A Strong Focus on Low Price When Selecting Software Providers Increases the Likelihood of Failure in Software Outsourcing Projects

Magne Jørgensen
Simula Research Laboratory &
University of Oslo
Martin Linges vei 17, Fornebu
Norway
+47 92433355
magnej@simula.no

ABSTRACT

Context: Bidding rounds are frequently used to select competent and cost-efficient providers for software projects. **Objective**: We hypothesize that emphasizing low price when selecting software providers in such bidding rounds substantially increases the likelihood the project will fail. **Method**: The hypothesis is tested by analyzing a dataset of 4,791,067 bids for 785,326 small-scale projects registered at a web-based marketplace connecting software clients and providers. **Result**s: We find evidence supporting our hypothesis. For example, selecting providers with bids 25% lower than the average bid is connected to a 9% increase in the frequency of project failures for the same level of provider skill. In addition, we found that clients emphasizing a low price, on average, selected providers with lower skill levels. This decrease in provider skill level further strengthened the negative effect of a strong focus on low price on project failures. For example, selecting a provider with a 15% failure rate for previous projects instead of 5% increased the failure rate by 33%. **Conclusion**: We interpret the findings to suggest that a client may substantially reduce the likelihood of project failure by reducing the emphasis on low price when selecting a provider.

Categories and Subject Descriptors

D.2.9 [**Management**]: Cost estimation, productivity, time estimation

General Terms

Management, Measurement, Economics

Keywords

Software cost estimation, bidding, provider selection, project failures, adverse selection, winner's curse, Dunning-Kruger effect

1. INTRODUCTION

Surveys of software projects show that many software projects have cost and time overruns, are delivered with quality problems, and are cancelled. Sauer et al. [1], for example, found in a survey of UK projects that 9% of the software projects were cancelled before being completed. Clearly, starting a software project that does not lead to useful products is a waste of resources and should be avoided. In this paper, we examine to what degree clients' focus on low price when selecting software providers affects the likelihood of project problems and failures. The effect of a strong focus on low price when selecting providers is, as far as we have experienced when reviewing relevant studies in [2], a neglected, potentially important, reason for software project failures.

Our hypothesis motivated and tested in this paper is as follows:

A strong focus on low price when selecting providers increases the likelihood of project failure.

The remainder of the paper is organized as follows: In Section 2, we describe three causal mechanisms potentially connecting a strong focus on low price with a higher likelihood of project failures. In Section 3, we describe the properties of the software outsourcing context and the dataset we use to test the hypothesis. In Section 4, we analyze the dataset. In Section 5, we discuss the limitations of the results, and in Section 6, we conclude the paper.

2. MOTIVATION

Providers participating in a bidding round inform the client about the prices they require and their competences in developing the software system. Based on this information, the client selects a provider to complete the work. Two complicating elements when selecting among competing providers based on this type of information are the following:

- Not all aspects of software are equally easy to specify. For example, it is seldom possible to completely specify the usability and quality aspects of software. Consequently, providers may produce quite different software solutions based on the same requirement specification. The contracting situation thus typically is a situation with a fixed-price for a, to some extent, not fixed product, and the client will not know exactly what it will receive for the price paid.
- Information about the software development skill of the provider, e.g., information related to education, training, and previous projects, may be difficult to translate into knowledge about the provider's ability to produce user-friendly, high-quality software on time for the project required by the client. Consequently, there may be high

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
EASE '13, April 14 - 16 2013, Porto de Galinhas, Brazil
Copyright 2013 ACM 978-1-4503-1848-8/13/04…$15.00

uncertainty regarding the provider skill level relevant for the client.

If low bids always were consequences of high skill, i.e., a consequence of high productivity, there would be no harm in a strong focus on low price when selecting providers. Frequently, however, this is not the case, and a client may have to select between a provider with a lower price but not the best skill and a provider with a higher price and higher skill. People tend to emphasize what is perceived as accurate knowledge (in our case, the price of the project) compared to what is inaccurately known (in our case, the provider skill) in decision situations with uncertain information; see, for example, [3]. A consequence of this tendency, which sometimes is termed the *adverse selection effect*, is that when there is a substantial degree of uncertainty about the skill, highly skilled providers are not selected as often as they would be in situations with accurate information about price, quality, and skill. Not surprisingly, selecting less skilled providers may increase the risk of project failure.

A strong emphasis on low price may also increase the risk of project failure due to two other phenomena: *the winner's curse* and *the Dunning-Kruger effect*. We briefly describe the relevance of these phenomena for software development contexts in Sections 2.1 and 2.2. In Figure 1, the connection between emphasis on low price and higher risk of project failure is summarized through three mechanisms: i) adverse selection, ii) the winner's curse, and iii) the Dunning-Kruger effect.

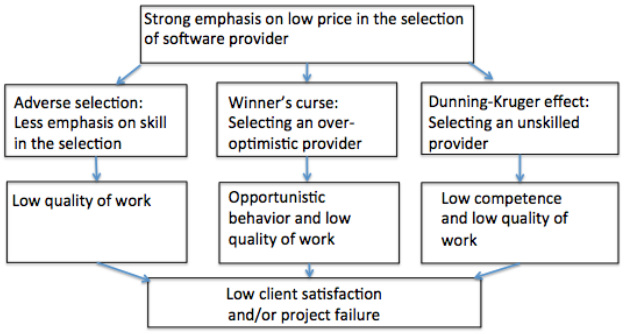

Figure 1. Connection between emphasis on low price and higher likelihood of project failure.

2.1 The winner's curse

The winner's curse mechanism, translated into our software project bidding context, is related to the situation in which a bidder is more likely to win a bidding round, assuming a client emphasizing low price, when being overly optimistic about the amount of work required to complete the development project and there is uncertainty in the estimates. In the study reported in [3], we found that the expected degree of effort overrun in a context where the software providers, on average, were unbiased; the mean estimated effort of the providers equals the mean actual effort, under certain assumptions (such as normally distributed estimated and actual effort), can be expressed as follows:

$$\textit{Expected \% effort overrun of selected provider} = \left(1 - \rho_{est,act}\frac{\sigma_{act}}{\sigma_{est}}\right)(1 - w),$$

where $\rho_{est,act}$ is the correlation between the estimated effort and the actual effort (the estimation accuracy), $\frac{\sigma_{act}}{\sigma_{est}}$ is the ratio of the standard deviation of the actual and estimated effort, and *w* is a measure of the client's emphasis on low price (defined as $w = \frac{estimated\ effort\ of\ selected\ bid}{mean\ estimated\ effort\ of\ all\ bids}$). As can be seen in the expression, the stronger the emphasis on low price (lower *w* value), the higher the expected percentage effort overruns, i.e., the stronger the winner's curse effect. Take, for example, a situation in which the correlation between the estimated effort and the actual effort ($\rho_{est,act}$) is 0.6, the estimated effort varies twice as much as the actual effort ($\frac{\sigma_{act}}{\sigma_{est}} = 0.5$), and the client selects a bid that is 50% of the mean bid of all the providers (corresponds to a *w* value of around 0.5 assuming a strong correlation between effort estimates and bids). We should then expect to select a provider that has underestimated the actual effort with (1-0.6*0.5)(1-0.5) = 0.7*0.5 = 35%, a quite overly optimistic provider. This example gives an indication of the expected size of the winner's curse in this scenario. We evaluated the validity of the expression on a real-world dataset and found that the expression predicted the actual effort overrun quite accurately. Notice that the expected degree of overoptimism in this scenario assumes no underlying bias toward overoptimism among the software providers, just uncertainty in estimating the effort and a focus on low price in the selection of provider. Any underlying tendency toward overoptimism among the providers will further add to the expected degree of effort overrun. If, for example, the providers' effort estimates, on average, are 30% too low, this scenario would lead to the selection of a provider with an expected 65% underestimation of effort.

In some domains, a fixed-price contract may imply that all the production risk is on the provider. Although there may be a winner's curse on the provider side, this causes no problems for a client who has a contract accurately stating what is to be produced, at what price, and when to deliver the product. This is, however, not necessarily the case in software development where the products are typically incompletely specified. A client will suffer if the product has to be accepted with lower-than-expected quality due to incomplete specification and a provider tries to avoid financial losses and maintain the planned delivery schedule through quality reductions. This latter phenomenon is frequently termed *opportunistic behavior* or *moral hazard* in economics and is a well-documented phenomenon; see, for example, [4]. In our context, these phenomena may imply that overly optimistic effort estimates and bids influence software providers to work differently, typically with lower quality, compared to situations with more realistic effort estimates. Although mechanisms that make providers behave less opportunistically in situations with the winner's curse exist, for example, the wish to get new projects from the same client, the winner's curse increases the risk of such behavior. In short, the winner's curse may lead to an increased risk of the client's curse (project failures or low-quality products).

2.2 The Dunning-Kruger effect

The Dunning-Kruger effect says that unskilled people overestimate their abilities more than those with higher skills, who in some contexts even underestimate their skills [5]. Translated into our software development context, this means that we may have a situation in which those with the lowest bids tend *not* to be those with the highest skills but rather are those with the lowest skills. In our studies related to this effect, we have had mixed

observations. In a context with experienced software developers performing the same maintenance tasks, we found that a lower estimate was connected with better development skill [3], an observation *not* in accordance with the Dunning-Kruger effect. In an outsourcing context with larger variance in provider skill, however, we found that a lower bid was connected with less experience with similar projects [6], observations in accordance with the Dunning-Kruger effect. The Dunning-Kruger effect seems to be context-dependent and, perhaps, to increase in importance when the complexity of the task increases.

3. THE PROJECT DATABASE

The dataset we use to test the hypothesis consists of bids and projects using the services of vWorker (now merged with Freelancer). vWorker is a web-based marketplace that connects clients and providers. The providers are typically single software developers or smaller outsourcing companies located in low-cost countries. The services offered by the marketplace include the means to search for and invite developers with appropriate skills, support for providers to bid on a project, arrangements that ensure that the developer is paid when the work is completed and that the client does not have to pay if the work is of too low quality to be accepted, processes for managing disagreements/negotiation between clients and developers regarding payment or quality of work (arbitration processes), skill tests for the developers, and skill ratings for the providers and the clients based on previous clients' ratings and provider satisfaction. Typically, the developers and the clients never physically meet and do all their communicating through the communication means provided by vWorker and other Internet-based communication means.

Especially interesting for the analysis in this paper is the price and skill information vWorker offers the client, to support the selection of a software provider:

- The price offered to complete the specified work by the bidding providers.
- Previous clients' satisfaction with a provider. This includes aggregated satisfaction measures and individual ratings with text comments from each previous client.
- The number of previous projects by the providers, including completed and cancelled projects.

Figure 2 displays an example of a subset of the information the client has available when choosing among bidders. The names of the bidders have been hidden for anonymity reasons.

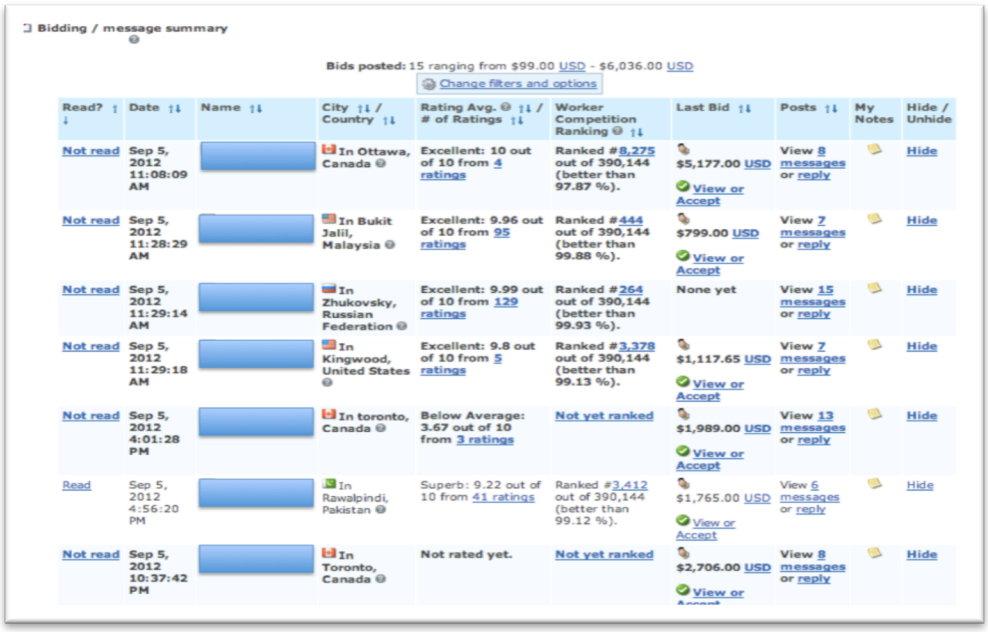

Figure 2. Example of information about the bids and bidders at vWorker.

The project database included in our analysis has the following characteristics:

- Project data registered between May 2001 and October 2012
- Number of projects: 785,326
- Number of bids placed: 4,791,067
- Mean number of bids per project: 6.1
- Number of software provider nationalities: 187
- The ten largest provider countries (sorted by decreasing number of projects): India, the US, Romania, Pakistan, the UK, Russia, Ukraine, Canada, Bangladesh, and the Philippines
- Number of client nationalities: 177
- The ten largest client countries (sorted by decreasing number of projects): the US (with more than 50% of the projects), the UK, Australia, Canada, India, Germany, the Netherlands, Israel, Sweden, and France
- Cost of selected work in started fixed-price projects: Between $1 and $30,000, with a mean of $146 and a median of $50.

4. ANALYSIS OF THE DATASET

4.1 The Measures

Measures of project outcome:

- **Client Satisfaction (CS)**: The client's rating of the provider's performance on a project. The client's rating ranges from 1 (horrible) to 10 (excellent), where the values 1–3 are reserved for failed projects. Employees from vWorker can give the provider the score -3 if there has been improper behavior on the provider side.
- **Project Failure (PF)**: For our analyses, we define a project as a failure (PF=1) if the client gives the provider a satisfaction score of 3 or lower or the project is cancelled after being subject to an arbitration (negotiation) process between the client and the provider. A project is otherwise defined as a non-failure (PF=0).

Measures indicating the skill of the selected provider at the time of the bidding:

- **Mean Client Satisfaction (mCS)**: Mean client satisfaction (CS) on the provider's previous projects.
- **Proportion Project Failure (pPF)**: Proportion of the provider's previous projects that failed (projects with PF = 1).

A high proportion (72%) of the projects received the top client satisfaction score of 10. Many of those top scores hide a not fully satisfied client, and the difference between, for example, an mCS of 9.9 and 9.8 is not likely to give much information about the underlying skill. Therefore, for our regression model, we divided the mCS values into three broad categories (mCS-Cat): "Good" (better than the median (9.77) client satisfaction on previous projects), "Medium" (mean client satisfaction on previous projects between 8.0 and 9.77), and "Poor" (mean client satisfaction on previous projects lower than 8.0). This categorization is based on reading what clients perceive as a good, medium, and poor provider performance. We code the mCS-intervals Good=1, Medium=2, and Poor=3.

Measure of the client's focus on low price when selecting a provider:

- **Focus on Low Price (FLP):** The relative difference between the selected price and the mean price of the bids:

$$FLP = \frac{mean\ price\ of\ all\ bids - price\ of\ selected\ bid}{\max\left(mean\ price\ of\ all\ bids, price\ of\ selected\ bid\right)}$$

A high positive FLP value indicates a strong focus on low price by the client. We divide by the maximum of the mean price for all bids and the price of the selected bid instead of the mean price of all bids to avoid an asymmetric distribution of FLP (to avoid a situation in which the maximum FLP is 1, the FLP when the selected bid equals the mean bid is 0, and the minimum FLP has a value much lower than -1). The decision to use the maximum instead of the minimum function in the denominator of the formula is based on the wish to limit the range of the values, which may ease the interpretation. FLP has a value between -1 and 1.

For the analysis, we also divided the FLP values into four categories with about the same number of observations. We used the Q1 (the lower quartile boundary value), the median, and the Q3 (the upper quartile boundary value) as the boundary values for the FLP category intervals. Consequently, FLP category 1 includes projects with the least focus on low price in selecting providers, while FLP category 4 includes projects with the strongest focus on low price. This measure is termed FLP-Cat.

Measures to support the analysis of consequences of different client priorities on low price and skill:

- **Distance to Best mean Client Satisfaction (DBmCS)**: The difference between the mean client satisfaction (mCS) of the bidder with the best bidder and that of the selected bidder.
- **Distance to Lowest mean Project Failure (DLpPF)**: The difference between the proportion of failures (pPF) of the bidder with the lowest bidder and that of the selected bidder.

Table 1: Summary of the measures

Measure	Description	Values
CS	Client's satisfaction with provider's project performance	1 (worst) to 10 (best); in rare cases, use of the value -3
PF	Project failure, defined as cancelled or with client satisfaction of 3 (poor) or less	1=failed project 0=non-failed project
mCS	A project provider's mean client satisfaction rate for previous projects	1 (worst) to 10 (best)
mCS-Cat	Categories of mCS values	1 (good) = mCS between 9.77 and 10 2 (medium) = mCS between 8.0 and 9.77 3 (poor) = mCS lower than 8.0
pPF	A project provider's proportion of failures on previous projects	0 (best) to 1 (worst)
FLP	A client's focus on low price, based on how much the bid of the selected provider deviates from the mean bid	-1 (least price focus) to 1 (most price focus), gives positive (negative) values when a bid lower (higher) than the mean bid is selected
FLP-Cat	Category of client's price focus (FLP) when selecting a provider	1 (very low): FLP between -1 and -0.11 2 (low): FLP between -0.11 and 0.06 3 (medium): FLP between 0.06 and 0.34 4 (high): FLP between 0.34 and 1.0
DBmCS	Distance to best mean client satisfaction on previous projects (among the bidders for the project)	0 (the provider with the best mCS is selected) to 10 (the maximum distance to the provider with the best mCS)
DLpPF	Distance to lowest failure rate on previous projects (among the bidders for the project)	0 (the provider with the lowest pPF is selected) to 1 (the maximum distance to the provider with the lowest pPF)

To illustrate the use of the measures, assume that a project, which is our unit of study, is to be started and that the client has selected a provider with a bid of $1,000 among a total of four bids. The selected provider has completed 20 projects. The client satisfaction (CS) on these 20 projects is 10 for 18, 9 for one, and 1 for one project (that project is then defined as failed, with PF=1). This gives the selected provider of the project a mean client satisfaction (mCS) of 9.5 [(18*10+9+1)/20], which belongs to the medium mCS category (mCS-Cat=2), and a proportion of project failures (pPF) of 0.05 (1/20). The four bidders offer 500, 1,000 (the selected bid), 2,000, and 2,500 (the mean bid is 1,500), have mean client satisfaction values (mCS) of 9.3, 9.5 (the selected provider), 9.8, and 9.8, and failure rates (pPF) of 0.1, 0.1 (the selected provider), 0.0, and 0.0. These values give a client focus on low price (FLP) of 0.33 [(1500-1000)/max(1500,1000), a distance between the selected and the best mean client satisfaction (DBmCS) of 0.3 (9.8–9.5), and a distance between the selected and the lowest failure rate (DLpPF) of 0.1 (0.1–0.0).

4.2 The Dataset

The main aim of this study is to assess the effect of the client's focus on low price when selecting among providers in bidding round contexts. Thus, we excluded all projects that had only one bidder or were smaller than $100.

Excluding projects with only one bidder was motivated by the purpose of our analysis. We cannot study the impact of a price focus when selecting providers when there is only one bidder, i.e., when there is no provider selection. In addition, in bidding rounds with only one bid, the client had collaborated with the bidding provider in 76% of the cases and the failure rate was only 7%. Comparing this with the bidding rounds with more than one bidder, in which the client and the selected provider have collaborated earlier in only 11% of the cases and the failure rate is 28%, we see a substantial benefit of selecting a provider with which the provider has previously collaborated.

We excluded projects smaller than $100 for two reasons. First, our price focus measure FLP is a ratio-based measure and could give rather misleading values for the smallest projects, e.g., where a difference between a price of $10 and $20 is perceived by the client to be minor, but our ratio-based measure will give that the difference is very large. Second, while there is a substantial increase in the failure rate as the project increases to about $100, the effect of project size on project failure seems to decrease after that (the failure rate of projects up to $100 is 12%, while the failure rates of the categories $100–500, $500–1000, and larger than $1,000 are all in the range 18%–24%). Avoiding projects smaller than $100, consequently, eases the analysis, as it is less likely that the observed differences in failure rates are related to project size and not to measurement and project size challenges.

After projects with only one bidder or price less than $100 were excluded, 111,487 projects remained for the analysis. Notice that the data about a provider's previous scores for client satisfaction and failure rates are based on all 785,326 projects, not only those included in the reduced dataset.

Table 2 provides information about the mean, Q1, the median, and Q3 of the measures for the subset of 111,487 projects. The measures mCS and pFP reflect the skill values for the providers selected for the projects, while FLP is a measure of the price of the selected bid compared to the average of all the bids.

Table 2: Project outcome, skill, and FLP characteristics

Measures		Mean	Q1	Median	Q3
Project outcome	CS	8.44	9.00	10.0	10.0
	PF	0.28	n.r.	n.r.	n.r.
Skill of selected provider	mCS	9.36	9.31	9.77	10.0
	pPF	0.12	0.00	0.07	0.16
Focus on low price	FLP	0.12	-0.11	0.06	0.34

Several observations from Table 2 are of interest for our analysis:

- Client satisfaction (CS) is highly skewed toward higher values. Figure 3 shows the distribution of client satisfaction scores, where as much as 72% of the projects were given the maximum satisfaction score of 10. This indicates that clients frequently did not distinguish between projects for which the provider did a good job and those for which the job was outstanding.
- The mean PF value shows that 28% of the projects failed (was cancelled or had a client rating of 3 or less). This is substantially higher than the average of all projects conducted in this marketplace (14%), and a consequence of our filtering of projects (a consequence of our excluding the smallest projects and most of those in which the client and provider had collaborated on an earlier occasion). Notice that a failed project sometimes was restarted with a new provider until successful. The same client request for a project may consequently have had failed and successful (a maximum of one) completions as outcomes.
- The typical (median) provider selected by a client had a mean client satisfaction score on earlier projects of 9.77 and had failed in 7% of the previous projects. In 25% of the projects, a provider with a mean client satisfaction score of 9.31 or lower was chosen.
- The FLP measure shows that the price of the selected provider in 25% of the cases (the Q3 value of FLP) was at least 34% lower than the average bid, in 50% of the cases (median value of FLP) 6% lower than the average bid, and in 25% of cases at least 11% higher than the average bid. Further examination of the data shows that the lowest-priced bid was selected in 24% of the projects.

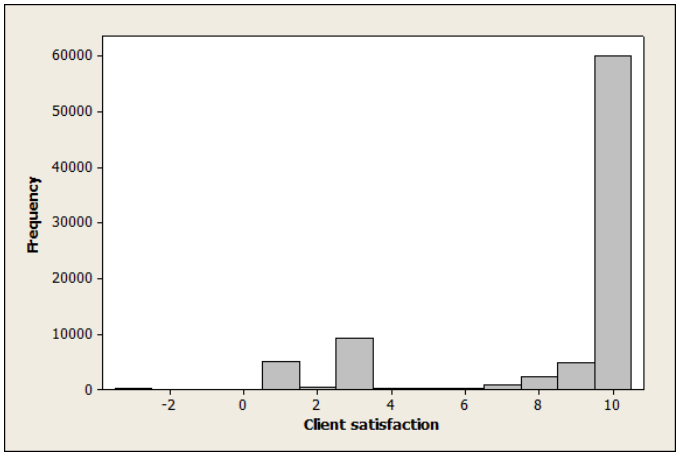

Figure 3. Client satisfaction scores.

4.3 The Analyses

The usual measures of statistical significance become less meaningful as indicators of interesting relationships in our analyses of more than 100,000 projects, since even the smallest effect size will become statistically significant. Therefore, we are more interested in effect sizes than in statistical significance in our analysis. In addition, our analysis differs from analyses in which we have a sample of a population and try to infer the characteristics of the larger population. In our case, we more or less have the population of interest in our dataset, and we need non-statistical argumentation to generalize to other populations. This means that we should interpret the measured differences and effects as characteristics of the observed population rather than use them to infer, with statistical means, from the observed to a larger population.

The analyses in this paper are related to testing the hypothesis of how much the likelihood of failure is affected by a client's focus on low price when selecting a provider. We separated this into analyses of the following:

- How much does the focus on low price *in itself*, given the same level of provider skill, affect the frequency of project failures? This question is related to the *winner's curse* effect, that a focus on low price leads to a higher risk of selecting an overly optimistic provider, which in turn increases the risk of low-quality (opportunistic) behavior.
- To what extent does the focus on low price lead to selecting providers with lower skill, and, as a consequence, to a higher frequency of project failure? This question is related to the Dunning-Kruger effect *and* the consequence of that a higher emphasis on low price in many cases leads to a provider with lower skill than otherwise as a consequence of the *adverse selection* effect.

For our analyses, we used a binary logistic regression model with project failure (PF) as the dependent variable and the focus on price (FLP) together with the provider skill variables (mCS and pPF) as independent variables. The interpretation of a binary logistic regression-based odds ratio in our analysis was that it indicates the increase or decrease in likelihood of project failure given one unit change in the independent variable, assuming that all other variables are held constant. One unit change, however, is a very large change when using our FLP and pPF measures. We therefore multiplied the FLP and pPF values by 10 and made the odds ratios reflect what happens if we changed the FLP and pPF variables by 0.1.

The following is an excerpt of the output of the binary logistic regression analysis (Somer's d = 0.33):

			Odds	95% CI	
Predictor	Coef.	P	Ratio	Lower	Upper
Constant	-2.08	0.00			
FLP*10	0.035	0.00	1.04	1.03	1.04
mCS-Cat	0.361	0.00	1.44	1.40	1.48
pPF*10	0.288	0.00	1.33	1.32	1.35

Interesting implications of the above odds ratios include the following:

- For the same level of previous skill (the same mCS and pPF), there was a higher frequency of failure with increased price focus by the client (higher FLP). The odds ratio of 1.04 suggests, for example, that there is a 4% increase in the likelihood of project failure when going from selecting the average-priced bid to a bid that is 10% lower than the average bid, when assuming the same level of provider skill. Although this effect is not large, it supports our hypothesis that selecting a provider with a low bid price can have a negative impact on project failure in itself, i.e., even when the level of skill of the selected provider is good, the overoptimism have a negative effect. If, for example, we change the FLP value to FLP*4 and re-run the regression analysis, we get the odds ratio 1.09; we find that there is a 9% increase in the likelihood of failure when selecting a bid that is 25% lower than the average bid. Thirty-two percent of the clients in our dataset selected a provider with a bid that was at least 25% lower than the average bid.
- Previous client satisfaction, assuming our categories (mCS-Cat) Poor, Medium, and Good, explained much of the variance in project failure. For example, there was a 44% higher likelihood of failing with a project when selecting a provider with mean client satisfaction in the category "Medium" (8.0-9.77) compared to the top category "Good" (9.77-10.0).
- The providers' failure rate for previous projects (pPF) explained much of the variance in project failure. A difference in 0.1, e.g., selecting a provider with a pPF of 15% rather than one with a pPF of 5%, leads to a 33% higher likelihood of project failure. Thirty-six percent of the clients selected a provider with a failure rate at least 0.1 lower than the best bidder.

The next step in our analysis was to examine the connection between a strong focus on low price and the selection of providers with lower skill. We used the measures Distance to Best mean Client Satisfaction (DBmCS) and Distance to Lowest proportion Project Failure (DLpPF), as defined in Section 4.1. Bidders with very few (fewer than three) previous client evaluations were not included in these measures.

The data in Table 3 suggest that stronger focus on low price (lower FLP-Cat value) was connected with selection of providers with lower skill (lower mCS and higher pPF) and an increased distance between the selected provider and the provider with the highest skill or the lowest failure rate (a higher DBmCS and DLpPF). Especially the increase in the DBmCS and the DLpPF provides strong evidence supporting that selecting a bidder with a relatively low price is connected with a lower focus on selecting the provider with the best skill.

Table 3: Focus on low price vs. provider skill

FLP-Cat	mCS	DBmCS	pPF	DLpPF
1: very low price focus	9.45	0.31	0.11	0.06
2: low price focus	9.38	0.31	0.12	0.06
3: medium price focus	9.36	0.39	0.12	0.07
4: high price focus	9.25	0.53	0.14	0.09

The average effect on the increase in focus on low price on decreased provider skill may not look impressive when examining the differences in Table 3. However, the effect is important. To illustrate this importance, we compared the projects with the following:

- An mCS between 9.4 and 9.5 and a pPF between 0.10 and 0.12 (a project with providers with mCS and pPF values

around the mean mCS and mean pPF values of the FLP category 1 (very low price focus)) and those with

- An mCS between 9.2 and 9.3 and a pPF between 0.13 and 0.15 (a project with providers with mCS and pPF values around the mean mCS and mean pPF values of FLP category 4 (high price focus)).

We found that this seemingly minor increase in the pPF and decrease in the mCS, when going from a very low to a high price focus, was connected to a substantial increase in failure rate. This corresponded to an increase in the failure rate from 22% to 30%, i.e., a 36% increase in the failure rate. Thus, the effect of decreased focus on high skill, i.e., selecting a provider with lower skill than otherwise because of the focus on low price, creates an even larger increase in risk of project failure than the effect of the winner's curse alone.

We proposed in Section 2 three potential explanations for the connection between a focus on low price and increased frequency of project failure: i) adverse selection, ii) the winner's curse, and iii) the Dunning-Kruger effect. We provided evidence supporting the validity of explanations i) and ii) earlier in this section.

To indicate the validity of reason iii), i.e., whether a lower bid for a project, on average, is connected with lower skill, we introduced two additional measures:

- The correlation between the bid amount and the skill of the bidding providers. This measure does *not* analyze the correlation between low bid and low skill per project, but rather assumes that there would be a correlation between bid and price over *all* projects if this correlation is present on most individual projects.
- The correlation between the bid amount and the provider skill for a random set of 10 projects with many (more than 100) bidders. This analysis gives the correlations between low bid and low skill within the projects.

We did not find indications of a systematic connection between low skill and low bids through these measures. The first measure, which includes all projects, provided a correlation between bid amount and mCS of 0.0, and between bid amount and pPF of 0.0. The second measure, i.e., the same correlation within individual projects, gave projects with negative correlation and projects with positive correlations between bid amount and mCS or pPF. The average correlation was, however, 0.0 for the projects—the same result as for the first analysis. This indicates that in the bidding context analyzed in this study the Dunning-Kruger effect is on average not essential, and a low price cannot be used to indicate low or high skill. Of course, this does not mean that the effect is not present in some of, perhaps the most complex, projects or that other measures of skill would have resulted in observations of the Dunning-Kruger effect.

5. DISCUSSION OF THE RESULTS

It is understandable that clients, frequently without much of their own competence in software development, find it hard to evaluate the difference in different providers' skills and the importance of skill differences for the success of the projects. The marketplace, the subject of our analysis, supports the client with previous client evaluations and failure rates. Although this information is clearly useful information in support of the selection, the information is far from perfect about the provider's skills. Clients tended to give the maximum satisfaction score on most projects, several quality aspects of software were not likely to be evaluated properly by the client at the time of the evaluation, the similarity of the previous projects and the current project was hard to compare, and the provider may not have been the one to blame for a project failure. The uncertainty in this information means that, even if the skill measures of this marketplace are better than in several other bidding contexts, there may nevertheless have been a tendency for the client to put too little emphasis on uncertain information about skill and too much on the more accurately known bid price when selecting a provider. Our analysis suggests that many project failures could have been avoided if the clients had put less emphasis on low price and more on high skill, even in a situation with uncertain indicators of skill, such as those provided by the online marketplace.

Our context of relatively small software development projects means that the results cannot necessarily be generalized to contexts with much larger software projects. We found, however, that the identified relationships remain stable when only including the largest projects, e.g., only the 10% largest projects, in our statistical analysis. This suggests that the mechanisms we study are the same for larger projects, but there is nevertheless a need to repeat our study with larger scale projects to be more confident in the relevance for larger projects.

In addition, our analyses are statistical. Although the results are consistent with two of the three proposed causal mechanisms connecting a focus on low price with higher likelihood of project failures, other causes may explain or at least contribute to the observed connections. If, for example, a strong focus on low price is strongly correlated with clients with low competence in software development, the core reason could be the low competence of the client rather than the strong focus on low price.[1] A further limitation of the study is our mechanical use of the skill measures in the analyses. In practice, a client may give different previous client evaluations different weights due to the relevance for the new context. In addition, the clients may use the dialogue with competing providers to guide the selection, e.g., the quality of the question asked, with the providers preceding the selection as essentially important. This means that our mechanical use of skill measures may not fully reflect clients' actual skill evaluation process.

In total, we think that the large population studied and the quality of the data make our results quite robust, but one should be careful when translating the results to other contexts than the one studied, e.g., to contexts with much larger projects and other skill measures.

[1] We analyzed this potentially competing explanation by adding a variable containing the frequency of a client's previous failed projects in the binary logistic regression model. The odds ratio of that variable suggests that a 0.1 increase in the clients' failure rate increased the frequency of project failures by 13%. This increased the ability of the model to explain project failure (Somer's d) from 0.34 to 0.39. Adding this variable, however, did not remove and gave only minor changes in the odds ratio values of the already included variables, which suggests that the characteristics of the client are essential for explaining project failures, but do not create an alternative explanation to the previously suggested causal mechanisms.

6. CONCLUSION

We have provided evidence supporting a connection between a client's focus on low price when selecting a provider and higher likelihood of project failure. We argue that this increase in project failures has two main sources:

- The winner's curse (a low price, based on an over-optimistic estimate, in itself makes the provider perform worse) and
- A decreased focus on high skill, which leads to selecting lower-skilled providers.

The main implication of our findings for clients' selection of software providers is that software clients will substantially decrease the likelihood of project failure when they emphasize the high skill of the provider rather than the low price of the bid.

7. ACKNOWLEDGEMENT

Thanks to the founder and CEO of vWorker, Mr. Ian Ippolito, who generously provided the data and gave supporting input.

8. REFERENCES

[1] Sauer, C., A. Gemino, and B.H. Reich. 2007. The impact of size and volatility on IT project performance. *Communications of the ACM*. 50(11), 79-84.

[2] Jørgensen, M. and K. Moløkken-Østvold. 2004. Reasons for software effort estimation error: impact of respondent role, information collection approach, and data analysis method. *IEEE Transactions on Software Engineering*. 30(12), 993-1007.

[3] Jørgensen, M., 2013. The Influence of Selection Bias on Effort Overruns in Software Development Projects. *Information and Software Technology (to appear)*. simula.no/publications/Simula.simula.1478.

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

[5] Kruger, J. and D. Dunning. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*. 77(6), 1121-1134.

[6] Jørgensen, M. and G. Carelius, 2004. An empirical study of software project bidding. *IEEE Transactions of Software Engineering*. 30(12), 953-969.