An Attempted Replication of Peter Congdon's "Latent variable model for suicide risk in relation to social capital and socio-economic status"

James Lester

Introduction

A major element of scientific research concerns the exploration of highly complex systems, with patterns of variables measured produced by numerous processes, which are in turn governed by a multitude of variables. However, typically for reasons of feasibility, resources and indeed awareness, rarely, if ever, are all the possible factors influencing an observed pattern examined in detail, with instead those most relevant to the investigator's hypotheses, and deemed most important to the system explored, being prioritised. However, whilst this more classical and focussed approach to a scientific question still prevails, and remains entirely suitable in many cases, there are other situations when a more lateral and cross-disciplinary approach can be immensely valuable. In the case of human disease epidemiology (infectious or otherwise) this lateral consideration of factors is intrinsic to a rational exploration of patterns of disease, and given the sheer complexity of human populations, from their social structures to the underlying economics and indeed even individual-level factors of genetic make-up , all of which can influence disease, this can often be an immensely challenging proposition. Nonetheless, through robust experimental design disease epidemiology can (and indeed, often has) offered important and practical insights into patterns of disease within the human population, and how to reduce the burden of experience experienced.

Of course, whilst disease epidemiology has tremendous potential to enable a deeper understanding of the distribution of disease, its aetiology and how to manage it, the sheer complexity of the system at hand renders the challenge of isolating correlation from causation an immense one. However, it is one which can be tackled with the aid of a suitable number of samples from a population, along with detailed information on, and theoretical consideration of factors which could play a role. Peter Congdon's 2012 study "Latent variable model for suicide risk in relation to social capital and socio-economic status" [1] illustrates many of the aforementioned points - isolating a range of distinct factors from across several disciplines (sociology, psychiatry, demography) for the robust examination of the epidemiology of psychiatric health and making use of a dataset which provides over 7000 immensely detailed individual records.

Theoretical considerations

Social capital is a concept which possesses a range of specific definitions, but broadly it refers to the benefits that an individuals receives from social contacts within a population. So, for instance, an individual possessing a high social capital would be someone with a number of close friends and family members, deeply embedded within their local community. Intuitively, given that a high social capital would be expected to have a considerable positive impact on mental health, the inverse could also be anticipated. However, past work has produced a very mixed picture of the relationship between social capital and mental health [2], in part perhaps due to the difficulties of and variations in recording of social capital. Additionally, the lack of past research looking simultaneously at suicidal behaviour and social capital has rendered this particular area highly challenging to empirically investigate. As such Congdon's detailed assessment of the relationship between social capital and mental health, specifically the largely unexplored area of suicidal behaviour, serves to add useful evidence for assessing the question of the impact of social capital on psychiatric illness.

Data

The dataset used during this study consists of the results of the 2007 Adult Psychiatric Morbidity Survey, a survey which aimed to collect data on mental health in adults across England, with the eventual objective of using such data to analyse trends in mental health in England, along with exploring the possibility for protective social factors. In addition to asking questions about specific mental health diagnoses, the survey included questions on subjects ranging from social functionality and capital to household composition and cognitive performance, lending it a degree of breadth and detail ideal for epidemiological analysis. Whilst the large size of the dataset could render it challenging, particularly with the wide range of questions and the responses that they allow, the comprehensive documentation of the dataset ensures that handling it is not unnecessarily difficult.

Methods

The R programming language was used for the replication of the most basic figures and table present in the first half of the paper, as this required only basic recoding and numerical manipulation. Whilst the actual names of the variables were not included by Congdon, most descriptions were typically suitably clear to allow for their identification in the dataset. An important aspect of his manipulations of the dataset are conversions from data columns containing multiple categories to binary responses. Whilst methodologically reasonable, in practice Congdon does not clarify the process employed to do this, beyond a somewhat vague description of 0 for disagreement, and 1 for agreement, rendering this challenging to implement in light of the multi-point scales used, which incorporate an option for neutrality. Additionally, it appears that Congdon equated only the highest agreement responses (e.g. strongly agree) with agreement, with other responses "below" this being counted as disagreement, although the rationale for this is not discussed. A table based upon that in the original paper is included, further detailing the values used for recoding into a binary format, along with the variable names used. These now binary responses are then combined in a clearly described calculation to produce a "social capital score", to provide a broad indication of the relationship between social capital and suicidal behaviour in this dataset.

Following this, latent variable analysis is performed in WinBUGS. Whilst the overall framework is described by Congdon, the sparse details provided rendered it highly challenging to try to replicate this part of the analysis. The method used, Bayesian structural equation modelling, is a highly useful but somewhat rarely used technique of which there exists relatively few examples (with notable examples of those that do exist being a book [3], an article [4] and a lecture [5]), and so in the absence of more detailed explanation or code, I was unable to replicate this portion of the paper effectively, and so the results of this portion of the replication will not be explored further (code is nonetheless included).

	Variable name	Dataset variable
Suicidal behaviours		
Y1 ideation	dshlife	There may be times in everyones life when they become very miserable and depressed and may feel like taking drastic action because of these feelings. Have you ever thought of taking your life, even if you would not really do it?
Y2 attempt	dshtry	Have you ever made an attempt to take your life, by taking an overdose of tablets or in some other way?
Y3 self-harm	dshharm	Have you ever deliberately harmed yourself in any way but not with the intention of killing yourself?
Social capital		
X1	dlss1	Family and friends do things to make me happy (certainly true = 1, not true/partly true = 0)
X2	dlss2	Family and friends make me feel loved (certainly true = 1, not true/partly true = 0)
X3	dlss3	Family and friends can be relied on no matter what happens (certainly true = 1, not true/partly true = 0)
X4	dlss4	Family and friends would see that I am taken care of if I needed to be (certainly true = 1, not true/partly true = 0)
X5	dlss5	Family and friends accept me just the way I am (certainly true = 1, not true/partly true = 0)
X6	dlss6	Family and friends make me feel an important part of their lives (certainly true = 1, not true/partly true = 0)
X7	dlss7	Family and friends give me support and encouragement (certainly true = 1, not true/partly true = 0)
X8	Belong	I feel like I belong around here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X9	Trust	I trust people around here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X10	Enjoy	I enjoy living around here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X11	Realhme	I think of the area around here as a real home not just a place (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X12	Safe	I feel safe around here in the day time (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X13	Move	Given the opportunity I would like to move away from here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X14	Resident	The area around here is nicely kept by its residents (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X15	Litter	Litter is a problem around here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X16	Graffit	Graffiti or vandalism is a problem around here (strongly agree = 1, somewhat agree/disagreement/neutrality = 0)
X17	ComGrp	Participate in community group (at least once a month = 1, less than once a month/never = 0)
X18	noclub	Involved in clubs/activities (no = 1, yes = 0)

Table 1: Indicators of suicidal behaviour and social capital

Results

The initial table of suicide outcomes split by age fundamentally retains much the same significance as it does in Congdon's paper, indicating a particularly high prevalence amongst the youngest age category of 18-34. Whilst different in absolute values do exist, this table also retains the slight increase in suicidal thoughts in the 35-54 age category compared to the 18-34 category, indicating that whilst some underlying different treatment of the data has clearly occurred, it has not been suitably extreme as to disrupt fundamental patterns present within the dataset.

	18-34	35-54	55-74	75+	All
Suicidal thoughts	21.1	23.0	13.8	4.0	17.3
Suicide attempts	8.2	7.5	4.8	1.3	6.0
Self-harm	9.4	5.2	1.2	0.3	4.2

Table 2: Percentage suicide outcomes

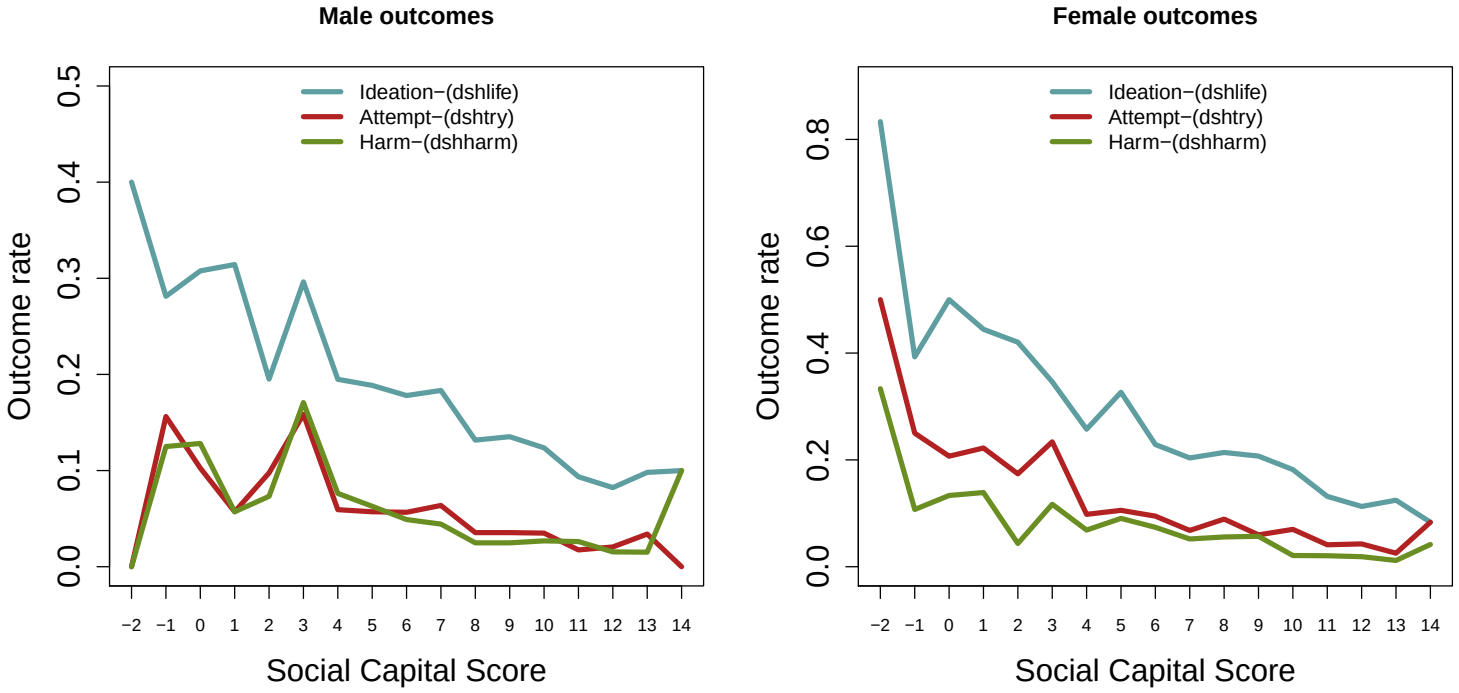

Figure 1: Suicide response against ad-hoc social capital score

Next is the comparison of the ad-hoc "social capital score" (hereafter SCS) with suicidal outcomes. Here we see that within the female portion of the population, suicide ideation is more abundant than suicide attempts, and both are more common than self-injury. It is also clear that there is an overall trend towards a lower outcome rate for all of these in the presence of a higher SCS, although this is far more pronounced for suicide ideation than self-injury. These conclusions are very similar to those from Congdon's figure, although in the replication the outcome rates are consistently higher than in the original paper. For male outcomes however, the figures diverge

rather more considerably. Overall, the replication indicates that whilst suicide ideation in males does show a very small negative correlation with higher SCS, distinguishing such a trend for suicide attempts or self-injury is almost impossible. Additionally, the rates of suicide outcomes are consistently lower in males than in females across all SCS. This makes for a dramatic comparison with the original plot which indicates an overall trend much like that for females, but with consistently lower outcome rates. As a result of this, rather different conclusions could be drawn on the basis of the replication, namely that whilst suicidal ideation in males does reduce with higher levels of social capital, the low levels of suicide attempts and self-injury which are reported are relatively unaffected by changes in social capital. The rationale for this difference (even with many different approaches being taken for its formulation), could indicate an error on my, or Congdon's behalf, or simply differences in our approaches to calculating SCS.

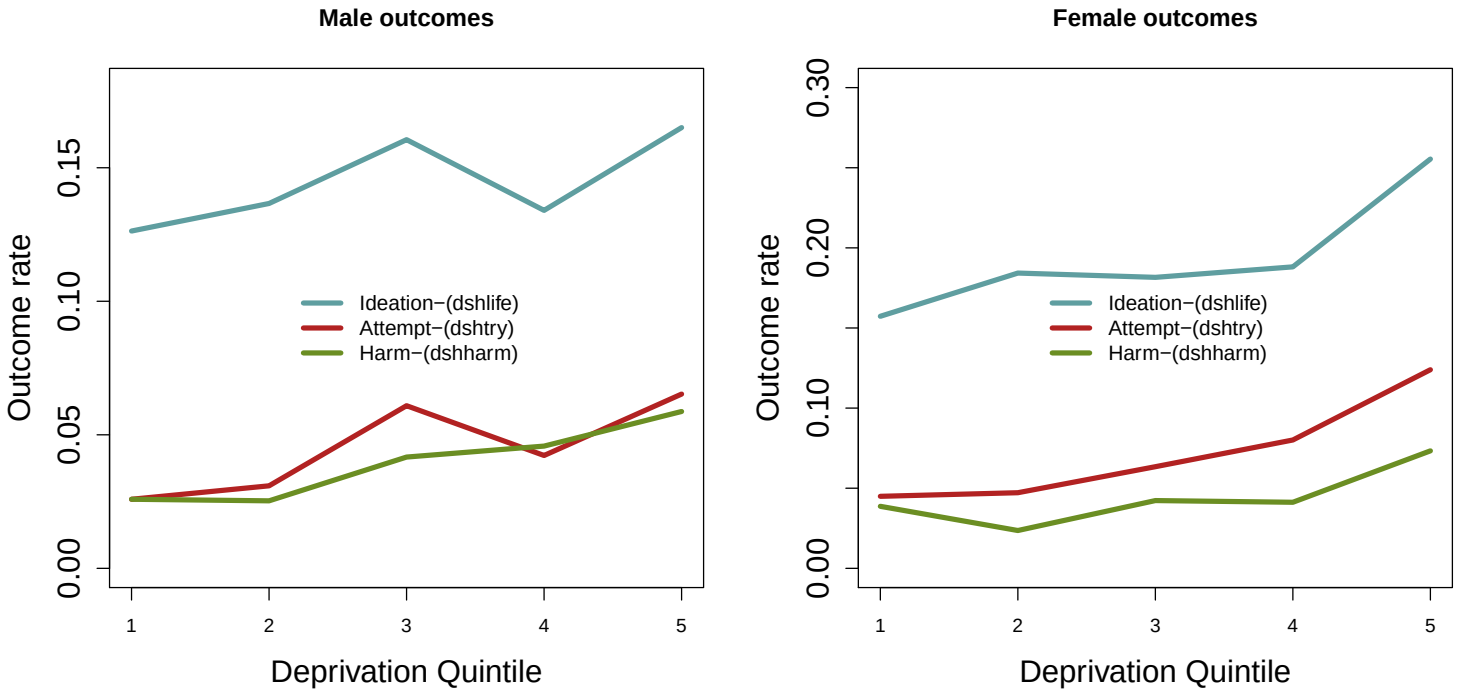

Figure 2: Suicide response against economic quintile

Further supporting this suggestion of some form of cryptic calculation difference being the problem is the effective replication of the figure assessing gender differences in suicide outcomes. Just as in Congdon's paper, this figure indicates a correlation between higher neighbourhood deprivation quintiles and higher rates of suicide outcomes. The efficacy of this element of the replication in comparison to the others further clarifies that the social capital score calculation is the key area of discrepancy, whilst also indicating that the age subsetting of the first table may also have differed, given that clearly the relevant suicide outcome measures were indeed used.

Conclusions

Congdon's paper both adds an additional, and seemingly strong piece of evidence to the debate over the impact of social capital on psychiatric illness in some specific areas, and offers an example of the value of complex Bayesian methods to do so. However, in both of these approaches the relative lack of explanation and detail provided for the "nitty gritty" of methodological choice renders a paper which could have additionally been a valuable and helpful tool for those seeking to apply similar methodology to new datasets, considerably less helpful. For the most part I was indeed able to reproduce Congdon's methodology, but the lack of fundamental detail renders one of the contrasting conclusions, that certain male suicidal outcomes may vary little with changing SCS, impossible to further clarify. Furthermore, whilst offering a robust applied example of a Bayesian SEM model which also captures the effects of spatial autocorrelation is a highly valuable contribution to the literature, it is unfortunate that further detail was not provided in the appendix, or supplementary to the paper, given the potential value of this paper to a great many areas of research.

References

[1] Peter Congdon. Latent variable model for suicide risk in relation to social capital and socio-economic status. *Social psychiatry and psychiatric epidemiology*, 47(8):1205–1219, 2012.

[2] Mary J De Silva, Kwame McKenzie, Trudy Harpham, and Sharon RA Huttly. Social capital and mental illness: a systematic review. *Journal of Epidemiology and Community Health*, 59(8):619–627, 2005.

[3] Sik-Yum Lee. *Structural equation modeling: A Bayesian approach*, volume 711. Wiley, 2007.

[4] Sik-Yum Lee, Xin-Yuan Song, and Nian-Sheng Tang. Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling*, 14(3):404–434, 2007.

[5] Peter Congdon. *Bayesian factor and structural equation models in spatial applications. Specification, identification and model assessment, with case study illustrations*,. Lecture, 2009.