

On the external validity of construction bidding experiment

Bee Lan Oo

University of New South Wales, Australia

Abstract

The external validity of experimental studies and in particular, the subject pool effects have been much debated among researchers. The common objections are that the use of student as experimental subjects is invalid as they are likely to be unrepresentative. This paper addresses this methodological aspect in building economics research. It compares the bidding behavioural patterns of experienced construction executives (professionals) and student subjects through replication of a bidding experiment that aimed at testing theories. The results show that the student subjects' bidding behavioural patterns, in terms of decision to bid and mark-up decision, are sufficiently similar to that of the professionals. This suggests that the subject pool *per se* is not a threat to the external validity of the bidding experiment. In addition, the demonstrated practicality of an experimental approach in testing theories should lead to more use of experimental studies with student subjects in building economics research. It is suggested that experimental and field findings should be seen as complementary in building economics research, as advocated in social sciences.

Keywords: Construction bidding, experiment, external validity, methodological issue.

Paper type: Research article

Introduction

Since the critiques made by Runeson back in 1997 on the importance of theory in construction management and economics research (1997a), and on the slow progress in its methodological component - in the sense of using what is appropriate from related, well-established disciplines (1997b), it seems that there has been little progress on these aspects in the discipline over the last two decades. This is evidenced in his most recent review (Runeson and de Valence 2015) in the *Construction Management and Economics* journal with comments on poor standard of current research in construction or building economics research. In their review that focused on research on tendering theory and innovations in construction, they cannot emphasize enough the importance of using tried and tested theories and methodologies in progressing science in the discipline, with one of the key advantages being that appropriate research methods would have been established. This study focuses on the methodological aspect in research on construction bidding, which is conventionally referred to as part of building economics. Specifically, it examines the external validity of bidding experiments where students were used in place of professionals to establish behavioural patterns. This is a commonly used research method in social sciences although there are very few examples in construction management.

Hence, while there is no study specifically on the external validity of experimental studies in building economics, the problem has been much debated in the social sciences of which

Copyright: Construction Economics and Building 2016. © 2016 Bee Lan Oo. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 Unported (CC BY 4.0) License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: Oo, B. L. 2016. On the external validity of construction bidding experiment, *Construction Economics and Building*, 16(1), 64-75. DOI: <http://dx.doi.org/10.5130/AJCEB.v16i1.4818>

Corresponding author: Bee Lan Oo; Email - bee.oo@unsw.edu.au

Publisher: University of Technology Sydney (UTS) ePress

construction management and economics research are parts. By external validity, I refer to the ability to generalize results (or behavioural observations) from the laboratory to the non-laboratory environments (typically called field or real world), i.e., the problem on generalizability (Campbell and Stanley 1966). Specifically, it is the ability of a causal relation $x = f(y)$ to be generalized over subjects and environments (Frechette 2015). Alm, Bloomquist and McKee (2015) referred the subjects and environments as subject pool and context effects, respectively. In terms of subject pool effects, a common objection among construction management researchers, sometimes echoed by other social scientists is that the use of student as experimental subjects is invalid as they are unlikely to be representative of the population that is tested (Falk and Heckman 2009). Falk and Heckman (2009) have examined the five most commonly mentioned issues related to subject pool effects, namely: (i) stakes or monetary rewards in experiments are trivial; (ii) the number of subjects is too small; (iii) subjects are inexperienced; (iv) the possibility that subjects behave differently because they perceive that they are observed; and (v) the self-selection of subjects may bias results. For context effects, critics refer to the extent to which the context in the laboratory decision resembles the context in the field for the same decision (Alm, Bloomquist and McKee 2015). This study focusses on the subject pool effects with a very precise question: *do student subjects behave different to nonstudent subjects in an identical construction bidding laboratory experiment?* Nonstudent subjects here refer to construction executives with experience in bidding.

To the author's knowledge, there are no studies that have compared the behavioural patterns in construction bidding between student and nonstudent subjects in an experimental setting. However, the study on auction theory by Dyer et al. (1989) may be considered as the closest work that compared the bidding decisions between construction executives and students in common value auctions. In answering the question, this study proposes a direct replication of the bidding experiment in Oo (2007) by replacing its nonstudent construction executives (professionals) subjects with student subjects. Replication, despite its unpopularity in construction management and economics research, has been seen as a key self-correcting force (together with peer review) towards ensuring a high standard in research in the discipline (Runeson and de Valence 2015). Here, the findings are important for two reasons: (i) because they may validate the use of an experimental approach in testing theory, and (ii) because they test the applicability of the use of students as experimental subjects. Notably, undergraduate students are typical subjects in social science research using experiments, but little is known about their use in construction management and economics research.

Resistance to experimental approach in building economics research

Before progressing any further, it may be worthwhile to highlight that experiments have been remarkably successful in terms of extending, among many other things, the theoretical framework for auction theory in economics. Indeed, the use of laboratory experiments in economics are common and the first specialty journal - *Experimental Economics* - was founded in 1998 (Falk and Heckman 2009). In contrast, the practicability of an experimental approach has been demonstrated in a very limited number of papers in construction management and economics discipline. For the two well-known journals in the discipline - *Construction Management and Economics* and *Journal of Construction Engineering and Management* - the fraction of experimental studies in relation to all published papers between 1983 and 2015 (with online access) are as low as 1% (26 out of 2597) and 3.6% (99 out of 2788), respectively (author's calculations). Recently, however, there is a small collection of experimental studies on construction bidding as presented in the next section. Thus, it can be assumed that many researchers in the discipline are still reluctant to accept laboratory evidence, as indicated by the overwhelming use of other research methods, especially surveys and case studies. This also explains why the debate on the resistance

to experimental studies, and in particular the debate on the external validity is absent in the discipline. The subsequent review is, thus, based on literature in social sciences that applies to building economics research.

Falk and Heckman (2009) noted that the perceived lack of realism and generalizability have caused considerable resistance among social scientists to accept laboratory evidence. However, they argued that many recent objections against laboratory experiments are misguided due to a misunderstanding of the nature of evidence in science and of the kind of data collected in the laboratory, and that more laboratory experiments should be conducted. This argument is consistent with the claims of those who argued that many experiments do not have to represent the 'real world' in any direct way, and that a proper approach to address external validity questions that suit the various goals of experimentalists should be developed (e.g., Vissers 2001; Guala and Mittone 2005; Schram 2005). In fact, many experiments are aimed at contributing to a body of experimental knowledge to be applied case by case, referred as a 'library of robust phenomena' by Guala and Mittone (2005). Similarly, Camerer (2011) argued that external validity is crucial for experimental studies that aim to inform policy, but not for experimental studies aiming at understanding general principles, in what he referred to as the policy view and science view, respectively. Kessler and Vesterlund (2015), on the other hand, pointed out that the resistance to experimental studies has centered on the extent to which the quantitative results (e.g., the magnitude of response, the point predictions) are externally valid. They argued that for most laboratory studies it is only relevant to ask whether the qualitative results (e.g., the direction of response) are externally valid, but not whether an exact quantitative result can be found in experimental data. Above all, it is interesting to note that, in defense of external validity, many authors have concluded that laboratory and field results are highly complementary, and that both are important to the progress of knowledge in social sciences (e.g., Levitt and List 2007, Falk and Heckman 2009, Kessler and Vesterlund 2015).

In terms of subject pool effects, there is a number of studies that specifically compare student and nonstudent experimental subjects through replications. There is for example a collection of thirteen articles in Frechette (2015) which includes the study by Dyer et al. (1989) and the other twelve articles that he loosely classified into four thematic groups: (i) preferences; (ii) market experiments; (iii) information signals; and (iv) a miscellaneous group. It should be noted that his comparisons were focused on the comparative statistics and qualitative nature of the results rather than in tests of point predictions (i.e., in line with Kessler and Vesterlund's argument). In most (9 out of 13), although not in all cases, the comparisons showed no significant difference in behaviour between student and nonstudent experimental subjects in a way that would lead us to draw different conclusions when testing theories. In another recent study, Alm et al. (2015) examined both the subject pool and context effects in laboratory tax compliance experiments. Similarly, their results show that the behavioural patterns of student and nonstudent subjects in their experiments are sufficiently similar, and in particular, their comparisons were based on statistical modelling (or quantitative) results.

Experimental studies on construction bidding

According to Roth (1988), laboratory experiments in economics can be classified into three categories based on their primary goals: (i) 'Speaking to Theorists'; (ii) 'Search for Facts'; and (iii) 'Whispering in the Ears of a Prince. Experiments in the first category test hypotheses that have been derived from specified models or theories. 'Search for Facts' includes experiments that examine the effects of variables about which existing theory has little to say. Lastly, the third category includes experiments motivated by policy issues. To date, most experiments in building economics research, or construction bidding have fallen into the first two categories. For examples of experiments with nonstudent (experienced professionals) subjects, Hackemer (1970)

examined the effects of variability of estimate, number of competitors and mark-up on bidding strategies by asking five competitors to bid for 200 contracts. The experiment, what he referred to as ‘simulation’, produced some 200,000 bids via application of different variability factors of estimate. In Drew and Skitmore (2006), a bidding experiment was designed to test the applicability of Vickery’s revenue equivalence theorem in construction bidding. Perng et al. (2006), on the other hand, tested their conceptual model on economically most advantageous tender in construction procurement. Next, the bidding experiment in Oo (2007) aimed at testing the tenability of bidder homogeneity assumption of tendering theory.

Turning into examples of experiments with student (inexperienced) subjects, recently, there is a series of experiments that aim to test the effect of information feedback on bidders’ competitiveness and learning (Soo and Oo 2010, Oo, Abdul-Aziz and Lim 2011, Oo, Ling and Soo 2014, 2015). Also, there is an experiment on the effect of construction demand in construction bidding (Soo and Oo 2014, Soo 2015). Both the qualitative and quantitative results from these studies are consistent with empirical findings based on field data, suggesting that the subject pool *per se* may not be a threat to the external validity of bidding experiments aimed at testing theories. However, hitherto, there is no study specifically testing the subject pool effects in experimental studies on construction bidding. Oo’s (2007) experiment was selected in this study because the author has described all events in detailing the designed bidding experiment that make a replication effort possible.

Research method

The experiment in the present work is a replication of the bidding experiment in Oo (2007). While Oo’s experiments involved 18 and 31 experienced professionals from the Hong Kong (HK) and Singapore (SIN) construction industries respectively (hereafter nonstudents experiment); the present work involved a group of 24 inexperienced student subjects (hereafter students experiment). The student subjects were enrolled in a construction estimating and bidding course as part of their construction management undergraduate degree program in a tertiary institute in New South Wales (NSW), Australia. Although a small number of the student subjects (5 out of 24) were working in the construction industry as revealed in a survey before commencing the students experiment, none of them was involved in the bidding decision making process of their organizations – the selection criterion for experimental subjects in the present work. This justified the inclusion of their responses in the experimental dataset. In addition, it should be noted that there were eight international students from Asian countries (e.g., China, Singapore, and Myanmar) in the group. Their presence was not engineered for the students experiment, and indeed their presence was random and inevitable as all Australian universities have comparatively high international student numbers. Above all, these international students were inexperienced subjects, and thus eligible for inclusion in the sample.

The instruments from the non-students experiment, consisting of an instruction page and a bid response form, were used in the students experiment with only one change. That is, to revise the descriptions of the twenty hypothetical general building projects using public sector project information obtained from the NSW e-tendering website. This was done to localize the project descriptions in controlling the effects of extraneous variables, including: (i) contract type; (ii) project type; (iii) project location; and (iv) client type. Similar to the nonstudents experiment, the hypothetical projects were lump sum contracts for conventional buildings, such as schools and institutional buildings located in the Sydney region with the NSW government as project client. The unbiased cost estimates given in the bid response form were also expressed in the local currency (AUD).

Replicating the nonstudents experiment, the student subjects were invited to participate in the

experiment by (i) acting as senior managers of their construction firms; and (ii) bidding for a total of 20 hypothetical projects. For every hypothetical project, there were eight number of bidders' scenarios (ranging from 4 to 30) with the estimated number of competing bidders, N , increasing in the levels of 4, 6, 8, 10, 14, 18, 24 and 30. With the given project information (location, duration, client and contract type) and an unbiased cost estimate for each hypothetical project, they were required to decide which project to bid for, and bid up to the bidding scenarios of N bidders that they wish to bid by completing the bid response form. The student subjects were informed that the lowest bidder would win the job, and that their ultimate aim is to survive and prosper. The experiment was arranged in two rounds according to two extreme market conditions scenarios, i.e., (i) boom times with low need for work, and (ii) recession times with high need for work. The same twenty hypothetical projects were used in both rounds of the experiment to establish a strong basis for comparison of results. However, the sequence of the projects was randomly revised in the second round in order to avoid contamination of responses.

It is noted that the findings from the nonstudents experiment have been reported in a number of publications (e.g., Oo, Drew and Lo 2007a, 2007b, 2008a, 2008b and 2010), examining the nonstudent subjects' decision to bid and their mark-up decisions using different statistical modelling techniques (i.e., quantitative results). However, it was necessary to obtain the nonstudents experimental dataset from the author for comparing the student and nonstudent subjects' bidding responses in the present work. This is because the focus of the analysis is on the comparative statistics and qualitative nature of the results rather than a test of point predictions - a principle of analysis adopted from Frechette (2015). To illustrate this principle, consider the following hypothetical example. Suppose that the probability of a 'bid' decision is 0.4 for recession scenario, and 0.3 for booming scenario from the student subjects, and that similar trend was observed from nonstudent subjects with the recorded probabilities of 0.6 and 0.45, respectively. Here, the interesting results is the fact that both the student and nonstudent subjects recognized the need to submit more bids in "recession with high need for work" as reflected by the higher probabilities of a 'bid' decision by each group. The conclusion from these results is that the findings are robust, and the results are classified as the same for the two groups. Even if all the numbers are different between the two groups the qualitative response is the same. To be considered different, the two groups should produce results which lead to a different interpretation of bidding behavior with respect to the respective theoretical predictions.

Data sample

The comparisons between the student and nonstudent subject groups are based on the subjects' decision to bid and mark-up decisions. Appropriate statistical tests for dependent samples or matched-pairs (each observation in the booming scenario pairs with an observation in the recession scenario) were selected in the respective sections by testing whether the subjects' bidding decisions are statistically different across the market conditions and number of bidders' scenarios. With the adopted principle of analysis, no test was performed to test whether there is statistically different between the student and nonstudent subject groups. That is, the focus is on the comparative statistics and qualitative results (i.e., the directional effects). Table 1 shows the experimental datasets used for the comparisons. It should be noted that the outliers or non-serious bids have been removed from the analysis on the subject groups' mark-up decision, which is expressed as a percentage above the unbiased cost estimate [$MU\% = (\text{bidder's bid} - \text{unbiased cost estimate}) / \text{unbiased cost estimate} \times 100$]. The removal of outliers was based on criterion set that include all non-serious bids that are 25% above the individual project unbiased cost estimate.

Table 1: The experimental datasets

Experiment	Datasets (number of bids)					
	Booming			Recession		
	Bid	No-bid	Outliers	Bid	No-bid	Outliers
Decision to bid						
Students	1921	1919	n.a.	2229	1611	n.a.
Nonstudents - HK	1290	1590	n.a.	1851	1029	n.a.
Nonstudents - SIN	1645	3315	n.a.	1915	3044	n.a.
Mark-up decision						
Students	1025	1919	896	2012	1611	217
Nonstudents - HK	1224	1590	66	1851	1029	0
Nonstudents - SIN	1417	3315	228	1720	3044	196

Results

Table 2 shows the various metrics used to compare the subjects' decision to bid according to the two market conditions. It can be seen that the sample proportions of 'bid' decision are higher for both subject groups in recession than the booming scenario, providing suggestive evidence of their willingness to compete for jobs despite the expected strong competition with fewer jobs are available. Although the difference of the sample proportions of 'bid' between the two market conditions for HK nonstudent subjects (0.20) is higher than the student (0.08) and SIN nonstudent (0.06) subject groups, the changes in the willingness to bid in these three subject groups are comparable as demonstrated by the respective 95% confidence intervals for the true change in probability of a 'bid' decision that yield negative values. That is, it is 95% certain that the probability of a 'bid' decision is lower in booming than the recession scenario.

Table 2: The experimental subjects' decision to bid according to market conditions

Decision to bid metric	Experimental subjects		
	Students	Nonstudents - HK	Nonstudents - SIN
1 Booming: sample proportions of 'bid' decision	0.50	0.45	0.33
2 Recession: sample proportions of 'bid' decision	0.58	0.65	0.39
3 95% CI for true change in probability of a 'bid' decision	(-0.07, -0.09)	(-0.18, -0.22)	(-0.05, -0.07)
4 McNemar test <i>p</i> -value	< 0.01	< 0.01	< 0.01
5 Estimated odds ratio exp (β) in logit model	0.465	0.235	0.481
6 Odds of a 'bid' decision in recession	2.15	4.26	2.08

A more critical test of the decision to bid trends involved testing the null hypothesis that the probability of a 'bid' decision for the booming and recession scenarios is identical. To do this, the McNemar test was used to test the matched-pairs and the results are reported as the fourth metric in Table 2. The test results for the student and nonstudent groups all yield a *p*-value less than 0.01 providing further strong evidence of increase in willingness to bid during recession,

and thus the null hypothesis is rejected. In terms of the relationship between market conditions and the probability of a ‘bid’ decision, the estimated odds ratios $\exp(\beta)$ that are not equal to zero in the respective logit models (see Oo, Drew and Lo 2008a) show that there is a statistical significant relationship between market conditions and number of bids submitted by the student and nonstudent groups.

Taking the student group as an example, the estimated odds ratio $\exp(\beta)$ is found equal to 0.465, or equally the odds of a ‘bid’ decision are 2.15 (i.e. $1/0.465$) times higher in recession than in booming scenario. Despite the use of various metrics in the comparison, it is interesting to note that decision to bid trend of the student group is closer to that of SIN nonstudent group. This is true for the difference in the sample proportions of ‘bid’ between the two market conditions (metrics 1 and 2), metrics 3, 5 and 6.

Table 3: Sample proportions of ‘bid’ according to market conditions and number of bidders

Experimental subjects	Sample proportions of 'bid'							
	No. of bidders							
	4	6	8	10	14	18	24	30
Students								
Booming	0.69	0.67	0.63	<i>0.52</i>	<i>0.40</i>	<i>0.38</i>	<i>0.36</i>	<i>0.34</i>
Recession	0.73	0.71	0.68	<i>0.63</i>	<i>0.58</i>	<i>0.50</i>	<i>0.42</i>	<i>0.41</i>
95% CI	(.01, -.07)	(.00, -.08)	(.00, -.08)	(-.06, -.14)	(-.13, -.22)	(-.08, -.17)	(-.02, -.11)	(-.02, -.10)
Nonstudents - Hong Kong								
Booming	<i>0.74</i>	<i>0.74</i>	<i>0.65</i>	<i>0.52</i>	<i>0.28</i>	<i>0.23</i>	<i>0.21</i>	<i>0.21</i>
Recession	<i>0.83</i>	<i>0.83</i>	<i>0.83</i>	<i>0.70</i>	<i>0.61</i>	<i>0.47</i>	<i>0.43</i>	<i>0.43</i>
95% CI	(-.13, -.04)	(-.13, -.04)	(-.23, -.13)	(-.24, -.12)	(-.39, -.27)	(-.30, -.19)	(-.28, -.17)	(-.28, -.17)
Nonstudents - Singapore								
Booming	0.54	0.54	<i>0.48</i>	<i>0.42</i>	<i>0.31</i>	<i>0.15</i>	<i>0.11</i>	<i>0.11</i>
Recession	0.56	0.55	<i>0.53</i>	<i>0.45</i>	<i>0.38</i>	<i>0.27</i>	<i>0.18</i>	<i>0.16</i>
95% CI	(-.08, .03)	(-.07, .04)	(-.10, .01)	(-.09, .02)	(-.12, -.01)	(-.17, -.08)	(-.11, -.03)	(-.09, -.01)

Any pairs with italicized sample proportions of 'bid' indicate the difference in probability of 'bid' between the two market conditions is statistically significant at $p < 0.01$.

Table 3 summarizes the subject groups’ sample proportions of ‘bid’ according to the two market conditions and number of bidders, N . The student subjects’ sample proportions of ‘bid’ decrease as the number of bidders increase for both the booming and recession scenarios, consistent with that of both the HK and SIN nonstudent groups. In considering the difference of probability of a ‘bid’ decision between the two market conditions, the results from the McNemar test show that the respective probabilities of the student group are statistically significant ($p < 0.01$) when number of bidders is considered large ($N \geq 10$). Interestingly, a similar trend was detected among the SIN nonstudent group when $N \geq 8$. However, it can be seen that the sample proportions of ‘bid’ when N is considered small for both the student and SIN nonstudent groups are still higher in recession than in the booming scenario, although there is insufficient evidence to reject the null hypothesis in these particular scenarios. In terms of true change in probability of a ‘bid’ decision, the trend for the student group is again similar to that of SIN nonstudent group. Here, the probability of a ‘bid’ decision may be higher in booming than recession when N is considered small (students: $N \leq 8$; SIN nonstudents: $N \leq 10$) as reflected in

the upper 95% confidence intervals which contain some positive values (students: from -8 to 1%; SIN nonstudents: -10 to 4%).

It is noted that a rather different decreasing trend of sample proportions of 'bid' was observed for HK nonstudent group, where the probability of a 'bid' decision between the two market conditions is statistically significant ($p < 0.01$) for all the number of bidders' scenarios. Also, the probability of a 'bid' decision of HK nonstudent group in booming is lower than recession for all the number of bidders' scenarios with all negative values recorded in the 95% confidence intervals. Nonetheless, the decreasing decision to bid trends as the number of bidders increase for both the market conditions' scenarios are consistent among the student and nonstudent groups as clearly illustrated in Figure 1, in which a second order quadratic regression form is found to be the curve of best fit for all six decreasing trends (R^2 values above 0.9). The best-fit trend lines further demonstrate that the decision to bid trend of the student group is closer to that of SIN nonstudent group when the sample proportions of 'bid' were considered according to market conditions and number of bidders, consistent with the metrics in Table 2.

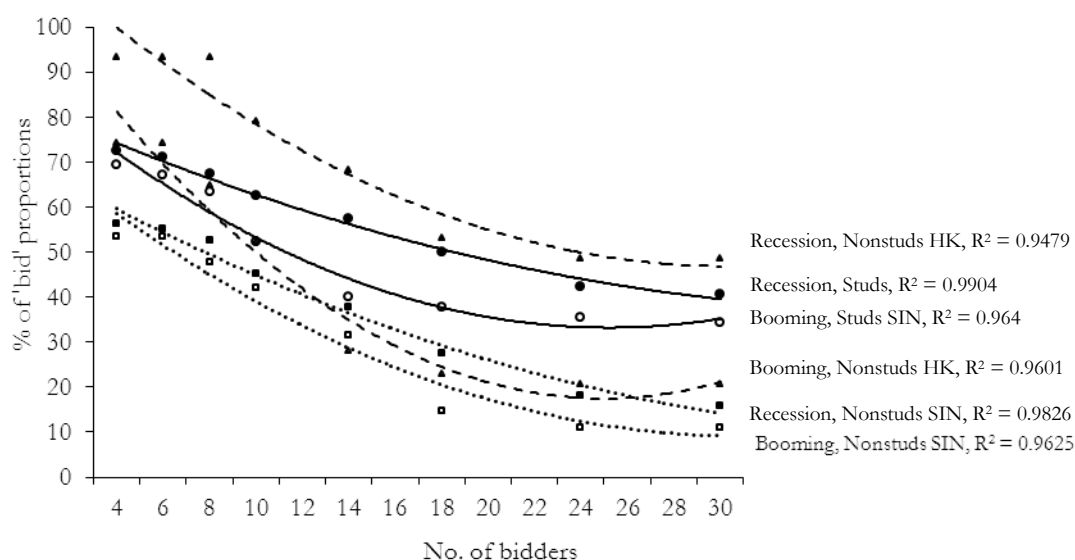


Figure 1: Decision to bid trends of the experimental subjects according to market conditions and number of bidders

Turning into the experimental subjects' mark-up decision, Table 4 shows the various metrics used to compare their mark-up decision according to the two market conditions. Although the percentage of serious bids from the student group is considerably lower in the booming scenario compared to the HK and SIN nonstudent groups, this observation is explainable by the fact that there are different motives for submitting non-serious bids among bidders as identified in the literature (e.g., Drew 1994, Collins and Pasquire 1996). The percentages of serious bids from both the student and nonstudent groups in the recession scenario, on the other hand, are close to each other, and suggest that they recognize the need to bid competitively in order to win jobs.

In terms of the subjects' mark-up size, the mean percentage mark-up as most commonly used measure of central tendency might be of more interest to the readers, but it is influenced by extreme values (i.e., exceptional low and high mark-ups) in the respective datasets. For example, the minimum mark-up sizes recorded for HK and SIN nonstudent groups in recession are as low as around -20% (i.e. a mark-down strategy). In this case, the median is the appropriate

central tendency measure (the mean measure is still included in Table 4 for reference). It can be seen that the median percentage mark-ups for both the student and nonstudent groups are close to each other in the booming scenario (i.e., 8 to 10%). Their median percentage mark-ups in the recession scenario, on the other hand, are rather varied from around 1 to 7%. Although the median measure does not alter when there are extreme values in a dataset, the varied median percentage mark-ups can partly be explained by the high occurrences of negative mark-ups for both the HK and SIN nonstudent groups as indicated by the recorded minimum mark-up sizes. A further examination on the datasets based on the mode does reveal that both the student and nonstudent subjects' mark-up are dominated by zero (HK) and positive mark-ups (i.e., 5% for student and SIN nonstudent group). Here, both the median and mode measures suggest that student and nonstudent experimental subjects were all sensible in their mark-up decision and aiming to make a profit or at least to breakeven.

Table 4: The experimental subjects' mark-up decision according to market conditions

Mark-up decision metric	Experimental subjects		
	Students	Nonstudents - Hong Kong	Nonstudents - Singapore
% of serious bids: booming vs. recession	53% vs. 90%	95% vs. 100%	86% vs. 90%
Booming: MU%			
mean (std. dev.)	10.38 (5.98)	8.53 (6.71)	9.36 (6.71)
median	9.09	8.00	10.00
mode	10.00	0.00	15.00
minimum	-5.00	-10.00	-7.50
maximum	25.00	24.91	25.00
Recession: MU%			
mean (std. dev.)	8.98 (6.36)	0.67 (6.10)	3.08 (5.58)
median	7.27	1.25	3.00
mode	5.00	0.00	5.00
minimum	-5.00	-19.64	-20.00
maximum	25.00	21.05	20.62
Skillings-Mack test p -value	< 0.01	< 0.01	< 0.01

Next, the analysis aimed to test if there is a statistical difference in percentage mark-ups between the two market conditions for each individual subject groups, using a Skillings-Mack test. The Skillings-Mack statistic is a general Friedman-type statistic that can be used in almost any block design with an arbitrary missing-data structure. In this case, the number of bids in each market conditions scenario is different (i.e., there are missing data), which is considered as an unbalanced incomplete block design (see Skillings and Mack, 1981). As shown as the last metric in Table 4, the results show that there is a statistically significant difference in percentage mark-ups between the two market conditions, with higher mark-ups in booming than in recession, for both the student and nonstudent subject groups at $p < 0.01$ level. Combining the test results with the central tendency measure, it is therefore concluded that the mark-up behaviour of the student and nonstudent subject groups are comparable, and so are their decision to bid behaviour.

Discussion

Consistent with the findings in Dyer et al. (1989), the student group in the present work has demonstrated a bidding behavioural pattern similar to the HK and SIN nonstudent groups. It is evident that both the student and nonstudent groups exhibited sufficiently similar decision to bid and mark-up trends, including: (i) the higher number of bidding attempts in recession than in booming market conditions; (ii) the decreasing number of bidding attempts as the number of competing bidders increase; (iii) the strategy to apply negative mark-ups in order to win jobs; and (iv) the higher number of serious and competitive bids with lower mark-ups in recession than in booming market conditions. These bidding trends are, in fact, similar to that of empirical findings based on field data as reported in a number of publications on the nonstudent groups resulted from Oo's (2007) experiment. Although these trends do, to some extent, address the threat to the external validity of the bidding experiment in terms of context effects, the main focus here is on the subject pool effects. Specifically, in answering the question: *do student subjects behave different to nonstudent subjects in an identical construction bidding laboratory experiment?* the results suggest that the subject pool *per se* is not a threat to the external validity of the experiment. While there is no similar work that examines the subject pool effects in experimental studies of construction bidding, the results in the present work are comforting and plausible. The implication here is that those considering similar studies in future could use student subjects in their experiments, especially since empirical analyses of bidding behaviour using field data are limited by the difficulty of obtaining data. However, such use depends largely upon the purpose of an experiment. In addition, the sample selections and analytical methods are important in order to give credibility to reported results (Runeson and de Valence 2015).

Another important implication is that the results address the concern on the effect of monetary incentives in experimental studies as noted in the literature. It is worth noting that there is no monetary incentive in any of these experiments, students or nonstudents. Both subject groups were only offered a copy of the findings in return for their participation. Nonetheless, the results suggest that the student subjects engaged in the experiment seriously, with behavioural patterns that are similar to the nonstudent subjects. Again, this does not mean that there is no need for studies in future to decide on whether or not to introduce monetary incentives in experimental designs. Although offering payoffs to experimental subjects based on subjects' performance is common for experiment studies in economics (see Lee (2007) for a comprehensive review), one should note that it is difficult to decide on the comparable payoffs to both student and nonstudent subjects (Frechette 2015).

In summary, the reported results demonstrate the practicality of using experimental approaches in testing theory in building economics. In addition, they suggest that the use of student subjects in experiments in the area of construction bidding should not be seen as a threat to the external validity of the experiments. This is further supported by some recent bidding experiments that used student subjects, and have reported robust findings as highlighted in the above review. However, it is not the intention here to argue that experimental dataset can be used to provide quantitative results such as point estimates. Here, the reported qualitative results allow one to safely predict bidding behavioural patterns that would arise in the field. Indeed, as contended by Kessler and Vesterlund (2015), there is significantly less (and possibly no) disagreement among researchers on the extent to which the qualitative results of an experimental study are externally valid.

Conclusion

The main conclusion from this replication effort is that the use of student subjects should not be seen as a threat to the external validity of standard experiments on construction bidding

aimed to test theories. With no similar study in the literature, the findings are important, as they address the common concerns of most researchers on the practicality of using student subjects in building economics experiments. It is recognized that the stakes are obviously smaller in an experimental setting, and the decision settings are unavoidably less rich, but the findings should lead to more confidence in the use of experimental approaches in future studies. Indeed, causal knowledge requires controlled variations, and this can only be achieved via active manipulation of variables in an experimental setting. It is suggested that experimental and field findings should be seen as complementary in building economics research, as advocated in other social sciences. Lastly, peer reviewers should obviously question the relevance of an experimental approach for the phenomena being investigated, but not reject papers on the grounds only that use experiments. As with other approaches, they need to consider the various goals of researcher and how they may be best achieved without violating the external validity of the results.

References

- Alm, J., Bloomquist, K. and McKee, M. (2015). "On the external validity of laboratory tax compliance experiments." *Economic Inquiry*, 53(2), 1170-1186. doi: <http://dx.doi.org/10.1111/ecin.12196>
- Camerer, C. (2011). "The promise and success of lab-field generalizability in experimental economics: a critical reply to Levitt and List." Available at SSRN: <http://ssrn.com/abstract=1977749>. doi: <http://dx.doi.org/10.2139/ssrn.1977749>
- Campbell, D. and Stanley, J. (1966). *Experimental and Quasi-Experimental Designs for Research*, Rand McNally, Chicago.
- Collins, S. and Pasquire, C. (1996). "The effect of competitive tendering on value in construction." *RICS Research Paper Series*, 2(5).
- Drew, D. and Skitmore, M. (2006). "Testing Vickery's revenue equivalence theory in construction auctions." *Journal of Construction, Engineering and Management*, 132(4), 425-428. doi: [http://dx.doi.org/10.1061/\(ASCE\)0733-9364\(2006\)132:4\(425\)](http://dx.doi.org/10.1061/(ASCE)0733-9364(2006)132:4(425))
- Drew, D.S. (1994). "The Effect of Contract Type and Size on Competitiveness in Construction Contract Bidding." *PhD thesis*, University of Salford.
- Dyer, D., Kagel, J. H., and Levin, D. (1989). "A comparison of naive and experienced bidders in common value offer auctions: A laboratory analysis." *The Economic Journal*, 108-115. doi: <http://dx.doi.org/10.2307/2234207>
- Falk, A. and Heckman, J.J. (2009). "Lab experiments are a major source of knowledge in the social sciences." *Science*, 326(5952), 535-538. doi: <http://dx.doi.org/10.1126/science.1168244>
- Frechette, G. R. (2015). "Laboratory experiments: professionals versus students." in Frechette, G.R. and Schotter, A. (Eds.), *Handbook of Experimental Economic Methodology*, Oxford University Press, Oxford, UK. doi: <http://dx.doi.org/10.1093/acprof:oso/9780195328325.003.0019>
- Guala, F. and Mittone, L., (2005). "Experiments in economics: External validity and the robustness of phenomena." *Journal of Economic Methodology*, 12(4), 495-515. doi: <http://dx.doi.org/10.1080/13501780500342906>
- Hackemer, G.C. (1970). "Profit and competition: estimating and bidding." *Building Technology and Management*, Dec, 6-7.
- Kessler, J. and Vesterlund, L. (2015). "The external validity of laboratory experiments: The misleading emphasis on quantitative effects." in Frechette, G.R. and Schotter, A. (Eds.), *Handbook of Experimental Economic Methodology*, Oxford University Press, Oxford, UK. doi: <http://dx.doi.org/10.1093/acprof:oso/9780195328325.003.0020>
- Lee, J. (2007). "Repetition and financial incentives in economics experiments." *Journal of Economic Surveys*, 21(3), 628-681. doi: <http://dx.doi.org/10.1111/j.1467-6419.2007.00516.x>
- Levitt, S.D. and List, J.A. (2007). "What do laboratory experiments measuring social preferences reveal about the real world?" *The Journal of Economic Perspectives*, 21(2), 153-174. doi: <http://dx.doi.org/10.1257/jep.21.2.153>
- Oo, B.L. (2007). "Modelling individual contractors' bidding decisions in different competitive environments." *PhD thesis*, Hong Kong Polytechnic University.

- Oo, B.L., Abdul-Aziz, A.R. and Lim, Y.M. (2011). "Information feedback and learning in construction bidding." *Australian Journal of Construction Economics and Buildings*, 11(3), 34-44. doi: <http://dx.doi.org/10.5130/ajceb.v11i3.2173>
- Oo, B.L., Drew, D. S., and Lo, H.P. (2010). "Modeling the heterogeneity in contractors' mark-up behavior." *Journal of Construction Engineering and Management*, 10.1061/(ASCE)CO.1943-7862.0000186. doi: [http://dx.doi.org/10.1061/\(ASCE\)CO.1943-7862.0000186](http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000186)
- Oo, B.L., Drew, D.S. and Lo, H.P. (2007a). "Applying a random-coefficients logistic model to contractors' decision to bid." *Construction Management and Economics*, 25(4), 387-398. doi: <http://dx.doi.org/10.1080/01446190600922552>
- Oo, B.L., Drew, D.S. and Lo, H.P. (2007b). "Modelling contractors' mark-up behaviour in different construction markets." *Engineering, Construction and Architectural Management*, 14(5), 447 – 462. doi: <http://dx.doi.org/10.1108/09699980710780755>
- Oo, B.L., Drew, D.S. and Lo, H.P. (2008a). "A comparison of contractors' decision to bid behaviour according to different market environments." *International Journal of Project Management*, 26(4), 439-447. doi: <http://dx.doi.org/10.1016/j.ijproman.2007.06.001>
- Oo, B.L., Drew, D.S. and Lo, H.P. (2008b). "Heterogeneous approach to modelling the contractors' decision-to-bid strategies." *Journal of Construction, Engineering and Management*, 10.1061/(ASCE)0733-9364(2008)134:10(766). doi: [http://dx.doi.org/10.1061/\(ASCE\)0733-9364\(2008\)134:10\(766\)](http://dx.doi.org/10.1061/(ASCE)0733-9364(2008)134:10(766))
- Oo, B.L., Ling, F.Y.Y., and Soo, A. (2014). "Information feedback and bidders' competitiveness in construction bidding." *Engineering, Construction and Architectural Management*, 21(5), 571-585. doi: <http://dx.doi.org/10.1108/ECAM-04-2013-0037>
- Oo, B.L., Ling, F.Y.Y., and Soo, A. (2015). "Construction procurement: modelling bidders' learning in recurrent bidding." *Construction Economics and Building*, 15(4), 16-29. doi: <http://dx.doi.org/10.5130/AJCEB.v15i4.4653>
- Peng, Y.H., Juan, Y.K. and Chien, S.F. (2006). "Exploring the bidding situation for economically mostadvantageous tender projects using a bidding game." *Journal of Construction Engineering and Management*, 132(10), 1037-1042. doi: [http://dx.doi.org/10.1061/\(ASCE\)0733-9364\(2006\)132:10\(1037\)](http://dx.doi.org/10.1061/(ASCE)0733-9364(2006)132:10(1037))
- Roth, A.E. (1988). "Laboratory experimentation in economics: a methodological overview." *The Economic Journal*, 98, 974-1031. doi: <http://dx.doi.org/10.2307/2233717>
- Runeson, G. (1997a). "The role of theory in construction management research: comment." *Construction Management and Economics*, 15(3), 299-302. doi: <http://dx.doi.org/10.1080/014461997373033>
- Runeson, G. (1997b). "The methodology on building economics research." *Journal of Construction Procurement*, 3(2), 3-18.
- Runeson, G. and de Valence, G. (2015). "The critique of the methodology of building economics: trust the theories." *Construction Management and Economics*, 117-125, 10.1080/01446193.2015.1028955. doi: <http://dx.doi.org/10.1080/01446193.2015.1028955>
- Schram, A. (2005). "Artificiality: The tension between internal and external validity in economic experiments." *Journal of Economic Methodology*, 12(2), 225-237. doi: <http://dx.doi.org/10.1080/13501780500086081>
- Skillings, J.H. and Mack, G.A. (1981). "On the use of a Friedman-type statistic in balanced and unbalanced block designs." *Technometrics*, 23, 171-177. doi: <http://dx.doi.org/10.1080/00401706.1981.10486261>
- Soo, A. (2015). "The effect of construction demand on inexperienced bidders' bidding behaviour." *PhD thesis*, The University of Sydney.
- Soo, A. and Oo, B.L. (2010). "The effect of information feedback in construction bidding." *Australasian Journal of Construction Economics and Building*, 10(1/2), 65-75. doi: <http://dx.doi.org/10.5130/ajceb.v10i1/2.1589>
- Soo, A. and Oo, B.L. (2014). "The effect of construction demand on contract auctions: an experiment." *Engineering, Construction and Architectural Management*, 21(3), 276-290. doi: <http://dx.doi.org/10.1108/ECAM-01-2013-0010>
- Visser, G., Heyne, G., Peters, V. and Guerts, J., (2001). "The validity of laboratory research in social and behavioral science." *Quality and Quantity*, 35(2), 129-145. doi: <http://dx.doi.org/10.1023/A:1010319117701>