



Centre for Time Use Research

<http://www.timeuse.org/>

An Implementation of an Experimental Time Use Methodology: an 'Australianised Counterfactual' Estimate of National Subjective Wellbeing in the USA

Technical Paper 2012-01

Centre for Time Use Research
University of Oxford, United Kingdom

Roger Patulny and Kimberly Fisher

Abstract

Kahneman and Krueger's landmark Princeton Affect and Time Survey (PATS, 2006), which popularised collection of emotional responses to daily activities previously pioneered by Juster and Stafford (1985) and Robinson and Godbey (1997), reveals the emotional context of American time schedules. While only a handful of countries have collected national sample time use and affect data, those surveys available offer opportunities to measure the degree to which people in these countries find various activities 'pleasant' or 'unpleasant'. We can use time and affect data to construct national well-being accounts linked to lived daily experiences. Such accounts in turn create the opportunity to test how populations might feel should they adopt alternative lifestyles encouraged by national government policies. This paper outlines the technical procedures undertaken in producing a 'counterfactual' estimate of time and affect in America. We do this by comparing time use patterns in Australia and the USA and simulating what would happen to wellbeing if the latter country adopted the time patterns of the former. Even though we expect differences in unpleasant time to emerge between the two countries, a hypothetical examination of what would happen if Americans were to shift to Australian-style time schedules is revealing for policy in both countries.

Introduction

This technical paper explains the process of developing a counterfactual analysis of the emotional well-being consequences for people in the United States were they to shift their lifestyles to live more like Australians. We have published this research separately in the *Australian Journal of Social Issues* (2013). This technical paper also sets out the policy case for adopting such a research strategy, and then provides more extensive detail of the process than appears in the published paper.

Aggregate emotional experiences (how happy, satisfied, stressed or bored national populations feel) reflect the general quality of life in societies (Seligman et al. 2005; Layard 2005; Diener and Seligman 2004). Common survey measures that ask about happiness or satisfaction with life *in general*, however, do not reveal much about the dynamics of emotions that play out in *particular social situations*. Kahneman and Krueger draw a distinction between life-domain/generalised vs. experienced/hedonic wellbeing. They find in experiments and in reviews of experimental research (Kahneman et. al. 1997, 2006, Krueger and Schkade 2007) that life-domain assessments are over-weighted by extreme and recent experiences, shifting moods, and question-order effects. These complications raise doubts in the reliability of life-domain measures over time. Kahneman and Krueger distinguish between measures that ask ‘How happy/satisfied are you in general?’ (as collected by Easterlin 1974, Heady and Wearing 1992, Oswald 1997, Helliwell 2003) and those that ask ‘How much time do you spend doing enjoyable activities?’ (as collected by Juster 1985, Robinson and Godbey 1997, Kahneman and Krueger 2006, Gershuny 2011), emphasising the importance of the latter measures.

Time-based measures of wellbeing capture activity that takes place in private settings, such as households, and often occur within gendered contexts. Recent surveys of time and activities have extended the generalised approach to measuring wellbeing to examine emotions in a variety of social and private contexts, and this has added considerable richness to our understanding of wellbeing (Kahneman and Krueger 2006). An impressive body of new evidence on emotional wellbeing is emerging from studies of emotions and time-use, which inquire into how people feel about what they do with their time and with whom they spend their days. Such research comes from a tradition of time-diary analysis going back many years (Michelson 1977, Csikszentmihalyi and Hunter 2003). The more sophisticated of these investigations have inquired into the enjoyment associated with types of activities in the context of whole day experiences (Juster and Stafford 1985), and during particular episodes (Gershuny and Halpin 1996; Robinson and Godbey 1997).

Such approaches build on the Experience Sampling Method (ESM) (Csikszentmihalyi and Larson 1992, Zuzanek 2012), which continues to be refined and tested in relation to other

time use data collection techniques (Soupermas et. al. 2005, Sonnenberg 2012). This method involves respondents wearing pagers – or more recently having apps downloaded to smart-phones – that message the participant at random times, whereupon respondents record what they are doing and how they feel at that moment. This method has the advantage of instantaneous collection of emotional responses, but the disadvantage of missing out on the sequences of activities that define the context in which emotions arise.

A promising alternative to the ESM is the Day Reconstruction Method (DRM) employed by Kahneman and Krueger (2006). After completing a conventional time diary, participants in DRM then report six emotional responses (happiness, sadness, pain, stress, tiredness and interest) to their activities during three randomly selected 15 minute intervals. The DRM contextualises emotions within activities and social interactions, examines a greater range of emotions (or affect-types) than just the happiness, satisfaction and stress measures used in the older time use and affect surveys which asked one or two emotional questions of all episodes in the diary (Gershuny 2012). The US Bureau of Labor Statistics has adapted the DRM, using the six affect dimensions developed by Kahneman and Krueger, but asking these six questions of three randomly chosen episodes rather than three fixed time intervals. In contrast with the 1 or 2 dimension questions asked of all episodes or the BLS approach, the Kahneman and Krueger approach does capture how emotions change at different points in long-duration episodes. Research has yet to determine which of these three approaches produces the more sophisticated micro-picture of the feelings, or which provides the best value for money policy-relevant utility.

The potential for cross-national emotional wellbeing analysis remains limited, largely as only four countries have regional or national sample time and affect surveys at the moment, Canada, France, the United Kingdom and the USA. Also, the Canadian regional surveys asked two questions of all episodes, while the British, one French and one USA survey asked one question of all episodes (using different scales), in contrast with the other USA national surveys and regional surveys in France which use the DRM. Though Gershuny concludes that the different size of the scales as in the UK and USA in the mid-1980s (2012) have little impact on the comparability of these surveys (similar affect pictures emerge for women and men in relation to the same domains of activity), no research yet compares the design affects of other differences between these surveys.

As no time diary and affect survey has been conducted (yet) in Australia, we cannot directly compare emotional well-being in this country and in the USA. Nevertheless, we can use the time-affect data generated by Kahneman and Krueger's work to engage in a 'counterfactual' picture of aggregated 'unpleasant' time to reveal how American's might feel were they to shift their daily activities to live more like Australians. The approach we outline

here is more modest than that developed by Gershuny (2012), who uses counterfactual analysis based on a combination of the surveys collected in both the USA and the UK in the mid-1980s to assess how behaviours in a range of countries with different welfare regimes might affect emotional well-being. This paper outlines the four stages undertaken to determine whether Americans would experience a net gain or loss in ‘pleasant’ time from shifting to Australian time patterns, and in what areas/activities they would gain and lose.

Stage 1: Selecting Relevant Data

We make use of three surveys, all collected in 2006: Kahneman and Krueger’s Princeton Affect and Time Survey, and the Multinational Time Use Study (MTUS) version of both the 2006 American Time Use Survey and the 2006 Australian Bureau of Statistics National Time Use Survey. We generate affect scores using the PATS, a nationally representative telephone survey of 11,905 respondents in the 48 contiguous states of the USA, adjusted with sample weights to resemble the national population distribution captured in that same year in the Current Population Survey. The survey team conducted phone interviews, and asked participants to describe activities from the previous 24-hour day. After completing the diary, participants then rated their perceptions of the six affect measures (happiness, sadness, pain, stress, tiredness and interest) on a 0-6 scale for the three randomly selected 15-minute episodes. Table 1 displays a sample PATS diary (where one written in a paper form).

Table 1 – Example of how a PATS diary would appear in paper form

Time	Main Activity	Where	Who with	Happy	Sad	Stressed	U-index
04:00-07:00	Sleep	Home	Partner				
07:00-07:30	Shower, dress	Home	Alone	4	2	2	0
07:30-08:00	Eat breakfast	Home	Partner				
08:00-09:00	Commute	Train	Strangers				
09:00-12:00	Work	Work	Colleagues				
12:00-13:00	Eat lunch	Work	Alone	5	1	1	0
13:00-17:00	Work	Work	Colleagues	3	2	4	1
17:00-18:30	Drink, socialise	Pub	Friends				
18:30-19:00	Cook	Home	Partner				
19:00-19:30	Eat dinner	Home	Partner				
19:30-21:00	Watch TV	Home	Partner				
21:00-21:30	Read	Home	Partner				
21:30-04:00	Sleep	Home	Partner				

The negative emotion rating exceeds the positive rating in **around 20%** of PATS episodes.

The Multinational Time Use Study offers an enhanced and harmonised version of the 2006 American Time Use Survey and 2006 Australian National Time Use Survey. In the MTUS episode file, each row case represents a change in main activity, secondary activity, location, mode of transport or who else was present. The harmonised main and secondary activity variables cover 69 categories that appear in a large majority of time use surveys. Episodes include the start and stopping time on the clock. The PATS file has a similar structure. All three surveys contain the same key demographic characteristics of respondents and their households. The distribution of the samples is shown below. The PATS sample is of episodes of activity with affect-information based activity (with 3 episodes per person) while the MTUS samples are of persons (see Table 2).

Table 2 – Sample characteristics in the PATS and MTUS

	PATS		American MTUS		Australian MTUS	
	N	%	N	%	N	%
Male	7,204	41%	5,140	42%	6,028	47%
Female	10,395	59%	7,060	58%	6,786	53%
Single, no children	5,755	33%	3,867	32%	2,713	21%
Single, has dependent children	1,422	8%	1,122	9%	457	4%
Married, no children	7,331	42%	2,635	22%	5,429	42%
Married, has children	3,023	17%	4,575	38%	4,215	33%
Aged 35 years or less	2,039	12%	2,948	24%	3,266	25%
Aged 34 to 64	8,099	46%	7,035	58%	7,235	56%
Aged 65 years or more	7,395	42%	2,217	18%	2,313	18%
Employed	6,644	41%	7,468	61%	7,157	56%
Student	585	4%	720	6%	1,360	11%
Retired	6,728	41%	2,003	16%	2,500	20%
Unemployed	2,301	14%	2,009	13%	1,797	12%
Not in labour force	1,191	7%	1,648	14%	4,417	34%
Less than secondary education	4,359	25%	3,409	28%	4,637	36%
Secondary education only	11,925	68%	7,143	59%	3,760	29%
More than secondary education	7,204	41%	5,140	42%	6,028	47%

There are a few key demographic differences between the PATS and MTUS surveys. A greater proportion of older people participated in the PATS survey than in the contemporaneous ATUS. This is a deliberate feature of the survey design, in that the PATS survey over-sampled older people. We retained the over-sample retiree diaries to increase the range of reactions to experiences for this age group. There also is a greater proportion of household with children in the MTUS sample. This is because the PATS sample did not include a specific question asking for the presence of dependent children in the household, forcing us to impute such a variable from data on household structure and co-presence

reported in the diaries. This estimated variable likely underestimates the number of diarists with dependent children in 2006. This should be considered in conducting analysis.

The PATS survey collected a full activity diary, similar in design to the American time use survey, though the affect question element relates to only three fifteen minute periods during that diary period. We do not consider potential emotional reactions to sleep, as the PATS survey excluded regular sleep, and only included periods where the diarist reported being awake when sampling the intervals for emotion data collection (though the PATS data do sample periods of insomnia and sleeplessness). Table 3 displays the breakdown of the samples of episodes associated with each study according to these activities.

Table 3 – Number of episodes of sample activities, amalgamated to 15 categories, PATS and MTUS versions of the USA and Australian surveys

	PATS		American MTUS		Australian MTUS	
	N	%	N	%	N	%
Sleep (including sleeplessness)	258	1.4	27,158	10.20%	29,706	8.20%
Personal care	1,808	10.1	40,571	15.20%	77,182	21.20%
Paid work and searching for work	2,018	11.3	15,678	5.90%	16,009	4.40%
Education	246	1.4	1,287	0.50%	1,546	0.40%
Unpaid domestic work	3,203	17.9	43,486	16.30%	83,398	22.90%
Child care	495	2.8	14,999	5.60%	18,195	5.00%
Adult	98	0.5	3,859	1.40%	927	0.30%
Pet care	158	0.9	2,765	1.00%	5,844	1.60%
Civic, voluntary, religious activity	321	1.8	3,385	1.30%	7,075	1.90%
Out of home free-time & leisure	624	3.5	5,182	1.90%	5,531	1.50%
Sports, Exercise, outdoors activity	976	5.5	4,735	1.80%	7,834	2.20%
In home free time & leisure	1,406	7.9	16,135	6.00%	9,249	2.50%
Media & computing (not TV)'	1,477	8.2	12,067	4.50%	24,591	6.80%
Watch TV	3,072	17.2	22,641	8.50%	27,718	7.60%
Travel	1,646	9.2	52,204	19.50%	48,565	13.40%

Stage 2: Estimate the Likelihood of an Episode Being ‘Unpleasant’

We next determined which episodes diarists rated as unpleasant, and examined the association of these episodes with demographic characteristics and with activity patterns in the PATS data. We start this process using the approach developed by Kahneman and Krueger (2006) to identify an episode as unpleasant: where the scores for any of the negative emotions outweighs the score for the positive emotions for that particular episode, we code the episode with the value 1, representing being unpleasant. Otherwise, we code the episode as 0, representing not being unpleasant. This is a conservative approach to measuring unpleasantness, and as a consequence produces fewer unhappy episodes than other potential strategies. Roughly 20 per cent of activities are coded as ‘unpleasant’.

We next use regression modelling to predict this unpleasantness attribute, as undertaken by Gershuny (2011), who used activity-level emotion information collected in time diary surveys in the USA in 1985 and in the UK in 1987 to make similar projections about potential implications for national time accounts (a measure of national well-being based on how happy the population feels across an average day measured by the mean utility of activity profiles). In that paper, Gershuny used non-contemporary (much older) utility data from full 24-hour diaries rather than the more detailed utility measures for three random activities that we make use of from the PATS survey. Nevertheless, it is worth noting the Gershuny found little difference in the utility profiles for British and US men and women - the gender differences are greater than the country differences (Gershuny 2011). We have no data at present to say how closely Australian and US utility profiles might map, though the Australian Bureau of Statistics has been investigating including a utility dimension on a future time use survey.

We used a Linear Probability Regression (LPM) model to predict unpleasantness of the basis of the demographic qualities of the person and their main activity during the episode, described by the following equation:

$$\mathbf{Pu}_e = \mathbf{a}_k\mathbf{X} + \mathbf{b}_j\mathbf{T}$$

Here, \mathbf{Pu}_i is the probability of each episode \mathbf{e} being rated as unpleasant. \mathbf{T} is a vector of \mathbf{j} activity variables (such as “sleep”, “watching television” and so on) that correspond to the “main activities” in the MTUS time-diary survey that may influence the unpleasantness of activities, while \mathbf{a}_i is the set of coefficients relating this vector to each episode. \mathbf{X} is a vector of \mathbf{k} (demographic) control variables that may influence the enjoyment of activities, and \mathbf{a}_i is the set of coefficients relating this vector to each to each episode.

The results suggest that the model provides a reasonable estimation of the affect and the unpleasantness associated with activities in America. The mean affect scores predicted from the model, shown in Table 4, closely match the actual means in the PATS data. Table 5 compares the application of these predicted scores with the imputation to the datasets drawn from the MTUS.

Table 4 details the betas obtained from this model, these represent the chance of the episode being unpleasant given a one unit change in the relevant independent variable, or (since all independents are binaries) given that it has a ‘yes’ (1) response.

Table 4 – Beta coefficients for LPM predicting unpleasantness, PATS

Demographic quality	Predicted % change in U-index
Male	-0.022
Single, no children	0.023
Single, has dependent children	0.006
Married, no children	-0.008
Aged 35 years or less	-0.004
Aged 65 years or more	-0.036
Employed	-0.056
Student	-0.045
Retired	-0.024
Less than secondary education	0.025
More than secondary education	-0.032
Episode activity	Predicted % change in U-index
Sleep and naps	0.098
Wash, dress, care for self	0.100
Meals at work or school	-0.010
Paid work-main job (not at home)	0.147
Paid work at home	0.122
Second or other job not at home	0.117
Unpaid work to generate household income	-0.035
Other time at workplace	-0.035
Look for work	0.361
Regular schooling, education	0.096
Episode Activity	Predicted % change in U-index
Homework	0.428
Leisure/other education or training	0.035
Food preparation, cooking	0.048
Set table, wash/put away dishes	0.129
Cleaning	0.095
Laundry, ironing, clothing repair	0.105
Maintain home/vehicle	0.118
Other domestic work	0.158
Purchase goods	0.039
Consume personal care services	-0.051
Consume other services	0.154
Physical, medical child care	0.048
Teach, help with homework	-0.009
Read to, talk or play with child	-0.039
Supervise, accompany, other child care	0.056
Adult care	0.165
Voluntary, civic, organisational act	0.041

Demographic quality	Predicted % change in U-index
Worship and religion	-0.023
General out-of-home leisure	0.070
Attend sporting event	-0.045
Cinema, theatre, opera, concert	-0.062
Other public event, venue	0.055
Party, social event, gambling	-0.058
General sport or exercise	-0.040
Walking	-0.090
Cycling	0.160
Other outside recreation	-0.067
Gardening/pick mushrooms	0.025
Walk dogs	-0.007
Receive or visit friends	0.007
Conversation (in person, phone)	0.038
Other in-home social, games	-0.024
General indoor leisure	0.168
Correspondence (not e-mail)	-0.116
Knit, crafts or hobbies	-0.002
Relax, think, do nothing	0.047
Reading	0.020
Listen to music etc	-0.092
Listen to radio	-0.018
Watch TV, video, DVD	0.068
E-mail, surf internet, computing	0.030
Travel to/from work	0.094
Education travel	0.039
Episode Activity	Predicted % change in U-index
Voluntary/civic/religious travel	-0.036
Child/adult care travel	0.130
Shopping, personal or household care travel	-0.017
Other travel	0.054
No recorded activity	0.102

Stage 3: Imputing the ‘Unpleasantness’ Likelihood to Australian and American Time-Diary Episodes

We next used the beta coefficients derived from the previous regression to impute the likelihood of episodes in the Australian and American MTUS data being ‘unpleasant’ (were American's to live the pattern reflected in that diary). In this way, we mapped affect information onto the MTUS version of the data for Australia and the USA in 2006, creating an Australian counterfactual dataset to compare with actual contemporary American data. We

additionally tested our imputation by mapping the scores to the whole diary from the PATS survey (see Table 5).

We opted to use the MTUS version of the 2006 American Time Use Survey rather than the PATS diary file for two reasons. The 2006 Australian Time Use Survey is a full national sample survey. The PATS survey concentrates on the contiguous states. The ATUS covers all states, and includes a larger sample, and hence better reflects the range of variation in behaviour across the whole USA, and thus is the more comparable survey to include with the Australian survey. Second, by imputing the emotion ratings to both surveys, we avoid the problem of comparing people's personal responses to their own diary with imputed data - we have a more level comparison field with two imputed datasets. While the ATUS collected PATS style affect data in a 2010 round, this affect data reflects a period after the global financial crisis, and hence may reflect difference in response to changing circumstances. Hence, we stick to the three surveys collected in the same year.

Table 5 – Comparison of selected predicted and actual affect scores in PATS

	Mean actual affect score, PATS	Mean predicted affect score, PATS	Mean predicted affect score, MTUS, USA	Mean predicted affect score, MTUS, AUS
All persons	0.180	0.177	0.177	0.190
Men	0.160	0.162	0.158	0.170
Women	0.190	0.1903	0.188	0.203

The results from the imputation (shown in Table 5) suggest that the model provides a reasonable estimation of the affect and the unpleasantness associated with activities in America. Table 5 shows that predicted scores imputed in the 2006 American MTUS data are very similar to actual and predicted scores from the 2006 PATS data. The mean episode u-index estimated for the American MTUS data is nearly identical to actual episode u-index in the PATS data: around 0.18. Similar scores were also obtained across all datasets on the basis of gender (0.16 for American men and 0.19 for American women).

Stage 4: Estimate Unpleasant Minutes Per Day

In our final step, we use the predicted affect scores in the MTUS data to find ‘unpleasant’ and ‘pleasant’ episodes (for Americans) in the time-use datasets, and then aggregate these episode-level scores to produce a sum total of ‘pleasant’ and ‘unpleasant’ minutes per day that can be compared across the two national lifestyle patterns. Given that the modelling produced estimates of the likelihood of episodes being pleasant or unpleasant

(rather than a definitive yes or no) we had to choose a cut-off for defining episodes as falling into one status or the other. We coded episodes in MTUS as ‘unpleasant’ if the predicted affect score for an episode was higher than the mean u-index score in the PATS data of 0.18, or ‘pleasant’ if equal to or lower than 0.18. After assigning an ‘unpleasant’ or ‘pleasant’ rating to each episode, we summed total ‘pleasant’ and ‘unpleasant’ minutes per day generated by each set of daily patterns. These scores in turn reveal that shifting Australian lifestyles would reduce unpleasant time in certain activity categories for Americans, such as paid work, but at the cost of losing more pleasant time in other activities, including socialising (described more fully in Patulny and Fisher 2013). Curiously, as men would lose out more than women shifting to Australian lifestyles, the gender gap in scores that favour men with more pleasant and less unpleasant time would decrease - making emotional well-being more balanced between the genders. The gender gap finding has a qualifier; middle class American women with secondary-school (but not tertiary) education might feel happier leading an Australian lifestyle (Patulny and Fisher 2013).

Discussion

The context of the experience of emotions associated with everyday activities contributes to emotional wellbeing. Comparing the emotion profile that Americans lived in 2006 with the profile they might have experienced living like Australians give us some sense of whether policies that change behaviours might leave the population better - or worse - off. As this approach remains an experimental technique, we conclude by highlighting areas for future refinement.

1. *Choice of variables used in running models* – the number and choice of variables used in predicting affect scores in PATS to then map into MTUS was limited. More independent variables might enhance model efficiency.
2. *Coding pleasant and unpleasant time* – the choice of coding MTUS episodes as pleasant or unpleasant if they lie below or above the PATS mean was the most straightforward method of deriving pleasant and unpleasant time. We did not use the median given the binary nature of the original affect variable in the PATS data. However, there were alternatives to using the total mean from all episodes in the PATS data, such as for example calculating separate means and pleasant/unpleasant episodes for men and women. More fine-grained approaches are possible. This paper presents a straightforward and parsimonious basic approach.
3. *Lack of internationally comparative emotions data* – Thus far, only France, the United Kingdom and the USA have national sample affect data. As more countries

collect this data, future research can explore the potential influence of cultural differences on the experience of activities.

4. *Lack of time series data* – At the time of writing, France had one national time and affect sample, the UK had two, and the USA had three, each of which captured affect differently. The 1985 USA survey collected one emotion for all episodes. The 2006 survey collected six emotions for three 15-minute intervals, while the 2010 ATUS collected the six emotions for three episodes. Future work will need to determine which method produces the more robust data, allowing the collection of time series data in the future.
5. *Choice of Emotions* – Future research additionally must explore which emotions best capture national utility. We suggest that there is a need for a ‘audit’ of emotions in the context of gathering diaries about a daily time-based activities.

Emotions matter in wellbeing research. Studies such as those undertaken by Kahneman and Krueger demonstrate the limitations in conventional wellbeing research, and highlight a need for more sophisticated and contextually-based measures of subjective wellbeing. Time and affect diaries measure wellbeing in a range of contexts in terms we understand; our daily activities.

References

- Csikszentmihalyi, M. and Hunter, J. (2003) "Happiness in everyday life: The uses of experience sampling" *Journal of Happiness Studies* 4(2), 185-199.
- Csikszentmihalyi, M. and Larson, R. (1992) 'Validity and reliability of the experience sampling method' In M. W. de Vries and M. Csikszentmihalyi (Eds.) *The experience of psychopathology: Investigating mental disorders in their natural settings* Cambridge, UK: Cambridge University Press.
- Diener, E., and Seligman, M.E.P. (2004) 'Beyond money: Toward an economy of well-being'. *Psychological Science in the Public Interest*, 5, 1-31.
- Easterlin, R.A. (1974) 'Does economic growth improve the human lot? Some empirical evidence', in A.D. Paus & W.R. Melvind (Eds.), *Nations and households in economic growth: Essays in honor of Moses Abramovitz*, New York, Academic Press, Inc.
- Gershuny, J.I. (2000) *Changing times: Work and leisure in postindustrial society*. Oxford, UK, Oxford University Press.
- Gershuny, J.I. (2012) 'National utility: measuring the enjoyment of activities' *European Sociological Review*, On-line first, DOI 10.1093/esr/jcs077.
- Gershuny, J.I. (2011) *Time-use surveys and the measurement of national well-being*. Swansea, UK: Office for National Statistics.

- Gershuny, J.I., and Halpin, B. (1996) 'Time use, quality of life, and process benefits', In A. Offer (Ed.), *In pursuit of the quality of life* (pp. 188-210). Oxford, UK: Clarendon Press, Oxford University Press.
- Headly, B., and Wearing, A. (1992) *Understanding happiness: A theory of subjective well-being*. Longman Cheshire, Melbourne
- Helliwell, J.F. (2003) How's life? Combining individual and national variables to explain subjective well-being. *Economic Modelling*, 20(2), 331-360.
- Juster, F. T. (1985) 'Preferences for work and leisure'. In F. T. Juster & F. P. Stafford (Eds.), *Time, goods, and well-being* (pp. 333–351). Ann Arbor: Institute for Social Research, University of Michigan.
- Juster, F. T. and Stafford, F. P. (1985) *Time, goods, and well-being*. Ann Arbor: Institute for Social Research, University of Michigan.
- Kahneman, D. and Krueger A.B. (2006) 'Developments in the measurement of subjective well-being', *The Journal of Economic Perspectives*, 20, 3-24.
- Kroll, C. and Pokutta, S. (2012) 'Just a perfect pay? Developing a happiness optimised day schedule', *Journal of Economic Psychology*, On-line first, DOI 10.1016/j.joep.2012.09.015.
- Krueger, A. and Schkade, D.A. (2007) 'The reliability of subjective well-being measures'. *Journal of Public Economics*, Elsevier, 92 (8-9), 1833-1845.
- Kahneman, D., Wakker, P., and Sarin, R. (1997) 'Back to Bentham? Explorations of experienced utility'. *Quarterly Journal of Economics*, 112(2), 375–405.
- Layard, P.R.G. (2005) *Happiness: Lessons from a new science*. New York, Penguin Press.
- Michelson, William. (1977) *Environmental choice, human behavior, and residential satisfaction* Oxford, UK: Oxford University Press.
- Multinational Time Use Study, Versions World 5.5.3, 5.80 and 6.0 (released October 2011). Created by Jonathan Gershuny and Kimberly Fisher, with Evrim Altintas, Alyssa Borkosky, Anita Bortnik, Donna Dosman, Cara Fedick, Tyler Frederick, Anne H. Gauthier, Sally Jones, Jiweon Jun, Aaron Lai, Qianhan Lin, Tingting Lu, Fiona Lui, Leslie MacRae, Berenice Monna, José Ignacio Giménez Nadal, Monica Pauls, Cori Pawlak, Andrew Shipley, Cecilia Tinonin, Nuno Torres, Charlemagne Victorino, and Oiching Yeung. Centre for Time Use Research, University of Oxford, United Kingdom.
- Oswald, A. (1997) 'Happiness and economic performance', *Economic Journal*, 107, 1815-1831.
- Patulny, R., and Fisher, K. (2013) 'Advancing wellbeing research: would Americans be happier if they lived like Australians?' *Australian Journal of Social Issues*, forthcoming.
- Robinson, J., and Godbey, G. (1997) *Time for life: the surprising ways Americans use their time*, University Park, PA: Pennsylvania State University Press, second edition.
- Seligman, M.E.P., Steen, T.A., Park, N., and Peterson, C. (2005) 'Positive psychology progress: Empirical validation of interventions'. *American Psychologist*, 60(5), 410-421.

- Sonnenberg, B., Riediger, M., Wrzus, C. and Wagner, G. G. (2012) 'Measuring time use in surveys – concordance of survey and experience sampling measures' *Social Science Research*, On-line first 10.1016/j.ssresearch.2012.03.013.
- Soupourmas, F., Ironmonger, D., Brown, P. and Warner-Smith, P. (2005) 'Testing the practicality of a personal digital assistant questionnaire versus a beeper and booklet questionnaire in a random-time experience sampling method context' *Annals of Leisure Research*, 8(2/3), 142-152.
- Tetlock, P. E. and Belkin, A. (Eds.) (1996) *Counterfactual thought experiments in world politics: Logical, methodological, and psychological perspective*. Princeton, NJ: Princeton University Press.
- Zuzanek, J. (2012) 'Does being well-off make us happier? Problems of measurement' *Journal of Happiness Studies* on-line first, DOI 10.1007/s10902-012-9356-0.