoughts of where one is going to and where one has just been. Indeed, people might not even need formal evidence for things like this; their own past experience of being less excited by a new good after the first few months could inform their later decisions. However, a considerable body of evidence reviewed in George Loewenstein and David Schkade (1999) indicates that people make serious mistakes in predicting future affect. Though it is larger in size and scope, one can view a mistake in predicting future affect as akin to a mistake in predicting whether one will like a particular type of ice cream. It is not clear whether people are making the right decisions or not until they are well-informed about the modification of the time-path of affect that will actually result from a purchase.

Some psychologists have gone further, to maintain that the fact that reported happiness with a new car is often high in the first few months after purchase and much lower thereafter—in a way that people are bad at predicting—*necessarily* means that someone has made a mistake in purchasing it. To our way of thinking, this is going too far. The key question is whether a correct knowledge of the modification of the time-path of affect that will actually be induced by a good would make a material difference to a decision. The issue becomes clearest if we consider the choice of a purchaser who is fully conversant with the Hedonic Psychology literature and carefully observant of the pattern of his or her own affective reactions. As long as the purchaser is aware of and thoughtfully considers the fact that happiness with a new durable is likely to fade after a time (and absent the kind of inconsistency discussed below), it seems appropriate to defer to that well-informed, thoughtful decision in judging utility, regardless of the time-path of subjective experience with the new car. Indeed, there is every reason to think that people care about many attributes of a car other than its price and the subjective experience they will have with it—such as its ability to get reliably to work and back. Even the set of all indirect effects of a purchase on affect (for example, including the reduced likelihood of sorrow from being scolded for getting to work late) should not necessarily be dominant in the decision of whether to make a purchase. The concept of utility (or equally in this context, preferences) involves deferring to each individual's own view of how much to factor in the modification of the time-path of affect that would result from a purchase when making the decision of whether to buy or not to buy. Since the consequences for affect are only one aspect of a good, it would not necessarily be irrational to give those consequences only a small weight even after understanding them fully.

A fourth type of mistake is described by Barry Schwartz (2004) in his well-publicized recent book *The Paradox of Choice.* He emphasizes the mistake of trying "too hard" to optimize. The "maximizers" he identifies by an abbreviated personality test seem to optimize without regard to the costs of the time and effort devoted to deliberation about a choice. Of course, this is not true optimization in a larger sense; a fully optimal choice must take into account deliberation costs. However, this raises an important issue for the measurement of utility. We argue that the utility function for everything other than deliberation costs should be measured by the choices that would be made by exactly such an agent who disregards deliberation costs.[26] Where people's

[26] Calculation and deliberation costs should be recognized just as much as any other costs an agent faces. The difficulties in modeling "bounded rationality" problems due to the "infinite regress" problem discussed by John Conlisk (1996) among others should not blind us to the obvious fact of deliberation costs. Because we do not view

preferences are similar to one another, this concept of utility approximates the utility that can be achieved when a small jury selected from a large group of people with similar preferences pays the deliberation costs for everyone in the whole group. For example, *Consumer Reports* is a practical effort to help people approximate this level of utility.

The general point that arises from thinking about mistakes is that, even when underlying preferences are fully consistent, choices arise from the interaction of preferences, ordinary constraints *and information structures.* The relevant information structures can include both external information constraints and internal cognitive constraints. In principle, internal cognitive constraints are no harder for economic theory to deal with than external informational constraints. For example, Woodford's (2002) model of monetary nonneutrality (based on the model of rational inattention in Christopher Sims, 2002) has been criticized for relying on extremely low bit transmissions rates. One way to defend Woodford's model is to locate the low bit transmission rates inside the small portion of the typical decision-maker's psyche devoted to thinking about macroeconomics.

Rayo and Becker (2005) provide an interesting example of explicitly modeling decision-making as the outcome of an underlying utility function ("evolutionary efficiency") filtered through internal informational constraints. Assuming a limit on the total number of gradations into which values of the underlying utility function can be distinguished, they show that an optimal deployment of that limited available total precision is to make fine distinctions in the neighborhood of values where an agent will actually be operating but only gross distinctions at outlying values.[27]

Could consideration of mistakes overturn the conclusion of subsection A that utility is rising? It is instructive to consider again the choice hundreds of thousands of people make every year to migrate from a poor home country to a foreign rich country. Certainly some regret their decision *ex post* and wish they had never migrated, and some even pay the fixed cost of returning to their home country to a situation no richer than before. But among those who are able in the end to bring their families as well, there is no evidence of widespread regret at migrating to a rich country. There is even less regret among the grandchildren of those who migrated, who escape most of the large fixed costs of migrating.

C. Inconsistent Preferences. Inconsistent preferences are more difficult to deal with than mistakes. Just as mistakes involving one's own preferences can be modeled as an underlying utility function together with internal information acquisition, transmission and processing constraints, inconsistent preferences are now routinely modeled as an intra-psychic game between multiple agents within the same person, each having a distinct set of preferences. However, if there is more than one set of preferences operating within a single individual, normative analysis faces a version of the Social Choice problem even for evaluating individual

---

the recognition of deliberation costs as a departure from "rationality" at all, we favor the more neutral term "bounded cognition" for what has traditionally been called "bounded rationality."

[27] In their paper, Rayo and Becker (2005) call the filtered version of underlying utility "happiness." However it is unclear in what way it would relate to happiness as we are using the term. In particular, like visual processing (which they use as an analogy), the filtered version of utility they discuss might operate at a very early unconscious or "automatic" stage in the sense of Colin Camerer, George Loewenstein and Drazen Prelec (2005) and so could be several cerebral processing steps prior to "happiness" in the experiential sense.

welfare. Unlike the standard Social Choice problem, there is no reason for a presumption of equal ethical value for all the different intra-psychic agents.

In order to do normative welfare analysis even at the individual level, one must take some stand on this intra-psychic social choice problem. Our proposal in this regard would be to rely, in all ordinary cases, on the psyche's own dispute resolution system. As Camerer, Loewenstein and Prelec (2005) discuss, the brain's controlled cognitive system is often brought into action to use deliberation to resolve disputes between other systems in the brain. Thus, we consider well-informed, thoughtful, revealed preference to be the best practical gold standard for an individual's preferences for the purposes of welfare analysis at the individual level.[28]

When multiple preferences coexist within the same individual, the whole-person utility function that is the solution to the intra-psychic social choice problem may not be the utility function that has the tightest relationship with happiness. Here, a key issue is which "selves" give affect reports to an interviewer. A relatively straightforward case is when the problem is short-sightedness in the sense of hyperbolic discounting of the sort described by David Laibson (1997), where there is one "self" in command at each point in time, in a known sequence. Other cases could be more complex. Although we consider it a high priority for the future, modeling the relationship between reported happiness and either whole-person utility or the set of utility functions within the psyche is beyond the scope of this paper. From here on, in discussing the relationship between happiness and utility, we will assume that the individual has only one set of preferences, which are internally consistent.

D. Is Happiness in the Utility Function? The principle of revealed preference indicates that happiness is in the utility function. Hundreds of thousands of people spend thousands of dollars each on therapy that is not reimbursed by insurance in hopes of becoming happier or at least less unhappy. Millions of people endure the significant negative side effects of chemical antidepressants in order to feel happier. Self-help books and magazines featuring cover articles on happiness sell briskly. Moreover, many products that may not actually make one's psychological state significantly more positive are advertised as if they will, as described in great detail by Melinda Davis (2002). Advertising aimed at suggesting that a product will improve one's brain state would not be so prevalent if a desire for positive affect were not an element of preferences. In sum, many people want to feel happier and are willing to sacrifice other things in order to attain that psychological state.

On the other hand, it seems clear that happiness is neither the only thing in the utility function nor a sufficient statistic for all of the goods that are in the utility function. To make this clear it is best to use the technical term "affect" for happiness as a reminder that we are talking about feelings. People care about things other than how they feel. Most obviously, they sometimes sacrifice current affect for a later benefit; for example (to take one of people's lowest affect activities according to Kahneman, Krueger, Schkade, Schwarz and Stone, 2004) it makes sense to spend time on household chores despite the low momentary affect associated with that activity because of the later benefits of having a clean house. There might be a later increase in affect as

[28] Insisting on transitivity is one aspect of "thoughtfulness" here.Thus, in principle, in assessing preferences, we would rely on an individual's deliberative choices for an entire menu of decisions at once, with an iterative process where the individual is forced to resolve non-transitivities.

a result, but it is not clear that the later benefits show up in affect and utility in the same way.[29] Second, people often sacrifice their own affect to benefit their children, as when one spends long hours at a grueling job to finance college educations for one's children. In standard economic models, the benefits to one's children show up in one's own utility function, but it is not clear that the benefits to others show up in one's own affect in the same proportion as in utility. Third, some people genuinely care about things that contribute to their lives but do not on average contribute to affect. For example, it would not necessarily be a bad decision to pursue excellence even if one knew that the effort would lower one's expected level of affect over a lifetime. Even striving for social rank—a dimension in which affect and utility track each other especially well—provides good examples of the divergence between affect and utility. Think of how many people have knowingly and deliberately sacrificed happiness (affect) for the sake of ambition. Some of these people would do it all over again if they had their lives to live over.

E. Persuasion about Preferences. Since most people do care about happiness at least somewhat to begin with, extolling the wonders of happiness and exhortations to value happiness more highly can often be effective tools for those who desire to persuade others to change their priorities—that is, to change their preferences and the utility function that would be needed to represent those preferences. To the extent that people *are* genuinely persuaded by such arguments, their utility function will shift to be somewhat more tightly related to happiness. The adverb "genuinely" in "genuinely persuaded" is needed to subtract out the effects of social pressure in which people are brought to outwardly assent to something they do not really agree with. From the standpoint of Pareto optimality, there is no reason to question the new, post-persuasion utility function if the persuasion is, indeed genuine. However, it seems only fair that people be made aware that it is not *illogical* to put a low valuation on happiness in one's preferences, if one so chooses in a top-down process of concretizing one's own preferences.[30]

5. Evidence that Happiness is Not the Same Thing as Flow Utility

Having laid out the definitions of felt happiness and choice-based utility both conceptually and operationally let us look more closely at their relationship. It would be convenient for many reasons if happiness in the sense of current affect were proportional to flow utility. Not only would this make a welfare measure based on a discounted sum of current affect equal to lifetime

[29] This is a very interesting empirical question. In testing whether intertemporal tradeoffs in utility match intertemporal tradeoffs in happiness alone, one must address the problem that people are not good at predicting their future happiness, as pointed out by Loewenstein and Schkade (1999). It may be possible to address this problem with some combination of educating people about the likely consequences of a decision for future feelings and eliciting what their expectations about future feelings are after that education to control for any remaining mispredictions of future feelings.

[30] Of course, in a large fraction of cases of attempted persuasion about preferences, the desired preferences for the other person will be given the rhetorical label "happiness," "true happiness," "genuine happiness," or "authentic happiness," regardless of how important happiness in the narrow sense of positive affect is in those preferences. For logical clarity (which can be at variance with persuasive power), the phrase "recommended preferences" can be substituted in place of "true happiness" or similar phrases. Aristotle's use of *eudaimonea* (the Greek word for happiness) in the *Nicomachean Ethics* (fourth century B.C.E.) can be seen as an example of using "happiness" as a label for such recommended preferences. Saying this in no way diminishes the cogency of Aristotle's recommendations.

utility, but such a simple measure of flow utility would make utility empirics more like production empirics, where one actually gets to see output directly. However, there are two reasons why it is very difficult to maintain that happiness and flow utility are even close to the same thing: the Easterlin Paradox and Hedonic Adaptation.

A. The Easterlin Paradox. As Easterlin (1974, 1995, 2003) observes, real per capita GDP and real consumption expenditure in the United States has risen dramatically in the last fifty years, but the percentage of people saying they are "very happy" has been falling slightly. The story is even more dramatic in Japan, where the percentage rise in per capita GDP is even more rapid, but the graph of subjective well-being is essentially flat. In other words, in developed societies, profound increases in average real income and in the objective standard of living over the last 50 years have not been associated with increases in the average happiness of their citizens. As argued above, it is not just a matter of money not buying happiness, since there are many other positive trends. In short, there is strong evidence that utility has gone up, but happiness has not.

B. Hedonic Adaptation. In addition to the difference in the trend behavior of utility and affect, the shorter-run dynamic behavior of affect is also quite different from the dynamic behavior of flow utility as normally modeled. As an empirical matter, affect in response to a number of important categories of changes in circumstances is subject to *hedonic adaptation*--regression of affect toward its previous level. Some of this evidence is surveyed in Frederick and Loewenstein (1999). In response to discrete negative events with lasting practical consequences, significant hedonic adaptation over time is observed for incarceration (Zamble and Proporino, 1990; Zamble, 1992), the loss of the use of limbs, (Wortman and Silver, 1987) and for serious burns (Patterson, et al., 1993). The death of a spouse seems to have a particularly long-lasting effect on affect. But Kaprio, Koskenvuo, and Rita's (1987) finding that suicide rates the week after a spouse's death are elevated almost tenfold for women and almost seventyfold for men suggests especially low affect immediately after the loss, which then moderates to some extent.

Some of the most striking data is that on lottery winners. Less than a year after winning the lottery, Brickman, Coates and Janoff-Bulman (1978) find that winners of large lotteries displayed only slightly higher life satisfaction. Frederick and Loewenstein (1999) interpret this as evidence suggestive of substantial hedonic adaptation since it is likely that many winners of large lotteries are ecstatic immediately after winning. More recently, Gardner and Oswald (2001) look at people receiving a windfall--primarily lottery winners--in the British Household Panel Survey. They find that winning £10,000 raises affect by *six times* as much in the first year as £10,000 per year in additional income. This comparison is suggestive of income having been subject to greater hedonic adaptation than the hedonic adaptation to the relatively recent windfall.

Brickman and Campbell (1971) refer to the implications of hedonic adaptation for the trend in affect the *hedonic treadmill*. Because of the close apparent connection between the Easterlin Paradox and the phenomenon of hedonic adaptation, it seems appropriate to search for a joint explanation.

C. The Implications of *Affect* Data for Hedonic Adaptation and the Easterlin Paradox. In unpublished work, Kahneman and Schwarz seem to have discovered another important fact

about hedonic adaptation: measures of current affect such as data from experience sampling show even stronger hedonic adaptation (mean reversion) than life-satisfaction or global happiness measures. (There is some discussion of this in Kahneman, Krueger, Schkade, Schwarz and Stone, 2004.) As discussed above in footnote 18, because life satisfaction and global happiness evaluations incorporate an element of autobiography *and* people's ideas about how they "should" feel, they will tend to show more permanent effects of events such as unemployment, as Lucas, Clark, Georgellis and Diener (2004) find.

What is even more serious, the likely influence of people's folk theories of how they "should" feel on life satisfaction and global happiness evaluations may account for some of the modest relationships with income that these measures show. Also, some of the variance in income comes from recent enough income innovations that the dependence of happiness on news we argue for below could account for some of the remainder of the correlation observed between income and life satisfaction. (Recall that life satisfaction and global happiness measures are significantly influenced by current affect.)

Given these hints, we predict that future research focusing on affect data as opposed to life satisfaction and global happiness evaluations will deepen the Easterlin Paradox and raise estimates of the extent of hedonic adaptation.

D. Hedonic Adaptation vs. Habit Formation. Note that hedonic adaptation is not the same thing as habit formation. Hedonic adaptation is a statement about happiness, as measured by psychologists. Habit formation is a statement about utility, as measured by revealed preference. For example, habit formation often refers to a tendency to do something more if you have done it in the past—an effect of past consumption on marginal utility. Of course, *if* happiness were proportional to flow utility *then* hedonic adaptation and habit formation would be tightly linked. This could be empirically problematic, because data on hedonic adaptation might then imply extremely strong habit formation. For example, suppose utility was of the form made popular by George Constantinides (1990), which can be represented as

$$v_t = E_t \sum_{j=0}^{\infty} \beta^j U(C_{t+j} - \theta H_{t+j}),$$

where v_t is lifetime utility, β is the discount factor for flow utility U, current consumption is denoted C, the "habit" H is a weighted average of past consumption levels, and θ is a parameter between zero and one. Given this form of the utility function, *if* happiness were proportional to flow utility, then evidence of complete hedonic adaptation would only be consistent with $\theta = 1$. For comparison, Joseph Lupton (2002) estimates $\theta \approx .75$ when estimating based on data for life-cycle portfolio choices and a value of θ close to zero when looking at consumption choices. The reason consumption data does not support a high value of θ is that, unless the lags in the habit H are quite long, a high value of θ implies there should be a strong autocorrelation for consumption growth rates that is absent in the data. If happiness were proportional to flow utility, matching the speed of hedonic adaptation would require a fast-moving habit, for which consumption data point to a value of θ near zero. Moreover, even the higher value of $\theta \approx .75$ would not match the observed extent of hedonic adaptation.

A more subtle discussion is required if—to match happiness data—someone suggests a type of habit formation that would not show up in empirical data *other* than happiness data. Suppose that everyone agreed, based on empirical results, that current affect A_t was given by $A_t = f(C_t) - f(C_{t-1})$. To make things even better, suppose that lifetime utility v_t could be represented as

$$v_t = E_t \sum_{j=0}^{\infty} \beta^j [f(C_{t+j}) - f(C_{t+j-1})].$$

One could then claim that affect was equal to flow utility, where flow utility was given by

$$U(C_t, C_{t-1}) = f(C_t) - f(C_{t-1}),$$

with lifetime utility v_t being given by

$$v_t = E_t \sum_{j=0}^{\infty} \beta^j U(C_{t+j}, C_{t+j-1}).$$

But a bit of algebra shows that

$$v_t = -f(C_{t-1}) + (1-\beta) E_t \sum_{j=0}^{\infty} \beta^j f(C_{t+j}).$$

Since C_{t-1} is already fixed at time t, and the multiplicative factor $(1-\beta)$ does not affect preferences, this utility function represents the same preferences over choices from time t on as the lifetime utility function

$$v_t = E_t \sum_{j=0}^{\infty} \beta^j f(C_{t+j}).$$

There are enough degrees of freedom in this example to force flow utility to be equal to affect. We argue, however, that in this instructive case it is clearer and more evocative of the existing economic literature to represent the lifetime utility function in the equivalent, but simpler, more convenient, and more familiar, form $v_t = E_t \sum_{j=0}^{\infty} \beta^j f(C_{t+j})$, where $f(C_t)$ is thought of as the flow utility function $U(C_t)$. The complexity in affect can then be represented in the *relationship* between flow utility and affect. In particular, the stipulated equation $A_t = f(C_t) - f(C_{t-1})$ can then be described by saying that "affect is equal to the first difference of flow utility."[31]

[31] Note that with a finite horizon, the two formally similar versions of the utility function would no longer represent exactly the same preferences. The lifetime utility function $v_t = E_t \sum_{j=0}^{T-j} \beta^t [f(C_{t+j}) - f(C_{t+j-1})]$ would imply a

While the two flow utility functions $U_t=f(C_t)$ and $U_t=f(C_t)-f(C_{t-1})$ are equivalent in the preferences they represent over choices at time *t* and beyond, could the difference between them bear on the hypothetical choice between the consumption bundle now and the consumption bundle fifty years ago discussed in Section 4 A? One answer is to point out that the individual and social choices we really face are those of the next fifty years, not of the past fifty years. Looking toward the future, we have the habits that we have from the past, and must take those as given. From this point of view, the two utility functions are fully equivalent.

Another answer is to carefully pose the hypothetical choice between different comprehensive consumption bundles in a way that takes into account all relevant habit formation. For example, imagine that one were forced to put one's newborn child up for adoption in one of two worlds, where one world has the comprehensive consumption bundle of fifty years ago, while the other world has the comprehensive consumption bundle we have now. Alternatively, assuming that per capita GDP and other objective circumstances improve as much in the next fifty years as they have in the last fifty years, would you rather put your newborn child up for adoption in the world that has the comprehensive consumption bundle we have now or the world that has the comprehensive consumption bundle of fifty years from now? Because it is hard to imagine the future in detail (even after conditioning on the values of some key statistics, as here) this is a more difficult question, but an important one.[32]

The closely related choice of which society one would wish to be born into is a crucial tool in John Rawls's (1971) extremely influential book of political philosophy *A Theory of Justice* (anticipated by Rawls 1951, 1958). In the Economics literature, choices between societies are also a crucial tool in John Harsanyi's (1953, 1955) theory of social welfare (discussed ably by Pattanaik, 1968.) Of course, these are very difficult choices to make. Nevertheless, revealed preference gives some guidance here, while a simple model of 100% hedonic adaptation would guarantee that happiness data could give *no* guidance for such choices.

E. Local and Global Marginal Thinking vs. Focusing Illusion. Just as the distinction between utility and happiness breaks any tight link between hedonic adaptation and habit formation, the

---

greater tendency to consume in the period immediately before death than $v_t = E_t \sum_{j=0}^{T-j} \beta^t f(C_{t+j})$. However, the lifetime utility function $v_t = E_t \sum_{j=0}^{T-j} \beta^t [f(C_{t+j}) - f(C_{t+j-1})]$ would still be equivalent to the lifetime utility function $v_t = E_t \left\{ \left[\sum_{j=0}^{T-j-1} \beta^t f(C_{t+j}) \right] + \frac{\beta^{T-t}}{1-\beta} f(C_T) \right\}$. This equivalent form with "flow utility" depending only on current consumption might easily be more convenient, despite the odd-looking coefficient on $f(C_T)$.

[32] Note that while there is good reason to hope that utility will be higher in the future, it is not clear that the Easterlin Paradox will continue into the future. It is possible that average long-run happiness will be significantly higher in the future.

distinction between utility and happiness should make one cautious in using happiness data to assert that people are making systematic optimization mistakes.[33]

David A. Schkade and Daniel Kahneman (1998) consider a thought experiment familiar to readers of David Lodge's (1978) comic novel, *Changing Places*, in which a professor from the gray English industrial city of Birmingham has the opportunity to spend a sabbatical year at Euphoric State on the shores of San Francisco Bay while a California professor takes his place as a visitor at Birmingham. Schkade and Kahneman (1998) study two groups of students, one residing in a gray Midwestern climate and the other in the brilliant sunshine of California. When surveyed, students in both locations have the same distribution of subjective well being. Both Midwestern and California students also predict that either they themselves or a student like them would be more satisfied with specific aspects of California including climate, outdoor activities, social life and cultural opportunities. Schkade and Kahneman explain their results in terms of a *focusing illusion*:

> When a judgment about an entire object or category is made with attention focused on a subset of that category, a focusing illusion is likely to occur, whereby the attended subset is overweighted relative to the unattended subset. In particular, when attention is drawn to the possibility of change in any significant aspect of life, the perceived effect of this change on well-being is likely to be exaggerated. (p. 340.)

While they do not conduct such an experiment, it appears that Schkade and Kahneman believe that a person who actually moved to California would not experience a permanent increase in measured happiness or satisfaction. This would seem logical in light of the equality of overall life satisfaction they observe among Midwestern and California students. Moreover, they cite other instances involving paraplegics, lottery winners and widowed spouses in which the positive or negative effect of these events on measured happiness is transient.

What do these results imply? Tacitly assuming that happiness can be set equal to flow utility, Schkade and Kahneman suggest that people mispredict utility for two reasons. First, because of the focusing illusion they overemphasize the importance of a particular aspect of life in California—say, climate—among the determinants of overall satisfaction. Second, and perhaps for the same reason, people fail to predict that their mood will adapt to local circumstances within a relatively short period of time. This appears to be consistent with Schkade and Kahneman's interpretation when they write, "At the individual level, the focusing illusion may lead to unnecessary initiatives. For example, it is not unlikely that some people might actually move to California in the mistaken belief that this will make them happier. (p. 345)

The theory we advance in this paper would predict the same pattern of survey results about happiness, both cross-sectionally and longitudinally, but the interpretation of the results would be different. When the prospect of relocating, say, from Ann Arbor to Berkeley arises, conventional economic theory suggests that an individual needs to consider global utility maximization by comparing the (ordinal) heights of two utility mountains, one corresponding to attainable levels

[33] In the absence of an adequate theory of the relationship between utility and happiness, it is best to be cautious about asserting that people are making systematic optimization mistakes even when it is clear that people *are* making mistakes in predicting the dynamics of happiness. We return to this issue after presenting our theory of the relationship.

of utility in Ann Arbor and the other to attainable levels of utility in Berkeley. The heights of these mountains depend on location-specific nontraded goods such as climate, topography and culture but also on variations in location-specific traded and partially traded goods that would be available to the person given wages, prices, employment opportunities, family and friends, leisure possibilities and so on. In conventional economic theory, an individual would make a migration decision by comparing the heights of the two utility mountains. Once in Berkeley, the utility mountain in Ann Arbor becomes irrelevant and an individual's decisions are concentrated on finding allocations of income and time to alternative bundles of traded, partially traded and non-traded goods that place her as close as possible to the summit of the local utility mountain.

We suggest that focusing is best understood not as an illusion, but rather as a mental act that plays a familiar role in economic theory. Conventional economic theory suggests that a consumer chooses an allocation that maximizes his utility subject to a budget constraint, a time constraint and other relevant constraints such as distance from family and friends. In finding this optimum, the consumer compares the marginal utility gained from a good with the utility value of its marginal cost in dollars, time, or social interaction. A mental calculation of marginal utility—a partial derivative—requires focusing because it asks how much utility would change holding everything else constant. The empirical evidence of focusing described by Schkade and Kahneman suggests that people are readily able to think about the positive or negative impact of a particular event or state, holding other aspects of life constant.

As is often noted both by economists and non-economists, the optimization task assumed to take place in standard economic theory is daunting in the complexity of its cognitive demands on both information and calculation. To find the local optimum associated with a given utility mountain corresponding to a given location and a given time, an individual may need to consider only variations in a small number of aspects of life because many others are already settled through past decisions, trials and errors. In most day-to-day decisions, focusing on the few dimensions at issue yields a large savings in deliberation costs. The person already has a job, a spouse and children, a home and, perhaps, the only significant decision at the moment is whether to go to a Chinese or Italian restaurant tonight. By contrast, a large decision involving changing a location, choosing a spouse or changing jobs will cause many aspects of life to change simultaneously. To find the optimum in such cases, the person needs some way to discover the highest utility mountain in a vast range of (high dimensional) mountains, each associated with a particular discrete choice. Just knowing that the next step has higher altitude is not enough.

Even in deciding about an actual move, a fully rational *homo economicus* might conduct a series of thought experiments, similar to those on Schkade and Kahneman's questionnaire concerning satisfaction with aspects of life in the Midwest or California, for each relevant aspect of life. If we assume that utility is additively separable in different aspects of the locations, then the total difference in utility from a move is

$$\Delta U = \Delta U_1 + ... + \Delta U_n .$$

Some of these aspects will be essentially the same in both locations, so the individual can focus on just those that are different, together with whatever combinations of aspects interact in a nonseparable way.

Daydreaming in a focused way may be a very helpful way of sorting through particular aspects of a location choice before getting on to the difficult task of making an actual location decision—which entails a summary valuation. After all, Hawaii, New Zealand or the South of France may have even better climates than California. But, among these, perhaps France and California are the best of these in culture and cuisine, on which a given person places a higher marginal utility value. But, after considering the value of these particular aspects of other mountains, it may be that the advantages of the current mountain dominate because it is close to family and friends, its properties are more certain and staying avoids the costs of moving. In long run equilibrium, migration takes place until the expected utility of individuals in the place they reside is at least as high as it is in other places. Moreover, in equilibrium, location-specific advantages such as climate will tend to generate offsetting compensating disadvantages such as high housing prices or low wages (Sherwin Rosen, 1986). It would not be surprising to find that utility is nearly equated in those locations that seem like relevant alternatives.

In this subsection, we have argued that our theory would yield the pattern of survey results reported by Schkade and Kahneman (1998), but that our interpretation of these results would be quite different from their theory of focusing illusion. This raises the question of whether there are any testable differences between the two theories. The most obvious concerns regret. If focusing creates an illusion that leads to the misprediction of utility, we would expect that, on average, people who actually moved to California would experience regret. In our theory, focusing is just an intermediate mental step in forming a summary judgment involving weighting a broad range of relevant issues and aspects of life. While the summary judgment might be erroneous from an ex post point of view, there is no reason to think that the errors are in one direction or the other—California might turn out to be even better than one imagined in a Midwestern college classroom. Similar testable differences between focusing illusion and our theory could be sought from data on regret from other sources such as new car purchases, dating behavior and many other areas of life.

F. Is Choosing Lower Long-Run Happiness Evidence of a Mistake or Evidence that Happiness and Utility are Not the Same Thing?

Equating happiness with utility is a key assumption in what has become an established theoretical consensus among happiness researchers in Economics as well as Psychology. This consensus challenges the validity of the foundations of conventional Welfare Economics which lie in revealed preference theory. In this section, we briefly describe the established consensus in the context of a specific empirical application by Frey and Stutzer (2004b) which examines the relationship between happiness and time spent commuting. Their assumption that happiness and utility are the same thing, in conjunction with the empirical relationship between happiness and commuting, leads them to conclude that individuals systematically mispredict utility. This conclusion, in turn, calls into question the key assumption of revealed preference theory: namely, that the chosen alternative yields higher utility to the consumer than those which are not chosen. Instead, in the spirit of "Subjective Well-Being is Desirable, But Not the Summum Bonum," (Diener and Scollon, 2003), we argue that Frey and Stutzer's (2004b) findings provide evidence that utility and happiness are empirically distinct, but do not bear on the validity of welfare theory based on revealed preference.

In the consensus theory, as summarized by Frey and Stutzer (2003, 2004b), reported subjective well-being is taken as a proxy measure for utility. Maintaining this very strong assumption opens up a wide range of empirical applications and allows for direct tests of conventional theory, as Frey and Stutzer (2004b) illustrate with their analysis of commuting time. Most people find commuting time unpleasant, but endure it as a necessary evil in order to work at a more interesting or better paying job while living in a nicer or cheaper location. In equilibrium, along the lines of Rosen (1986), they argue that individuals should sort themselves among locations such that the disutility of additional commuting is offset by compensating monetary or nonpecuniary benefits associated with a better job or residential location. In such an equilibrium, they argue, total utility should not be related to total commuting time.

Frey and Stutzer (2004b) test this hypothesis in a regression of happiness on commuting time using data from the German Socioeconomic Panel (GSOEP), holding a number of socioeconomic characteristics constant but leaving labor income free to vary. They find a significant negative coefficient for commuting time, contrary to the zero coefficient expected under a Rosen-esque theory. Of course, persons with higher (non-labor) wealth might have higher utility and choose both a better job or house and a shorter commute, thus creating a spurious negative correlation between happiness and commuting time. However, Frey and Stutzer find a significant (although somewhat smaller) coefficient on commuting time in an alternative specification in which permanent differences in wealth or other differences are controlled with the use of individual fixed effects. While their econometric model might be subject to other criticisms, for our purposes we provisionally accept their empirical finding of a negative relationship between happiness and commuting time. We also note that other investigators have suggested similar results for other kinds of decisions. For example, Gruber and Mullainathan (2002) find that cigarette tax increases raised the happiness of potential smokers; Schorr (1991) argues that people mismanage the balance between work and leisure, tending to overwork; and Loewenstein, Ted O'Donoghue and Matthew Rabin (2002) suggest that misguided purchases of consumer durables such as fancy cars occur because people overestimate the future satisfaction the purchase will bring.

The theoretical explanation advanced by Frey and Stutzer (2004b) and other happiness researchers for such findings is that people systematically mispredict the future utility or, equivalently in this view, the future happiness they will obtain by taking a given action. In particular, Frey and Stutzer (2004b) hypothesize that misprediction is most severe for goods or activities with extrinsic attributes that can be purchased in the market relative to those with intrinsic attributes involving nonmarket social interactions. Thus, while a commuter may choose his home and job with the expectation that the extra money he gets from lower rent or a higher wage will offset the utility loss resulting from spending less time with his family and friends, the negative relationship between happiness and commuting time is interpreted to imply that people systematically overestimate the future relative utility of the things they obtain with the extra income.

Standard economic theory can easily accommodate unsystematic mistakes by consumers, but has a much more difficult time making sense of systematic mistakes. If people get lower net utility from long commutes, why don't they learn this and change the location of their home or job

accordingly? Frey and Stutzer (2004b, p. 9) explicitly address this issue. They argue that the formation of expectations about future utility depends on reconstructions of feelings in the past. Failure to learn from mistakes results because "…remembered utility and predicted utility become similar and relatively independent of actually experienced utility." For example, they cite studies in which participants on vacation or holiday trips enjoyed the actual trip less than they had predicted, but report enjoyment levels similar to the ones predicted when they recall the experience afterward.

This is a remarkable argument. Revealed preference suggests that the people who took a vacation gained expected utility. Moreover, in recalling the trip they believe they actually received as much satisfaction as they had expected to get. Presumably, they felt no regret. Their decision appears to involve no mistake according to standard revealed preference arguments and certainly nothing in their recalled experience would cause them to be less likely to take a similar trip in the future. Despite all that, it is alleged that these people actually experienced less utility during the trip than they had expected to receive. This discrepancy is interpreted as a mistake and the failure to notice it after the fact is regarded as the reason that people do not learn from experience and correct their mistakes. Hence, misprediction of utility is common, causing people to make wrong decisions repeatedly which, against their own interest, result in lower levels of experienced utility than could be achieved by alternative decisions. The hypothesis that utility misprediction is relatively greater for actions with extrinsic aspects suggests that materialistic people will be most harmed by these mistakes.

In our view, to make a convincing argument that the individuals were making a mistake, one would need either to find evidence of regret or indications that being presented with the purported evidence of misprediction of utility caused people to want to change their decisions. In the case of commuting, we do not think people would be surprised to be told that commuting is quite unpleasant. Learning evidence that it is difficult to buy much happiness with money or that the effects of additional money on happiness are transient could have more impact on people's decisions. In the case of trips, forgetting some of the annoyances of travel may, in fact, distort people's decisions; being reminded of these annoyances might affect their decisions to some extent. However, some of the most important benefits of a trip are precisely the memories one brings back. To the extent those memories are positive, the traveler has achieved one of the main objectives of a trip—with the forgetting of annoyances serving as a helpful aspect of the household production function for vacation memories. The incidence of regret and second-thoughts after being presented with relevant data is ultimately an empirical matter for which the quantitative size of effects is just as important as the qualitative direction of effects.

In the absence of evidence of regret or second-thoughts upon being presented with relevant data, the other possibility (which we highlight) is that utility and happiness are not the same thing. Under this alternative, the interpretation of much of the evidence cited by happiness researchers about utility misprediction and systematic mistakes in decisionmaking is simply misleading. While we present our argument using a formal model in the second half of this paper, it is useful to provide some informal intuition now for our contention that evidence from the happiness literature is not inconsistent with a conventional economic model of rational (albeit not omniscient) utility maximizing consumers. It seems quite reasonable, as Diener and Scollon (2003) argue, to assume that maximizing subjective happiness is not the only goal of many

consumers because happiness competes with other values or objectives, some of which do not have positive effects on affect. Concretely, much like Becker (1965) or Lancaster (1966), we think of happiness as the outcome of one of a number of household production processes each of which combines inputs of goods, time, and social and physical environment to generate outputs of final commodities according to a household technology. For instance, in the commuting example of Frey and Stutzer (2004b), an individual may endure unpleasant commuting in part because it affords additional money or a more desirable residential location that enables him to buy nice things for his wife and children, to have his children attend a better school or to be able to contribute to a charity to relieve the suffering of others. An empirical question, mostly not addressed in the happiness literature, is whether each of these ways to use money has the same effect on subjective mood or happiness. It seems possible that they do not but, nonetheless, that the individual would be willing to sacrifice his own happiness to benefit others. If so, the negative correlation between commuting distance and happiness observed by Frey and Stutzer (2004b) is quite consistent rational with utility maximizing behavior by persons whose preferences include goals beyond narcissistic fixation on their own pleasure. *Though an altruistic motivation makes the example especially clear, the same logic applies if the objective the individual is pursuing in preference to happiness is a non-altruistic goal.*

G. Summary of the Argument that Utility and Happiness are Empirically Distinct. Here is the underlying structure of the argument that utility and happiness are empirically distinct. First, using standard utility representations, utility has a strong upward trend, while happiness has very little trend. Moreover, happiness is strongly mean-reverting even after permanent changes in circumstances, while utility is not. Second, if one is willing to use nonstandard utility representations (including the flexibility one has in choosing flow utility functions that add up to equivalent lifetime utility functions), one can say the following:

> (a) On one hand, if changes or innovations in lifetime utility were the *only* component of happiness, then maximizing happiness and maximizing lifetime utility would be essentially equivalent; indeed, happiness could even be viewed as an exotic way of representing lifetime utility *except* that since happiness is focused on changes, it still could not represent preferences over initial levels or initial paths. To put the issue dramatically, though happiness is quite tightly linked to utility in this case, because it is focused on changes, happiness provides no representation of people's views over which society it is best to be born into.
>
> (b) The frequent use of the concepts of utility and happiness to make social welfare statements makes it ill-advised to dismiss the representation of preferences over which society to be born into as unimportant or meaningless. Indeed, this kind of preference is closely related to important conceptions of social justice. These preferences over different comprehensive social situations do not necessarily line up with measured happiness.
>
> (c) Any evidence for persistent, predictable effects of choice variables on happiness imples that changes or innovations in lifetime utility are *not* the only component of happiness.

(d) The fact that at times people knowingly, thoughtfully and without regret make choices that predictably lower their mood, day after day, implies that utility and happiness are empirically distinct.

All of the statements (a—d) remain true regardless of what utility representation one uses for a given set of preferences. Further discussion of these arguments must wait until we have laid out our model.

6. An Integrated Theoretical Framework for Utility and Happiness

Given the empirical evidence that utility and happiness behave very differently, what is the relationship between happiness and utility? In this section we will explain the main elements of our answer in a model with more structure than is really needed. Within that structure we propose that

$$affect = baseline\ mood + elation,$$

where *elation* is short-run happiness—which depends on recent news about lifetime utility and *baseline mood* is long-run happiness—which is a subutility function much like health, entertainment, or nutrition. There is a two-way linkage between affect and utility in this theory. First, baseline mood is an argument of the flow utility function. Second, elation is a function of news about lifetime utility.

One weakness of the approach in this section—using relatively well-defined functional forms—is that, because of the illusion of cardinality for von Neumann-Morgenstern utility, it does not sufficiently emphasize the ordinal nature of utility. The Appendix takes a more general axiomatic approach, which allows us to more clearly demonstrate the consistency of our theory with Ordinalism.

A. News and Short-Run Happiness. To motivate the mathematical model below, let us begin with the observation that—although the relationship between circumstances and happiness is weak in the long run—all the evidence suggests that subjective well-being responds in an intuitive and important way to *news* about objective circumstances. For example, subjective well-being rises significantly after experimental subjects find a dime and falls significantly after experimental subjects are given negative test results (e.g., Schwarz, 1987). The theoretical outline we propose builds on these observations by positing that a major component of affect depends directly on *news* about objective life circumstances that has arrived over the last few months rather than on the *level* of circumstances. This assumption is consistent with the general observation that people evaluate changes rather than states, an assumption that is also central to Prospect Theory (Kahneman and Tversky, 1979).

We call the component of happiness due to recent news about lifetime utility *elation*. *Dismay* is the algebraic opposite of elation: *dismay* = -*elation*. If expectations are rational, standard results about rational expectations imply that *news*—dynamic revisions to rational expectations—will be zero-mean and unpredictable. As a result, *elation*—which is a function of recent news—will be strongly mean reverting. Intuitively, news doesn't stay news for very long. At the psychological level, the initial burst of elation dissipates once the full import of news is emotionally and cognitively processed.

B. Happiness in the Utility Function. Since Gary Becker's (1965) pioneering work, much of the activity of a household outside of paid work has been reconceived as household production of goods. Becker (1965) emphasized the concept of household production as a way to study the structure of the household's utility function. For example, a household may undertake many activities and purchases all focused on preserving health, such as buying and consuming

vitamins, exercising, and going to the doctor. It often aids intuition to think of the health subutility function as giving the output of a household production function for health. We think of the part of happiness *not* due to recent news about lifetime utility as this kind of subutility function—or equivalently as the output of a household production function.

We call the part of happiness *not* due to recent news about lifetime utility *baseline mood*. In particular,

> 1. Any predictable aspect of happiness is part of baseline mood. This includes any persistent aspect of happiness.
>
> 2. Any aspect of happiness that *would* be predictable if the relevant arguments of the subutility function were predictable is a part of baseline mood. The pleasantness of one's current activity falls into this category. [34]

Physical health provides a good analogy for baseline mood. Like health, baseline mood

- can be measured independently of its arguments (inputs);
- is only one argument of the flow utility function;
- depends on different things than flow utility does—or on the same things with different; weights
- has a complex household production function or subutility function.

Ultimately, it is an empirical matter what baseline mood depends on, but provisionally, we view baseline mood as depending on factors such as:

a. genes[35]
b. psychologically active drugs, such as Prozac
c. sleep
d. exercise[36]
e. eating habits
f. time spent with friends[37]
g. social rank[38]
h. the pleasantness of one's current activity.[39]

Viewing baseline mood as one of the arguments of flow utility allows the powerful language of price theory to be applied to baseline mood, just as to health. For example, Hall and Jones (2004) argue that health is a luxury good in the sense that continuing increases in per capita income will increase the budget share devoted to health-related expenditures. Similarly, one

[34] See Kahneman, Krueger, Schkade, Schwarz and Stone (2004) on the average level of affect experienced during different activities. As one unsurprising example, people experience higher affect while eating than the affect they experience while doing housework.
[35] See Diener and Lucas (1999).
[36] See Thayer (1989), Biddle and Murtrie (1991), Steptoe, Kimbell and Basford (1996) and Argyle (1999).
[37] See Lewinsohn, Sullivan and Grosscup (1982), Reich and Zautra (1981) and Argyle (1999).
[38] See Luttmer (2004).
[39] See Kahneman, Krueger, Schkade, Schwarz and Stone (2004).

might argue that continuing increases in per capita income are likely to increase the budget share devoted to baseline-mood-related expenditures.[40]

A key limitation on our ability to apply price theory to baseline mood is the possibility that people may not have accurate knowledge of the production function for baseline mood. People's expenditures of time and money will depend on their beliefs about the production function for baseline mood rather than the true function. Pursuing the analogy to health again, it seems reasonable that, just as people don't know the true production function for health, they may not know the true production function for baseline mood. In principle, the discovery and dissemination of facts about the determinants of baseline mood could have large positive welfare effects.[41]

One factor that could make it especially difficult for people to figure out the determinants of baseline mood is the salience of the component of happiness due to elation. Although the elation mechanism has its own functions, from the standpoint of figuring out the determinants of baseline mood, elation acts as noise.

To the extent that people do understand the determinants of baseline mood, price theory can contribute in important ways to an understanding of long-run happiness. Consider, for example, the negative correlation that has sometimes been found between "materialism" and happiness. Robert Lane (2000) gives a discussion of the mixed empirical evidence for such a negative correlation. In assessing the evidence, it is also important to be aware of the partial tautology in relating measures of unhappiness to materialism indices that contain many survey items measuring dissatisfaction and griping. Nevertheless, in order to make the logical point as clearly as possible, suppose it could be documented conclusively that materialism, in the narrow sense of valuing material goods highly, lowers happiness. Price theory suggests that as long as there is any tradeoff between happiness and material goods, those who value material goods more compared to happiness will choose a bundle with more material goods (as often found for those who are more materialistic) and less happiness. The mechanics of the tradeoff could, for example, be due to decisions such as the decision of whether to commute further to a higher paying job discussed in Section 5F. Materialism lowering happiness would be similar to the effect preferences have on any choice between two distinct goods—such as when those who place an extremely high value on career success have worse physical health because they do not make time to exercise or see the doctor.

---

[40] The hypothesis that in the future of rich countries baseline mood will be a luxury good is inspired by Maslow (1943), who argues that once basic needs (such as physiological and safety needs) are satisfied, higher needs (such as needs for love, belonging, esteem and actualization) come to the fore. Both long-run happiness at home and long-run happiness at work might exhibit strong income effects. However, one bit of evidence running contrary to this idea that baseline mood is a luxury good is that in the Hindhu and Buddhist traditions a great deal of time and effort were often devoted to baseline-mood-raising meditation even thousands of years ago at much lower levels of per capita income than today.

[41] This view of the value of pinning down the determinants of baseline mood is consistent with the program of Positive Psychology, as described by Seligman (2002).

[43] Television may have enhanced the negative effect of social rivalry on happiness by leading people to believe the distribution of income and other advantages in society is higher than it actually is, leading people to underestimate their true social rank. See O'Guinn and Shrum (1997).

Another important application of price theory is to the Easterlin Paradox itself. Even after accounting for the elation mechanism, since baseline mood is likely to be a normal good, there is still a version of the Easterlin Paradox that we must confront. With people much richer now, why don't they purchase more baseline mood? Trends in the externalities related to community and social rivalry[43] and any exacerbations of internal conflicts can certainly contribute toward an explanation, since most of these externalities and internal conflicts are likely to figure into happiness at least as strongly as they figure into utility. Lack of knowledge of the true production function for baseline mood could also contribute in an important way toward explaining this version of the Easterlin Paradox. But there may also be a price-theoretic element to the explanation. *Although income has gone up, the price of baseline mood may have risen.* The most likely reason for this is if many of the inputs into baseline mood are time-intensive, such as exercise or time spent with friends. With the price of baseline mood higher, people may choose to expand their consumption of other goods rather than baseline mood. The greater people's willingness to substitute between baseline mood and other goods, the smaller the price rise necessary to explain the Easterlin Paradox.

C. A Formal Model of Utility and Affect. The formal model in this subsection assumes a fully rational optimizing agent, with an internally consistent utility function, who is well informed about the nature of his or her own preferences. Indeed, we posit a lifetime utility function that is totally standard in how it is built up from flow utility U. The one difference from the standard case is that the flow utility function U is a comprehensive function that includes baseline mood M as an argument:

$$v_t = E_t \sum_{j=0}^{T-t} \beta^j U(K_{t+j}, X_{t+j}, M(K_{t+j}, X_{t+j})).$$

K_t is a potentially large vector of state variables encoding every aspect of the past that carries over to affect the present in a way that matters for utility, such as wealth, weight, level of fatigue, one's spouse being alive, oneself being alive, genes, etc., concatenated with a vector of exogenous variables (variables over which the individual has no control) such as the weather, the state of the entire economy and the level of consumption of the average person in society (to allow for direct social rivalry in consumption) and other external determinants of social rank. X_t can be a large vector of control variables representing aspects of the current actions that can be chosen, such as time allocations (including exercise, time with friends, and sleep), consumption (including psychologically active drugs and the services of psychotherapists), and portfolio choices. Baseline mood $M(K_t, X_t)$ is written as a general function of K_t and X_t—which also appear as direct arguments in the flow utility function. Thus, in one sense, the flow utility function is no different from a function of the vectors K_t and X_t, directly:

$$U(K_t, X_t, M(K_t, X_t)) = \mathcal{U}(K_t, X_t).$$

However, in applications, the dependence on the baseline mood subutility function M can provide additional structure to the flow utility function. Moreover, the specification of M makes predictions about what will be observed in affect data.

The lifetime utility function can also be written recursively as

$$v_t = U(K_t, X_t, M(K_t, X_t)) + \beta E_t v_{t+1},$$

pinned down also by the terminal condition that lifetime utility is uniformly zero after the end of all things the agent cares about: $v_{T+1} \equiv 0$. (This recursive form takes a step toward the Bellman equation without yet assuming optimization.)

We define the lifetime utility innovation ι_t ("iota") as

$$\iota_t = v_t - E_{t-1} v_t.$$

The lifetime utility innovation ι_t is a precise way of formalizing the concept of "news about lifetime utility." Since the lifetime utility innovation is the surprise in lifetime utility at time *t*, rational expectations implies that the lifetime utility innovation ι_t is mean-zero and unpredictable by all information available at time *t-1* or earlier.

The recursive expression for lifetime utility can be lagged and rearranged to yield this equation for the lifetime utility innovation:

$$\iota_t = v_t - \beta^{-1}[v_{t-1} - U(K_{t-1}, X_{t-1}, M(K_{t-1}, X_{t-1}))].$$

Thus, the lifetime utility innovation is almost, but not quite, equal to a simple change in lifetime utility. It differs by removing the predictable part of the movement in lifetime utility due to the passage of time, whether from discounting or from flow utility becoming "water under the bridge."

Elation, in turn, is an increasing function of current and past lifetime utility innovations:

$$e_t = e(\iota_t, \iota_{t-1}, \iota_{t-2}, ...).$$

Finally, affect A_t itself ("happiness") is the sum of baseline mood and elation:

$$A_t = M(K_t, X_t) + e(\iota_t, \iota_{t-1}, \iota_{t-2}, ...).$$

Notice that in this framework, the utility function is defined first, in a way that is a straightforward extension of a standard form. Then elation (the news component of happiness) is modeled as dependent on lifetime utility innovations. Baseline mood (the non-news component of happiness) is modeled in a fairly agnostic way as a function of current variables and implicitly of lagged variables through the state vector K_t. To match empirical evidence about baseline mood, it is important to include non-marketed goods such as social rank in the arguments of baseline mood.

D. Evidence that Expectations Matter for Affect. One of the central predictions of this model is that expectations will matter for affect, since the lifetime utility innovations are given by $\iota_t = v_t - E_{t-1} v_t$, and elation is a function of current and past lifetime utility innovations. The importance of expectations for affect is indicated by the evidence surveyed in Frederick and

Loewenstein (1999) that advance notice of the death of a spouse reduces the size and duration of the drop in affect after the actual death of the spouse. The following passage from Frederick and Loewenstein (1999, p. 315) is especially close to the spirit of the model here: "Even if advance notice does improve post-outcome well-being, its *overall* effect on well-being is ambiguous, since receipt of the bad news may diminish the well-being of the person between the time the notice is received and the time the event actually occurs." In the model here, it is the processing of bad news that generates a period of lower affect, whether the primary bad news occurs before the actual death of the spouse or only at the time of the actual death.

Camerer, Loewenstein and Prelec (2005, p. 28) give a good summary of some remarkable neurobiological research relevant to the role of expectations in determining affect:

> An important feature of many homeostatic systems is that they are highly attuned to changes in stimuli rather than their *levels*. A dramatic demonstration of such sensitivity to change came from single-neuron studies of monkey responding to juice rewards (see Wolfram Schultz and Anthony Dickinson 2000). These studies measured the firing of dopamine neurons in the animal's ventral striatum, which is known to play a powerful role in motivation and action. In their paradigm, a tone was sounded, and two seconds later a juice reward was squirted into the monkey's mouth. Initially, the neurons did not fire until the juice was delivered. Once the animal learned that the tone forecasted the arrival of juice two seconds later, however, the same neurons fired at the sound of the tone, but *did not* fire when the juice reward arrived. These neurons were not responding to reward, or its absence … [ellipses and all italics in original] they were responding to deviations from expectations. (They are sometimes called "prediction neurons.") When the juice was expected from the tone, but was not delivered, the neurons fired at a very low rate, as if expressing disappointment.

These results are just the tip of the iceberg in the neurobiology literature. A great deal of evidence points to machinery in the human brain that generates sophisticated short-run expectations—expectations that people are not always consciously aware of. See for example John O'Doherty et al. (2003), Jay Gottfried, O'Doherty and Raymond Dolan (2003), Ben Seymour et al. (2004), Seymour et al. (forthcoming) and O'Doherty (2005).

E. The Evolutionary Significance of Elation. Though any such claim is highly speculative at this point, we are inclined toward Randolph Nesse's (2000, 2001, 2004, forthcoming) functional interpretation of affect as part of the motivational system for processing utility-relevant information. If something good happens, elation motivates the individual to think about what went right (in case there is a way to make it happen again) and how to take advantage of any new opportunities that may have arisen. If something bad happens, dismay (negative elation) motivates the individual to think about what went wrong (in case there is a way to avoid it in the future), and how to mitigate the harm of the new situation. On this view, elation and dismay are in the same genus as *curiosity,* which is part of the motivational system for processing information that is neither obviously good nor bad, but for which there may be value to finding out more. Indeed, experimental inductions of elated and depressed moods have been found to change individuals' strategy of information processing across a variety of tasks (for reviews see Schwarz, 1990, 2002 and William Morris, 1999). Elated people are especially good at seeing opportunities, while dismayed people are especially good at seeing dangers.

F. The Evolutionary Significance of Hedonic Adaptation. Thinking of a temporary jump in affect occurring after utility-relevant news as functionally related to information-processing makes the functional significance of hedonic adaptation similar to the functional significance of adaptation in other aspects of perception. Frederick and Loewenstein (1999, p. 303) make this comparison explicit:

> "Adaptive processes serve two important functions. First, they *protect* organisms by reducing the internal impact of external stimuli…. Second, they *enhance perception* by heightening the signal value of changes from the baseline level…."
>
> "Hedonic adaptation may serve similar protective and perception-enhancing functions…. persistent strong hedonic states (for example, fear or stress) can have destructive physiological concomitants … Thus, hedonic adaptation may help to protect us from these effects."
>
> "Hedonic adaptation may also increase our sensitivity to, and motivation to make, local changes in our objective circumstances…."

Rayo and Becker (2005) construct a formal model that spells out the logic of Frederick and Loewenstein's (1999) claim.

G. Speculations on the Evolutionary Significance of Baseline Mood. Certain kinds of persistent situations could call for heightened sensitivity toward opportunities or toward dangers. For example, moderately high social rank or good physical health may make it safe to look more for opportunities than for dangers. Thus, it could make sense for these situations to stimulate the same machinery that is turned on by the receipt of good news. The high variance of persistent individual differences in baseline mood suggests a frequency dependence in which there is an advantage to being a pessimist looking for dangers when most of the surrounding people are optimists who might miss dangers, while there is an advantage to being an optimist who sees opportunities if there are plenty of pessimists around to alert one to possible dangers, and few other optimists around to boldy seize opportunities.

One of the most interesting possibilities is that important aspects of the determination of baseline mood are just quirks in the affective system that have no functional significance. The mixed-strategy evolutionary equilibrium in which the fitness of moderately happy and moderately unhappy people is equal would reduce the strength of any evolutionary pressure against such quirks.

Regardless of how the "production function" for baseline mood arose, now that it is present, it makes sense to exploit it, just as Stephen Pinker (1997) argues that we exploit our sense of taste (designed, say, to motivate the search for nuts and ripe fruits) with cheesecake and our musical sense (designed, say, to help us distinguish the sounds of different kinds of objects) with symphonies and Rock and Roll.

H. Implications of the Integrated Framework for Utility and Happiness. There are three key implications of this benchmark model for the relationship between affect and utility. First, there is a clear distinction between the psychological concept of affect and the economic concept of flow utility. Affect is *not* equal to either flow utility or to the overall objective function.

Second, the elation component of affect depends primarily on unexpected *changes* in lifetime utility. For applications, the most important point about elation is that the theory here contradicts the notion that a temporary movement in affect is unimportant because of its short duration. To the contrary, *a temporary movement in affect may be extremely important as a signal of important utility-relevant news related to the long-term welfare of the individual.*

Third, baseline mood, while not a summary measure of flow utility, is something that people care about. As with health, the relative concern with raising baseline mood compared to raising consumption of other goods may increase along with per capita income, implying that the average share of effort and expenditures devoted to raising baseline mood may increase in the future.

Since elation depends on (mean-zero) *news* about lifetime utility, rather than on the *level* of lifetime utility, elation has no trend. Thus, utility can rise with per capita income while happiness has only the trend imparted by the growth rate of baseline mood. This guarantees that the economic concept of lifetime utility and the psychological concept of the temporal sum of affect over time put forward by Kahneman (1999) will be numerically distinct approaches to assessing overall welfare. Distinguishing clearly between utility and happiness allows scientific questions about utility and happiness to proceed in a way that respects the insights of both Psychology and Economics without prejudging the ethical question of the proper contribution of each concept to the assessment of overall welfare–an ethical question that revolves fundamentally around the extent to which one should trust people's immediate feelings and the extent to which one should trust people's choices as indications of what most enhances their welfare. In this ethical debate, traditional Welfare Economics has implicitly staked out a position in favor of utility as the better measure of overall welfare, but the case for Kahneman's (1999) proposal deserves to be thoughtfully considered as well.[44]

Maintaining a clear distinction between affect and flow utility also makes it possible to see where the psychological approach toward welfare assessment and the economic approach toward welfare assessment are pulling in the same direction. For example, social rank—whether appearing as an effect of other people's consumption or time use on baseline mood or on flow utility directly—will matter for both the psychological and economic measures of overall welfare. As another example, as long as baseline mood is an argument of the flow utility function, any advance in scientific understanding of determinants of baseline mood, and the dissemination of scientific knowledge about baseline mood to individuals in society will be important for both measures of overall welfare.

7. Elation Theory and the Confusion Between Utility and Happiness

Any adequate account of the relationship between utility and happiness must explain why these two concepts are often confused. Why is it that they often seem to mean the same thing? To answer this question, it is useful to compare maximizing lifetime utility with Kahneman's (1999)

[44] The strength of Kahneman's case depends in important measure on whether, as he argues, there is no way to construct a consistent underlying set of preferences from the contradictory decisions people make, even after following the approaches discussed above in Section IV, "Measuring utility."

proposal of maximizing the true mathematical expectation of the present discounted value of happiness[45] in the context of the theory presented above.

A. Maximizing the Present Discounted Value of Happiness versus Maximizing Lifetime Utility. To the extent that baseline mood is different from flow utility and to some extent controllable, maximizing the expected present discounted value of happiness as Kahneman (1999) recommends will be different on that account from maximizing lifetime utility. But what about maximizing the expected present discounted value of happiness when baseline mood is beyond the individual's control? In that case only elation will matter in maximizing the presented discounted value of happiness. Proposition 1 addresses this case:

Proposition 1: Given (i) rational expectations, (ii) perfect memory, (iii) happiness that is the sum of baseline mood and elation, (iv) baseline mood that is exogenous to the individual, and (v) elation that is a positive linear combination of lifetime utility innovations, as of time t, maximizing the expected present discounted value of affect is equivalent to maximizing lifetime utility.

Proof: Let elation e_t be given by

$$e_t = \sum_{\ell=0}^{n} a_\ell \iota_{t-\ell}.$$

Then the expected present discounted value of happiness is

$$E_t\left\{\sum_{j=0}^{T-t}\beta^j A_{t+j}\right\} = E_t\left\{\sum_{j=0}^{T-t}\beta^j M_{t+j} + \sum_{j=0}^{T-t}\beta^j e_{t+j}\right\} = E_t\left\{\sum_{j=0}^{T-t}\beta^j M_{t+j} + \sum_{j=-n}^{T-t} b_{j,t}\iota_{t+j}\right\},$$

where

$$b_{j,t} = \sum_{\ell=-j}^{n} \beta^{j+\ell} a_\ell$$

as long as time t is at least n periods away from death, and somewhat less if t is less than n periods from death. Using the definition of lifetime utility innovations, perfect memory and the fact that the expectation of lifetime utility innovations conditional on previous information is zero, one can simplify the expected present discounted value of happiness further, to

$$E_t\left\{\sum_{j=0}^{T-t}\beta^j A_{t+j}\right\} = \sum_{j=0}^{T-t}\beta^j E_t M_{t+j} + b_{0,t}(v_t - E_{t-1}v_t) + \sum_{j=-n}^{-1} b_{j,t}\iota_{t+j}.$$

[45] The extension of Kahneman's proposal to the true mathematical expectation in uncertain situations is not explicit in Kahneman (1999), but it seems a reasonable interpretation.

Given the exogeneity of baseline mood M and the perspective of time t, everything in this expression is fixed except for $b_{0,t}\, v_t$. Thus, maximizing the expected present discounted value of happiness is equivalent to maximizing $b_{0,t}\, v_t$, which in turn is equivalent to maximizing v_t.[46]

B. Maximizing Current Happiness. Note that under the assumptions of Proposition 1, maximizing *current* happiness alone is also equivalent to maximizing lifetime utility, since

$$A_t = M_t + e_t = M_t + \sum_{\ell=0}^{n} a_\ell \iota_{t-\ell} = M_t + a_0 v_t - a_0 E_{t-1} v_t + \sum_{\ell=1}^{n} a_\ell \iota_{t-\ell}.$$

Given the assumed exogeneity of baseline mood M_t, the only thing that is not fixed in this expression as of time t is the term $a_0 v_t$, so one does the same thing to maximize current happiness as to maximize lifetime utility. The reason a present discounted value of happiness is not required is that elation is already forward-looking.[47]

C. Why Utility and Happiness are Often Confused. Psychological evidence is accumulating that baseline mood can in fact, be modified deliberately—and in ways that go beyond the zero-sum game of acquiring social rank. But a lack of understanding of the determinants of baseline mood can make baseline mood seem exogenous. As noted above, one reason for this lack of understanding may be that a large fraction of the time-series variance of happiness may be accounted for by elation and dismay. To the extent that elation and dismay dominate people's perception of happiness, Proposition 1 indicates why people might think that utility and happiness are essentially the same thing.

It is when people do begin to recognize that baseline mood might be controllable that the distinction between utility and happiness becomes crucial. Understanding the ways in which baseline mood is controllable clearly matters for optimization. Understanding the distinction between utility and happiness is becoming important precisely because we are beginning to see a wider variety of ways to raise *utility* by raising *happiness* rather than being limited to raising *happiness* (temporarily) by raising *utility*.

8. Utility of Elation.

To the extent that people value transient happiness as well as lasting happiness, elation may enter the utility function. Because elation depends in turn on news about lifetime utility, putting elation in the utility function requires one to solve simultaneously for elation and lifetime utility. For that reason, we have delayed the discussion of elation in the utility function until this point.

A. Adding Elation when Elation is a Linear Function of Lifetime Utility Innovations. One key result, showing the robustness of the model of Section 6 to the addition of elation, is the following:

[46] Note that only exogeneity of the conditional *mean* of baseline mood is needed for this result. An ability to control the variance of baseline mood, with no effect on the mean, would still leave elation totally dominant in the expected present discounted value of happiness.

[47] In an analogy to exotic financial securities due to George Akerlof when he first heard about elation, elation provides a kind of *tranche* of current and future effects on flow utility.

Proposition 2: Given rational expectations, adding to the flow utility function a linear function of lifetime utility innovations (with positive coefficients summing to less than one) has no effect on the preferences represented by the utility function.

Proof: Using an asterisk to represent the modified flow utility and lifetime utility functions, let

$$U^*(t) = U(t) + \theta e_t = U(t) + \theta \sum_{\ell=0}^{n} a_\ell \iota^*_{t-\ell}.$$

where $\theta \sum_{\ell=0}^{n} a_\ell < 1$. Note that the relevant lifetime utility innovations will be those for the modified lifetime utility function. Modified lifetime utility is then

$$v^*_t = v_t + \theta E_t \sum_{j=-n}^{T-t} b_{j,t} \iota^*_{t+j},$$

where, as above, $b_{j,t} = \sum_{\ell=-j}^{n} \beta^{j+\ell} a_\ell$ as long as time t is at least n periods away from death, and somewhat less if t is less than n periods from death. The essential structure here is that modified lifetime utility v^*_t is equal to the original lifetime utility v_t plus the expected value of a linear combination of the modified lifetime utility innovations with positive coefficients running from n periods back, up to the lifetime utility innovation in the agent's last period. Because lifetime utility innovations have mean zero conditional on previous information, one can simplify this further to

$$v^*_t = v_t + \theta b_{0,t} \iota^*_t + \theta \sum_{j=-n}^{-1} b_{j,t} \iota^*_{t+j} = v_t + \theta b_{0,t} (v^*_t - E_{t-1} v^*_t) + \theta \sum_{j=-n}^{-1} b_{j,t} \iota^*_{t+j}.$$

The condition that $\theta \sum_{\ell=0}^{n} a_\ell < 1$ guarantees that $\theta b_{0,t} < 1$. Therefore, one can solve for v^*_t:

$$v^*_t = \frac{v_t - \theta b_{0,t} E_{t-1} v^*_t + \theta \sum_{j=-n}^{-1} b_{j,t} \iota^*_{t+j}}{1 - \theta b_{0,t}}.$$

Because $-\theta b_{0,t} E_{t-1} v_t + \theta \sum_{j=-n}^{-1} b_{j,t} \iota^*_{t+j}$ is fixed as of time t, as a representation of preferences over choices at time t, v^*_t is equivalent to $v_t/(1-\theta b_{0,t})$, which in turn is equivalent to v_t itself.

To recap the proof, when a linear combination of lifetime utility innovations is added to the lifetime utility function, (1) the future lifetime utility innovations do not affect decisions because their expectation is zero, (2) the past lifetime utility innovations do not affect decisions because they are predetermined and (3) the current lifetime utility innovation does not affect decisions because, to the extent it is not predetermined, it is perfectly correlated with the original lifetime utility function.

B. Manipulating the Timing of News and Manipulating Expectations. In the proof of Proposition 2, have we tacitly assumed a fixed information structure? Does adding elation that is a positive linear combination of lifetime utility innovations to the utility function affect preferences over information structures, even when both information structures would lead to the same decisions over other variables? For example, could it make people want to delay when they hear news in order to manipulate their own feelings? The answer is no. Because rational expectations take into account the information structure, there is no way to game the system with any rule set up in advance. Suppose for example, that you told your friend to tell you good news right away, but to withhold bad news. The Bayesian inference in rational expectations would cancel out any effect on the expected lifetime utility innovation, though it would certainly affect the *ex post* distribution of lifetime utility innovations. Formally, v^*_t can be expressed as a linear function of v_t and *past* expectations about lifetime utility. Choosing a different information structure *now* can only affect current and future expectations about lifetime utility. That includes choosing an information structure when your friend says "I know what happened, do you want me to tell you or not?" since any revelation is still in the future, if in the near future.

Of course, even a mean-zero effect on the distribution of lifetime utility innovations will affect lifetime utility when added elation is a nonlinear function of lifetime utility innovations or elation enters the utility function nonlinearly, as we discuss below. Also, imperfect memory of past expectations may provide an opening for gaming the system by trying to reduce one's remembered past expectations. This may be particularly relevant for the memories of past expectations parents transmit to a child about the child's prospects: the gap between parent and child can be one source of imperfect memory in a dynasty. More generally, an attitude of gratitude (whose value is not diminished by the triteness of the phrase) can serve the same purpose as manipulable memory. It often involves substituting comparisons with others in a worse situation for comparisons with one's own remembered past expectations or one's own deductions of what one ought to have expected in the past.

Given perfect memory, but irrational expectations, it may be harder to beneficially manipulate expectations than one might at first think, since then lowering one's expectations adds to flow utility in the future, but subtracts from flow utility now.[48] (As long as one is more than n periods away from death, this will have no effect on lifetime utility. Closer to death, pushing one's expectations down is harmful and pushing one's expectations up is beneficial.) It is when one can manage high expectations now, but remember them in the future as if they were low expectations that there is a real opening for beneficial manipulation of beliefs.

C. Do Mistakes about the Rate of Hedonic Adaptation Matter? We argued above that because utility and happiness are distinct, the psychological phenomenon of hedonic adaptation does not have any necessary implications for the shape of the utility function. In particular, if

[48] Nevertheless, there is evidence people do some of this kind of manipulation of expectations. Nisan (1972) finds that study participants taking an immediate test were less confident than those taking a test in 4 weeks. Similarly, Shepperd, Ouellette, & Fernandez (1996) find that college seniors were more muted in estimated first-job salaries than sophomores and juniors. (See also Shepperd, Findley-Klein, Kwavnick, Walker and Perez (2000).) In each case, confidence was reduced when proximity to performance outcomes was more immediate. We are grateful to Norbert Schwarz for cluing us in to this evidence.

flow utility depends only on baseline mood and not on elation, as in the model above, the determination of elation, including the rate of hedonic adaptation, has no effect on the lifetime utility function. Thus, when elation is not an argument of the utility function, misprediction of hedonic adaptation causes no material harm to utility maximization, contrary to the claims of Schkade and Kahneman (1998).

To pursue the question further, consider how much harm there is to mistakes about the rate of hedonic adaptation in the context of the model of subsection A, with a positive linear combination of lifetime utility innovations added to the flow utility function. Mistakes about the rate of hedonic adaptation are mistakes about the true values of the coefficients a_ℓ. Since the modified lifetime utility function is equivalent to the original utility function regardless of the values of the coefficients a_ℓ (as long as they are positive and add to less than one), mistakes about the rate of hedonic adaptation will not distort decisions at all and so will be costless!

In the light of the lack of harm to optimization from misperception of the rate of hedonic adaptation in this benchmark case, any serious claim of quantitatively significant harm to optimization from misperception of the rate of hedonic adaptation would require careful modeling. For example, when elation is a nonlinear function of lifetime utility innovations, or flow utility is a nonlinear function of elation, there is likely to be at least some harm from misperception of the rate of hedonic adaptation, but it is not clear how large this harm would be. In the case of imperfect memory, misperception of the rate of hedonic adaptation might cause one to exert too much or too little effort toward manipulating one's memories, but whether this results in a serious reduction in lifetime utility depends on how great the scope is for manipulation of memory.

One of the most important effects of underestimating the rate of hedonic adaptation is that it will cause an overestimation of the unconditional variance of elation, since the effects of unforeseen increases or decreases in lifetime utility seem like they will be long-lasting. An overestimation of the unconditional variance of elation should, in turn, cause an individual to overestimate the fraction of the variance of happiness due to elation and underestimate the fraction of the variance of happiness due to baseline mood. As shown above, this overestimation of the persistence of elation does not necessarily interfere with maximizing lifetime utility, but it *would* tend to push Kahneman's suggested alternative of maximizing the expected present discounted value of happiness in the direction of maximizing lifetime utility. Since elation embodies movements in lifetime utility, anything that exaggerates the importance of elation in happiness is likely to make maximizing happiness more like maximizing utility, as indicated by the extreme case of Section 7, where elation is the only controllable component of happiness.

D. Elation Nonseparable in the Utility Function. Given Proposition 2, the key issues arising from elation in the utility function are (1) imperfect memory and departures from rational expectations, discussed briefly above, (2) nonseparability of elation in the utility function, and (3) nonlinearity of the utility function in lifetime utility innovations once elation is substituted out.[49] Here we will barely mention the possible consequences of nonseparability.

[49] There are some other possible extensions of the model that we cannot give a serious discussion to here. One of the more interesting is the possibility that elation responds more to news about whether one's choices worked out

Nonseparability of lifetime utility innovations could make manipulating the timing of news optimal. For example, an altruistic person might want to throw a surprise party to take advantage of a complementarity in the recipient's utility function between elation and the presence of friends. Nonseparability of elation in the utility function can also generate wealth effects that modify the size of the effects of imperfect memory, nonrational expectations and nonlinearity.

E. Elation Nonlinear in Lifetime Utility Innovations as a Foundation for Prospect Theory. It is worth discussing in some depth the effects of nonlinearity of the flow utility function in lifetime utility innovations, because this can lead very naturally to a version of Prospect Theory. Rather than discuss the effects of nonlinearity in general, we focus on a particular type of nonlinearity, motivated by the evolutionary interpretation of elation in Section 6E. Consider the following set of features one might wish a model to exhibit:

a. Elation is proportional to the rate of cognitive processing of news.
b. Within bad or good news, the total amount of processing needed is proportional to the magnitude of the news.
c. Bad news requires more processing per unit of lifetime utility innovation than good news. (This implies a partial separation of the cerebral system for working through good news from the cerebral system for working through bad news.)
d. It takes longer to process a big chunk of news than a small chunk of news.

Let us present a toy model that illustrates how these features could arise. It is in continuous time to provide detail of how news is being processed and for convenience. News arrives once, at time 0. G is a positive state variable representing the cumulative amount of unprocessed good news. B is a negative state variable for the cumulative amount of unprocessed bad news magnified by the parameter $\psi>1$ to represent the additional difficulty of processing bad news.[50] Formally, at the instant of news, either G or B jumps according to

$$\Delta G = \max(\iota, 0)$$
$$\Delta B = \psi \min(\iota, 0).$$

The capacity for cognitive processing is γ, so that in the absence of additional news, and as long as $G>0$,

$$dG / dt = -\gamma .$$

Similarly, in the absence of additional news, and as long as $B<0$,

$$dB / dt = \gamma.$$

---

than to news about things beyond one's control. That would make it possible to manipulate elation by labeling good events as due to one's efforts, while labeling bad events as beyond one's control.

[50] We choose ψ –Greek psi in the shape of a pitchfork—to symbolize the hellishness of working through bad news.

Elation is γ for some time after good news, while elation is $-\gamma$ for some time after bad news.[51] The time it takes to get over a chunk of news is proportional to the size of the news. It takes ψ times as long to get over bad news.

The von Neumann-Morgenstern lifetime utility function at the moment the news is received at time zero is given by

$$\int_0^\infty \exp(-\rho t)[f(C_t) + \theta e_t],$$

with $\theta<\psi^{-1}$. If the lifetime utility innovation ι_0 is greater than zero, the present discounted value of elation $g(\iota_0)$ is given by

$$g(\iota_0) = \int_0^\infty \exp(-\rho t) e_t\, dt = \int_0^{\iota_0/\gamma} \exp(-\rho t)\gamma\, dt = \frac{\gamma}{\rho}\left[1 - \exp\left(-\rho \iota_0 / \gamma\right)\right].$$

If $\iota_0<0$,

$$g(\iota_0) = \int_0^\infty \exp(-\rho t) e_t\, dt = \int_0^{-\psi\iota_0/\gamma} \exp(-\rho t)[-\delta] dt = -\frac{\gamma}{\rho}\left[1 - \exp\left(\psi\rho \iota_0 / \gamma\right)\right].$$

Graphically, $g(\iota_0)$ is concave for positive lifetime utility innovations and convex for negative lifetime utility innovations, with a kink at zero. (The slope is 1 for small positive values, but $\psi>1$ for small negative values.) $g(\iota_0)$ has an asymptote at γ/ρ at $+\infty$ and an asymptote at $-\gamma/\rho$ at $-\infty$. This is a shape familiar from Prospect Theory (Kahneman and Tversky, 1979).[52]

Let v_- be the level of lifetime utility immediately before the news at time zero, while v_+ is the lifetime utility immediately after the news. The news tells the constant level of consumption C that will prevail from time zero on. Therefore,

$$v_+ = \frac{f(C)}{\rho} + \theta g(v_+ - v_-),$$

which can be solved uniquely for v_+, since the maximum slope of θg is $\theta\psi<1$:

---

[51] A more realistic model might make the flow of elation and dismay increasing in the magnitude of the news, and greater in absolute magnitude for bad news than good news, but this assumption will do for our example.

[52] One facet of this model worth pointing out is that, other than the kink at zero, the curvature of $g(\iota_0)$ depends on the discount rate ρ. It may be that this will not provide sufficient curvature to match the observations that motivate Prospect Theory. But in a more complex model, hyperbolic discounting could come to the rescue by providing a high discount rate in the first little while. In any case, it is easy to add mechanisms that generate more curvature. Most obviously, one can posit that the amount of cognitive processing required goes up less than proportionately with the magnitude of news, so that $\Delta G = \varphi(\iota)$ and $\Delta B = -\psi\varphi(-\iota)$, where the function $\varphi(\cdot)$ is increasing and concave for positive values and uniformly zero for negative values.

$$v_+ = h\left(\frac{f(C)}{\rho}, v_-\right).$$

Also,

$$v_- = E_- h\left(\frac{f(C)}{\rho}, v_-\right),$$

which has a unique solution for v_- since h is decreasing in v_-. To get a little more intuition, think of what happens for small θ. If θ is small, then

$$v_+ = \frac{f(C)}{\rho} + \theta g\left(\frac{f(C)}{\rho} - v_-\right) + O(\theta^2).$$

Thus, the lifetime utility function takes some of its curvature from $f(C)$ and some of its curvature from g. If $f(C)$ has the functional form of decreasing absolute risk aversion, then as the agent becomes richer, more and more of the curvature of the lifetime utility function will come from g. Also, since g is kinked at zero, the function g—which comes from the agent's awareness of the affective consequences of good and bad news—will dominate the agent's choices between small risks. For choices among large enough risks, the fact that g has a limited range, bounded between two asymptotes, could make the curvature of f the dominant factor.

One aspect of these equations that may seem esoteric is the self-referential aspect of lifetime utility depending on its own innovation. In order to see more clearly how that self-dependence is resolved, it is helpful to look at a different approximation: the approximation for small lifetime utility innovations. For small positive lifetime utility innovations,

$$v_+ \approx \frac{f(C)}{\rho} + \theta(v_+ - v_-),$$

so that

$$v_+ \approx (1-\theta)^{-1}\left[\frac{f(C)}{\rho} - \theta v_-\right]$$

and

$$\iota_0 \approx (1-\theta)^{-1}\left[\frac{f(C)}{\rho} - v_-\right].$$

For small negative lifetime utility innovations,

$$v_+ \approx \frac{f(C)}{\rho} + \psi\theta(v_+ - v_-),$$

so that

$$v_+ \approx (1-\psi\theta)^{-1}\left[\frac{f(C)}{\rho} - \psi\theta v_-\right]$$

and

$$\iota_0 \approx (1-\psi\theta)^{-1}\left[\frac{f(C)}{\rho} - v_-\right].$$

There are two interesting results apparent from this approximation. First, the realization of consumption at which there is no surprise is where $\frac{f(C)}{\rho} = v_-$, just as it would be if elation were not in the utility function. Second, the kink at this level is made sharper by the way ψ appears in the factor $(1\text{-}\psi\theta)^{-1}$. This factor looks formally like a Keynesian multiplier—reflecting the self-referential aspect of lifetime utility depending on its own innovation.

The model above is only a toy model, but we think it accurately indicates the likely flavor of a more general treatment of nonlinear von Neumann-Morgenstern preferences over elation: given the structure of elation, Prospect Theory can easily arise from rational preferences over one's own emotions.[53] Such an affective foundation for Prospect Theory puts Prospect Theory in context. For example, this kind of model predicts that Prospect-Theory-like behavior will arise where the affective consequences of a choice are much larger than the non-affective consequences of a choice. Also, to the extent that Prospect Theory arises from the affective consequences of choices, affect data will be helpful in understanding people's choices, even though it will not be the whole story.

9. Implications of the Theory for Happiness Empirics

Even without the extensions discussed in Section 8, the integrated framework for utility and happiness laid out in Section 6 has many important implications for empirical work using happiness data. One of the most basic tests of the value of our framework is whether pursuing these implications for empirical work turns out to be fruitful.

A. The Time-Series Properties of Happiness Matter. The most obvious implication of our framework is the need for more research on the time-series properties of happiness. For example,

- How fast and how extensive is hedonic adaptation for affect, as compared to the (so far) better-studied hedonic adaptation for life satisfaction and global happiness measures?
- Do the time series properties of happiness have any implications for econometric practice in research to identify the determinants of happiness?

We have work in progress along both of these lines, but the details must to left to other papers.

[53] These are psychological preferences in the same sense as those appearing in the "psychological games" studied by Geneakoplos, Pearce and Stacchetti (1989).

B. Price Theory Can be Used to Study Baseline Mood. Second, the theory of baseline mood implies that standard price theoretic tools can be applied to the low-frequency movements of happiness. For example, the dollar value people place on feeling permanently happier can be gauged by how much they are willing to pay for psychotherapy in time and money (beyond what insurance pays for), divided by the effect of the psychotherapy on happiness. The less effective psychotherapy is at actually raising happiness, the higher the implied valuation on happiness. In the case of antidepressants, in addition to the monetary and time cost, one would have to determine how much people would be willing to pay to have an antidepressant free of side effects and add that value before dividing by the effect on happiness. Such ratios can begin to identify the marginal value of happiness.

Several other price-theoretic issues have been discussed above. Even in the context of our framework, normality of baseline mood still leads to a version of the Easterlin Paradox. It is important to construct measures of the price of happiness over time to see if an upward trend in that price can explain why people are not choosing higher baseline mood in their ever-expanding consumption bundle.

As mentioned above, one limitation in the use of these price theoretic tools is that they depend on knowing people's beliefs about the household production function for baseline mood. Would people do more things that add to happiness if they knew what they were? For example, there are some hints that, in addition to its other benefits, getting more sleep might add significantly to happiness.[54] If this is true, and people knew it, this could place a strong upper bound on the value people place on happiness (the hourly wage divided by the effect of an extra hour's sleep on happiness), but such a conclusion would only be warranted if people really knew exactly how much benefit an extra hour's sleep would have for happiness. On might obtain a more reasonable estimate of the value of happiness by conditioning on people's reports of how much they *believe* an extra hour of sleep each night would add to happiness.

Given measures of the marginal value of happiness, any evidence about the determinants of happiness should be included in cost-benefit analyses. If the marginal dollar value of happiness is high, it could motivate ever more careful empirical work to measure the strength of the effects of variables on long-run happiness. In particular, it could motivate many clever minds to look for good instruments for the possible determinants of long-run happiness.[55]

In principle, the application of price-theoretic tools to baseline mood should yield tests of the theory as well as applications. This kind of test of the theory is likely to emerge over time as the measures of the relevant concepts are refined.

C. The Elation Theory is Readily Testable. Because the theory of elation is the most highly structured aspect of our theory, it is also the most readily testable. In particular, we hope to test whether or not people's hypothetical choices between alternatives A and B always match their

[54] Norbert Schwarz, personal communication, and Kahneman, Krueger, Schkade, Schwarz and Stone (2004). The results are not definitive because of the lack of a good econometric instrument for hours of sleep that is known *a priori* not to affect happiness directly.

[55] Kerwin Charles (2002) is a good example of the kind of attention to exogeneity in happiness research that we mean.

predictions of how happy they would be immediately after receiving the news that A had happened or that B had happened with no action on their part. Our theory predicts that people will choose the alternative that would seem like the best news to them (as indicated by their happiness immediately after hearing). To the extent that this does not seem like a very daring prediction, it indicates that the reader has a prior belief in favor of one of the key linchpins of our theory. Nevertheless, it *is* a testable prediction. It is *not* true by definition.

D. Elation Provides Information about Preferences and Expectations. The theory of elation implies that, if it is possible to control for variation in baseline mood, the response of happiness to news will give direct information about preferences. Indeed, the elation component of happiness is essentially an "excess returns" measure for lifetime utility. Therefore, in principle, happiness data can serve as the basis for exactly the same kinds of "event studies" as those carried out using data on excess financial returns.

To implement this insight about the use of happiness data for "event studies," it would be very helpful to have a regular monthly, or even daily, time series on average aggregate affect. This would allow a test of average preferences over aggregate events. In particular, after accounting for the lagged effects of the previous months' news, the theory of elation implies that whether affect goes above baseline or below baseline indicates whether the month's news has been on average good news overall or bad news overall. Since many things happen in a month, each month's data would give information about whether a different innovation vector for the expected consumption bundle represented an increase or a decrease in lifetime utility. Over time, this would tell a great deal about average preferences for aggregate events. Information about preferences for aggregate events is particularly valuable because many of these things do not have regular markets. For example, one might want to know about the relative importance people put on geopolitics compared to economics. Any month in which there is good economic news but bad geopolitical news, or the reverse, would provide relevant information. Election returns are often read as giving this kind of information, but affect data immediately on the heels of news may give more detail. At a minimum, high-frequency happiness data serves as a kind of general-purpose poll question that can give useful insight into how people feel about any big event that comes along.

Kimball, Helen Levy, Fumio Ohtake and Yoshiro Tsutsui (2006) report a pilot study using a few months' worth of happiness data on the University of Michigan Surveys of Consumers. They find a significant dip in measured happiness both in the first week of September 2005, right after Hurricane Katrina, and in the week after the earthquake in Pakistan that occurred in October 2005. Adaptation to the hedonic effects of these national and international news events was close to complete after two weeks. The dip in happiness after Hurricane Katrina was significantly greater in the South Central region of the United States, closest to the hurricane's landfall. The size of the average U.S. reaction to the earthquake in Pakistan is almost as great as the size of the average U.S. reaction to Katrina. Although the human toll from the earthquake in

Pakistan was much greater than from Katrina, this still indicates a surprisingly high degree of concern for people on the other side of the world if we are correct in our hypothesis that the size of short-run spikes in happiness indicates the magnitude of the implications of news for lifetime utility. Tsutsui, Kimball and Ohtake (2005) apply a similar event-study methodology to the hedonic reactions in a Japanese sample to the overwhelming electoral victory of Japanese Prime Minister Koizumi in October 2005.

In situations where preferences are clear, the theory of elation draws a strong link between happiness and expectations. This provides another avenue for testing the theory. Section 6D discusses some of the supporting evidence that has already been established on this score. More can be done in this area. One of the most interesting tests would be in areas where people are known to violate rational expectations or where the memory of past expectations is likely to fade. Here the test would be to see if the pattern of people's reported happiness matched the quirks in their expectation formation and memory.

Assuming that the elation theory is valid, it may have relevance for the survey measurement of preferences and expectations quite broadly. On the preference side, since the elation mechanism seems to be fairly automatic as a psychological process, it may be that it is easier and more reliable for respondents to predict their happiness after option A and after option B than it is for them to make a direct choice. On the expectations side, given the unfamiliarity of precise probabilities compared to the familiarity of happiness, it may be easier and more reliable for respondents to report happiness than for them to directly report probabilities. For example, after setting the stage by asking how happy a sample of people on one side of the political divide would feel (A) if their preferred presidential candidate won or (B) if their less preferred presidential candidate won, their average happiness in the days before the election might be an efficient way to assess their subjective probabilities of victory for their preferred candidate.

We have an example from personal experience of using elation to gather information about the strength of preferences. One of us was present when a daughter opened letters from the admissions departments of the colleges she had applied to. The evident strength of the daughter's positive affective reaction was persuasive in establishing the extra value she placed on going to her much more expensive first-choice college, as opposed to her much less expensive second-choice college. Of course, this did not indicate what the ultimate wisdom of each choice would be, but it did indicate her preferences given her beliefs about what it would be like to go to each college.

E. Sufficient Statistics. As we have argued at length, elation and dismay measure the effect of news on expected lifetime utility. Somewhat more formally, elation provides a sufficient statistic that captures the effect on the (expected) lifetime utility of current or future events that had not previously been anticipated. This interpretation of elation is similar to the more familiar idea that in analyzing lifecycle maximization problems the marginal utility of income, a scalar quantity denoted by λ_t, "serves as the sufficient statistic which captures all information from other periods that is necessary to solve the current-period maximization problem." (Blundell and MaCurdy, 1999, p.1594). In particular, λ_t measures the utility value of a dollar saved for expenditure in the future relative to a dollar spent on consumption today and also the marginal utility value of additional time spent on leisure or household production per unit of foregone

earnings. Unanticipated changes in the marginal utility of income provide a signal about changes in the optimal allocation of consumption and leisure over the lifecycle.

It is interesting to consider the roles of these two quantities for consumer behavior. The role of the marginal utility of income is well known. Any news about changes in expected future income or price which causes the marginal utility of income to fall serve as a signal to a utility-maximizing consumer to increase current consumption and leisure by reducing saving; news that leads to an increase in the marginal utility of income is a signal to reduce consumption and leisure and increase saving. Both elation and λ_t are derived from the (expected) lifetime utility function. It follows that news that affects probability beliefs about future incomes, prices, health or any other variables that affect preferences or constraints that causes changes in λ_t will also tend to cause elation or dismay.

Mathematically, there are two differences between elation in our theory and λ_t as sufficient statistics. First, elation is about the *total* lifetime utility, while λ_t is about the *derivative* of lifetime utility with respect to wealth. Second, it is *temporary spikes* in the level of elation that indicate a change lifetime utility, while it is *permanent changes* in the level of λ_t that indicate a change in the first derivative of lifetime utility.[57]

In principle, economists could gain insight into the effects of news by studying the dynamics of consumption and labor supply jointly with longitudinal data on subjective well-being. Observed changes in savings, consumption or labor supply respond to news that influences the marginal value of a dollar while spikes in subjective well-being reflect changes in total lifetime utility. It is easy to show that elation and changes in the marginal utility of income are not necessarily correlated in a simple way. For example, a person who receives news that he has been promoted and will be receiving a higher salary next year will feel elation—a signal that his lifetime utility has gone up—and his marginal utility of income will fall—a signal that he should allocate more of his wealth to current consumption. Today, he might choose to celebrate his promotion with dinner at a fancy restaurant. Consider a less fortunate person who has just heard a jury convict him of a long prison sentence. Obviously, this person suffers dismay. However, given the difference in the availability of fine food in prison and in town, his marginal utility of income also falls and he may also choose a fancy restaurant meal today, assuming that his appeal allows him to stay out of prison for a while. As another example, a person (with full medical insurance) who has had a successful operation on a malignant tumor will experience elation and an increase in his marginal utility of income, signaling that he should increase his saving to accommodate his longer life expectancy.

Although elation and changes in the marginal utility of income could be correlated in either direction, it is likely that most news in the economic domain reflects good or bad news about

[57] A similarity between elation and λ_t is that econometric identification of both spikes in elation and changes in λ_t requires subtracting out an individual fixed effect. It may also be necessary to control for a few other factors that have predictable effects on changes in subjective well-being or behavior. For example, the real interest rate can have predictable effects on the evolution of consumption and labor supply, even in the absence of news, while time-varying determinants of baseline mood can have predictable effects on overall subjective well-being.

future income and wealth. That is, macroeconomic news about expansions or recessions or microeconomic news about one's own promotion or layoff tend to produce both unexpected gains or losses in lifetime utility and signals to increase or decrease current consumption. This might be consistent with psychological evidence that has been interpreted as suggesting that an up mood is a signal to move forward with bold plans, putting aside worries, while a down mood is a signal to focus on things that might go wrong and to proceed with caution. It would be interesting to examine whether the correlation between mood and consumption is usually positive and also to test whether this correlation is reversed in the less usual situations such as those described above when mood and the marginal utility of income move together rather than in opposite directions.

10. Implications of the Theory for Policy

A good way to discuss the policy implications of our framework is to contrast the views that we believe follow from our framework with those of Layard (2005). Layard (2005) is very bold in making policy recommendations based on happiness theory and empirics. Although he is especially bold, we consider the general tenor of his recommendations to be reasonably representative of views expressed in much of the existing happiness literature.

Layard explicitly accepts Kahneman's (1999) proposal to use the expected present discounted value of measured happiness averaged across people as the social welfare function. Besides the issues we discuss in this paper, Layard is assuming a solution to interpersonal comparability issues that we think have not been solved, but let us leave that aside, since all of our social welfare measures share that difficulty in all but the easiest applications.[58]

Many of Layard's recommendations depend only on happiness being more valuable than current public policy recognizes. The general discounting of intangibles in policy discussions makes this likely. Generating and popularizing happiness accounts in parallel to GDP accounts is a reasonable step to rectify insufficient attention to these intangibles. Taking happiness more seriously also suggests many other concrete steps, such as fighting the stigma to antidepressants and psychological treatment, and devoting more resources to mental health care, mental health research, happiness research, and public education about the determinants of long-run happiness.

Other recommendations depend on the externalities inherent in people caring about social rank. Since both revealed preference and happiness data indicate that social rank is important, these

[58] The Ordinalist Revolution made it clear that the key philosophical issues in judging social welfare for purposes of public policy could not be avoided even if a perfect direct measure of individual welfare existed. Most notably, there is no easy escape from the difficulties surrounding interpersonal comparison. For example, should those with more refined tastes who can distinguish more minute differences in quality therefore be accorded greater weight in social choice? See Stigler (1950).

recommendations remain on the mark. Quantitatively, a revealed preference measure of the importance of social rank may be different from an affective measure, but qualitatively, the implications of social rank mattering are the same either way.

Affective data also provide a good reminder of the importance of many other externalities and public goods—an importance that can be verified by revealed preference. The sense of community matters, the strength and quality of marriages and families matter: and the responsiveness of government matters. Again, a revealed preference measure may differ quantitatively in the importance it suggests for these externalities, but it is likely to agree qualitatively.

There are two areas where we differ with Layard. First, Layard makes many recommendations based on Kahneman's (1999) social welfare measure, where we would turn to utility based on revealed preference as the appropriate welfare measure. This leads to stark differences in recommendations about tradeoffs between economic growth and other values. For example, Layard argues that since economic growth does very little to raise happiness, while being forced to move from one city to another lowers happiness significantly, it is worth sacrificing a great deal of economic growth in order to slightly reduce the need for mobility. To us, this either tacitly assumes that feeling happy is the only thing people care about (which we dispute), or it requires forcing upon them the objective of maximizing happiness when, given the choice, they reject this objective for themselves. There are many indications that economic growth is, in fact, important to people, even if it does not raise happiness. In principle, the dollar value of happiness could be high enough to make such a sacrifice worthwhile even if happiness is not the only thing in the utility function, but it would require an extremely high value. Even given existing lay knowledge about the determinants of baseline mood, if the value people place on happiness were high enough to make this kind of sacrifice worthwhile, we should see many more people seeking psychological treatment, sleeping more, exercising more, eating better, pursuing meditative practices, and so on, than we observe. Of course, if stability of residence enters the utility function beyond its effect on happiness, its valuation could be higher, but this is not Layard's argument.[59]

It is worth being very explicit about why Kahneman's (1999) social welfare measure differs so much in its implications from standard social welfare measures in Economics. Our theory implies that the present discounted value of affect Kahneman points to is the sum of two very different components: the present discounted value of baseline mood and the present discounted value of elation. Conceptually, we view the present discounted value of baseline mood as something like the present discounted value of any other sub-utility function, such as the present discounted value of a health measure. While likely to be correlated with lifetime utility, this present discounted value of baseline mood represents only one of the things people care about.

[59] Note that for the set of things that only enter utility through happiness, valuing them at their effect on happiness times the revealed preference dollar value of happiness is a very different procedure from the common valuation procedure of dividing the effect of a variable on happiness by the effect of *income* on happiness. This procedure is wrong because it assumes that income only affects utility by affecting happiness—something we know to be false for income to the extent that people thoughtfully sacrifice happiness for higher income. In order to use income as a *numeraire*, all the benefits of income on utility need to be accounted for, not just the (possibly small) fraction of the benefits of income that show up in a higher level of happiness.

(Less importantly, this present discounted value also has in it no representation of any decreasing returns to baseline mood in the utility function.) By contrast, the present discounted value of elation is a very interesting quantity that (at least approximately) represents the cumulative innovation in lifetime utility over the interval of time covered in the discounted sum. In other words, the present discounted value of elation answers the question of how well one's life has turned out compared to what one expected at an earlier time, perhaps long ago, when the discounted sum began. If for example, one could separate out elation from baseline mood in measured affect and constructed present discounted sums of elation from a long panel of the adult population over 18, the average discounted sum of elation would represent how much better or worse people's lives turned out than they expected at age 18. As a social welfare measure, this intriguing quantity has one serious problem: it does not credit as social improvement any improvement in how people's lives in a society look as of age 18. Even if the panel were extended back to five-year olds, the present discounted sum of elation from that age on would not give due weight to improvement in life prospects as life prospects appear as of age 5. In our view, many of the most valuable aspects of progress over the past few centuries, or even the past few decades, are ones that would be highly valued by five-year olds, not just in the moment, but as they think about what their lives will be like when they grow up.

Second, we consider Layard too quick to believe that people are making systematic mistakes in optimization.[60] People no doubt do make mistakes, but because happiness is not the only thing people care about, happiness data alone is seldom enough to identify optimization mistakes. The key types of evidence we would point to for identifying mistakes are regret and people changing their minds on a decision after thinking more carefully or getting better information. Also, not all factual mistakes lead to optimization mistakes. In Section 8 C, we argue at length that mistakes in predicting the dynamics of affect do not necessarily lead to optimization mistakes.

11. Conclusion

Happiness research matters because—even if economic progress continues unabated over the next 50 years in the U.S. and other advanced countries—whether the citizens of these countries end up rich and happy or rich and unhappy depends on whether money *can* buy happiness *and* on whether the additional economic resources will, in fact, be used to obtain additional happiness. To the extent there is a tradeoff between happiness and other values, the increases in income and wealth that accompany economic progress are likely to make improvements in subjective well-being increasingly important for welfare compared to further improvements in other areas.[61]

In order for happiness research to fully tap into the vast accumulated human capital of the Economics profession, we consider it important to develop a theory that respects the canons of

[60] In this, Layard follows Gruber and Mullainathan (2002) and Frey and Stutzer (2004b).

[61] One area where trends in happiness could have important macroeconomic effects is in the area of happiness on the job. For example, it is possible that, in the coming decades, advances in subjective well-being at work could alter people's relationship to work in a way that significantly raises the average retirement age. Happiness on the job is likely to be an increasingly important element of competitive advantage—particularly for firms that need to attract skilled workers who may place a higher dollar value on happiness.

Economics as well as the findings of Psychology. One of the most important canons of Economics is Ordinalism, or the principle of Revealed Preference.

Revealed Preference, applied to the Easterlin Paradox of nontrending happiness in the face of dramatic improvements in per capita income and many other areas of life, clearly distinguishes utility from happiness. Utility is the extent to which people achieve what they care about, as indicated by their choices; happiness is how they feel. This distinction is important. In particular, the distinction between utility and happiness leads to many insights and productive questions that would be difficult to see as long as utility and happiness are confused with one another.

Appendix: Axiomatics

This paper is not the place to deal thoroughly with axiomatic issues about the link between happiness and utility, but it is useful to briefly address some issues about the preference for baseline mood and the link between innovations in lifetime utility and elation.

Preference for Baseline Mood: One element of our discussion above is the postulate that people value long-term happiness positively. This postulate cannot be made meaningful without auxiliary hypotheses distinguishing things that matter for utility only through their effect on happiness—which is a statement outside of strict revealed-preference analysis over the goods in the utility function more elementary than happiness. To illustrate this logical issue, suppose the flow utility function can be described by the function $U(K_t, X_t, M_t)$, where baseline mood M_t is given by the function $M_t = \Phi(K_t, X_t)$ and the partial derivative $U_M > 0$. This flow utility function is obviously equivalent to the alternative flow utility function Ω defined by $\Omega(K_t, X_t, M_t) \equiv U(K_t, X_t, 2\Phi(K_t, X_t)-M_t)$, but the partial derivative $\Omega_M < 0$.

This is not a new issue. It arises for any Becker-esque treatment of goods produced by a household production function. For example, it would not be unreasonable to say a priori (at least as an approximation) that, on the benefit side, a washing machine is only valued for its laundering services, which in turn are only valued for their contribution to clothing services. This stipulation would then be important to an analysis of the demand for washing machines.

In an applied context, we think auxiliary hypotheses of the type needed to study the place of baseline mood in the utility function can, in fact, be reasonable. For example, it might be reasonable to assume that other than its time and money costs, talk therapy enters the utility function only through its effect on baseline mood. This auxiliary assumption, together with evidence on the size and duration of the effect of talk therapy on baseline mood would allow what is otherwise a revealed preference evaluation to be placed on happiness. To take a more complex example, one might assume that besides its time and money cost, an antidepressant medication enters the utility function only through baseline mood and its medical side-effects. Preferences over another medication that generates similar side effects but has few benefits for the individual might make it possible to evaluate the cost of the side effects. Then, other than time and money costs, the vector good of switching from the other ineffective medication to the effective antidepressant would enter the utility function only through its effect on baseline mood.

Happiness and News about Lifetime Utility. In connection with elation, the key axiomatic issue is whether it is possible to express our main claims about happiness and news in a way that is independent of any particular representation of the lifetime utility function. Let the stochastic process of the control variable vector X and the state variable vector k be the fundamentals that lifetime utility depends on. Define K as a vector giving the history of k through time t and the history of X through time t-*1*. That is,

$$K_t = (k_t, k_{t-1}, \ldots k_0; X_{t-1}, X_{t-2}, \ldots X_0).$$

Call lifetime utility v_t, as before. Note that v_t depends on the information at time t about the future, so news is reflected in changes in v. We propose the following three axioms relevant to happiness and news:

1. *Happiness A at time t is a function of the realized history of X, k and v up through time t. That is,* $A_t = h_t(K_t, X_t, v_t, v_{t-1}, \ldots, v_0)$.

2. *Holding fixed the history of realized X and k through time t, and holding fixed the past history of realized v through time t-1, happiness at time t is increasing in current lifetime utility* v_t. *That is, if* $v_t' > v_t$, *then* $h_t(K_t, X_t, v_t', v_{t-1}, ..., v_0) > h_t(K_t, X_t, v_t, v_{t-1}, ..., v_0)$.
3. *Holding fixed current lifetime utility* v_t *and the realized history of X and k through time t, happiness at time t is decreasing in previous realized values of lifetime utility. That is, for any integer* $j>0$, *if* $v_{t-j}' > v_{t-j}$, *then*
$$h_t(K_t, X_t, v_t, v_{t-1}, ..., v_{t-j}', ..., v_0) > h_t(K_t, X_t, v_t, v_{t-1}, ..., v_{t-j}, ..., v_0).$$

Remarks: These axioms are all ordinal. The would not be changed in meaning by monotonically increasing transformations of *v* and *h*.

These axioms can be applied readily to any lifetime utility function that can be expressed by the terminal condition $v_{T+1} \equiv 0$ (perhaps with $T \to \infty$ in the end) and the recursive relationship $v_t = \Psi(k_t, X_t, F_{v_{t+1}})$, where $F_{v_{t+1}}$ is the probability distribution function for v_{t+1}. (Although there would be excess baggage in using the expanded state vector K_t, this recursive equation can also be written $v_t = \Psi(K_t, X_t, F_{v_{t+1}})$.) Note that expected utility maximization is not required for the axioms to be meaningful.

Assuming the simultaneous equations ultimately yield a well-defined and well-behaved lifetime utility function, these axioms can also be applied to lifetime utility functions like those in section 8 for which, in light of the expression for happiness given by News and Happiness Axiom 1, one can represent the inclusion of overall happiness in the utility function by an intertemporal equation of the form
$$v_t = \Psi(K_t, X_t, F_{v_{t+1}}, v_{t-1}, v_{t-2}, ..., v_0),$$
together with the terminal condition $v_{T+1} \equiv 0$.

In conjunction with the additively time-separable intertemporal expected utility function of Section 6, for which using the reduced form flow utility function $\mathcal{U}$, and *k* for the state variable vector that directly matters for flow utility, the equation $\iota_t = v_t - \beta^{-1} v_{t-1} + \beta^{-1} \mathcal{U}(k_{t-1}, X_{t-1})$ means that we can define a function H_t so that
$$\begin{aligned} &h_t(K_t, X_t, v_t, v_{t-1}, ..., v_{t-j}, ..., v_0) \\ &= H_t(K_t, X_t, v_t - \beta^{-1} v_{t-1} + \beta^{-1} \mathcal{U}(k_{t-1}, X_{t-1}), ..., v_1 - \beta^{-1} v_0 + \beta^{-1} \mathcal{U}(k_0, X_0), v_0) \\ &= H_t(K_t, X_t, \iota_t, \iota_{t-1}, ..., \iota_1, v_0). \end{aligned}$$
In words, the entire history of *k* and *X* included in *K* allows one to calculate the history of flow utility, which allows one to back out the history of lifetime utility from the history of lifetime utility innovations and v_0. This equation makes it easy to apply Happiness and News Axioms 2 and 3. Axiom 2 implies that $\frac{\partial H_t}{\partial \iota_t} > 0$. In addition to $\frac{\partial H_t}{\partial v_0} < 0$, Axiom 3 implies that $\beta^{-1} \frac{\partial H_t}{\partial \iota_{t-j}} > \frac{\partial H_t}{\partial \iota_{t-j-1}}$ for any integer j from 0 to *t-1*. This says that recent news about future events will have a bigger effect on happiness than older news about future events. (The factor β^{-1}

merely puts the comparison between the effects of lifetime utility innovations at different lags on a present-value rather than a current-value basis.) This inequality allows the possibility that distant enough lags of lifetime utility innovations could have a negative effect on happiness. Though we do not think this possibility is empirically relevant, we also do not think it should be ruled out *a priori*.

Finally, we argue that, other than for the application of Happiness and News Axiom 3, it is reasonable to include the initial value of lifetime utility in the comprehensive history state variable vector *K*. We interpret v_0 as the view of lifetime utility in the instant before birth begins, when the individual has no information about her or his life prospects other than the information that is embodied in genes and body structure at that point. Because the individual's information set is biologically limited up until birth, it is appropriate to view v_0 as an element of the state variable vector that is not subject to subtle expectational effects. After this inclusion, we can write happiness as

$$A_t = H_t(K_t, X_t, \iota_t, \iota_{t-1}, ..., \iota_1).$$

Happiness A_t depends on the current (expanded) state variable vector K_t, the current control variable vector X_t, and the history of lifetime utility innovations. This is our essential claim about the nature of happiness given an additively time-separable intertemporal expected utility function. The additivity in the main text equation $A_t = M(K_t, X_t) + e(\iota_t, \iota_{t-1}, \iota_{t-2}, ...)$ between the function $M(K_t, X_t)$ of *K* and *X* and the function $e(\iota_t, \iota_{t-1}, \iota_{t-2}, ...)$ of lifetime utility innovations is only a mathematical and expositional convenience.

References

Argyle, Michael, 1999. "Causes and Correlates of Happiness," chapter 18 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Aristotle, fourth century B.C.E. *Nicomachean Ethics*, available in many modern editions.

Becker, Gary, 1965. "A Theory of the Allocation of Time," *Economic Journal* 75, pp. 492-517.

J. Bentham (1789) Introduction to the Principles of Morals and Legislation. 1948 edition, New York: Hafner.

Biddle, S., and N. Murtrie, 1991. *Psychology of Physical Activity and Exercise,* London: Springer/Tavistock/Routledge.

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

Brickman, P., and D. Campbell, 1971. "Hedonic Relativism and Planning the Good Society," in M. H. Appley ed., *Adaptation Level Theory: A Symposium,* New York, Academic Press, pp. 287-302.

Brickman, P., D. Coates and R. Janoff-Bulman, 1978. "Lottery Winners and Accident Victims: Is Happiness Relative?" *Journal of Personality and Social Psychology*, 36 (8), pp. 917-927.

Camerer, Colin, George Loewenstein and Drazen Prelec, "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature* 43 (March 2005), pp. 9—64.

Charles, Kerwin K., 2002. "Is Retirement Depressing?: Labor Force Inactivity and Psychological Well-Being in Later Life" NBER working paper #9033.

Clark, A., 1999. "Are wages habit-forming? Evidence from micro data," *Journal of Economic Behaviour and Organisation, 39*, 179-200.

Conlisk, John, 1996. "Why Bounded Rationality," *Journal of Economic Literature* XXIV:2 (June), 669-700.

Constantinides, George, 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy* 98 (June), 519-43.

Davis, Melinda, 2002. *The New Culture of Desire: The Pleasure Imperative Transforming Your Business and Your Life***,** Free Press.

Di Tella, Rafael, Alberto Alesina, and Robert MacCulloch. "Happiness and Inequality: Are Europeans and Americans Different?" *Journal of Public Economics* 88, nos. 9-10 (August 2004): 2009-42.

Di Tella, Rafael, and Robert MacCulloch, 1999. "Partisan Social Happiness," *Review of Economic Studies* (forthcoming).

Di Tella, Rafael, Robert MacCulloch, and Andrew J. Oswald, 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics* 85 (4), 793-809.

Di Tella, Rafael, Robert MacCulloch, and Andrew Oswald, 2001. "Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness," *American Economic Review* 91 (March).

Diener, E., and Lucas, R. E., 1999. Personality and subjective well-being. In D. Kahneman, E. Diener and N. Schwarz (Eds.), *Well-Being: The Foundations of Hedonic Psychology* (pp. 213-229). New York: Russell Sage Foundation.

Diener, Ed, and Christie Scollon, 2003. "Subjective Well-Being is Desirable But Not the Summum Bonum," mimeo, University of Illinois at Urbana-Champaign.

Easterbrook, Gregg, 2003. *The Progress Paradox: How Life Gets Better While People Feel Worse*, Random House, New York.

Easterlin, Richard A., 1974 . "Does Economic Growth Improve the Human Lot?: Some Empirical Evidence," in P. A. David and W. R. Levin eds., *Nations and Household in Economic Growth*, Stanford University Press, pp. 98-125.

Easterlin, Richard A., 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* 27, pp. 35-47.

Easterlin, Richard A., 2003. "Explaining Happiness," *Proceedings of the National Academy of Science*, 100 (19), pp. 11176-11183.

Frank, R. H., 1985. *Choosing the right pond*. New York: Oxford University Press

Frank, R., 1999. *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess*, New York: N.Y. Free Press.

Frederick, Shane, and George Loewenstein, 1999. "Hedonic Adaptation," chapter 16 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Frey, Bruno S., Simon Luechinger and Alois Stutzer, 2004. "Valuing Public Goods: The Life Satisfaction Approach," Center for Economic Studies and Ifo Institute for Economic Research Working Paper #1158, University of Zurich.

Frey, Bruno S. and Stutzer, Alois, 2003, "Testing Theories of Happiness" In *Paradoxes of Happiness in Economics,* Conference held at the University of Milano-Bicocca, March 21-23, 2003.

Frey, Bruno S., and Alois Stutzer, 2000. "Happiness, Economy and Institutions," *The Economic Journal*, 110 (October), 918-938.

Frey, Bruno S., and Alois Stutzer, 2002. "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature*, 40 (June), pp. 402-435.

Frey, Bruno S., and Alois Stutzer, 2004a. "Happiness Research: State and Prospects," Institute for Empirical Research in Economics Working Paper #190, University of Zurich.

Frey, Bruno S., and Alois Stutzer, 2004b. "Economic Consequences of Mispredicting Utility," Institute for Empirical Research in Economics Working Paper #218, University of Zurich.

Gardner, Jonathan, and Andrew Oswald, 2001. "Does Money Buy Happiness? A Longitudinal Study Using Data on Windfalls," mimeo, available at http://www2.warwick.ac.uk/fac/soc/economics/staff/faculty/oswald/

Geneakoplos, John, David Pearce and Ennio Stacchetti, 1989. "Psychological Games and Sequential Rationality," G*ames and Economic Behavior* 1, 60-79.

Gottfried, Jay A., John O'Doherty and Raymond J. Dolan, 2003. "Encoding Predictive Reward Value in Human Amygdala and Orbitofrontal Cortex," *Science* 301 (22 August), 1104-1107.

Gruber, Jonathan, and Sendhil Mullainathan, 2002. "Do Cigarette Taxes Make Smokers Happier?" NBER Working Paper #8872.

Hamermesh, Daniel S., forthcoming. "Subjective Outcomes in Economics," *Southern Economic Journal*.

Hall, Robert E., and Charles I. Jones, 2004. "The Value of Life and the Rise in Health Spending," National Bureau of Economic Research Working Paper #10737.

Harsanyi, John C., 1953. "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, 61 (October), pp. 434-435.

Harsanyi, John C., 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, pp. 309-321.

Heckman, James J. and MaCurdy, Thomas E. (1980) "A Life Cycle Model of Female Labor Supply." *Review of Economic Studies* 47: 47-74.

Helliwell, John F., 2002. "How's Life? Combining Individual and National Variables to Explain Subjective Well-Being," NBER Working Paper #9065.

Hicks, J. R. and Allen, R. G. D., "A Reconsideration of the Theory of Value", Economica, February/May, 1934, pp.52-60 and 196-219.

Hirsch, F. (1976). *Social limits to growth*. Cambridge, MA: Harvard University Press.

Hirshleifer, David, and Tyler Shumway, 2003. "Good Day Sunshine: Stock Returns and the Weather," *Journal of Finance*, 1009-1032.

Kahneman, Daniel, Ed Diener and Norbert Schwarz, 1999. *Well-Being: The Foundations of Hedonic Psychology*, edited volume, Russell Sage Foundation, New York.

Kahneman, Daniel, 1999. "Objective Happiness," chapter 1 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz, and Arthur A. Stone, 2004. "Toward National Well-Being Accounts," *American Economic Review*, 94, pp. 429-434.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz and Arthur A. Stone, 2004. "A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method (DRM)," *Science* (December), 1776-1780.

Kahneman, Daniel, and Amos Tversky, 1979. "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263-291.

Kaprio, J., M. Koskenvuo, and H. Rita, 1987. "Mortality after Bereavement: A Prospective Study of 95,647 Widowed Persons," *American Journal of Public Health* 77, pp. 283--287.

Kapteyn, A., and Wansbeek, T., 1985. "The individual welfare function: A review," *Journal of Economic Psychology, 6,* 333-363.

Kezdi, Gabor, and Robert J. Willis, 2003. "Who Becomes a Stockholder? Expectations, Subjective Uncertainty, and Asset Allocation," mimeo, University of Michigan.

Kimball, Miles, Helen Levy, Fumio Ohtake and Yoshiro Tsutsui, 2006. "Unhappiness after Hurricane Katrina," NBER Working Paper #12062 (March).

Kreps, David, and Evan Porteus, 1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185-200.

Laibson, David, 1997. "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 62 (2), pp. 443-478.

Lane, Robert E., 2000. *The Loss of Happiness in Market Democracies*, Yale University Press, New Haven, pp. 141-158.

Larsen, Randy J., and Barbara L. Fredrickson, 1999. "Measurement Issutes in Emotion Research," chapter 3 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Layard, Richard, 2005. *Happiness*. London, UK: Penguin Press. This book is based to a large extent on Layard's 2003 Lionell Robbins Memorial Lectures, London School of Economics.

Lewinsohn, P. M., J. M. Sullivan and S. J. Grosscup (1982). "Behavioral Therapy: Clinical Applications," in A. J. Rush (ed.), *Short-Term Therapies for Depression* (pp. 50-87), New York: Guilford.

Lodge, David. (1978) *Changing Places*. Penguin Books.

Loewenstein, George, Ted O'Donoghue and Matthew Rabin (2002), "Projection Bias in Predicting Future Utility," mimeo. Carnegie-Mellon University.

Loewenstein, George, and David Schkade, 1999. "Wouldn't It Be Nice? Predicting Future Feelings," chapter 5 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Lucas, Richard E., Andrew Clark, Yannis Georgellis, and Ed Diener, 2004. "Unemployment Alters the Set Point for Life Satisfaction," *Psychological Science* 15 (1), pp. 8--13.

Lupton, Joseph, 2002. "The Habit Liability in Life-Cycle Consumption and Portfolio Choice," mimeo, Federal Reserve Board.

Luttmer, E.F.P., 2004. "Neighbors as negatives: Relative earnings and well-being," NBER Working Paper, No. w10667.

Lykken, D., and Tellegen, A., 1996. Happiness is a stochastic phenomenon. *Psychological Science, 7*, 186-189.

Marmot, Michael, 2004. *How Our Position on the Social Gradient Affects Longevity and Health*, London: Bloomsbury.

MaCurdy, Thomas E. (1980) "An Empirical Model of Labor Supply in a Life-Cycle Setting." *Journal of Political Economy* 89: 1059-1085.

MaCurdy, Thomas (1985) "Interpreting Empirical Models of Labor Supply in An Intertemporal Framework with Uncertainty." In J.J. Heckman and B. Singer, eds. Longitudinal Analysis of Labor Market Data (Cambridge University Press, Cambridge UK).

MaCurdy, Thomas E. and Blundell, Richard (1999) "Labor Supply: A Review of Alternative Approaches." In O. Ashenfelter and D. Card., eds. Handbook of Labor Economics, Vol. 3A. Elsevier: 1559-1695.

Maslow, Abraham, 1943. "A Theory of Human Motivation" *Psychological Review,* 50, pp. 370-396.

Morris, Desmond, 2004. *The Nature of Happiness*, Little Books Ltd., London.

Morris, W. N., 1999. The mood system. In D. Kahneman, E. Diener and N. Schwarz (Eds.), *Well-Being* (pp. 169-189). New York: Russell Sage Foundation.

Nesse, Randolph M., 2000. Is depression an adaptation? *Archives of General Psychiatry, 57*, 14-20.

Nesse, Randolph M., 2001. Motivation and melancholy: a darwinian perspective. *Nebr Symp Motiv, 47*, 179-203.

Nesse, Randolph M., 2004. Natural selection and the elusiveness of happiness. *Philos Trans R Soc Lond B Biol Sci, 359*, 1333-1347.

Nesse, Randolph M., forthcoming. "Natural Selection and the Regulation of Defenses," *Evolution and Human Behavior.*

Nickerson, Carol, Norbert Schwarz, Ed Diener, and Daniel Kahneman, 2003. "Zeroing in on the Dark Side of the American Dream: A Closer Look at the Negative Consequences of the Goal for Financial Success," *Psychological Science*, 14 (November), 531-536.

Nisan, M., 1972. "Dimension of time in relation to choice behavior and achievement orientation," *Journal of Personality and Social Psychology,* 21, 175-182.

O'Doherty, John P., 2005. "Reward Representations and Reward-Related Learning in the Human Brain: Insights from Neuroimaging," Wellcome Department of Imaging Neuroscience.

O'Doherty, John P., Peter Dayan, Karl Friston, Hugo Critchley and Raymond Dolan, 2003. "Temporal Difference Models and Reward-Related Learning in the Human Brain," *Neuron* 28 (April 23), 329-337.

O'Guinn, T.C. and Shrum, L. J., 1997. "The role of television in the construction of consumer reality". *Journal of Consumer Research,* 23*,* 278-294

Pareto, Vilfredo, 1906. Manuel of Political Economy trans. by Ann S Schwier, edited by Ann S. Schwier and Alfred N. Page (New York: Augustus M. Kelley, 1971).

Pattanaik, P. K., 1968. "Risk, Impersonality and the Social Welfare Function," *Journal of Political Economy*, 76 (December), 1152-1169.

Patterson, D. R., J. J. Everett, C. H. Bombardier, K. A. Questad, V. K. Lee, and J. A. Marvin, 1993. "Psychological Effects of Severe Burn Injuries," *Psychological Bulletin* 113, pp. 362--378.

Pinker, Stephen, 1997. *How the Mind Works,* Norton.

Putnam, Robert D., 2000. *The Collapse and Revival of American Community,* Simon & Schuster.

Rawls, John, 1951. "Outline of a Decision Procedure for Ethics," *Philisophical Review*, 60, pp. 177-197.

Rawls, John, 1958. "Justice as Fairness," *Philosophical Review*, 67, pp. 164-194.

Rawls, John, 1971. *A Theory of Justice*, revised edition 1999, Belknap Press of Harvard University Press, Cambridge, MA.

Rayo, Luis and Gary S. Becker, 2005. "Evolutionary Efficiency and Happiness," mimeo, University of Chicago.

Reich, J. W. and J. Zautra, 1981. "Life Events and Personal Causation: Some Relationships with Satisfaction and Distress," *Journal of Personality and Social Psychology*, 41, 1002-1012.

Robbins, Lionel C., 1932. *An Essay on the Nature and Significance of Economic Science*. 1984 edition, New York: New York University Press.

Robinson, Michael D. and Gerald L. Clorem 2002. "Belief and Feeling: Evidence for an Accessibility Model of Emotional Self Report," *Psychological Bulletin*, 128, 934-960.

Rosen, Sherwin (1986) "The Theory of Equalizing Differences," in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*

Samuelson, Paul A., 1938. "A Note on the Pure Theory of Consumer's Behaviour," *Economica* February), 61—71.

Samuelson, Paul A., 1947. *Foundations of Economic Analysis.* Cambridge: Harvard University Press, (enlarged edition, 1983), esp. chapters V, VII, and VIII.

Schkade, D. A. and Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science*, 9 (5), 340-529.

Schor, Juliet B. (1991) *The Overworked American: The Unexpected Decline of Leisure*. New York. Basic Books.

Schultz, Wolfram, and Anthony Dickinson, 2000. "Neuronal Coding of Prediction Errors," *Annual Review of Neuroscience*, 23: 473—500.

Schwartz, Barry, 2004. *The Paradox of Choice: Why More is Less*, New York: HarperCollins.

Schwarz, N. (1987). *Stimmung als Information: Untersuchungen zum Einfluß von Stimmungen auf die Bewertung des eigenen Lebens*. [Mood as information.] Heidelberg, Germany: Springer Verlag.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins and R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. 2; pp. 527-561). New York, NY: Guilford Press.

Schwarz, N., 1996. *Cognition and communication: Judgmental biases, research methods and the logic of conversation.* Hillsdale, NJ: Erlbaum.

Schwarz, N., 1999. "Self-reports: How the questions shape the answers," *American Psychologist, 54*, 93-105.

Schwarz, N. (2002). Situated cognition and the wisdom of feelings: Cognitive tuning. In L. Feldman Barrett and P. Salovey (Eds.), *The wisdom in feelings* (pp.144-166). New York: Guilford.

Schwarz, N., and Bohner, G., 2001. "The construction of attitudes," In A. Tesser and N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (Vol.1, pp. 436-457). Oxford, UK: Blackwell.

Schwarz, N., and Clore, G.L., 1983. Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513 - 523.

Schwarz, Norbert and Fritz Strack, 1999. "Reports of Subjective Well-Being: Judgmental Processes and Their Methodological Implications," chapter 4 in Daniel Kahneman, Ed Diener and Norbert Schwarz eds., *Well-Being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Seligman, Martin E. P., 2002. *Authentic Happiness: Using the New Positive Psychology to Realize Your Potential for Lasting Fulfillment*, Free Press, New York.

Sen, Amartya, 1983. "Poor, relatively speaking," *Oxford Economic Papers, 35,* 153-169.

Seymour, Ben, John P. O'Doherty, Martin Koltzenburg, Katja Wiech, Richard Rackowiak, Karl Friston and Raymond Dolan, forthcoming. "Opponent appetitive-aversive neural processes underlie predictive learning about pain relief," *Nature Neuroscience*.

Seymour, Ben, John P. O'Doherty, Peter Dayan, Martin Koltzenburg, Anthony K. Jones, Raymond J. Dolan, Karl J. Friston, 2004. "Temporal Difference Models Describe Higher-Order Learning in Humans," *Letters to Nature* 429 (June), 664-667.

Shepperd, J. A., Findley-Klein, C., Kwavnick, K. D., Walker, D., and Perez, S., 2000. "Bracing for loss," *Journal of Personality and Social Psychology*, 78, 620-634.

Shepperd, J. A., Ouellette, J. A., and Fernandez, J. K., 1996. "Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback," *Journal of Personality and Social Psychology*, 70, 844-855.

Sims, Christopher, 2002. "Implications of Rational Inattention," unpublished paper, Princeton University.

Solnick, Sara J., and David Hemenway, 1998. "Is More Always Better?: A Survey of Positional Concerns," *Journal of Economic Behavior and Organization*, 37, 373--383.

Steffick, Diane E., 2000. "Documentation of Affective Functioning Measures in the Health and Retirement Study," HRS Documentation Report DR-005, available online at http://hrsonline.isr.umich.edu/docs/userg/dr-005.pdf

Steptoe, A., J. Kimball and P. Basford, 1996. "Exercise and the Experience and Appraisal of Daily Stressors: A Naturalistic Study," *Journal of Behavioral Medicine,* 21, 363—374.

Stigler, George, 1950. "Development of Utility Theory," Journal of Political Economy 58, pts 1-2 (August-October), pp. 307-327, 373-396.

Strack, F., Martin, L.L., and Schwarz, N. (1988). Priming and communication: The social determinants of information use in judgments of life-satisfaction. *European Journal of Social Psychology, 18*, 429 - 442.

Strack, F., Schwarz, N., Chassein, B., Kern, D., and Wagner, D. (1990). The salience of comparison standards and the activation of social norms: Consequences for judgments of happiness and their communication. *British Journal of Social Psychology, 29*, 303-314.

Strack, F., Schwarz, N., and Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, mood, and mode of thinking. *Journal of Personality and Social Psychology, 49*, 1460 -1469.

Sulls, J., and Wills, T.A. (Eds) (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum.

Thayer, R. E., 1989. *The Biopsychology of Mood and Emotion*, New York: Oxford University Press.

Tsutsui, Yoshiro, Miles Kimball and Fumio Ohtake, 2005. "Koizumi Carries the Day: Happiness and Unhappiness after the Japanese Election," Osaka University and University of Michigan.

Wilson, Timothy D., 2002. *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Harvard University Press, Cambridge, MA and London, England, especially chapter 6, pp. 137-159.

Wolfers, Justin, 2003. "Is Business Cycle Volatility Costly? Evidence from Surveys of Subjective Well-Being," NBER Working Paper #9619.

Woodford, Michael, 2002. "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps,* Princeton Univ. Press.

Wortman, C. and R. Silver, 1987. "Coping with Irrevocable Loss," in *Cataclysms, Crises and Catastrophes: Psychology in Action*, Master Lecture Series, v. 6, Washington, D. C., American Psychological Association, pp. 189-235.

Zamble, E., 1992. "Behavior and Adaptation in Long-Term Prison Inmates: Descriptive Longitudinal Results," *Criminal Justice and Behavior* 19, pp. 409--425.

Zamble, E., and F. Proporino, 1990. "Coping, Imprisonment, and Rehabilitation: Some Data and Their Implications," *Criminal Justice and Behavior* 17, pp. 53-70.

NBER WORKING PAPER SERIES

DO PEOPLE SEEK TO MAXIMIZE HAPPINESS? EVIDENCE FROM NEW SURVEYS

Daniel J. Benjamin
Ori Heffetz
Miles S. Kimball
Alex Rees-Jones

Working Paper 16489
http://www.nber.org/papers/w16489

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2010

We are extremely grateful to Dr. Robert Rees-Jones and his office staff for generously allowing us to survey their patients and to Cornell's Survey Research Institute for allowing us to put questions in the 2009 Cornell National Social Survey. We thank Gregory Besharov, John Ham, Benjamin Ho, Erzo F. P. Luttmer, Michael McBride, Ted O'Donoghue, Matthew Rabin, Antonio Rangel, and Robert J. Willis for especially valuable early comments and suggestions. We are grateful to participants at the CSIP Workshop on Happiness and the Economy held at the Federal Reserve Bank of San Francisco, participants at the Cornell Behavioral/Experimental Lab Meetings, and seminar audiences at Cornell University, Deakin University, Syracuse University, NBER, the Stanford Institute for Theoretical Economics (SITE), and Florida State University for helpful comments. We thank Eric Bastine, J.R. Cho, Kristen Cooper, Isabel Fay, John Farragut, Geoffrey Fisher, Sean Garborg, Arjun Gokhale, Jesse Gould, Kailash Gupta, Han Jiang, Justin Kang, June Kim, Nathan McMahon, Elliot Mandell, Cameron McConkey, Greg Muenzen, Desmond Ong, Mihir Patel, John Schemitsch, Brian Scott, Abhishek Shah, James Sherman, Dennis Shiraev, Elizabeth Traux, Charles Whittaker, Brehnen Wong, Meng Xue, and Muxin Yu for their research assistance. We thank the National Institute on Aging (grant P01-AG026571/01) for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Do People Seek to Maximize Happiness? Evidence from New Surveys
Daniel J. Benjamin, Ori Heffetz, Miles S. Kimball, and Alex Rees-Jones
NBER Working Paper No. 16489
October 2010
JEL No. D03,D60

ABSTRACT

Are subjective well-being (SWB) measures a good empirical proxy for utility? We evaluate one necessary assumption: that people's preferences coincide with what they predict will maximize their SWB. Our method is to present survey respondents with hypothetical scenarios and elicit both choice and predicted SWB rankings of two alternatives. While choice and predicted SWB rankings usually coincide, we find systematic reversals. Furthermore, we identify factors—such as predicted sense of purpose, control over one's life, family happiness, and social status—that help explain choice controlling for predicted SWB. We explore how our findings vary with the SWB measure and the choice situation.

Daniel J. Benjamin
Economics Department
Cornell University
480 Uris Hall
Ithaca, NY 14853
and NBER
daniel.benjamin@gmail.com

Ori Heffetz
Cornell University
324 Sage Hall
Ithaca, NY 14853
oh33@cornell.edu

Miles S. Kimball
Department of Economics
University of Michigan
Ann Arbor, MI 48109-1220
and NBER
mkimball@umich.edu

Alex Rees-Jones
Economics Department
Cornell University
Ithaca, NY 14853
arr34@cornell.edu

All things considered, how satisfied are you with your life as a whole these days?

Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?

Much of the time during the past week, you felt you were happy. Would you say yes or no?[1]

Economists increasingly use survey-based measures of subjective well-being (SWB) as an empirical proxy for utility. In some applications, SWB data are used for conducting welfare evaluations that are difficult to conduct with standard revealed preference methods. Examples include measuring the negative externality from neighbors' higher earnings (Luttmer, 2005), individuals' tradeoff between inflation and unemployment (DiTella, MacCulloch, and Oswald, 2003), and the effect of health status on the marginal utility of consumption (Finkelstein, Luttmer, and Notowidigdo, 2008). In order for such SWB-based evaluations to assess welfare as conceived in the standard economic approach, a necessary assumption is that an agent's preferences over alternatives and her measured SWB under each alternative yield identical rankings of the alternatives. This paper provides evidence for evaluating that assumption. Put differently but equivalently, the goal of this paper is to assess whether individuals seek to maximize SWB (as commonly measured).

In pursuing that goal, this paper also empirically addresses two opposing theoretical views regarding the relationship between SWB and utility. The first, reflected at least implicitly in a large number of papers (e.g., Gruber and Mullainathan, 2005; Frey, Luechinger, and Stutzer, 2008; Welsch, 2009), is that SWB data represent idealized revealed-preference utility in the sense of what individuals would choose if they were well-informed about the consequences of their choices for SWB. The second view, explicitly laid out in Kimball and Willis (2006) and Becker and Rayo (2008), is that even well-informed agents will be willing to trade off SWB for other things they care about, making SWB and utility distinct.[2] This paper assesses these two

[1] The first of these three questions is from the World Values Survey (WVS); similar questions appear in the Euro-Barometer Survey, the European Social Survey, the German Socioeconomic Panel (G-SOEP), and the Japanese Life in Nation survey. The second question is from the U.S. General Social Survey (GSS); similar questions appear in the Euro-Barometer survey, the National Survey of Families and Households (NSFH), and the WVS. The third question is from the University of Michigan's Survey of Consumers; similar questions appear in the Center of Epidemiologic Studies Depression Scale (CES-D), the Health and Retirement Study (HRS), and the Gallup-Healthways Well-Being Index (GHWBI).

[2] For other elaborations of this view, see also Kimball, Levy, Ohtake and Tsutsui (2006), Tsutsui, Kimball, and Ohtake (2010), and Kimball, Nunn and Silverman (2010).

views empirically by measuring the concordance between choices and what individuals think will maximize their SWB.[3]

We pose a variety of hypothetical decision scenarios to three respondent populations: a convenience sample of 1,066 adults, a representative sample of 1,000 adult Americans, and 633 students. Each scenario has two alternatives. For example, one scenario describes a choice between a job that pays less but allows more sleep versus a job with higher pay and less sleep. We ask respondents which alternative they would choose. We also ask them under which alternative they expect greater SWB; we assess this "predicted SWB" using measures based on each of the three commonly-used SWB questions posed in the epigraph above. Importantly, since people often mispredict the happiness consequences of their choices (e.g., Gilbert, 2006), we compare an individual's choices with her *ex ante beliefs* regarding her choices' effects on SWB.[4]

We have two main results. First, we find that overall, respondents' SWB predictions are a powerful predictor of their choices. On average, predicted SWB and choice coincide in our data 83 percent of the time. We find that the strength of this relationship varies across choice situations, subject populations, survey methods, questionnaire structure variations, and measures of SWB, with coincidence ranging from well below 50 percent to above 95 percent.

Our second main result is that discrepancies between choice and predicted SWB rankings are systematic. Moreover, we can identify other, non-SWB factors that help explain respondents' choices. In some surveys, in addition to measuring participants' choices and predicted SWB, we also ask participants to predict how each alternative would affect particular aspects of life other than SWB. The aspects that systematically contribute most to explaining choice, *controlling for SWB*, are sense of purpose, control over life, family happiness, and social status. At the same time, and in line with our first main result above, when we examine how well SWB compares with the other factors we measure, we find that across our scenarios, populations, and methods,

[3] Some researchers adopt yet a third view, the philosophical position that happiness *per se* is the appropriate maximand for a policymaker (e.g., Bentham, 1789; Layard, 2005; Bronsteen, Buccafusco, and Masur, forthcoming). See Bernheim (2009) for a categorization and discussion of alternative philosophical approaches to welfare economics. Throughout this paper, we adopt the standard economic perspective (common to the two views mentioned in the text) that the appropriate maximand for welfare analysis is an individual's preferences. However, the value of our results does not hinge on this perspective. Even if one believes that happiness is the appropriate maximand, one would like to know under what circumstances people choose to maximize it on their own, and under what circumstances a policymaker would have to intervene in order to maximize it.

[4] In the terminology of Kahneman, Wakker, and Sarin (1997), we compare "decision utility" (what people choose) with "predicted utility" (what people predict will make them happier).

SWB is by far the single best predictor of choice.

These results provide a nuanced answer to our motivating question. People do not exclusively seek to maximize SWB (as operationalized by common SWB measures), yet SWB does seem to be a uniquely important argument in the utility function. Furthermore, as discussed shortly, the choice-SWB relationship varies systematically with scenario and SWB measure.

We use a variety of survey versions and empirical approaches in order to test the robustness of our main results to alternative interpretations. For example, while most of our data are gathered by eliciting both choice and predicted SWB rankings from each respondent, in some of our survey variations we elicit the two rankings far apart in the survey, or we elicit only choice ranking from some participants and only SWB ranking from others. To take another example, we assess the impact of measurement error by administering the same survey twice (at two separate times) to some of our respondents. While these different approaches affect our point estimates and hence the relative importance of our two main results, both results appear to be robust.

As steps toward providing practical, measure-specific and situation-specific guidance to empirical researchers as to *when* the assumption that people try to maximize predicted SWB is a better or worse approximation, we analyze how our results differ across SWB measures and across scenarios. Comparing SWB measures, we find that a "life satisfaction" measure (modeled after the first question in the epigraph) is a better predictor of choice than either of two "happiness" measures (modeled after the second and third questions in the epigraph) which perform similarly to each other. Comparing scenarios, we find that in scenarios constructed to resemble what our student respondents judge as representative of important decisions in their lives, predicted SWB coincides *least* often with choice, and other factors add relatively more explanatory power. This suggests that caution is especially warranted in drawing welfare conclusions from SWB data in such situations. Importantly, we also find that in scenarios where one alternative offers more money, respondents are systematically more likely to choose the money alternative than they are likely to predict it will yield higher SWB. This suggests that survey-based SWB measures may understate the utility benefits of higher income.

Our work is related to a literature in philosophy that poses thought experiments in hypothetical scenarios in order to demonstrate that people's preferences encompass more than their own happiness. That literature focuses on extreme situations such as being hooked up to a

machine that guarantees happiness.[5] In contrast, by focusing on realistic choice situations rather than on extreme ones, and by focusing on commonly-used SWB measures rather than on the concept of happiness, we seek to assess to what extent measured happiness is a good proxy for preferences in empirically-relevant situations.[6] As far as we are aware, the work whose methodology is most similar to ours is one of the experiments in Tversky and Griffin (1991), but their research question is different.[7]

The paper is organized as follows. Section I discusses the survey design and subject populations. Section II asks whether participants choose the alternative in our decision scenarios that they predict will generate greater SWB. Section III asks whether aspects of life other than SWB help predict choice, controlling for SWB, and compares the relative predictive power of the factors that matter for choice. Section IV presents robustness analyses. Section V discusses additional results. Section VI concludes, and discusses other possible applications of our methodology. Finally, the Appendix lists our decision scenarios, and the Web Appendix provides a chronology of all pilots; details information on all survey instruments; describes all of our robustness analyses and reports all the resulting tables; and reports an extensive set of additional analyses and results.

---

[5] In what is perhaps the most famous thought experiment, a reader is asked to introspect whether he or she would choose to be hooked up to an "experience machine" that would guarantee complete happiness for the rest of life, without awareness that the experience is not real (Nozick, 1974, pp.42-45). It is argued that if a reader would refuse the machine, he or she must not care exclusively about happiness. Less fancifully, extreme examples such as sacrificing one's life for a cause one believes in may also suggest that happiness is not the only maximand.

[6] Like us, Becker and Rayo (2008) propose empirical tests of whether things other than happiness matter for preferences in empirically-relevant choice situations. Relatedly, Truglia (2010) tests empirically whether the utility function inferred from consumption choices is distinguishable from the estimated happiness function over consumption. In contrast to our approach, where we measure individuals' predictions about SWB, these alternative tests rely on *ex post* measured happiness. Hence these tests assume that people correctly predict the happiness consequences of their choices, while our tests do not. However, we still interpret individuals' choices as revealing their "true preferences"—i.e. their assessment of their welfare—given their (potentially mistaken) beliefs about the consequences of their choices. This interpretation may be unjustified in situations where, e.g., due to social or moral pressure, choices do not reveal "true preferences" (see Koszegi and Rabin, 2008). To the extent that the decision scenarios we study involve such situations, our finding that people sometimes make choices that do not maximize predicted SWB may overstate the discrepancy between "true preferences" (or *welfare*) and SWB. However, our finding that people often choose an option that generates higher income despite predicting lower SWB for that option leans against this view; social or moral pressure would probably dictate choosing the *lower*-income option.

[7] In a between-subjects experiment with 66 undergraduates, Tversky and Griffin's participants were presented with a hypothetical choice between a job with a higher salary where co-workers earn even more versus a job with a lower salary where co-workers earn even less. Consistent with what we find in a similar choice problem, Tversky and Griffin find that participants are more likely to choose the higher absolute salary but say they would be happier with the higher relative salary. Tversky and Griffin interpret the result as supporting their theory that payoff levels are weighted more heavily in choice, while contrasts between payoffs and a reference point are weighted more heavily in SWB judgments.

I. Survey Design

While our main evidence is based on 29 different survey-versions, all versions share a similar underlying structure. Respondents are presented with a sequence of hypothetical scenarios, and in each they face a choice between two options. To illustrate, our 'Scenario 1' highlights a tradeoff between sleep and income. Followed by its SWB and choice questions, it appears on one of our questionnaires as follows:

> Say you have to decide between two new jobs. The jobs are exactly the same in almost every way, but have different work hours and pay different amounts.
>
> **Option 1:** A job paying $80,000 per year. The hours for this job are reasonable, and you would be able to get about 7.5 hours of sleep on the average work night.
>
> **Option 2:** A job paying $140,000 per year. However, this job requires you to go to work at unusual hours, and you would only be able to sleep around 6 hours on the average work night.
>
> Between these two options, taking all things together, which do you think would give you a happier life as a whole?

Option 1: Sleep more but earn less			**Option 2: Sleep less but earn more**		
definitely happier	*probably happier*	*possibly happier*	*possibly happier*	*probably happier*	*definitely happier*
X	X	X	X	X	X
Please circle one X in the line above					

> If you were limited to these two options, which **do you think you would choose**?

Option 1: Sleep more but earn less			**Option 2: Sleep less but earn more**		
definitely choose	*probably choose*	*possibly choose*	*possibly choose*	*probably choose*	*definitely choose*
X	X	X	X	X	X
Please circle one X in the line above					

In within-subject questionnaires, respondents are asked both the SWB question and the choice question above. In between-subjects questionnaires, respondents are initially or exclusively asked only one of the two questions.

I.A. Populations and Studies

We conducted surveys among 2,699 respondents from three populations: 1,066 patients at a doctor's waiting room in Denver who voluntarily filled out our questionnaires while waiting for their appointment; 1,000 adults who participated by telephone in the 2009 Cornell National Social Survey (CNSS) and form a nationally representative sample;[8] and 633 Cornell students who were recruited on campus and participated for pay or for course credit.[9] The Denver and Cornell studies include both within-subject and between-subjects survey variants, while the CNSS study is exclusively within-subject.[10]

Table 1 summarizes the design details of these studies. It lists each study's respondent population, sample size, scenarios used (see I.B below), types of questions asked (see I.C below), and other details such as response scales, scenario order, and question order.[11] The rest of this section explains the details summarized in the table.

I.B. Scenarios

Our full set of 13 scenarios is given in the Appendix, starting with Scenario 1 ("sleep versus income") from the example above. Referring to scenarios by their scenario number (1 through 13), table 1 reports which scenarios are used in which studies, and in what order they appear on different questionnaires. As detailed in the Appendix, some scenarios are asked in different variations (e.g. different wording, different quantities of money/sleep/commute time, etc.) and some scenarios are tailored to different respondent populations (e.g., while we ask students about school, we ask older respondents about work). In constructing the scenarios, we were guided by four considerations.

First, we chose scenarios that highlight tradeoffs between options that the literature suggests might be important determinants of SWB. Hence, respondents face choices between jobs and housing options that are more attractive financially versus ones that allow for: in

[8] The CNSS is an annual survey conducted by Cornell University's Survey Research Institute. For details: https://sri.cornell.edu/SRI/cnss.cfm.

[9] We conducted additional, smaller, studies among another 924 participants. Briefly, they included a pilot of the Cornell surveys with University of Chicago Students (102 respondents), a survey of Cornell students eliciting important decisions they face (171 respondents), as well as two surveys among the Denver sample (429 respondents) and a survey among Cornell students (222 respondents) that measured willingness-to-pay for greater SWB, among other things not related to this paper. For additional details, see the Web Appendix.

[10] In our Cornell between-subjects surveys, we use a design that allows us to also elicit within-subject data from our respondents, as described in section III.

[11] The median age in our Denver, CNSS, and Cornell samples is, respectively, 49, 49, and 21; the share of female respondents is 75, 53, and 60 percent. For summary statistics, see Web Appendix table A3.

Scenario 1, more sleep (Kahneman et al., 2004; Kelly, 2004); in Scenario 12, a shorter commute (Stutzer and Frey, 2008); in 13, being around friends (Kahneman et al., 2004); and in 3, making more money *relative* to others (Luttmer, 2005; see Heffetz and Frank, forthcoming, for a survey).

Second, since some of us were initially unsure we would find any divergences between predicted choice and SWB, in our earlier surveys we focused on choice situations where one's SWB may not be the only consideration. Hence, in Scenario 4 respondents choose between a career path that promises an "easier" life with fewer sacrifices versus one that promises posthumous impact and fame, and in Scenarios 2 and 11 they choose between a more convenient or "fun" option versus an option that might be considered "the right thing to do."

Third, once we found divergences between predicted SWB and choice, in our later surveys (the Cornell studies) we wanted to assess the magnitude of these divergences in scenarios that are representative of important decisions faced by our respondent population. For this purpose we asked a sample of students to list the three most important decisions they made in the last day, the last month, the last two years, and in their whole lives.[12] Naturally, decisions that were frequently mentioned by respondents revolved around studying, working, socializing and sleeping. Hence, in the resulting Scenarios 7-10, individuals have to choose between socializing and fun versus sleep and schoolwork; traveling home for Thanksgiving versus saving the airfare money; attending a more fun and social college versus a highly selective one; and following one's passion versus pursuing a more practical career path. To these scenarios we added Scenario 6, where respondents choose between giving up two hours a week versus paying a weekly sum of money. With this relatively pedestrian scenario we also hoped to get an insight into our subjects' time-versus-money tradeoff.

Fourth, as an informal check on our methods, we wanted to have one falsification-test scenario where we expected a respondent's choice and SWB ratings to coincide. For this purpose, we added Scenario 5, in which respondents face a choice between two food items (apple versus orange) that are offered for free and for immediate consumption. Since we carefully attempted to avoid any non-SWB differences between the options, we hypothesized that in this scenario, predicted SWB would most strongly predict choice. (Notice that this scenario has the additional attraction of being similar to prevalent decisions in almost everyone's

[12] The sample included 102 University of Chicago students; results were subsequently supported by surveying another 171 Cornell students. See the Web Appendix for full details (including the classification of responses by "decision type").

life, which is our third consideration above.)

I.C. Main Questions

(a) Choice question: In all studies, for each scenario, the choice question is presented as in our example above. In our analysis, we convert the horizontal six-point response scale into an intensity-of-choice variable, ranging from 1 to 6, or into a binary choice variable ("choose Option 1" versus "choose Option 2").[13] We view answers to this question as measures of preference, and the binary variable as a measure of (hypothetical) pairwise-choice.

(b) SWB question: While the language in the choice question is always kept the same, we vary the SWB question across questionnaire versions in order to examine how choice relates to several different, common SWB measures. In our Denver within-subject study we ask three different versions of the SWB question, modeled after what we view as three "families" of SWB questions that are commonly-used in the literature (see examples in the epigraph):

(i) life satisfaction: "Between these two options, which do you think would make you more satisfied with life, all things considered?";

(ii) happiness with life as a whole: "Between these two options, taking all things together, which do you think would give you a happier life as a whole?"; and

(iii) felt happiness: "Between these two options, during a typical week, which do you think would make you feel happier?"[14]

Since our between-subject tests have less statistical power than our within-subject tests, we ask only version (*i*) in our Denver between-subject surveys. In the CNSS study, where we were limited to one version of the SWB question due to design constraints, we ask only version (*ii*). As with the choice question, in the Denver study there are six possible answers (see the example

[13] Responses in the CNSS study are elicited as a binary choice rather than as a six-point scale, because in telephone interviews the binary format is both briefer for interviewers to convey and easier for respondents to understand.

[14] A fourth family of SWB questions, notably seen in the Gallup World Poll and Pew Global Attitudes surveys, and more recently on the HRS and its sister studies in Europe, asks respondents to rank their lives on an objective scale without using happiness or satisfaction language. For example, the Gallup World Poll asks: "Please imagine a ladder/mountain with steps numbered from zero at the bottom to ten at the top. Suppose we say that the top of the ladder/mountain represents the best possible life for you and the bottom of the ladder/mountain represents the worst possible life for you. If the top step is 10 and the bottom step is 0, on which step of the ladder/mountain do you feel you personally stand at the present time?" This family of SWB questions is used when difficulties in translation make standard SWB questions unreliable. We did not include this type of measure in our experimental surveys, but consider studying ladder/mountain measures with the tools developed here a promising topic for further research.

response-scales above) which we convert into either a six-point variable or a binary pairwise-SWB-ranking variable; in the CNSS study, response is binary.

Finally, as described shortly, in our Cornell studies we ask respondents about twelve different aspects of life, of which (one's own) happiness is only one. In that study we use versions of (*ii*) and (*iii*) that are modified to remain meaningful, with fixed wording, across aspects. The modified (*ii*) and (*iii*) result in these two new versions:

(iv) own happiness with life as a whole: "Between these two options, taking all things together, which option do you think would make your life as a whole better in terms of … [your own happiness]"; and

(v) immediately-felt own happiness: "Between these two options, in the few minutes immediately after making the choice, which option do you think would make you feel better in terms of … [your own happiness]."

The modified response scale now includes a middle "no difference" response, and has seven possible answers (*Option 1 definitely better; Option 1 probably better; Option 1 possibly better; no difference; Option 2 possibly better*, etc.). Respondents are allowed to indicate "no difference" for aspects of life that, in a given scenario, they judge irrelevant. We ask both (*iv*) and (*v*) in the Cornell within-subject surveys, and only (*iv*) in the Cornell between-subjects surveys.

(c) Questions about meta-choice and about other aspects of life: For completeness, we briefly mention here that some of our surveys contain the following additional questions. First, in all questionnaires of the Denver and Cornell within-subject studies, the choice question is followed by what we refer to as a meta-choice question: "If you were limited to these two options, which **would you want yourself to choose**?" Second, as mentioned above, the SWB question in all Cornell studies is modified to elicit ratings of the two scenario options in terms of "own happiness" as well as eleven additional aspects of life. For example, in versions (*iv*) and (*v*) of the SWB question (see language above), [your own happiness] may be followed by [your family's happiness], [your health], [your romantic life], etc.[15] We discuss these additional

[15] In some questionnaire versions, we separate "own happiness" from the other eleven aspects, and ask respondents first only about own happiness in each scenario, and then, re-presenting the scenarios, we ask about the other aspects. In these versions, we refer to the question on own happiness as an "isolated" measure of SWB (see table 1).

questions in sections III and IV, where we analyze the data they yield.

II. Do People Respond to the Choice and SWB Questions in the Same Way?

In this section we look at respondents' ranking of Option 1 versus Option 2 in terms of hypothetical choice compared with their ranking of the options in terms of predicted SWB. Viewing responses to the choice and the SWB questions as binary variables, we ask: *Do respondents answer the two questions as if they were answering the same question, or do they answer them systematically differently?* (We postpone analyzing responses as multi-point variables to section III.) Our analysis provides a first test of the null hypothesis that individuals seek to maximize SWB.

To answer our question, table 2 reports the distribution of binary responses to our surveys' choice and SWB questions by study and scenario, along with *p*-value statistics from equality-of-proportions tests. The table pools responses across SWB question variants (see I.C(b) and table 1 above); we discuss results by specific SWB measure below. The rest of this section interprets the figures reported in the table.[16]

II.A. Within-Subject Analysis

We start with table 2's leftmost column, which reports Scenario 1 figures from the Denver within-subject questionnaires (our "sleep versus income" scenario from the example in section I). The column's top four cells report a vertically-stacked 2×2 contingency matrix, consisting of the joint binary distribution of subjects who favor Option 1 in the choice question, and those who favor it in the SWB question. Looking at these four cells, we point out two facts that illustrate the present section's two main findings. First, the top two cells reveal that SWB response is highly predictive of choice response: between the two cells, 87 percent of respondents rank Option 1 versus Option 2 in the choice question the same as they do in the

---

In other versions, where the twelve aspects appear together either in the order specified above or in reverse order, we refer to the own happiness question as a "first/last in a series" measure.

[16] Non-response in our surveys was generally low. In the Cornell studies, virtually all questions had a non-response rate below 2 percent (one Cornell respondent was excluded due to obvious confusion with instructions). In the CNSS, less than 5 percent of respondents answered "Do not know" or refused to answer in any of the questions. Due to the less-structured recruiting method used in our Denver doctor's office studies, some questions from those studies had non-response rates as high as 20 percent. However, the majority of this non-response is driven by respondents being called in for their appointments, alleviating concerns of selection bias. Comparing the completed responses of subjects who did not finish the complete survey to the responses of those who finished the entire survey, we find no evidence of a difference in average responses.

SWB question. Second, the next two cells reveal systematic differences across the two questions among the remaining 13 percent of respondents: while 12 percent rank Option 1 (sleep) above Option 2 (income) in the SWB question and reverse this ranking in the choice question, only 1 percent do the opposite. This asymmetry suggests that on average, respondents react to the two questions systematically differently. The fifth cell reports the *p*-value from a paired Liddell exact test, a (nonparametric) equality-of-proportions test (Liddell, 1983).[17] The null hypothesis—namely, that the proportion of respondents who rank Option 2 above Option 1 does not depend on whether they answer the choice or the SWB question—is easily rejected.

Examining the top five rows in table 2 for the rest of the Denver within-subject columns verifies that the two main findings above are not unique to Scenario 1: in the remaining five scenarios, 81 to 90 percent of respondents rank the two options identically across the choice and SWB questions; yet in four out of five cases, choice-SWB reversals among the remaining 10 to 19 percent of respondents are asymmetric, and the equality-of-proportions null hypothesis across the two questions is easily rejected. In these cases, respondents rank income above legacy, concert above duty, low rent above short commute, and income above friends in higher proportions in the choice question than in the SWB question. Interestingly, there appears to be a systematic tendency among respondents to favor money in the choice question more than in the SWB question.

Similarly, the CNSS column suggests that qualitatively, Scenario 1's findings carry over from our Denver study—a pencil-and-paper survey with six-point response scales administered to a convenience sample—to the CNSS study—a telephone survey with binary response scales administered to a nationally representative sample. While the proportion of participants with no choice-SWB reversals increases to 92 percent, almost all of the rest—7 out of the remaining 8 percent—favor Option 1 (sleep) in the SWB question and Option 2 (income) in the choice question. The direction of this asymmetry is hence the same as in the Denver sample, and equality of proportions is again easily rejected.

Last among our within-subject data, results from the Cornell surveys are reported on the

[17] Conceptually, the Liddell exact test works as follows. For each respondent who answered both the SWB and the choice questions, one of his or her two responses is randomly (and hence, possibly falsely) reassigned to be treated as choice-response and the other as SWB-response. The difference between the proportions who favored Option 1 over Option 2 in the (reassigned) choice versus SWB questions is recorded. This procedure is then repeated for every possible permutation of assignments. The Liddell *p*-value reports the fraction of repetitions that resulted in a difference in proportions equal to or greater than the actual difference in proportions in our data. Computationally, *p*-values are calculated using an exact formula.

second page of table 2. The structure of this part of the table is similar to the corresponding Denver and CNSS parts, with the following three differences that result from the fact that the Cornell questionnaires allow for an additional "no difference" response in the SWB question: (a) an additional row below the top four rows reports the proportion of respondents who choose the "no difference" response; (b) the top four rows report vertically-stacked contingency matrices as before, only here they exclude these "no difference" responses (their sum is normalized to 100 percent after dropping these responses); and (c) the "no difference" responses are excluded from the Liddell tests.[18]

Starting again with Scenario 1 in the leftmost column, choice-SWB reversals (in the third and fourth rows, 24 percent together) are still a minority, although they are almost twice to three times more common in the Cornell sample than in the Denver and CNSS samples. Importantly, consistent with the Denver and CNSS data, in virtually all of these reversals—23 of the 24 percent—Option 1 (sleep) is ranked above Option 2 (income) in the SWB question and below it in the choice question. Equality of proportions is, again, strongly rejected for this scenario.[19]

Moving to the rest of the Cornell within-subject columns reveals a similar story. Equality of proportions is strongly rejected for all the remaining nine scenarios (2-10) as well, with the exception of Scenario 5. Recall that we constructed Scenario 5 ("apple versus orange") as a falsification test, to alert us if our methods produced choice-SWB reversals they should not. We predicted that, barring problems with our methods, choice and SWB should largely coincide for this scenario. Reassuringly, the results support this prediction. Indeed, only 5 percent of responses exhibit reversals in this scenario, by far the lowest fraction among the ten scenarios.[20]

---

[18] Since respondents must choose between Option 1 and 2 in the choice question, we cannot observe whether individuals indicating "no difference" for SWB support or weaken the equality-of-proportions null hypothesis. These individuals' responses in the choice question mirror the distribution of choice-responses among the rest of the respondents reasonably closely (Web Appendix table A5), and, reassuringly, the choice proportions in table 2 are virtually unaffected by excluding these individuals. Moreover, under the null hypothesis that choice is determined solely by predicted SWB, the distribution of choice-responses should be closer to 50-50 for individuals indicating SWB "no difference." Hence, the responses of these respondents actually provide additional suggestive evidence against the null hypothesis.

[19] Comparing each of the top four cells in the scenario 1 column across the three within-subject samples reveals that the reported proportions differ dramatically between the samples. Given the very different populations and, in the CNSS study, the very different survey methods, this finding in itself is not surprising. (For example, we speculate that since a telephone survey is harder to understand, more respondents answered the two questions in the same way, taking the "artificial consistency" mental shortcut discussed in II.B below.)

[20] Furthermore, we find no evidence that these reversals are in one systematic direction. At the same time, a sizeable 37 percent of respondents indicate "no difference" in the SWB question in scenario 5—by far the highest. This may suggest that scenario 5 is "cleaner" than we intended it to be: not only non-SWB aspects of life, but even own

As to the two other scenarios that are used in both the Denver and Cornell studies—Scenarios 3 and 4—choice-SWB reversals maintain their direction: in both studies, (absolute) income is ranked above relative income (Scenario 3) and above legacy (Scenario 4) in the choice questions more often than in the SWB questions. (While equality of proportions is rejected in the Cornell data but not in the Denver data in Scenario 3, it is rejected in both studies in Scenario 4.)

Finally, in Scenarios 6 and 8, which are used only in the Cornell studies and include a "money" option, we once again find that respondents favor money in the choice question more than in the SWB question. That this tendency holds in all seven scenarios that trade off more money/income for something else—be it more sleep, higher relative income, a legacy, a shorter commute, being around friends, having more time, or visiting family—suggests that survey-based SWB measures understate the utility benefits of higher income.

Our two main findings—that the ranking of the two options is identical across the choice and SWB questions for most respondents and in most scenarios, but that respondents react to the two questions systematically differently—hold not only in the pooled data, but also for each SWB question variant (*i*)-(*v*) separately. We show this in Web Appendix table A4, which reports versions of table 2 by SWB measure. Interestingly, we find some differences across the measures in the prevalence of choice-SWB reversals. In the Denver sample, the life satisfaction question variation (*i*) comes closest to matching choice, with only 11 percent of reversals, averaged across all scenarios. In comparison, happiness with life as a whole (*ii*) and felt happiness (*iii*) yield more reversals—17 percent each. In the Cornell sample, own happiness with life as a whole (*iv*) and immediately felt own happiness (*v*) both yield 22 percent reversals. We return to the comparison between different SWB measures in section V.A below.

II.B. Between-Subjects Analysis

Our within-subject analysis above is based on both choice and SWB responses elicited from each individual. However, empirical work that uses SWB data relies on surveys that measure SWB alone, not together with choice. Thus, two potential biases could compromise the relevance of our findings to existing SWB survey data and their applications. On the one hand,

---

happiness is deemed by many respondents irrelevant in what they may perceive as a context of *de gustibus non est disputandum*.

asking a respondent both questions might generate an "artificial consistency" between the two responses. For example, respondents might think they ought to give consistent answers, or might give consistent answers as an effort-saving mental shortcut. On the other hand, an "artificial inconsistency" bias is also possible if respondents infer from being asked more than one question that they ought to give different answers, or if the presence of the other question focuses participants' attention on the contrast between the wordings.

To address these concerns, we compare the above results from the Denver and Cornell within-subject studies with their counterpart between-subjects studies, in which respondents are asked either just the choice or just the SWB question. For ease of comparison, the between-subjects results are reported below the within-subject results on the two pages of table 2.

We start with the Denver sample. While the Denver within-subject questionnaires use the three SWB variants (*i*)-(*iii*), their between-subjects counterpart uses only variant (*i*), the life satisfaction measure. Looking again at Scenario 1, of the between-subjects respondents who are presented with this SWB question, 34 percent rank Option 2 (income) above Option 1 (sleep). Of the respondents who are presented with the choice question, the proportion is 44 percent. The direction of the discrepancy is the same as in the within-subject data and, as reported by the Fisher exact test *p*-value, the difference is statistically distinguishable from zero.[21] In contrast, the other columns show that there is essentially no difference between the responses to the choice and the SWB questions in the rest of the scenarios in the Denver between-subjects study.

Before comparing these proportion point estimates across the within- and between-subjects data, we first compare statistical results. Since the unpaired Fisher test has less power than the paired Liddell test, table 2 also reports Fisher test results for the within-subject data. They are calculated by randomly assigning our subjects into two equal groups, and examining only either their choice or their SWB response (hence "unpairing" these data). The reported average *p*-values (and average *n*) from 500 such repetitions show that a Fisher test rejects the no-reversals null in (on average) one or two of our six within-subject scenarios (compared with one of the five between-subjects scenarios). In other words, when using a comparable low-power

[21] We use the Fisher exact test as a non-parametric unpaired equality-of-proportions test. Conceptually, it works as follows. Remember that each respondent answers either the choice or the SWB question. These responses are pooled and reassigned (possibly falsely) into the two groups, and the proportion of those who rank Option 1 above Option 2 in each group is recorded. This procedure is repeated for all possible permutations of assignments. The reported p-value is the fraction of repetitions that resulted in difference in proportions at least as great as that between the two actual groups. Computationally, p-values are calculated using an exact formula.

test, overall statistical results seem comparable.

In order to compare the *magnitudes* of the choice-SWB reversals in the between- and within-subject data, we focus on only the three scenarios that are identical across the two studies[22] and on only those of the Denver within-subject respondents who answer the life satisfaction question (*i*) (reported in Web Appendix table A4). In this comparison, we find that while in Scenario 1 (income vs. sleep) discrepancies increase from 3 percentage points in the within- to 10 in the between-subjects data, in Scenario 13 (friends vs. income) they decrease from 7 to virtually nil, and in Scenario 3 (absolute income vs. relative income) they are essentially nil in both studies. Our overall interpretation of our Denver data with the life satisfaction measure—the measure with the fewest choice-SWB reversals—is that while there are clear differences across the within- and between-subjects studies, the evidence in either supports the same two main findings, albeit with differences in relative emphasis.

Finally, our Cornell between-subjects questionnaires ask the "own happiness with life as a whole" SWB question variant (*iv*). Starting again with the first scenario, results are strikingly similar to those in the Cornell within-subject data. On magnitudes, 66 percent favor Option 2 (income) in the choice question while 46 percent favor Option 2 in the SWB question, similar to the corresponding proportions—68 and 45 percent—among the Cornell within-subject respondents who are asked SWB question variant (*iv*) (see Web Appendix table A4).[23] On statistical results, in spite of the reduced power of the unpaired Fisher test in the bottom row compared with the paired Liddell test directly above it, equality of proportions between the choice and SWB questions is still easily rejected in the between-subjects study.

Looking at the rest of the scenarios (in both tables) suggests that the finding in Scenario 1—that the choice-SWB difference in proportions in the between-subjects data is similar to but slightly smaller than that in the within-subject data—is typical of only some scenarios. Of the ten scenarios in the Cornell studies, the difference maintains its sign but increases in one scenario (from 7 to 12 percent); decreases in four (from 23 to 20; 10 to 4; 11 to 9; and 11 to 8 percent); disappears in two (10 to 0; and 12 to 0 percent); and changes sign in three, including two very

[22] As reported in table 1 and the Appendix, only scenarios 1, 3, and 13 are comparable across the studies: scenario 2 is only included in the between-subjects study, and different versions of scenarios 4 and 12 are presented across the within- and between-subjects studies. For further explanation see the Web Appendix.

[23] Remember that in both studies, these figures exclude "no difference" responses in the SWB question. We note that such "no difference" responses seem more prevalent in the within- than in the between-subjects data. See Web Appendix table A4, where we report their prevalence by design variations.

large changes (1 to -5; 20 to -13; and 24 to -11 percent). The Fisher test at the bottom row rejects the null hypothesis in three of the ten scenarios in the Cornell between-subjects data—strikingly similar to the Fisher test reported in the "Average *p*-value from 500 repetitions" row for the within-subject data.

As with the Denver data, our overall interpretation is that while there are differences across the between- and the within-subject studies—in particular, choice-SWB reversals are on average less pronounced in the between-subjects study—either set of studies supports our two main findings.

II.C. Measurement Error

Our analysis above suggests that in many scenarios, individuals do not respond to the choice and SWB questions as if they were responding to the same question. However, in a given scenario, such rejection of the null hypothesis could be explained by differences in measurement error across the two questions—for example, because it is easier to introspect about choice than about SWB, or vice versa. An individual whose "true" ranking of the options is identical across the questions is more likely to mistakenly rank the "wrong" option higher in a question with greater measurement error, leading to ranking proportions closer to 50-50 for that question.

Looking across table 2's columns reveals that cross-question differences in the measurement error variances for choice and SWB in the same direction in all scenarios in a study cannot explain our data. For example, in the Denver within-subject data, choice proportions are closer to 50-50 in Scenarios 1, 11, and 13, but SWB proportions are closer in Scenarios 4 and 12. We postpone further discussion of measurement error to the end of section III.

To summarize, the two main findings in this section are (a) that most respondents in most scenarios do not exhibit choice- versus SWB-ranking reversals, and (b) that when they do, their pattern of reversals is systematic. Overall, the two findings hold up well—although with differences in relative strength—across scenarios, populations, and designs. Furthermore, these findings cannot be explained by a measurement error structure that is stable across scenarios.

III. Do Non-SWB Factors Help Predict Choice?

In this section we take a second approach to assessing whether individuals seek to

maximize SWB. We ask: *Can we identify non-SWB factors that help explain choices, controlling for predicted SWB?* We also analyze *by how much* respondents' choices can be explained by their predicted SWB and non-SWB aspects of life together, compared with their predicted SWB alone.

We address these questions using data from the Cornell sample, where we ask respondents to rank the options on a set of eleven non-SWB aspects of life, in addition to ranking them on choice and SWB (see section I.C). Specifically, in addition to being asked about "your own happiness," respondents are also asked about: your family's happiness, your health, your romantic life, your social life, your control over your life, your life's level of spirituality, your life's level of fun, your social status, your life's non-boringness, your physical comfort, and your sense of purpose. While still a limited list, we chose these aspects of life in an attempt to capture what else, besides one's own happiness, might enter one's preferences and hence matter for choice. Our list is guided by economists' and philosophers' enumeration of "capabilities" (Sen, 1985; Nussbaum, 2000), non-hedonic components of SWB proposed by psychologists (White and Dolan, 2009), and our own introspection (for example, we tried to include all aspects that money can buy, such as physical comfort).

As mentioned above, in our Cornell between-subjects surveys we use a design that allows us also to elicit within-subject data from our 201 participants. This is done by presenting subjects with the between-subjects part of the survey, followed by an additional, within-subject part.[24] In the between-subjects analysis in section II we used only data from the first, between-subjects part. In contrast, in this section we pool data from both parts, treating them as within-subject data. Further pooling these data with the original Cornell within-subject data (432 respondents) yields an augmented sample of 633 Cornell within-subject respondents. In this section we analyze this pooled data set. In section IV.E and in the Web Appendix we show that our main results hold in the constituent subsamples.

[24] Specifically, we present the entire sequence of ten scenarios three times. First, each scenario is presented and is followed by only a choice question (for half the respondents) or only a SWB question (for the other half). Second, after respondents finish answering their question for each of the ten scenarios, the ten scenarios are presented again, each followed by only the question (SWB or choice) respondents have not seen yet. Finally, the ten scenarios are presented for a third time, with each scenario followed by the non-SWB aspect questions. Respondents are specifically instructed to answer the surveys in exactly the order questions are presented, and the experimenters verify that they do (in the rare cases where a respondent was observed to flip through the pages, she/he was promptly reminded of this instruction). With this design, excluding data collected after the first round of scenario-presentation results in between-subjects data.

III.A. Response distributions

Figure 1 displays, by scenario, the histograms of raw, multi-point responses to the choice, SWB, and eleven non-SWB-aspect questions. Note first that the choice responses—and also the SWB responses, although to a lesser extent—tend to be bimodal with most of the mass on "definitely" or "probably," suggesting that the choice-SWB reversals discussed in section II are *not* the result of widespread near-indifferences. Second, notice that we were rather successful in constructing Scenario 5 (apple versus orange): almost everyone indicates "no difference" on non-SWB aspects of the options. While 37 percent also indicate "no difference" on SWB, the low count of reversals in Scenario 5 suggests that for the other respondents, variation in choice is strongly related to variation in SWB. Finally, note that in many other scenarios, there is substantial variation in the non-SWB aspect rankings, and that the histogram of choice responses sometimes looks rather different from the histogram of SWB responses.

III.B. Predicting choice

Table 3 presents a variety of specifications in which we regress choice on SWB and non-SWB aspects of life, aggregating data across the ten scenarios (we discuss regressions by scenario in section V.B below). We want to estimate the relationship from the within-scenario—rather than the between-scenario—variation in responses. For this purpose, in the probit and ordered probit specifications, we include scenario fixed effects.[25] In the OLS specifications, we demean all variables at the scenario level. Doing so yields identical coefficients to a fixed-effects OLS specification but has the advantage that the R^2's reflect only the within-scenario explanatory power of the regressors. Interpreting the first three columns in table 3 as an attempt to predict choice responses in our data, we first focus only on R^2's. We return to interpreting coefficients in the next subsection.

The first column of table 3 reports an OLS regression of six-point choice on seven-point SWB. As reported by the R^2, 0.38 of the variation in choice is explained by SWB (own happiness) alone. In comparison, a regression of the same choice measure on our eleven non-SWB aspects (each as a seven-point variable) yields an R^2 of 0.21 (second column of table 3).

[25] Including fixed effects in non-linear models is problematic in cases where the number of fixed effects is asymptotically infinite. Since our asymptotics rely on an infinite number of respondents, not of scenarios, the inclusion of scenario fixed effects does not harm the consistency of our estimates.

Hence, we find that SWB predicts choice substantially better than all of the other aspects combined. In the third column we regress choice on both SWB and the eleven non-SWB aspects. The R^2 of 0.41 is again substantially higher than that in the second column but is only slightly higher than that in the first column.[26] The pattern in these three columns is similar when we relax the linear functional form, replacing each regressor with a set of six dummy variables (not reported): the respective R^2's are 0.46, 0.33, and 0.50. In summary, when we pool data across scenarios we find that adding the non-SWB aspects to the regression of choice on SWB increases explanatory power, but the increase is quite modest. (However, the increase is substantial in some of the individual scenarios, as we report in section V.B.)

III.C. Comparing the coefficients

In order to interpret—and compare—the regression coefficients in table 3, we need to provide more structure. We hence assume that choices result from maximizing a utility function,

$$U(H(\boldsymbol{X}), \boldsymbol{X}), \tag{1}$$

where H is SWB and $\boldsymbol{X}$ is a vector of non-SWB factors that might affect utility both directly, and indirectly through H. If people seek to maximize SWB alone (as opposed to trading off SWB for other, non-SWB factors), then the (vector) partial derivative $\partial U/\partial \boldsymbol{X}$ will be identically zero. To a first-order approximation, the difference in utility between two options in a decision problem is

$$\Delta U \approx (\partial U/\partial H)\, \Delta H + (\partial U/\partial \boldsymbol{X})\, \Delta \boldsymbol{X}, \tag{2}$$

where ΔH and $\Delta \boldsymbol{X}$ are the differences between the options in predicted SWB and predicted non-SWB aspects, respectively. Naturally, the smaller are these differences, the more justified is the approximation—a point we return to shortly.

We can now interpret the OLS regression from the third column of table 3 as estimating the implied regression equation,

[26] Bootstrapped standard errors yield the following 95-percent confidence intervals around the three respective R^2's: [0.36, 0.40], [0.19, 0.23], and [0.39, 0.43].

(3) $$\Delta U_{is} = \beta_H \Delta H_{is} + \boldsymbol{\beta}_X \Delta \boldsymbol{X}_{is} + \varepsilon_{is},$$

where ΔU_{is}, ΔH_{is}, and $\Delta \boldsymbol{X}_{is}$ are, respectively, respondent i's choice, SWB, and non-SWB responses in Scenario s (raw responses demeaned at the scenario level); and ε_{is} is an error term, which captures the direct effects of unmeasured differences between the options as well as choice mistakes. Under this interpretation, the twelve regression coefficients estimate marginal utilities.

Under this approximation, the null hypothesis $\partial U/\partial \boldsymbol{X} \equiv \mathbf{0}$ can be rewritten as the hypothesis that the vector $\boldsymbol{\beta}_X = \mathbf{0}$, or that all eleven non-SWB coefficients in the third column are zero—a hypothesis we can easily reject (F-test $p = 0.000$). This result is robust to treating the choice measure as ordinal rather than cardinal, or as binary (the ordered probit and probit specifications in the fifth and sixth columns); to relaxing the linearity of our SWB measure by replacing it with a set of six dummy variables; and to combinations of these specifications (see tables A7-A9 in the Web Appendix for these and other specifications). Furthermore, with the exception of Scenario 8 (where F-test $p = 0.086$), the result holds in each individual scenario.[27] This and the robustness to relaxing the linearity of ΔH suggest that not all the marginal utilities $\partial U/\partial \boldsymbol{X}$ are zero, regardless of how good an approximation we use in equation (2).

To move from testing the null hypothesis to interpreting the magnitudes of coefficients requires additional assumptions. Econometrically, the coefficients in regression equation (3) are consistently estimated if $E(\varepsilon_{is} \mid \Delta H_{is}, \Delta \boldsymbol{X}_{is}) = 0$. We measure a long list of non-SWB aspects as an attempt to make this assumption more plausible; an unmeasured factor will bias the coefficient on the regressors it is correlated with. Psychologically, the coefficients are comparable if respondents respond to the seven-point scales similarly across the twelve aspects.

Comparing the coefficients in the third column of table 3, the coefficient on SWB is by far the largest. A one-point increase in our seven-point measure of predicted SWB is associated with a highly significant 0.46-point increase in our six-point choice measure. The largest non-SWB coefficients are on sense of purpose (0.12), control over one's life (0.08), family happiness (0.08), and social status (0.06). The relative sizes of the coefficients are similar in alternative

[27] We show in table A10 in the Web Appendix that this result holds even when the regressions include only aspects for which more than a trivial fraction of respondents (e.g. 15 percent) indicate answers other than "no difference." In other words, it holds even when we include only the most reliably-estimated coefficients. Interestingly, table A10 shows that the only large and robust non-SWB coefficient in the "apple versus orange" scenario is that on "physical comfort"; this seems consistent with the *de gustibus* (or taste-related) interpretation of this scenario.

specifications (e.g., the ordered probit column), but we remind the reader that the data are pooled across surveys that use two opposite orders in which aspects are presented, and order matters for the coefficient estimates (see section IV.E). While the rejection of $\boldsymbol{\beta}_X = \mathbf{0}$ suggests that SWB is not the only argument in the utility function, a comparison of the coefficients suggests that the marginal utility of SWB is several times larger than the marginal utilities of even the most significant non-SWB aspects we measure.[28]

Finally, related to the question to what extent people seek to maximize SWB, we can ask to what extent people prefer a higher level of some non-SWB aspect *because* it increases SWB. Rewriting equation (2) as $\Delta U \approx [(\partial U/\partial H)\,(\partial H/\partial \boldsymbol{X}) + (\partial U/\partial \boldsymbol{X})]\,\Delta \boldsymbol{X}$ decomposes the net effect on utility of a change in $\boldsymbol{X}$ into an indirect effect that operates through ΔH and a direct effect. The coefficients in the second column of table 3 estimate the sum of the indirect and direct marginal effects, $(\partial U/\partial H)\,(\partial H/\partial \boldsymbol{X}) + (\partial U/\partial \boldsymbol{X})$. In the third column, the coefficient on SWB estimates $\partial U/\partial H$, while the coefficients on the non-SWB aspects estimate the direct marginal effects, $\partial U/\partial \boldsymbol{X}$. Under this causal interpretation of the coefficients, comparing the second and third columns suggests that, for example, while the effects of health and life's non-boringness on utility operate only through their effects on SWB, the entire effect of social status operates directly (without affecting SWB). In between these extremes, aspects such as family happiness, control over one's life, life's level of fun, physical comfort, and sense of purpose seem to have roughly equal direct and indirect effects on utility.

III.D. Measurement error

Measurement error in ΔH_{is} or $\Delta \boldsymbol{X}_{is}$ will cause the assumption $E(\varepsilon_{is} \mid \Delta H_{is}, \Delta \boldsymbol{X}_{is}) = 0$ to fail, therefore biasing the coefficient estimates. Moreover, measurement error in ΔH_{is} could also invalidate our tests of the null hypothesis that $\boldsymbol{\beta}_X \equiv \mathbf{0}$; even if predicted non-SWB aspects do not enter utility directly, if they affect predicted SWB, and if predicted SWB is measured with error, we could find—as we do—that the predicted non-SWB aspects are statistically significant predictors of choice. In order to address these concerns, we collected repeated observations on a sub-sample (of 230) of our Cornell respondents. This enables us to estimate a measurement-error-corrected equation (3). Specifically, we use Simulation-Extrapolation (SIMEX) (Cook and

[28] However, we believe that the most plausible bias from unmeasured factors exaggerates the coefficient on SWB. Specifically, an unmeasured factor whose effect on ΔH has the same sign as its direct effect (i.e., not through ΔH) on ΔU will bias upward the coefficient on ΔH.

Stefanski, 1994), a semi-parametric method that assumes homoskedastic additive measurement error but does not make assumptions about the distribution of the regressors.[29] Intuitively, the SIMEX method proceeds in two steps. First, it simulates datasets with additional measurement error and uses them to estimate the function describing how the regression coefficients change with the amount of measurement error. Then the algorithm extrapolates in order to estimate what the coefficients would be if there were no measurement error in the original data. Table 3 displays the results of applying this method to our regressions. As expected, relative to the OLS results, the coefficient on own happiness increases, and remains by far the most predictive regressor. However, the non-SWB aspects with largest coefficients and statistical significance in the OLS regressions remain statistically significant and also increase.

These results suggest that SWB is not the only argument in the utility function and that non-SWB aspects of life enter preferences as well. At the same time, among the aspects we measure, SWB has by far the highest marginal utility.

IV. Robustness

To examine the robustness of our results from sections II and III, we conduct a long list of additional analyses. Full details, including all tables, charts, and statistics, are available in the Web Appendix. In this section we briefly summarize our findings. Unless stated otherwise, they are based on our within-subject data from either the Denver or the (augmented) Cornell samples.

IV.A. Are results driven by only a few individuals?

A natural interpretation of our evidence from sections II and III is that most people's preferences assign high but not exclusive weight to SWB, and therefore people usually but not always choose the option they predict would maximize SWB. However, an alternative interpretation is that most people exclusively seek to maximize SWB, and our results are driven by a small minority of participants who do not. In contrast with this latter interpretation, we find

[29] We choose to use the SIMEX method over several more common measurement error correction methods (such as IV or regression disattenuation) for several reasons. Primarily, the other methods are highly inefficient. Moreover, the SIMEX method is flexible in its treatment of the measurement error structure, it accommodates misclassified categorical data, and it easily accommodates non-linear models such as probit or ordered probit regressions. For additional discussion of SIMEX see the Web appendix, and for IV results see table A12 there. For example, in our IV estimates, standard errors are so large that we cannot reject that any of the OLS or SIMEX coefficients reported in table 3 are equal to the IV coefficients, with one exception: the 0.46 coefficient on own happiness in the OLS regression (third column) is significantly lower than the corresponding IV coefficient, which is 0.75.

that most respondents (both Denver and Cornell) exhibit at least one reversal and that very few exhibit reversals in half or more of the scenarios.

Moreover, we explore whether some of the respondents who do not exhibit a choice-SWB reversal in a given scenario *would have* done so if that scenario's tradeoff between SWB and non-SWB factors had assigned a different "price" to SWB. We do this by presenting some of our Denver respondents with four different versions of Scenario 4, a choice between a career as an artist who earns $40,000 annually and will have a lasting legacy versus a higher-paying career as a commercial artist who will leave no legacy. The versions differ from each other in the commercial artist's annual income: it is $42,000, $60,000, $80,000, or $100,000. We find that the fraction of participants who exhibit a choice-SWB reversal varies from as low as 3 percent in the $42k version to as high as 17 percent in the $80k version. Overall, 24 percent exhibit a reversal in at least one income-version, suggesting that the fraction of individuals for whom we observe a reversal in a given scenario is a lower bound on the total number who would exhibit a reversal in the same scenario with *some* "price of SWB."

IV.B. Scenario-order effects and participant fatigue

We investigate the effects of scenario order on responses with our Denver sample, where respondents face the six scenarios in one of two opposite orders (see table 1). Scenario-order effects could arise, for example, due to increasing fatigue or boredom among respondents. They could either increase or decrease the number of choice-SWB reversals through increasing response variance, decreasing respondents' effort, and in general changing artificial consistency or inconsistency patterns. While we indeed find evidence of scenario-order effects on responses in individual scenarios, they do not seem to influence our main findings. Averaged across the six scenarios, 15 percent of respondents exhibit a choice-SWB reversal in a scenario, regardless of whether it appears in the first or second half of the survey (Fisher exact test $p = 0.87$).

IV.C. Self-reported artificial consistency/inconsistency and mistakes

In section II we address artificial consistency and inconsistency concerns, as well as measurement error that in all scenarios is either larger for choice than for SWB, or vice versa. In section III, our measurement-error-corrected regressions address the possibility of response mistakes that are uncorrelated across repetitions of the survey. As an additional check on these

and other concerns, after our Cornell respondents finish responding to all the decision scenarios, we directly ask them whether their choice-SWB reversals were a mistake; whether they think they would regret them; and whether they were trying to make their choice and SWB responses consistent. Interestingly, only 7 percent report their reversals as a mistake, 23 percent report that they would regret them, and 20 percent report consciously trying to make their choice and SWB consistent. We repeat section III's analysis excluding groups of respondents based on their responses. We find, in line with our findings from our earlier approaches, that while mistakes and artificial consistency/inconsistency are likely present in our data, they alone cannot explain our results.

IV.D. Self-control

A related concern is that choice-SWB reversals may reflect a self-control problem (as in Laibson, 1997 and O'Donoghue and Rabin, 1999), rather than a preference for non-SWB aspects of life.[30] To address this concern, in some versions of the survey, in addition to asking participants what they would choose, we also ask them what they would want themselves to choose (the meta-choice question mentioned in Section I.C(c)). We reasoned that if, for example, a participant preferred Option 1 but would choose Option 2 due to a self-control problem, the participant would indicate favoring Option 2 in the choice question but Option 1 in the meta-choice question. Aggregating across all surveys that include the meta-choice question (see table 1), we find reversals between choice and meta-choice in 28 percent of the cases. However, while self-control problems may be relevant in these cases, our main conclusions from section III appear to be robust to excluding these observations. For example, in our benchmark specification from the third column of table 3, while excluding observations with a choice/meta-choice inconsistency increases the coefficient on own happiness, it also increases the coefficient on sense of purpose.

[30] In most of our decision scenarios, it is not obvious how a self-control problem would be implicated, but there are a few where it could be. For example, in scenario 7, one of the alternatives is staying out later with friends, while the other is going to bed earlier to feel better and be more productive the next day. A respondent who correctly anticipates having a self-control problem might respond that she would choose to stay out late, even though her welfare would be maximized by going to bed earlier.

IV.E. Context of choice, SWB, and non-SWB aspect questions

Closely related to concerns of artificial consistency and inconsistency, it is possible that our results are affected by the organization of our choice, SWB, and non-SWB aspect questions. Specifically, respondents' interpretations of the questions or their understanding of the meaning of the related concepts may be context-dependent.[31] For example, when the choice and SWB questions appear close together (as in our example scenario in section I), respondents may interpret each of the questions differently from how they interpret each when presented separately. Additionally, interpretations of the questions may depend on whether the choice question appears before or after the SWB question in each scenario.

It is similarly possible that respondents' interpretation of "own happiness" depends on whether it appears as a stand-alone question or as merely one aspect of life in a list of twelve. For example, when "own happiness" appears as an aspect in a list, respondents may interpret it to refer to only the part of own happiness that is not correlated with—or affected by—the other aspects on the list ("happiness controlled for other aspects of life"). Finally, the order in which the twelve aspects are presented may affect responses, for example because order affects the aspects' interpretation, or because respondents pay less attention to aspects at the bottom of the list.

As mentioned in sections I (see table 1) and III, different versions of our surveys vary in whether the choice and SWB questions are asked close together or far apart, and in the order the questions are asked; they also vary in the distance between the SWB and non-SWB aspects, and in the order of aspects. Repeating our analysis in section III by questionnaire organization verifies that such order and context effects do indeed matter. We briefly summarize the main findings. First, aspects listed earlier have larger coefficients. Second, own happiness as part of a twelve-aspects list has a smaller coefficient than as a stand-alone question. Third, a design where we present each scenario to solicit only choice for that scenario, then present each scenario again to solicit only happiness, and then present each scenario yet a third time to solicit only non-SWB aspect ratings yields the lowest coefficient on own happiness (0.31). That design also yields, in one of the aspect orderings, the highest coefficient on sense of purpose (0.27), making it the highest coefficient on any of our non-SWB aspects in any of the designs. Yet, in all designs, non-

[31] Notice the important difference between this possibility and the possibility of cross-respondent differences in the interpretations or understanding of the *scenarios*. The latter possibility is a lesser concern as long as a respondent's interpretation or understanding of a scenario remains the same across the choice and SWB questions.

SWB aspects are statistically significant, and the coefficient on own happiness has a higher point estimate than any of the non-SWB aspects. While the context in which we ask our key survey questions definitely matters, our basic results appear to be robust.

IV.F. An alternative approach: willingness-to-pay questions

Our empirical methodology in section III was to assess the marginal utility of SWB and other factors by confronting participants with a series of scenarios, and estimating the coefficients from a regression of choice on predicted SWB and non-SWB aspect ratings of the choice options. As another source of evidence on marginal utilities, which we can use to cross-validate our regression-based findings, we also ask questions aimed to directly elicit participants' willingness-to-pay (WTP) for improving SWB and other aspects of their lives.

Near the end of the Cornell surveys, after participants had faced all of the scenarios, we ask them, for the list of twelve aspects, the maximum amount of minutes per week they would be willing to invest in improving their rating of their life on that aspect by one point on a ten-point scale. Consistent with our regression-based findings, own happiness has the largest WTP.[32] Moreover, the correlation between mean WTP and the coefficient from a univariate OLS regression of choice on each aspect individually is 0.67 ($p < 0.02$).[33] SWB has a large impact on this correlation coefficient because it has both a high WTP and a high regression coefficient. To assess the robustness of the correlation, we omit SWB from the set of aspects; in that case, the correlation is 0.49, economically sizeable but not statistically significant ($p = 0.12$).[34]

We interpret these WTP-based findings as broadly supportive of our scenario-based findings.

---

[32] We measure WTP in units of time because we expect time to be more comparable across respondents than money for our student sample. Response categories are 0, 5, 10, 15, 20, 30, 45, 60, 75, and 120 minutes. Both mean WTP for own happiness (69 minutes) and its median (60 minutes) are well below the highest response option of 120 minutes (34 percent chose the latter). Across all of our survey populations, we similarly find that WTP for SWB is typically below the highest response option in a variety of questions including, for example, asking respondents what fraction of their income or what amount of meditation time they would be willing to devote to raise happiness. Although there are well-known reasons to be skeptical about contingent valuation questions, especially those far removed from everyday experience, this response pattern may be interpreted as suggesting that respondents think about a tradeoff between SWB and other aspects of their life when answering these questions.

[33] The univariate coefficient, as opposed to the coefficient from the multivariate regression reported in table 3, is the appropriate one to use for comparability with the WTP question because the WTP question does not partial out spillover effects on the other characteristics when eliciting WTP to improve one characteristic.

[34] These correlations might be inflated if aspects listed earlier have both larger coefficients (see section IV.E) and higher WTP; however, we find high correlations even between coefficients measured in one aspect order and WTPs measured in the opposite order.

V. Additional Results

To provide additional results that may be useful for happiness researchers and policy makers, we conduct the analysis from section III separately across SWB measures, scenarios, and respondent characteristics. This section is a brief summary of a more thorough treatment in the Web Appendix.

V.A. Comparing SWB measures

Across our surveys, we ask five different SWB question variants, based on three families of SWB questions asked in large-scale surveys used for empirical work: life satisfaction, happiness with life as a whole, and felt happiness (see section I.C(b) and table 1). In section II.A we compare the frequency of choice-SWB reversals across our different SWB measures. We now compare how well the measures predict choice using R^2's from univariate OLS regressions of our multiple-point choice variable on our multiple-point SWB measures.

As in section III, we demean our variables at the scenario level. In the Denver sample, the life satisfaction question variant (*i*) is the best predictor of the choice question, with $R^2 = 0.65$. Happiness with life as a whole (*ii*) and felt happiness (*iii*) come second and third, respectively, with $R^2 = 0.59$ and 0.55. The felt happiness R^2 is statistically significantly lower than the life satisfaction R^2 ($p = 0.02$ calculated using bootstrapped standard errors), and the R^2 for happiness with life as a whole is not statistically distinguishable from the other two. In the Cornell sample, own happiness with life as a whole (*iv*) and immediately felt own happiness (*v*) have $R^2 = 0.39$ and 0.37, not statistically distinguishable from each other.

These R^2's and our findings in II.A paint a consistent picture. While in the Denver data the life-satisfaction-type SWB question is more predictive of choice than the happiness-type SWB questions, in both Denver and Cornell the felt happiness and the happiness with life as a whole questions predict choice similarly. One possible hypothesis as to why some SWB measures predict choice better is that they encourage participants to report the present value of SWB flows over time. However, our finding that variant (*v*)—about happiness "in the few minutes immediately after making the choice"—is as predictive of choice as variant (*iv*)— about happiness in "life as a whole"—is inconsistent with this view.

V.B. Heterogeneity across decision scenarios

For applied work, it is useful to know in *which* situations the assumption that people try to maximize only SWB is a better or worse approximation. Table 4 shows the benchmark OLS specification from table 3, conducted separately for each of the ten scenarios in the Cornell data. The "Incremental R^2" row reports the difference between the reported R^2's (from the reported multivariate regressions) and R^2's from *univariate* regressions of choice on only SWB (which are not reported).

Mechanically, if utility is increasing in SWB, then a ranking of decision alternatives resulting from utility-maximization and a ranking resulting from SWB-maximization will coincide in choice settings where predicted SWB happens to be the only aspect that varies across the choice alternatives. As discussed above, Scenario 5 (apple versus orange)—which was designed to be as close as possible to such a case—has very little variance in non-SWB aspects and the fewest choice-SWB reversals (see figure 1 and table 2). As expected, the R^2 in a univariate regression of choice on predicted SWB is the highest (at 0.56) in Scenario 5, and the incremental R^2 from adding all other aspects is the lowest (at 0.02). If this type of minor decision—which possibly comprises most decisions in life—generally features low variance in non-SWB aspects, then the assumption that people try to maximize predicted SWB is a good approximation in such settings. The next highest univariate R^2 is for Scenario 3 (absolute income vs. relative income, at 0.49). This result is a surprise to us. We did not expect predicted SWB to capture the issues involved in the choice between absolute and relative income as well as it does.

Interestingly, at the other extreme, the four scenarios we designed to be representative of typical important decisions facing our college-age Cornell sample—Scenarios 7-10 (socialize versus sleep, family versus money, education versus social life, and interest versus career)—are among the scenarios with the lowest univariate R^2 and, correspondingly, the highest incremental R^2 from adding non-SWB aspects as regressors. Indeed, in Scenarios 7 and 10, where univariate R^2 is the lowest—at 0.25 and 0.24, respectively—incremental R^2 is 0.07 and 0.13. Here, non-SWB aspects increase predictive power (as measured by R^2) substantially, by 28 and 54 percent respectively. This in turn suggests that one should be especially cautious in interpreting measured SWB as utility in empirical applications that focus on important life decisions.

The rest of the scenarios fall somewhere in between. They include the scenarios that were designed to explore common themes from the happiness literature and, surprisingly, those designed as situations where we most expected to find tensions between SWB and non-SWB

factors (see section I.B).

Finally, table 4 reveals a heterogeneity in coefficients across scenarios (Chow test p = 0.000). Moreover, this heterogeneity appears to be systematic, with an increased coefficient on SWB and decreased coefficients on non-SWB factors in scenarios where there is less variance across the choice alternatives in non-SWB factors. In contrast, our simple theoretical framework from section III predicts constant marginal utilities and hence constant coefficients across scenarios. We speculate that this empirical heterogeneity results from the fact that in scenarios where some aspect varies less, that aspect becomes less salient as an input into utility, and respondents tend to take it into account less when making a choice.[35] Regardless of the explanation for the heterogeneity, it reinforces the assessment that predicted SWB may be a better approximation of utility in minor than in important decisions.

V.C. Preference heterogeneity across respondents

For a policymaker deciding whether to use SWB data as a proxy for utility for a particular demographic group, it may be useful to know how the propensity to maximize predicted SWB varies with demographic characteristics, such as gender, age, race, education, and income. In the Web Appendix, we report regressions that predict choice-SWB reversals as a function of respondent characteristics. In the Cornell sample, black respondents are 25 percentage points more likely than white ones to exhibit a choice-SWB reversal (with a standard error of 10 percentage points). However, on the whole across the Denver, CNSS, and Cornell samples, there is relatively little evidence for heterogeneity across demographic groups.

To see how heterogeneity in maximizing predicted SWB relates to well-understood dimensions of individual heterogeneity, we also predict the probability of a choice-SWB reversal

[35] In the Web Appendix we examine and rule out several alternative explanations. First, since our regression equation (3) relies on a first-order approximation of the utility function (1), one possible explanation of the above empirical heterogeneity is that the approximation breaks down in scenarios where the choice alternatives differ greatly in some aspects. Second, coefficients may be different in scenarios where both choice alternatives are far from the respondent's current situation, either because the relevant marginal utilities are not the same or because the respondent has difficulty predicting choice in fanciful hypothetical settings. We test these two explanations by running separate regressions for scenarios where we expect the local approximation to hold well or not; we find little evidence for differences in coefficients across these categories. Third, the coefficients may be identified off of different respondents in different scenarios. However, Cook's influence statistic (calculated at the respondent level) is highly correlated across most scenarios, and influence-robust iteratively reweighted least squares regressions still yield coefficients that vary across scenarios. This suggests that identification off of different respondents is not driving the variation in our coefficients across scenarios. We cannot rule out other explanations, such as respondents scaling their responses differently in different scenarios.

as a function of scores on the "Big 5" personality traits, which we measure in the Cornell within-subject sample using John and Srivastava's (1999) BFI scale. We find that a one standard deviation increase in conscientiousness is associated with an 8 percent lower likelihood of a choice-SWB reversal, while a one standard deviation increase in neuroticism (i.e., moody, tense) is associated with a 7 percent higher likelihood.

VI. Discussion

When aggregating our results across scenarios, we find that choices maximize predicted SWB (especially "life satisfaction") for most people most of the time. This may partially allay one source of concern about using SWB data as a welfare measure. However, care must be taken in interpreting our findings. The amount of choice-SWB concordance we find represents an upper bound on the degree to which the use of SWB measures is justified as a welfare measure in empirical applications. Applications always require additional assumptions that we do not test. For example, a typical assumption is that SWB measures are comparable and can be aggregated across individuals.[36]

When comparing scenarios, our results suggest that, first, researchers should be especially cautious about using SWB data as a proxy for utility in settings that are perceived to involve personally-important decisions. Second, in settings where one alternative involves higher income or more money, predicted SWB systematically understates the utility benefits of the money alternative. This in turn suggests that the increasingly-common practice of estimating willingness to pay for more sex, less pollution, etc., by comparing the coefficient on income with that on another variable in multivariate SWB regressions biases these estimates upwards.

Our scenario-based methodology could be usefully applied in several new directions. First, the method of assessing the correspondence between choice and predicted SWB could be used to assess new SWB measures. In the Web Appendix, we describe pilot data we collected on two such measures, neither of which appears to predict choice any better than existing measures. Measures that explicitly encourage integration of SWB over time, as well as ladder/mountain-

[36] Our results may also overstate the extent to which standard SWB questions provide a good measure of preferences because standard questions are asked absolutely ("How satisfied are with your life?"), while our SWB questions are asked relatively ("Between these two options, which do you think would make you more satisfied with life?"). Different individuals may apply different scales to a greater extent for an absolute measure, making it more difficult to translate an absolute SWB measure into a meaningful utility number than might be suggested by our results.

type measures, seem especially promising areas for further research.[37]

Second, our method could be used to provide more tailored guidance for applied work by asking about scenarios that are intended to address specific issues of interest. For illustrative purposes, we pilot four such scenarios at the end of our Cornell repeat-survey. For example, to reconcile the intuition that Americans today are better off than in the past with the finding that average SWB has remained flat in the U.S. over the past decades (Easterlin, 1974, 1995; see Stevenson and Wolfers, 2008, for a recent assessment), we ask respondents to rank being born in 1950 versus being born in 1990 in both choice and SWB questions. Although our 209 respondents overwhelmingly favor being born in 1990 in both questions (87 and 79 percent, respectively), the 9 percent who choose 1990 despite believing that they would be happier in 1950 is statistically significantly larger than the 1 percent exhibiting the reverse response pattern (Liddell test p-value < 0.001). This result indeed suggests that some people prefer being born later even if it does not make them happier. For another example, to reconcile the intuition that expanding political and economic freedoms for women have made women better off with the finding that average SWB among women has declined in the U.S. since the 1970s, both absolutely and relative to men (Stevenson and Wolfers, 2009), we ask respondents to rank living in a world with or without these expanded freedoms for women. Again, significantly more respondents choose a world with these expanded freedoms for women in spite of believing that a world without them would make them happier than the reverse (Liddell $p = 0.007$). For further examples and full details, see the Web Appendix.

Finally, some researchers have attempted to identify the key non-SWB aspects of life that are associated with greater welfare (e.g., Sen, 1985). Others have called for an SWB-based "national well-being index" to provide a measure of welfare that captures factors not represented in economic indicators such as GDP (e.g., Diener et al., 2009). To the best of our knowledge, our paper is the first attempt to empirically estimate the appropriate weights on SWB and other factors for combining them into an overall index of welfare. Our method could be applied more systematically for this purpose.

[37] Since different SWB questions seem to capture distinct dimensions of well-being that correlate differently with income and other variables (e.g. Kahneman and Deaton, 2010), future research could also explore whether a combination of SWB questions predicts choice better than any individual SWB question alone.

References

Becker, Gary S., and Luis Rayo. 2008. "Comment on 'Economic growth and subjective well-being: Reassessing the Easterlin Paradox' by Betsey Stevenson and Justin Wolfers." *Brookings Papers on Economic Activity*, Spring, 88-95.

Bentham, Jeremy. 1789. *An Introduction to the Principle of Morals and Legislations*.

Bernheim, Douglas B. 2009. "Behavioral Welfare Economics." *Journal of the European Economics Association*, 7(2-3), 267-319.

Blanchflower, David G. and Andrew J. Oswald. 2004. "Money, Sex and Happiness: An Empirical Study." *Scandinavian Journal of Economics*, 106(3), 393-415.

Bronsteen, John, Christopher J. Buccafusco, and Jonathan S. Masur. Forthcoming. "Well-being analysis." *Georgetown Law Journal*.

Cook, J. R. and Leonard A. Stefanski. 1994. "A simulation extrapolation method for parametric measurement error models." *Journal of the American Statistical Association*, 89, 1314-1328.

Diener, Ed, Richard Lucas, Ulrich Schimmack, and John Helliwell. 2009. *Well-Being for Public Policy*. New York: Oxford University Press.

DiTella, Rafael, Robert J. MacCulloch, and Andrew Oswald. 2003. "The Macroeconomics of Happiness." *Review of Economics and Statistics*, 85(4), 809-827.

Ellison, Christopher G. 1991. "Religious involvement and subjective well-being." *Journal of Health and Social Behavior*, 32, 80–99.

Easterlin, Richard A. 1974. "Does economic growth improve the human lot? Some empirical evidence." In Paul A. David and Melvin W. Reder (eds.), *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press, Inc.

Easterlin, Richard A. 1995. "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization*, 27, 35-47.

Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowigdo. 2008. "What good is wealth without health? The effect of health on the marginal utility of consumption." *NBER W.P. 14089*, June.

Frey, Bruno S., Simon Luechinger, and Alois Stutzer. 2008. "The life satisfaction approach to valuing public goods: The case of terrorism." *Public Choice*, 138, 317-345.

Gilbert, Daniel T. 2006. *Stumbling on happiness*. New York: Knopf.

Gruber, Jonathan, and Sendhil Mullainathan. 2005. "Do cigarette taxes make smokers happier?" *Advances in Economic Analysis & Policy*, 5(1) Article 4.

Heffetz, Ori, and Robert H. Frank. Forthcoming. "Preferences for Status: Evidence and Economic Implications." *Handbook of Social Economics*, Jess Benhabib, Alberto Bisin, Matthew Jackson, eds., Elsevier.

John, Oliver P., and Sanjay Srivastava. 1999. "The Big Five trait taxonomy: History, measurement, and theoretical perspectives." In Lawrence A. Pervin and Oliver P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford.

Laibson, David I. 1997. "Golden eggs and hyperbolic discounting." *Quarterly Journal of Economics*, 112(2), 443-477.

Layard, Richard. 2005. *Happiness*. Penguin Press.

Liddell, F. D. K. 1983. "Simplified exact analysis of case-referent studies: matched pairs; dichotomous exposure." *Journal of Epidemiology and Community Health*, 37(1), 82-84.

Luttmer, Erzo F.P. 2005. "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics*, 120(3), 963-1002.

Kahneman, Daniel, and Angus Deaton. 2010. "High income improves evaluation of life but not emotional well-being." *Proceedings of the National Academy of Sciences*, 107(38) 16489-16493.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Nobert Schwarz, and Arthur A. Stone. 2004. "A survey method for characterizing daily life experience: The Day Reconstruction Method." *Science*, 306, 1776-1780.

Kahneman, Daniel, Peter P. Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of experienced utility." *Quarterly Journal of Economics*, 112(2), 375-405.

Kelly, William. 2004. "Sleep-Length and Life Satisfaction in a College Student Sample." *College Student Journal,* 38(3).

Kimball, Miles S., Helen Levy, Fumio Ohtake, and Yoshiro Tsutsui, 2006. "Unhappiness After Hurricane Katrina." University of Michigan mimeo.

Kimball, Miles S., Ryan Nunn, and Daniel Silverman, 2010. "Accounting for Adaptation in the Economics of Happiness." Univeristy of Michigan mimeo.

Kimball, Miles S., and Robert J. Willis. 2006. "Happiness and Utility." University of Michigan mimeo.

Koszegi, Botond, and Matthew Rabin. 2008. "Choices, Situations, and Happiness." *Journal of Public Economics*, 92, 1821-1832.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

Nussbaum, Martha. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.

O'Donoghue, Ted, and Matthew Rabin. 1999. "Doing it now or later." *American Economic Review*, 89(1), 103-124.

Oreopoulos, Philip. 2007. "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling." *Journal of Public Economics*, 91, 2213–2229.

Sen, Amartya. 1985. *Commodities and Capabilities*. Oxford: Oxford University Press.

Stevenson, Betsey and Justin Wolfers. 2008. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox." *Brookings Papers on Economic Activity*, Spring, 1-87.

Stevenson, Betsey and Justin Wolfers. 2009. "The Paradox of Declining Female Happiness." *American Economic Journal: Economic Policy*, 1(2), 190-225.

Stutzer, Alois and Bruno S. Frey, 2008. "Stress that Doesn't Pay: The Commuting Paradox," *Scandinavian Journal of Economics,* 110(2), 339-366.

Stutzer, Alois. 2007. "Limited Self-Control, Obesity and the Loss of Happiness," IZA Discussion Papers 2925, Institute for the Study of Labor (IZA).

Truglia, Ricardo Nicolás Pérez. 2010. "A Samuelsonian Validation Test for Happiness Data," Harvard University mimeo, August 17.

Tsutsui, Yuichi, Miles Kimball, and Fumio Ohtake. 2010. "Koizumi Carried the Day: Did the Japanese Election Results Make People Happy and Unhappy?" *European Journal of Political Economy*, 26, 12-24.

Tversky, Amos, and Dale Griffin. 1991. "Endowments and Contrast in Judgments of Well-Being." In Richard J. Zeckhauser (ed.), *Strategy and Choice*, Cambridge, MA: MIT Press. Reprinted in Kahneman, Daniel, and Amos Tversky (eds.), *Choices, Values, and Frames*, Cambridge, UK: Cambridge University Press.

Welsch, Heinz. 2009. "Implications of happiness research for environmental economics." *Ecological Economics*, 68, 2735-2742.

White, Mathew P., and Paul Dolan. 2009. "Accounting for the richness of daily activities." *Psychological Science*, 20(8), 1000-1008.

Table 1: Study-specific Information

Study Location	**Denver**		**CNSS**	**Cornell**	
Choice vs. SWB: Within- or Between- Subjects	Within	Between	Within	Within	Between
Sample Population	Volunteers at a doctor's waiting room		Nationally representative	Cornell Students	
Observations	497	569	1000	432†	201
Scenarios used	1, 3, 4, 11, 12, 13	1, 2, 3, 4, 12, 13	1	1-10	
SWB Question Format	Observations for each SWB question variant in parenthesis				
(*i*) Life Satisfaction					
(Isolated)	(164)	(569)			
(*ii*) Happiness with Life as a Whole					
(Isolated)	(162)		(1000)		
(*iii*) Felt Happiness					
(Isolated)	(171)				
(*iv*) Own Happiness with Life as a Whole					
Isolated				(107)	(201)
First/Last In Series				(107)	
(*v*) Immediately Felt Own Happiness					
Isolated				(110)	
First/Last In Series				(108)	
SWB Response Scale	6-point		Binary	7-point	
Choice Response Scale	6-point		Binary	6-point	
Meta-Choice Question?	Yes	No	No	Yes	No
Order variations					
Scenario order	4-1-11-12-13-3 3-13-12-11-1-4	1-2-12-13-3-4 3-13-12-2-1-4‡	1	1-2- … -9-10	
Question order	Choice-Meta-SWB SWB-Choice-Meta		SWB-Choice	Choice-SWB	
Aspects of life order				Two opposite orderings of 11 non-SWB aspects	
Summary: number of questionnaire versions	12	4	1	8	4

Notes: See section I for details and for the framing of the choice, SWB, and meta-choice questions; see the Appendix for the list of scenarios (1-13).
† Of these, 230 were surveyed twice, allowing us to conduct measurement-error-corrected estimation.
‡ Scenario 4 is always presented last because it is followed by both a choice and a SWB question. In order to have a clean between-subjects design, we did not want subjects to know we were interested in both choice and SWB until after subjects were done with the rest of the scenarios. We also note that this scenario is presented in four different order-versions, so strictly speaking the Denver between-subjects study includes the four questionnaire versions reported in the table's bottom row, times four (sixteen versions in total). We report the number four because in terms of the between-subjects part of this study, there are four versions.

Table 2: Choice and SWB Responses Across Studies and Scenarios *(continued on next page)*

		Denver Study							CNSS
	Choice Scenario	1	2	3	4	11	12	13	1
	For exact phrasing see Appendix	Sleep vs Income	Concert vs Birthday	Abs. Inc. vs Rel. Inc.	Legacy vs Income	Concert vs Duty	Low Rent vs Short Commute	Friends vs Income	Sleep vs Income
Within	Higher SWB: Option 1 Chosen: Option 1	58%		48%	24%	16%	52%	50%	74%
	Higher SWB: Option 2 Chosen: Option 2	29%		42%	60%	65%	32%	34%	18%
	Higher SWB: Option 2 Chosen: Option 1	1%		6%	2%	12%	11%	2%	1%
	Higher SWB: Option 1 Chosen: Option 2	12%		4%	14%	7%	5%	14%	7%
	p-value from Liddell Exact Test	0.000		0.350	0.000	0.024	0.002	0.000	0.000
		n = 425		n = 420	n = 422	n = 422	n = 425	n = 422	n = 980
	Average p-value from 500 repetitions of Fisher Exact Test treating within-subject data *as if* they were between-subjects	0.111		0.527	0.044	0.359	0.315	0.071	
		n = 427		n = 423	n = 427	n = 423	n =427	n = 424	
					Version 2		Version 2		
Between	Higher SWB: Option 2	34%	86%	51%	***		54%	54%	
	Chosen: Option 2	44%	84%	48%	***		55%	53%	
	p-value of Fisher Test	0.020	0.716	0.433	***		0.930	0.793	
		n = 525	n = 524	n = 525			n = 526	n = 525	

Notes: Response distribution by study and scenario. For the complete text of each scenario, see the Appendix. If a scenario's phrasing changed meaningfully between surveys, the version of the scenario is indicated in the first row of the study block. For between-subjects data, we report the Fisher Exact Test *p*-value testing the null hypothesis that mean response to choice question = mean response to SWB question (an unpaired equality-of-proportions test). For within-subject data, we report the analogous Liddell Exact Test *p*-value (a paired equality-of-proportions test). In cases where respondents could indicate SWB indifference, responses indicating indifference were dropped from these tests. To conduct hypothesis tests with equal power for the within- and between-subjects data, we treat the within-subject data as if they were between-subjects by randomly assigning half of the observations to choice and half to SWB, and only looking at each respondent's assigned response. The reported *p*-value and *n* are the average across 500 repetitions of this algorithm. See details in section II. Since Scenario 4 ("legacy vs income") was not presented in a between-subjects design in spite of being included in the between-subjects part of the Denver study, its results from that part are not reported. For a description of this scenario's results, see section IV.A.

Table 2: Choice and SWB Responses Across Studies and Scenarios *(continued from previous page)*

Cornell Study

	Choice Scenario *For exact phrasing, see Appendix*	1 Sleep vs Income	2 Concert vs Birthday	3 Abs. Inc. vs Rel. Inc.	4 Legacy vs Income	5 Apple vs Orange	6 Money vs Time	7 Socialize vs Sleep	8 Family vs Money	9 Education vs Social life	10 Interest vs Career
		Version 2									
Within	Higher SWB: Option 1 Chosen: Option 1	29%	29%	41%	44%	45%	44%	62%	68%	53%	27%
	Higher SWB: Option 2 Chosen: Option 2	46%	49%	43%	31%	50%	37%	15%	15%	22%	35%
	Higher SWB: Option 2 Chosen: Option 1	1%	7%	14%	8%	2%	14%	17%	5%	22%	3%
	Higher SWB: Option 1 Chosen: Option 2	23%	15%	2%	17%	3%	5%	6%	12%	3%	35%
	Indifference for SWB	8%	14%	13%	10%	37%	22%	10%	5%	6%	6%
	p-value of Liddell Exact Test	0.000	0.002	0.000	0.001	0.424	0.000	0.000	0.001	0.000	0.000
		n = 397	n = 368	n = 375	n = 387	n = 270	n = 333	n = 385	n = 409	n = 402	n = 402
	Average p-value from 500 repetitions of Fisher Exact Test.	0.025	0.387	0.234	0.348	0.531	0.333	0.184	0.377	0.035	0.002
		n = 197	n = 196	n = 193	n = 196	n = 168	n = 194	n = 198	n = 198	n = 198	n = 198
		Version 2									
Between	Higher SWB: Option 2	46%	67%	49%	38%	56%	34%	26%	14%	16%	61%
	Chosen: Option 2	66%	71%	40%	38%	51%	34%	18%	26%	29%	50%
	p-value of Fisher Test	0.006	0.643	0.195	1.000	0.537	1.000	0.234	0.048	0.027	0.117
		n = 197	n = 196	n = 193	n = 196	n = 168	n = 194	n = 198	n = 198	n = 198	n = 198

Notes: Response distribution by study and scenario. For the complete text of each scenario, see the Appendix. If a questions phrasing changed meaningfully between surveys, the version of the question is indicated in the first row of the study block. For between-subjects data, we report the Fisher Exact Test *p*-value testing the null-hypothesis that mean response to choice question = mean response to SWB question (an unpaired equality-of-proportions test). For within-subject data, we report the analogous Liddell Exact Test *p*-value (a paired equality-of-proportions test). In cases where respondents could indicate SWB indifference, responses indicating indifference were dropped from these tests. To conduct hypothesis tests with equal power for the within- and between-subjects data, we treat the within-subject data as if they were between-subject by randomly assigning observations to choice or to SWB in numbers that match the between-subjects test, and only looking at each respondent's assigned response. The reported *p*-value and *n* are the average across 500 repetitions of this algorithm. See details in section II.

Table 3: Regressions of Choice on Aspects of Life

	OLS				Ordered Probit	Probit	
Measurement error correction	None	None	None	SIMEX Corrected Additive	None	None	SIMEX Corrected Additive
Own happiness	0.54***		0.46***	0.60***	0.37***	0.37***	0.49***
	(0.009)		(0.010)	(0.015)	(0.009)	(0.012)	(0.021)
Family happiness		0.15***	0.08***	0.14***	0.06***	0.09***	0.17***
		(0.017)	(0.015)	(0.032)	(0.012)	(0.017)	(0.040)
Health		0.07**	0.00	0.02	0.01	0.01	0.05
		(0.021)	(0.019)	(0.041)	(0.016)	(0.022)	(0.055)
Life's level of romance		-0.00	-0.01	0.04	-0.00	-0.00	0.08
		(0.024)	(0.021)	(0.047)	(0.018)	(0.025)	(0.061)
Social life		-0.01	-0.03	-0.07*	-0.02	-0.02	-0.05
		(0.020)	(0.018)	(0.035)	(0.015)	(0.021)	(0.044)
Control over your life		0.17***	0.08***	0.12***	0.06***	0.09***	0.14***
		(0.017)	(0.015)	(0.029)	(0.012)	(0.017)	(0.033)
Life's level of spirituality		-0.08**	-0.02	-0.03	-0.02	-0.04	-0.04
		(0.024)	(0.021)	(0.054)	(0.018)	(0.025)	(0.066)
Life's level of fun		0.13***	0.05*	0.05	0.04*	0.04*	0.04
		(0.021)	(0.018)	(0.039)	(0.015)	(0.021)	(0.045)
Social status		0.07***	0.06***	0.08**	0.05***	0.07***	0.12***
		(0.016)	(0.014)	(0.027)	(0.012)	(0.016)	(0.033)
Life's non-boringness		0.07***	-0.01	-0.00	0.00	0.00	0.02
		(0.020)	(0.017)	(0.037)	(0.014)	(0.020)	(0.047)
Physical comfort		0.09***	0.04**	0.03	0.04**	0.05**	0.05
		(0.017)	(0.014)	(0.029)	(0.012)	(0.017)	(0.037)
Sense of purpose		0.21***	0.12***	0.14***	0.10***	0.12***	0.16***
		(0.015)	(0.013)	(0.026)	(0.011)	(0.015)	(0.029)
Observations	6285	6220	6217	6217	6217	6217	6217
(pseudo) R^2	0.38	0.21	0.41		0.19	0.35	

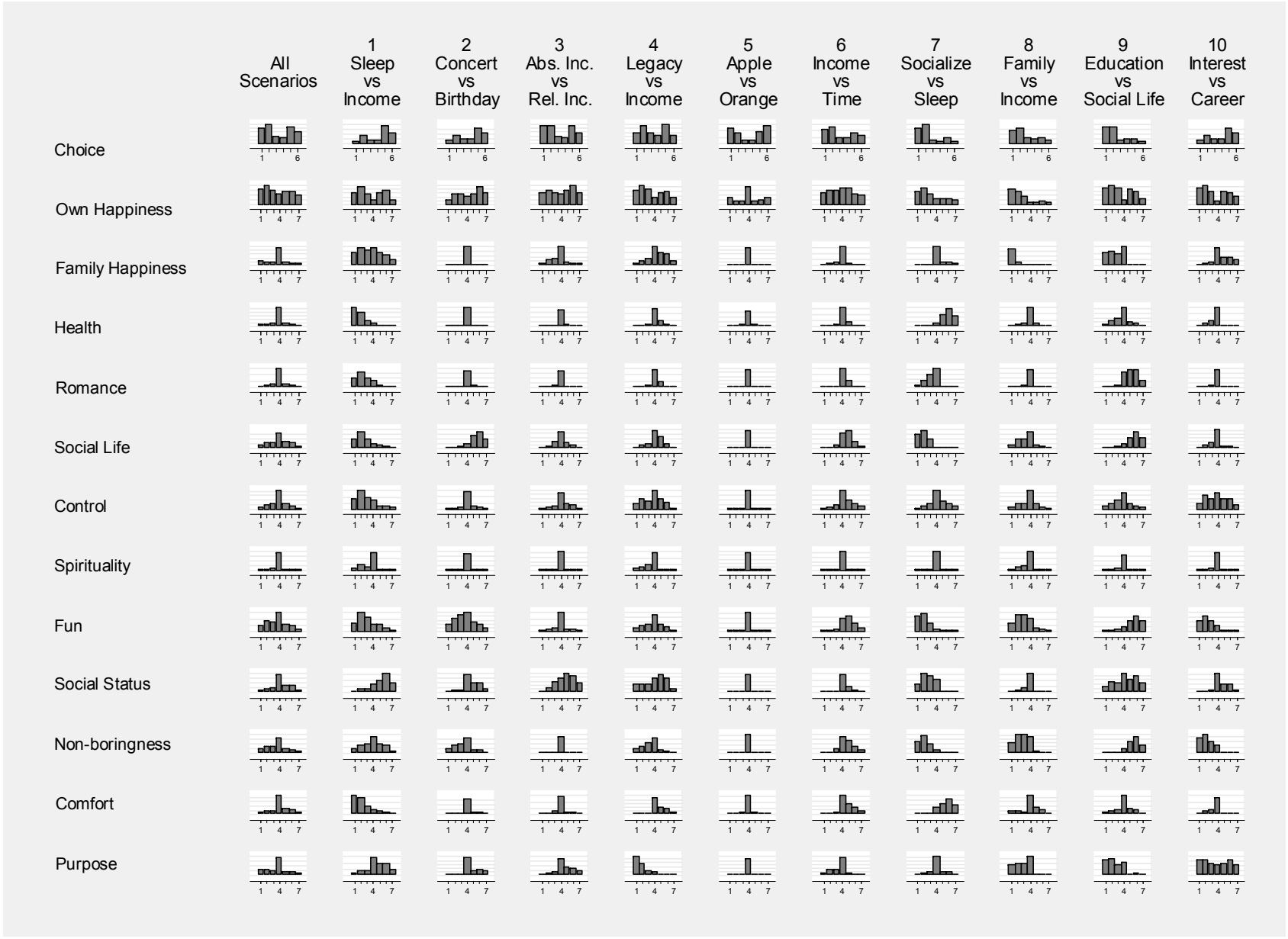
Notes: Standard errors in parentheses. In the OLS and ordered probit regressions, the dependent variable is 6-point choice. In the probit regressions the dependent variable is binary choice. All regressions use 7-point ratings of aspects. Based on 633 Cornell respondents. Each observation is a respondent's ratings for one scenario; there are 10 observations per respondent corresponding to the 10 scenarios in the questionnaires. Probit and ordered probit regressions include (unreported) scenario fixed effects. OLS regressions' variables are demeaned at the scenario level, generating coefficients equivalent to including scenario fixed effects. Measurement error corrections are done using the simulation extrapolation method described in section III, under the assumption of additive measurement error. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: OLS Regressions of Choice on All Aspects of Life, by Scenario

Choice Scenario *For exact phrasing, see Appendix*	All questions	1 Sleep vs Income	2 Concert vs Birthday	3 Abs. Inc. vs Rel. Inc.	4 Legacy vs Income	5 Apple vs Orange	6 Money vs Time	7 Socialize vs Sleep	8 Family vs Money	9 Education vs Social life	10 Interest vs Career
Own happiness	0.46***	0.38***	0.44***	0.52***	0.44***	0.73***	0.53***	0.31***	0.53***	0.35***	0.27***
	(0.010)	(0.031)	(0.031)	(0.032)	(0.031)	(0.036)	(0.036)	(0.032)	(0.033)	(0.029)	(0.030)
Family happiness	0.08***	0.07*	0.01	0.16***	0.05	0.16	0.15*	-0.09	0.05	0.14***	0.21***
	(0.015)	(0.032)	(0.071)	(0.046)	(0.041)	(0.159)	(0.059)	(0.053)	(0.050)	(0.037)	(0.041)
Health	0.00	-0.05	-0.07	-0.11	-0.04	0.05	0.06	0.18***	0.05	-0.03	-0.06
	(0.019)	(0.055)	(0.076)	(0.077)	(0.058)	(0.065)	(0.075)	(0.054)	(0.057)	(0.044)	(0.063)
Life's level of romance	-0.01	0.08	-0.02	0.07	-0.00	-0.67**	-0.10	0.02	-0.03	0.01	0.01
	(0.021)	(0.059)	(0.064)	(0.078)	(0.066)	(0.228)	(0.086)	(0.054)	(0.068)	(0.053)	(0.072)
Social life	-0.03	-0.02	0.02	-0.01	0.00	0.02	0.04	-0.00	-0.05	-0.04	0.01
	(0.018)	(0.055)	(0.043)	(0.056)	(0.058)	(0.225)	(0.071)	(0.065)	(0.053)	(0.053)	(0.054)
Control over your life	0.08***	0.02	0.05	0.04	0.08*	-0.00	0.07	0.15***	0.05	0.06	0.07*
	(0.015)	(0.042)	(0.053)	(0.056)	(0.039)	(0.093)	(0.052)	(0.043)	(0.049)	(0.038)	(0.035)
Life's level of spirituality	-0.02	-0.04	-0.00	-0.16	0.13*	0.31	-0.15	-0.01	-0.15*	-0.00	-0.01
	(0.021)	(0.049)	(0.061)	(0.090)	(0.055)	(0.221)	(0.091)	(0.076)	(0.062)	(0.054)	(0.068)
Life's level of fun	0.05*	0.06	0.15**	0.04	0.05	-0.08	0.13	-0.03	0.03	0.06	-0.00
	(0.018)	(0.042)	(0.051)	(0.066)	(0.047)	(0.127)	(0.068)	(0.073)	(0.059)	(0.057)	(0.057)
Social status	0.06***	-0.00	0.04	0.05	0.04	-0.27	-0.01	0.06	0.11	0.06*	0.16***
	(0.014)	(0.036)	(0.045)	(0.040)	(0.036)	(0.227)	(0.061)	(0.059)	(0.060)	(0.029)	(0.043)
Life's non-boringness	-0.01	0.05	-0.03	0.22**	-0.01	0.09	-0.03	0.18**	-0.05	-0.02	0.05
	(0.017)	(0.037)	(0.054)	(0.078)	(0.047)	(0.121)	(0.060)	(0.062)	(0.061)	(0.055)	(0.055)
Physical comfort	0.04**	0.09*	0.00	-0.05	0.00	0.21**	-0.00	0.05	-0.10*	0.06	-0.02
	(0.014)	(0.036)	(0.060)	(0.054)	(0.042)	(0.066)	(0.049)	(0.048)	(0.041)	(0.040)	(0.049)
Sense of purpose	0.12***	0.17***	0.12**	0.12**	0.12**	0.29*	0.05	0.04	0.09*	0.17***	0.17***
	(0.013)	(0.038)	(0.047)	(0.044)	(0.041)	(0.119)	(0.050)	(0.044)	(0.046)	(0.037)	(0.029)
Observations	6217	615	621	620	624	624	619	622	625	626	621
R^2	0.41	0.46	0.43	0.53	0.41	0.58	0.42	0.32	0.38	0.43	0.37
Incremental R^2	0.03	0.06	0.03	0.04	0.04	0.02	0.02	0.07	0.02	0.08	0.13

Notes: Standard errors in parentheses. OLS regressions of 6-point choice on 7-point aspects of life. Based on 633 Cornell respondents. The leftmost column aggregates data across choice scenarios; each of the other columns corresponds to a specific scenario. Each observation is a respondent's ratings for one scenario; there are 10 observations per respondent corresponding to the 10 scenarios in the questionnaires. All variables are demeaned at the scenario level, generating coefficients equivalent to including scenario fixed effects. "Incremental R^2" is the difference in R^2 between the reported multivariate regression and a univariate regression of choice on happiness; it represents the increased percentage of variation in choice that can be explained by including the additional aspects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Raw Response Distributions (Choice and Aspects of Life)



Notes: Based on 633 Cornell respondents. See details in text. The histograms show the distribution of 6-point responses to the choice question (top row) and 7-point responses to the aspect questions (bottom twelve rows). The leftmost column aggregates data across choice scenarios; each of the other columns corresponds to a specific scenario.

Appendix: Scenarios Presented in Surveys

Scenario 1: Sleep vs. Income

Say you have to decide between two new jobs. The jobs are exactly the same in almost every way, but have different work hours and pay different amounts.

Option 1: A job paying $80,000 per year. The hours for this job are reasonable, and you would be able to get about 7.5 hours of sleep on the average work night.

Option 2: A job paying $140,000 per year. However, this job requires you to go to work at unusual hours, and you would only be able to sleep around 6 hours on the average work night.

Scenario 2: Concert vs. Birthday

Suppose you promised a close friend that you would attend his or her 50th ["21st" in student samples] birthday dinner. However, at the last minute you find out that you have won front row seats to see your favorite musician, and the concert is at the same time as the dinner. This is the musician's last night in town. You face two options:

Option 1: Skip your friend's birthday dinner to attend the concert.

Option 2: Attend your friend's birthday dinner and miss the concert.

Scenario 3: Absolute Income vs. Relative Income

Suppose you are considering a new job, and have offers from two companies. Even though all aspects of the two jobs are identical, employees' salaries are different across the two companies due to arbitrary timing of when salary benchmarks happened to be set. Everyone in each company knows the other employees' salaries. You must choose one of the two companies, which means you must decide between the following two options:

Option 1: Your yearly income is $105,000, while on average others at your level earn $120,000.

Option 2: Your yearly income is $100,000, while on average others at your level earn $85,000.

Scenario 4: Legacy vs. Income

(*Phrasing in study 1)***:** Suppose you are a skilled artist, and you have to decide between two career paths for your life.

Option 1: You devote yourself to your own style of painting. This would require a number of sacrifices, such as having less time for friends and family, and making less money. For example, you expect that selling your paintings will give you an income of $40,000 a year. If you choose this path, you don't expect that your work will be appreciated in your lifetime, but posthumously you will make an impact on the history of art, achieve fame, and be remembered in your work.

Option 2: You become a graphic designer at an advertising company. This would give you more money and more time with friends and family than Option 1. The company is offering you a salary of $60,000 a year, which will afford you a much more comfortable lifestyle, but you will have no impact and leave no legacy to be remembered.

(*Phrasing in studies 2 and 4***) :** Suppose you are a skilled artist, and you have to decide between two career paths for your life. There are two styles of painting that you consider to be your own style, and you enjoy both equally. Style 1 happens to be much less popular than Style 2 today, but you know it will be an important style in the future.

Option 1: You devote yourself to Style 1. You expect that selling your paintings will give you an income of $40,000 a year. If you choose this path, you don't expect that your work will be appreciated in your lifetime, but posthumously you will make an impact on the history of art, achieve fame, and be remembered in your work.

Option 2: You devote yourself to Style 2. You expect that selling your paintings will give you an income of $60,000 a year, but you will have no memorable impact. [In Study 2, each subject saw this question three times, with different salaries in option 2. This number took a value of

either $42,000, $60,000, $80,000, or $100,000.]

Scenario 5: Apple vs. Orange

Suppose you are checking out a new supermarket that just opened near where you live. As you walk by the fresh fruit display, you are offered your choice of a free snack:

Option 1: A freshly sliced apple.

Option 2: A freshly sliced orange.

Scenario 6: Money vs. Time

Suppose that due to budget cuts, the school implements a "student activities fee" of $15 dollars a week to help pay for maintenance of facilities used for extracurricular student activities. However, the school allows you to not pay the fee if instead you put in 2 hours of service a week shelving books at the library. You face two options:

Option 1: Spend 2 hours a week shelving books.

Option 2: Pay $15 a week.

Scenario 7: Socialize vs. Sleep

Say you are hanging out with a group of friends at your friend's room. You are having a really good time, but it is getting to be late at night. You have to decide between two options.

Option 1: Stay up another hour. It is likely you will feel tired all day tomorrow, but this particular evening you are having an especially fun time.

Option 2: Excuse yourself from the group, and go to bed. You will be disappointed to miss the fun, but you know you will feel better the next day and be more productive at paying attention in class and doing your homework.

Scenario 8: Family vs. Money

Imagine that for the first time in three years, your parents (or if your parents are gone, your closest relatives who are older than you) have arranged for a special family gathering that will happen the day after Thanksgiving, with everyone also invited to Thanksgiving dinner. You face two options. Would you choose to go to the family gathering the day after Thanksgiving (and maybe to Thanksgiving dinner) if getting there required a $500 roundtrip plane ticket for plane flights that were 5 hours each way?

Option 1: Go to the thanksgiving gathering, which requires a $500 round trip plane ticket.

Option 2: Miss the thanksgiving gathering, but save the money.

Scenario 9: Education vs. Social Life

Suppose you have decided to leave Cornell, and are transferring to a new school. You have been accepted to two schools, and are deciding where to go. The first school is extremely selective and high quality, but is in a small town out in the country with a less active social scene. The second school is in a major city with a great social scene, but is slightly less renowned. Which would you choose?

Option 1: Highly selective school, isolated socially and geographically.

Option 2: Less selective school, socially active and in a major city.

Scenario 10: Interest vs. Career

Suppose you are considering two summer internships. One is extremely interesting and involves work you are passionate about, but does not advance your career. The other will likely be boring, but will help you get a job in the future. Which would you choose?

Option 1: Interesting internship which does not advance career.

Option 2: Boring internship which will help you get a job.

Scenario 11: Concert vs. Duty

Say you are driving by yourself to see your favorite musician in concert on their last day in town.

You are five minutes away, and the concert starts in ten minutes. On the drive, you witness a truck hit a parked car, causing roughly $500 in damages, and then drive away without leaving their information. You notice the truck's license plate, and you are the only witness. You face two options:

Option 1: Keep driving and get to the concert on time.

Option 2: Call the police, in which case you will have to wait around the parked car to give a testimony. This would take about half an hour. You would have trouble finding a seat and might miss the whole concert.

Scenario 12: Low Rent vs. Short Commute

(*Phrasing in study 1):* Say you are moving to a new town. You are trying to decide between two similar apartments which you could rent. The two apartments are identical in almost everything – including floor plan, amenities, neighborhood character, schools, safety, etc. However, they have different rents and are located at different distances from your work.

Option 1: An apartment which requires a 45-minute drive to work. The rent is about 20% of your monthly income.

Option 2: A similar apartment, with only a 10-minute drive. The rent is about 40% of your monthly income.

(*Phrasing in study 2):* Say you are moving to a new town. The new town is known for its terrible traffic jams, and driving there is widely considered to be unpleasant. You are trying to decide between two similar apartments which you could rent. The two apartments are identical in almost everything – including floor plan, amenities, neighborhood character, schools, safety, etc. However, they have different rents and are located at different distances from your work.

Option 1: An apartment which requires a 45-minute drive each way to work. The commute has heavy traffic almost the whole way. The rent is about 20% of your monthly income.

Option 2: A similar apartment which requires a 10-minute drive each way to work. The commute has heavy traffic almost the whole way. The rent is about 40% of your monthly income.

Scenario 13: Friends vs Income

Say you have been reassigned at your job, and will be moved to a new location. There are two offices where you could request to work. One office is in a city where many of your friends happen to live, and pays 20% less than your current salary. The other office is in a city where you don't know anyone, and pays 10% more than your current salary. Your job will be exactly the same at either office. You must decide between the following two options:

Option 1: Make 20% less than your current salary and move to the city with your friends.

Option 2: Make 10% more than your current salary and move to a city where you do not know anyone.