The Effects of the Format of Software Project Bidding Processes

Magne Jørgensen, Simula Research Laboratory

Abstract: *This study investigates how differences in format of the bidding process affect companies' bids for software projects. Thirty outsourcing companies from different Asian and European countries participated in the bidding for a software project. A participating company either started with the provision of a bid based on a reduced version of the specification and then continued with a bid based on the full specification (the Increase situation), or started with the full specification and then continued with the reduced one (the Decrease situation). We observed important differences in bids for the same project as a result of different bidding sequences. Our results constitute evidence that in situations similar to the one we studied, the client will typically select a provider with about a 40% lower price in the Decrease situation than in the Increase situation. The difference in bids seems to be explained by whether the first bid was provided on the full or the reduced specification, not by the process of updating, i.e., increasing or decreasing, the bids. We warn against manipulation of the bidding processes to receive lower bids. This increases the risk of over-optimistic bids. Over-optimistic bids frequently have as a consequence "the winner's curse", which leads easily to "the client's curse".*

Keywords: Software project bidding, over-optimism, "the winner's curse"

1. Introduction

Organizations that develop software have, in general, a bad reputation for cost overruns. According to (Moløkken and Jørgensen 2003), the average software project cost overrun is about 30%[1]. Large cost overruns may affect how software developers

[1] The more frequently reported overrun of 189%, as reported by the Standish Group, seems to be based on a flawed study, as documented in Jørgensen, M. and K. Moløkken-Østvold (2006). "How Large Are Software Cost Overruns? Critical Comments on the Standish Group's CHAOS Reports." <u>To appear in Information and Software Technology (Spring 2006)</u>. As pointed out in Grimstad, S. and M. Jørgensen (2005). "Software Effort Estimation Terminology: The Tower of Babel." <u>To appear in Journal of Information and Software Technology</u>. most estimation surveys seem to have

work (Jørgensen and Sjøberg 2001) and lead to low quality software, frustrated software developers and dissatisfied customers. In addition, over-optimistic cost estimates may cause over-optimistic bids and severe losses or low profit for the software providers.

One possible way to reduce or eliminate the strong bias towards over-optimistic cost predictions is to use formal estimation models. Proper formal prediction models are unbiased and affected less by political issues or wishful thinking. There has, consequently, been much research on formal cost estimation models, e.g., the COCOMO-model (Boehm 1981). In spite of this extensive research on formal cost estimation models, empirical studies suggest that the more flexible method "expert estimation" is just as accurate (Jørgensen 2004a). Possibly for that reason, most software companies rely on expert estimation and seldom use formal cost estimation models (Heemstra and Kusters 1991; Hihn and Habib-Agahi 1991; Paynter 1996; Hill, Thomas et al. 2000). It is, for this reason, clearly relevant to study expert judgment-based cost estimation and bidding processes, as we do in this paper.

Expert estimation of cost is, to a large extent. based on non-explicit, unconscious, and non-reviewable processes. As reported in (Jørgensen 2005), the step from understanding the task to quantifying the cost required to complete the task is poorly understood, by both researchers and the estimators themselves. The evidence suggests that the estimators are not aware of the factors that influence their own judgment-based cost estimates. This lack of awareness may, in turn, have the consequence that changes in the estimation process or the inclusion of information irrelevant to the actual cost of developing the software can have unexpected impacts on the estimated costs (Jørgensen and Sjøberg 2004) and, consequently, the bids

An example of process that relates cost estimation and bidding is described in (Kerzner 2003, p 513). As can be seen from that description, the bidding process depends on, but also differs from, the cost estimation process. A company may decide to provide bids lower than the most likely internal costs for strategic reasons, e.g., for the purpose of establishing a good relationship with a client. Consequently, findings related to bidding may not be transferred to an estimation context automatically. On the other hand, our

---

estimation accuracy measurement problems, i.e., the interpretation of the measured 30% cost overrun is not obvious.

own experiences of studying estimating and bidding processes suggest that the correlation between the cost estimate and the bid is high. In fact, as reported in (Grimstad and Jørgensen 2005), several companies do not clearly separate the price (the bid) from the cost estimate, so what they actually estimate is the price to the client. Hence, bidding and the cost estimation process may, despite their different goals and concerns, be rather similar in practice. A further point to note in this context is that an over-optimistic bid is, in our experience, typically based on an over-optimistic cost estimate.

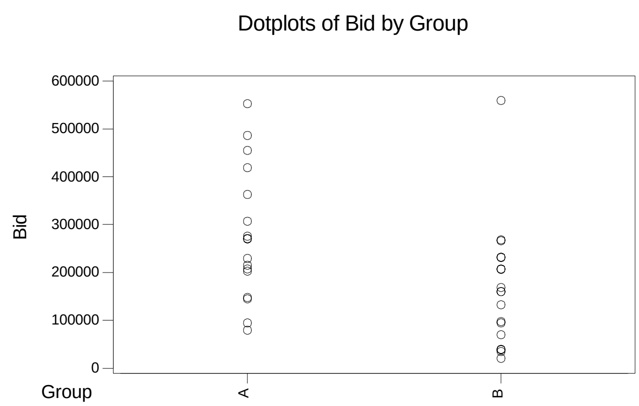
Systematical experimentation and evaluation of the effects of bidding process format on the bids is not feasible in typical bidding contexts where only one bidding process is followed. This means that in most real-world software bidding contexts, any impact that results from irrelevant information and changes in the bidding process goes unnoticed. Laboratory experiments on bidding, on the other hand, may impose a context on the participants that is too artificial to have sufficient external validity. Consequently, the approach chosen for the study described in this paper is to conduct experiments in a real-world setting.

The remaining part of this paper is organized as follows: Section 2 briefly describes a previous study on software bidding processes that motivated the hypotheses and studies described in this paper. Section 3 describes the design and results of the study reported in this paper. Section 4 discusses the importance and implications of our results. Section 5 concludes.

2 Previous Study

In a previous study (Jørgensen and Carelius 2004), we made the surprising observation that software companies going from a price indication based on a preliminary, vague and reduced specification to a more complete and precise specification (Group A companies) delivered, on average, much higher bids than those that started with the final requirement specification (Group B companies). The cost estimates, used as input to the bids, were based on expert judgment (bottom-up estimation methods) or a combination of expert judgment and informal use of estimation models, e.g., a combination of expert estimation and informal use of the use case point estimation method (Anda, Angelvik et al. 2002). The mean bid of Group A companies in our study was about 280 000 Norwegian Kroner (NKR), while that of Group B companies was about 170 000 NKR (1 Euro is, as of August 2005, about 8 NKR). The distribution of bids is displayed in Figure 1. A t-test of difference in mean values gives the p-value 0.01. Cohen's d is 0.8, i.e., the size of effect is large.

Figure 1: Bids (in NKR) of Group A and Group B Companies



We discuss possible reasons, based on bidding results in other domains and the bidders' own reflections, in (Jørgensen and Carelius 2004). The main explanation we provide in that paper is based on the difference in level of uncertainty between Group A and Group B companies, i.e., it is mainly based on the conjecture that software companies in bidding situations tend to over-compensate for perceived uncertainty and under-compensate for decrease in uncertainty. From that discussion, it was clear that more studies on the effect the bidding process were needed to better understand the underlying mechanisms. In particular, we wanted to identify any impact from the increase and decrease in functionality, in situations with similar levels of uncertainty. It was for this purpose that we designed the study described in this paper.

3 The Bidding Experiment

3.1 Research Question and Hypotheses

In the previous study, as stated earlier, we studied the impact of starting with a preliminary, vague, reduced specification on the bids for the more precise, fully specified project. Consequently, the changes in bidding process concerned both the level of specification detail (preliminary, high uncertainty vs more precise, lower uncertainty) and the completeness (reduced vs complete) of the specification. In this study we focus on the completeness factor only. Our main research questions are these:

RQ: *Is there a difference in bids, for the same project specification, between companies going from a reduced specification to a full specification and companies going from a full specification to a reduced specification. If there is a difference in bids, what are the underlying reasons for this difference?*

A major concern when designing the study was to separate (i) the effect of whether the first bid was based on a reduced or full specification (difference in "starting point"), from (ii) the effect of whether the bids were updated as a response to increase or decrease in requirements. Formulated differently, we wanted to be able to examine whether it was differences in the bid updating process or the impact from the first bid that led to changes, if any, in the bids.

To enable this separation of effects we designed a two-phase bidding process and two versions of the requirement specification: reduced and full. A software company was requested to either: (i) Bid for the reduced project (reduced requirement specification) in phase one and then to update the bid in phase two for the full project (full specification), or (ii) Bid for the full project (full requirement specification) in phase one and then to update the bid in phase two for the reduced project (reduced requirement specification).

The following abbreviations are used to describe the hypotheses we wanted to test:

- *Fs: Full software specification*
- *Rs: Reduced software specification (Rs is a subset of Fs)*

- *Ph1: Phase where the first bid is provided (Phase 1)*
- *Ph2: Phase where the update of bid is provided (Phase 2)*
- *GrRsFs: Group of participants with bids based on Rs in Ph1 and Fs in Ph2*
- *GrFsRs: Group of participants with bids based on Rs in Ph1 and Fs in Ph2*
- bid_i*(specification, phase): Bid provided by participant i based on specification Rs or Fs in phase Ph1 or Ph2*

H1: The mean bid when starting with a reduced specification is higher for both the reduced and the full specification.

- **H1a**: $mean(bid_i\ (Fs, Ph2)) > mean(bid_i\ (Fs, Ph1))$
- **H1b**: $mean(bid_i\ (Rs, Ph1)) > mean(bid_i\ (Rs, Ph2))$

H2: The mean increase in bids is larger than the corresponding decrease

- **H2a** (absolute difference): $mean[bid_i(Fs,Ph1)- bid_i(Rs,Ph2)] < mean[bid_i(Fs,Ph2)- bid_i\ (Rs,Ph1)]$
- **H2b** (relative difference): $mean[(bid_i(Fs,Ph1)- bid_i(Rs,Ph2))/ bid_i(Rs,Ph2)] < mean[(bid_i(Fs,Ph2)- bid_i(Rs,Ph1))/ bid_i(Rs,Ph1)]$

The hypotheses are formulated to be consistent with the preliminary explanations and findings in the previous study.

3.2 Experiment Design

Our research institute (Simula Research Laboratory) specified a tool to support our web-based bidding experiments (SIMBID). The specified software tool is computation-intensive and required a user interface that possesses greater than average complexity. There are two versions of this specification, both written by the author of this paper: a full (Fs) and a reduced (Rs) software specification. Fs has all functionality specified in Rs, plus additional functionality. The length of Fs is 27 pages, while the length of Rs is 23 pages. The decision of our research institute to implement Fs or Rs would depend on the increase in cost for the full version, i.e., the bids for both versions included information that we would use to make an informed decision. We required that the tool should be implemented using Java and MySQL. (Interested readers may request copies of the requirement specifications from the author.)

The bidding process was conducted as follows:

1) **Preparation**: We invited 50 outsourcing companies to bid for the development project (SIMBID) and divided them randomly into two groups (GrFsRs and GrRsFs). The invited companies were found through a search on the web applying the search terms "Java" and "outsourcing" and an informal evaluation of competence based on information on their web-pages. The companies varied a great deal in size and experience, from small 3-4 person companies to large companies with more than 1000 employees.

GrFsRs companies were requested to provide their first bids based on Fs and then to update their bids in accordance with the reduced requirements in Rs. GrRsFs companies were asked to begin with Rs and then update their bids in accordance with the increased requirements in Fs. Thirty of the 50 companies accepted the invitation and were willing to bid for our project. All of them were evaluated by us to be serious software providers, based on their proposals and company descriptions.

Thirteen of the companies that accepted the invitation belonged to GrFsRs and seventeen belonged to GrRsFs. Thirteen of the companies were from India, while the other companies were from the Ukraine (4), Russia (3), Pakistan (2), Romania (1), Moldovia (1), Phillipines (1), Thailand (1), Poland (1), Belarus (1), and China (1). We found no obvious bias regarding the allocation of company to the groups, i.e., the distribution of countries, the size of the organizations and the experience level seemed to be similar for the two groups. The analysis of some properties, e.g., the level of experience was, however, based on incomplete information and we cannot exclude the possibility that biases exist due to group allocation.

2) **Phase 1**: The GrFsRs companies received the Fs specification and the GrRsFs companies the Rs specification. All companies were asked to provide a proposal within 14 days. The proposal was to include the following elements:

- A brief description of how the company intended to construct the system, including:
    - the company's understanding of the requirement specification,
    - a high-level design and technical solution
    - a description of the development process
    - a description of the estimation process
- Total cost for the development (the bid)
- The *curriculum vitae* of the person(s) that were supposed to complete the project, including a summary of their relevant project experience.
- A brief description of the company (size, location, history).

3) **Phase 2**: When the first bids were received and evaluated we started Phase 2. The GrFsRs companies were then provided with the following request:

"*To enable us to make a more informed decision about cost/benefits of different amounts of functionality included in the first version of the SIMBID system, could you inform us about the total development cost (your bid) given that the following functionality is not included:* [a description of the reduced functionality followed]". The reduced specification (Rs) was included in the mail to the companies.

The GrRS-FS companies were provided with this request:

"*To enable us to make a more informed decision about cost/benefits of different amounts of functionality included in the first version of the SIMBID system, could you inform us about the total development cost (your bid) given that the following functionality is included:* [a description of the added functionality]". The full specification (Fs) was included in the mail to the companies.

Similarly to Phase 1, the companies were allowed 14 days to update their bids.

4) **Decision + Implementation**: We analyzed the proposals, decided to implement the Fs version of the specification and selected four companies. The selection

of four companies, instead of only one, is related to considerations of research ethics (see discussion in Section 3.4) and research goals extraneous to those described in this paper.

3.3 Results

One of the companies decided not to increase its bid in response to the increase in requirements, i.e., there was no update of the bids. With no update of bid the data was clearly not useful for our research purpose and we removed that company's bid from the analysis. Twenty-nine companies did update their bids. There were obvious differences in the quality of the proposals, but no obvious difference in the average level of quality between the two groups of companies.

The received mean bids for the reduced (Rs) and the full (Fs) specification, and the mean absolute and relative difference in bids, are shown in Table 1. The values are displayed in Figures 2-5.

Table 1: Results

Group	**N**	**mean bid(Rs)**	**Mean bid(Fs)**	**mean absolute difference**	**mean relative difference**
GrFsRs	17	$ 8841	$ 12589	$ 3740	0.52
GrRsFs	12	$ 11195	$ 15482	$ 4287	0.48

Figure 2: Bid Values Based on Reduced Specification (Rs)

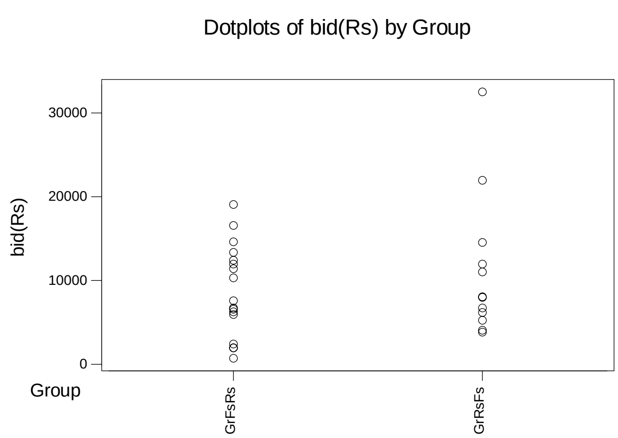


Figure 3: Bid Values Based on Full Specification (Fs)

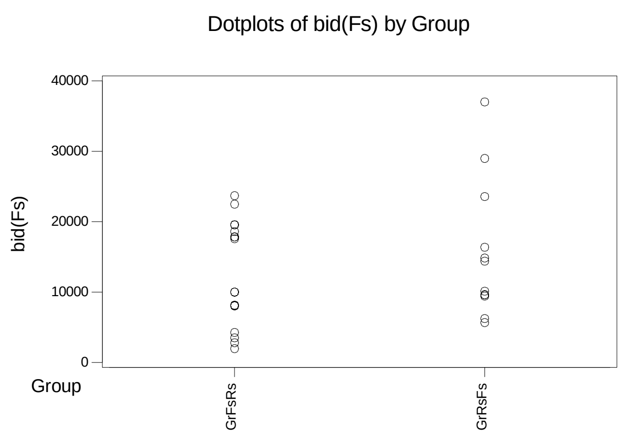


Figure 4: Absolute Difference Between Increase and Decrease in Bid

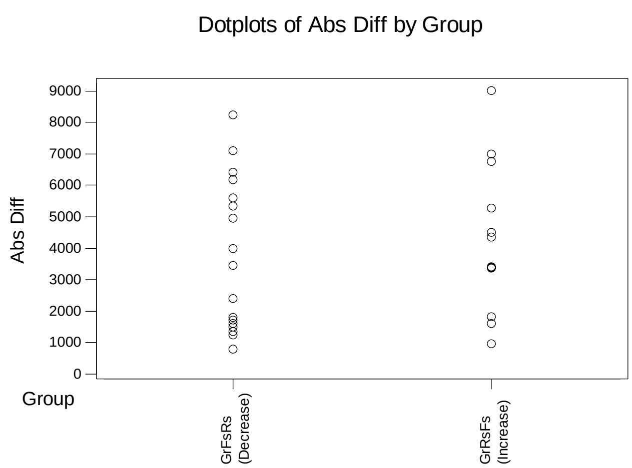
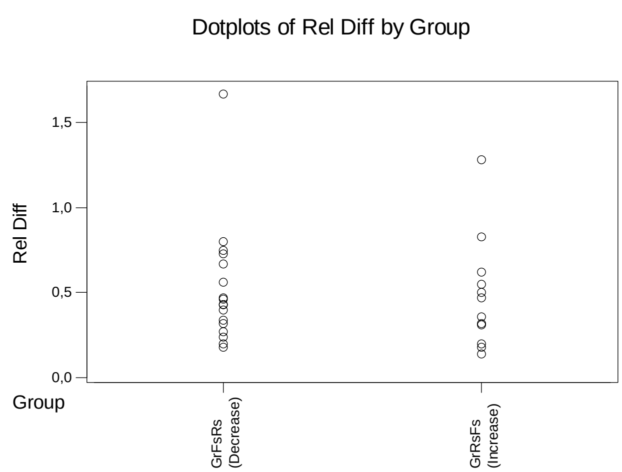


Figure 5: Relative Difference Between Increase and Decrease in Bid



Figures 2 and 3 show a high variance in bids. While the highest bid for the Fs was $37 000, the lowest bid was $2 000. For the Rs, the highest bid was $16600 and the lowest $750! This lowest bid was followed by a proposal that we deemed to be not of sufficiently high quality. The lowest bid with a acceptable-quality proposal was $2 000 for Rs and $ 2 800 for Fs, so even when considering only proposals with acceptable quality the variation in the bids is huge.

This huge variation is particularly amazing, given that we believe the requirement specification to be relatively precise and that none of the companies stated that they had pre-made software components they could use to strongly increase their development efficiency. The variance of bids is, however, similar to that found in the previous study and our experience from bid variance in other software bidding rounds.

Table 2 shows the p-values from one-sided t-test of difference in mean values, the size of effect (Cohen's d), and, the power of the tests for each of the observed differences given a significance level of 0.1.

Table 2: Statistical Tests of Hypotheses

Tests	**Difference in mean values**	**p-value (one sided t-test)**	**size of effect (Cohen's d)**	**Power (significance level 0.1)**
H1a	$ 2354[1]	0.2	0.3 (medium effect)	0.5
H1b	$ 2893[2]	0.2	0.3 (medium effect)	0.5
H2a	$ 539[3]	0.3	0.2 (low effect)	0.3
H2b	0.04[4]	0.6	0.1 (low effect)	0.2

1: Difference in mean bid for Fs for GrRsFs and GrRsFs
2: Difference in mean bid for Rs for GrRsFs and GrRsFs
3: Difference in mean bid increase (GrRsFs) and mean bid decrease (GrFsRs)
4: Difference in mean relative increase (GrRsFs) and mean relative decrease (GrFsRs)

As shown in Table 2, the power values of these tests, for the observed differences in mean values, are not high. This is caused by the large variation in bids within the groups. In particular, the tests of H2a and H2b have low power.

We find it meaningful to interpret the statistical results in Table 2 together with visual inspection of the data plots (Figures 2-5) and the results of the previous study. A correspondence between results in two similar studies can be interpreted as supporting evidence, even if the p-values of each individual study are not interpreted as significant.

We interpret the results displayed in the tables and figures as follows:

- There seems to be a systematic difference between the bids of companies in GrFsRs and GrRsFs. This result is consistent with the results in the previous study and we therefore find that H1a and H1b are, to some extent, supported.
- The absolute increase in bid of GrRsFs companies may be slightly higher than the corresponding decrease in bid of GrFsRs companies (H2a). There seems to be no (or a very small) difference in relative difference between increase and decrease in bids (H2b). The power of the tests for the observed difference is, however, low and it is not likely that our study would find significant differences for the observed size of difference even if the difference had been there. We interpret the findings as providing weak support for H2a and no support for H2b.

The analysis suggests that the main reason for the difference in bids between companies in GrRsFs and GrFsRs is the difference in starting point (whether the first bid was based on Rs or Fs), not the difference between updating a bid as a result of reduced or increased requirement specification. This suggest that updates of bids are *not* based on assessment of the relative change of size of specification, e.g., that there is a 20% increase of functionality and a corresponding 20% increase in bid. Instead, the results suggest that bid updates are based on separate estimation of the cost of each removed or additional feature.

An interesting observation, made both in the previous and present studies, is that the mean bid within a group is of similar size, regardless of whether the starting point was a reduced or full specification. Perhaps, the underlying reason is that there are factors other than size of specification that determine the first bid, while bids are updated more in response to the increase or decrease of specification. It may, for example, be the case that the first bid is, in practice, determined by the bidding policy that "projects of type X (which includes both the reduced and the full project) cost about $ Y". When requested to update, however, the actual difference in requirements must be considered. The underlying reason for the observed bid differences may consequently be that the provision of an initial bid and the updating of that bid are based on quite different cognitive processes and strategic concerns.

The actual cognitive and strategic processes involved in updating bids are, however, not well understood. We were unable to find any studies on this topic, and so research is required in this area.

As described earlier, we selected four companies to implement the software based on the full specification (Fs). We have no actual cost data from these companies and, consequently, no exact information about the level of cost estimation over-optimism of the companies. The data in Table 3 provides some information about the outcome of the implementations. To preserve anonymity, we have not disclosed the actual price, but instead provided a price interval that includes the actual price. All solutions met all specified requirements, although there were large differences related to user friendliness and system robustness.

As can be seen in Table 3, all companies had substantial time overruns due to lack of functionality and errors in the first deliveries. As we interpret it from our dialog with the companies, all companies had underestimated the complexity of the project, which led to their being over-optimistic regarding both time and cost. In particular, the company with the lowest price (Company C) strongly underestimated the complexity and took three times as long as the estimated time to complete the project. That company also delivered the solution with the lowest quality. Other than the high time overrun and low quality of the company with the lowest bid, we observed no clear connection between price, time overrun and quality, e.g., the company with the second lowest price (Company B) had the lowest time overrun and a solution with good quality. There were also differences in how much we, as clients, had to be involved to support the providers. The company with the lowest bid and the lowest quality, for example, required more supporting work from us than the other companies. Clearly, the price is only one element of the total cost of developing software.

Table 3: The Implementations of the Four Selected Companies

Company	**Group**	**Price (Fs) in $**	**Time Overrun (in weeks and % of planned weeks)**	**Quality (as assessed by an independent expert)**
A	GrRsFs	[8000-10000]	12 (86%)	Good
B	GrFsRs	[2000-4000]	8.5 (65%)	Good
C	GrFsRs	[2000-4000]	16 (213%)	Poor
D	GrFsRs	[8000-10000]	9 (70%)	Medium

3.4 Limitations and Research Ethics

Our experiment is a field experiment, in that the companies participated in a commercial bidding process and not a laboratory experiment. This increases the external validity of the results in relation to laboratory experiments. On the downside, it also imposes limitations on the interpretation of the results. The two most important limitations are, we believe, these:

- In field settings there are many variables that we do not control or measure. The identification of cause-effect relationships is therefore more complicated than in laboratory settings. There may, for example, be unobserved variables that are the direct cause(s) of the difference in bids. If that is so, our changes in bidding process will not be the causes of differences in bids, but merely correlating factors.
- The generality of our results is, by and large , unknown. Both the previous and the current experiment were performed using relatively small projects. It is possible that, for example, the same underlying mechanism would lead to other differences in bids for larger projects. There is clearly a need for more studies to understand the effect of bidding processes on projects larger than those studied by us.

However, the fact that we were able to replicate the findings of the previous bidding experiment (database-intensive software application, Norwegian companies) in a different bidding setting (algorithm-intensive software application and foreign outsourcing companies) demonstrates that the results are robust and supports our observation that the format of the bidding process matters.

We tried to further increase our knowledge about the underlying mechanism through two laboratory bidding experiments with similar study design. The participants were software professionals. The results of these studies were in line with those of the two real-life bidding experiments. However, the variation of bids was higher than in the real-life situation and there were many "strange" bids due to the artificiality and time limitations of the experiments. After a careful analysis of the quality of the data, we decided that they did not have the desired quality for inclusion as evidence. The problems we had interpreting results from the laboratory experiments illustrate the fact that using more controlled settings does not necessarily lead to easier identification of cause and effect, e.g., although the more controlled setting reduced the impact from some of the confounding factors, other confounding factors were introduced.

In comparison with other software project bidding situations, the number of bidders may be unusually high. Our reason for this high number of bidders was related to our research needs. For reasons concerning the ethical conducting of research, we decided to select four of the providers to implement the software. This resulted in a more favourable ratio of selected companies to bidding companies, i.e., the ratio was 4/30 = 13%. The selection process was based on business-related criteria, i.e., price and quality. As stated in the hypotheses, we expected the companies in GrRsFs to have higher bids than those in GrFsRs due to the format of the bidding process. To ensure that companies in the GrRsFs group did not suffer because of this, we decided to choose at least one company from that group. It is our opinion that the companies were not disadvantaged by participating in this bidding process, although the bids also constituted input to our research.

4 Discussion of Results

4.1 Practical Size of Effect

One may argue that the observed differences in bids are not very important, e.g., the difference in mean bids for the groups when based on Fs was only 19%. This is not very much compared with, for example, the difference between the lowest and highest bid of the companies. The difference in mean values is, however, misleading for the practical importance of the results. When a client selects a software provider, it typically does so from those with the lowest bids, not from those with average bids. The differences between the two groups (GrRsFs and GrFsRs) turned out to be much larger when only analyzing the lowest bids. Analyzing the bids for the full specification (Fs) we found that:

- If the client selects the provider with the lowest bid, the difference is 43%, i.e., the client would select a bid that is 43% lower in the GrFsRs group than in the GrRsFs group.
- If the client selects from the providers with the three lowest bids, the average difference is 41%, i.e., the client would select a bid that is 41% lower in the GrFsRs group than in the GrRsFs group.

(A similar analysis of the data in the previous bid study yields an even larger decrease: an 86% decrease when selecting the lowest bid, and an average of 70% decrease when selecting from the three lowest bids.)

We found software providers with proposals of sufficiently high quality and relevant experience among the three lowest bids in both groups, i.e., a difference in price of about 40% is a realistic difference between the two bidding processes in this case.

The main reason for this increase in effect when studying the lowest bids instead of the mean bids is, we believe, that the impact of the bidding format on the bids varies with different companies. The companies with the lowest bids in the GrFsRs group are likely to have been biased more than average as a result of the bidding process format. This exemplifies the danger of being too concerned about the mean values in research studies. Analyses of the mean bids are useful for studying a phenomenon itself, but not always useful for analyzing the industrial importance of the effect. The larger effect of the bidding process on those with the lowest bids also emphasizes the danger of selecting a provider among those companies. This danger is discussed in the following section.

4.2 Potential Implications

Our results constitute evidence that there are ways in which a client can design a bidding process to receive lower bids. This section discusses the implication of bids that are too low, and examines whether clients should exploit their opportunities to get a lower price by manipulating the bidding process.

In bidding situations, there is a well-known phenomenon called the "winner's curse". This phenomenon describes a situation where the winner of a bidding round is a company with strongly over-optimistic estimates of the value or the cost. In some cases a company deliberately chooses to provide bids lower than their cost estimates, for, e.g., strategic reasons. This deliberate provision of low bids is not considered to be part of the "winner's curse" phenomenon. Software development companies that have to complete a project based on over-optimistic cost estimates that leads to unexpected losses or low profit are the ones that suffer, thus the term "the winner's curse".

Typical conditions for the winner's curse are the following:

- There is a competitive environment, e.g., when many providers are trying to win a software development contract. The more competitors there are, the more likely it is that the provider that is awarded the contract will suffer from the winner's curse.
- Software providers differ in level of realism when estimating the most likely cost of developing software and/or the benefits of winning the contract. Notice that the winner's curse may be present even if most providers are realistic or even over-pessimistic. It may be sufficient that only one provider is over-optimistic. The greater the uncertainty about cost, the more likely it is that the winner's curse will strike.
- Software development project clients require fixed-price contracts.
- Software development project clients tend to select providers from those with the lowest bids. The stronger the focus on low price, the more likely it is that the "winner's curse" will strike.

These conditions were present in both the previous and the current bidding experiments and may have been the principal cause of the seemingly over-optimistic bids of the four companies that implemented the solution. We discuss the winner's curse in more detail in (Jørgensen and Grimstad 2005).

Our finding that changes in the bidding process can be used to achieve lower bids can be exploited by software clients to reduce the price. The finding can, consequently, add to an already unfortunate situation with a high occurrence of over-optimistic bids and greater risk of the winner's curse. A client that aims at the lowest possible price for a software project may, in accordance with the conditions for "the winner's curse" and our findings:

- Invite many bidders.
- Ask for a fixed price on the project.
- Indicate a strong focus on price.
- Start with a specification of larger size than that which is actually needed. Ask for bid updates for a decreased specification. Repeat this until the desired size of specification is reached.

Assume that it is, as we claim, possible for the client to design the bidding process so that it is highly likely that a low price based on very over-optimistic cost estimates wins the bidding round. Is this bidding process practice beneficial for the client?

We believe that the answer in most cases is a clear *no*. Reasons for this include the following:

- An over-optimistic low, fixed price without fixed delivery, which is the typical situation in software development projects, simply makes the providers look for shortcuts and be less flexible.
- Providers with the lowest bids have been found to have less relevant experience than those with medium or high bids (Jørgensen 2004b). Even worse, providers with appropriate experience and good skills may perform worse as a consequence of a bid that is too low. For example, we have observed that estimates that are too low may lead to too little work on design and user interaction as a result of an attempt to adhere to the initial,

over-optimistic cost estimates and plans (Jørgensen and Sjøberg 2001). This again, may lead to chaos, delays and poor quality.

- More work for the clients/requires greater competence. (More testing needed, more iterations due to quality assessment.)

In short, the strategy of attempting to get the lowest possible price is rarely a good one in software development. A low price strongly increases the risk of low quality, even when the provider normally delivers high quality software. The winner's curse can easily become the "client's curse". In most cases, we believe, both parties would be better off finding a higher price, which would enable delivery of the required quality.

5 Conclusions

Software project bids are, typically, based on expert judgments of price-to-win and most likely cost. Expert judgment is, at root, an intuition-based process that can be influenced by both relevant and irrelevant factors. The influence of these factors is not controlled by the expert and awareness about the influencing factors is probably low (Jørgensen 2005). This means that software costing and bidding processes should not be seen as mechanical, rational processes, but instead as processes that can easily be influenced by irrelevant factors. We believe that the study reported in this paper has the potential to contribute to higher awareness of the vulnerability of the bidding process, through our results on how seemingly neutral changes in bidding process can lead to substantial changes in bids provided by software providers, and,how much different starting points of the bidding process may affect the bids.

In addition, we describe a practice (that we do not recommend) for how to achieve bids based on over-optimistic cost estimates. The main message here is that there are ways to design the bidding process that increase the probability that prices will be too low, including an exploitation of the results from our experiments, but that the effect of selecting a provider subject to the winner's curse also increases the risk of the client's curse.

It is difficult to design bidding processes that lead to a good price for the quality required and selection of the provider most likely to deliver the required quality. Current bidding process practice by clients of software development projects may, as suggested in (Jørgensen and Grimstad 2005), be based on poor awareness of the risks involved in selecting companies with over-optimistic bids. Our results suggest that an important key towards improving the realism of bids (and, implicitly, the accuracy of cost estimation) is to improve the format of the bidding process. This process is typically controlled by the software clients. So far, research has focused on the cost estimation processes of the software providers. Our results suggest that there is a strong need for more research into the impact of the bidding processes that lie within the control of the clients.

Acknowledgment: We thank Siw Hove, Stein Grimstad and Hans Christian Benestad, all of them working at Simula Research Laboratory, for their invaluable contributions to the work on this paper.

References

Anda, B., E. Angelvik, et al. (2002). Improving estimation practices by applying use case models. Product Focused Software Process Improvement. 4th International Conference, Rovaniemi, Finland, Simula Res. Lab. Lysaker Norway.

Boehm, B. W. (1981). Software engineering economics. New Jersey, Prentice-Hall.

Grimstad, S. and M. Jørgensen (2005). "Software Effort Estimation Terminology: The Tower of Babel." To appear in Journal of Information and Software Technology.

Heemstra, F. J. and R. J. Kusters (1991). "Function point analysis: Evaluation of a software cost estimation model." European Journal of Information Systems **1**(4): 223-237.

Hihn, J. and H. Habib-Agahi (1991). Cost estimation of software intensive projects: A survey of current practices. International Conference on Software Engineering, IEEE Comput. Soc. Press, Los Alamitos, CA, USA.

Hill, J., L. C. Thomas, et al. (2000). "Experts' estimates of task durations in software development projects." International Journal of Project Management **18**(1): 13-21.

Jørgensen, M. (2004a). "A review of studies on expert estimation of software development effort." Journal of Systems and Software **70**(1-2): 37-60.

Jørgensen, M. (2004b). "Top-down and bottom-up expert estimation of software development effort." Information and Software Technology **46**(1): 3-16.

Jørgensen, M. (2005). The "Magic Step" of Judgment-Based Software Effort Estimation. International Conference on Cognitive Economics, New Bulgarian University, Sofia, Bulgaria.

Jørgensen, M. and G. J. Carelius (2004). "An Empirical Study of Software Project Bidding." IEEE Transactions of Software Engineering **30**(12): 953-969.

Jørgensen, M. and S. Grimstad (2005). Over-optimism in Software Development Projects: "The winner's curse". IEEE CONIELECOMP, Puebla, Mexico.

Jørgensen, M. and K. Moløkken-Østvold (2006). "How Large Are Software Cost Overruns? Critical Comments on the Standish Group's CHAOS Reports." To appear in Information and Software Technology (Spring 2006).

Jørgensen, M. and D. I. K. Sjøberg (2001). "Impact of effort estimates on software project work." Information and Software Technology **43**(15): 939-948.

Jørgensen, M. and D. I. K. Sjøberg (2004). "The impact of customer expectation on software development effort estimates." International Journal of Project Management **22**: 317-325.

Kerzner, H. (2003). Project management: a systems approach to planning, scheduling and controlling, Wiley.

Moløkken, K. and M. Jørgensen (2003). A review of software surveys on software effort estimation. International Symposium on Empirical Software Engineering, Rome, Italy, Simula Res. Lab. Lysaker Norway.

Paynter, J. (1996). Project estimation using screenflow engineering. International Conference on Software Engineering: Education and Practice, Dunedin, New Zealand, IEEE Comput. Soc. Press, Los Alamitos, CA, USA.