Time-varying differential misclassification is always possible, but highly unlikely
Response to commentary by Bentley, Baker, and Blakely.

We appreciate the interest our paper has generated, which takes advantage of a natural experiment design to better ascertain causality than conventional observational studies.¹ In the accompanying commentary, Bentley, Baker, and Blakely suggest that misclassification of depression symptoms (e.g., if someone with depression reports that they are not depressed) may explain our findings if misclassification rates changed due to the reduction in the Local Housing Allowance, one component of Housing Benefit in the UK for people in the private renter sector.²

Misclassification can bias estimated effects in two ways. First, if non-differential over time then it will dilute associations, tending to underestimate effect sizes. This is likely to occur from measurement error. Second, differential misclassification can impact the direction of findings. This can occur, for example, if the reduction in Housing Benefit itself leads to changes in how people report symptoms of depression. This is what the authors suggest may have happened and generated a spurious correlation.

Is there evidence to suggest that Housing Benefit reductions impacted how people reported depression in the Annual Population Survey?

We do not find any. It also seems highly unlikely. The main reason is that, according to one sample, 85% of those affected by the policy were unaware that they had lost income due to the reduction in Housing Benefit. Moreover, 75% of this same sample did not know how housing benefit was calculated.³ It is possible the remaining 15% changed how they responded but it is difficult to imagine why people would change the way they responded to one specific survey question because of this administrative change in how Housing Benefit is calculated.

Nevertheless, to address this possibility, we further investigated whether changes occurred in other subjective health problems, which could similarly have had misclassification bias but which would not plausibly be affected by the Housing Benefit reform. These health outcomes therefore serve as ‘falsification tests’ for the possibility of misclassification. The self-reported measures are ‘chest or breathing problems’, ‘allergies’, or a generic category measuring any ‘other health problems’. Each is ambiguous, open to subjective interpretation, and therefore possibly also subject to time-varying misclassification. But, unlike mental health, none of these are associated with the reform (see Table 1). If the reduction in Housing Benefit changed how people responded to surveys about self-reported health problems, then it only affected this single response (depression) to this one question (health). It seems much more likely to have been a real mental health effect.

Of course, as Bentley et al. note, even a strong natural experiment design cannot rule out potential misclassification bias. We agree. Yet, our evidence consistently suggests that any such bias is likely to be non-differential with respect to our hypothesis, leading to conservative estimates of the causal effect. Time-varying differential misclassification is always possible, but in this instance it is highly unlikely.
Table 1: Estimates of the difference-in-differences for various subjective measures of health comparing April 2009-March 2011 and April 2011-March 2013, private renters, matching model

<table>
<thead>
<tr>
<th>Probability of people reporting:</th>
<th>Depression (1)</th>
<th>Chest/breathing problems (2)</th>
<th>Allergies (3)</th>
<th>‘Other’ health problems (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference-in-difference estimate (after April 2011)</td>
<td>0.011** (0.0038)</td>
<td>-0.0047 (0.0038)</td>
<td>-0.0031 (0.0027)</td>
<td>-0.0020 (0.0030)</td>
</tr>
<tr>
<td>Observations</td>
<td>150,731</td>
<td>150,731</td>
<td>150,731</td>
<td>150,731</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are reported in parentheses. All models include a measure of change over time, the difference between housing benefit recipients and non-housing benefit recipients before April 2011, and the probability of depressive symptoms among non-housing benefit recipients before April 2011.

* p < 0.05   ** p < 0.01   *** p < 0.001
References

