

HOW QUANTILE REGRESSION RESCUED HAPPINESS – FOR THE RICH: AN R VINAIGRETTE

ROGER KOENKER

ABSTRACT. Killingsworth et al. (2023) have recently argued that earlier work of Kahneman and Deaton (2010) that found that the relationship between log income and self-reported happiness “flattened” at higher incomes was flawed due to overlooked heterogeneity. Quantile regression methods revealed that this effect was attributable to the least happy few, while those at upper quantiles of the conditional distribution of happiness found incremental income beyond \$100,000 just as joyful as below this threshold. Assertions in Killingsworth et al. (2023) that the conclusions of Kahneman and Deaton (2010) might have been different if the relationship between income and unhappiness were studied make no sense to me, nor do claims that there is a “correct” solution to questions raised by this line of research.

1. INTRODUCTION

I will resist the temptation to indulge in speculation about the measurement of happiness, or for that matter income. The data employed by Killingsworth et al. (2023) is available from <https://osf.io/qye4a/>. and consists of 33,391 individual observations with reported income and responses to the question, “How do you feel right now” at randomly selected times over the survey period. How these temporal responses were aggregated is not clear from the PNAS paper. Other covariates were also collected, but do not appear in the results reported or the data source provided.

2. REPLICATION

Figure 1 contains a scatter plot of the raw data and estimated decile estimates. If you squint carefully you will see that the the first and second conditional deciles have slightly flatter slopes after the (red) \$100,000 income threshold than before the threshold. The blue curve represents the ordinary least squares fit of the simple broken-line regression model. Note that the OLS fit *does not flatten* noticeably as reported in the original Kahneman and Deaton (2010) paper, and as claimed in new Killingsworth et al. (2023) paper. Since the two papers employ quite different data, and the former paper incorporates additional covariate effects this is hardly surprising, but it calls into question the claim of the latter authors that they have found a conclusive, “correct” explanation for the findings in the prior paper.

Figure 2 depicts the quantile regression point estimates with pointwise confidence bands. The red horizontal lines represent the least squares estimate and its associated confidence interval. The black lines and the associated shaded regions represent the corresponding quantile regression estimates. The indicator variable d takes the value 1 if income exceeds \$100,000,

August 5, 2023. Thanks to Angus Deaton for clarifying differences in the two data sources used in the earlier and later papers. The former paper used responses from the Gallup survey question: “did you experience a lot of happiness yesterday?” The Killingsworth et al. (2023) data came from responses to randomly timed “beeper” inquiries. A genre manifesto for R Vinaigrettes is available at <http://davidofmeaning.blogspot.com/2016/12/r-vinaigrettes.html>.

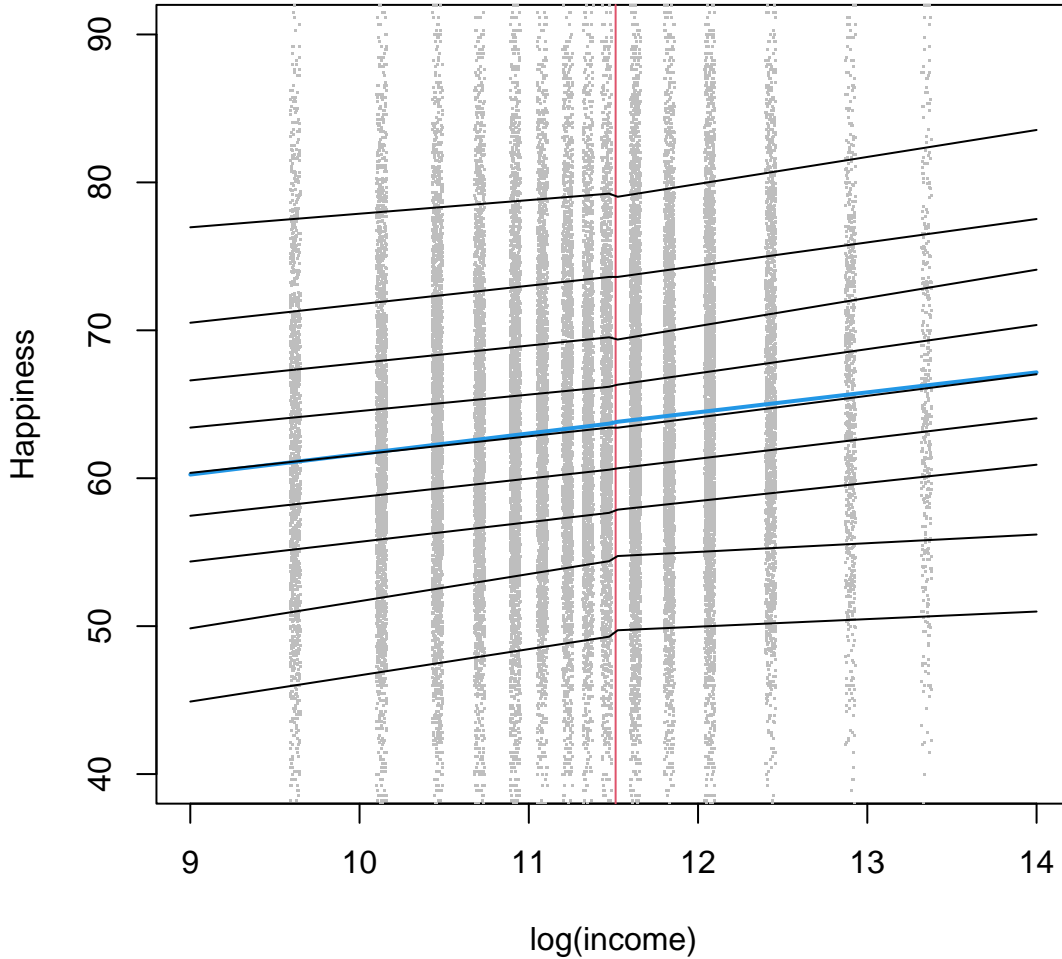


FIGURE 1. Raw Data and Regression Decile Estimates. The red vertical line is the \$100,000 threshold. The heavy blue line is the mean regression fit. Income responses have been dithered in an effort to convey a sense of their observed density.

and 0 otherwise, so the model permits two distinct linear segments: one for income below this threshold and another for income above it. Of primary interest in the present case is the difference in the slopes of these segments at the various quantiles. From the plots in Figure 2 it is apparent that at the lower deciles the difference in these slopes given by the panel labeled $x : d$ is considerably lower for the segment above the threshold than below. Although the OLS point estimate is essentially zero for this interaction term, for the 0.10 and 0.20 quantile regression estimate it is significantly below 0, as revealed by the “flattening” observed in

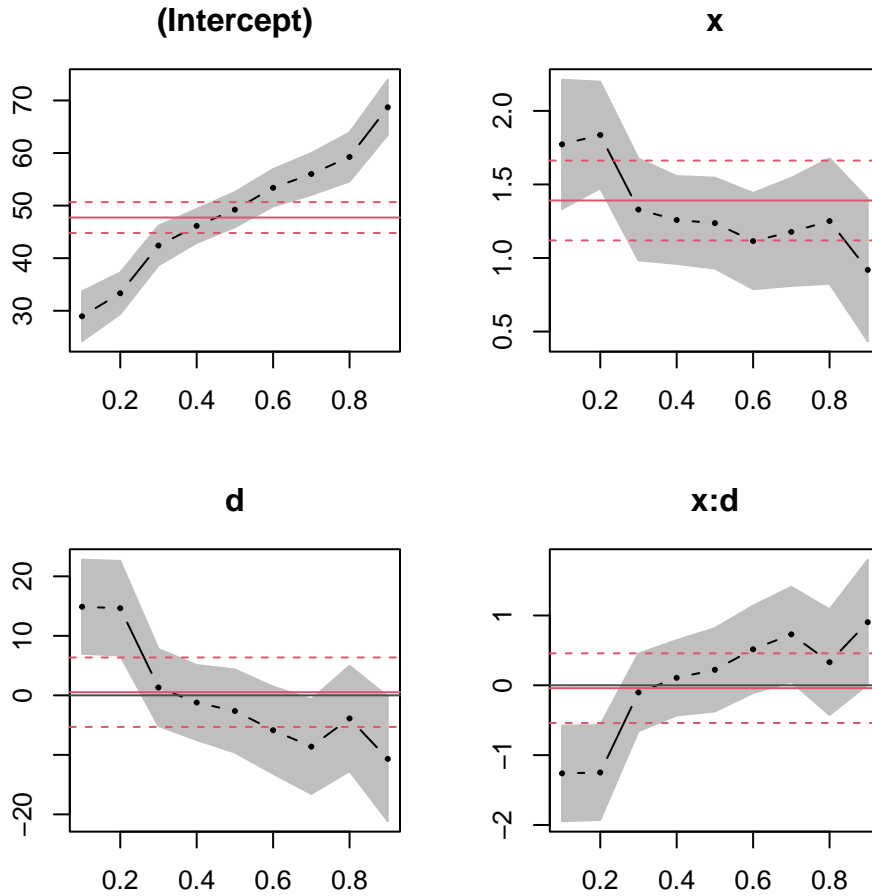


FIGURE 2. Pointwise Confidence Bands for QR Estimates

Figure 1 for these quantiles. Note too that in the upper tail the situation is reversed; the slope of the income/happiness relationship is higher above the threshold than below it. This is also visible in Figure 1 where the 0.90 conditional quantile bends upward slightly at the threshold.

3. CONCLUSION

What does it all mean? I suppose if one were to take self-reported happiness scales seriously – as it happens I do not – the results would suggest that most respondents get approximately linear satisfaction over the entire log income range surveyed, as anticipated by Daniel Bernoulli. Those who are already quite happy at low to moderate incomes get even more joy from increments to log income at higher income levels. On the contrary, those that are relatively unhappy at lower income levels don't get much of a boost from log income increments beyond the \$100,000 threshold. For the naive utilitarian this might suggest redistribution schemes from the sad to the glad, but you would have to be mad to subscribe to that.

REFERENCES

- Kahneman, D. and Deaton, A. (2010), ‘High income improves evaluation of life but not emotional well-being’, *Proc. Natl. Acad. Sci.* **107**, 16489–16493.
- Killingsworth, M. A., Kahneman, D. and Mellers, B. (2023), ‘Income and emotional well-being: A conflict resolved’, *Proc. Natl. Acad. Sci.* **120**.