Delft University of Technology
Software Engineering Research Group
Technical Report Series

Pricing via Functional Size: A Case Study of 77 Outsourced Projects

Hennie Huijgens, Georgios Gousios, and Arie van Deursen

Report TUD-SERG-2014-012

TUD-SERG-2014-012

Published, produced and distributed by:

Software Engineering Research Group
Department of Software Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Mekelweg 4
2628 CD Delft
The Netherlands

ISSN 1872-5392

Software Engineering Research Group Technical Reports:
http://www.se.ewi.tudelft.nl/techreports/

For more information about the Software Engineering Research Group:
http://www.se.ewi.tudelft.nl/

Pricing via Functional Size: A Case Study of 77 Outsourced Projects

Hennie Huijgens
Delft University of Technology and
Goverdson, Delft, The Netherlands
h.k.m.huijgens@tudelft.nl

Georgios Gousios
Delft University of Technology
Delft, The Netherlands
g.gousios@tudelft.nl

Arie van Deursen
Delft University of Technology
Delft, The Netherlands
arie.vandeursen@tudelft.nl

***Abstract*—A medium-sized west-European telecom company experienced a worsening trend in performance, indicating that the organization did not learn from history, in combination with much time and energy spent on preparation and review of project proposals. In order to create more transparency in the supplier proposal process a pilot was started on Functional Size Measurement pricing (FSM-pricing). In this paper we evaluate the implementation of FSM-pricing in the software engineering domain of the company, as an instrument useful in the context of software management and supplier proposal pricing. We found that a statistical, empirical, evidence-based pricing approach for software engineering, as a single instrument (without a connection with expert judgment), can be used in distributed environments to create cost transparency and performance management of software project portfolios.**

Keywords—Software Economics; Software Pricing; Functional Size Measurement; FSM-pricing; Continuous Improvement.

I. INTRODUCTION

This story is about a company that experiences two problems in its software engineering outsourcing. First, a worsening trend is seen in productivity, indicating that the organization does not learn from historic projects. Second, much time and energy is spent on preparation and review of fixed price project proposals. Our case study explores whether a new project pricing method helps to solve these problems.

To arrive at a price that is acceptable for both parties involved, most companies rely heavily on expert judgment [1]; where the advice of knowledgeable staff is solicited [2]. Usually this is performed as a bottom up approach, where component tasks are identified and sized and then these individual estimates are aggregated to produce an overall estimate [2].

Yet, in practice effort and/or schedule overruns are business-as-usual [3], despite involvement of experts. Software development is characterized by high cost and schedule overruns [4]. Estimation errors are reported to be essential causes of poor management, due to lack of a solid baseline of size [5].

An alternative method for software project estimation is based on algorithmic cost models (COCOMO 2 is a well-known example) which take cost drivers representing certain characteristics of the target system and the implementation environment and use them to predict estimated effort [2]. In many of these statistic approaches size is assumed to be a key factor to estimate project cost [6] [7]. Usually size of software engineering projects is measured with a formal Functional Size Measurement (FSM) standard [8]. FSM is a method to measure size of software engineering activities by means of the functionality delivered to users [7], which lays the foundation for a statistical method of project pricing based on functional size.

Advantages of such a statistical method are that this will help to improve transparency of estimations and that it can be a good instrument to create continuous improvement of project performance.

However, our observation is that a purely statistic method is almost never used. If statistical analysis is used, this is usually supplementary to an expert judgment-based approach [1]. And practice shows that in most cases the expert opinion – in many cases supported by reasoning by analogy – is leading when it comes to decision making [9].

The goal for this paper is to answer the question whether a purely statistical approach to pricing is effective in an outsourcing context. We define an approach to be effective when a so-called win-win situation is achieved: meaning that both involved parties are satisfied. The supplier delivers a service for a price that is higher than the cost, and the customer gets higher value than the paid price. In addition to that the outsourcing context asks for a long-term (5 year) relation.

Based on this we focus on transparency as the factor we need to measure to determine success: transparency when setting the price and transparency when finalized. Transparency is important for 'next' projects; when pricing of actual projects is transparent, this can be re-used in future projects.

A long-term relation asks, in an outsourcing context, for intent of continuous improvement. When a supplier becomes more efficient and effective, the price can go down without a negative effect on the supplier's margin of profit. More value for the same amount of money represents a win-win situation for both the customer and the supplier.

For this purpose we define three research questions:

RQ1: To what extent are both parties involved in an outsourcing contract satisfied with FSM-pricing?

RQ2: To what extent does FSM-pricing help to improve transparency of project proposals?

RQ3: To what extent does FSM-pricing help to create continuous improvement?

In order answer these research questions, we describe the implementation and evaluation of FSM-pricing as a single instrument for software management, in a telecom company in a west-European country (in this paper indicated as COMPANY C), and the pricing approach agreed with its main Indian IT-supplier (in this paper indicated as SUPPLIER S). We study data collected from 77 projects conducted (completed as well ongoing) since 2012. Moreover, we conducted 25 interviews including structured as well as open-ended questions.

In Section II, we survey research on FSM-pricing and discuss the few empirical studies that do exist. In Section III, we chalk out the backgrounds of FSM-pricing. In Section IV, we describe our research method. We present results in Section V. In Section VI we put things together and discuss limitations. In Section VII the implications are described, and finally Section VIII includes a conclusion and a summary of the main recommendations for future research on FSM-pricing.

II. RELATED WORK

When it comes to software pricing, Sommerville [10] mentions that estimates are made to discover the cost of producing a software system. There is not a simple relationship between the price charged to the customer and the development cost, and the price charged is influenced by a number of broader organizational, economic, political, and business considerations. Two types of estimation techniques are distinguished; experience-based techniques such as expert judgment and algorithmic cost modeling where cost is estimated as a mathematical function of product, project and process attributes [10].

A well-known example of the latter is Boehm's COCOMO 2 [11] [12] [13]; more methods based on algorithmic software cost models with specific regression formula are widely used in industry, such as the Putnam Model [14], and SEER-SEM [15].

Studies covered in a review by Moløkken and Jørgensen on Surveys on Software Effort Estimation [3] mention a variety of estimation aids; such as system development method (SDM) [16], work breakdown structure [17], Functional Size Measurement (FSM) such as Function Point Analysis (FPA) [18] [9], parametric tools [19] [20], and qualitative methods [21].

Functional Size Measurement (FSM) is a method to measure size of software engineering activities. Five FSM methods are certified by ISO as an international standard; in our study IFPUG FPA (ISO 2003c) is used. FSM origins from Function Point Analysis, designed by Albrecht in 1979 [22] to estimate size by means of user functionality. An overview of FSM can be found in [7].

For a long time researchers and practitioners have been investigating the use of statistics in software estimation. A study by Fairley back in 1992 mentions as future trend on software estimation "*an increased reliance on statistical, rather than single point estimates of size, effort, and schedule*" [23]. Since the 90's a limited number of studies has been published on the subject of pricing of projects based on statistics [24] [25].

Despite all models and practices something seems wrong with actual software estimation and software pricing. Moløkken and Jørgensen [3] reveal that 60-80% of the projects encounter effort and/or schedule overruns. Estimation methods in most frequent use are expert based: expert consultation, intuition and experience, and analogy. The frequent use of expert judgment is grounded by a lack of evidence that formal estimation models lead to more accurate estimates [3].

In the research literature, it is hard to find case studies of organizations approaching IT-investments in a purely quantitative way, i.e., based on calculations derived from mathematical models built upon historic cross-industrial data sets [26]. We did not find studies that describe dedicated use of algorithmic cost models in practice, without interference of expert-judgment based methods. Besides that, very limited research is performed specifically on the topic of pricing software projects. We have not found any studies that emphasize the use of FSM as a single instrument for a company's pricing method. This is remarkable; several studies on FSM stress that software size is a primary predictor of project effort and thus project cost [7] [6]. We build on Abran et al. [27] arguing that "*in the software engineering literature, even though there is a large number of 'metrics' proposed, there is still very little discussion on the topic of measuring instruments (...)*".

The innovation of our study is that we – for the best of our knowledge for the first time in scientific research – raise the question to what extent a single, statistical, empirical approach to project estimation can reach the goal of transparent project proposals and continuous improvement. To do so, we provide an in depth case study of actual use in 77 completed and ongoing projects. This study is primarily descriptive, and not comparative: we do not have the data to see how other pricing approaches might have worked. Yet, we provide a rigorous analysis of what worked well, and what did not work well using Functional Size Measurement as an instrument for pricing.

III. FSM-PRICING

FSM-pricing, as described in this paper, is implemented in the software project department of COMPANY C, as part of a transformation program that includes a change from one large European IT-supplier to a large Indian IT-company (SUPPLIER S) for the majority of its software engineering activities for the Customer Relationship Management (CRM), Billing, and Data Warehouse (DWH) applications. FSM-pricing aims to implement Functional Size Measurement (FSM) based on Function Point Analysis [8] as an approach to improve the capability of the company to challenge SUPPLIER S's proposals for to-be-started software engineering activities.

A few months before FSM-pricing became operational within COMPANY C, through analysis of finalized software engineering projects we discovered two major disadvantages in the current expert-judgment-based estimation approach. First, COMPANY C showed a worsening trend on productivity, indicating that the organization did not learn from historic project data. Second, much time and energy was spent on preparation and review of fixed price project proposals. This led to ongoing discussions about proposed project costs, driven by a mismatch between expectations from COMPANY C's customers, and highly detailed effort estimations by SUPPLIER S's developers. To turn the tide on the worsening productivity, and to smoothen the proposal process, a decision was made to change towards an empirical, evidence-based, and analytical way of preparing fixed price project proposals. FSM-pricing was

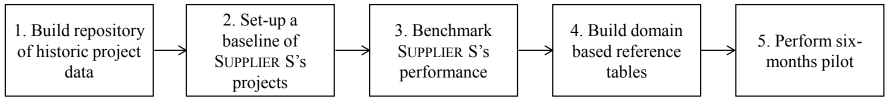

Fig. 1 The 5 steps in FSM-pricing

born, having two goals: 1) improve transparency of proposals, and 2) create continuous improvement of software delivery.

A. *Implementation of FSM-pricing in practice*

FSM-pricing consists of 5 steps (see Fig. 1):

1) *Build repository of historic project data*

A measurement team of COMPANY C collected data on historic software projects that finalized in 2012 and 2013. Both quantitative data (e.g. core metrics such as size, effort, cost, duration) and qualitative data (e.g. project backgrounds, factors that influenced a project) were collected in a measurement repository. Projects cover a mix of business domains (e.g. CRM, Billing, DWH), project types (e.g. newly built systems, enhancements, off-the-shelf packages), and sizes (e.g. small enhancements, large once-only projects). In most projects the design, build, and testing activities were performed by one or more external suppliers. Most software projects were combined in releases and delivered at one moment to the business organization; each year eight releases are rolled out under guidance of a portfolio management team.

2) *Set up a baseline of SUPPLIER S projects*

Once the transformation program was finalized in which all contractual agreements were made with the new main SUPPLIER S, analysis was performed, by a measurement team of COMPANY C, on the performance of projects that were performed during the transformation period by SUPPLIER S. All data used in the analysis were shared and thoroughly reviewed by measurement experts of SUPPLIER S.

3) *Benchmark SUPPLIER S's performance*

Once a representative set of finalized software projects performed by SUPPLIER S was collected, the results were internally and externally benchmarked; the performance was compared with 22 finalized projects within COMPANY C, and with a repository of 331 comparable projects from other companies that were in earlier research (see also [28] [29]). All compared projects conduct software engineering in business environments. Peer group projects were measured, collected, and recorded in the same way as conducted in this case study. The projects are compared according to an existing set of key performance indicators (KPIs) that are used within COMPANY C's performance dashboard: the realized productivity (in cost per FP and effort per FP) and time-to-market (in days per FP).

4) *Build domain specific reference tables*

Based on analysis and benchmarking of projects performed by SUPPLIER S, two domain-specific baselines on cost per FP were calculated. To create the baseline, we obtained the best fit after conducting a log-log transform. After performing a power regression, the resulting price calculation formula is:

$$Price = \alpha \times (FP)^{\beta} \quad (1)$$

The coefficients α and β may differ per application domain. In the portfolio under study, we typically have $\beta \approx 0.75$. Note that this formula is in line with COCOMO 2's effort estimation formula (which uses KLOC instead of function points) [13].

We use simple regression on size and cost with power fit. Our foundation of this argument is that such a model facilitates greater analyzability and thus helps improving transparency. For a statistics-based explanation we create a cross correlation table to determine, and filter the strongly dependent variables in our sample out from the regression model. We found that size and duration are all pair-wise highly correlated; we rejected duration and only used size as a predictor for cost. See the technical report for more details on statistics [30].

We prepared two baselines: 1) CRM/Billing and 2) DWH. CRM/Billing domain projects are combined in one baseline because the analysis shows no large differences between projects from both domains, many projects overlap domain borders, and because not enough data were available for proper individual trend lines for both domains. A separate DWH baseline was setup because these show a different pattern.

5) *Perform a six-months pilot*

Based on both baselines a tool was set up for cost calculation in project proposals by SUPPLIER S. For all to be started software projects the fixed price is calculated with this tool. Once the size of a project is counted and reviewed, the tool calculates the price for a project to be performed by SUPPLIER S based on the applicable domain baseline.

Stakeholders from COMPANY C opted strongly for a single pricing approach (only based on statistics), because ongoing discussions on project estimates were expected due to a variety of expert opinions if two approaches were to be used simultaneously. To reassure stakeholders of SUPPLIER S with doubts on this single method for supplier proposal pricing, a six month's FSM-pricing pilot was started. This pilot is the subject of the case study that is discussed in this paper.

IV. RESEARCH METHOD

We use a mixed methods methodology, as we are examining a phenomenon with multiple (qualitative and quantitative) tools. We perform a single-case, holistic case study that involves two instruments; a survey consisting of open and closed questions, and a quantitative analysis of actual project data. As reflected in Fig. 2 the research method is based on the three research questions with regard to satisfaction with FSM-pricing (RQ1), improvement of transparency of project proposals (RQ2) and the creation of continuous improvement (RQ3).

We created a combined 10-minute questionnaire survey. The survey topics and the survey approach were determined in a number of preparation sessions between management representatives and the measurement experts of both COMPANY C and SUPPLIER S. Our aim is to come up with a manageable set of topics that would represent the pilot effectively. The survey

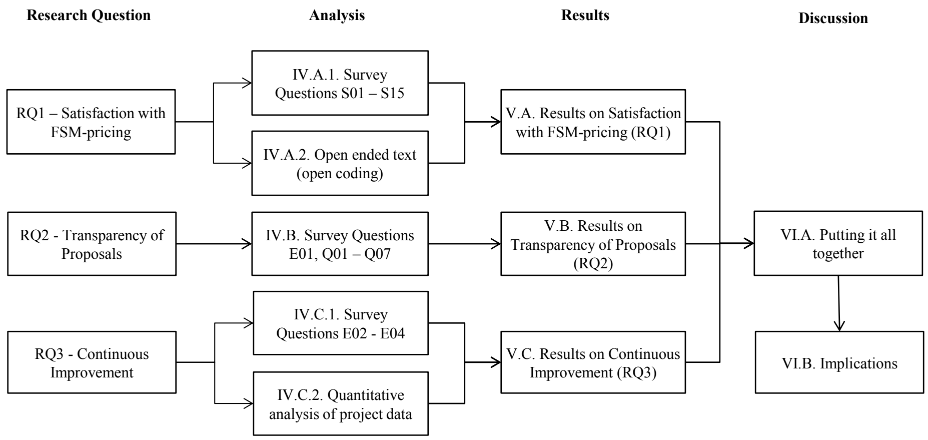

Fig. 2 Overview of the research method, incl. references to Sections in this paper.

consists of a number of closed questions; respondents are asked to rate these survey topics on a 5-point Likert scale. Next to the 5-point scale for each of the survey topics a choice of "Don't Know" as an answer is an option. Besides that the survey contains three open questions.

The survey starts with the collection of demographic information, and the answering of two partially closed questions: "What company are you working for?" and "What is your connection with FSM-pricing?" Both questions are intended to find out any differences in satisfaction with FSM-pricing within both the involved parties COMPANY C and SUPPLIER S, and between respondents with different roles.

A comprehensive overview of setup and respondent statements in the survey can be found in the technical report [30].

A. RQ1 - Satisfaction with FSM-pricing

For RQ1 (To what extent are both parties involved in an outsourcing contract satisfied with FSM-pricing?) we assess the satisfaction with FSM-pricing. As a response to the question "How satisfied are you with the following?" respondents are asked to rate 15 survey topics.

To find out whether respondents feel that FSM-pricing needs to be continued a question is asked to be answered with yes or no: "Should FSM-pricing be continued as an operational practice once the pilot is finalized?"

To understand possible reasons behind the closed questions we ask the stakeholders to answer three open questions (max 3 answers are allowed for each question):

1. What is going well during the FSM-pricing pilot that we want to continue?
2. What is not going well during the FSM-pricing pilot that we want to fix?
3. What can we do to improve FSM-pricing?

As is common in case studies, answers contain a substantial element of narrative [31]. As these are representatives of the complexities and contradictions of real life, we include a selection of statements made by the survey respondents in the section on open ended text analysis in our paper. We try to include examples of respondent statements that apply to differences as well as similarities. TABLE III gives an overview of all survey topics related to RQ1 - Satisfaction of FSM-pricing.

B. RQ2 – Transparency of Project Proposals

As an answer to RQ2 (To what extent does FSM-pricing help to improve transparency of project proposals?) we perform a survey with eight closed questions. The first seven (Q01 to Q07) are intended to find out how respondents experience the quality of artifacts and processes with regard to FSM-pricing. As a response to the question "How would you rate the quality of the following?" respondents are asked to rate these seven survey topics. Next to these questions one additional question (E01) is asked: "To what extent did you experience a change on the transparency of proposals during the FSM-pricing pilot?" TABLE IV gives an overview of all survey topics related to RQ2 - Transparency of Project proposals.

C. RQ3 – Continuous Improvement

RQ3 (To what extent does FSM-pricing help to create continuous improvement?) is answered by performing quantitative analysis of project data. We analyze the performance of 77 finalized or ongoing software engineering projects. For our study we use data of four categories of software engineering projects, all performed within COMPANY C:

1. *Repository*: project data of historic projects in the period preceding FSM-pricing, not performed by SUPPLIER S (n = 22);
2. *Baseline*: project data of finalized projects performed by SUPPLIER S that were used to prepare the FSM-pricing baseline (n = 16);
3. *Pilot*: project data of projects finalized during the pilot that are in scope of FSM-pricing (n = 10);
4. *Forecast*: project data of ongoing projects that are in scope of FSM-pricing (n = 29).

We collect data on finalized and still ongoing (forecasted) software engineering projects of three business domains; CRM, Billing, and DWH. During a one-year period, from Q4-

2012 to Q4-2013 the majority of finalized software engineering projects that are performed within these business domains of COMPANY C are in scope for measuring and analysis of the overall project performance. We exclude projects that are only about infrastructure, or that include only non-functional requirements (e.g. performance, security), from the analysis. No failed projects are included in the repository.

For all to-be-analyzed software engineering projects, size is measured in Function Points (FPs), according to FSM ISO/IEC 20926 guidelines [8]. We perform a so-called estimated or approximated function point analysis: a variant of function point analysis in which the number of functions is determined for each type of user function (user transactions and logical files), and which uses standard values for complexity: 'Average' for the user transactions and 'Low' for the logical files [32]. As add-on we used additional guidelines for Data Warehouse projects [33].

Function Point Analysis is performed by specialists either from a COMPANY C measurement team (in the period that SUPPLIER S is not in scope as main supplier yet), or by a SUPPLIER S measurement team (once SUPPLIER S is in scope as main supplier they perform all FPA). Every FPA is thoroughly reviewed on correct utilization of counting practices by an experienced IT-metrics expert who is also one of the authors of this paper, and on correct interpretation of requirements by an applicable subject matter expert of COMPANY C.

Besides project size, we collect a set of core metrics for each project; e.g. project cost, supplier cost, effort (per supplier), effort and cost for user acceptance testing, dates of project milestones. As a source for the project data we use the formal project administration. All project data is reviewed by the applicable project manager and the financial controller of COMPANY C, and adjusted where needed.

For each project we calculate and analyze the following performance indicators (based on the standard set of KPIs within COMPANY C):

1. *Project Productivity*; total project cost divided by the project size, expressed in Euros/FP;
2. *Build & Test Productivity*; cost of the Build & Test phase divided by the project size, in Euros/FP;
3. *Project Time-to-Market*; duration of the project from start of the Initiation phase to technical go live divided by the project size, in Days/FP.
4. *Build & Test Time-to-Market*; duration of the Build & Test phase divided by the project size, in Days/FP.

When in this study Productivity or Time-to-Market is mentioned without any prefix, the project version of each indicator is meant. For analysis purposes results of individual projects are aggregated to company level.

V. RESULTS

In this section we report results based on the three research questions of our study. We sent 41 survey requests by email to 17 employees of COMPANY C and 24 employees of SUPPLIER S. We selected these stakeholders because they are all involved in the FSM-pricing pilot. 27 surveys are returned, of which 2 are assessed to be incomplete (respondents only noted that they knew too less of the subject). 25 surveys are completed (completion rate 61%); the analysis in this study is based on these completed surveys only. TABLE I summarizes the backgrounds of the respondents that completed the survey.

TABLE I BACKGROUNDS FROM SURVEY RESPONDENTS

Respondent background	COMPANY C n=11 (44%)	SUPPLIER S n=14 (56%)
Overall IT-management	28%	29%
FPA Measurement Team	18%	14%
Portfolio Management	27%	0%
Data Warehouse Team	9%	14%
CRM/Billing Team	9%	36%
Other	9%	7%

Besides the results of the survey ratings we collected a large amount of open ended text from our survey. The first open question "What is going well during the FSM-pricing pilot that we want to continue?" resulted in 46 answers. The second open question "What is not going well during the FSM-pricing pilot that we want to fix?" resulted in 47 answers and 44 answers were given to the question "What can we do to improve FSM-pricing?" In total 2,007 words were produced. In this section we label respondents as P1 through P25 and we include results from the open text analysis where applicable.

To analyze the free text answers, we adopt the coding technique described by Runeson et al [34]. We applied high level codes and medium level codes and counted the frequency of each code. A summary of the results of this analysis is shown in TABLE II.

TABLE II SUMMARY OF THE OPEN ENDED TEXT ANALYSIS

Category Name / Medium Level Code
Interactions, communications, people
Improved proposal transparency
Improve knowledge of Function Point Analysis and FSM-pricing
Discussion on size when lower price is expected or on waivers
Organization, processes
Uniform, standard and simplified process
Too small projects; no focus on release-based working
Delay due to search for clarity and review
Improve pricing tables (e.g. benchmarking, more realistic figs.)
Promote release-based working based on size
Promote pricing tables based on applications (technology)
Measurements
Perform gap-analysis on FSM-price versus actual effort spent
Requirements
FSM-pricing does not cover non-functional requirements
Low reliability of FSM-pricing when compared to actual effort
Improved Requirement Management
Artifacts
Good quality of Function Point Analysis process and products

TABLE III SURVEY RESULTS FOR RQ1- SATISFACTION WITH FSM-PRICING

Survey Topic (How satisfied are you with the following?)	Nr	Mean Overall	Standard Deviation	Mean Company	Mean Supplier	Effect Size Company/ Supplier	Effect Size Management / Development
Function Point Analysis method (IFPUG, estimated count)	S09	3.96	0.81	4.00	3.92	0.08	0.11
FSM-pricing pilot period itself	S02	3.87	0.55	3.91	3.83	0.08	-0.20
Preparation of the FSM-pricing pilot	S01	3.75	0.90	3.82	3.69	0.13	0.00
Overall FSM-pricing	S15	3.72	0.74	3.64	3.64	0.00	0.08
Advantages of FSM-pricing for COMPANY C	S13	3.68	0.65	3.80	3.58	0.22	-0.30
Pricing table for DWH	S07	3.50	0.73	3.86	3.22	0.63	0.15
Proposal Process (with regard to FSM-pricing)	S12	3.42	0.88	3.70	3.21	0.49	0.06
Management Commitment on FSM-pricing	S04	3.42	0.83	3.64	3.23	0.41	0.25
Advantages of FSM-pricing for SUPPLIER S	S14	3.40	0.68	3.29	3.46	-0.18	0.18
Communication with regard to FSM-pricing	S03	3.39	0.66	3.36	3.42	-0.05	0.22
Setup of the SUPPLIER S Baseline	S06	3.30	0.93	3.55	3.08	0.46	0.13
Pricing table for CRM / Billing	S08	3.28	0.83	3.57	3.09	0.48	0.22
Reliability of the FSM-pricing	S05	3.28	0.94	3.55	3.07	0.47	0.09
Coverage of FSM-pricing	S11	3.26	0.92	2.70	3.69	-0.99	-0.45
Waiver procedure for Function Point Analysis (exclusions)	S10	3.25	1.03	3.00	3.46	-0.46	0.38

Sorted by Mean Overall; higher is better.

A. *Results RQ1 – Satisfaction with FSM-pricing*

TABLE III summarizes the survey results with regard to RQ1 (To what extent are both parties involved in an outsourcing contract satisfied with FSM-pricing?). The two last columns show Effect Size calculated as two measures; 1) for each survey topic the difference between the mean COMPANY C score and the mean SUPPLIER S score, and 2) for each survey topic the difference between the mean Management score (all scores of respondents with the profile Overall IT-management, FPA Measurement Team, Portfolio Management, and Other) and Development (all scores of respondents with the profile Data Warehouse Team, and CRM/Billing Team). A negative Effect Size indicates COMPANY C / Management respondents are less satisfied with a survey topic than SUPPLIER S / Development respondents. A positive Effect Size indicates COMPANY C / Management respondents are more satisfied with a survey topic than SUPPLIER S / Development respondents.

We calculated the p-value for each survey topic using a Wilcoxon rank-sum test. However, analysis shows that the p-values are not statistically relevant. Therefore we do not include these in our paper; all p-values can be found in the technical report [30]. The lack of statistical significance for the comparisons is due to the sample set of 25. We report our most striking findings; future work on a larger sample is needed to obtain statistical significance.

We found the following with regard to satisfaction with FSM-pricing based on analysis of the survey results:

1) *88% want FSM-pricing as operational practice*

On the question “Should FSM-pricing be continued as an operational practice once the pilot is finalized?” 80% answered “Yes”; 8% answered “Ok, but with improvement points (e.g. include effort of non-functional requirements”).

2) *Function Point Analysis is appreciated by both parties*

Both COMPANY C and SUPPLIER S respondents appreciate the applied Function Point Analysis method (IFPUG, estimated counts); based upon the highest overall mean score of the survey (3.96). Besides that both parties appreciate the quality of the function point analyses that are performed by SUPPLIER S (3.78), and the reviews done by COMPANY C (3.80).

Qualitative analysis confirmed this finding. Many respondents considered the quality of the function point analysis high:

> *Good Function Point review by COMPANY C and SUPPLIER S Function Point Analysis teams before proposal submission. (P10)*

> *Appreciate the way Function Point counting is done by SUPPLIER S. (P23)*

> *No big difference between COMPANY C and SUPPLIER S countings occur. (P14)*

Apparently a good Function Point Analysis, including proper review, is a prerequisite for efficient FSM-pricing.

Many remarks made by respondents were related to requirements; which makes sense since requirements usually are the basis for project proposals. A noteworthy side-effect of FSM-pricing is that respondents experienced an improvement of the requirement management process during the pilot.

> *Most of the details are sorted out at the time of proposals. Earlier these details were discussed in design phase. (P17)*

> *The solution is looked into more detail in order to get the right Function Points at the proposal stage itself. This helps in early detection of issues and resolution. (P2)*

This positive effect on requirements management might even be one of the main reasons for FSM-pricing success.

3) *COMPANY C management: coverage needs improvement*

Coverage is about the number of projects in COMPANY C‘s IT-portfolio that is subject of FSM-pricing. Based on a relatively low mean value for COMPANY C (2.70), combined with an

TABLE IV SURVEY RESULTS FOR RQ2 - TRANSPARENCY OF PROJECT PROPOSALS

Survey Topic (To what extent did you experience change on…?)	**Nr**	**Mean Overall**	**Standard Deviation**	**Mean Company**	**Mean Supplier**	**Effect Size Company/ Supplier**	**Effect Size Management / Development**
Transparency of Proposals	E01	3.88	0.65	3.82	3.93	-0.11	0.36
Survey Topic (How would you rate the quality of the following?)							
Function Point Analysis performed by SUPPLIER S	Q02	3.83	0.70	3.70	3.93	-0.23	-0.06
Function Point Analysis Review by COMPANY C	Q03	3.78	0.60	3.73	3.83	-0.11	-0.11
The Overall FSM-pricing method	Q07	3.64	0.57	3.55	3.71	-0.17	-0.22
The SUPPLIER S Proposals based on FSM-pricing	Q06	3.52	0.65	3.55	3.50	0.05	0.12
The CRM / Billing Baseline used for FSM-pricing	Q05	3.47	0.80	3.57	3.40	0.17	-0.05
Requirements delivered by COMPANY C	Q01	3.44	0.65	3.45	3.43	0.03	-0.01
The DWH Baseline used for FSM-pricing	Q04	3.43	0.76	3.71	3.14	0.57	0.55

Sorted by Mean Overall; higher is better.

Effect Size of -0.99 between COMPANY C and SUPPLIER S, we conclude that respondents from COMPANY C are more than average dissatisfied about the coverage of FSM-pricing. An Effect Size of -0.45 between Management and Development indicates that coverage is a management rather than a developer concern.

We conjecture a connection with low rating of the waiver procedure by COMPANY C respondents; this procedure allows SUPPLIER S to exclude a project from FSM-pricing. A standard waiver is applied for infrastructure projects, configuration projects, and projects executed by other external suppliers. Also qualitative analysis revealed indications that ongoing discussions tend to be related with waiver requests:

> *Many ongoing discussions on waiver requests occur. (P20)*

4) SUPPLIER S development: reliabiliy needs improvement

In the context of FSM-pricing by reliability we mean whether respondents experience the outcome of FSM-pricing to be in line with their own judgment. SUPPLIER S developers seem dissatisfied with FSM-pricing where it comes to reliability. Proposal process (Effect Size 0.49), both pricing tables (0.48 and 0.63), reliability of FSM-pricing (0.47), and setup of baselines (0.46) are all rated low. We believe these are connected, but we did not find evidence for this in our data.

Looking at this aspect further in the qualitative analysis shows a feeling of disagreement between the outcome of FSM-pricing and effort-based estimates. Many respondents, especially from SUPPLIER S, mention that FSM-pricing does not cover Non-Functional Requirements and complexity (technology).

> *Function point analysis is not applicable to projects where more testing efforts are required for less development changes. (P5)*

> *All the projects do have different non-functional requirements or technology; due to this the efforts differs. (P2)*

> *The complexity of the changed code does not match with the amount of functionality to be changed, causing a disparity. (P16)*

> *Refinement of the Function Point trend lines based on technology, in order to make them more realistic. (P7)*

We identified one specific measurement-related issue: the wish to perform a gap-analysis to find any differences between FSM-pricing proposals and actual effort spent in a project:

> *To keep the counting simple we are considering all the requirements are at average level; we may need to perform gap analysis if the requirements mix is really averaging out on efforts. (P17)*

> *Cross verification with actuals towards the end of project to revalidate the estimates would be an improvement. (P7)*

We identified a need for gap-analysis in order to identify differences between (estimated) project cost and actual effort. We consider conducting this gap-analysis as future research.

B. Results RQ2 – Transparency of Project Proposals

TABLE IV summarizes the survey results with regard to RQ2 "To what extent does FSM-pricing help to improve transparency of project proposals?" We observed one major finding here:

1) 84% experienced improved proposal transparency

Many respondents experienced an improvement of the transparency of project proposals during the FSM-pricing pilot (72% said transparency improved; 12% said greatly improved). Qualitative analysis confirmed this finding. Respondents mention improved transparency as a positive outcome of the FSM-pricing pilot:

> *A good point is that there is less discussion. (P8)*

Some respondents see improved transparency as a driver for better requirements or to solve disagreements between customer and supplier:

> *Instead of plain list of entities that we were maintaining in work-breakdown-structure entities, we now have clarity on what kind of functionality is getting delivered. (P17)*

> *Function points analysis sometimes is a constructive argument in case of disagreement. (P20)*

We observed the fact that FSM-pricing is experienced as a uniform, simplified process is on top of respondents' list:

> *FSM-pricing is a single point for the final estimation, answerable to all stakeholders. The estimation review process becomes very simple. A standardized process, which can be trusted from both vendor and client stakeholders. (P24)*

> *Uniformity in pricing approach as it does not depend on individual components to derive their efforts. (P2)*

> *Avoid delays and budget overruns as estimation can be done at an initial stage against task-based. (P13)*

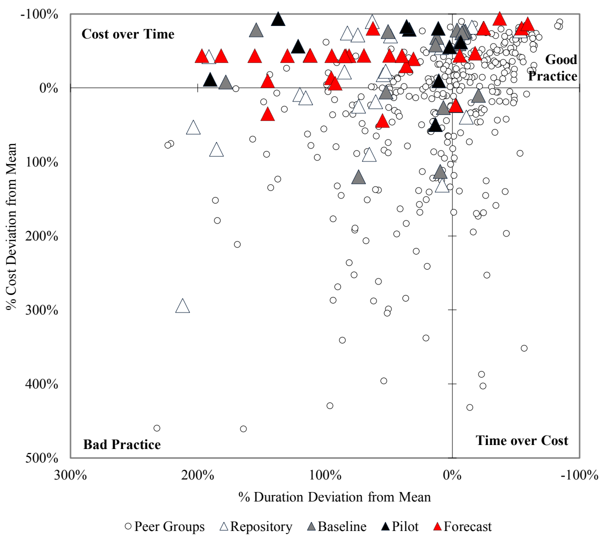

FIGURE 3 COST / DURATION MATRIX

C. Results RQ3 – Continuous Improvement

As described in Section IV.C quantitative data from four categories of 77 software engineering projects are used for quantitative analysis of project data (resp. Repository, Baseline, Pilot, and Forecast). In TABLE V we summarize the performance indicators for these four project categories. The analysis resulted in the following findings:

1) Time-to-Market not in sync with peer groups

Analysis of the performance of the software engineering projects of COMPANY C shows that, although the project cost are in line with the prevailing market, the organization suffers from project durations that are substantially longer than those of peer groups in industry. An external benchmark against historic data of 331 finalized software engineering projects [28] from different companies shows that a majority of the finalized projects of COMPANY C are cost effective (average Project Productivity is 46% better than the peer groups, see TABLE VI), yet project durations are longer than the average of the total research group (average Project Time-to-Market is more than twice that of the peer groups, see TABLE VI). This finding is applicable to all four categories of software projects performed within COMPANY C in our research repository, yet Time-to-Market is worsening during the pilot.

We plot both all COMPANY C and peer group projects in a Cost / Duration Matrix (see FIGURE 3) [28] [29]. This matrix shows for each project the measure of deviation from the average trend line (average of peer group projects plus COMPANY C projects) expressed in a percentage; negative when below the average trend line, positive when above the trend line. The matrix is divided in four quadrants. Each quadrant is characterized by the measure of negative or positive deviation from the average trend. When analyzed it shows that 80% of the projects is assessed to have a longer than average duration. 25% of the projects are in the Bad Practice quadrant; these projects perform in both cost and duration worse than average. 55% ends up in the quadrant Cost over Time; costs are less than average, yet project duration takes longer than average. Due to these deviating percentages we argue that Company A’s Time-to-Market, measured in days per FP, is not in sync with its peer groups; COMPANY C should improve its Time-to-Market in order to stay competitive in the market.

TABLE V PERFORMANCE OVER FOUR PROJECT CATEGORIES

Performance Indicator	Rp	Bl	Pi	Fc
Number of projects (n)	22	16	10	29
Average project Size (FP)	157	183	25	55
Project Productivity (EUR/FP)	2,651	1,485	2,560	1,539
B&T Productivity (EUR/FP)	1,338	813	1,537	1,123
Average B&T cost (K Euro)	209	148	39	61
Project Time-to-Market (Days/FP)	2.35	1.58	7.17	*2.95*
B&T Time-to-Market (Days/FP)	1.40	0.89	4.58	*2.87*
Average B&T Duration (Months)	7.21	5.35	3.81	*5.15*

Rp = Repository, Bl = Baseline, Pi = Pilot, Fc = Forecast, B&T = Build & Test.
Numbers printed in *italic* are forecasted and therefore subject to change once finalized.

TABLE VI PERFORMANCE COMPARED TO PEER GROUPS

Performance Indicator	COMPANY C	Peer Gr.	Delta
Number of Projects (n)	26	331	n.a.
Average Project Size (FP)	126	261	-52%
Project Productivity (EUR/FP)	1,604	2,983	-46%
Average Project Cost (K Euro)	203K	780K	-74%
Project Time-to-Market (Days/FP)	2.20	1.04	112%
Average Project Duration (Months)	9,14	8.92	2%

Performance of Company in comparison with peer group projects from our research repository. Only finalized projects that were performed by Supplier are incorporated (Baseline + Pilot)

TABLE VIII PERFORMANCE OVER TIME

Performance Indicator	2012-2013	2014-FC	Delta
Number of projects (n)	38	39	n.a.
Average project Size (FP)	168	68	-59%
Throughput (FP)	6,366	2,660	-29%[1]
Project Productivity (EUR/FP)	2,116	*1,679*	*-21%*
B&T Productivity (EUR/FP)	1,097	1,180	8%
Average B&T cost (K Euro)	184K	56K	-70%
Project Time-to-Market (Days/FP)	2.00	*3.52*	*76%*
B&T Time-to-Market (Days/FP)	1.17	*3.10*	*166%*
Average B&T Duration (Months)	6.43	*4.81*	*-25%*

Numbers printed in *italic* are forecasted and therefore subject to change once finalized.
[1]Throughput percentage is calculated based on extrapolation per year; B&T = Build & Test.

Our analysis is that the low Time-to-Market is caused by two problems. First; the combined release approach of COMPANY C causes waiting time (waste) and unnecessary dependencies between projects. Second; long average project duration, combined with mall project size cause a bad Time-to-Market as illustrated in the following.

2) Small projects block improvement

A finding with regard to project size is that from 2013-Q3 onwards substantially more very small projects (e.g. projects smaller than 30 FPs) are performed. We did not find any reason that could explain this reduction of project size. Although smaller projects are from a cost point of view advantageous

TABLE VII SURVEY RESULTS FOR RQ3 – CONTINUOUS IMPROVEMENT

Survey Topic (To what extent did you experience change on…?)	Nr	Mean Overall	Standard Deviation	Mean Company	Mean Supplier	Effect Size Company/ Supplier	Effect Size Management / Development
Productivity (Cost per FP)	E02	3.33	0.70	3.40	3.29	0.11	0.17
Process Quality (Defects per FP)	E04	3.22	0.65	3.25	3.20	0.05	0.67
Time-to-Market (Days per FP)	E03	3.00	0.76	2.78	3.15	-0.37	0.42

Sorted by Mean Overall; higher is better.

for SUPPLIER S, portfolio managers of COMPANY C are responsible for the construction of a specific release portfolio (a number of projects combined in one release; to be delivered at one specific moment). An analysis is performed on possibilities to join small projects within a release, in order to gain economy-of-scale advantages. This analysis shows that, measured over 2014 as a whole, a cost saving of 4% is achievable by carefully combining projects.

The idea that small projects from an economy-of-scale perspective should be combined is mentioned by some respondents in the open ended text as well:

SUPPLIER S divides the offer in small pieces; we must have release based funding to make use of economy-of-scale. (P8)

Too many small projects are negative for COMPANY C due to economy-of-scale effects. (P3)

We observed that in 2014 the throughput (total delivered number FPs) is approximately 29% lower than in the preceding years (see TABLE VII). One can argue that the maybe rather rigid approach of FSM-pricing is not sufficiently encouraging for SUPPLIER S due to a somewhat single-sided focus on cost reduction. However, COMPANY C promotes the idea that delivery of more throughput where applicable is desired. Looked upon from this side FSM-pricing underlines the delivery of more value for less money; and at the same time it rewards throughput enlarging by creating more turnover for the supplier.

3) Performance is not improving over time

Looking at Productivity and Time-to-Market over time (see TABLE V) we find that these performance indicators do not show an improving trend. Although the Project Productivity (the Productivity measured over the whole project lifecycle from initiation to technical Go Live) improves by 21% in 2014 onwards compared to the years before, Build & Test Productivity (the Productivity measured over the Build and Test period that was performed by SUPPLIER S) decreases with 8%.

Next to our finding that Time-to-Market is substantially higher than that of the peer groups, no sustained improvements with regard to project durations are seen when assessed over time. Both Project Time-to-Market and Build &Test Time-to-Market show a worsening trend. Especially worrying is the Time-to-Market of the Build and Test period; this is even 166% longer than in the preceding years; as discussed before the small size of many projects and the amount of waste in projects plays an important role here.

A noteworthy observation with regard to the measured deterioration of Time-to-Market over time (see TABLE V) is that 13 respondents experienced Time-to-Market as neutral; while 5 respondents experienced improvement. Only 3 respondents experienced Time-to-Market as being deteriorated. The background of these respondents is diverse; a slight majority of SUPPLIER S development rated Time-to-Market as improved and COMPANY C management rated this as deteriorated. A possible explanation for this inconsistency is that COMPANY C management receives quarterly performance updates based on the measurement approach that we used, while SUPPLIER S staff does not receive these reports, and therefore is not familiar with the applied key performance indicators.

VI. DISCUSSION

A. Putting it all together

Analysis with regard to RQ1 (To what extent are both parties involved in an outsourcing contract satisfied with FSM-pricing?) resulted in four findings:

First, 88% of the respondents of our survey want FSM-pricing as an operational practice once the FSM-pilot is finalized.

Second, the applied method for function point analysis, including the counting itself as performed by and SUPPLIER S and the review by COMPANY C, is appreciated highly by both respondents of both parties.

Third, coverage of FSM-pricing with regard to COMPANY C's IT-portfolio is experienced as too low, mainly by managers from COMPANY C. Additional analysis of the measure of coverage of FSM-pricing with regard to IT-portfolio shows that at finalization of the FSM-pricing pilot 27% of all IT-portfolio costs where calculated based on FSM-pricing. Analysis showed that a goal for 55% coverage within one year might be achievable. The remaining 45% is among others related to infrastructure (19%), support (17%), third party projects (5%) and small innovations (3%).

Fourth, developers from SUPPLIER S are dissatisfied with the reliability of FSM-pricing. The major reason for this seems to be that they experience little possibilities to incorporate non-functional requirements and complexity in project proposals. From a statistical point of view all projects are treated as average, where non-functional requirements and related complexity are incorporated in both trend lines. Apparently this approach is difficult to handle for developers, possibly because – in the period before the FSM-pricing pilot – they were consulted to come up with expert judgment for estimation purposes. To finalize our discussion on RQ1; an additional positive signal with regard to this is that after evaluation of the FSM-pricing pilot both COMPANY C and SUPPLIER S agreed upon continuation of the approach as an operational practice.

With regard to RQ2 (To what extent does FSM-pricing help to improve transparency of project proposals?) a noteworthy finding was that a large majority (84%) of the respondents of the survey experienced that transparency of project proposals is improved during the FSM-pricing pilot. We observed that the majority of discussions moved from effort (and price) estimate to waiver requests and getting requirements ready for function point analysis. Noteworthy is that function point analysis seems to have a positive effect on requirements management.

Looking at RQ3 (To what extent does FSM-pricing help to create continuous improvement?) quantitative analysis of the performance of the COMPANY C projects taught us that the performance (Time-to-Market and Productivity) is not improving over time.

A remark is in place with regard to economy-of-scale. Regression analysis of finalized projects shows that bigger projects are cheaper and faster than smaller ones. However, in our repository no projects larger than approximately 2,000 function points are found, and large projects are certainly not a majority. We therefore consider that estimation of larger projects as unreliable: FSM-pricing is limited to the maximum size in its repository.

A worrying issue is that Time-to-Market is not in sync with external peer groups. We assume that the most important reason for the high Time-to-Market is the fact that within projects extreme waiting time and waist is hidden due to the combination of projects in releases.

B. Implications

From our analysis of related work, it is clear that pricing in itself is a topic that has received little attention from the research community. Yet pricing is a topic of great practical value, which strongly affects the outcome (success or failure) of a software development project. The many budget overruns reported for such projects, may very well be more attributable to inadequate pricing than to poor project execution.

Our research shows that an evidence-based approach, in which historical data on key performance indicators are used in combination with a simple (power) regression, can lead to prices that are satisfactory to both suppliers and commissioning parties. Our research emphasizes a holistic approach, in which pricing is considered for the full IT portfolio of an organization, possibly in combination with a supplier in an outsourcing relation. While any organization can adopt this pricing approach, one prerequisite is the availability of historical project data. This implies that the approach is only applicable to organizations willing and capable to aim for a long term solution.

The need for historical project data is likely also one of the causes why pricing has received limited attention in the research community; few researchers have access to such data. A way out of this dilemma may be opening up performance data for government-funded projects, making them available for researchers. Besides bringing new research insights, this might also help governments to reach more adequate prices for their IT projects.

The research presented opens up a number of avenues for further research. From a benchmarking perspective, our current approach distinguishes between data-warehousing and CRM / Billing projects. Further research is needed to come up with general guidelines on how to group projects into sufficiently cohesive units to permit adequate pricing. Another concern that arose from our case study is dealing with non-functional requirements such as security or infrastructure.

Approaches like COCOMO 2 introduce factors to compensate for such project characteristics, but whether this works well in combination with the purely statistical approach investigated in the present paper calls for additional research.

VII. LIMITATIONS

The reader should consider several limitations when interpreting our results. First, the survey has limited generalizability due to the limitation of respondents to 25 stakeholders. Determination of survey topics was done by members of both measurement teams. The limited number of survey topics was determined by length of survey (10-minutes). Further, the results of the ratings within the survey have to be looked upon with low significance in mind. We did not ask respondents to connect their open ended text data with the answers given in the rating part of the survey.

Second, we conducted the study only within COMPANY C and SUPPLIER S, so the results may not generalize elsewhere. Since we did not find any other study on a comparable single, statistical pricing approach, we cannot predict what the outcome of our method will be in other companies [20].

Third, our study focused on transparency of proposals and continuous improvement. The respondents might have been influenced by this focus and emphasize these aspects in their answers.

VIII. CONCLUSIONS

The key contributions of this paper are:

RQ1: We demonstrate that FSM-pricing can successfully be used in practice, where customer and supplier are different companies, as a statistical, evidence-based pricing approach for software engineering project proposals.

RQ2: We show that using FSM-pricing as a single instrument, which means without intervention of expert judgment-based opinions, leads to an improved transparency of project proposals and satisfied stakeholders from both the customer and the supplier.

RQ3: We demonstrate that using FSM-pricing does not in the short term lead to continuous improvement. Although Productivity shows to be in line with external peer groups, Time-to-Market is too high when benchmarked externally and this shows a deteriorating trend. However, on a midterm continuous improvement is expected to be achievable due to contractual agreements on yearly adjustments of baselines.

ACKNOWLEDGMENTS

We thank both COMPANY C and SUPPLIER S for their generosity to agree on using project and survey data for our study. Furthermore we thank Philippe Kruchten, Frank Vogelezang, Kim Herzig, and all other reviewers for their valuable feedback.

References

[1] M. Jørgensen, "A review of studies on expert estimation of software development effort," *The Journal of Systems and Software,* vol. 70, no. IEEE, pp. 37-60, 2004.

[2] B. Boehm, "Software Engineering Economics," *IEEE Transactions on Software Engineering,* vol. 10, no. 1, pp. 7-19, 1984.

[3] K. Moløkken and M. Jørgensen, "A Review of Surveys on Software Effort Estimation," in *IEEE Proceedings of the 2003 International Symposium on Empirical Software Engineering (ISESE'03)*, 2003.

[4] C. Verhoef, "Quantitative IT Portfolio Management," *Elsevier - Science of Computer Programming,* vol. 45, no. 1, pp. 1-96, 2002.

[5] R. Glass, Facts and Fallacies of Software Engineering, Addison Wesley, 2002.

[6] A. Abran, I. Silva and L. Primera, "Field studies using functional size measurement in building estimation models for software maintenance," *Journal of Software Maintenenace and Evolution: Research and Practice,* vol. 14, no. John Wiley & Sons, Ltd., pp. 31-64, 2002.

[7] C. Gencel and O. Demirors, "Functional Size Measurement Revisited," *ACM Transactions on Software Engineering and Methodology,* vol. 17, no. 3, pp. 15:1-15:36, June 2008.

[8] IFPUG, IFPUG FSM Method: ISO/IEC 20926 - Software and systems engineering – Software measurement – IFPUG functional size measurement method, New York: International Function Point User Group (IFPUG), 2009.

[9] F. Heemstra, "Software cost estimation," *Information and Software Technology,* vol. 34, no. 10, pp. 627 - 639, 1992.

[10] I. Sommerville, Software Engineering (9th Edition), Boston, USA: Addison-Wesley, 2010.

[11] B. Boehm, Software Engineering Economics, Englewood Cliffs, NJ: Prentice-Hall, 1981, ISBN 0-13-822122-7.

[12] B. Boehm, C. Abts and S. Chulani, "Software development cost estimation approaches - a Survey," *Annals of Software Engineering,* vol. 10, no. J.C. Baltzer AG, Science Publishers, pp. 177-205, 2000.

[13] B. Boehm, C. Abts, A. Winsor Brown, S. Chulani, B. K. Clark, E. Horowitz, R. Madachy, D. J. Reifer and B. Steece, Software Cost Estimation with COCOMO II, Englewood Cliffs, NJ: Prentice-Hal, 2000, ISBN 0-13-026692.

[14] L. Putnam and W. Meyers, Five Core Metrics, The Intelligence Behind Succesfull Software Management, New York: Dorset House Publishing, 2003.

[15] L. Fischman, K. McRitchie and D. Galorath, "Inside SEER-SEM," *Crosstalk - The Journal of Defense Software Engineering,* vol. April, pp. 26-28, 2005.

[16] A. M. Jenkins , J. D. Naumann and J. C. Wetherbe, "Empirical Investigation of Systems Development Practices and Results," *North-Holland Information and Management,* vol. 7, no. Elsevier Science Publishers B.V, pp. 73-82, 1984.

[17] D. D. Phan, "Information systems project management: An integrated resource planning perspective model," The University of Arizona, USA, 1990.

[18] F. Heemstra and R. Kusters, "Function point analysis: evaluation of a software cost estimation model," *European Journal of Information Systems,* Vols. 1, 4, no. Operational Research Society Ltd., pp. 229-237, 1991.

[19] A. L. Lederer and J. Prasad, "Nine Management Guidelines for Better Cost Estimating," *Communications of the ACM,* vol. 35, no. ACM, pp. 51-59, 1992.

[20] A. L. Lederer and J. Prasad, "Information systems software cost estimating: a current assessment," *Journal of Information Technology,* vol. 8, no. Palgrave Macmillan, pp. 22-33, 1993.

[21] F. Bergeron and J.-Y. St-Arnaud, "Estimation of information systems development efforts: a pilot study," *Information & Management,* vol. 22, no. Elsevier, pp. 239-254, 1992.

[22] A. Albrecht, "Measuring Application Development Productivity," in *Joint Share Guide, and IBM Application Development Symposium 14-17 October 1979*, Monterey, California, 1979.

[23] R. E. Fairley, "Recent Advances in Software Estimating Techniques," *ACM ICSE '92 Proceedings of the 14th International Conference on Software Engineering,* 1992.

[24] C. Dekkers and P. Forselius, "Scope Management: 12 Steps for ICT Program Recovery," *CROSSTALK The Journal of Defense Software Engineering,* vol. January / February, pp. 16-21, 2010.

[25] B. Czarnacka-Chrobot, "Rational pricing of business software systems on the basis of functional size measurement: a case study from Poland," in *Proceedings 7th Software Measurement European Forum (SMEF)*, Rome, Italy, 2010.

[26] C. Verhoef, "Quantifying the value of IT-investments," *Science of Computer Programming , Volume 56,* pp. P.315-342, 2005.

[27] A. Abran, A. Sellami and W. Suryn, "Metrology, Measurement and Metrics in Software Engineering," in *IEEE Proceedings of the Ninth International Software Metrics Symposium (METRICS'03)*, 2003.

[28] H. Huijgens, R. van Solingen and A. van Deursen, "How To Build a Good Practice Software Project Portfolio?," *ICSE Companion 2014 Companion Proceedings of the 36th International Conference on Software Engineering (SEIP),* vol. 2014, no. IEEE, pp. 64-73, 2014.

[29] H. Huijgens and R. van Solingen, "Measuring Best-in-Class Software Releases," *IWSM-MENSURA 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 2013 Eighth International*

Conference on Software Process and Product Measurement, no. IEEE, pp. 137-146, 2013.

[30] H. Huijgens, G. Gousios and A. van Deursen, "Pricing via Functional Size: A Case Study of 77 Outsourced Projects - Technical Report TUD-SERG-2014-012," Delft University of Technology, Delft, The Netherlands, 2014.

[31] B. Flyvbjerg, "Five Misunderstandings About Case-Study Research," *Qualitative Inquiry,* Vols. 12, 2, no. Sage Publications, pp. 219-245, 2006.

[32] NESMA, NESMA functional size measurement method conform ISO/IEC 24570, version 2.2, Netherlands Software Measurement User Association (NESMA), 2004.

[33] NESMA, "FPA applied to Data Warehousing, version 1.2," Netherlands Software Metrics User Association, 2014.

[34] P. Runeson, M. Host, A. Rainer and B. Regnell, Case Study Research in Software Engineering; Guidelines and Examples, Hoboken, New Jersey. USA: John Wily & Sons, 2012.

I. Addendum

A. Survey Results - Summary

	Count	Completed / Started	Completed / Viewed	Started / Viewed
Completed	25	68.21%	78.13%	
Started	29			90.63%
Viewed	32			

What company are you working for?		
Company C	11	44.00%
Supplier S	14	56.00%
Other	0	0%
Total	25	

What is your connection with Evidence-Based Pricing?		
Overall IT Management	7	28.00%
FPA / Measurement & Analysis Team	4	16.00%
Portfolio Management	3	12.00%
DWH Team	3	12.00%
CRM / Billing Team	6	24.00%
Other (Release Management; IT Portfolio Management)	2	8.00%
Total	25	

1) Survey Results – Mean Likert Scores

Nr.	How satisfied are you with the following?	Very unsatisfied	Unsatisfied	Neutral	Satisfied	Very Satisfied	Don't know	Mean	Standard deviation	Variance
S01	Preparation of the FSM-pricing-pilot	0	3	4	13	4	1	3.75	0.90	0.80
S02	FSM-pricing pilot period itself	0	0	5	16	2	2	3.87	0.55	0.30
S03	Communication with regard to FSM-	0	1	13	8	1	2	3.39	0.66	0.43
S04	Management commitment on FSM-	0	2	13	6	3	1	3.42	0.83	0.69
S05	Reliability of the FSM-pricing method	0	6	8	9	2	0	3.28	0.94	0.88
S06	Setup of the Supplier baseline	0	6	5	11	1	2	3.30	0.93	0.86
S07	Pricing table for DWH	0	1	7	7	1	9	3.50	0.73	0.53
S08	Pricing table for CRM/Billing	0	3	8	6	1	7	3.28	0.83	0.68
S09	Function Point Analysis method	0	1	5	12	6	1	3.96	0.81	0.65
S10	Waiver procedure for FPA (exclusions)	0	7	7	7	3	1	3.25	1.03	1.07
S11	Coverage of FSM-pricing	0	6	6	10	1	2	3.26	0.92	0.84
S12	Proposal process (with regard to FSM-	0	5	5	13	1	1	3.42	0.88	0.78
S13	Advantages of FSM-pricing for Company	0	1	6	14	1	3	3.68	0.65	0.42
S14	Advantages of FSM-pricing for Supplier	0	2	8	10	0	5	3.40	0.68	0.46
S15	Overall FSM-pricing	0	1	8	13	3	0	3.72	0.74	0.54

Nr.	How would you rate the quality of the following?	Very poor	Poor	Average	Good	Excellent	Don't know	Mean	Standard deviation	Variance
Q01	Requirements delivered by Company	0	2	10	13	0	0	3.44	0.65	0.42
Q02	FPA performed by Supplier	0	1	5	15	3	1	3.83	0.70	0.49
Q03	FPA reviewed by Company	0	0	7	14	2	2	3.78	0.60	0.36
Q04	The DWH baseline used for FSM-pricing	0	2	4	8	0	11	3.43	0.76	0.57
Q05	The CRM/Billing baseline used for FSM-	0	2	6	8	1	8	3.47	0.80	0.64
Q06	The Supplier Proposals based on FSM-	0	1	11	12	1	0	3.52	0.65	0.43
Q07	The overall FSM-pricing method	0	0	10	14	1	0	3.64	0.57	0.32

Nr.	To what extent did you experience a change on the following performance indicators during the FSM-pricing pilot?	Strongly deteriorated	Deteriorated	Neutral	Improved	Greatly improved	Don't know	Mean	Standard deviation	Variance
E01	Transparency of Proposals	0	2	2	18	3	0	3.88	0.73	0.53
E02	Productivity (Cost per FP)	0	1	16	5	2	1	3.33	0.70	0.49
E03	Time-to-Market (Days per FP)	1	3	13	5	0	3	3.00	0.76	0.57
E04	Process Quality (Defects per FP)	0	1	13	3	1	7	3.22	0.65	0.42

2) Survey Results – Likert Scores per Respondent

Results on survey question X01: Should FSM-pricing be continued as an operational practice once the pilot is finalized?

topic Nr.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25
X01	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	1	3	3	1	1

Remarks with 3 above:

- P22: To be embedded in agile / scrum;
- P23: Ok but considering the improvement points mentioned above.

Results on survey questions S01 to S15: How satisfied are you with the following?

Topic Nr.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25
S01	4	5	4	4	3	3	5	4	2	4	4	4	3	4	4	4	2	6	5	4	4	3	4	5	2
S02	5	4	4	4	3	3	4	3	4	4	4	4	4	4	4	3	4	6	5	3	4	4	4	4	6
S03	4	5	4	4	4	3	4	3	3	3	3	3	4	3	4	3	3	6	4	3	3	3	3	2	6
S04	5	3	5	4	5	4	4	3	4	3	2	3	3	3	3	3	2	6	4	3	3	4	3	3	3
S05	4	3	4	4	2	4	4	3	2	3	3	4	2	4	3	2	2	2	4	3	4	3	3	5	5
S06	4	4	4	3	2	3	4	2	2	4	6	4	3	4	4	2	2	2	4	4	4	3	3	5	6
S07	5	4	4	4	4	3	3	3	2	4	6	4	6	3	6	3	6	6	6	3	4	6	3	6	6
S08	5	4	4	6	3	3	3	6	2	4	6	3	2	4	6	3	2	6	4	3	4	3	3	6	6
S09	4	5	5	4	3	4	5	4	4	4	4	4	3	5	4	2	4	3	5	4	4	3	3	5	6
S10	2	5	3	3	2	3	5	3	5	4	4	2	4	3	4	2	4	6	4	4	2	2	2	3	3
S11	2	5	2	6	4	3	4	3	3	4	3	4	3	4	2	2	3	6	4	4	2	2	4	4	4
S12	4	4	4	6	2	4	4	3	4	4	4	2	4	3	3	2	2	3	5	4	3	4	4	4	2
S13	5	3	4	4	2	4	3	4	6	3	3	4	3	4	3	4	4	3	4	4	4	6	4	4	6
S14	3	4	4	6	4	3	3	6	6	3	3	4	3	4	3	3	2	2	4	4	4	6	3	6	4
S15	5	5	4	6	3	3	4	3	4	4	3	4	3	4	3	3	2	3	5	4	4	3	4	4	4

Results on survey questions Q01 to Q07: How would you rate the quality of the following?

Topic Nr.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25
Q01	3	4	4	3	2	3	4	3	4	4	3	4	3	3	3	3	4	2	3	4	4	4	4	4	4
Q02	5	5	4	4	4	2	5	4	4	4	4	4	3	4	4	3	4	3	4	4	3	6	4	4	3
Q03	4	5	4	4	3	4	5	4	4	4	4	4	3	4	4	3	4	3	4	6	3	4	6	3	3
Q04	4	4	4	4	4	6	4	3	2	4	6	3	6	6	6	3	6	2	6	6	3	6	6	4	6
Q05	4	4	4	6	2	6	4	6	2	4	6	3	3	4	6	4	3	6	5	3	3	3	6	4	6
Q06	4	4	4	3	4	3	4	3	3	4	3	4	3	4	4	3	3	2	5	4	4	3	3	4	3
Q07	4	4	4	4	3	3	4	3	4	4	3	4	3	4	4	3	3	3	5	4	3	3	4	4	4

Results on survey questions E01 to E04: To what extent did you experience a change on the following performance indicators during the FSM-pricing pilot?

Topic Nr.	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25
E01	4	5	4	5	3	4	5	4	4	3	4	4	2	4	4	2	4	4	4	4	4	4	4	4	4
E02	5	3	3	3	3	6	3	3	4	4	3	4	3	5	3	2	3	3	4	3	4	3	3	3	3
E03	1	3	4	2	2	6	3	3	3	3	3	4	3	4	3	4	3	3	4	3	3	4	6	3	2
E04	3	3	3	6	2	6	3	3	6	3	3	4	3	6	3	3	3	3	5	3	4	6	6	6	6

3) Survey Results – Means

The survey questions S01 to S15, Q01 to Q07, and E01 to E04 were asked in the order as listed below.

Effect Size was calculated as two unstandardized measures; 1) for each survey topic the difference between the mean COMPANY C score and the mean SUPPLIER S score, and 2) for each survey topic the difference between the mean Management score (all scores of respondents with the profile Overall IT-management, FPA Measurement Team, Portfolio Management, and Other) and Development (all scores of respondents with the profile Data Warehouse Team, and CRM/Billing Team).

Effect size does not make any statement about whether the apparent relationship in the data reflects a true relationship in the population. In that way, effect size complements inferential statistics such as a p-value. For this purpose we calculated the p-value for each survey topic using a Wilcoxon rank-sum test. However, analysis showed that the p-values were statistically not relevant (none of the topics showed a statistically significant difference between COMPANY C and SUPPLIER S and Management and Development). Therefore we did not include these in our paper.

Nr.	Survey Topic	Mean Overall	Mean Company	Mean Supplier	P-value Company / Supplier	Effect Size Company / Supplier	Mean Management	Mean Development	P-value Management / Development	Effect Size Management / Development
S01	Preparation of the FSM-pricing pilot	3.75	3,82	3,69	0.744	0.13	3,75	3,75	0.645	0.00
S02	FSM-pricing pilot period itself	3.87	3,91	3,83	0.701	0.08	3,80	4,00	0.371	-0.20
S03	Communication with regard to FSM-	3.39	3,36	3,42	0.588	-0.05	3,47	3,25	0.663	0.22
S04	Management commitment on FSM-	3.42	3,64	3,23	0.475	0.41	3,50	3,25	0.975	0.25
S05	Reliability of the FSM-pricing method	3.28	3,55	3,07	0.217	0.47	3,31	3,22	0.905	0.09
S06	Setup of the Supplier baseline	3.30	3,55	3,08	0.817	0.46	3,36	3,22	0.385	0.13
S07	Pricing table for DWH	3.50	3,86	3,22	0.406	0.63	3,55	3,40	0.744	0.15
S08	Pricing table for CRM/Billing	3.28	3,57	3,09	0.192	0.48	3,36	3,14	0.500	0.22
S09	Function Point Analysis method	3.96	4,00	3,92	0.814	0.08	4,00	3,89	0.485	0.11
S10	Waiver procedure for FPA (exclusions)	3.25	3,00	3,46	0.223	-0.46	3,38	3,00	0.838	0.38
S11	Coverage of FSM-pricing	3.26	2,70	3,69	0.034	-0.99	3,13	3,57	0.057	-0.45
S12	Proposal process (with regard to FSM-	3.42	3,70	3,21	0.144	0.49	3,44	3,38	0.782	0.06
S13	Advantages of FSM-pricing for Company	3.68	3,80	3,58	0.855	0.22	3,57	3,88	0.432	-0.30
S14	Advantages of FSM-pricing for Supplier	3.40	3,29	3,46	0.488	-0.18	3,46	3,29	1.000	0.18
S15	Overall FSM-pricing	3.72	3,64	3,64	0.508	0.00	3,75	3,67	0.975	0.08
Q01	Requirements delivered by Company	3.44	3,45	3,43	0.878	0.03	3,44	3,44	0.924	-0.01
Q02	FPA performed by Supplier	3.83	3,70	3,93	0.926	-0.23	3,81	3,88	0.541	-0.06
Q03	FPA reviewed by Company	3.78	3,73	3,83	0.376	-0.11	3,75	3,86	0.155	-0.11
Q04	The DWH baseline used for FSM-pricing	3.43	3,71	3,14	0.953	0.57	3,55	3,00	0.251	0.55
Q05	The CRM/Billing baseline used for FSM-	3.47	3,57	3,40	0.608	0.17	3,45	3,50	0.977	-0.05
Q06	The Supplier Proposals based on FSM-	3.52	3,55	3,50	0.879	0.05	3,56	3,44	0.659	0.12
Q07	The overall FSM-pricing method	3.64	3,55	3,71	0.530	-0.17	3,56	3,78	0.456	-0.22
E01	Transparency of Proposals	3.88	3,82	3,93	0.703	-0.11	4,11	3,75	0.299	0.36
E02	Productivity (Cost per FP)	3.33	3,40	3,29	0.461	0.11	3,44	3,27	0.843	0.17
E03	Time-to-Market (Days per FP)	3.00	2,78	3,15	0.789	-0.37	3,29	2,87	0.104	0.42
E04	Process Quality (Defects per FP)	3.22	3,25	3,20	0.881	0.05	3,67	3,00	0.164	0.67

4) Survey Results – Open Ended Text per respondent

In the overview below all open text answers given per respondent (P1 to P25) are listed.

Resp	Org	Role	What is going well during the FSM-pricing pilot that we want to continue?	What is not going well during the FSM-pricing pilot that we want to fix?	What can we do to improve FSM-pricing?
P1	Company C	FPA / Measurement & Analysis Team	Preparation of FPA's by Supplier S; Review of FPA's by Company C; Pricing of proposals based on FP's;	Too small projects are proposed (instead of bundling to larger projects / releases); Duration of planned (forecasted) projects is too long;	Plan releases as a whole based on FP's; Improve project / release administration (better tooling needed); Beside cost, focus at duration too (improve time-to-market);
P2	Supplier S	FPA / Measurement & Analysis Team	The solution is looked into more detail in order to get the right function points at the proposal stage itself. This helps in early detection of issues and resolution; Uniformity in pricing approach as it does not depend on individual components to derive their efforts;	Lack of understanding on how Function points are calculated often raises questions of why higher Function points are generated even though less work is foreseen and this is looked upon as higher cost for low efforts; All the ZIPs do not have NFRs or technology wise the efforts differ but by including all the ZIPs in benchmarking exercise and by measuring them with same scale we may have compared apples with oranges;	NFRs should not be included in benchmarking exercise as function points are not counted for NFRs and not all ZIPs have NFRs; Technology wise Benchmarks would give fair FP based price;
P3	Company C	Overall IT-management	Quick pricing definition; clear overview of productivity; Interesting pricing for both pricing;	Too much small project - negative for Company C; Need to extend footprint of FP pricing;	Extend knowledge of FP within Company C;
P4	Company C	DWH Team	Transparency on WBS (entities) versus FP; FP count knowledge (both sides);	Keeping to agreements. Always reopening the discussion; current FP pricing setup maybe not fully aligned with Agile set-up we are looking into (story points); Cost discussion (past delivery vs current);	
P5	Supplier S	Overall IT-management	Efforts for some projects are reduced; Efforts for some projects are more towards realistic figures; Standard process;	FP is not applicable to some projects where more testing efforts are required for less development changes; Delaying overall proposal submission process; For some projects cost also increased due to FP calculation;	FP mapping should be different for different applications; Should be possible to get waiver project to project and not only for support projects;
P6	Company C	Portfolio Management	We are enforcing a uniform and transparent way of measuring complexity of projects/releases; The measurements are doing twice so Supplier S can be challenged on their numbers;	Not everyone is (fully) aware of how the measurements are done;	Provide some more insight in how these measurements are being done. That way people understand how and why this is being done;

P7	Supplier S	FPA / Measurement & Analysis Team	For calculating FP, clarity of requirement is essential. As FP is being calculated at the beginning of Project it ensures better transparency about expectations, minimizes scope creep and can be used for revalidating actual work delivered; Pilot phase helps in gauging the validity of pricing table. Analysis would show whether separate pricing table should be defined for different components currently considered within a single category. Detailed scrutiny would help define more realistic pricing tables reducing estimation effort in future; Easy to capture outliers and carry out root cause analysis on such cases;	Ensuing complete clarity towards the beginning of Project occasionally results in delay in finalizing the WO; Concerns on FP based Trendline being "realistic" is prevalent. The gap would have to be reduced to ensure more realistic Trendline;	Complete requirement clarity should be available before D2J phase; Refinement of Pricing table based on Technology etc to make it more realistic; Cross verification with actuals towards the end of Project to revalidate the estimates;
P8	Company C	Portfolio Management	Transparency; Less discussion;	Supplier S divides the offer in small pieces we must have release based funding; Not everybody understands that this is a statistical method more info must be provided towards all involved stakeholders;	
P9	Company C	Other: External IT-Sourcing Consultant			
P10	Supplier S	Overall IT-Management	Simplified pricing mechanism;	End-2-End Test efforts; Non-functional percentages; Technology specific FP;	
P11	Supplier S	Portfolio Management	FP review with Company C and Supplier S FP teams before proposal submission;	Sometimes FP amount is not consistent / comparable with WBS pricing matrix;	Not sure but may be detailed analysis of cases where FP pricing and WBS pricing matrix there is huge difference and come up with some rationalization if possible;
P12	Supplier S	DWH Team	Calculation for overall DV and PA area working out well; Reviews are detail and comprehensive;	To work out FP estimates, detailed entities and flows required in solution; FP Cost gets apportioned across non FP components as well due to which FP fetching component get reduced efforts. Overhead areas to be considered separately; Reviews takes somewhat longer time to conclude;	Agreed and simpler form of template would be advisable so that estimates can be quick; IA/ODS area FPs are complex due to technology, mechanism;
P13	Company C	Portfolio Management	Timely delivery of FP calculations for proposals;		Transparency on how FPs are calculated; Introduce FP-like calculation for infra;

P14	Supplier S	CRM / Billing Team	Because of product sizing estimates are more accurate most of the time; Avoid delays and budget overruns as estimation can be done at initial stage as against task base estimates (provided requirement is cleared up to certain level);	Currently all the application within specific domain are having same cost/FP. But some application are doing functional and technical design and some are doing only technical;	FP should consider only for factory efforts. Non Functional efforts and post development efforts can also be excluded from FP for more accurate estimation; Good to exclude efforts required for functional design from FP (especially for small application) as sometime functional design/analysis required more efforts. Even though it is average out for medium size application (for small size application or requirement within release is small then difficult to accommodate efforts using FP); Would be good to have some provision to increase FP size up to certain percentage (i.e arrived FP*some factor) in case functionality to be delivered is more complex;
P15	Company C	Overall IT-Management	No big difference between Company C and Supplier S counting;	Involvement of SPOC per team not yet done;	Number of projects per release included in the process; Agreement on KPI related to FP;
P16	Supplier S	Overall IT-Management	Rolling out of standard FP calculation across every domain;	FP Pricing calculation (Mapping of FP to exact CI which will be delivered);	Identify right mapping of CIs' to FP points; FP waivers candidates;
P17	Supplier S	CRM / Billing Team	Keep on benchmarking the FP data to refine the baseline;	Most of the complexity of the code to be changed does not match with the amount of functionality to be changed causing a disparity; Resource loading would need to be done as per the WBS itself and not as per FP. This normally gives out a difference in the resource spend v/s demand. The law of average may not work out for each release; Measuring productivity for COTS is a challenge and grouping of various technologies into a similar baseline might also hurt later;	Gap analysis are a must to understand the outlier data points; Regular review of the data points and keep on correcting the baseline as per the increase data points; Performance testing / NFR's / Deployment should be excluded from the scope of FP;
P18	Supplier S	DWH Team	Most of details are sorted out at the time of proposals, Earlier these details were discussed in design phase; Instead of plain list of entities that we were maintaining in WBS entities we now have clarity on what kind of functionality is getting delivered;	To keep the counting simpler we are considering all the requirements are at average level. But we may need to perform gap analysis if the requirements mix is really averaging out on efforts. Earlier in WBS we were having complexity for each entity which was providing support for effort estimations; Area wise we still have different rules agreed upon which are not uniform in nature. Earlier, team was working on only DV and PA areas but we are also estimating on ODS and other areas like SFTP / DRS etc. Though we have worked out on some rules for counting no clarity on how we have set the base lining for these areas;	Bring uniformity in counting; Understand that in DWH we need to combine multiple flows so efforts are not uniform for each entity, so there will always be deviations;

P19	Supplier S	CRM / Billing Team	It's making standardization in estimation process; Process is well defined across domain wise; It would be helpful across deliveries in standard format instead of expert judgments to avoid deviations & dependencies;	FP counting is happening across deliverables so could be involved each component POC's instead of only Senior Management;	More Socialization; FP counting for configuration related work; FP counting for COTS products;
P20	Supplier S	CRM / Billing Team	Discussion between Product SME and FP Expert before arriving on FP count;	Not always the FP count justifies the development efforts;	Conduct more trainings to Product SME;
P21	Company C	Overall IT-Management	Process being followed seriously; Most of the proposal falling in the category of FP and delivered using FP pricing; Constructive argument in case of disagreement;	Shortage of expertise, one person responsible for validating and reviewing FP proposals; Multiple waiver discussion, it should not according to me; Still lot of projects are not covered in FP based pricing;	Awareness & training; Finalize the scope to avoid waiver related discussion; Released based FP based funding;
P22	Company C	CRM / Billing Team	Supplier offers include FP based quotations, consistently;	Billing scope is limited to small initiatives; Still very much reliant upon external expertise to validate the FP; SME's not really involved;	To be embedded in agile / scrum;
P23	Supplier S	CRM / Billing Team	FP as an additional yardstick for WBS-based estimation;	Certain actual on-the-ground efforts need to be secured. E.g. Consulting/component level Analysis, proposal and design-level support to components like DWH, test support, etc; Efforts in technical enhancement wherein end user is not directly impacted should be secured;	Efforts in technical enhancement wherein end user is not directly impacted should be secured; Certain actual on-the-ground efforts need to be secured. E.g. Consulting/component level Analysis, proposal and design-level support to components like DWH, test support, etc; Baseline definition as per actual implementation at Company C;
P24	Company C	FPA / Measurement & Analysis Team	The general FP principles; The way FP counting is done by Supplier S;	Lack of adherence to the principle of using FP for evaluations; Lack of trust at Company C in the counting made by Supplier; The review mechanism is not transparent;	Increase buy-in at Company C thanks to training, better communication, etc; Create transparency on the FP mechanisms and their execution; Create a group of reference people, assigned to pilot all the FP subject: 'reference' in case of discussion regarding FP, 'controller' on FP, etc;
P25	Supplier S	Overall IT-Management	A single point for the final estimation, answerable to all stakeholders. Estimation review process becomes very simple; A standardized process, which can be trusted from both vendor and client stakeholders; As there is less manual intervention unlike traditional WBS based pricing, there is less scope of error;	The strict timeline for proposal submission process, And the FP estimation and approval process is another link being introduced in the already tight schedule. That is leading to delay in proposal submission timelines; There is lack of awareness among the stakeholders in understanding the FP based estimation process. Awareness among the stakeholders needs to increased;	Making FP based estimations mandatory for all stakeholders who are directly involved with FP pricing process. That would reduce lots of un-necessary questions being asked on FP estimation data points; Increasing the proposal submission timeline;

B. Project Details

Due to compliancy reasons no details of individual projects such as Project Cost, Project Duration, Build & Test Cost are included in this Technical Report. We give an overview of portfolio data, summarized at the level of the four different Project Categories mentioned in the paper:

1. *Baseline*; projects that were performed within COMPANY C; yet with other external suppliers than SUPPLIER S.
2. *Repository*; projects that were performed prior the FSM-pricing pilot within COMPANY C; Build & Test was performed by SUPPLIER S.
3. *Pilot*; projects that were performed in the scope of the FSM-pricing pilot within COMPANY C; Build & Test was performed by SUPPLIER S.
4. *Forecast*; projects that were forecasted (ongoing projects) during the FSM-pricing pilot within COMPANY C; Build & Test was performed by SUPPLIER S.

1) Portfolio Summary – Project Size

Project Size Measurements (Function Points)	Overall	Baseline (Bl)	Repository (Rp)	Pilot (Pi)	Forecast (Fc)
Number of projects (n)	77	22	16	10	29
Minimum Project Size (FP)	4.00	11.00	17.00	4.00	5.00
1st Quadrant Project Size (FP)	16.00	44.00	38.25	6.75	9.00
Median Project Size (FP)	38.00	90.00	71.50	17.50	22.00
Mean Project Size (FP)	117.20	182.60	156.59	36.40	79.17
3rd Quadrant Project Size (FP)	117.00	206.50	115.25	55.25	37.00
Maximum Project Size (FP)	1089.00	773.00	1089.00	130.00	716.00

2) Portfolio Summary – Project Cost

Project Cost Measurements (Euros)	Overall	Baseline (Bl)	Repository (Rp)	Pilot (Pi)	Forecast (Fc)
Minimum Project Cost (Initiation to Technical Go Live)	4,281	18,893	45,592	4,281	13,842
1st Quadrant Project Cost (Initiation to Technical Go Live)	37,570	46,992	124,357	10,270	25,168
Median Project Cost (Initiation to Technical Go Live)	76,957	88,544	258,145	20,569	59,681
Mean Project Cost (Initiation to Technical Go Live)	212,408	271,133	415,120	93,190	67,337
3rd Quadrant Project Cost (Initiation to Technical Go Live)	245,435	311,218	350,327	58,495	79,293
Maximum Project Cost (Initiation to Technical Go Live)	1,867,856	1,683,060	1,867,856	559,460	281,703

3) Portfolio Summary – Build & Test Cost

Build & Test Cost Measurements (Euros)	Overall	Baseline (Bl)	Repository (Rp)	Pilot (Pi)	Forecast (Fc)
Minimum B&T Cost (Execution to Technical Go Live)	4,281	12,948	15,620	4,281	13,842
1st Quadrant B&T Cost (Execution to Technical Go Live)	24,621	24,264	66,779	10,270	22,930
Median B&T Cost (Execution to Technical Go Live)	64,783	69,805	130,790	20,569	54,653
Mean B&T Cost (Execution to Technical Go Live)	118,818	148,369	209,449	38,884	61,324
3rd Quadrant B&T Cost (Execution to Technical Go Live)	133,145	162,889	188,072	52,289	76,957
Maximum B&T Cost (Execution to Technical Go Live)	1,312,624	689,249	1,312,624	159,790	281,703

4) Portfolio Summary – Project Duration

Project Duration Measurements (Months)	Overall	Baseline (Bl)	Repository (Rp)	Pilot (Pi)	Forecast (Fc)
Minimum Project Duration (Initiation to Technical Go Live)	2.920	4.400	4.040	2.920	3.550
1st Quadrant Project Duration (Initiation to Technical Go Live)	6.930	6.772	9.178	5.327	6.740
Median Project Duration (Initiation to Technical Go Live)	8.120	7.365	10.415	7.770	7.360
Mean Project Duration (Initiation to Technical Go Live)	9.434	9.489	12.115	8.576	7.667
3rd Quadrant Project Duration (Initiation to Technical Go Live)	10.520	11.445	14.555	9.988	9.300
Maximum Project Duration (Initiation to Technical Go Live)	26.820	19.780	26.820	21.590	12.290

5) Portfolio Summary – Build & Test Duration

Build & Test Duration Measurements (Months)	Overall	Baseline (Bl)	Repository (Rp)	Pilot (Pi)	Forecast (Fc)
Minimum Build & Test Duration (Execution to Technical Go Live)	0.490	2.330	1.710	0.490	3.06
1st Quadrant Build & Test Duration (Execution to Technical Go Live)	3.980	4.025	5.695	2.155	4.27
Median Build & Test Duration (Execution to Technical Go Live)	5.420	5.190	7.295	3.300	5.19
Mean Build & Test Duration (Execution to Technical Go Live)	5.606	5.354	7.209	3.808	5.15
3rd Quadrant Build & Test Duration (Execution to Technical Go Live)	6.640	6.460	8.950	5.405	5.59
Maximum Build & Test Duration (Execution to Technical Go Live)	13.670	10.220	13.670	7.690	7.07

TUD-SERG-2014-012
ISSN 1872-5392