



## *Sociology Working Papers*

*Paper Number 2009-07*

**Activities, durations, and the empirical estimation of utility**

**J Gershuny**

***Department of Sociology***

*University of Oxford*

*Manor Road*

*Oxford OX1 3UQ*

[www.sociology.ox.ac.uk/swp.html](http://www.sociology.ox.ac.uk/swp.html)

# **Activities, durations, and the empirical estimation of utility**

J Gershuny

Centre for Time Use Research,  
Department of Sociology,  
University of Oxford  
September 2009

## **ABSTRACT**

This explores empirically the issue of diminishing marginal utility of specific consumption experiences with respect to time. It deploys a little-used pair of national sample surveys, from the UK and the USA in the mid-1980s, consisting of diaries in which respondents provide straightforward affect ratings for a day-long sequence of activities, as the empirical basis for an estimation of the differing levels of enjoyment (and hence utility) derived from various categories of activity by different sorts of people. It demonstrates some striking similarities between the two countries' mean relative ratings of the enjoyment of various daily activities, and produces direct estimates of the extent of diminishing marginal utility with respect to time in the case of some leisure activities, again with remarkably similar estimates in the two countries. It concludes that simple diary instruments recording enjoyment levels may be a suitable basis for, among other applications, aggregation into national accounts of utility.

# Activities, durations, and the empirical estimation of utility

## Introduction

The concept of utility comes to us from Bentham via Mill (1861, Chapter 2), who unambiguously identifies it as the extent of enjoyment of specific activities. Marshall derived the proposition of diminishing marginal utility (DMU) from a German source (von Wieser 1884). In what follows I pursue the idea that diary-based ratings of the enjoyment of activities correspond reasonably directly to utility. The special topic of the diminishing marginal utility of national income is the deeply contentious subject of an ongoing empirical controversy (see for example Easterlin 1974, Veenhoven, 1991, Easterlin 2005). But my concern here is with the more general underlying issue of diminishing marginal utility in individuals' consumption experiences. Though some consider DMU in this broader sense to be a *law* this is, I think, not well established empirically—because it is not generally speaking observed directly, but only inferred from prices on the basis of assumptions about the shapes of indifference/preference orderings. In what follows I suggest building an empirical bridge from modern survey evidence back to classical utility theory using time-diary measures.

Measurement of utility was not really an issue for Marshall. And discussion of utility *in itself* largely disappeared with the emergence of indifference/preference analysis. But something which—I shall argue—may be considered as an approach to the direct measurement of utility has nevertheless emerged over the last third of a century. Juster and Dow (1984) introduced the concept of the “process benefit”—the affective consequence (specifically enjoyment) accruing directly and immediately as a result of the individual's participation in an activity. Their methodology combined diary-based measurements of the duration of activities with separate questionnaire-derived enjoyment ratings of each of the activity categories. Robinson<sup>1</sup>, seeking a more direct measurement of affect, collected a 1985 national sample of self-completion diaries both to describe

---

<sup>1</sup> Initially as a member of Juster's group at the University of Michigan Institute for Social Research, but by this time Director of the Survey Research Centre at the University of Cleveland.

activities and to rate their enjoyability<sup>2</sup>—nearly 20 years before the somewhat similar methodological proposal for the measurement of instantaneous enjoyment contained in Kahneman *et al* 2004a. Robinson’s 1985 US national diary sample, together with a separate UK sample from 1986, forms the empirical basis for this paper.

An apparently quite independent parallel literature on “happiness” began to emerge in economics during the 1990s (eg Easterlin 1995, Oswald 1997, Layard 2005). Kahneman (1999) provides an elegant summarising discussion of “hedonic psychology”, which distinguishes between reflections on past affective states over extended periods (“How happy were you in March?”), described as “subjective happiness”, and the instantaneous affect more reliably recorded close to the moment of experience which he calls “objective happiness”.

In a long sequence of publications Czikszentmihalyi (eg 1990) developed the “experience sampling” methodology (initially using remote controlled beepers responding to radio signals, subsequently free standing personal recorders) registering affect related to specific activities, virtually simultaneously with those activities. This approach provides the prospect of an optimal quality of subjective measure for specific instances of Kahneman’s objective happiness. Though it suffers from the disadvantage that it measures affect discontinuously through the day it can still produce results somewhat resembling the DMU findings that follow here.

Kahneman *et al* (2004a) announced a new method—“day reconstruction”—for estimating instantaneous enjoyment, evidently unaware of Robinson’s (eg 1993) prior operationalization deploying the “yesterday” diary instrument (a format established well before the classic exposition in Szalai 1972), or of other prior approaches to diary-based affect measurement in the UK and elsewhere (Sullivan 1997, Gershuny and Halpin 1995).

---

<sup>2</sup> These form part of the American Harmonised Time Use Study, (<[www.timeuse.org/ahtus/](http://www.timeuse.org/ahtus/)> though the 1985 affect vectors, which have no equivalents in the 1965, 1975 or the 1990s and 2000s materials, are not yet downloadable from this source.

This paper deploys the “yesterday diary” or “day reconstruction” approach to measurement of the instantaneous enjoyment or utility of specific events or activities. My starting intuition is that this approach allows us to observe *directly* the phenomenon of diminishing marginal utility with respect to total recent (24 hour, 120 hour) durations in consumption (leisure) but not in work (production) activities.

## Theory

Utility is often considered in an atemporal manner, as related to volumes or quantities of consumption of a commodity. But all acts of consumption take time (Becker 1965). Utilities are produced in the minds of actors as the consequence of activities which generate affect in real time. Utility derives, that is to say, not from the meal itself, but from the act of eating it (though the benefit may vary with the quality of the food). The extent of activities is measured by their duration, and the affect may differ over elapsed time in the activity. So utility is properly to be regarded as a *rate of delivery of a benefit in time*—just *so much* enjoyment over a specified period of time—and marginal utility is the *change in this rate over successive increments of time*.

This temporal dimension of utility is implied by Kahneman (1999) who identified utility (as distinct from other “favourable states of being”) as linked to “the pleasantness or unpleasantness of particular moments” (p4). Utility generated by activities may therefore be estimated from diaries which are kept continuously throughout a given observation period, which list unbroken sequences of activities together with their durations and the associated enjoyment ratings of each activity. We can assert that:

$$1) \quad e_{ij} = a_{jk}X + b_j t_j + c_j t_j^2$$

where  $e_{ij}$  is the enjoyment rating of each element  $i$  in a continuous sequence of diary events which are classified into  $j$  exclusive categories of activity. The  $j$  types of activity

(such as “sleep”, “watching television” and so on) correspond to the “primary activities” in a time-diary survey.  $\mathbf{X}$  is a vector of  $\mathbf{k}$  (demographic and economic) control variables which may influence the enjoyment of activities,  $\mathbf{a}_{jk}$  is the set of coefficients relating this vector to each of the categories of activity, and  $\mathbf{t}_j$  is the *total* of all time devoted to activity-type  $\mathbf{j}$  by the diary respondent *over the entire observation period*. Note that since each event  $\mathbf{i}$  is classified into just one of the  $\mathbf{j}$  activity categories, Equation 1 may be estimated as a set of  $\mathbf{j}$  separate regressions. Hence, identifying enjoyment with utility, we may estimate marginal utility with respect to time for each of the  $\mathbf{j}$  categories as:

$$2) \quad \delta e_{ij} / \delta t_j = b_j + 2c_j t_j$$

This approach to modelling enjoyment of diary activities corresponds closely to the Kahneman *et al* (2004b), Krueger *et al* (2008) formulation of overall time-weighted enjoyment (or utility):

$$3) \quad \mathbf{u} = (\sum_i^i \sum_i^j \mathbf{d}_{ij} \cdot \mathbf{e}_{ij}) / \mathbf{T}$$

where  $\mathbf{d}_{ij}$  represents the duration of each of the diary events and  $\mathbf{T}$  is the total length of the observation period (ie the sum of the  $\mathbf{t}_j$  durations). Note that  $\mathbf{e}_{ij}$  and  $\mathbf{d}_{ij}$  vary with each diary event, whereas the  $\mathbf{t}_j$  coefficient varies only with the identity of the diary respondent. Equation 3 is not discussed extensively in what follows, but its use in the estimation of process benefit-type national accounts is an important application of the theory.

We may speculate as to the expected shape of the utility ( $\mathbf{e}_j$ ) and marginal utility ( $\delta \mathbf{e}_{ij} / \delta \mathbf{t}_j$ ) functions. Our expectations, at least in relation to consumption activities, are that:

- utility will have an inverted-U shape in relation to time, the initial instant in the activity (ie crossing the y-axis) with a positive value, and accordingly
- the marginal utility for each consumption activity will have a negative slope, and cross the x-axis (ie marginal utility equals zero) at some positive total of time,
- and the marginal-utility-equals-zero points will vary with the nature of the activity.

There is epidemiological evidence of a U-shaped relationship between sleep time and mortality, with significantly higher adjusted death rates below 5 hours and to the aggregated category sleeping above 8 hours per night (Ferrie et al 2008), while the overall average of sleep time per 24 hour period for people aged between 18 and 60 is around 530 minutes<sup>3</sup>. Our “sleep plus personal care” activity category includes around 30-60 minutes devoted to personal toilet and medical wants, so we would expect that, as a whole, the marginal utility of sleep will only fall to zero well beyond the daily physiological plus social requirements—with washing and other personal care, at least 600 minutes. By contrast, the zero-point for leisure at home will come somewhat before the normal expectation for onset of cabin fever for the housebound—given the 600 minutes of sleep plus personal care, together with a requirement (as in Jahoda 1982) for 4-6 hours of paid or unpaid work time to maintain a proper psychological balance, plus some minutes of leisure time outside the home—we might expect marginal utility for leisure at home to reach the zero point at around 400 minutes/day.

Work-type activities (paid and unpaid), may be distinguished from non-work by the “third person criterion”—they can be undertaken by some third party without losing the final utility derived from them (Walker and Gauger 1973, Hawrylyshyn 1977). It is not clear that reported enjoyment levels for these should show any significant time dependencies. On the basis that paid and unpaid work are *means* rather than *ends-in-themselves*, we might not expect to find any significant relation of enjoyment of work to work duration net of the effect of the relevant control variables. In particular education (since human capital relates positively to intrinsic job quality) and wage levels (since the knowledge of achieved earnings might be expected to generate some satisfaction) should

---

<sup>3</sup> <[www.statistics.gov.uk/timeuse/summary\\_results/activities\\_age\\_gender.asp#sleep](http://www.statistics.gov.uk/timeuse/summary_results/activities_age_gender.asp#sleep)>

be important explanators of variation in expressed levels of enjoyment of work, as well as family circumstances (since enjoyment might reflect satisfaction from meeting of gender- and family-role norms and expectations). If work is indeed purely instrumental, then any apparent temporal dependency should be accounted for by the characteristics that motivate it. The model of time allocation in Becker (1965) has the choice of the work/leisure balance purely on the basis of optimal *combinations* of income and consumption time, and these optima are a function of human capital and cultural characteristics fixed at the individual level. The implication is that individuals may be indifferent among various amounts of work having controlled for these factors.

I am not yet clear about the relationship *between* the utilities attached to work and to leisure—except that, again on the assumption of the purely instrumental nature of work activities, there may well *not* be a constant level of marginal utility across particular work and leisure activities.

## **Data**

What is the relevant observation or measurement period? It is likely that different activities have different periodicities of decay of affect (compare, for example, Wagner opera cycles and bars of chocolate). Nevertheless randomly sampled observation periods of various different lengths should reveal the same decay-behaviour for any particular activity. I will assume that for the broad categories of activity discussed in the following sections (“out of home leisure”, “watching television” and so on) the appropriate answer is provided by *The Tempest* prologue—“our little life is rounded by a sleep”—suggesting good operational reasons for the use of single-day diaries (though longer period estimates are discussed briefly below). Shakespeare’s methodological proposal implies that given a conventional 24 hour observation period (with, in accordance with the Harmonised European Time Use Survey guideline, a 4am start time), there will be virtually no left- or

right-censoring of any events except sleep, and the censorship pattern for sleep is a symmetrical one from which representative total durations can be estimated<sup>4</sup>.

The empirical work relies on two completely independent large scale surveys (independent in the sense that the survey designers were mutually unaware of each others' virtually simultaneous work), both completed just under 25 years ago, in the US and the UK. Both used the same underlying diary methodologies, both collected scaled information on the respondents' enjoyment of the primary diary activity. They collected similar total numbers of days of diary materials. But otherwise the designs were quite different.

(1) Robinson collected his US mail-back diary sample in 1985 using a slightly modified version of the Szalai (1972) methodology. This survey is a constituent part of American Heritage Time Use Study 1965-2007 dataset (<[www.timeuse.org/ahtus/](http://www.timeuse.org/ahtus/)>), but, unlike the other surveys included in the study, it asked respondents to record for each event, in addition to primary, secondary, location (but, exceptionally, not co-presence), "how much did you enjoy this activity?" on an 11 point, 0-10 positive scale. (I suspect that this scale may perhaps be too long—note the sparse population of the lower scores in Appendix A). It provides 31948 diary events with non-missing primary activity and enjoyment records from respondents aged 18-60 who are members of couples. The version used here has 90 distinct activity categories coded from open text responses (a more detailed activity classification is also available). Data are weighted to correct for non-response, and to adjust for the varying durations of events (and are rescaled to maintain the original sample N of cases), so as to produce a mean enjoyment score 6.99, with a standard deviation of 2.43.

(2) Ehrlich (1987) working on behalf of Unilever Research collected a 1986 face-to-face UK diary sample with primary and multiple secondary activities, "with whom?" together with, for each successive event, a "how much did you enjoy this?" on a 5 point 5-1

---

<sup>4</sup> ...unless, of course, "sleep" is used metaphorically by Shakespeare, to relate to the state of unbeing before birth and after death, in which case the observation is better applied to conventional event history analysis,

negative scale. (In this case I think the scale was too short: note again in Appendix A that 7 to 8 of the scale points in the Robinson data are well populated). The survey also collected additional diary affect data (“would you have liked more/less time?” and “how stressed?”) questions not used here. The study yields 40910 events with good primary and enjoyment data from members of couples aged 18-60. It records 70 distinct activity categories from a pre-coded list. To make comparisons easier within this paper, the UK scale ratings are transformed by subtracting them from 5.5 and doubling, to produce a 5-point, 1-9 positive scale, which has means and standard deviations very similar to the US sample. Data are again weighted to adjust for the varying durations of events and then rescaled to maintain the original sample N of cases, producing a mean enjoyment score of 6.86 with a standard deviation of 2.13.

### **Analysis strategy**

I use Ordinary Least Squares regression despite the ordinal nature of the scales (a procedure that may be justified by the regular intervals between categorical enjoyment levels revealed by ordinal logit models—see the Appendix to this paper). I analyse each of the two datasets as separate event files for each of the nine categories of time use (which in aggregate produces results equivalent to a “saturated model” with interactions between each activity category and all of the demographic/economic control variables). The case is the diary event, events are weighted proportionally to their duration (and also, in the US case to the selection probability of the diary day), weights being adjusted to give approximately the same total N of cases as in the original analysis sample. I attach each respondent’s daily *totals* of time in the relevant activities as appropriate to the nature of each event, and regress the relevant totals and totals-squared, together with controls, on the enjoyment ratings of each event. The overall saturated version of the model has a multiple correlation coefficient of .448 for UK, .385 for US.

Potential **problems** with this approach include the evident fact that, since most of the controls are, from a longer-term perspective, elective states, the selection processes might raise issues about the model specification—there are *some* things that might be done

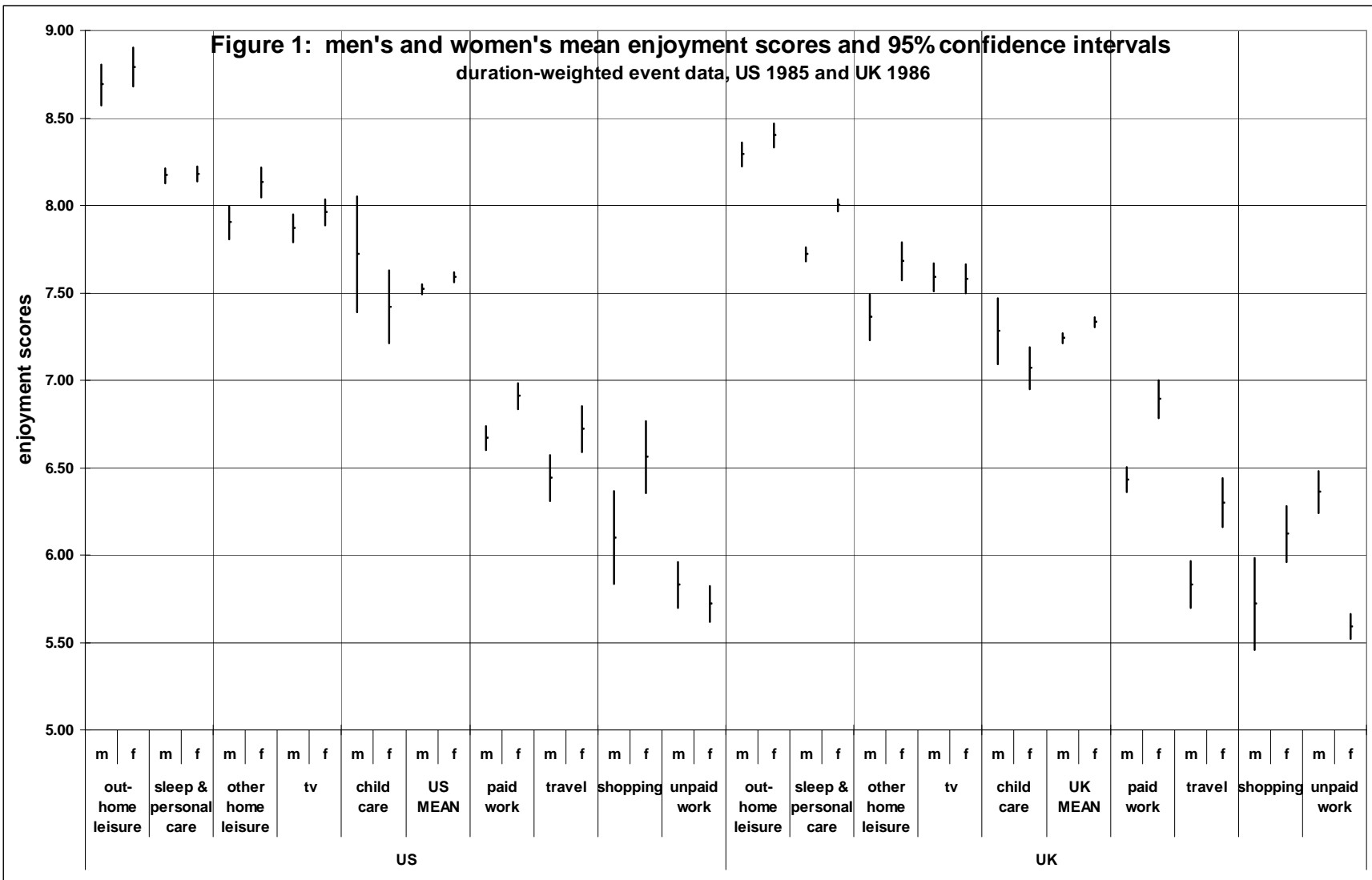
about this, particularly since the UK sample is in fact a five-day diary—but I choose for the moment to press forward with the simple demonstration that follows.

This analytic approach implies that the population from which the sample is drawn is not of persons but of events through the year. In any day recorded by a single diarist there may be more than one single occurrence-event of any given type of activity, so the 30-40K events in the two random samples surveys are slightly clustered within the days. And the UK diarists each kept their diaries for five days, implying some more substantial clustering of observations. Under these circumstances, the UK events—but, since the modal numbers of events per day in each activity category is unity, probably *not* the US events—are more appropriately modelled using OLS estimates clustered by diarist, which leaves the coefficients unchanged but may substantially reduce the significance estimates. Significance results based both on the simple and the robust indicators of significance are therefore reported in what follows: the main findings remain despite the application of the more stringent test.

## **Results**

Primary activities for each event-case are classified into nine general categories: *out-of-home leisure* including eating and drinking, cinema, visiting friends; *non-television leisure at home*, including reading and listening to music; *sleep and personal care* including washing and dressing and consumption of medical and similar personal services; *watching television* as a category in itself; *paid work* including related travel; *domestic and other unpaid work*, for the diarist's own household or volunteering for others; *shopping* including the consumption of retail services such as hairdressing; *child care* including play with children; and all *travel* not included elsewhere. The nine activity categories are exhaustive, so the total duration of time devoted to all of them sums to the 1440 minutes of the day. Figure 1 shows the mean enjoyment ratings and 95% confidence intervals, separately for men and women for these nine activities in the two countries. It lists the activities in descending order of the mean enjoyment scores of US

men—though as we see, US women, or either sex in the UK would have yielded pretty much the same ordering.



Out of home leisure is the most-enjoyed activity, unpaid domestic work the least-enjoyed, with child care and paid work in the middle of the enjoyment distribution. Not only are both sexes' enjoyment of particular categories rather similar in the US, but they also correspond quite closely to those in the UK. Note that the 0-9 scale for the UK was chosen rather arbitrarily, just in order to give integer values for each point on the five-point UK scale: if a 1-11 rescoring with non-integer intermediate values had been adopted, the US and UK patterns would have looked even more similar. Some national differences would surely be expected, reflecting the economic and cultural variations in the two populations. The fact that two entirely independent surveys, using separately developed designs and coding systems, should nevertheless produce such strikingly similar results, speaks strongly to the robustness of the common underlying methodology.

In each country the enjoyment rating of each category of activity is in general similar for men and women whereas the different activities have significantly dissimilar ratings. In particular, levels of enjoyment of out-of home leisure are clearly and distinctively higher than those of any other activity. Sleep and personal care at home are enjoyed substantially more than television, with other home leisure occupying an intermediate position between these two. Child care, which would qualify at least in part as work according to the third person criterion, lies close to the overall mean enjoyment level in both countries, with, in both cases, men having just above the mean and women just below the mean level—the difference reflecting in both cases women's higher proportional contribution to physical care, where men's contribution is more focussed on play<sup>5</sup>. Both sexes' enjoyment of the four remaining activities (paid work, travel, shopping and unpaid domestic work) all fall below the overall country mean enjoyment levels, with women in each country enjoying the first three of these significantly more than men. Men in the UK claim to enjoy unpaid housework much more than women—this being the only substantial difference between the US and the UK in these orderings.

---

<sup>5</sup> In the US, for example, 19.7% of the childcare time of women with children age <6 was interactive (play, reading etc.) while for men 34.2% of childcare fell into this category.

Table 1 sets out the OLS regression results for the nine activities for each of the countries. To simplify comparisons of the US (where clustering may not be appropriate), and the UK (where it is required), summaries of significance calculated from both simple and robust standard errors are presented simultaneously (since identically the same regression coefficients must emerge from both the simple and the clustered OLS procedures); so for example the formulation  $**(*)$  implies that the degree of certainty of difference between a regression coefficient and its default comparator value has  $p < .0005$  as judged from simple standard errors but only  $p < .005$  as judged from robust standard errors.

**Table 1 OLS regressions. Dependent variable: enjoyment ratings: answers to “how much did you enjoy current activity?”**

11 point (0-10) scale for US, 5 point (1-9) scale for UK, \* p<.05 \*\* p<.005 \*\*\*p<.0005 Parentheses where significance levels are achieved with simple but not robust standard errors.

	leisure out	Non-tv home leisure	sleep, personal	watching tv	paid work	domestic, other.unpaid	shopping	child care	travel
US (1985)									
<i>Multiple R</i>	0.242	0.121	0.061	0.154	0.157	0.086	0.222	0.258	0.140
<i>Mins in activity/100</i>	<b>0.308</b> **(*)	<b>0.296</b> **(*)	<b>0.253</b> (**)	-0.073	0.047	0.088	0.242	-0.083	-0.179
<i>Mins in acty sq/1000</i>	<b>-0.002</b>	<b>-0.003</b> *(*)	<b>-0.002</b> (**)	0.002 (*)	-0.001	-0.001	0.065	0.008	0.003
Woman	0.221 (*)	0.273 **(*)	0.029	0.092	0.266 *(**)	-0.096	0.273	-0.106	0.265 *
Age	0.014	-0.035	-0.013	0.009	-0.019	0.028	0.077	0.028	-0.115 *(**)
Age sq /100	0.010	0.357	0.226	-0.035	0.514 (*)	-0.320	-1.124	-0.387	1.555 *(**)
Employed fulltime	0.193	-0.073	-0.060	0.031	-0.070	0.163	-0.458	0.141	0.057
Has cores. partner	-0.228 (*)	0.206 (*)	0.081 (*)	0.033	-0.063	0.096	-0.330	0.566	0.051
Has child aged <5	0.326 **(*)	0.074	0.150 *(**)	0.228 *(*)	0.327 (***)	0.115	-0.179	0.025	0.089
Has child aged 5-15	-0.033	0.007	0.033	0.020	-0.047	-0.151 *	-0.136	-0.258 *	0.143 (*)
Complete sec'ry ed	0.267	-0.078	0.078	-0.276 *(*)	0.411 (***)	-0.106	-0.317	-0.367	0.507 (**)
Some tertiary educ	-0.150	-0.077	0.055	-0.595 ***	0.706 **(*)	-0.241	-0.735 (*)	-0.668	0.109
Log hourly wage	0.005	0.019 (*)	0.008 (*)	0.010	0.023 (**)	0.005	0.007	0.013	0.001
Constant	7.471 ***	8.169 ***	7.330 ***	7.873 ***	6.041 ***	5.231 ***	5.777 ***	6.997 ***	8.098 ***
UK (1986)									
<i>Multiple R</i>	0.167	0.201	0.193	0.214	0.240	0.280	0.287	0.231	0.256
<i>Mins in activity/100</i>	<b>0.248</b> *(**)	<b>0.347</b> **(*)	<b>0.494</b> *(**)	<b>0.483</b> ***	-0.075	-0.002	<b>1.164</b> ***	<b>0.467</b> *(*)	0.188
<i>Mins in acty sq/1000</i>	<b>-0.002</b> *(**)	<b>-0.004</b> ***	<b>-0.004</b> *(**)	<b>-0.006</b> *(**)	0.002 (*)	0.002 (*)	<b>-0.016</b> ***	<b>-0.010</b> **(*)	-0.001
Woman	0.167 (*)	0.535 *(**)	0.228 (***)	0.135	0.509 ***	-0.703 ***	0.440 (*)	-0.160	0.277 (*)
Age	0.072 (**)	-0.072	-0.044 (***)	-0.030	-0.061 (*)	-0.058 (*)	-0.134 (*)	-0.237 **(*)	-0.202 *(**)
Age sq /100	-0.774 (*)	1.005 (*)	0.524 (**)	0.500	1.013 (**)	1.063 (**)	1.819 (*)	3.389 **(*)	2.698 *(**)
Employed fulltime	0.023	0.408 (***)	-0.049	0.148	-0.167	0.254 (**)	0.134	-0.016	-0.374 (*)
Has cores. partner	-0.041	1.948 ***	1.010 *(**)	1.698 *(**)	1.804 (***)	-0.629 (*)	-0.698	0.525	-0.804
Has child aged <5	0.194 (**)	0.286 (*)	0.244 *(**)	0.361 *(**)	-0.174 (*)	0.001	0.424 *	0.217	0.332 (**)
Has child aged 5-15	0.065	-0.082	0.143 (***)	0.023	0.570 (***)	-0.190 (*)	0.097	0.215	0.539 **(*)
Complete sec'ry ed	0.029	0.012	-0.072 (*)	0.135	0.240 (**)	-0.153	0.191	-0.564 **(*)	-0.442 *(**)
Some tertiary educ	0.023	-0.212	-0.119 (**)	0.002	0.493 (***)	-0.564 **(*)	0.387 (*)	-0.478 *(*)	-0.466 *(*)
Log hourly wage	0.097	-0.004	-0.004	-0.010 (*)	0.051 *(**)	-0.015 (**)	-0.014	0.015 (*)	-0.011
Constant	6.172 ***	5.813 ***	6.076 ***	5.334 ***	4.679 *(**)	7.409 ***	7.211 ***	10.163 ***	10.042 ***

Women (controlling for all the other variables) seem generally to enjoy most activities more than men, with the exception of unpaid work and childcare (significant only for UK women's housework). The enjoyment of most activities has a generally U-shaped distribution with respect to age in the UK, enjoyment levels growing substantially in later life; the single exception, for no immediately apparent reason, seems to be out-of-home leisure. Full-time employment status has some significant effects in the UK, but not in the US: I suspect that this reflects weaknesses in the measurement of the UK hourly wage variable (estimated from a categorical labour income-range measure), which means that the positive correlation of full-time employment with income is reflected to some degree in these coefficients.

The presence of a co-resident partner has a substantial positive effect on the enjoyability of home leisure consumption, particularly in the UK—a result noted previously by Sullivan (1997) using this same data. Out of home leisure, contrarily, seems to suffer somewhat—perhaps as a result of the partner's relative absence? Having children, and particularly younger children, has a positive effect on enjoyment of leisure consumption. Hourly wage levels (logged) have some significant positive relation to the enjoyment of work. We might perhaps have expected the wage effects for work to be larger—but we would also predict some straightforward cognitive dissonance effects on the diarists' perceptions: respondents who are relatively well-paid for their paid work efforts, have less reason to represent their jobs to themselves as pleasanter than they really are. Education has the most substantial positive effect on the enjoyment of paid work, of a similar order in both countries: since the effect of pay rates are controlled for, this presumably reflects the relatively high levels of intrinsic satisfaction derived from the sorts of jobs undertaken by the better-educated.

The very first pair of coefficients, representing respectively the time spent in the activity over the entire observation period, and that time squared, provide the basis for the consideration of utility in its proper temporal context. Seven of the eight leisure-consumption coefficients present the inverted-U, diminishing marginal utility, results

foreshadowed in the brief theoretical discussion. The eighth pair, concerning the enjoyment of television watching in the USA, showing a much smaller scale of temporal dependency, not significant even under the less-demanding simple standard error criterion, may perhaps be explained, either by the relative ubiquity of the television switched on (even if not actually watched) in US households, or (from my chauvinistic UK perspective) the rather higher average standard of UK television programming!

The hypothesis that work-type activities should not exhibit this inverted-U, DMU result, is partially contradicted in the UK case. Shopping and childcare in the UK both show a substantial and strongly significant DMU results. My interpretation is that this is a symptom of the very imperfect nature of the strong means/ends dichotomy implicit in the “third party” criterion for the definition of work—a dichotomy which is important in the construction of traditional national accounts, and crucial to the estimation of national accounts extensions (to include production outside “the economy”) of the sort most recently championed in Stiglitz *et al* (2009 pp. 14, 35-38). This result is also strongly consistent with a quite separate time-diary-based finding (not discussed further here) that fully one-third of all diarists do not in fact consider such unpaid production activities to be work (Gershuny 2009 forthcoming).

Table 2 calculates (on the basis of equation 2 above) the turnover points for the inverted-U temporal utility curves for those cases where we have significant coefficients. We see that marginal utility for leisure out of home does not become negative until 10 to 14 hours per day—a level so much beyond the normal levels as to be considered an out-of-sample prediction. The marginal utility of non-television leisure at home reaches zero at 7 hours and 27 minutes in the US, and 7 hours and 21 minutes in the UK—remarkably similar, and close to the back-of-envelope estimate of the onset of cabin fever in the previous section. The marginal utility of sleep plus personal care reaches zero at 10 hours and 20 minutes in the UK, where the (less strongly significant) US coefficients imply a cross-over at 11 hours and 24 minutes—consistent with the expectations set out in a previous section.

**Table 2 Where marginal utility equals zero.....**

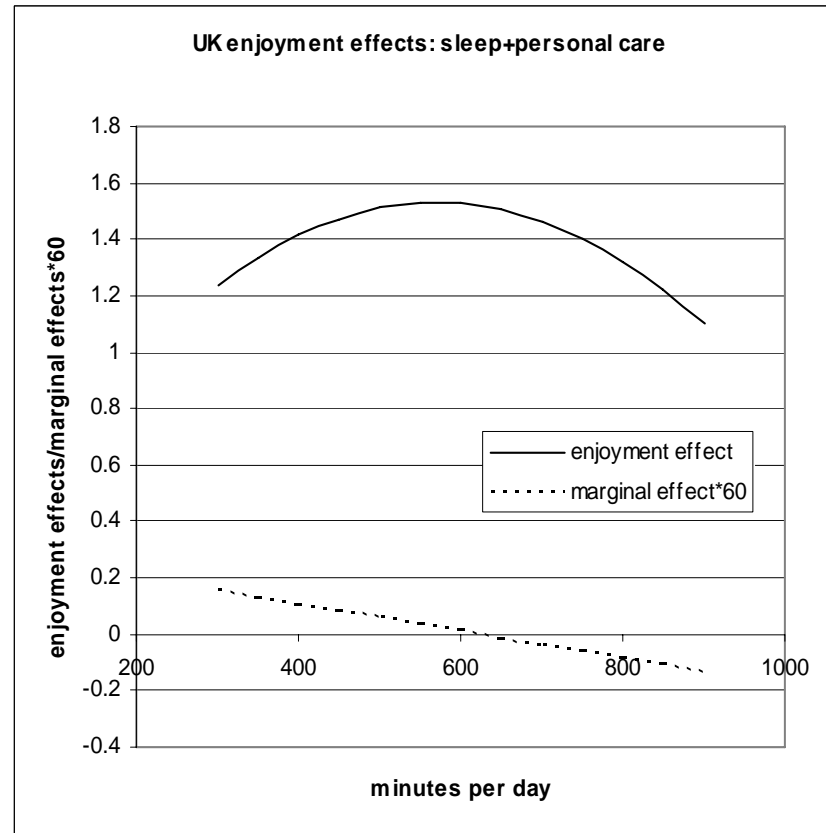
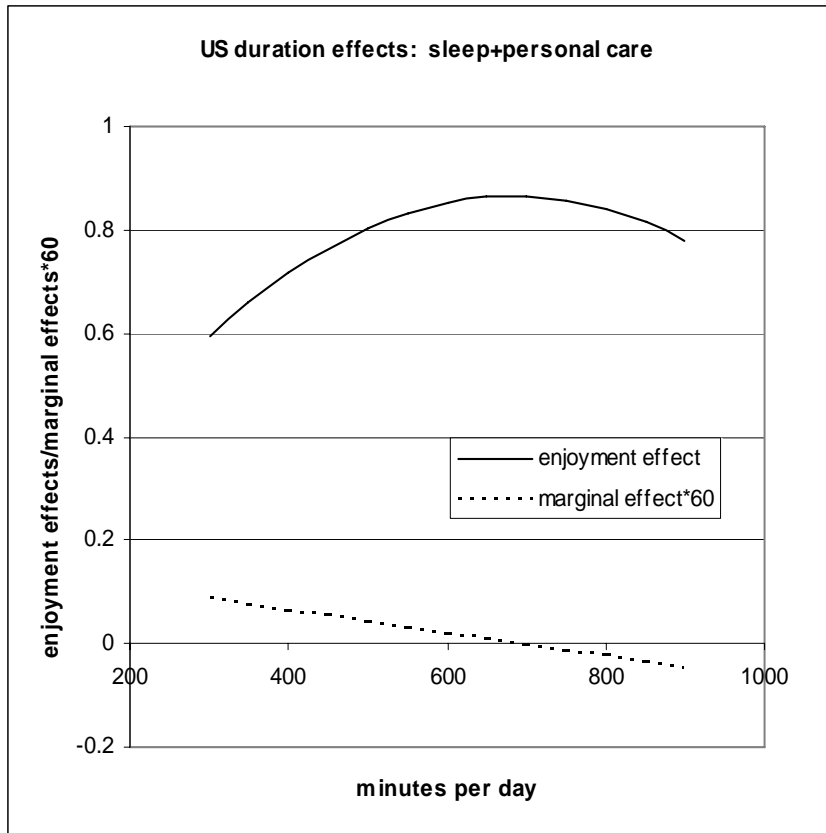
**bold** where temporal coefficients ( $t$ ,  $t^2$ ) are significant with robust standard errors  
 small and non-bold where both coefficients significant only with simple standard errors

	leisure out	non-tv home leisure	sleep, personal	tv	paid work	domestic, unpaid work	shopping	child care	travel
<b>US</b>	<b>840</b>	<b>447</b>	684	ns	ns	ns	ns	ns	ns
<b>UK</b>	<b>563</b>	<b>441</b>	<b>620</b>	<b>434</b>	ns	ns	<b>369</b>	<b>235</b>	ns

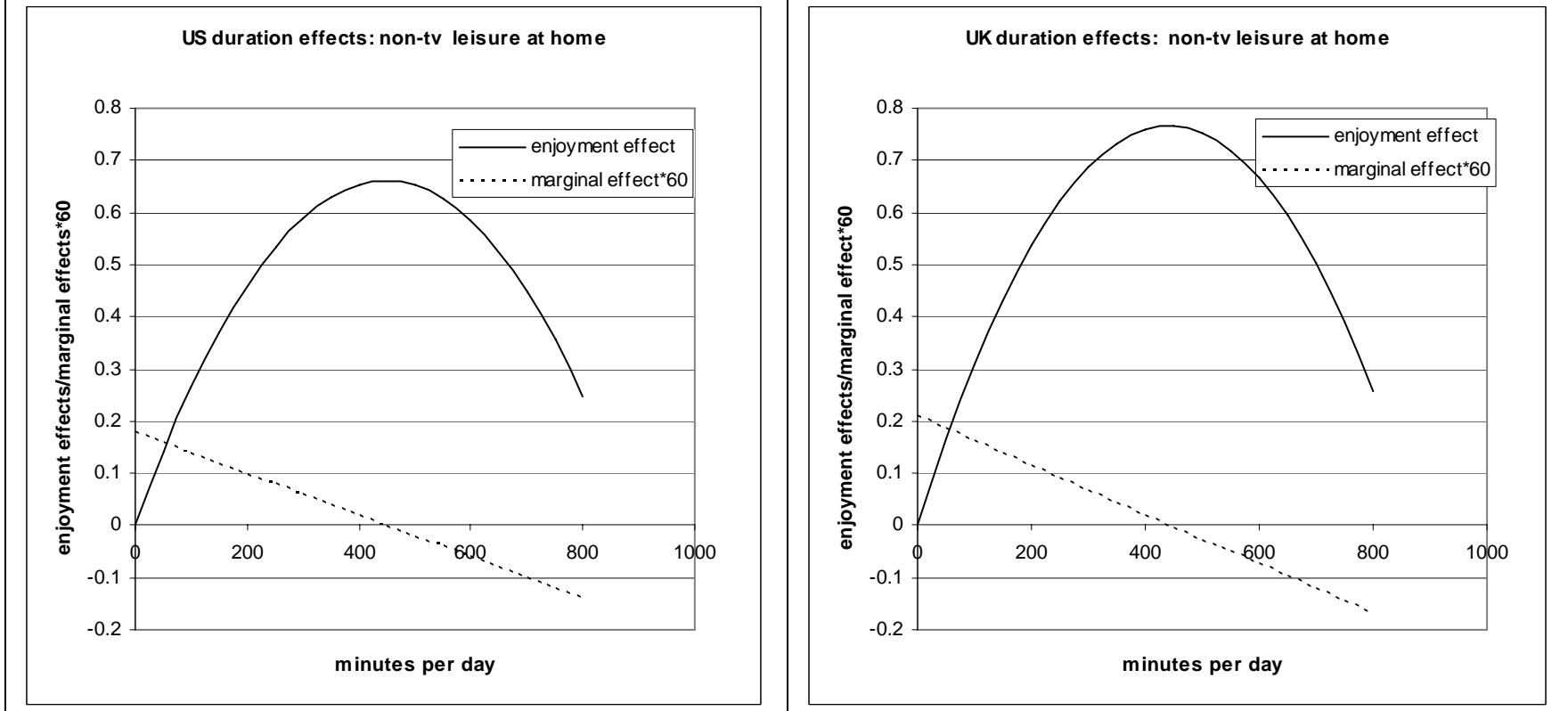
Remember that the different enjoyment scales used in the two countries have no impact whatsoever on the time-point at which their two marginal enjoyment curves cross the zero axis. The results are suggestive of a common affective reality across the two countries, a finding that is made all the more dramatically by comparison of the utility and marginal utility curves for the three cases for which we have reasonable estimates for both countries (Figures 2a to 2c).

Finally, consider the length of the observation period appropriate for estimating utilities. We have so far taken our cue from Shakespeare’s “little life” of the day. Should we expect to find different results from longer observation periods? This is in fact not the case. All the non-sleep daily activities appear uncensored within the 4am to 3.59am diary days. So when, instead of treating the 5 days of the UK survey separately, including the diarists’ identity in the clustering variable to cope with sampling design effects, we combine the five days to total time in the activity, we calculate  $t_j$  variables across the five diary days, and then estimate the same regressions, we arrive at a set of coefficients and standard errors virtually identical to those in the UK pane of Table 1.

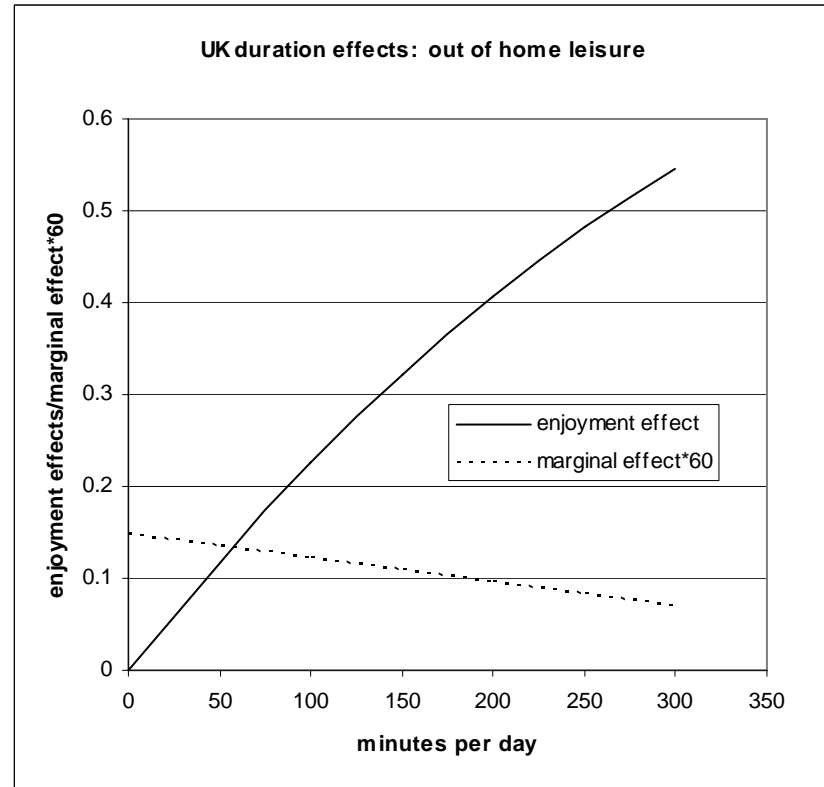
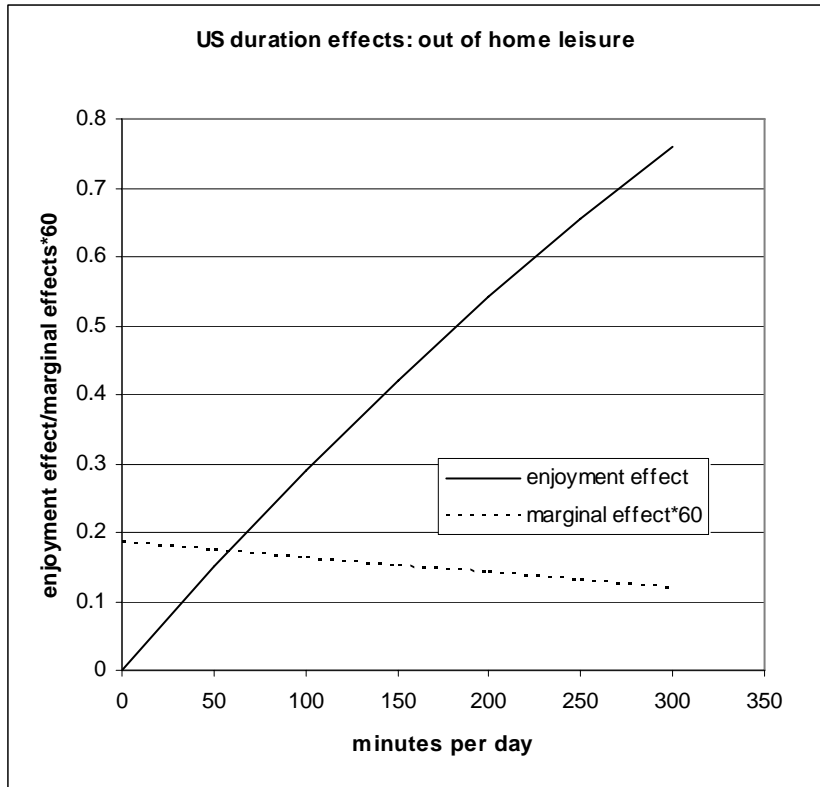
**Fig 2a Total and marginal utility: effect of time in activity on enjoyment of sleep and personal care )**



**Fig 2b Total and marginal utility: effect of time in activity on enjoyment of non-tv leisure at home**



**Fig 2c Total and marginal utility: effect of time in activity on enjoyment of out of home leisure**



## Conclusion

The instruments presented in this paper seem to work rather well considering their rough-and-ready nature. Enjoyment as implemented here is undoubtedly a somewhat noisy dependent variable, insofar as any diary respondent might deploy several different criteria for “enjoyability” across the various activities of the day, and certainly different diarists will have varying norms and modes of expression when asked to rate their enjoyment levels. But the signals that emerge seem nevertheless, to be quite sensible.

The descriptive statistics summarized in Figure 1 exhibit behaviour that is intuitively recognisable, (1) allowing us to distinguish among significantly differing levels of enjoyment produced by various types of activity, with expected similarities and dissimilarities between men and women, and (2) showing similar patterns of differentiation by sex and activity-type between two countries established by independently developed instruments. And the measurements of the differing marginal rates of enjoyment for various categories of activity (sleep, home and out of home leisure) with respect to time correspond remarkably well both to our theoretical expectations of marginal utility and to evidence from other sources (Diener *et al* 2010 for example, while assessing affect over an entire day rather than associated with particular activities, and mostly focussing on DMU of income, also suggest differential rates of marginal utility associated with different activities). Enjoyment is being *plausibly* measured by the diaries. And enjoyment, of specific activities at specific points in time, as Mill and Kahneman alike tell us, *is* utility.

New scientific measures of phenomena that are otherwise un- or only indirectly-observable always present problems of circularity. We believe such measures in part *because* they produce plausible results—insofar as they confirm our theoretical expectations. Nevertheless, because the enjoyment measures do seem (subject to further testing and experiment suggested below) to behave in the way we might expect from classical utility theory, we can properly conclude that they are appropriate to be used to

national accounts in the manner suggested by Stiglitz and colleagues<sup>6</sup>, perhaps in the form of the “national time accounts”—estimates of historical change in aggregate time-weighted enjoyment levels—proposed by Kahneman *et al* (2004b) and Krueger *et al* (2008) and constructed along the lines of Equation 3 above.

Such national utility measures potentially complement GNP in a useful way. We might, for example question whether GNP growth, *if* accompanied by increased time in less enjoyable jobs and by less enjoyable leisure, could properly be accounted as social progress. The particular benefit of the diary approach to measurement of utility is that it allows us to look directly at “links between various quality of life domains for each person” (Stiglitz *et al* p.15); historical change in overall utility estimated in this manner potentially reflects shifts in time devoted to both production and consumption. Note the much higher enjoyment levels of highly educated people’s paid work in Table 1; UK and US time in high-end jobs has approximately doubled, over the last 40 years, while time in paid work for the low educated has halved, partly because of the reversal of the status/leisure gradient (well educated people now work much longer hours) and partly because of the rapid disappearance of low-end jobs. Historical development may also bring change in the content of leisure—perhaps shifts towards the more enjoyable activities such as out-of-home leisure, or alternatively to less-satisfying pursuits such as watching television. Diary measures of enjoyment, combined with historical series of national sample surveys of time use will allow us to *decompose historical change in utility*, by types of activity and by types of person. They will allow us to examine (another Stiglitz recommendation) the *changing distribution of utility* across particular societies.

Both the UK and the US (and perhaps a dozen other developed societies) possess the sorts of historical series of time diary data necessary to implement this sort of broader socio-economic accounting. No countries other than the UK and the US (to my knowledge) *currently* possess the sorts of national-level diary enjoyment surveys discussed here, though France will now collect a substantial sample as an add-on to the 2009-12 round of the Harmonised European Time Use Study. But—this is part of the

---

<sup>6</sup> In particularly recommendations 5 and 10....

strength of the Krueger *et al* proposal—it may be that, to start with at least, we can make do with single measurements of enjoyment. It has after all been something close to a postulate of happiness research, going back at least to Durkheim (1893), that the enjoyment of particular sorts of activities has natural maximum and minimum levels which do not and indeed cannot vary over time, and that historical change in aggregate societal utilities reflects just the change of quantities and distributions of time among these activities and across the population.

So new diary enjoyment surveys conducted in, say, Netherlands, Denmark, Norway, Italy, Australia, Sweden or Finland might allow us, depending on the constancy of enjoyment postulate, to extend the analysis much wider. And the existence of the mid-1980s heritage of enjoyment diary studies in the UK and the US does of course mean that we can now, perhaps for the first time, test the postulate of historical constancy by repeating those studies. Opportunities for further development include the following:

- There are other areas of economic and social theory involving individual utility that provide potential test applications for these data. One likely candidate is to bargaining and other models of household decision-making. The 1986 UK dataset is collected from heterosexual *couples*, *both* of whom completed 5-day diaries. Can we use the evidence of utility differences within couples alongside human capital differentials to provide improved bargaining models? Or alternatively, perhaps we may find evidence of marital partners “doing gender” in the manner proposed by some sociologists.
- The evidence of DMU presented in this paper is strictly cross-sectional (it reflects mostly differences **between** people with different durations of participation in activities). To take the finding further we need to collect longitudinal data—which we might achieve by attaching diary instruments to existing longitudinal studies such as Understanding Society (the continuation of the British Household Panel Study) or the Panel Study of Income Dynamics in the US.
- There are also new measurement opportunities associated with GPS and mobile telephony, which would allow the collection of experience sampling-type data on a

scale sufficient to rival diary-based approaches but—because of their greater immediacy of recording—with higher levels of reliability.

And many others: this is clearly a fertile field of research.

## References

- Becker, Gary (1965) "A Theory of the Allocation of Time" **The Economic Journal**, Vol 75, No.299 (Sep.,1965), 493-517
- Csikszentmihalyi, Mihalyi (1990). **Flow: The Psychology of Optimal Experience**. New York : HarperCollins
- Diener Ed, Weiting Ng ,Will Tov (2010) "Balance in Life and Declining Marginal Utility of Diverse Resources" **Applied Research Quality of Life** (forthcoming)
- Durkheim, Emile (1893) [1933] **The Division of Labour in Society** (tr George Simpson) Illinois: The Free Press of Glencoe
- Easterlin, R. A. (1974). 'Does economic growth improve the human lot?' in (P. A. David and M. W. Reder, eds.), **Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz**. New York: Academic Press Inc.
- Easterlin, R. A. (1995). 'Will raising the incomes of all increase the happiness of all?' **Journal of Economic Behavior and Organization**, vol. 27, pp. 35-47.
- Easterlin, R. A.: 2005, 'Diminishing marginal utility of income? caveat emptor', **Social Indicators Research** 70, pp. 243–255.
- Ehrlich, Alma 1978 **Households Research Project, Research Report Vol 1**, London: Uniliver Research.
- Ferrie, Jane E., Martin J. Shipley, Francesco P. Cappuccio, Eric Brunner, Michelle A. Miller, Meena Kumari and Michael G. Marmot, "A Prospective Study of Change in Sleep Duration: Associations with Mortality in the Whitehall II Cohort" **Sleep**, Volume 30(12); December 1, 2007 pp 1659–1666
- Gershuny, Jonathan and Brendan Halpin. 1996. "Time Use, Quality of Life and Process Benefits", **In Pursuit of the Quality of Life**, Avner Offer (ed.), Oxford: Clarendon Press.
- Hawrylyshyn, Oli (1977) "Toward a Definition of Non-Market Activities", **Review of Income and Wealth** 23 (1):79-96
- Jahoda Marie 1982 **Employment and unemployment : a social-psychological analysis** Cambridge : Cambridge University Press
- Juster, Frank Thomas and Gregory Dow (1985) "Process Benefits and The Problem of Joint Production", in Juster, F. T., & Stafford, F. P. (eds.) (1985). **Time, goods, and well-being**. Ann Arbor: Institute for Social Research.

- Kahneman, Daniel, (1999) “Chapter 1: Objective Happiness” in Daniel Kahneman, Ed Diener, and Norbert Schwartz (eds.) **Well-Being: The Foundations of Hedonic Psychology** New York: Russell Sage Foundation.
- Kahneman Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz, Arthur A. Stone (2004a) “A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method” **Science** 306, 1776 –1780
- Kahneman Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz, Arthur A. Stone (2004b) “Towards National Well-Being Accounts” **The American Economic Review** Vol 92 No 2, 429-434.
- Krueger Alan B, Daniel Kahneman, David Schkade, Norbert Schwarz, Arthur A. Stone (2008) **National Time Accounting: The Currency of Life** Working Paper 523, Princeton University, Industrial Relations Section
- Layard, Richard. (2005). **Happiness: Lessons from a New Science.** London: Penguin.
- Mill, John S (1861) [1871] **Utilitarianism** Fourth edition London : Longmans, Green, Reader & Dyer
- Oswald, Andrew (1997) “Happiness and Economic Performance” **The Economic Journal** 107 (November) pp 1815-1831
- Robinson, J. P. (1993). “As we like it”. **American Demographics**, 15 (2), 44-48.
- Stiglitz, Joseph E, Amartya Sen, Jean-Paul Fitoussi (2009) **Report by the Commission on the Measurement of Economic Performance and Social Progress**  
[www.stiglitz-sen-fitoussi.fr](http://www.stiglitz-sen-fitoussi.fr)
- Sullivan Oriel, (1996) 'Time Co-ordination, the Domestic Division of Labour and Affective Relations', **Sociology** 30/1, 79-100
- Szalai, Alexander (ed) (1972) **The Use of Time** The Hague and Paris: Mouton Press.
- Veenhoven, R.: 1991, ‘Is happiness relative?’ **Social Indicators Research** 24, pp. 1–34.
- Walker, Katharine E and William H Gauger (1973) “Time and its dollar value in Household Work”, **Family Economics Review** Fall 1973 pp8-13
- Wieser, Friedrich von (1889) **Der natürliche Werth** [*Natural Value*] Bk I Ch V “Marginal Utility”

## APPENDIX:

### OLS modeling of “how much did you enjoy this activity?” evidence.

In both the UK and the US surveys the lower, and particularly the bottom enjoyment ratings seem relatively underpopulated, while in the much longer US ratings, the penultimate rating also seems sparsely populated. My preliminary conclusion is that a seven or nine point scale would be preferable for future exercises of this sort.

**Table A1 US 1985 distribution of ratings**

rating	Frequency	Percent
0.00	1319	1.89
1.00	789	1.13
2.00	1737	2.49
3.00	2669	3.82
4.00	3413	4.88
5.00	9359	13.39
6.00	7221	10.33
7.00	9106	13.03
8.00	12714	18.19
9.00	6852	9.81
10.00	14701	21.04
Total	69880	100.00

**Table A2 UK 1986 distribution of ratings**

rating	Frequency	Percent
1.00	931	1.73
3.00	3724	6.91
5.00	15948	29.58
7.00	11026	20.45
9.00	22288	41.34
Total	53917	100.00

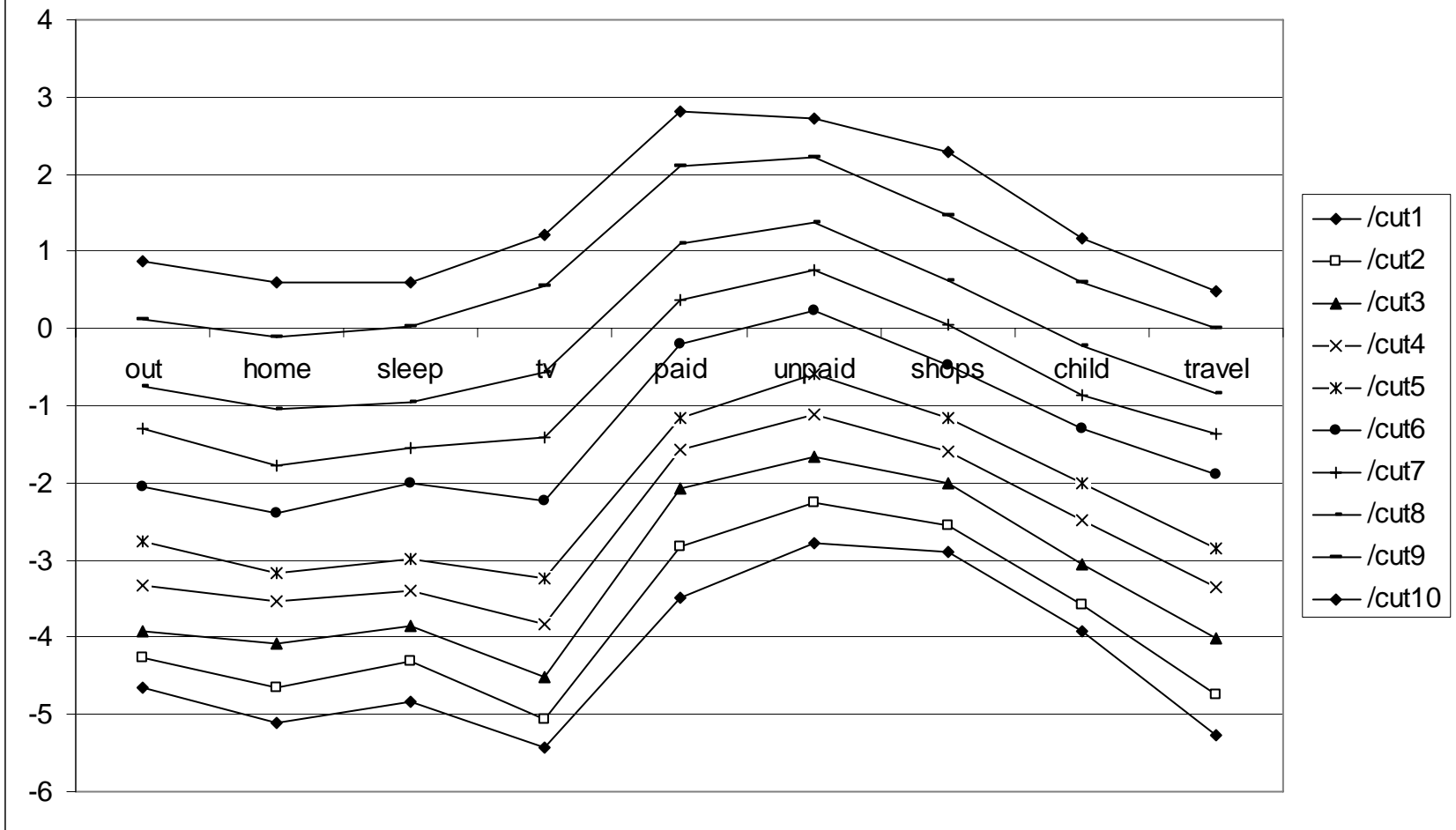
These are ordinal-level measurements, so it is formally inappropriate to use them as dependent variables in OLS regressions. But since—particularly in the case of the US data—the more formally appropriate ordered logistic regression approach seems to reveal through the rather evenly spaced cut points at least in the central range of the scale, implying that increments in the independent variables have a more or less constant effect on successive rating intervals (Table A3, Figure A1), it nevertheless seems reasonable to use the more easily interpretable coefficients of the OLS analysis deployed in the body of this paper. The UK ordered logit results (Table A4, Figure A2) do show, in relation to some of the activities, increasing-sized intervals through the shorter rating scale, perhaps implying that a small negative exponential rescaling might be appropriate, but this possibility is not pursued here.

**Table A3 OLS Ordered logit regression on activity enjoyment ratings.**

\* p<.05 \*\* p<.005 \*\*\*p<.0005, parentheses where significance levels are achieved with simple but not robust standard errors.

	leisure out	Non-tv home leisure	sleep, personal	watching tv	paid work	domestic, other.unpaid	shopping	child care	travel
US (1985)	0.027	0.005	0.001	0.008	0.007	0.002	0.017	0.021	0.005
<i>Multiple R</i>									
woman	0.245	0.310 **(*)	0.054	0.107	0.246 *(*)	-0.042	0.167	-0.098	0.209 *(*)
age	-0.005	-0.039	-0.017	0.017	-0.008	0.020	0.046	0.002	-0.094 *(*)
age sq /100	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.001 **(*)
employed fulltime	0.102	-0.135	-0.038	0.100	-0.065	0.128	-0.262	0.107	0.046
has cores. partner	-0.205	0.144	0.070	-0.025	-0.082	0.060	-0.233	0.500 *	0.046
has child aged <5	0.424 *	0.057	0.140 (*)	0.219	0.325	0.086	-0.091	-0.025	0.065
has child aged 5-15	-0.018	0.031	0.043	0.005	-0.009	-0.098 (*)	-0.109	-0.197 *	0.107 (*)
complete sec'y ed	0.179	-0.029	0.060	-0.313	0.332 (*)	-0.080	-0.502	-0.104	0.375 *(*)
some tertiary educ	-0.432	-0.082	0.037	-0.615	0.612 **(*)	-0.173	-0.817 *	-0.392	0.078
log wage	0.016	0.022 (*)	0.006	0.005	0.022	0.004	0.004	0.016	0.001
<i>mins in actvy/100</i>	0.309 *	0.256 **(*)	0.019	-0.184 (*)	0.024	-0.001	0.302 (*)	-0.197	-0.114
<i>mins in act sq/1000</i>	-0.001	-0.003 *	0.000	0.004 *	-0.001	0.001	-0.001	0.010 (*)	0.002
/cut1	-4.656	-5.108	-4.832	-5.439	-3.483	-2.791	-2.891	-3.931	-5.280
/cut2	-4.271	-4.659	-4.308	-5.064	-2.826	-2.247	-2.548	-3.573	-4.739
/cut3	-3.917	-4.091	-3.850	-4.516	-2.064	-1.665	-2.005	-3.044	-4.014
/cut4	-3.330	-3.532	-3.396	-3.833	-1.565	-1.114	-1.595	-2.492	-3.353
/cut5	-2.765	-3.161	-2.979	-3.241	-1.160	-0.578	-1.169	-2.010	-2.841
/cut6	-2.055	-2.396	-2.008	-2.232	-0.192	0.230	-0.466	-1.303	-1.898
/cut7	-1.304	-1.775	-1.551	-1.420	0.361	0.762	0.044	-0.855	-1.370
/cut8	-0.739	-1.056	-0.947	-0.556	1.091	1.382	0.621	-0.228	-0.834
/cut9	0.127	-0.113	0.024	0.542	2.098	2.219	1.473	0.601	0.005
/cut10	0.864	0.593	0.596	1.221	2.816	2.724	2.278	1.160	0.488

**Figure A1 US 1985:  
inferred intervals between successive rating points**



**Table A4 OLS Ordered logit regression on activity enjoyment ratings.**

\* p<.05 \*\* p<.005 \*\*\*p<.0005, parentheses where significance levels are achieved with simple but not robust standard errors.

	leisure out	Non-tv home leisure	sleep, personal	watchin g tv	paid work	domestic, other.unpaid	shopping	child care	travel
Uk (1986)	0.023	0.019	0.015	0.021	0.021	0.029	0.030	0.023	0.024
<i>Multiple R</i>	0.439 *	0.443 (***)	0.291 (**)	0.168	0.462 (**)	-0.598 (***)	0.380 (*)	-0.132	0.225
woman	0.120	-0.050	-0.065 (*)	-0.033	-0.043	-0.052 (*)	-0.122 (*)	-0.288 (***)	-0.181 (*)
age	age sq /100	-0.001	0.001 (*)	0.001 (*)	0.001	0.001 (**)	0.002 (*)	0.004 (***)	0.002 (**)
employed fulltime	0.137	0.317 (*)	-0.094	0.137	-0.130	0.229 (**)	0.171	0.032	-0.341 (*)
has cores. partner	-0.146	1.566 (***)	1.154 (***)	2.124 (**)	1.468 (**)	-0.554 (*)	-0.511	0.555 (*)	-0.728
has child aged <5	0.335 (*)	0.182	0.273 (***)	0.356 (*)	-0.159	0.011	0.337 (*)	0.150	0.298 *
has child aged 5-15	0.096	-0.091	0.185 (**)	0.024	0.457 (**)	-0.148 (*)	0.092	0.253 (*)	0.456 (***)
complete sec'ry ed	0.041	0.032	-0.070	0.141	0.230	-0.125	0.172	-0.553 (***)	-0.390 (*)
some tertiary educ	0.066	-0.301 (**)	-0.192 (*)	-0.005	0.377 (**)	-0.482 (***)	0.335 (*)	-0.482 (***)	-0.398 (**)
log wage	-0.002	-0.002	-0.004	-0.010	0.039 (*)	-0.013 (**)	-0.014	0.015 (*)	-0.010
<i>mins in actvy/100</i>	0.449 *	0.381 (***)	0.374	0.558 (***)	-0.052	-0.018	1.025 (***)	0.447 *	0.164
<i>mins in act sq/1000</i>	-0.004	-0.004 (***)	-0.003	-0.006 *	0.002	0.002	-0.014 (***)	-0.009 *	-0.001
/cut1	-1.388	-1.917	-3.612	-4.842	-1.818	-4.712	-4.590	-8.749	-7.533
/cut2	-0.425	-1.034	-2.844	-1.524	-0.285	-2.884	-2.617	-6.947	-5.438
/cut3	1.389	0.381	-0.210	1.365	1.234	-1.138	-1.069	-4.779	-3.416
/cut4	2.604	1.296	0.462	2.793	2.268	-0.049	0.309	-3.528	-2.530

**Fig A2 UK 1986:  
inferred intervals between successive rating points**

