

## Reply to Dan Turton

Helen Johns and Paul Ormerod [Volterra Consulting, UK]

Copyright: Helen Johns and Paul Ormerod, 2009

Our original article <http://www.paecon.net/PAEReview/issue46/JohnsOrmerod46.pdf> in *Real-World Economics Review* has been downloaded well over 10,000 times. The editor of the journal invited a range of leading happiness researchers to respond to our critique of time-series data on average happiness. The only response, however, has been from Dan Turton, whose contribution is the focus of this short note.

We should make it clear beyond doubt that we are not criticising the entire corpus of happiness economics research. Far from it. We regard the careful, micro-econometric panel data analysis carried out under the banner of happiness research as a part of valuable attempts in behavioural economics to construct better models of boundedly rational agent behaviour.

Indeed, in our original monograph on the topic (Johns and Ormerod, 2007) we devoted the opening chapter to this aspect of happiness research and economic theory, concluding that its findings fit in with a wide range of evidence in the more general field of experimental and behavioural economics that economics needs different postulates on individual behaviour.

The focus of our critique is time series data on average happiness at the national level, and specifically in developed countries. The notorious chart which plots GDP rising over time and average happiness remaining flat has been enormously persuasive, especially in policy circles. Carrying out simple correlations of average happiness with other variables makes clear that the simple lack of correlation between GDP and happiness over time has no meaning.

Turton agrees with this, pointing out quite correctly that multiple regression analysis is needed whenever there may be more than one hypothesised causal factor. But the community of happiness economists is largely silent when it comes to making clear that no credibility can be accorded to the simple GDP/happiness time series plot. Why? We can only pose the question and speculate on the answer.

However, our critique of time series data on average happiness goes much further than this.

Turton's main argument is that the basic premise that happiness levels remain flat in Western societies is incorrect. He cites evidence from Inglehart et al. (2008). But to use this as a counter-argument to our paper is disingenuous. The claim we make is simply replicating those made by some of the very leading researchers on happiness economics, recent examples being Easterlin (2005) and Kahneman and Krueger, (2006).

Clearly there is a difference of opinion within happiness research itself. In Britain, the idea that happiness levels have not increased over several decades has been presented as established fact and used as the basis for influential policy arguments (Layard, 2005) The idea that this flatness is due to the inability of economic growth to generate happiness in already affluent societies has rooted itself in public debate. If some happiness researchers

are now claiming, as the data presented in Inglehart (2008) appears to show, that in fact happiness has been increasing for some time in the USA and most European countries and this simply was not spotted before, this hardly stimulates confidence in the discipline or the quality of the data.

Further, if the average happiness scores presented in the Annexes of Inglehart et al. (2008) are representative and definitive, one wonders why many of the most prominent names in happiness research have been labouring under this misapprehension for so long. Far from clinching the argument, we suggest that this shows that totally different interpretations of happiness evidence have emerged within the discipline, and that this raises serious doubts that the analysis of happiness evidence is reliable and robust enough for policy purposes.

Turton cites evidence from Inglehart et al. (2008) that Lithuania, Mexico, Slovenia and Great Britain have all experienced 10% increases in average happiness over relatively short timescales on a four-point scale. We suggest that to draw such definitive statements from selected pairs of data points at face value without estimating sampling error is poor scientific practice and lacks statistical rigour. In addition, many of the arguments Turton derives from Inglehart et al. (2008), as well as studies such as Alesina et al. (2004), are based on microeconomic analysis. As our arguments relate solely to national-level average happiness time series and not to microeconomic studies, a great deal of Turton's arguments are at total cross-purposes to our own.

Our paper by no means claims to be a comprehensive review of happiness time series sampling error; Turton correctly states that it examines only one three-point time series. However, if it is the case that the estimation of sampling error has not been attempted as a matter of routine during several decades' collation and analysis of nation-level happiness data, we consider this to be the greater omission, . We agree that sampling error *should* be estimated for other time series and scales and would encourage happiness researchers to do so.

We do not recognise Turton's digest of our interpretation of sampling error and confidence intervals. Our paper is not intended as a formal test of the hypothesis that all points in the US happiness 3-point time series are not significantly different from the time series mean, as Turton implies. Nor do we claim that all variation is due to sampling error; that there is no possibility of a trend showing in an aggregate happiness time series (a possibility we explicitly allow for); or make the rather bizarre suggestion that only some data points are affected by sampling error.

For clarity, we argue that over short time periods true variation in average happiness time series (at least on the 3-point scale examined) is of comparable magnitude to sampling error and that this suggests the indicator is unresponsive. Physical scientists question whether true signals in their measurements are at risk of being drowned out by statistical noise, and so should social scientists.

It is self-evident that where random samples are taken, every data point is a composite of a real value and some sampling error. If the error of an indicator is routinely of comparable or greater magnitude to real changes in it – even if it sluggishly lurches a little way beyond a particular confidence interval from time to time - then the ability of the indicator to convey useful information is thrown into doubt. If, however, it routinely displays movement

which far exceeds statistical noise, it is far likelier that it is telling us something about underlying processes.

Turton's detailing of which specific data points are within or without a particular confidence interval is therefore not especially key, because the confidence interval used in our analysis is *approximate*, a best guess. We don't actually *know* the distribution of people in each happiness category in the underlying population for any given year, we only have the samples to guide us.

As it happens, the application of random matrix theory to time-series happiness data suggest that the data is dominated by noise rather than signal<sup>1</sup>. The eigenvalues of the correlation matrix of an appropriate delay matrix formed from UK time series data fall entirely within the theoretical range of the eigenvalues of a purely random matrix. For US data, the leading eigenvalue lies slightly outside this range. These results suggest that the UK data is indistinguishable from noise, whilst the US data contains a small amount of true information but is dominated by noise. These results are available on request.

We further argue that "true variation" is not equivalent to "useful information". We do not doubt that even a three-point scale could register increased happiness in Zimbabwe were daily life there to dramatically improve. But is this useful information?

Just as one might question why it is necessary to prove that bereavement and poor health make people less happy, one might ask what value is added for policy-makers in quantitatively demonstrating that people in a prosperous, free society are happier than those in a tyrannous, disease-ridden nation with no functioning economy.

Who exactly are the political decision-makers whose views on the benefits of democracy and economic prosperity will pivot on this proof? What extraordinarily weak faculties of qualitative reasoning they would need to have. In so many areas of public policy, it is not diagnosis which is the challenge but prescription; how to put decent leaders, institutions and policies in place to achieve a desired outcome.

Inhabitants of any country know that the reality is far more complex than a single quantitative score can convey. Turton states that happiness has been increasing in the UK; but there is a significant strand of British public opinion which considers British society to have grown more aggressive, ill-at-ease and socially fragmented over the period cited. Prominent experts have argued over recent years that Britons are suffering such an epidemic of mental illness that a drastic rethink of our cultural values and/or public policy is required (James, 2007; Layard, 2005), the latter making extensive use of happiness evidence to make his case.

Why are these arguments being made if Britons are demonstrably getting happier? What is the British public to make of the suggestion that happiness experts were just misinterpreting the data and that happiness is, despite everything previously said, actually increasing? They could be forgiven for concluding that expert opinion should be taken with more than a pinch of salt.

---

<sup>1</sup> An application of this, with a description of the technique, is given in Ormerod and Mounfield (2000)

If happiness data is of high quality and robust then it should be as clear as crystal whether happiness in Western societies is increasing or not. This is a basic question on which happiness researchers have drawn totally different conclusions.

Social scientists should not expect political decision-makers to base policies on indicators which give such conflicting signals. An omission to routinely estimate sampling error in average happiness time series is a further interpretative flaw which raises questions over whether national-level average happiness scores can be taken at face value.

## References

Alesina, A., Di Tella, R. and MacCulloch, R. (2004), 'Inequality and happiness: Are Europeans and Americans different?', *Journal of Public Economics*, 88, 2009-2042

Easterlin, R.A. (2005), 'Feeding the illusion of growth and happiness: a reply to Hagerty and Veenhoven', *Social Indicators Research*, 75, 429-443

Johns, H. and Ormerod, P. (2007) *Happiness, Economics, and Public Policy*, IEA Monograph 62, IEA, London

Inglehart, R., Foa, R., Peterson, C. and Welzel, C. (2008), 'Development, freedom, and rising happiness: a global perspective (1981-2007)', *Perspectives on Psychological Science*, 3(4), 264-285

James, O. (2007), *Affluenza*, Vermilion

Kahneman, D. and Krueger, A.B. (2006), 'Developments in the measurement of subjective well-being', *Journal of Economic Perspectives*, 20, 3-24.

Layard, R. (2005), *Happiness: Lessons from a New Science*, Allen Lane

Ormerod, P. and Mounfield, C. (2000), 'Random matrix theory and the failure of macro-economic forecasts', *Physica A*, 280, 497-504

[general.hj@googlemail.com](mailto:general.hj@googlemail.com) pormerod@volterra.co.uk

---

### SUGGESTED CITATION:

Helen Johns and Paul Ormerod, "Reply to Dan Turton", *real-world economics review*, issue no. 49, 12 March 2009, pp. 90-93, <http://www.paecon.net/PAERreview/issue49/Johns49.pdf>