Behavioural and Experimental Economics: Discussion Questions

Title: *The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List*

Author: Colin Camerer

Publication Date: 30 December 2011

Thomas A. Stephens

13 June 2012

1 What does it mean when distributions of responses are similar in lab and field experiments, but individual correlations are relatively low?

On pages 31–32, Camerer refers to Benz and Maier's (2008) results comparing naturally occurring donations to student funds in Switzerland with a similar lab experiment. He provides 'good news' of similar distributions, but with the caveat that individual correlations are relatively low, in the 0.22 to 0.31 range.

The above result is particularly interesting to me because in the NLA vignette study we did as part of iLEE2/3, there is a correlation of about 0.25 between the two NLA index measures taken a year apart. It also reminds me of results in Dulleck et al. (2011), who find similar distributions of risk preferences when they are measured multiple times, but argue that a relatively low correlation of 0.39 within subjects casts doubt on the idea of persistent risk preferences.

We covered this topic to some extent in the discussion of Dulleck et al., but is noise sufficient to explain these low correlations? Camerer's claim that correlations of trait-like behaviour are often not much larger than 0.30 is encouraging, particularly

since it applies to field-field comparisons, as well as to lab-field and lab-lab. If a strong case can be made that these correlations are reasonable for traits that are widely accepted as 'real', it would be very encouraging. It is not clear where Camerer's claim comes from. It may be Andreas Leibbrandt's (2011) ESA presentation, but unfortunately this does not appear to be available online.[1]

2 Can mistaken guesses about research hypotheses by subjects be ignored?

On page 18, in a discussion of potential experimenter demand effects, Camerer refers to results by Lambdin and Shaffer (2009), who find that most subjects (80%) in their experiment are confident that they can identify the research hypothesis the experiment is designed to investigate. Most of them are wrong in their guesses, but does that necessarily matter?

If financial incentives are weak, and demand effects can influence results, it is not obvious to me that demand effects based on incorrect beliefs are necessarily less problematic than demand effects based on correct beliefs. They would still distort behaviour and potentially corrupt results. As long as incentive are sufficient to overcome demand effects (condition (b) in Camerer's discussion), this arguable does not matter, but it is not clear to me why demand effects motivated by mistaken beliefs are necessarily benign, if demand effects motivated by correct beliefs are not.

3 If subject behaviour differs in the lab and the field, is the direction of the difference important?

On pages 29, Camerer notes that local card dealers in Levitt and List (2008) appear to be more reciprocal in the field than in the lab. This goes against Levitt and List's hypothesis that lab subjects are more co-operative than individuals in the field. However,

[1] Camerer spells the surname as Liebbrandt, but there are no results under either spelling.

as a criticism of lab experiments, does consistency with Levitt and List's hypothesis matter, or is the more important issue simply that behaviour differs?

A lab-field difference that goes in the opposite direction of the one predicted by Levitt and List naturally presents a problem for their story, but does this really undermine their argument? Could they be wrong on the details of the story but right on the general criticism?