Science of Computer Programming 77 (2012) 314–354

Contents lists available at SciVerse ScienceDirect

Science of Computer Programming

journal homepage: www.elsevier.com/locate/scico

Quantifying forecast quality of IT business value

J.L. Eveleens [a,*], M. van der Pas [b], C. Verhoef [a]

[a] VU University Amsterdam, Department of Computer Science, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
[b] Vodafone, Portfolio & Innovation Management, Am Seestern 1, D-40547 Dusseldorf, Germany

ARTICLE INFO

Article history:
Received 24 August 2010
Received in revised form 31 March 2011
Accepted 6 July 2011
Available online 28 July 2011

Keywords:
IT business value
Forecast
Estimation
Reference cone
Forecast to actual ratio
Estimating Quality Factor
EQF
IT

ABSTRACT

This article discusses how to quantify the forecasting quality of IT business value. We address a common economic indicator often used to determine the business value of project proposals, the Net Present Value (NPV). To quantify the forecasting quality of IT business value, we develop a generalized method that is able to account for asymptotic cases and negative valued entities. We assess the generalization with real-world data of four organizations together consisting of 1435 IT assets with a total investment cost of 1232+ million Euro for which 6328 forecasts were made. Using the generalized method, we determine the forecasting quality of the NPV, along with the benefits and cost using real-world data of another 102 IT assets with a total business value of 1812 million Euro. For the real-world case study, we will find that the quality of the forecasted NPVs is lower than the forecasted benefits, which is again lower than the forecasting quality of the cost. Also, we perform a sensitivity analysis to investigate the impact on the quality of an asset's forecasted NPV when the forecasting quality of benefits or cost improves. Counterintuitively, it turned out in this case study that if the quality of cost forecasts would improve, the overall quality of its NPV predictions would degrade. This underlines the importance of both accurate cost and benefit predictions. Finally, we show how to use the quantified forecast information to enhance decision information using two simulation examples.

1. Introduction

Organizations need to undertake IT projects every year to consolidate and expand their business. These IT projects often play a profound role in modern companies due to their size and impact. This stresses the importance of their adequate management by the Chief Information Officer (CIO).

More and more, voices are raised [43,20] that the CIO must manage IT projects to maximize their business value and return, instead of controlling their cost. For instance, in 1998 a research project known as Beyond Budgeting was started to change the current management model [23]. Where previously the management was focussed on planning and control of cost and the technical aspect of IT, the research emphasizes the importance of the business value.

Pisello et al. [48] argue that the CIO needs to become the Chief Financial Officer (CFO) of IT to improve the organization's value. But, what does managing the business value of IT entail? An article by Lorie et al. [40] states that executives face three tasks in achieving good financial management, among which the correct forecasting of the expected cash flows. In this article, we focus on that particular task.

In information technology, the META group [20] showed that forecasting, especially of benefits, is far from common practice. Of the organizations that were surveyed, 84% indicated that no business cases were made for their IT projects or only for a select few projects. And if forecasts of the cash flows are made, the question is how to assess their validity.

* Corresponding author. Tel.: +31 20 5987760.
E-mail addresses: laurenz@few.vu.nl (J.L. Eveleens), mark.vanderpas@vodafone.com (M. van der Pas), x@cs.vu.nl (C. Verhoef).

0167-6423/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.scico.2011.07.010

That proper forecasting is indeed a challenging task, is shown, for instance, by a survey of the Kellogg School of Management [36]. The survey found that 82% of the responding CIO's regarded forecasting of IT benefits as a major challenge. On top of that, the survey showed that 68% do not track project benefits at all and only 26% track actual financial metrics after having made an investment decision. So, how do we know that forecasts of the IT business value are accurate, unbiased and reliable enough to support the decision making process? And are organizations able to adequately predict the business value of IT investments?

In this article we will address these questions. First, we discuss common economic indicators often used to determine the business value of proposals. We elaborate on one of these indicators, the Net Present Value (NPV). For this indicator, we will assess the forecasting quality. Since the overall forecasting quality of the NPV is determined by the forecasting quality of its components, such as benefits and cost, their forecasting quality is assessed as well.

We assess the quality of NPV forecasts, benefits and cost using real-world data of 102 IT assets. We obtained this data from a telecommunication organization, Z, that structurally makes business cases for its IT investments. Moreover, this organization uses procedures to evaluate completed investments. The data represents an NPV value of 1812 million Euro, with discounted benefits of 4714 million Euro and an investment value of 173 million Euro. Combined with data from other organizations, this article discusses 1620 IT assets with an investment cost of 1232+ million Euro for which in total 6513 forecasts are made.

For the data of organization Z, we determine the accuracy of the forecasts and check for potential biases. We found that the quality of the forecasted NPVs is lower than the forecasted benefits, which is again lower than the forecasting quality of the cost. Also, there turned out to be a significant difference in forecasting quality between assets classified as Cost Reduction or New Product Development. The NPV, benefits and cost forecasts of Cost Reduction assets were more accurately predicted than that of the New Product Development assets. Moreover, whereas the forecasts of Cost Reduction assets showed no biases, the New Product Development forecasts of benefits and cost were generally overestimated in the real-world case study.

We also performed a sensitivity analysis to investigate the impact on the quality of an asset's forecasted NPV when the forecasting quality of benefits or cost improves. Counterintuitively, it turned out that if the quality of the cost forecasts would improve, the overall quality of the NPV predictions would degrade. This is caused by the bias toward overestimation of both benefits and cost forecasts. The overestimation of the benefits is compensated for by an overestimation of the cost. This illustrates that increasing control over the cost without measures to ensure the quality of the benefits, may yield a lower forecasting quality of the overall NPV.

Finally, we will illustrate how it is possible to use the quantified forecasting quality further to enhance decision information. This is done by describing two basic simulation examples. The first example shows how to acquire additional information for rationing the capital budget over various project proposals. The second example provides insight in the forecasted business value that will be generated from project proposals when accounting for the forecasting quality and bias.

1.0.0.1. Generalized method. This article is self-contained, but also forms the final article in a triptych. To assess the forecasting quality of the NPV in this article, we make use of an existing method to quantify and visualize the quality of IT-forecasts [14]. That first article or left pane covered a method that assumes positive values, like budgets, durations, and size in function points or lines of code.

That research led to a second article or middle pane in IEEE Software [15] that allowed us to question existing research on the forecasting quality of important key performance indicators for IT-projects (cost, time, amount of functionality). In particular, we were able to question the validity of the Standish Chaos Report Figures.

This third and final article, the right pane, covers the case for forecasts and actuals that can take any value. By applying the method of our first article to real-world data of NPV calculations, it turned out the model was inadequate to deal with zero's and negative values. Indeed, it is natural for an NPV to be zero valued or negative.

Next to that, the actuals of all entities are only known with certainty after some time. For instance, the actual project cost are only known after the project is completed. Or, the actual NPV of an asset is known after the economic life span of an investment, for instance, 5 years after the initial forecast. However, to support decision making, we prefer recent information about the current forecasting practice; something that the quality of 5-year old forecasts will not provide for. Moreover, in case of some entities, for instance, the NPV, it is possible the actual will never be objectively measured. Therefore, we cannot always compare forecasts with actuals, making other reference points necessary. For these reasons, we need adaptations of the forecast assessment method of Eveleens et al. [14]. Note that the existing method is perfectly fit for positively valued indicators, so that method is not at all outdated by this article. But for business value a more sophisticated method is necessary.

Therefore, to assess the quality of NPV forecasts, in this article we extend and generalize the existing method that is intuitive for positively valued forecasts and actuals. We will develop a method that incorporates other reference points than the actual and allows for negative valued reference points or forecasts. Moreover, we resolve visual limitations in asymptotic cases.

At first sight, these issues appear trivial, but they are not. Adjusting the model to accommodate for these issues, entails an extensive and detailed discussion. For readers interested in the use of the model rather than the elaborate discussion of the adjustments, it is possible to skim the generalization of the method.

1.0.0.2. Data. In this article, we will make use of extensive data of four organizations. In total, we obtained data of 1620 IT assets with an investment cost of 1232+ million Euro for which 6513 forecasts were made.

Of a large financial service provider, Y, we use data consisting of 667 forecasts of 140 project costs and 83 functionality forecasts of 83 assets. A multinational organization, X, provided us with 3767 forecasts for 867 project costs. From Landmark Graphics, LGC, we obtained data containing 923 forecasts of 121 project durations. Finally, a large telecommunication organization Z provided 971 forecasts made for 307 project costs.

For the purpose of evaluating the forecasting quality of IT business value, we analyze another data set from the large telecommunication organization, Z. This data consists of 102 NPV forecasts made for 102 IT assets that together represent discounted benefits of 4714 million Euro and an investment value of 173 million Euro.

1.0.0.3. Related work. In statistical mathematics, assessing the quality of estimation methods is a well-discussed topic [16]. There are well-defined criteria that determine the quality of these methods. The generalized method we will develop makes use of such criteria and does not provide new statistical ways to determine forecasting quality. The generalized method is a conservative extension of the existing method, as it is in parts identical, yet extends the possibilities.

However, the statistical methods and metrics are often not accessible to IT governors. Therefore, in this article, we discuss how to present, summarize and visualize the IT forecasting accuracy in such a way that executives are able to assess their quality and use it to enhance decision information. With the generalized method we aim to make quantifying the forecasting quality more readily available to governors. Furthermore, the method allows executives to acquire knowledge on the forecasting quality of their organization.

IT forecasting methods and their accuracy are frequently discussed in the literature to achieve correct forecasting of project proposals. For example, many books [4,27,42,9,8,34] have been written describing issues and guidelines to achieve accurate estimates. Moreover, numerous estimation tools exist and are used in practice, among others COCOMO [4], SLIM [49], SEER, SPQR/20 and KnowledgePlan [28]. These tools assist in forecasting relevant project values, such as cost, effort and durations.

Numerous articles [6,30,32,54] compare different estimation methods to determine which of them are most accurate under certain circumstances. An article by Eveleens et al. [14] proposes a method to quantify the forecasting quality by assessing the accuracy and potential bias of predictions.

Yet, these articles do not address business value forecasts, but forecasts such as cost, size, functionality or duration. We are unaware of articles that quantify the quality of business value forecasts of IT investments. A book by Bower [35] did quantify the quality of NPV predictions for 50 assets in another industry.

In this article, we analyze forecasted and re-estimated NPVs of 102 IT assets to assess the accuracy of the initial forecasts of IT business value. We will compare our case study to the one described by Bower.

1.0.0.4. Organization of this article. In Section 2, we introduce terms and notations that will be used throughout the article. In Section 3 we discuss economic indicators that are used to quantify the business value of project proposals. More particularly, we cover in detail the well-known Net Present Value. For the interested reader, in Section 4 we extend an existing method to assess IT forecasting quality and make it fit to assess economic indicators. We generalize the method to allow for different reference points and negatively valued entities. We assess the impact of the generalization using data of four organizations. Section 5 explores the real-world NPV data we obtained from the telecommunication organization, Z. We compare this data to benchmark data from the literature. We also investigate the data for possible heterogeneity. After the exploration, we commence with the assessment of the forecasting quality of the NPV, benefits and cost in Section 6. Moreover, we perform a sensitivity analysis to investigate the impact on the quality of NPV forecasts when the forecasting quality of benefits or cost improve. Section 7 illustrates how the quantified forecast information is used to enhance decision information. This is illustrated by using two simulation examples. In Section 8 we discuss limitations of our research. Finally, Section 9 concludes the article.

2. Terminology

In this section, we introduce terms and notations that we use throughout the article. We recall some terms and notation that were elaborately discussed in another article [14]. Since we will make extensive use of these definitions and methods, for the sake of completeness, we summarize them here. Moreover, we introduce new terms that will often occur in this article. In Fig. 1, some of the terms and their relations are depicted. Although most terms are used frequently in common English, the meaning and understanding of them differs among people. Therefore, we clearly state how we comprehend them.

2.0.0.5. Asset. An asset is defined by the Oxford dictionary [55] as an item of property owned by a person or company, regarded as having value and available to meet debts, commitments, or legacies. In this article, we assume that projects need to be executed to create these assets. Even if a project modifies an existing asset, we will consider the altered asset as being a new asset. The assets can be both tangible and non-tangible.

2.0.0.6. Entity. In this article, we will use the word *entity* to denote any quantifiable aspect of an asset that is of interest. For instance, entities that we will consider are the project cost or benefits. Other examples are economic indicators, such as the Net Present Value or Internal Rate of Return.

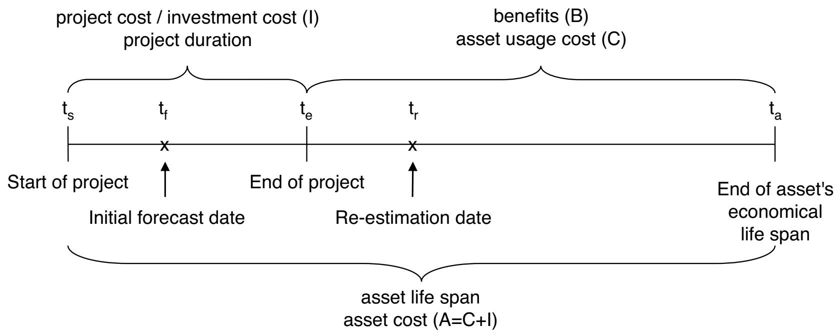

Fig. 1. Relevant terms in the life span of an IT asset.

2.0.0.7. Asset usage cost and project cost. We will use the term cost in different ways. The total asset cost, denoted by A is the full range of cost that is made for an asset. These cost consist of the project cost, denoted by I for investment cost, and the asset usage cost, denoted by C. We will consider project cost to constitute merely the cost for executing the project. The asset usage cost are all cost excluding the project cost. This entails, for instance, marketing, network usage and maintenance cost.

Note, that in this article we will only discuss cost that are computed in their present value. When we refer to A, I or C, we refer to the discounted asset cost, discounted project cost or discounted asset usage cost.

We make a distinction in cost as most organizations try to manage, control and contain project cost. This is done mainly due to the relative ease with which the project cost can be measured. In this article, we investigated to which extent control of this subset of the cost is useful in the context of the entire benefits and cost of an asset.

Note that for this article, it is not important how the asset cost are derived. We consider the asset usage cost as given. We will not question how the cost are computed, what precisely should be quantified and how it should be quantified. We assume that the estimators within an organization use equal definitions of how and what to incorporate in determining the cost.

2.0.0.8. Benefit. In the Oxford dictionary [55], a benefit is defined as an advantage or profit gained from something. In the context of this article, we consider a benefit, denoted by B, as a *quantified monetary* advantage or profit gained from something. Thus, we are only discussing benefits that have been quantified and represent a monetary gain.

Note, that in this article we will only discuss benefits that are computed in their present value. When we refer to B, we refer to the discounted benefits.

In this article, we will not dive into the question how the benefits have been quantified. Similar to the asset cost, we do not consider what precisely should be quantified and how it should be quantified. We assume that estimators within an organization apply the same definitions to compute the benefits.

2.0.0.9. Cash flow. In this article, a cash flow CF of a certain time period p is equal to the benefits B minus the asset cost A in that period. Discounting the cash flows of all time periods to the present time and summing them, leads to the Net Present Value, which will be discussed later on.

2.0.0.10. Duration and life span. As with cost, there are a number of durations we will consider. The different durations are displayed in Fig. 1. The first is the project duration, given by $t_e - t_s$. The second duration that is of interest is the economical life span of an asset, given by $t_a - t_s$. We define this life span of an asset as the period over which benefits and asset cost are forecasted.

2.0.0.11. Forecast. We define a forecast (or forecasting) of a certain entity as the prediction of the value that entity will have in the future. A forecast consists of the ex-post and the ex-ante part. The ex-post part is the part of the forecast that is already known—it is what has been done thus far. The ex-ante part is a prediction of what lies in the future.

Since we will address forecasts of different entities, we introduce a notation for forecasts. When we discuss forecasts, we will denote forecast f of entity e as f_e. If there is no ambiguity about the entity in question in a particular paragraph or it is of no relevance, we will simply use f.

2.0.0.12. Point forecast. The forecasts we discuss in this article are point forecasts. A point forecast is a single prediction that is often a summary of a large range of possible outcomes.

When a prediction is made, an estimator considers multiple scenarios that may occur and all relevant risks for the entity in question. For example, consider an estimator that needs to estimate the business value of an IT asset. The estimator should consider the risk that the project required to create the IT asset gets canceled. A study by Capers Jones found that of software applications in the 10.000 function point size range, about 36% are canceled and never completed [28]. The cancelation of a project significantly impacts the project outcome up to a swap to negative business value. This risk should be accounted for in the range of possible outcomes when forecasting the business value early on.

Table 1
Ten potential risks that can influence the outcome of software projects [29].

Potential risk
Risk of dilution of ownership due to multiple funding rounds
Risks of difficult data migration from legacy applications
Risk of significant layoffs of project team
Risk of inadequate warranties for quality and security
Risk of security flaws in application
Risk of late start in deploying risk solutions
Risk of estimates being rejected due to lack of benchmarks
Risk of software raising hardware warranty costs
Risks from disconnected "stove pipe" applications
Risk that requirements are not kept updated after release

Another example is the possibility of litigation when the asset is developed under a contract. In 2001, Capers Jones and his colleagues at SPR observed that 5% of projects within the United States that were outsourced, were probable to result in litigation or had litigation in progress [26]. Capers Jones states that an average lawsuit in the US costs both the plaintiff and the defendant so much money that all applications ending up in court have negative values. If an IT application is going to be developed under contract, a formal risk assessment is needed plus very strong contracts with penalties for non performance.

If one third of large applications are canceled, and 5% of outsourced projects may result in litigation, the CIO needs more certainty than exists today that applications receiving funds will be developed using best practices. This implies that an early risk analysis should be part of the funding equation.

A risk analysis considers the likelihood the risk will occur and its impact on the entity to be forecasted. Through personal communication from Capers Jones, we received a list containing 200 potential risks that can influence the outcome of software projects [29]. Table 1 describes 10 risks from that list.

These scenarios and their chance of occurrence lead to a range of possible outcomes. This range or interval of possibilities is the prediction of the value of interest and provides information on the risks related to the project. The interval allows the management to set adequate targets and make commitments based on their risk averseness or appetite.

However, in practice this interval is rarely given to the management. In many cases, the interval is summarized to a single point forecast, for instance, the most likely scenario to occur. As the management is confronted with these point forecasts, we assess their quality in this article. Next to that, we discuss ways to recreate the interval based on historic point forecasts in Section 7.

2.0.0.13. Actual. We define an actual of a certain entity to be the final realization of that entity. That is, the actual is the true value of that which has been forecasted. The notation that we will use for actuals is the same as with the forecasts, a_e with a an actual of entity e. Again, we also use the shorter version a when it is clear which entity is referred to.

We make two assumptions about the actual. First, we assume that IT governors want estimators to provide a prediction of the final realization. That is, the executives are interested in the true value. Second, we assume the actual is objectively measurable and thus that manipulation of the final realization is not possible.

2.0.0.14. Reference point. We define a reference point as the reference to which we compare the value of forecasts. An example of a reference point is the actual itself. In this article, we also use other reference points. For instance, we will use re-estimations of the benefits and the asset usage cost made a year after project completion as reference point. In Section 4, we will discuss the implications of using different reference points.

2.0.0.15. Bias. A bias is defined by the Oxford dictionary [55] as a systematic distortion of a statistical result due to a factor not allowed for in its derivation. Since we assume that the statistical result or forecast should predict the reference point, any systematic distortion, for instance, general overestimation, is considered a bias. There are several reasons why systematic distortions occur, both consciously and unconsciously. For instance, if an estimator wants to present an IT project proposal positively in order to get it approved, the estimator may underestimate the cost or duration of the project. Or, with a forecasting tool, parameter settings can unintentionally be inadequately set. In this article, we will not elaborate on the reasons for biases.

2.0.0.16. Forecast to actual ratio. A forecast to actual ratio or f/a ratio is a measure to assess the quality of forecasting. In the IT-context, this ratio was introduced by Barry Boehm [4]. It measures the forecasting quality by dividing the forecast by its actual. We will analyze f/a ratios of several entities. When necessary, we denote f_e/a_e also as $(f/a)_e$ for forecast f and actual a of entity e.

2.0.0.17. Forecast to actual plot. To assess the quality of forecasts, we plot the f/a ratios in what is known as the forecast to actual plot, or f/a plot. The f/a plot depicts f/a ratios against the relative time at which the forecasts are made. The f/a ratios that are depicted should be a homogeneous set of forecasts.

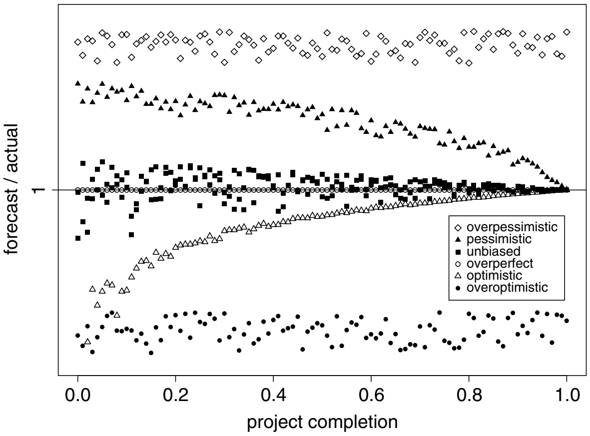

Fig. 2. Typical patterns in an f/a plot.

We recall Fig. 2 [14] that depicts several typical f/a patterns. The horizontal axis of the figure depicts the percentage of project completion at which the forecast is made. The vertical axis shows the f/a ratios on a logarithmic scale.

In the figure, a number of patterns are illustrated that may be found when plotting f/a ratios. If forecasts mostly overestimate the actual, the f/a plot will reveal the optimistic or over-optimistic pattern, or variations thereof. In case forecasts are aimed at predicting the actual value, one will find the f/a ratios equally above and below the value 1, indicating the unbiased pattern. If one finds many forecasts extremely close to 1, the data may be manipulated or fixed price agreements are present, leading to the overperfect pattern. When forecasts are mostly underestimated, the f/a plot will resemble the pessimistic or overpessimistic pattern. These patterns illustrate the usefulness of the f/a plot as it allows to obtain an impression of potential biases in the forecasts made.

Note that the naming conventions of the different patterns are ambiguous. For instance, for entities as project cost or project durations, a forecast larger than the actual is a pessimistic forecast. That is, the forecast is a pessimistic projection of what really happened. However, in case of entities such as benefits, a forecast larger than the actual is an optimistic forecast. Although the labels are ambiguous, the patterns remain the same.

2.0.0.18. Estimating Quality Factor. The f/a plot provides a means to distinguish potential biases in the forecasts made. Yet, it does not allow for quantifying and benchmarking their quality with others. To this end, we use the Estimating Quality Factor, or EQF, developed by Tom DeMarco [11]. The EQF is a measure of the deviation between the forecast and actual. It is computed with the following equation.

$$\text{EQF} = \frac{\text{Area under actual value}}{\text{Area between forecast and actual value}}$$

$$= \frac{\int_{t_s}^{t_a} a\, dt}{\int_{t_s}^{t_a} |a - e(t)|\, dt} \tag{1}$$

$$= \frac{\int_{t_s}^{t_a} 1\, dt}{\int_{t_s}^{t_a} |1 - e(t)/a|\, dt}. \tag{2}$$

In this formula, a is the actual value, t_s the start date of the asset, t_a the end date of the asset and $e(t)$ the value of the forecast at time t ($t_s \le t \le t_a$). In Section 4, we will generalize the EQF for other reference points than the actual and discuss the impact of this generalization.

An assumption is that $e(t)$ is known for the range $[t_s, t_a]$. That is, we know at all times what the value of the most recent forecast is. However, in some circumstances the initial forecast is not made at the beginning. In this case, we assume the first forecast made at time v is actually made at the start. Mathematically, this means we assume that $e(t)$ on range $[t_s, v)$ equals $e(v)$.

In statistics, other measures are known that quantify the quality of forecasts, such as the MSE, MAPE and MRE [10,16,24]. It is possible to use these measures instead of the EQF. However, the benefit of the EQF is that it is defined as a time-weighted average deviation to the actual. In our analyses, we assess the forecasting quality of assets that can have multiple forecasts. For decision making, it is important for these forecasts to be as quickly as accurately as possible. Therefore, it is important to account for the timing of subsequent forecasts. The EQF is defined to incorporate this effect.

2.0.0.19. Reference cone. The reference cone is a tool that compares f/a ratios against a benchmark forecasting quality. Given certain assumptions it is possible to compute lines that represent this forecasting quality. Consider the following assumptions.

– Ex-post inclusion: We assume that each consecutive forecast incorporates the ex-post part and we assume this part is known with certainty.
– Ex-post growth: The growth of the ex-post part is assumed to be described by a constant function.
– Ex-ante accuracy: The accuracy of the ex-ante part is assumed to remain constant as the project progresses.
– Goal: The goal of the forecast is to predict without bias and quickly as accurately as possible the actual value of interest for the project.

Recall that the ex-post part is the part of the forecast that is what has been done thus far. The ex-ante part is the prediction of what still lies in the future. The above assumptions are captured, assuming infinitely many forecasts are made, in the following formulas taken from an article by Eveleens et al. [14].

$$l(x) = x + \left(1 - \frac{2}{\mathrm{EQF}_l}\right) \cdot (1 - x) \tag{3}$$

$$u(x) = x + \left(1 + \frac{2}{\mathrm{EQF}_u}\right) \cdot (1 - x). \tag{4}$$

In this formula, x is the project's progression relative to its project duration with $x = (t_f - t_s)/(t_a - t_s) \in [0, 1]$. The lower reference line $l(x)$ is defined by the value EQF_l that represents the EQF quality of the lower line. The upper reference line $u(x)$ is given by the value EQF_u, which determines the EQF quality of the upper line. These reference lines are defined under the assumption that $\mathrm{EQF}_l \geq 2$ and $\mathrm{EQF}_u > 0$.

When we draw reference lines with these formulas, we use the notation $c(l, u)$ for reference cone c with lower bound of EQF quality EQF_l and a quality of EQF_u for the upper bound. If $\mathrm{EQF}_l = \mathrm{EQF}_u$ we also use the shorter notation $c(l)$.

Using these terms and notations, we commence with our investigation of the forecasting quality of IT asset business value.

3. Asset business value

Pisello et al. [48] state that the Chief Information Officer (CIO) must become the Chief Financial Officer (CFO) of IT. To do so, the CIO must manage IT projects in such a way that their business value is maximized. But how to determine the business value of IT assets?

To specify the business value of IT assets, in their decision making executives are provided with information about the advantages, disadvantages and risks of each project proposal. Project proposals contain two kinds of information: qualitative and quantitative. Qualitative arguments are, for instance, corporate social responsibility or strategic alignment. According to multiple surveys [2,17,47], many organizations consider these arguments as important criteria for project selection.

The contribution of this article is to the quantitative kind of information. This part of the available information consists, among others, of forecasts of quantified benefits and cost. Often, these predictions are summarized using economic indicators. Well-known examples of such indicators are the Net Present Value, the Internal Rate of Return, the Return on Investment, and the Payback Period. We describe these indicators briefly below.

– Net Present Value (NPV). The Net Present Value is a summation of the predicted monetary benefits and cost of a project discounted to current value. If the NPV is positive this indicates the project is estimated to provide for monetary gain given the discount rate. If the NPV is zero, the project is a neutral investment: it generates enough discounted benefits to cover the discounted cost. If the NPV is negative the project proposal is expected to result in a monetary loss given the rate used. In the next subsection we will discuss the NPV in more detail.
– Internal Rate of Return (IRR). The Internal Rate of Return is the discount rate for which the NPV is equal to 0. The IRR is compared to the rate we would receive for similar investments, which is also known as the opportunity cost of holding capital. If the IRR is lower than the opportunity cost of capital, the investment should not be funded from a quantitative point of view. If the IRR is high compared to other rates, the project will generate more yield than other similar projects in the market.
– Return on Investment (ROI). In a book by Bierman [21], the Return on Investment is defined as an average income after depreciation divided by its investment. The ratio is also known as rate of return and there are many different ways of computing it. In all cases the higher the ratio the more profitable the project.
– Payback Period (PBP). The Payback Period is the amount of time needed after project completion to generate an equal amount of cumulated cash flows to cover the initial investment. Often these calculations are not discounted to today's currency. The shorter the Payback Period, the sooner the initial investment is paid back.

These economic indicators give IT executives an indication of the business value of the project proposals. It is possible to use them to make predictions and to evaluate the final realization. Therefore, we are able to use these indicators to answer our question: how do we know that the quantitative business value forecasts are accurate, unbiased and reliable enough to support the decision making process?

We will answer this question by assessing the forecasting quality of one of the indicators, the NPV. We will only use the NPV, simply because our case study of organization Z uses this indicator to support its decision making. However, the generalized method we will propose later on, is applicable to the other indicators as well.

But, before answering the question, we need to better understand the NPV. First, we discuss the limitations of the different indicators with respect to each other. Each indicator has its theoretical and/or practical disadvantages. Then, we will identify the different components of the NPV and discuss each of them.

3.1. Indicator limitations

We briefly discussed the NPV, IRR, ROI and PBP. Brealey et al. [5] state that theoretically the use of the NPV leads to better investment decisions than these other well-known indicators. They argue that in some situations the IRR leads to different results than the NPV. We note that although both the IRR and NPV are derived from the same formula, the way they are derived can causes different outcomes. One situation in which the IRR is ineffective, is when there is a mixture of multiple positive and negative cash flows. For instance, consider an asset that, besides the initial investment, has some final fee to be paid at the end of its life span. In that case it may occur that there are two realistic IRR ratios for which the NPV is 0. This is also observed by, among others, Lorie et al. [40]. Moreover, Brealey et al. discuss that the IRR does not discern between borrowing or lending money, and has problems when the opportunity cost of capital is different over several years.

Brealey et al. also discuss that the Return on Investment leads to worse decisions, since it does not account for the timing of the cash flows. Since the cash flows are generally averaged, the indicator places no importance on whether the cash flows are earned in the first year or the last year. Yet, in reality this is a crucial aspect of the investment decision.

Finally, Brealey et al. argue that the Payback Period ignores all cash flows generated after the initial investment is paid back. However, these cash flows can make an investment highly lucrative or not. Therefore, Brealey et al. suggest to use the NPV to justify investment decisions.

However, the NPV is also not without limitations. A practical disadvantage of the NPV is that it does not consider the scarceness of the available resources. For instance, an asset with a predicted NPV value of 100 Euro is considered a better investment than one with a predicted 80 Euro, even though the former asset may involve investment cost of 10 million Euro and the latter 0.1 million Euro. Note that this aspect is accounted for by the IRR as well as in the ROI indicator by dividing by the investment cost.

Another disadvantage of the NPV is that the determination of the discount rate, which is needed in the calculations, is difficult. Later on in this section, we discuss this discount rate in more detail. It is possible to derive organization-specific discount rates. However, determining the required project-specific discount rate is not trivial.

Moreover, one should consider that a discount rate that is applicable now, may not be applicable next month as is pointed out by Ross [51]. A project proposal with a negative NPV given the discount rate this month, can be a very interesting opportunity some time later. Ingersoll et al. [25] developed a method to account for the value of optionality with respect to the uncertainty of interest rates. An investment should only be made when the NPV is sufficiently positive to forego the option to delay the investment.

3.1.0.20. Indicators in practice. A few decades ago, numerous surveys [17,19,45,47,53] found that the NPV was not the method of choice of most Chief Financial Officers (CFO). In 2002, an article by Ryan et al. [52] shows that till 1996 studies generally indicated the method most used by organizations was the IRR. Both Ryan et al. [52] and Arnold et al. [2] find that only just after the year 2000, organizations have adopted the NPV as preferable indicator. The survey of Ryan et al. [52] found that 85% of the respondents of the survey indicated to use the NPV often.

Still, all these surveys indicate that the organizations frequently use multiple indicators to support decision making. Although theory suggests the NPV should suffice to make investment decisions, in practice a combination of the NPV and other common indicators are used. Ryan et al. [52] also found that there is a correlation between the capital budget and the use of NPV and IRR. The larger the budget the more likely the use of either one of these methods. The Payback Period is found to be more frequently used by organizations with smaller capital budgets. These surveys show that many economic indicators are used for decision making.

3.2. Net Present Value

In this section we elaborate on the Net Present Value. First, we show how to compute the NPV. Then, we discuss its components, assumptions and interpretation.

Informally stated, the NPV determines the monetary value an asset adds to an organization. The cornerstones of this economic indicator are the predicted benefits, cost and economic life span. Simply put, the NPV determines whether the benefits outweigh the cost, both of which are computed in todays worth. If the NPV is positive, it means the asset will generate value for the organization. If it is negative, creating the asset will result in an overall loss.

Formally, the NPV is described by Brealey et al. [5] as follows. Denote CF_p as the cash flow predicted for time period p and r_p the discount rate of time period p. Let N be the total amount of time units that are used. Then, the NPV is calculated in the following way:

$$\text{NPV} = \sum_{p=1}^{N} \frac{CF_p}{(1+r_p)^{p-1}}.$$

Below we discuss the elements of the formula in more detail.

3.2.1. Discount rate

The discount rate r_p is also known as the opportunity cost of capital. Often, the discount rate is chosen identical for each time unit. That is, $r = r_p, \forall p$.

The purpose of the discount rate is two-fold. First, it accounts for the time value of money. It is better to have 100 Euro today than it is to have 100 Euro tomorrow, since the former can be invested immediately to generate additional income. By discounting the future cash flows, we acknowledge this time value of money.

Second, any investment must be funded with capital. The providers of this capital require compensation for making their capital available. This is the cost of the capital that the organization intends to use. Any investment should aim to have a higher return than the cost of capital. If not, the organization would waste the capital of the investors. They would have done better to return the money to the investors and let them invest it otherwise.

But how to determine this cost of capital or discount rate? A number of books and articles [1,5,21,50,53] explain methods, among others the weighted average cost of capital or WACC, to find the discount rate. A survey [7] found that the WACC is used most often in practice. The WACC is organization-specific.

However, the discount rate used in the calculations of the NPV is investment-specific. Not all assets of an organization will have the same risk as the entire organization. Some assets will have higher risk and other assets will have lower risk. The organization-specific WACC should be changed to account for the particular asset risk. For instance, Dewan et al. [12] and Verhoef [56] suggest to increase the WACC in case of IT assets.

Determining this correction to the WACC is not a trivial task. Moreover, it is often difficult to establish which other assets are equivalently risky. It is therefore not surprising that the survey by Bruner et al. [7] found that many organizations do not adjust the WACC for individual investments. Petty et al. [45] contended that the use of sophisticated risk-adjustment techniques would be limited until risk can be measured more precisely and one can show the impact of additional risk upon the firm's cost of capital.

3.2.1.1. Discount rate forecast. The discount rate that is used to compute a particular NPV, is based on an assessment of the risk of that investment. In most cases the risk involved is not objectively measurable and is thus only a prediction of the actual risk.

The accuracy of this forecasted discount rate directly impacts the forecasting accuracy of the NPV. However, in this article, we will not assess the forecasting quality of the discount rate. We consider the cost of capital as given and will not question its derivation. In our case study of organization Z, all calculated NPVs of a particular asset are based on the same discount rate.

3.2.2. Cash flow

Another crucial element of the NPV calculations are the forecasts of the cash flows. The cash flows CF_p in the equation are the expected cash flows for each time period p. These predictions should account for the likelihood and impact of the risk of different scenarios on the benefits and cost, such as cost overrun, project failure and/or late delivery. These scenarios lead to a probability distribution of possible cash flows.

Surveys of Gitman et al. [19] and Fremgen [17] found that estimating the cash flows is considered the most critical stage of the capital budgeting process. In 1978, Schall et al. [53] surveyed that individual project risk is assessed by means of a probability distribution of cash flows by 25% of the respondents and another 10% using sensitivity analysis. In most cases, the risks were assessed implicitly. That is, the distribution is not made explicit, but is implicitly incorporated in the predictions of the cash flows by the estimator. In 2000, Arnold et al. [2] found that 94% of the organizations required a formal risk evaluation. This was done in 85% of the cases using sensitivity analysis, often in conjunction with a subjective assessment. A probability analysis was performed by 31% of the organizations.

In stark contrast are the findings in the information technology sector. In 2002, the Meta Group [20] surveyed that 84% of the organizations do not use business cases at all for their IT-projects or only for selected projects. The Kellogg School of Management [36] found that 68% do not track benefits. The numbers show that organizations have difficulty determining the benefits of IT assets, let alone to formally account for the risks in the forecasts.

3.2.2.1. Unbiased. A critical assumption of the cash flow predictions is that they are unbiased. However, forecasts of, for instance, duration and cost can be biased [3,4,14,46]. It is conceivable that the same applies to forecasts of cash flows. The decision to invest in a project is highly dependent on these forecasts. Those making the proposal and those with interests in executing it, may be inclined to overestimate the cash flows to make the proposal more appealing. It is therefore crucial to check whether the assumption of unbiased cash flow forecasts holds.

To investigate this assumption, we have to consider what the cash flow is composed of. A cash flow is the resultant of the projected benefits minus the forecasted cost. If we find the cash flows to be unbiased, this may imply both benefits and cost are unbiased. But it could also mean that they are both highly biased, but counter each other's effect. Although the effect is overall the same, the latter situation is unwanted. In that case, it is mere luck and not good forecasting practice. Luck may change any instant, a good forecasting practice not. Therefore, to assess whether cash flows are unbiased, we should not only analyze the cash flows, but consider the components that it consists of.

To better illustrate the different components of the cash flow, we reformulate the above equation of the NPV. We define b_p to be the predicted benefits of period p, c_p the predicted asset usage cost and i_p the project cost, with $CF_p = b_p - c_p - i_p$. Let N be the total number of periods in which the economical asset life span is divided and r_p the discount rate of time period p. Then, we write the NPV using the following equation:

$$\begin{aligned} \text{NPV} &= \sum_{p=1}^{N} \frac{b_p - c_p - i_p}{(1+r_p)^{p-1}} \\ &= \sum_{p=1}^{N} \frac{b_p}{(1+r_p)^{p-1}} - \sum_{p=1}^{N} \frac{c_p}{(1+r_p)^{p-1}} - \sum_{p=1}^{N} \frac{i_p}{(1+r_p)^{p-1}} \\ &= B - C - I. \end{aligned} \tag{5}$$

In Formula (5), B amounts to the cumulated discounted benefits, C the cumulated discounted asset usage cost and I the cumulated discounted project cost. Recall that in Fig. 1 in Section 2, we illustrated that the benefits and asset usage cost are not present during project execution. That is, the summation of the benefits and asset usage cost usually only have values over the interval $p \in y, y+1, \ldots, N$, where y is the period in which the end of the project, t_e, falls. The summation of project cost usually has values over the interval $p \in 1, 2, \ldots, y$.

To assess whether the cash flows and NPV are unbiased, we have to investigate each of these components. We make a distinction between project cost and asset usage cost, since we wish to know whether the quality of their forecasts are different. Many organizations record and have insight in the project cost. However, these cost often amount to only a relative small portion of the entire asset cost. We want to see whether accurate forecasts of the project cost pays off, or that it is wise to put more effort in the correct prediction of the asset usage cost.

Note that it remains helpful to analyze the NPV directly. Assessing the quality of the forecasted NPVs shows the impact of the interactions of the individual elements. If we find both the benefits and the cost to be overestimated, we do not know their combined effect on the accuracy of the NPV. Therefore, a combination of the individual analyses and the overall quality of the NPV is most insightful. In this situation, the interdependences are contained in the NPV analysis and the individual analyses provide answers as to where the variance comes from.

In the case study of organization Z, we will assess the forecasting quality of the NPV. Above, we discussed that it is useful to increase the depth of the analysis by also investigating the components of the NPV, that is, B, C and I. It is possible to further increase the depth of the analysis by also considering the components of B, C and I. For instance, B is the sum of b_p over all time periods p. Therefore, it is possible to investigate the forecasting quality of each b_p separately.

The level of detail that is required depends on the goal of the analysis. In this article, the primary goal is the forecasting quality of the NPV. Moreover, we wish to determine whether control of the project cost is useful without sufficient control over the other elements. For these purposes, we will only analyze the NPV, B, C, and I.

3.2.3. Time

Besides the discount rate and the predicted cash flows, time is another variable in the equation. Time is captured in the predicted total number of time periods denoted by N. This total amount is a forecast of the asset's economical life span. The time periods are commonly expressed in years, but can also be different, for instance, in months.

An accurate prediction of this variable is important for the resulting NPV. If the life span is too long, we may unjustly attribute additional benefits and cost to the asset. On the other hand, if the predicted life span is too short, we will ignore future cash flows of the asset in our calculations. For instance, suppose the initial forecast predicted the economic life span to be 5 years. It may occur that when the NPV is re-estimated, it is estimated the economic life span is 5.5 years. When in this half year a positive cash flow is generated, the initial NPV forecast will be underestimated. These cash flows can make the difference between a positive or negative NPV.

However, in this article, we will not assess the forecasting quality of the asset's economic life span. We consider the life span, N, as given and will not question its derivation. In our case study of organization Z, the two estimated NPVs of a particular asset used the same economic life span. Moreover, all estimated NPVs of a particular asset are discounted to the same present moment.

3.2.4. Indirect influences

Besides the described elements in the NPV equation, there are other factors that influence the outcome of the NPV indirectly.

3.2.4.1. Indirect time effect. Apart from the direct impact, time also influences the NPV calculation in indirect ways.

For example, consider an asset that is delivered three months later than expected. In the first three months no benefits or asset usage cost occur, making their forecasts overestimations. Moreover, all predicted cash flows will become different. Namely, because the cash flows occur later, they will be discounted differently. Also, the timing of the forecasts can be crucial, for example, due to growing competition on the market. Therefore, delays can severely influence the NPV, benefits

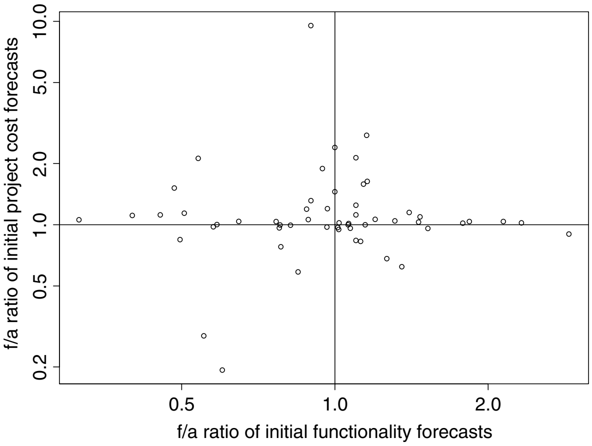

Fig. 3. No evidence of correlation between forecasting accuracy of initial project cost and functionality forecasts.

and cost, making their accuracy forecasting difficult. These inaccuracies in the forecasts are also discussed in an article by Peters et al. [44].

Moreover, in an article by Putnam [49] and a book by Boehm [4, p. 472] it was found that shortening the project duration beyond its optimum can significantly increase the project cost. Boehm observed a similar effect when the project duration was stretched beyond its optimum.

Therefore, the accuracy of the prediction of the project duration is reflected in the accuracy of the forecasts of benefits, asset usage cost and project cost. Any inaccuracy we find in the analysis of the other variables may partly be caused by the inaccurate forecast of the project duration.

3.2.4.2. Functionality. Although functionality is not mentioned in the formula of the NPV, it has an indirect impact on the variables in the equation. For instance, if a project is executed and delivers less functionality than anticipated, this may prevent certain forecasted benefits to materialize (solution underdelivery). Similarly, if the resulting IT program has more functionality, it is possible additional benefits are generated as a result. An increase in requirements is also known as requirements creep [27,8]. On the other hand, an increase in functionality may also result in higher project and asset usage cost.

These effects should be weighted in the overall forecasts of the benefits, asset usage cost and project cost. This was done, for instance, in an article by Peters et al. [44] by considering requirement creep scenarios. Due to these effects, it is possible the forecasts of functionality are correlated with the forecasts of the other components. For instance, if the functionality is underestimated, the benefits and cost may be underestimated as well.

To investigate a possible correlation between the accuracy of the project cost forecasts and the functionality forecasts, we analyzed data from a large financial organization Y. In Fig. 3, we depicted the initial functionality, denoted by F, $(f/a)_F$ ratios against the initial project cost $(f/a)_I$ ratios of 55 projects of which we had all the relevant data. The functionality forecasts were measured using function point countings [13,18].

The correlation coefficient of the two data sets is -0.05. Since values of -1 or 1 represent perfect correlation and 0 depicts no correlation, the coefficient shows that there is no correlation. This result reveals that the relation between functionality and project cost may only be marginal with respect to their forecasting quality. That is, underestimating the functionality does not directly cause an underestimation of the project cost.

Note that Boehm's cone of uncertainty [4] showed that the time at which the forecasts are made is relevant for their quality. Although the forecasts in our real-world data set of project cost and functionality are not made at the same time, the differences between the moments they are made are small. The median of the difference between the moment the project cost forecast was made and the functionality forecast was made, divided by their project duration is 0 and the average difference is 0.08. Therefore, there is no indication that the results of this analysis are influenced by the different timing of the forecasts.

3.2.4.3. Software quality. Software quality is highly relevant to value prediction of both the benefits and the cost. The benefits derived from an asset are based on the perceived value of the customer. If the software contains many defects, the customer may value the asset less, thereby generating less benefits.

Capers Jones [28] found that the US average for IT software quality is about 5.0 defects per function point combined with 85% defect removal efficiency. This results in delivery of about 0.75 bugs or defects per function point. Best in class IT software quality combines less than 3.0 defects per function point combined with more than 95% defect removal efficiency. This results in delivery of about 0.15 defects per function point.

Jones argues that increasing the defect removal efficiency from 85% to 95% saves money and shortens development schedules. This is supported by others [8]. Applications using state of the art software quality methods have development cost about 15% lower than average projects, and maintenance cost about 55% lower than average projects. Total cost of ownership is about 40% lower than average projects [29]. Therefore, it is beneficial to produce assets with a relative high quality.

This illustrates that the forecasts of benefits and cost are influenced by the delivered software quality of the asset. However, in this article, we will not consider the software quality of the assets. In this article, we focus on assessing the resulting forecasting quality of the NPV, benefits and cost.

3.3. Summary

To support decision making, IT executives have access to numerous economic indicators that summarize the expected business value of project proposals. Despite its limitations, theory suggests that the NPV method is superior compared to other methods. However, our discussion of the indicator shows that the NPV is far from certain. Each of its components, needs to be predicted. Therefore, it is not surprising that in practice many organizations use multiple economic indicators to gain insight in the value of an asset proposal.

Like the NPV, any economic indicator is highly dependent on the accurate forecasting of its elements. There are numerous components that influence the final NPV either directly, such as the benefits or cost, or indirectly, such as project duration or functionality. Therefore, in this article, we assess the forecasting quality of the NPV, and its components: the benefits, asset usage cost and project cost.

In the next section, we discuss a method with which it is possible to assess the forecasting quality of the NPV and its components. Those not interested in the mathematical elaborations, may skip the next section and continue with Section 5.

4. Generalized method

In this article, we wish to determine whether organizations are able to adequately assess the business value of project proposals. In the previous section, we discussed the NPV and other economic indicators, which represent the business value of assets. To assess whether organizations are capable to accurately forecast the business value of IT proposals, we need to investigate the forecasting quality of the NPV and its components. Recall that the Net Present Value is a summation of the predicted monetary benefits and cost of a project discounted to current value.

A method to assess forecasting quality is described in an article by Eveleens et al. [14]. As discussed for the sake of completeness in Section 2, that method uses the f/a plot, the reference cone and the EQF. Recall that the EQF is a measure of the deviation between forecast and actual. The method is applicable for entities that are positively valued, that is, $f, a > 0$. We wish to evaluate the quality of the forecasts of the NPV, benefits and cost in the same manner. However, if we want to do so, a number of problems arise.

4.0.0.4. Asymptotic behavior. The first problem is a visual complication in case of asymptotic behavior. With asymptotic behavior we refer to situations in which forecast and/or actual are zero. For project cost, usually the forecasts and actuals are greater than zero. In practice, project proposals that cost nothing or assets that cost nothing hardly ever occur. However, in the context of economic indicators, a forecast of zero is important. For instance, a NPV of zero is the turning point between an asset being yield or loss generating. Also, it is not unlikely to find assets that have no benefits. For example, consider an asset that was developed, yet on completion it turned out there was no longer a market for it.

Zero actual benefits cause visual problems for the f/a plot. Let us explain. The f/a plot visualizes potential biases by plotting f/a ratios on a logarithmic scale. If the actual is zero, the ratio becomes infinite, making the logarithm also infinite. Therefore, in the f/a plot this point cannot be visualized in a normal way.

This problem also arises if the forecast is zero. In this case, the f/a ratio is zero. However, the logarithm of zero is not defined. Thus, in that case the ratio can also not be depicted in the plot.

Normally, not many of such zero forecasts will be made. In most cases, a forecast of zero indicates no forecast is made at all. If this is the case, it is best to remove these forecasts from the analysis all together, as they will not reveal information on the quality of the forecasting process. However, sometimes an entity is truly forecasted to be zero, or remains interesting for analysis in conjunction with other forecasts. For instance, consider the case of a forecast of benefits and no forecast made for the cost of the project. In that case, it may be interesting to incorporate the data point in the analysis, to assess the quality of the resulting NPV forecast.

Why is it a problem that the f/a plot does not visualize ratios with a forecast or actual of zero? If there are many such ratios in a data set, the f/a plot may point to a potential bias in the data that does not exist. For example, consider an extreme situation in which 51% of the data in a data set consists of zero actuals and the remainder of the forecasts are underestimations, that is $f/a < 1$. The f/a plot would only depict the latter half of the data, which will show a bias toward underestimation. In reality, there is no particular bias given the data, since 51% of the data, the zero actuals, are overestimations. Clearly, the ability of the f/a plot to detect biases is hampered by zero forecasts and/or actuals.

4.0.0.5. Reference point. A second technical problem is caused by the reference point with which we compare the quality of the forecasts. With the f/a ratio this reference point is the actual. However, it is questionable whether the actual is useful to support decision making. For instance, suppose one finds that 2-year old project cost forecasts or 5-year-old forecasted benefits were generally overestimated. Although this sheds light on the forecasting quality of 2 or 5 years ago, in most cases

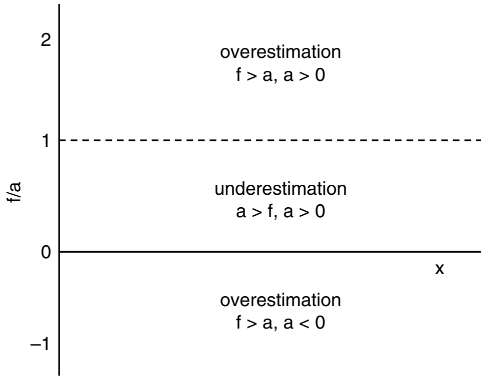

Fig. 4. Illustration of an f/a plot with a linear vertical axis, positive forecasts and allowing for positive and negative actuals.

this is hardly information that is useful to apply to today's forecasts. By the time the actuals are known, the forecasting practice may already have been changed. To support decision making, we prefer more recent information about the current forecasting practice. Moreover, in case of, for instance, the NPV, the actual may never be objectively measured.

A solution is to re-estimate, for instance, the benefits before the end of the economic life span. For instance, it is possible to re-estimate the benefits a year after project completion. At that time, it is more clear to the estimator which of the many possible scenarios is unfolding. The ex-post part is considerably larger than in the previous forecasts and the ex-ante part becomes smaller and smaller. This way we are able to approximate the actual, which allows us to derive more recent forecast information that we are able to use for today's project proposals.

However, in this case the value of the re-estimation is no longer objectively measurable. In fact, it is a forecast in itself. If we compare the forecasting quality of earlier forecasts with this approximation of the actual, we use a different reference point than the actual. We will investigate how the assessment of forecasting quality is affected by such alternative reference points.

4.0.0.6. Negative values. Finally, the NPV can be both positive and negative. In practice, most negative NPVs will occur in re-estimations. For example, consider an asset with a positive forecasted NPV that grossly overestimated the benefits. Afterward, the asset was recomputed and the cost turned out to be greater than its benefits, resulting in a negative NPV.

It is also possible to have negative forecasted NPVs. For example, mandatory assets or assets with a negative NPV that are interesting for their qualitative features. Note that mandatory assets can have a positive value as well. If the asset would not be performed, the organization risks a fine or other sanctions, which potentially make such assets beneficial to undertake. Even in case of negatively forecasted NPVs, their forecasting quality remains interesting to investigate to contain the predicted losses.

The negative values cause two problems for the f/a plot. The first problem is that the f/a plot uses a logarithmic axis to depict f/a ratios. However, the logarithm is not defined for negative values. Therefore, we are unable to depict them. We could abandon the logarithmic scale and use a linear axis, allowing us to depict negative ratios as well. However, a linear axis does not allow for easy distinction of biases.

Let us explain. Assume we have positive forecasts and both positive and negative actuals, that is $f > 0$ and $a \in \mathbb{R}$. Suppose we would use a linear axis for the f/a plot and allow for negative actuals. In this case, all the negative ratios would be depicted below the line $f/a = 0$. This situation is illustrated by Fig. 4. These negative ratios indicate that, since forecast f is positive, the forecast is larger than the actual and is thus an overestimation. For all f/a ratios in the interval [0, 1], the f/a ratio indicates that the forecasts are smaller than their actuals, or equivalently, they are underestimations. Finally, the f/a ratios above the line $f/a = 1$ depict forecasts that are larger than the actuals, which are again overestimations. Thus, when we use a linear axis and assume non-negative forecasts, and both positive and negative actuals, the f/a plot consists of three sections of which the middle one shows underestimations and the remainder overestimations.

A similar explanation applies when we assume non-negative actuals, and both positive and negative forecasts. Thus, such figures make it difficult to adequately distinguish biases, not to mention obtaining a visual impression of the quality of the forecasts.

The second problem of the negative values is that the f/a ratio becomes ambiguous. For instance, consider an f/a ratio of -1. Such a ratio is possible when $f < 0 < a$, but also when $a < 0 < f$. Thus, a ratio of -1 can both mean the forecast is larger than the actual as well as the actual is larger than the forecast.

A similar problem arises with positive f/a ratios. Consider an f/a ratio of 2. This occurs when $f > a$ with $f, a > 0$ and when $f < a$ with $f, a < 0$. Thus, it is no longer possible to determine whether the forecast is smaller or larger than the actual value. To cope with these problems, we will propose other ratios than Boehm's f/a ratio that deal satisfactorily with negatively valued entities.

In the remainder of this section, we address each of these issues separately. This results in a generally applicable method to assess forecasting quality of entities that range over the entire real numbers.

4.1. Asymptotic behavior

Above we stated that in case of asymptotic behavior, that is, forecast or actual is zero, a problem of visualization arises. There are three possible scenarios. First, suppose we have a positive valued forecast and an actual of zero. In this case the f/a ratio is infinite. In an f/a plot such points cannot be depicted in a normal way.

Second, consider a positive valued actual and a forecast of zero. Now, the f/a ratio is zero. Since the f/a plot uses a logarithmic axis and the logarithm of zero is undefined, this ratio cannot be drawn.

Finally, it is possible to have a forecast and actual of zero. In this situation, the f/a ratio of 0/0 itself is undefined. Again, we are unable to visualize this ratio in the f/a plot.

If many zero forecasts and/or actuals are present in a data set, this hampers the ability of the f/a plot to visualize biases. Therefore, we define the desired behavior of an f/a ratio in the described asymptotic cases.

4.1.0.7. Zero actual. If the actual is zero, we wish to depict a ratio in the f/a plot that is infinite. A possible solution is to draw these points an order greater than the maximum value of the other ratios in the data set that are smaller than infinity. For instance, suppose the maximum f/a ratio smaller than infinity is 20. Then, we plot the infinite f/a ratios as f/a ratios one order larger, that is, 200. To allow for a better distinction of the infinite f/a ratios, we visualize them differently from the other ratios. Suppose the normal ratios are shown as dots, then we plot the infinite ratios, for instance, as squares. In this way, we are able to visualize these ratios, resulting in an f/a plot to adequately depict potential biases.

4.1.0.8. Zero forecast. Suppose we have a forecast of zero and positive actual. In this situation, the f/a ratio is zero. The f/a plot uses a logarithmic scale for the f/a ratios. However, the logarithm of zero is not defined. Therefore, the logarithm prevents us from drawing f/a ratios of zero.

We could draw these points without taking the logarithm. However, this is unwanted as then the points are no longer scaled, which is very useful in recognizing potential biases and evaluating the accuracy of the forecasts. Namely, this scaling creates equal spacing for under- and overestimation. That is, in the figure the distance between 1/2 and 1 is equal to that of 2 and 1. This property is useful to assess the accuracy of both under- and overestimation.

Therefore, we propose to visualize these points similarly to f/a ratios with a zero actual. When the forecast is zero and the actual is positive, this indicates that the forecast was smaller than the actual. Thus, we draw these points one order smaller than the minimum value of the f/a ratios in the data set that are larger than zero. For instance, suppose the minimum value of f/a ratios larger than zero is 1/20. Then, we plot all f/a ratios that are zero at an f/a ratio of 1/200. Again, to allow for better distinction, we depict these points differently from the other f/a ratios.

4.1.0.9. Zero forecast and actual. Finally, the situation could occur that both forecast and actual are zero. The f/a ratio of 0/0 is undefined. In this scenario, the forecast that was made was extremely accurate. It precisely predicted the final outcome. Therefore, a solution is to define such an f/a ratio as 1, that is, $0/0 \equiv 1$. For better distinction, these points should be visualized differently from the other ratios.

Besides these solutions, we point out that there is a statistic that, in conjunction with the f/a plot, is useful to assess biases. This is the median f/a ratio of the data set. The median f/a ratio is not influenced if either a forecast or an actual is zero. Only if they are both zero, there is a potential influence, since the f/a ratio of 0/0 is not defined and is therefore not used to compute the median.

Nonetheless, it is useful, besides the f/a plot, to consider the median f/a ratio to confirm one's findings. We will visualize this median in the f/a plot by adding a horizontal dashed line at the value of the median.

4.2. Reference point

All f/a ratios use the actual to compare the forecasting quality with. However, when we want to use the forecasting quality to enhance decision information, the actual is not necessarily the most useful reference point. For instance, in our case with the NPVs, we know their value with certainty as soon as the economic life span is over, and then it is too late to enhance decision information.

Then, we need an alternative: a reference point that is not the actual, but most probably a better prediction of the actual than the initial forecast. For instance, it is possible to approximate the actual two or three years after the exploitation of the asset.

In our case study, the telecommunication organization, Z, is accustomed to approximate actuals by re-estimating them after, say, a year upon project completion. This is in itself a new forecast, usually of better quality than of other previous forecasts. Namely, at the time of the re-estimation, there is more information available. For the estimator it becomes more clear which of the many possible scenarios is unfolding. The project is finished and operational, initial benefits are received and initial asset usage cost are made. This should give a more reliable indication of the final realization than at the start of the project. Of course, in theory this may not be true, but in most cases this approach may be the most relevant prediction of the actual we are able to obtain.

If we wish to use such re-estimates, an important consideration is how to determine a reasonable period after which to perform the re-estimation. For instance, one may wonder whether it is beneficial to re-estimate the NPV as soon as possible.

If the re-estimation is made earlier, it will potentially deviate less from other forecasts as things may not have changed as much. Of course, the closer the reference point r is to the actual a, the better the results of the analysis using r should compare to the analysis using the actual.

To minimize such influences, preferably the ex-post part of the re-estimation must be as large as possible. Unfortunately, there is no easy guideline for what a reasonable period is.

Later on, we investigate the influence on the analysis of the forecasting quality when we use other reference points than the actual. But first, we discuss the technical consequences for the f/a plot, EQF and reference cone in case of alternative reference points.

4.2.1. Generalization

To account for reference points instead of the actual, we need to generalize the f/a ratio. So far, we have used the term *f/a ratio* to refer to a forecast divided by its actual. We now replace the actual by a more general reference point r. This leads to the notation f/r ratio to denote a forecast divided by a reference point r. When this point is the actual a, this is equal to the f/a ratio as proposed by Boehm [4].

4.2.1.1. f/r plot. We use the notation f/r plot, instead of f/a plot, to denote a forecast to reference point plot. The replacement of the actual by a reference point affects both the horizontal and the vertical axis.

In the current f/a plot, the horizontal axis was scaled using t_a, which indicates the date of the actual. For instance, if the actual occurred after two months and a forecast was made after one month, the f/a ratio would be depicted at the horizontal axis at $(t_f - t_s)/(t_a - t_s) = 1/2$. In the generalized method, we have obtained the reference point at t_r, some time before the date of the actual t_a. Since we refer to reference point r in the f/r plot, we will also use t_r as scaling factor instead of t_a. That is, the horizontal axis of the f/r plot depicts the project progression based on the reference point, defined by $(t_f - t_s)/(t_r - t_s)$.

For the vertical axis, the f/r plot depicts f/r ratios instead of f/a ratios.

4.2.1.2. EQF$_r$*.* The change of reference point not only affects the f/r plot, but also the way we compute the EQF. Let us explain. The EQF is defined by the following formula:

$$\text{EQF} = \frac{\int_{t_s}^{t_a} 1\,dt}{\int_{t_s}^{t_a} |1 - e(t)/a|\,dt}.$$

However, in this formula we do not have the values for the actual a or the value of t_a. We *do* have the values of the reference point r and the time at which this re-estimation was made t_r. Since we use this re-estimation as reference, we substitute a with our more general reference point r and t_a with t_r. This leads to the following notation of the EQF:

$$\text{EQF}_r = \frac{\int_{t_s}^{t_r} 1\,dt}{\int_{t_s}^{t_r} |1 - e(t)/r|\,dt}. \tag{6}$$

We denote this EQF with EQF$_r$ to indicate the difference between the formulas. Note that for EQF$_r$, the deviations to the reference point are not time-weighted over the entire economic life span of the asset, yet, only over the time span that we have seen thus far.

4.2.1.3. Reference cone. The reference cone is not affected by the change of reference point. The only necessary change is that the original EQF should be replaced by EQF$_r$. This leads to the following reference lines.

$$l(x) = x + \left(1 - \frac{2}{\text{EQF}_{r,l}}\right) \cdot (1 - x) \tag{7}$$

$$u(x) = x + \left(1 + \frac{2}{\text{EQF}_{r,u}}\right) \cdot (1 - x). \tag{8}$$

Note that x, given by $x = t_f - t_s/t_r - t_s$, is relative in these formulas and lies in the interval [0, 1]. These lines use t_r instead of t_a. Still, the slopes of the lines remain the same even though we changed the reference point.

4.2.2. Interpretation

Above we described the technical changes of the generalization, yet not their interpretation. With the f/a plot, EQF and reference cone, we are able to quantify forecasting quality by assessing the data for potential biases and its accuracy. These tools use an objectively measurable actual as reference point. Since it is objectively measurable, the actual is therefore unbiased.

With the f/r plot, EQF$_r$ and reference cone, we use an alternative reference point. This reference point is only partially objectively measurable. That is, it is only possible to objectively measure the ex-post part. The remainder, the ex-ante part, needs to be predicted. Therefore, the reference point is also a forecast. As we know, the result of a forecast is uncertain. Moreover, forecasts are potentially biased. Therefore, it is important to determine the influence of the change of reference point on the analysis of the forecasting quality.

For instance, is the ability of detecting potential biases affected by the change of reference point? If the reference point itself is biased, the overall bias we find in the analyses may be lessened or even nullified. And, how does using another reference point affect the outcome of the EQF_r? For instance, it is possible the change of reference point creates a positive as well as a negative effect on the EQF_r compared to the EQF. Namely, when we use t_r instead of t_a, we give more weight to the earlier forecasts. These earlier forecasts may appear more accurate if the reference point is closer to the forecast than the actual. On the other hand, it is also possible they appear less accurate when the reference point is further away.

Beforehand, in an unbiased situation there is no way to determine whether the reference point will be below or above the actual. Namely, any forecast and therefore the reference point should be centered around the actual. Therefore, it is equally likely to be above or below, or oscillate around the actual. Due to this oscillating behavior, the EQF_r will not necessarily improve with respect to the EQF. It may just as well be lower than the EQF.

To investigate the effect of the generalization, we analyze data of four organizations, of which we have multiple forecasts and of which we also know the actual [14]. This allows us to compute the forecasting quality using the actual as reference point. Subsequently, we use forecasts made before the final realization as alternative reference points. Using these reference points, we are again able to assess the forecasting quality. The differences between the outcomes of both analyses, provides insight in the effect of replacing the actual by a reference point in real-world cases.

As said, the data was obtained from four organizations. The first data set from a large financial organization, Y, consist of 667 forecasts of in total 140 project costs. The second organization is a multinational organization, X, for which we have 3767 forecasts for in total 867 project costs. The third data set is of a large telecommunications provider, Z, consisting of 971 forecasts of in total 307 project costs. The final data set is of Landmark Graphics containing 923 forecasts of in total 121 project durations.

In the data sets, we aim to use forecasts as reference points that are in the range of 20%–60% project completion. We chose this range to mimic the other data set of organization Z containing NPV calculations, which we will analyze in the next sections. There, we find that many re-estimations of benefits and asset usage cost are made between 0.5 and 1.5 years after project completion. With an average project duration of 1 year and an average life span of 5 years, this computes to roughly the same 20%–60%. The chosen range is thus a realistic reflection of the reference points that we will use in our case study to assess the forecasting quality of NPVs.

Some projects may not have a forecast available in the range 20%–60% to be used as reference point. In most cases, the actual is then replaced by a forecast in the range 60%–100%.

Below, we show for each data set the analysis using the actual and alternative reference points, and discuss the differences that occur in these analyses.

4.2.2.1. Organization Y. Of the total 140 projects, 125 projects had enough forecasts, to allow us to use one of their forecasts as reference point. The range in which 80% of these reference points were made, is 36%–74% project completion. Of the 125 projects, we have a total of 352 forecasts for the f/r plot. Since we wish to compare the quality of the assessments, we use the same forecasts for the f/a plot. Note that in the f/a plot, we compare these forecasts with the actual and in the f/r plot with an alternative reference point.

In Fig. 5 we depict both the f/a plot and the f/r plot of the forecasts. The f/a plot shows no bias, whereas the f/r plot reveals a potential optimistic bias. However, the f/a ratios with a median of 1.00 and the f/r ratios with a median of 0.97 indicate there is no particular bias in either case.

The forecasting quality of the f/r plot is slightly higher than that of the f/a plot. Although the 20% quantile of 3.3 versus 3.6, is slightly lower, the median EQF_r of 7.8 versus 8.1 for the EQF is higher. Interestingly, the differences between EQF and EQF_r are a median 20% of the original EQF, indicating the EQF is generally higher. In fact, only 38% of the projects have a higher EQF_r than the EQF value. Most of these projects have relatively small EQF values. Not surprisingly, in these cases, the EQF_r almost always improves, thereby increasing the overall EQF_r quality. For individual projects, the EQF and EQF_r differ significantly. It is not unlikely to find the EQF_r being half to two times the original EQF. Compared to benchmarks from the literature [14,37,11], the EQF values in both analyses are relatively high.

In this organization with unbiased and relatively accurate forecasts, the f/a plot and f/r plot lead to the same conclusion. Both the bias and the forecasting accuracy show similar results. Therefore, the analyses do not appear to be hampered by changing from reference point in this case. However, due to the large differences between the EQF and EQF_r for individual projects, it is more difficult to determine which projects have inaccurate forecasts. Projects with low EQF values are likely to have a better EQF_r value when a different reference point is used.

4.2.2.2. Organization X. From the 867 projects, we selected 713 projects that had sufficient forecasts to allow for a change of reference point. Of these alternative reference points, 80% are in the range 40%–87%. For the f/r plot we have 1373 forecasts available. We use the same forecasts in the analysis of the f/a ratios. Fig. 6 depicts the results of both analyses.

The plots show a different bias. The f/a plot reveals an overpessimistic pattern supported by a median f/a ratio of 2.16. However, the f/r plot shows no particular bias with a median f/r ratio of 1.00. The explanation is that the reference point itself is equally biased as the other forecasts. By comparing the forecasts to their biased reference point, the bias disappears.

Note that, in this case, the bias is equal for both forecasts and reference points, since there is no convergence to the reference point. This result is evident from both plots. If there is convergence, the bias of the reference point should be less than that of the forecasts. In that case we would still see the bias, as we will see in the next case studies. However, if there

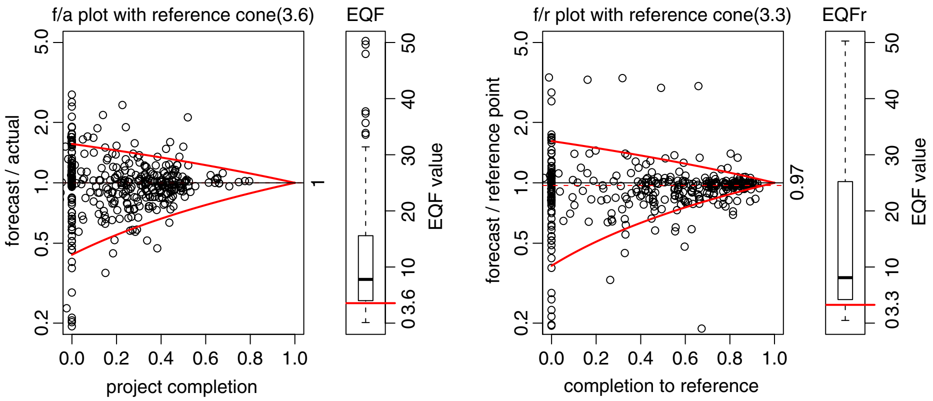

Fig. 5. The f/a plot and f/r plot of 352 forecasts of in total 125 project cost. Both are unbiased, show convergence to the reference point and have similar median forecasting quality.

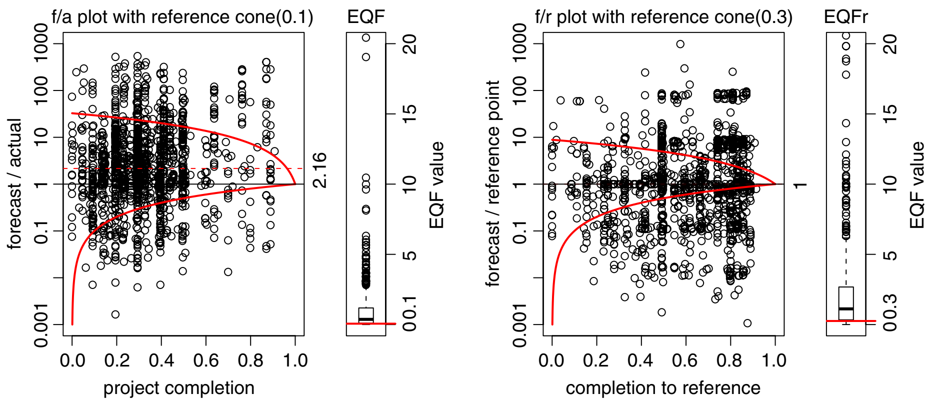

Fig. 6. The f/a plot and f/r plot of 1373 forecasts of together 713 project cost. The former shows an overpessimistic bias, while the latter reveals no bias. Both plots show no convergence to the reference point and their median forecasting accuracy is similar.

is no convergence to the reference point, the forecasts will have the same bias as the reference point, thereby nullifying the bias in the f/r plot.

The assessment of the forecasting accuracy shows considerable differences. The EQF_r is higher than the EQF, in this case for both the 20% quantile, with 0.3 and 0.1 respectively, and the median value, with 1.11 versus 0.37. For 67% of all projects, the EQF_r value is higher than their EQF value. For individual projects, the EQF_r can be up to 0.7 to 10 times the original EQF.

This example shows an organization with low forecasting quality and an overpessimistic bias. By using an alternative reference point, the overall forecasting quality improves, yet remains relatively low. Generally, the EQF_r values are higher than their corresponding EQF values. The f/r plot no longer reveals any bias. The reason is that the alternative reference point itself is equally biased as the forecasts. When the f/r plot shows no convergence to the actual, the plot no longer reveals the true bias of the forecasts. Of course, in this case, the most important problem is the absence of convergence, which the f/r plot still shows.

4.2.2.3. Organization Z. In this data set, 253 projects had more than one forecast. We selected 253 forecasts as alternative reference points, of which 80% are in the range of 11%–60%. In total, 554 forecasts remain for the f/r plot and the f/a plot. Fig. 7 depicts the results.

Both plots show no sign of a particular bias. This is supported by the median f/r ratio of 1 and the median f/a ratio of 1. Since the forecasts have no particular bias, the chosen reference points do not have a bias either.

The forecasting accuracy in terms of EQF_r is lower than that in terms of EQF. Both the 20% quantile, with 0.96 versus 1.5, and the median of 2.6 versus 4.3, illustrate that for this organization, the accuracy using an alternative reference point is lower than the accuracy compared to the final realization.

For this organization, the effect of changing the reference point is noticeable in the forecasting quality. The EQF_r values are considerable lower than the EQF values. However, both analyses show that the forecasts are unbiased and converge to the reference value.

4.2.2.4. Landmark graphics. In this data set, all the 140 projects had sufficient forecasts that could be used as alternative reference points. The alternative reference points are for 80% contained in the 31%–71% interval. In total, 352 forecasts remained that we used in both the f/r plot and f/a plot.

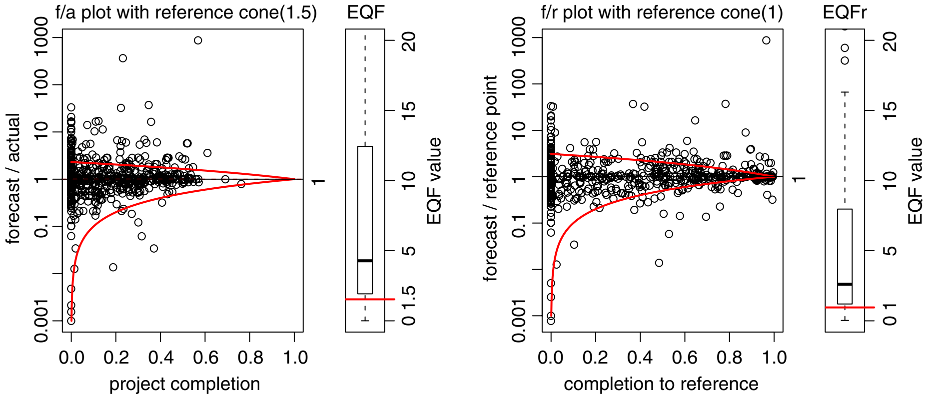

Fig. 7. The f/a plot and f/r plot of 554 forecasts of in total 253 project cost. Both plots indicate that no particular bias is present and the forecasts converge to their reference point. The median forecasting quality is about the same in both analyses.

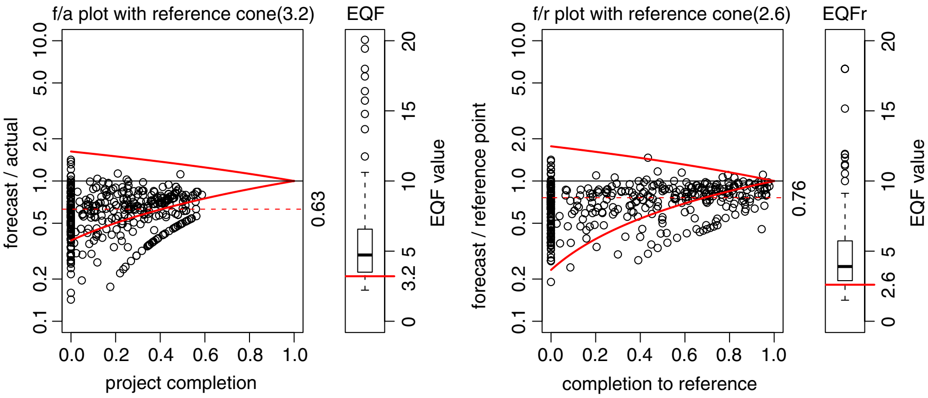

Fig. 8. The f/a plot and f/r plot of 381 forecasts of 121 project durations. Both plots illustrate an optimistic pattern and convergence to the reference point. The EQF_r values are slightly lower than the EQF values.

Fig. 8 depicts the results of the analyses. The f/r plot and the f/a plot indicate the forecasts have an optimistic bias. This result is confirmed by the median f/r ratio of 0.76 and the median f/a ratio of 0.63. Although both show the bias, the bias of the f/a plot is larger than in case of the f/r plot. This is not surprising as the forecasts that we use as reference points in the f/r plot are themselves biased. Since the forecasts converge to the reference points, visual in both f/a plot and f/r plot, the bias of the reference point is smaller than for earlier forecasts. Thus, the bias remains visible, but is less pronounced than in the f/a plot.

For this organization, the EQF_r values are generally lower than the EQF values. The f/r ratios show a 20% quantile of 2.6 and a median of 3.9, whereas the f/a ratios have a 20% quantile of 3.2 and a median of 4.7. The impact for individual projects is relatively small in this organization. The difference between EQF and EQF_r is in 80% of the cases between 0.3 and 1.5 times the EQF.

The analyses using the actual and alternative reference points, lead to similar conclusions. The forecasts have an optimistic bias and converge to the reference value. The bias is less conspicuous in the f/r plot than in the f/a plot. The EQF_r values are generally lower than the EQF values. Of all projects combined, the forecasting accuracy in both analyses is relatively similar.

4.2.2.5. Summary. The replacement of the actual by a reference point affects the analysis of forecasting quality. In both cases it is possible to distinguish whether the forecasts converge to the reference point or not. If the forecasts do not converge, it is no longer possible to distinguish biases in the f/r plot. If the forecasts do converge, the f/r plot will show a bias similar to the f/a plot. However, the bias derived from the f/r plot will be less than that of the f/a plot. Therefore, the true bias will generally be worse than the one we find in the f/r plot.

With respect to the effect on the forecasting accuracy measured in EQF and EQF_r, there is no way to determine the impact based on the f/r plot. For some of our case studies, the EQF_r values were generally higher than the EQF values, while for others they were lower. Beforehand, in an unbiased situation there is no way to predict this behavior, as the reference point may be both higher and lower than the actual value. Therefore, for individual projects it is no longer possible to identify outliers with respect to forecasting accuracy. Still, when we analyze the forecasting accuracy of all projects combined, the overall conclusions remain similar.

The results show that care should be taken when using the EQF_r values for benchmarking purposes. The EQF_r values do not lend themselves particularly well for benchmarking between different organizations. Such comparisons are hampered by the moment at which re-estimations are made and the biases of the organizations. For benchmarking between organizations, the EQF based on the actual as reference point should be used. The EQF_r values, however, do have their value for comparisons within an organization.

4.3. Negative values

Finally, a problem arises for the f/a ratio and f/r ratio in case of negative values of f, a, and/or r. If we allow for negative reference points, the f/r plot becomes dissected into three parts of which one indicates underestimation and two overestimation. If negative forecasts f are possible as well, the f/r ratio itself becomes ambiguous, as it is not clear whether the forecast or reference point is larger than the other. To this end, we need to find a different way of dealing with negative values.

This different way should preferably allow for an identical assessment of the forecasting quality as the f/r ratio. Therefore, before considering alternatives, we first enumerate the characteristics the f/r ratio has.

- Divisible: The ratio shows whether the forecast is larger than its reference point or vice versa. The ratios are strictly separated by some value that divides the ratios into two groups of under- and overestimation. For the f/r ratio this means that all values greater than 1 indicate the forecast is larger than the reference point and all values smaller than 1 indicate the forecast is smaller. This is useful as it allows to detect potential biases in the forecasts made.
- Difference: The f/r ratio shows the difference between the forecast and its reference point. It allows us to distinguish between accurate forecasts, in this case ratios close to 1, and inaccurate forecasts, which are very large ratios or ratios close to zero. Thus, the ratio displays the accuracy of the forecast with respect to its reference point.
- Scaling: Not only does the f/r ratio show the difference, but it also shows how accurate or inaccurate the forecasts are. Using the logarithmic scale, the f/r plot depicts equal inaccuracies for under- and overestimation at an equal distance. That is, both a ratio of 1/2 and 2, which represent an equal under- and overestimation for positive entities, have an equal distance to the reference point in the plot. This equal distance is relevant as it allows for a visual assessment of the quality of forecasts and the severity of potential biases.
- Relative: By dividing the forecast with its reference point, the ratio becomes independent of the size of the asset. Namely, in case of positive valued entities, the reference point is often a good indication of the size. For instance, the higher the actual cost, the larger the project. Similar arguments can be made, for instance, for the project duration and effort of a project. The advantage of accounting for the size is that both small and large projects are assessed equally by allowing similar leniency. Thus, a ratio of 2 indicates the same for different sized projects.

  Note, that we assume r to be a good approximation of size. If in certain cases r is not, the f/r ratio may not have the relative property.

This enumeration gives us a guideline to develop alternatives to deal with negative values. We wish to find an alternative that retains the listed properties and resolves the problems caused by negative values. Below, we address a number of ways, among others, alternative ratios, to cope with negative forecasts and reference points.

4.3.1. Alternative approaches

Below, we discuss a number of alternative approaches. Note that this enumeration is not exhaustive.

4.3.1.1. Mathematical translation. A possible solution to deal with negative values is to perform a mathematical translation. A mathematical translation is known as the movement of a point, line, or shape a constant distance in a specified direction.

By adding a large number M to each forecast and reference point, it is possible to remove the negative values. For instance, let M be defined as the absolute of the minimum value of forecast and reference point in the data set plus 1. That is, $M = |\min(f, r)| + 1$. Then, we add M to both forecast and reference point, effectively making the ratio $(f + M)/(r + M)$.

This translation solves the problems that arise from negative values by translating the ratios to the solution space of only positive valued entities. The translation has the properties difference and divisible, as it enables distinction between forecast and reference point.

However, due to adding a large number to both numerator and denominator, the ratio does not have the scaling or relative property. Although the ratio is translated to the solution space of positive entities, using the logarithmic scale does not yield similar results. To again allow for equal spacing, we would need to somehow translate the ratios back, so that it retains the same interpretation as the f/r ratio. However, how to do this is not clear.

4.3.1.2. Taking the difference. An alternative approach is to take the difference between the forecast and the reference point. This alternative does not have a problem with negative values like the f/r ratio. If $f > r$ then $(f - r) > 0$ regardless of the signs, resulting in a positive valued ratio. If $f < r$ then $(f - r) < 0$ resulting in a negative valued ratio. Thus, the difference is either positive or negative depending on whether the forecast is larger or smaller than the reference point.

Besides the divisible property, the value also complies with the difference and scaling property. For a number of economic indicators, such as the ROI and IRR, it is sufficient to use this value to assess the forecasting quality. Namely, the relative property for the $(f - r)$ value is needless as the underlying indicators are already relative themselves. For instance, the ROI relates the net benefits of the asset to its investment cost. Therefore, an ROI value of 10% means the same for an asset of 100 dollar and of 1 million dollar. As a result the $(f - r)_{\mathrm{IRR}}$ value is automatically relative. It is thus possible to analyze the forecasting quality of these entities using the value $f - r$.

However, other economic indicators, such as the NPV, are not relative themselves. Namely, if the difference between forecasted NPV and reference NPV is 2 Euro, the interpretation is completely different when the reference point NPV is 10 Euro or 1 million Euro. Therefore, we need another alternative ratio that is relative for all entities.

4.3.1.3. Taking the relative difference. An alternative to the $(f - r)$ value, is to divide the difference between forecast and reference point by an approximation of size, denoted by s. By dividing by an approximation of size, we make the differences between forecast and reference point relative. This means that the $(f - r)/s$ ratio has all the properties we listed for the f/r ratio and is thus suitable as an alternative. Furthermore, the $(f - r)/|s|$ ratio has the same interpretation as the f/r ratio. Namely, overestimations are represented above the horizontal line, while underestimations are depicted below this line.

There are many candidates for s. One example is the absolute size of the reference point, $|r|$. However, in that case the ratios do not necessarily indicate the same for a small and a large asset. For instance, the re-estimated NPV value does not adequately describe the size of an asset. Namely, suppose the re-estimation of the NPV results in an NPV of 1. It is possible for a large multi million Euro asset to result in an NPV of 1, but this can also apply to an asset of 100 Euro. Still, the ratio would yield the same result if the re-estimated NPV would turn out to be 2, even though the difference in the forecast is relatively small for the multi million Euro asset and relatively large for the 100 Euro asset. Therefore, in case of the NPV, this is not the best approximation.

For the $((f - r)/|s|)_{\mathrm{NPV}}$, other alternatives are, for example, the forecasts or re-estimations of the benefits, project cost, asset usage cost or asset cost. In case of entities that are already relative themselves, such as the IRR or ROI, the approximation of size s can be chosen 1, or $s = 1$.

4.3.1.4. Notation. Above we introduced a family of alternative ratios that is usable for positive and negative valued NPVs. Since the notation is cumbersome, we introduce a new notation here. We will use the term $(f - r)_e/s$ ratio, when we analyze the $(f - r)_e/|s_d|$ ratio for entity e. In this case, s_d is the size of entity d and should always be further specified. We will use $(f - r)/s$ ratio, when it is clear which entity is referred to.

4.3.1.5. Approximating size. We stated that candidates for approximations of size for the NPV are, for example, the forecasts or re-estimations of the benefits, project cost, asset usage cost or asset cost. The choice between these discounted benefits, project cost, asset usage cost and asset cost is arbitrary as long as they are adequate approximations of the size of an asset. In this article, we will choose the asset cost, as it incorporates all the cost that are made for the asset over its entire life span.

Then, the question remains whether we use the forecasted or re-estimated asset cost. That is, do we use f_{A} or r_{A}? This choice mainly depends on the purpose for which we wish to use the ratio.

The use of f_{A} is best if we wish to enhance decision information. Let us explain. With the $(f - r)/s$ ratio and using $s = f_{\mathrm{A}}$, it is possible to apply historic information to new project proposals, for instance, by means of a confidence interval. Suppose we have an 80% confidence interval of $[-0.25, 0.75]$ of historic $(f - r)_{\mathrm{NPV}}/s$ ratios for the benefits with $s = f_{\mathrm{A}}$. Consider a new project proposal that forecasted the NPV to be 5 million Euro and the total asset cost to be 2 million. This would allow us to enrich our decision information in the following manner. We want to know r. We know that $(f - r)_{\mathrm{NPV}}/s = -0.25$ leading to $(5 - r)/2 = -0.25 \rightarrow r = 5.5$. For the upper bound of the range, we find $(5 - r)/2 = 0.75 \rightarrow r = 3.5$. Thus, based on historic data we would expect this new project proposal in 80% of the cases to yield an NPV between 3.5 and 5.5 million Euro.

This enhancement of decision information is not possible if we would use $s = r_{\mathrm{A}}$, instead of $s = f_{\mathrm{A}}$. For example, suppose we have a 80% confidence interval of the $(f - r)_{\mathrm{NPV}}/s$ ratio with $s = r_{\mathrm{A}}$ of $[-0.25, 0.75]$ for the NPV. Suppose we have a new project proposal that forecasts the NPV to be 5 million Euro and the asset cost of 2 million Euro. We want to assess r, thus we would try $(2 - r)_{\mathrm{NPV}}/r_{\mathrm{A}} = -0.25$. Unfortunately, we cannot find r_{NPV}, since we do not know r_{A} at that time. Therefore, it is best to use f_{A} when we want to enhance decision information.

In Section 5 we will use $s = r_{\mathrm{A}}$ and in Section 6 we use $s = f_{\mathrm{A}}$.

4.3.1.6. Economic interpretation. We propose to use the $(f - r)/s$ ratio with either $s = f_{\mathrm{A}}$ or $s = r_{\mathrm{A}}$ to analyze the forecasting quality of the NPV. This ratio is not entirely arbitrary in the economical context of the NPV.

Suppose we would choose as approximation of size s the forecasted investment cost I of the asset. In this case, the $(f - r)_{\mathrm{NPV}}/s$ ratio is closely related to the ROI. Note, that the ROI can be calculated using the formula $(f_{\mathrm{B}} - f_{\mathrm{C}})/f_{\mathrm{I}}$.

Let us explain. When we dissect the $(f - r)_{\mathrm{NPV}}/s$ ratio, we find that it consists of the difference between $f/|s|$ and $r/|s|$. Rewriting the ratio gives $f_{\mathrm{NPV}}/|s| = \frac{(f_{\mathrm{B}} - f_{\mathrm{C}} - f_{\mathrm{I}})}{f_{\mathrm{I}}} = f_{\mathrm{ROI}} - 1$. Similarly, $r_{\mathrm{NPV}}/|s| = \frac{r_{\mathrm{B}} - r_{\mathrm{C}} - r_{\mathrm{I}}}{f_{\mathrm{I}}}$. This fraction has the interpretation $r_{\mathrm{ROI}} - 1$. This means that using the $(f - r)_{\mathrm{NPV}}/s$ ratio with this approximation of size, we basically analyze $(f - r)_{\mathrm{ROI}}$.

Therefore, the proposed $(f - r)/s$ ratio for the NPV makes sense given the economical context. In this article, we will choose either the forecasted or re-estimated asset cost for s.

4.3.2. Extension

The proposed $(f - r)/s$ ratio makes it necessary to evaluate the way we construct the tools that we use to analyze forecasting quality in case of negative values. In the previous section, we showed how to use the f/r plot, EQF and reference cone when assessing forecasts and reference points with non-negative values. There we altered the f/a ratio to account for other reference points than the actual, leading to the f/r ratio. However, the tools discussed there are not applicable in case of negative values. Therefore, we discuss how the f/r plot, EQF and reference cone should be altered to account for the proposed $(f - r)/s$ ratio in situations that negative values occur.

4.3.2.1. $(f - r)/s$ plot. In the $(f - r)/s$ plot we depict the $(f - r)/s$ ratios against their progression relative to the reference point. Therefore, the $(f - r)/s$ plot has the same horizontal axis as the f/r plot. We use the date of the reference point t_r to scale the dates of the other forecasts.

The vertical axis of the $(f - r)/s$ plot depicts the $(f - r)/s$ ratios. Note, that for these ratios the line $(f - r)/s = 0$ is the line between underestimation and overestimation. With the f/r ratio this is the line $f/r = 1$.

Furthermore, since negative values are possible for this ratio, the logarithmic transformation that we used for the f/r plot no longer applies. However, there is no problem abandoning the transformation in this case and use a linear axis, since the $(f - r)/s$ ratio allows for scaling. That is, for this ratio the distance between -2 and 2 on the linear axis indicates an equal accuracy of under- or overestimation.

4.3.2.2. $(f - r)/s$ ratio versus f/r ratio. We described that we will use the $(f - r)/s$ ratio for entities with both positive and negative valued forecasts or reference points, and the f/r ratio for non-negative entities. One may wonder, why we do not used the $(f - r)/s$ ratio for the non-negative entities as well. For instance, instead of f/r we could use the $(f - r)/s$ ratio.

However, this is not always a good alternative for non-negative valued entities. Let us explain. Suppose we use the f/r ratio to assess the forecasting quality of project cost. With the f/r ratio, all the underestimates, that is a forecast is smaller than its actual, will be in the range $[0, 1]$, and all overestimations are in the range $[1, \infty]$. Using the logarithm, these ranges are made equal. Namely, the logarithm of the underestimations transforms the range to $[-\infty, 0]$, and for the overestimates to $[0, \infty]$.

Now suppose we would use the $(f - r)/s$ ratio to assess the forecasting quality of project cost. As the reference point of the project cost is a good approximation of size, the most likely choice for s is the reference point itself, thus $s = r$. With the $(f - r)/r$ ratio all the underestimates, that is forecast smaller than the actual, will be in the range $[-1, 0]$, and all overestimation are in the range $[0, \infty]$. This range has no equal scaling. Since there are now negative values, it is not possible to use the logarithm to make the ranges comparable. Therefore, in such cases, the f/r ratio is more appropriate for non-negative valued entities.

Note that this argument holds when r is a good approximation of the size of the entity.

4.3.2.3. EQF_s. The original EQF divided the size of the actual by the difference between the forecast and its actual. In this case, we use an alternative size measure to compare the forecasting quality with. Therefore, we need to adapt the EQF for the new ratio. The EQF of the $(f - r)/s$ ratio, which we denote with EQF_s, is given by the following formula.

$$\mathrm{EQF}_s = \frac{\text{Area under } s}{\text{Area between forecast and reference value}}$$

$$\mathrm{EQF}_s = \frac{\int_{t_0}^{t_r} s\,dt}{\int_{t_0}^{t_r} |e(t) - r|\,dt} \tag{9}$$

$$\mathrm{EQF}_s = \frac{\int_{t_0}^{t_r} 1\,dt}{\int_{t_0}^{t_r} |(e(t) - r)/s|\,dt}. \tag{10}$$

This formula calculates 1 divided by the distance between the $(f - r)/s$ ratios and 0. This is also illustrated in Fig. 9 that shows an example EQF_s calculation. The computation is similar to the EQF_r, which computes 1 divided by the distance between the f/r ratios and 1.

4.3.2.4. Reference cone. Finally, we need to re-evaluate the reference cone as discussed in Section 2. The original reference cone no longer holds in this situation, since its reference lines are centered around an f/r ratio of 1 and are positive for both under- and overestimation. However, the $(f - r)/s$ plot is centered around 0 and depicts underestimations as negative ratios. Therefore, we need to recompute the reference lines.

Most assumptions discussed in Section 2 that were used to draw the reference lines, are still applicable. That is: we assume that each consecutive forecast incorporates the ex-post part and we assume this part is known with certainty; the accuracy of the ex-ante part is assumed to remain constant as the project progresses; and the goal of the forecast is to predict without bias and as quickly and as accurately as possible the actual value of interest for the project.

However, the ex-post growth assumption needs further consideration. We assumed the growth of the ex-post part to be described by a constant function. In many cases, the NPV will not comply to this assumption. Namely, during project execution investment cost are made until the time the project is completed. From that point on, benefits are potentially

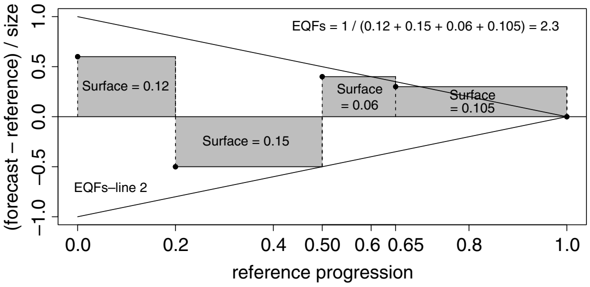

Fig. 9. An example of an EQF_s calculation for four forecasts.

generated with the asset that together are hopefully higher than the asset usage cost. This causes the NPV to grow, which continues over the economic life span of the asset. This common growth of the NPV is not a constant growing function.

However, we will make the crude assumption that the NPV does grow following a constant function. We assume the initial project cost are made right at the start of the project and the benefits are generated immediately. That is, the NPV will grow constant from the moment the project starts. This means the growth follows a constant growth line from $-I$ to r. Even though this is a crude assumption, we use it since, due to its simplicity, it is useful to illustrate the calculations required to compute the reference lines. If other growth functions are used the calculations become more involved, but the method not.

We are able to derive alternative reference lines as follows. First, we need to specify the forecast at relative time x, $e(x)$. Note that $x = (t - t_s)/(t_r - t_s) \in [0, 1]$. Each forecast consists of an ex-post and an ex-ante part. For the ex-post part, we incorporate as much information as possible and we know this information exactly. We describe the ex-post part with a linear function that is given by $p(x) = (r + I)x - I$. Moreover, we define $p(0) = 0$.

The ex-ante part is not known with certainty. We assume we are able to predict the ex-ante part with a forecasting accuracy of c_1 for underestimation and c_2 ($c_2 \geq c_1$) for overestimation. This indicates that we are able to predict the ex-ante part within c_1 and c_2 times its value. The size of the ex-ante part is defined by the reference value minus the ex-post part. Mathematically, the ex-ante part is given by $r - p(x)$.

With the ex-post and ex-ante part defined, we are able to specify $e(x)$. In case of solely underestimations, $e_l(x)$ is given by $p(x) + c_1 \cdot (r - p(x))$. For solely overestimations, we find $e_u(t) = p(x) + c_2 \cdot (r - p(x))$.

However, the $(f - r)/s$ plot does not depict $e(x)$, but, it illustrates $(f - r)/s$ ratios. Therefore, we need to formulate the reference lines in terms of $(f - r)/s$ ratios. This leads to the following formulas.

$$
\begin{aligned}
l(x) &= \frac{e_l(x) - r}{|s|} = \frac{(r + I)x - I + c_1 \cdot (r - (r + I)x + I) - r}{|s|} \\
u(x) &= \frac{e_u(x) - r}{|s|} = \frac{(r + I)x - I + c_2 \cdot (r - (r + I)x + I) - r}{|s|}.
\end{aligned}
$$

Preferably, similar to the current reference lines, we wish to express these lines in terms of EQF_s. We find the relation between EQF_s, c_1 and c_2 by computing the EQF_s. Below, we illustrate the calculations in case of solely overestimations.

$$
\begin{aligned}
EQF_{s,l} &= \frac{\int_0^1 1\,dx}{\int_0^1 |(r - e_l(x))/s|}\,dx \\
&= \frac{\int_0^1 1\,dx}{\int_0^1 \frac{r - (rx + Ix - I + c_1 \cdot (r - rx - Ix + I))}{|s|}}\,dx \\
&= \frac{\int_0^1 1\,dx}{\int_0^1 \frac{r - rx - Ix + I - c_1 r + c_1 rx + c_1 Ix - c_1 I}{|s|}}\,dx \\
&= \frac{1}{[rx - \frac{1}{2}rx^2 - \frac{1}{2}Ix^2 + Ix - c_1 rx + \frac{1}{2}c_1 rx^2 + \frac{1}{2}c_1 Ix^2 + c_1 Ix/|s|x]_0^1} \\
&= \frac{|s|}{r - \frac{1}{2}r - \frac{1}{2}I + I - c_1 r + \frac{1}{2}c_1 r + \frac{1}{2}c_1 I + c_1 I} \\
&= \frac{2|s|}{(r + I)(1 - c_1)}.
\end{aligned}
$$

From this formula, we find the following.

$$
\frac{2|s|}{\text{EQF}_{s,l}} = (r + I)(1 - c_1)
$$

Table 2
Overview of the tools that are applicable in different situations.

Entity value	Reference point	Type of plot	EQF formula	Reference cone formulas
Non-negative	Actual	f/a	(1)	(3) and (4)
Non-negative	Reference point	f/r	(6)	(7) and (8)
Positive or negative	Reference point	$(f-r)/s$	(10)	(11) and (12)

$$\frac{2|s|}{(r+I)\mathrm{EQF}_{s,l}} = 1 - c_1$$
$$\frac{2|s|}{(r+I)\mathrm{EQF}_{s,l}} - 1 = -c_1.$$
$$1 - \frac{2|s|}{(r+I)\mathrm{EQF}_{s,l}} = c_1.$$

Via similar calculations, we find that $c_2 = 1 + \frac{2|s|}{(r+I)\mathrm{EQF}_{s,u}}$. If we substitute these values in the formulas of our reference lines, we find the following expressions.

$$\begin{aligned} l(x) &= \frac{(r+I)x - I + \left(1 - \frac{2|s|}{(r+I)\mathrm{EQF}_{s,l}}\right) \cdot (r - (r+I)x + I) - r}{|s|} \\ &= \frac{(r+I)x - I + r - (r+I)x + I - r}{|s|} - \frac{\frac{2|s|}{(r+I)\mathrm{EQF}_{s,l}}(r - (r+I)x + I)}{|s|} \\ &= \frac{-2(r+I)(1-x)}{(r+I)\mathrm{EQF}_{s,l}} \\ l(x) &= \frac{-2}{\mathrm{EQF}_{s,l}}(1-x) \qquad (11) \\ u(x) &= \frac{2}{\mathrm{EQF}_{s,u}}(1-x). \qquad (12) \end{aligned}$$

These reference lines describe the reference cone that is usable for the $(f-r)/s$ plot. The calculations lead to relatively simple reference lines as also illustrated in Fig. 9. This is caused by the modest assumptions that we used for the growth of the ex-post part. Using the above steps, it is possible to derive other reference lines given different assumptions for the growth of the ex-post part. As calculations can become involved with more complex assumptions, we recommend using computer algebra packages like Maple [41] to compute the results.

With this generalized method, we derived at an $(f-r)/s$ plot, EQF_s and alternative reference cone that allow for an assessment of forecasting quality for negative valued entities.

4.4. Summary

In this section, we discussed that three problems may arise in applying the f/a ratio when we use it to quantify forecasting quality. The f/a ratio runs into visual complications in asymptotic behavior; fails if the actual is not objectively measurable; and does not cope with the case of non-negative valued entities. To overcome these problems, we extended and generalized the method.

This generalized method is applicable for any entity. The tools to be used depend on the entity in question. In Table 2, we give an overview of the tools that are applicable in each situation. Below, we summarize these different situations.

4.4.0.5. Non-negative valued actual. The method for non-negative valued entities with the actuals as reference point consists of an f/a plot, *EQF* and reference cone. The f/a plot depicts f/a ratios against the moment the forecast was made relative to the actual a. The f/a ratios on the vertical axis are drawn on a logarithmic scale.

In case the forecast f is zero, the resulting f/a ratio is visualized by plotting the ratio one order below the minimum f/a ratio. That is, if the minimum f/a ratio is 1/2, the ratio is drawn at 1/20. For better distinction, this point should be depicted differently from the other ratios. When the actual a is zero, the f/a ratio is drawn a clear distance above the maximum f/a ratio. Finally, when both forecast and reference point are zero, the point is visualized by $f/a = 0/0 = 1$.

4.4.0.6. Non-negative valued reference point. The method for non-negative valued entities consists of an f/r plot, EQF_r and reference cone. The f/r plot depicts f/r ratios against the moment the forecast was made, relative to the reference point r. The f/r ratios on the vertical axis are drawn on a logarithmic scale.

Comparing the forecasting quality against a reference point r affects the analysis with respect to the f/a plot. In both the f/r plot and the f/a plot, it is possible to distinguish whether the forecasts converge to the reference point or not. If the forecasts do not converge, it is no longer possible to distinguish biases in the f/r plot. In case of convergence, the f/r plot shows a similar bias, which is generally less clearly pronounced than in the f/a plot.

Using data of four organizations, we found that the EQF_r is sometimes higher than the EQF and sometimes lower. Still, when we analyze the forecasting accuracy of all projects combined, the conclusions remain similar.

However, for individual projects it is no longer possible to identify outliers with respect to the forecasting accuracy. For this reason, the EQF values computed using reference points other than the actual are not suitable for benchmarking purposes between organizations. For assessing forecasting quality, preferably the actual is used for the analyses instead of a more general reference point.

4.4.0.7. Positive and negative valued. When entities can take both positive and negative values for both their forecasts and reference points, the method consists of the $(f - r)/s$ plot, EQF_s and an alternative reference cone given by Formula (11) for the lower reference line and Formula (12) for the upper reference line. The $(f - r)/s$ plot depicts $(f - r)/s$ ratios against the relative progression with respect to the reference point r. The approximation of size s of an asset needs to be specified for each figure. For the NPV, candidates are the forecasts or re-estimations of the benefits, asset usage cost, project cost or asset cost. In the analyses of the NPV, we will in this article use both f_{A} and a_{A}.

5. Prerequisite data analysis

In the previous section, we discussed the necessary tools to assess the forecasting quality of the NPV, benefits, asset usage cost and project cost. In this section, we describe the data that we obtained from the large telecommunication organization Z. An investigation of this data is prerequisite for analyzing the forecasting quality of the organization in the next section.

First, we give an overview of the data we obtained. Since this data involves reference points and not actuals of the NPV, we discuss at which point in time the re-estimations are made. Next, we assess the data of organization Z by comparing it to benchmark data from the literature. Finally, since our tools assume homogeneous data, we check for heterogeneity in the data set provided.

In this section we will make use of statistical tests to statistically verify possible heterogeneity. For all the tests we will perform, we use a threshold of $\alpha = 0.05$ for accepting or rejecting the hypothesis that is being tested.

5.1. Data overview

From a large telecommunication organization Z, we received forecasts and re-estimations of NPVs of 102 projects. These projects were executed in the period of 2001 till 2009. In total, the overall benefits amount to 4714 million Euro and asset cost of 2908 million Euro, which consist for 173 million Euro of investment cost. The median project duration is 327 days and the assets have a median expected economic life span of 5.5 years.

More precisely, the data consists of the following components for each project.

– the actual startdate of the project (t_s).
– the actual enddate of the project (t_e).
– the forecasted enddate of the asset (t'_a).
– the initial forecast date of the NPV (t_f).
– the re-estimation date of the NPV (t_r).
– the forecasted asset life span (N).
– the forecasted and re-estimated NPV.
– the forecasted and re-estimated benefits (B).
– the forecasted and re-estimated asset usage cost (C).
– the forecasted and actual project cost (I).
– An asset classification: New Product Development, Cost Reduction or Sales Expansion.

In the above enumeration, the benefits, asset usage cost and project cost refer to the cumulated discounted totals. In Formula (5) in Section 3.2, we denoted them with B, C and I.

The data contain a large number of assets, that in their re-estimation have zero benefits. A number of IT projects was canceled or was completed successfully, yet the assets they created no longer had any business value. For instance, it is possible projects were stopped during implementation due to changes in the market. These assets are simple to recalculate, namely the benefits are zero in the post calculation. They comprise a relative large portion of the re-estimated NPVs. This confirms that forecasts and actuals with a value of zero are more common than you would expect.

We do note, that if a project was stopped, for instance, due to changes in the market place, in our analyses the forecasts of its entities may be considered inaccurate. Since the estimators provide single point forecasts, these predictions cannot account for the possibility of stopping the project. These forecasts will be regarded as inaccurate, even though the cancelation of the project may be beneficial to the organization. Those responsible may be disinclined to cancel a project to avoid

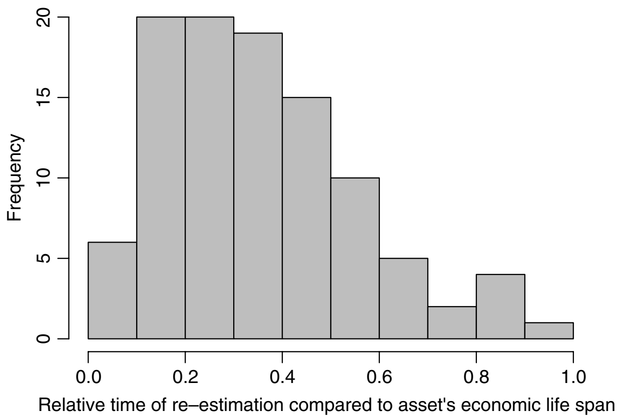

Fig. 10. Distribution of the relative time of the re-estimations with respect to the predicted asset's economic life span.

Table 3
Summary of a/f ratios of NPVs for in total 50 Cost Reduction, Sales Expansion and New Product Development assets as described in a book by Bower [35].

Type of asset	Mean a/f ratio	Standard deviation
Cost reduction	1.1	Narrow
Sales expansion	0.6	Reasonable ≈ 0.2
New products	0.1	Very large

inaccurate forecasts. Organizations should make sure that the accuracy of forecasts does not discourage the cancelation of projects when necessary.

We emphasize that, for the NPV, benefits and asset usage cost, the re-estimations are available and not the actuals. Although the data spans a period of 2001 till 2009, it does not contain actual NPVs as the organization only recomputes them before the economic life span has ended. This emphasizes once more that collecting data with truly finalized NPVs is rather scarce.

5.2. Re-estimation date

The data we obtained consists of re-estimated NPVs, benefits and asset usage cost. To assess the forecasting quality, it is preferable that the ex-post part of the re-estimations is as large as possible. That is, the re-estimation should be made shortly before the final realization. However, from the perspective of gaining relevant decision information, we would prefer the reference point to be made more closely after project completion. Therefore, it is interesting to investigate at which moment the re-estimations are made in the data set.

Organization Z re-estimated the benefits and asset usage cost some time after their initial forecast. The re-estimations were made a median 1.4 years after the initial forecast and a median 0.71 years after project completion. The date was also a median 3.5 years before the asset's predicted economic life span.

To further investigate the re-estimation date, we also assess the relative moment of re-estimation. We define the relative moment of re-estimation as the relative time of the re-estimation over the entire life span of the asset. Mathematically, this moment is given by $(t_r - t_s)/(t_a - t_s)$ with t_s the start date of the asset, t_r the date of re-estimation and t_a the end date of the assets life span. Note, that in this article t_a is given by the number of years over which benefits and cost are predicted.

In Fig. 10, we show a histogram of the relative time of the re-estimations compared to the predicted asset's economic life span. For instance, a value of 0.4 means the re-estimation of the NPV was made at 40% of the predicted economic life span of the asset. The figure shows a large spread in the moment of re-estimation. Many re-estimations are made in the range 20%–60%.

5.3. Literature benchmark

To assess the quality of the data of our real-world case study, we benchmark it against a case study found in the literature. The only benchmark related to the forecasting quality of NPVs we are aware of, is given in a book by Bower [35]. In this book, results of the difference between forecasted and actual NPVs are described for a case study consisting of 50 assets.

For the sake of completeness, we show Table 3 taken from the book. The table contains the mean a/f ratio of the NPV for three asset classifications. Unfortunately, the book reveals little detail about the analysis. Through personal communication,

Table 4
Summary of the r/f ratios of in total 100 assets for our telecommunications organization. For the sake of comparison two assets are left out, since their NPV forecast is zero. These results should not be used for other comparisons.

Type of asset	Mean r/f ratio	Standard deviation	Median r/f ratio	Number of assets
Cost reduction	2.32	5.00	1.03	35
Sales expansion	1.07	0.66	0.81	4
New products	0.42	4.95	0.11	61

we understood that the underlying data is no longer available. Still, the author was able to give an indication of the variation of the figures, allowing us to add the column *standard deviation*.

However, we must be cautious about this benchmark for three reasons. First, since the book was published in 1970, the assets are not IT assets. Therefore, these findings may have limited application in our IT-context.

Second, the case study examines the a/f ratio of the NPV. In Section 4, we discussed that this ratio is ambiguous when used for entities with both positive and negative outcomes, such as the NPV. Bower communicated to us, that the case study contained negative actuals, yet, not how many or their impact on the results. Therefore, the results depicted in the table do not reveal anything about potential biases or the quality of the forecasts.

Third, the case study of Bower uses the actual as reference point. We assume that these are objectively measured final realizations. In our case study, we use re-estimations of the NPV as our reference points.

However, we are still able to use the benchmark to assess whether our data behaves similarly for this ratio. The data we obtained contains twice the amount of assets as Bower and uses the same asset classification. In Table 4, we show summaries of the r/f ratios of our case study. We left two assets out of this overview, since there was a Sales Expansion and a New Product Development asset with a forecasted NPV of zero. Note that we use the r/f ratio and not the $(f - r)/s$ ratio we proposed as alternative ratio for the NPVs for the sake of comparison.

Both the mean and the standard deviations of the r/f ratios of our case study do not resemble the a/f ratio benchmarks of Bower. For all classifications, the mean of the ratios is higher than those observed by Bower. Moreover, in our case study, the standard deviation of the Cost Reduction assets and the New Product Development assets are comparable. The literature benchmark on the other hand, found this standard deviation to be quite different for these classifications. This indicates that our case study involves more variation than found in the literature benchmark. Because the means in our case study are significantly influenced by the variation in the data, we also incorporated the median value of the r/f ratios.

Based on the findings, the data behaves differently to the benchmark of the literature. Both the mean and the variation of the distributions in our data set are considerable larger.

Still, similar to Bower, we find that the categories are different. In both studies, the mean r/f ratio or a/f ratio of the Cost Reduction assets is larger than that of the Sales Expansion and New Product Development assets. This indicates that our data set is possible heterogeneous.

5.4. Heterogeneity

Finally, we analyze the data for heterogeneity. The proposed tools to analyze forecasting quality, assume that the underlying data is homogeneous. Therefore, before we are able to apply the tools, we verify if the data complies to this assumption.

Namely, forecasts may have different uncertainties for different types of assets. For instance, a forecast for the cost of a standard application may have less uncertainties than a forecast for a newly developed product. Therefore, a deviation of 2 times the actual may be a high quality forecast for one asset, whereas it is a poor forecast for another given its range of possible outcomes. In case of a significant difference between certain assets, the forecast quality should be assessed separately. Of course, a discussion is in order as to whether the found difference is justified or is simply caused by making less accurate predictions.

The data we obtained, contains a number of variables that are interesting to investigate in this respect. In the comparison with a data set from the literature, we found that there is a potential significant difference between the asset classifications. Therefore, first, we determine whether the different asset classifications cause heterogeneity in the data set. Moreover, we investigate whether the forecasting quality is correlated with the moment of re-estimation or contains a yearly trend.

5.4.1. Asset classification

Organization Z assigned each asset to one of three categories, being: *Cost Reduction*, *Sales Expansion* or *New Product Development*. The categories consist of 35 Cost Reduction assets, 5 Sales Expansion assets and 62 New Product Development assets. Unfortunately, the amount of assets available in the Sales Expansion category is insufficient to investigate for significant differences with the other categories. Therefore, we leave this category out of the analysis.

To detect potential heterogeneity, we plot the $(f - r)/s$ ratios of the NPV forecasts for both Cost Reduction and New Product Development assets in Fig. 11. In this case, we use $s = r_A$, since we only want to assess whether differences exist.

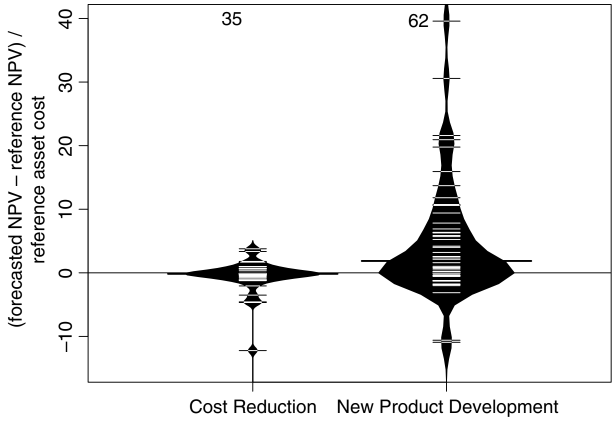

Fig. 11. Potential heterogeneity between Cost Reduction assets and New Product Development assets with respect to the forecasting quality of the NPV.

Table 5
An overview of the outcomes of the statistical tests for the locations of the distributions for Cost Reduction and New Product Development assets.

	Kruskal–Wallis test	Reject/accept	Wilcoxon test	Reject/accept
Same locations NPV	0.00018	Reject	0.000009	Reject
Same locations benefits	0.0004	Reject	0.0002	Reject
Same locations asset usage cost	0.003	Reject	0.002	Reject
Same locations project cost	0.51	Accept	0.26	Accept

Table 6
An overview of the outcomes of the statistical tests for the variation of the distributions for Cost Reduction and New Product Development assets.

	Brown–Forsythe test	Reject/accept
Same variance NPV	0.20	Accept
Same variance benefits	0.03	Reject
Same variance asset usage cost	0.42	Accept
Same variance project cost	0.08	Accept

At this time, we are not interested in enhancing decision information. Therefore, we do not need to use the initial forecast of the asset cost, but are able to use the best approximation of the size that is available.

The figure depicts two so-called bean plots [31], one for each category. The vertical axis shows the $(f - r)/s$ ratios of the NPVs. Each line in the beans represents a single forecast. The bean shape is an approximation of the underlying probability distribution of the data. The large horizontal line in each of the bean plots illustrate the median ratio of the data. A couple of extreme values are not depicted in the range we chose for this figure.

The bean plots illustrate that the location of the distributions is different. The Cost Reduction assets appear unbiased, whereas the New Product Development assets show a slight bias toward overestimation. We used the Kruskal–Wallis test and the Wilcoxon test to verify whether the location of the distributions is significantly different. As shown in Table 5, with a p-value of 0.00018 and a threshold of $\alpha = 0.05$ for the Kruskal–Wallis test, we reject the hypothesis that the location of the distributions is equal. A one-sided Wilcoxon test results in a p-value of 0.000009, which rejects the hypothesis that the location of the Cost Reduction assets is greater than or equal to that of the New Product Development assets. Thus, we find there is a significant difference of location for the $(f - r)/s$ ratios of the asset classifications.

Also, the bean plots show the variance of the $(f - r)/s$ ratios for the Cost Reduction category is smaller than that of New Product Development. We tested whether the variance is significantly different using the Brown–Forsythe test. As shown in Table 6, with a p-value of 0.20 for the Brown–Forsythe test, we cannot reject the null hypothesis. Therefore, the variance of the asset categories is not statistically different.

We also analyzed the f/r ratios of benefits and asset usage cost and the f/a ratios of project cost for heterogeneity with respect to the asset classifications. The analyses of the f/r ratios of benefits show similar results as the $(f - r)/s$ ratio of the NPVs as illustrated in Tables 5 and 6. The tests indicate that the location of the benefit f/r ratios shows significant differences between the categories. Also, the variance of the asset categories for the benefits turns out to be significantly different. This indicates that not only the location, but also the accuracy of the benefit forecasts differs between the asset categories.

The tests for the asset usage cost f/r ratios show heterogeneity in location, but not for the variation. The f/a ratios of project cost show no heterogeneity for the asset classifications. Moreover, we do not find a statistical difference of the

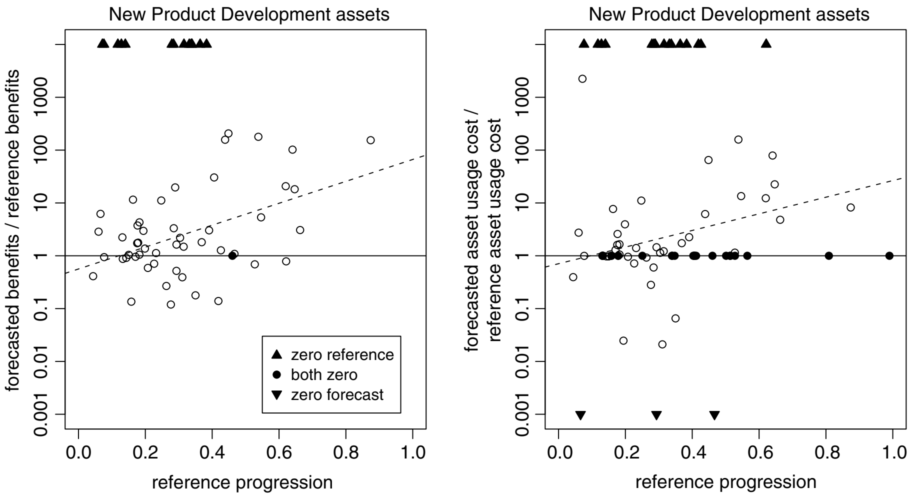

Fig. 12. No correlation between f/r ratios of benefits and asset usage cost and the relative time of re-estimation.

variance of the f/a ratios of project cost. In this case study, the forecasting quality of the project execution is not dependent on the type of assets that is being made.

Based on these findings, we conclude that the data of the forecasted NPVs, benefits and asset usage cost are heterogeneous for the asset classifications Cost Reduction and New Product Development. Therefore, we have to analyze the forecasting quality of the NPV, benefits and asset usage cost for each of these classifications separately.

When analyzing the forecasting quality of project cost, there is no necessity to assess the asset classifications separately. However, we will still do so, to make the different assessments analogue. Also, due to the low number of Sales Expansion assets, in the remainder of this article, we focus on the forecasting quality of Cost Reduction and New Product Development assets.

5.4.2. Re-estimation date

In Section 5.2, we discussed the re-estimation dates in the data set and found that the data has a large spread in the relative time of the re-estimations. Therefore, here we assess whether this spread results in heterogeneity of the data. For instance, it is possible that earlier re-estimations result in more accurate ratios as things may not have changed as much. We investigate whether the $(f - r)/s$ ratio of the NPV and the f/r ratios of benefits and asset usage cost are influenced by the timing of the re-estimations. Note, that we do not assess the f/a ratios of project cost, since these are actuals and not re-estimations.

To investigate for correlations, we plotted the ratios against the relative time and statistically confirmed the findings using Kendall's correlation coefficient test. Note that we performed the analyses for the combined asset classifications and the classifications separately. Most of these analyses show no correlations for the NPV, benefits and asset usage cost and the relative time of the re-estimations.

In two cases, Kendall's correlation coefficient indicated significant correlations. However, these correlations are influenced by the asymptotic behavior of numerous f/r ratios in the data set. Let us explain. We illustrate the two analyses in Fig. 12. The left-hand plot depicts on the vertical axis the f/r ratios of the benefit forecasts for New Product Development assets. The vertical axis of the right-hand plot shows the f/r ratios for the asset usage cost for New Product Development assets. The horizontal axes represent the relative time of re-estimation. The dashed diagonal line is the potential correlation of the data. This line is computed without taking the asymptotic cases of zero forecasts or zero reference points into account. Namely, the correlation test is unable to incorporate data points with value infinite.

Kendall's correlation coefficient test gives a p-value of 0.033 for the benefits and 0.015 for the asset usage cost. This indicates that the correlation is statistically significant given the threshold $\alpha = 0.05$.

However, the plots also illustrate numerous asymptotic f/r ratios that are not taken into account in the test. For instance, in the left-hand plot we find many assets that have no benefits in their re-estimations. In the right-hand plot, we also find numerous assets that both have a zero forecast and reference point. For instance, a number of IT projects were completed successfully, yet the assets they created no longer had any business value. These asymptotic f/r ratios highly influence our test for correlation.

Based on the analyses and accounting for the asymptotic cases, we concluded that there is no conclusive evidence for significant correlation between the forecasting quality of the NPV, benefits and asset usage cost, and the relative time at which the re-estimation is made. These analyses once more underscore the relevance of depicting ratios with asymptotic behavior when assessing forecasting quality.

Table 7
An overview of forecasted and re-estimated NPVs in millions of Euro.

	Cost reduction	New Product Development
Number of projects	35	62
Total initial forecasted NPV	194	2426
Total re-estimated NPV	179	1606
Median initial forecasted NPV	2.0	11
Median re-estimated NPV	2.3	1.5

5.4.3. Yearly trend

Besides the asset classification and re-estimation date, we also analyzed the data for a potential trend of the forecasting quality. Since we obtained data over 9 years, we analyzed whether forecasts of older assets had a significantly different forecasting quality than forecasts of newer projects. We do note that the available data for each year is relatively small. The figures and tests give no indication there exists a significant time-dependence in the forecasting quality for either the Cost Reduction projects or the New Product Development projects.

5.5. Summary

We found the data to be heterogeneous with respect to the asset classification. For Cost Reduction and New Product Development projects, the forecasting quality of the NPVs and its components, the benefits and asset usage cost are significantly different. Although the project cost showed no significant difference, to remain consistent, we will assess their forecasting quality separately for each asset classifications.

There were too few Sales Expansion assets to find differences with this category. Therefore, in the remainder of our analyses, we do not consider Sales Expansion assets.

The investigation of other variables did not result in further heterogeneity. With this knowledge we are able to commence with the analysis of the forecasting quality of the NPV, benefits, asset usage cost and project cost in the next section.

6. Case study

In the previous sections, we developed a generalized method and performed a prerequisite data analysis on the data that we obtained from organization Z. In this section, we will use the generalized method to assess the forecasting quality of the NPV, benefits, asset usage cost and project cost. Due to the heterogeneity caused by the asset classification, we consider Cost Reduction and New Product Development assets separately. Moreover, we leave the five Sales Expansion assets out of the analyses.

First, we investigate the forecasting quality of the NPV. Recall that the Net Present Value is a summation of the predicted monetary benefits and cost of a project discounted to current value. For this purpose, we use the $(f-r)/s$ ratio, the reference cone and the EQF_s. Recall that the EQF is a measure of the deviation between forecast and actual. As approximation of the asset size s, we use the first made forecast of the asset cost, or $s = f_A$.

Next to that, we assess the forecasting quality of the components of the NPV. We will analyze the f/r ratios for the benefits and asset usage cost, and the f/a ratios for the project cost.

Finally, we perform a sensitivity analysis to investigate the influence of the components on the forecasting quality of the NPV. We assess what the impact is on the forecasting accuracy of the NPV forecasts, when the forecasting quality of each of the components changes.

6.1. NPV

First, we investigate the forecasting quality of the NPV. In Table 7, we provide an overview of the forecasted and re-estimated NPVs of the 97 IT assets.

The table shows that the total initial forecasted NPVs in the Cost Reduction category are $194/179 \approx 1.1$ times higher than its re-estimated NPVs. For the New Product Development assets, the forecasts are $2426/1606 \approx 1.5$ times higher. The medians of the initial forecasts and the re-estimated NPVs show a larger difference. The median forecasted NPV of the Cost Reduction assets is slightly lower, namely $2.0/2.3 \approx 0.87$ times, than its median reference point. Yet, for the New Product Development assets, it is $11/1.5 \approx 7.3$ times higher. At first sight, these statistics indicate that the forecasted NPVs of the New Product Development assets follow an optimistic pattern. The Cost Reduction assets appear unbiased.

To further investigate the initial forecasted NPVs, we have drawn the $(f - r)/s$ plot with $s = f_A$ in Fig. 13. The reference cone that is drawn, is given by Formulas (11) and (12) and the EQF_s is as defined by Formula (10).

Let us explain the figure. The left-hand plot shows the forecasting quality of the Cost Reduction assets and the right-hand plot that of the New Product Development assets. The horizontal axis of each figure depicts the reference progression. For instance, suppose a re-estimation was made a year after the start of the asset. And suppose the initial forecast was made

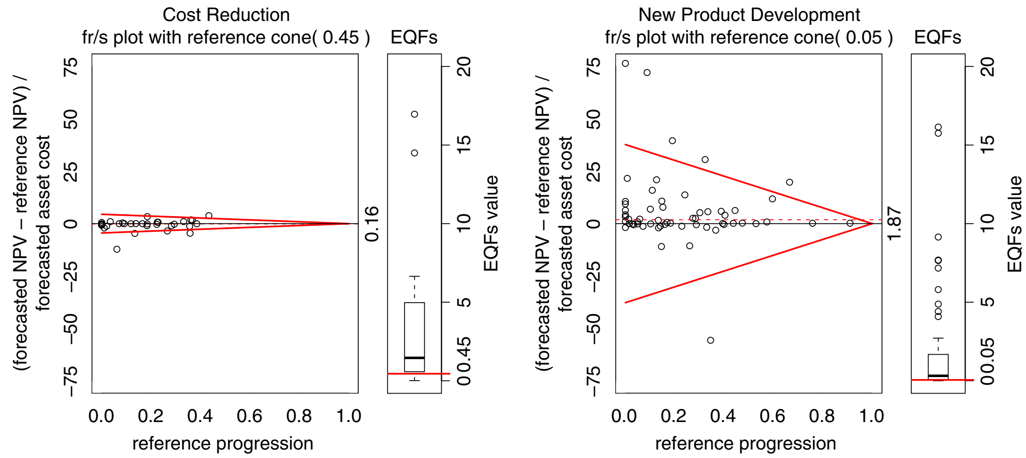

Fig. 13. $(f - r)/s$ plot, with $s = f_A$, reference cone and EQF_s box plot of the forecasted NPVs for Cost Reduction and New Product Development assets.

four months after the start. This forecast would be depicted at $4/12 = 0.3$ or 30% reference progression. The vertical axis displays the $(f - r)/s$ ratios.

The red lines in the figures are the reference lines drawn using Formulas (11) and (12). The EQF_s value that we chose is the 20% quantile of all EQF values. This value indicates the quality that 80% of the assets is able to obtain. The EQF_s box plot gives the EQF quality of the forecasts made.

Finally, the dashed horizontal lines in the figures represent the median $(f - r)/s$ ratios of the data. This is also indicated by the value to the right of each plot. For instance, in the left-hand plot, the median $(f - r)/s$ ratio is -0.16.

Most notable, the figure illustrates the difference in accuracy of the forecasted NPVs of the asset categories. The left-hand plot of the Cost Reduction assets show less spread in their $(f - r)/s$ ratios than the New Product Development assets. This is confirmed by the median EQF_s of 1.46 for the former category and 0.32 for the latter. The data shows that the organization is able to more accurately predict Cost Reduction assets than New Product Development assets.

For the Cost Reduction assets, we find a small pessimistic bias indicated by the figure and the median $(f - r)/s$ ratio of -0.16. On the other hand, the New Product Development assets show an optimistic bias in the figure combined with a median $(f - r)/s$ ratio of 1.87. Note, although it does not appear so in the figure, this is a relevant bias. For instance, suppose a new project proposal predicts an NPV of 3.5 million Euro with asset cost of 2 million Euro. The bias indicates that this forecast is possibly overestimated by $1.87 \cdot 2 = 3.74$ million Euro. A relevant deviation indeed, as this indicates the project proposal is likely to result in an NPV of $3.5 - 3.74 = -0.24$ million Euro in this example.

The above analyses show that the organization is able to predict NPVs of Cost Reduction assets almost without bias. Moreover, the accuracy of the Cost Reduction forecasts is better than the forecasts of New Product Development assets. The latter forecasts show an optimistic bias.

6.2. Benefits

One of the components of the NPV is the cumulated discounted benefits denoted by B in Formula (5). We analyze the forecasting quality of these benefits using the f/r plot, reference cone and EQF_r depicted in Fig. 14. These plots are similar to those we showed for the NPV. However, the reference lines are drawn using Formulas (7) and (8). Note, that for this lower reference line, a minimum EQF_r value of 2 is required. Therefore, in the plots we use $EQF_r = 2$ for the lower reference lines.

Moreover, the plots also visualize f/r ratios in asymptotic cases, which were not present for the NPV. The f/r ratios described by reference benefits of zero are shown as solid upper triangles. They represent ratios that are infinite, since the forecast is positive and the reference value is zero. If both forecast and reference are zero, the f/r ratio is defined by $0/0 = 1$ and is depicted with a solid dot. Finally, in case the forecast is zero, the f/r ratio is zero, indicating the forecast is underestimated, and is visualized by a downward solid triangle.

The left-hand f/r plot of the Cost Reduction assets shows no clear bias, similar as for the NPV. The median f/a ratio of the data reveals only a slight bias toward underestimation with a value of 0.93. The forecasts have a median EQF_r of 3.85 indicating that they have an average time-weighted deviation of $1/3.85 \approx 26\%$ to their reference point. Note, that in this case, we only have the initial forecast, thus the median deviation to the reference point is also 26%. It is not possible to compare this forecasting accuracy with that of the NPV, since the EQF_s and EQF_r are different. Moreover, we are unaware of other benchmarks to compare these EQF_r values with.

The right-hand plot of the New Product Development shows a larger spread of the f/r ratios than that of the Cost Reduction assets. This is reflected in the forecasting accuracy in terms of EQF_r. With a median EQF_r value of 0.50, the forecasts have a median time-weighted average deviation of $1/0.5 = 200\%$ from their reference point.

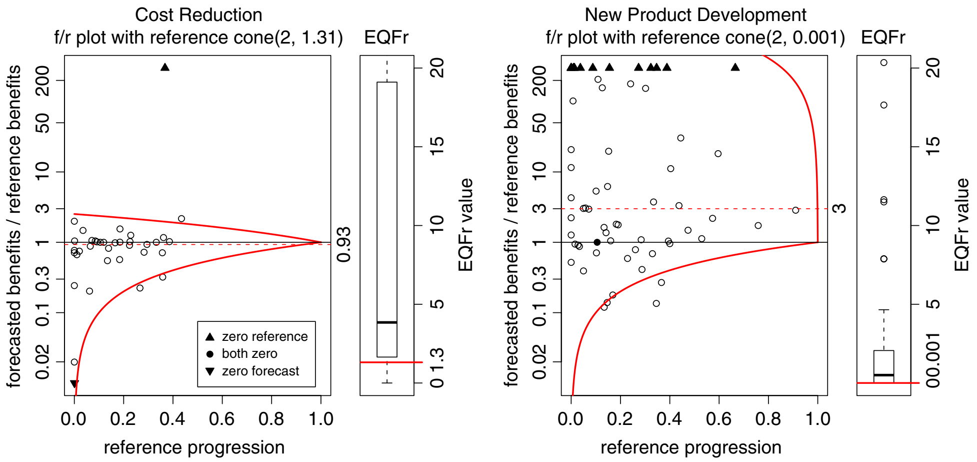

Fig. 14. f/r plot with reference cone and EQF_r box plot of the forecasted benefits for Cost Reduction and New Product Development assets.

The forecasts have an optimistic bias, which is illustrated by the figure and a median f/r ratio of 3. That is, the median of the initial forecasted benefits of New Product Development assets is three times as high as the median of their reference point.

Partly, this is caused by the fact that for a large number of assets the re-estimated benefits are zero. A number of IT projects were canceled or completed successfully, yet the assets they created no longer had any business value. For instance, it is possible projects were stopped during implementation due to changes in the market. The benefits of these assets are simple to recalculate, namely the benefits are zero in the post calculation. They comprise a relative large portion of the re-estimated assets. This confirms that forecasts and actuals with value zero are more common than you would expect.

We do note, that if a project was stopped, for instance, due to changes in the market place, in our analyses the corresponding forecasts may be considered inaccurate, even though the accuracy of the forecasts is not the reason for stopping the project. It can be considered unfair to hold the estimators accountable for an inaccurate forecast in this case. Therefore, an organization should carefully consider in these cases whether to judge the forecasts differently or not.

The results of this analysis compare well with the findings for the NPV. The Cost Reduction assets show no particular bias and are more accurately predicted than the New Product Development assets. The quality of the New Product Development forecasts is relatively low and shows a bias toward overestimation.

Note that poor benefit forecasts do not by definition imply forecasting errors of the business domain estimators or the IT domain estimators. The method indicates which forecasts were less accurate than others, yet do not answer the question what the reason is for these inaccuracies. For instance, an inaccurate forecast of the benefits may be caused by an underestimation or overestimation of the sales generated using the asset. Or, no benefits are generated because the asset could not be developed by the IT department. Therefore, inaccurate forecasts have different causes. The method only indicates which forecasts were inaccurate. To determine the cause of the inaccuracies, additional analyses are required.

6.3. Asset usage cost

Another component of the NPV is the cumulated discounted asset usage cost. In Formula (5) we denoted these usage cost by C. Fig. 15 shows the f/r plot, reference cone and EQF_r box plot of their forecasts for both the categories Cost Reduction and New Product Development.

The plots show a large number of f/r ratios for both categories, of which the forecast and reference points are zero. Again, we also find several assets with re-estimated zero reference points. For instance, assets that were stopped during implementation. In these cases, no asset usage cost needed to be made, making the re-estimations zero.

The plots show that the forecasts of the Cost Reduction asset have no bias which is supported by a median f/r ratio of 1. The forecasting quality with a median EQF_r value of 4.6 is higher than that of the benefits forecasts.

Similar to the benefits, the asset usage cost forecasts of the New Product Development assets show a bias toward overestimation in the figure and a median f/r ratio of 1.99. The similarity of the bias is partly caused by the relation between the asset usage cost and the benefits. The asset usage cost comprise of direct cost. These direct cost are cost that are incurred in proportion to the activity of the business and can be associated to a particular cost object. For example, if a telephone contract is sold, benefits are generated, yet also direct cost are incurred for, for instance, handset subsidies and interconnect cost. Due to these direct cost, a relation exists between the asset usage cost and the benefits. Therefore, it makes sense to have similar bias and forecasting quality for these components.

The quality of the New Product Development forecasts remains low with a median EQF_r value of 0.46. For this category, there is hardly any difference in accuracy between the benefits and asset usage cost.

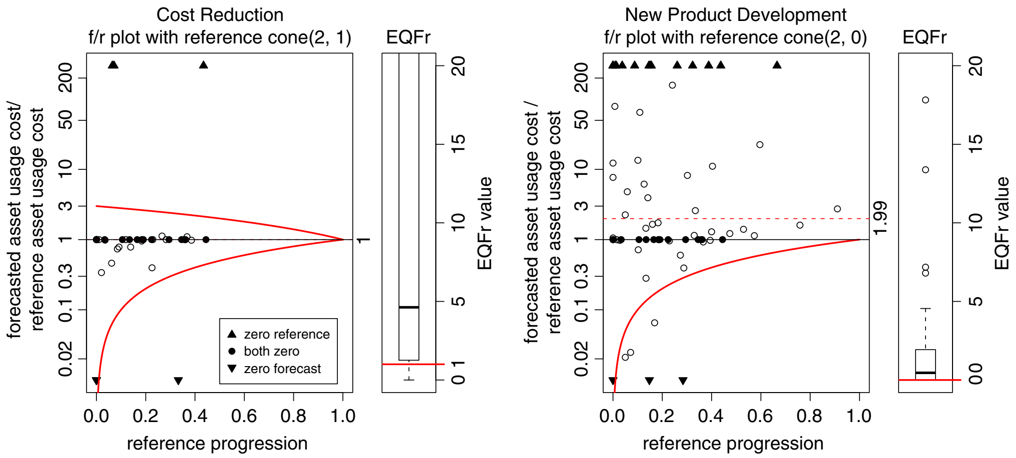

Fig. 15. f/r plot with reference cone and EQF_r box plot of the forecasted asset usage cost for Cost Reduction and New Product Development assets.

Summarizing, the analyses of the asset usage cost show similar results as found for the benefit forecasts. The assets of the Cost Reduction category have unbiased forecasts that are of reasonable quality. The New Product Development assets on the other hand have an optimistic bias and have a larger deviation of the forecasts to the reference points.

6.4. Project cost

Finally, we investigate the forecasting quality of the cumulated discounted project cost denoted by I in Formula (5). Fig. 16 depicts the f/a plot, reference cones and EQF box plots for both asset classifications. Note that for the project cost, we are able to use the actuals as reference point. Therefore, the horizontal axes of the f/a plots depict the project progression instead of reference progression.

The f/a plot of the Cost Reduction assets has a median f/a ratio of 1.05 indicating no significant bias in the forecasts. The forecasting accuracy shows a median EQF value of 10.3. This means the average time-weighted deviation is only 9.7% from the actual.

The f/a plot of the New Product Development assets shows a wider scatter of the f/a ratios than the f/a plot of the Cost Reduction assets. With a median EQF value of 2.0 the quality of these forecasts is significantly less than that of the Cost Reduction assets. The median deviation of the forecast to the actual is 50%. Still, the forecasts are unbiased as shown in the figure and the median f/a ratio of 1.

The forecasts of the project cost have a higher accuracy than those of the benefits and asset usage cost for both categories. However, while the forecasts of the benefits and asset usage cost for New Product Development showed an optimistic bias, we find no bias for the project cost forecasts. Apparently, the organization has found ways to adequately predict the project cost, yet has more difficulties assessing the benefits and asset usage cost.

6.4.0.1. Benchmark. Since we analyzed f/a ratios for the project cost, we are able to compare the quality of these forecasts against the forecasting quality of other organizations described in the literature. We will compare our case study with different benchmarks in the literature.

The first comparison is to another study [14] that gives a literature overview of forecasting quality in terms of EQF. For the sake of completeness, in Table 8 we reiterate that overview.

The benchmark figures in the overview concern only completed projects. However, in the case study of organization Z, there are also canceled projects. To make a fair comparison, we therefore removed the assets that were stopped.

Note, that in case of canceled projects, it may be beneficial to have very inaccurate forecasts. Namely, the sooner the project is stopped, the less money may have be spent and the more the forecast will deviate from the actual. Although stopping a project sooner rather than later results in less accurate forecasts, the organization will have lost less money.

We added the results of organization Z to the table. The label *Organization Z - NPD* shows the project cost forecasting quality of the New Product Developments assets and the label *Organization Z - CR* for the Cost Reduction assets. However, the case studies of the literature make no distinction in the asset classification. Therefore, we also added the combined quality of both classifications. This is labeled *Organization Z - combined* for the quality of the combined categories. The table is sorted by the median EQF value. It is not possible to use this table for comparisons with the EQF_r or EQF_s.

Note that in our case study, we analyze cumulated discounted project cost. The data in the literature does not contain discounted figures. However, discounting does not play a significant role for the results of organization Z, as the project duration is often smaller than a year. Therefore, it is possible to compare them to the literature data.

Also note that we are unable to give the mean EQF value for the different categories. This is caused by projects with both a forecast and an actual of zero. We defined the f/a ratio to be 1 in this case. Therefore, the forecast has an infinite EQF

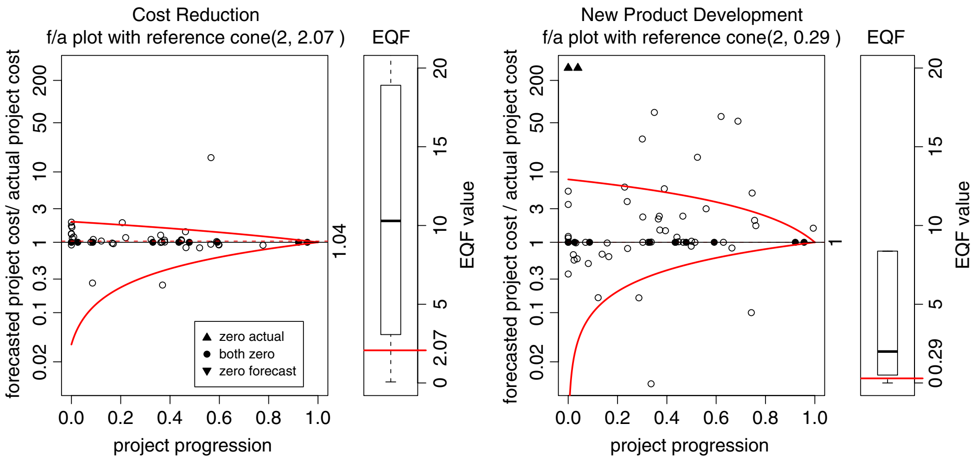

Fig. 16. f/a plot with reference cone and EQF box plot of the forecasted project cost for Cost Reduction and New Product Development assets.

Table 8
Summary of benchmark EQF values found in the literature and our case study.

Source	Median EQF	Average EQF	Number of projects
Organization Z - CR	10.3	–	34
Kulk et al. [33]	9.4	–	221
Organization Y cost	8.5	36.9	140
Landmark Graphics 2003 [38]	8.4	–	–
Landmark Graphics 2002 [38]	7.6	–	–
Landmark Graphics 2001 [38]	7.0	–	–
Organization Y functionality	6.4	9.9	83
Organization Z - combined	6.0	–	83
Lister [37]	4–9	–	–
Organization Z - NPD	4.7	–	49
Landmark Graphics 1999 [39]	4.7	6.3	121
DeMarco–Little [39]	1.9	4.2	20
DeMarco [11]	–	3.8	–
Organization X	0.43	1.6	867

value. Moreover, in our case study of organization Z, only the initial forecasts were available. Therefore, the EQF values of organization Z, are most likely lower than they could have been if multiple forecasts were made.

The table shows that the forecasts of project cost for the New Product Development category are of reasonable quality compared to other benchmarks. The forecasts of the Cost Reduction assets are of higher quality than the forecasts of other organizations in the overview. Especially considering the fact that we only account for initial forecasts, the quality of these forecasts are relatively high. We find the quality of the forecasts of the combined classifications comparable to the quality found by Lister and Organization Y functionality. This indicates that the telecommunication organization Z has reasonable forecasting quality of the project cost compared to the literature.

Another comparison that we make is with data from the ISBSG, or International Software Benchmarking Standards Group. The ISBSG collects benchmarking data from organizations in many countries. As of 2011, the data contains among others sizing information, cost and duration information for more than 5500 projects [22]. For 117 projects the data repository also contains both forecasted and actual cost.

In a book [22] an overview is given of the accuracy of these forecasts. In Table 9 these benchmark figures are compared with the forecasting accuracy of our case study. For the comparison, we only use the initial cost forecasts made of our case study. Since the ISBSG data contains different project types, we will compare the results with the combined projects in our case study. Moreover, the ISBSG data concerns completed projects, therefore, we again removed the canceled projects in the case study.

The table shows that our case study resembles the project cost forecasting accuracies reported by the ISBSG. The book [22] notes that the ISBSG repository is believed to reflect the best 25% of the industry. This implies that the project cost forecasts made in our case study are of high quality.

However, the possible comparisons between the figures reported by the ISBSG and our case study are limited. As we argue in another article [14], biases can significantly influence the accuracy of forecasts and can differ from organization to organization. The database of the ISBSG contains projects from many organizations, of which it is unknown what their organizational bias is. Without accounting for these biases, summarizing forecasting accuracies of different organizations does not result in meaningful benchmarks. Therefore, we restricted comparisons to the one we discussed above.

Table 9
Summary of benchmark values of the ISBSG repository [22] and our case study.

	ISBSG	Organization Z
Number of projects	117	87
Forecasts with EQF > 10	44%	46%
Forecasts with EQF > 5	64%	57%

Table 10
Overview of the total and median reference values in millions of Euros for the components of the NPV.

Reference	Benefits	Asset usage cost	Project cost
Total CR	380	129	72
Median CR	3.06	0.05	0.52
Total NPD	4104	2314	169
Median NPD	6.84	2.04	0.58

6.5. Sensitivity analysis

In the above analyses, we investigated the forecasting quality of the NPV, benefits, asset usage cost and project cost for organization Z. We found that the forecasting quality of the project cost for Cost Reduction assets is relatively high. However, that quality is not reflected in the equivalent forecasting quality of the NPV.

The reason for this is, that assets do not consist of an equal amount of benefits, asset usage cost and project cost. This is illustrated in Table 10, which depicts the total and median values of these components of the NPV. The table shows that the benefits for both asset classifications are more than twice as large as the asset usage cost and project cost.

Due to the differences in size, the impact on the quality of the NPV forecasts is not the same for each of the components. For instance, making highly accurate project cost forecasts hardly impacts the quality of the NPV, as it only forms a small part of the entire value. In fact, it is even possible that the overall quality of the NPV degrades when the accuracy of the project cost is increased.

To illustrate this, we perform a sensitivity analysis. We change the accuracy of the forecasts of one component and compute the effect on the forecasting quality of the NPV in terms of EQF_s. More precisely, we increase the accuracy of the benefits, asset usage cost and project cost forecasts in such a way that their EQF_r or EQF quality is increased by 10%. For instance, if the forecasted benefits of an asset have an EQF_r quality of 4, we adjust the forecast in such a way that its EQF_r quality becomes 4.4.

Let us explain precisely how to derive at these adjusted forecasts. Recall, that the EQF_r is defined in Formula (6) in Section 4.2 as the area under the reference value divided by area difference between forecast and reference point. Or mathematically, $\text{EQF}_r = \frac{\int_{t_s}^{t_r} 1\, dt}{\int_{t_s}^{t_r} |1-e(t)/r|\, dt}$. To increase the readability of the following calculations, we first rewrite this formula. Using $x = \frac{t-t_s}{t_r-t_s}$ and $f = e(t)$, we find $\text{EQF}_r = \frac{\int_0^1 r\, dx}{\int_0^1 |r-f|\, dx} = \frac{|r|}{|r-f|}$.

We wish to increase the EQF_r by 10%, or similarly the original EQF_r is multiplied by 1.1. Since the reference point does not change, this means we need to decrease the area between the forecast and the reference point. We are able to decrease the area by adjusting the forecast f. We denote the adjusted forecast by f'. The EQF_r based on the adjusted forecast should be equal to 1.1 times the original EQF_r value. To derive at the adjusted forecast f', we solve the following equation.

$$1.1 \cdot \frac{|r|}{|r-f|} = \frac{|r|}{|r-f'|}$$
$$1.1 \cdot |r| \cdot |r-f'| = |r-f| \cdot |r|$$
$$|1.1 \cdot r - 1.1 \cdot f'| = |r-f|$$
$$|-1.1 \cdot f'| = |-1.1 \cdot r + r - f|$$
$$|f'| = |r - \frac{1}{1.1} r + \frac{1}{1.1} f|$$
$$f' = \frac{0.1}{1.1} \cdot r + \frac{1}{1.1} \cdot f$$
$$f' = f + \frac{0.1}{1.1}(r-f).$$

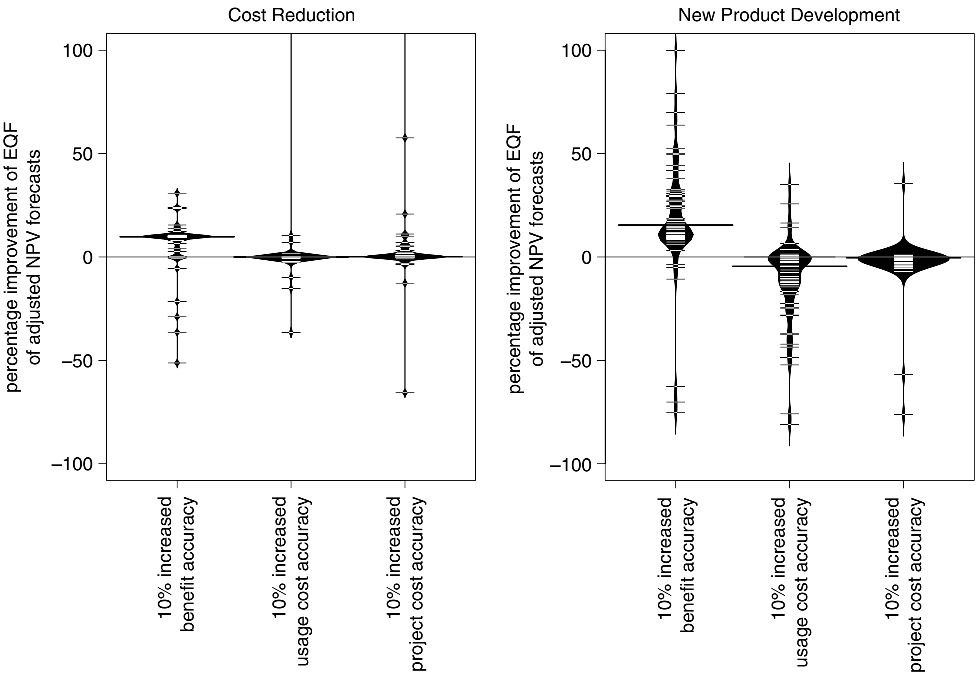

Fig. 17. The percentage increase of the EQF of the NPV forecasts when we improve either the benefit, usage cost or project cost forecast quality by 10%.

Table 11
Summary of the improvements of the forecasting quality of the NPV when each of its components increases their forecasting quality by 10%.

Component	% of assets with improved EQF_s quality	Median increase EQF_s (%)
Benefits CR	77	9.8
Asset usage cost CR	26	0.0
Project cost CR	63	0.19
Benefits NPD	87	15
Asset usage cost NPD	19	−4.5
Project cost NPD	10	−0.38

In this way, we are able to compute the adjusted forecasts for benefits, asset usage cost and project cost. Note that these computations are the same in case the reference point is the actual.

With these adjusted forecasts, we subsequently recompute the forecasted NPV. For instance, suppose the original NPV was computed using $\mathrm{NPV} = f_{\mathrm{B}} - f_{\mathrm{C}} - f_{\mathrm{I}}$ and we adjusted the forecasts of the benefits f'_{B}. Then, the adjusted NPV becomes $NPV' = f'_{\mathrm{B}} - f_{\mathrm{C}} - f_{\mathrm{I}}$. By replacing the original forecast of either the benefits, asset usage cost or project cost by their adjusted forecast, we find the recomputed NPV′ forecast. For the NPV′, we again compute the corresponding EQF'_s value.

For the adjusted NPV′ forecasts, we observe the percentage difference in EQF_s between the original and adjusted situation. We do this by computing $100\% \cdot (\mathrm{EQF}'_s - \mathrm{EQF}_s)/\mathrm{EQF}_s$ with EQF_s the original value of the NPV forecasts and EQF'_s the adjusted value.

In Fig. 17, we depict the percentage increase of the EQF_s of the NPV forecasts when we improve the forecasts of benefits, asset usage cost and project cost. In Table 11, we summarize these results.

For the Cost Reduction assets, we find that improving the benefits, results in a median improvement of 9.8% of the forecasting quality of the NPV. Indeed, in 77% of the Cost Reduction assets, an improvement of the benefit forecasts results in an improved NPV forecasting quality.

Note, that it is possible that the forecasting quality of the NPV decreases while improving other components. For example, consider an asset with forecasted benefits of 100 Euro and a forecasted asset cost of 95 Euro, resulting in an NPV of 5 Euro. Suppose the actual benefits turned out to be 25 Euro and the actual asset cost 19 Euro leading to an NPV of 6 Euro. That is, both the benefits and asset cost are overestimated. If we improve the benefits by 10% this means that the initial forecasted benefits of 100 Euro becomes $100 - (1 - 1/1.1) * (100 - 25) \approx 93$ Euro. Thus, the adjusted NPV forecast is $93 - 95 = -2$ Euro instead of the previous 5 Euro. This adjusted NPV forecast is less accurate than the original NPV forecast, even though the forecasting quality of the benefits increased.

Interestingly, when we improve the forecasting quality of asset usage cost or project cost for Cost Reduction assets, the NPV quality remains the same with 0.0% and 0.19%. The reason that the improved accuracy of the asset usage cost and project is hardly noticeable for the NPV forecasts is the relative size of these elements. The median of the reference benefits of the Cost Reduction assets are five times larger than the median of the combined asset cost. Therefore, a 10% improvement of only a small part does not significantly affect the overall quality. Moreover, there are numerous assets with zero forecast and actual. In these cases, the forecasts were perfect and cannot be further improved.

The plots of the New Product Development assets show that the forecasting quality of the NPV improves most by an increase of the accuracy of the benefit forecasts. For 87% of the assets, improved benefits results in an increased accuracy of the NPV. Moreover, improving the forecasting accuracy of the benefits by 10%, yields a median increase of 15% of the EQF_s for the NPV.

On the other hand, an increased accuracy of the asset usage cost and project cost forecasts leads to a *decrease* in the forecasting accuracy of the NPV. For 81% of the assets the NPV forecast accuracy decreases when the assets usage cost forecasts are improved, and 90% for the project cost. For the project cost, this decrease is hardly noticeable. However, for the asset usage cost, the forecasting quality of the NPV decreases a median 4.5%.

The reason for this decrease in forecasting quality of the NPV, are the biases we found for the benefit and asset usage cost of the New Product Development assets. Both these forecasts have an optimistic bias. Therefore, the overestimation of the benefits is balanced by the overestimation of the asset usage cost. By increasing the accuracy of the latter, we diminish the reduction of the overestimation of the former. Since the benefits are in general larger than the asset usage cost, it is more important for the NPV forecasting quality to remain overestimating the asset usage cost to compensate for the overestimated benefits.

These results support the idea [43,20,48,23] that CIO's must manage IT assets to maximize business value, instead of controlling their cost. The analysis shows that focus on control of the project cost or asset usage cost may not be beneficial to the forecasting quality of the NPV. In this case study, stressing the forecasting accuracy of the cost could even decrease the forecasting quality of the NPV. As the benefits often form the largest part of project proposals, IT governors should focus primarily on the accurate forecasting of the benefits and ultimately the NPV.

6.6. Case summary

In this section, we assessed the forecasting quality for the telecommunication organization Z. We found that Cost Reduction assets have no biases for the benefit, asset usage cost and project cost forecasts. However, the New Product Development forecasts showed optimistic biases for the benefits and asset usage cost forecasts. Moreover, the New Product Development forecasts were of lower quality in terms of EQF than the Cost Reduction forecasts. Compared to benchmarks from the literature, the forecasting quality of the project cost is reasonable.

Finally, we performed a sensitivity analysis to determine the impact on the forecasting quality of the NPV of the different components. The analyses showed that an improvement of the benefit forecasts results in the highest improvement of the NPV forecasting quality. Interestingly, if the quality of the asset usage cost forecasts improves, it results in a decreased forecasting accuracy of the NPV. These results support the idea that CIO's must manage IT assets to maximize business value, instead of controlling their cost.

7. Enhancing decision information

In the previous section, we illustrated how the management is able to investigate the forecasting quality of the IT business value. With the analyses, executives are able to find biases, and possible change the estimation process to increase the forecasting accuracy. However, the information of the analyses is also valuable to assist the management in their decision making.

In an article by Eveleens et al. [14], methods are discussed to use the information of the forecast quality and their bias to enhance decision making. With these methods, IT governors gain information on the accuracy of the forecasts and are able to make risk assessments on project proposals.

This is achieved, for instance, by using a confidence interval or an empirical distribution of f/a ratios. For instance, suppose a project proposal predicts the project cost to be 2 million Euro. Based on the forecasting quality of historically forecasted project costs, an approximate interval is determined that shows that the f/a ratios in 80% of the cases have a range of $[0.5, 1.25]$.

The IT governor is able to use this information to derive a likely range of possible project cost for new project proposal. This is done in the following way. The likely range is found by dividing the initial forecast by the bounds of the interval. This leads to $2/1.25 = 1.6$ and $2/0.5 = 4$. Thus, the likely range of possible project cost of the new project proposal is $[1.6, 4]$ million Euro. This additional information allows the executive to determine whether the range of possible outcomes is acceptable to fund the project. Note, that this method will also work for the ratios that we described in this article.

The methods to approximate forecast intervals are only necessary when the executives are given point forecasts. As discussed in Section 2, it is more interesting for IT governors to immediately receive an interval forecast. Estimators already consider multiple scenarios that may occur, and are therefore able to give a more accurate project-dependent interval. This should allow for better information than these approximated intervals.

In this section, we illustrate how to use forecasts to further enhance decision making. We will do so by describing two simulation examples. In these examples, we will use the data of our real-world case study of organization Z. The examples provide information for questions, such as: if a certain number of project proposals would be executed, what is the likelihood that the available budget is sufficient to finance them? Or, given a set of proposals, what is a likely range of business value that will be generated?

An important assumption of any simulation, and of any method that uses historical data to make predictions of the future, is that the future will behave similarly as in the past. In the examples we will describe, we will assume that the projects and assets are executed under similar circumstances as before. If this assumption does not hold in a certain situation, the outcomes of a simulation should be judged carefully.

Note, that even though we assume the projects will behave similar to the past, it is always possible a result occurs that was not simulated. Namely, it is always possible that the assumption does not hold as initially anticipated. Still, a simulation will provide for useful information.

Moreover, note that the examples are not meant to discuss in detail how to construct such simulation models. They are only meant to illustrate how it is possible to further utilize the information of the forecasting quality assessments to enhance decision information.

7.1. Rationing capital budget

In the first simulation example, we discuss how to utilize the quantified forecast information of forecasted project cost to assist with rationing the capital budget. Suppose that in organization Z, 30 of the 35 Cost Reduction projects and 58 of the 62 New Product Development projects have been completed and their actual project cost are known. For the remaining 5 newest Cost Reduction projects and 7 newest New Product Development, the project proposals have been submitted. These project proposals have a predicted project cost of 22.0 million Euro. However, the organization has determined that the maximum available budget for new project proposals is less, for instance, 20 million Euro.

The required predicted investment is larger than the budget the organization is willing to spend. Therefore, the IT governors are concerned with the question: How many and which project proposals can be executed, in such a way that the available budget is sufficient to finance them?

One simple solution to the question is that the IT governor selects the most promising project proposals that more or less sum up to the available budget. Yet, with the available information of the completed projects, the executive is able to do more, as we will demonstrate using a simulation model.

Note, that the complete answer to the question is not easily given. Namely, it depends on many factors, for instance, the strategic alignment of the proposals to the organization or the business values the proposals predict to yield.

In this example, we simply wish to illustrate the usefulness of information on the forecasting accuracy. Therefore, we simplify the problem by only considering the project cost and their chance of under- and overruns based on historic data. We will not consider other relevant factors necessary to fully answer the question.

Below, we give a basic description of the simulation model. Then, we discuss the results of the simulation and the information the management can derive from such models.

7.1.0.2. Creating the simulation. Using the historic information of completed projects, we are able to construct a simulation model. A simulation model virtually executes the 12 new project proposals under similar circumstances as the completed projects.

The procedure boils down to the following. We already have thrown with a dice 30 times for the Cost Reduction category and 58 times for the New Product Development projects. These are the projects that have already been completed. The outcomes of these throws determine the probability distribution of the dice. With the probability distribution of the dice, we are able to fictively throw with the same dice another 5 and 7 times, once for each new project proposal.

By repeatedly throwing the dice, each time we will have a different outcome of the throws. All these different outcomes will provide information about the possible variation and the likelihood of particular outcomes. Note, that since we throw with two different dice with different distributions, the combined distribution of the outcomes is not trivial to compute.

The main assumption of simulation models is that the new projects are performed similarly to those that are already executed. Without this assumption, it is impossible to make any extrapolations or predictions of the future. In this particular case, the assumption is reasonable, since no major changes in the estimating process took place.

Moreover, we assume that the f/a ratios are not correlated with the projects. That is, there is no relation between the f/a ratios and any aspect of the projects. This assumption allows us to create a simple simulation by assuming the f/a ratios are randomly distributed for the projects.

To create the simulation, we compute for each completed project its f/a ratio. For the Cost Reduction proposals, we randomly draw 5 f/a ratios out of the completed Cost Reduction projects. This is done similarly for the New Product Developments. Using these f/a ratios, we are able to simulate the actual project cost for project proposal i using a randomly chosen completed project j, by calculating $u_i = f_i/(f_j/a_j)$. In this formula, f_i is the original forecasted project cost, f_j/a_j the f/a ratio of a completed project and u_i the simulated actual project cost. Summing these simulated outcomes computes to an aggregated simulated project cost for the combined new project proposals. This procedure is iterated 1000 times, resulting in a distribution of possible outcomes.

7.1.0.3. Simulation results. Fig. 18 illustrates the result of the above described simulation. The figure shows a cumulative distribution function of the simulated outcomes for the new project proposals. The horizontal axis depicts the total simulated

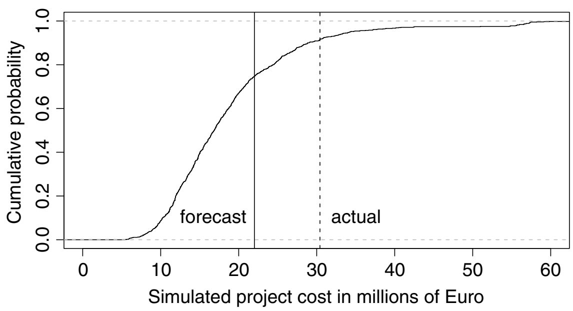

Fig. 18. Simulation results in the form of a cumulative distribution function of simulated project cost of 5 Cost Reduction project proposals and 7 New Product Development project proposals.

project cost for the new project proposals. For example, $x = 30$ indicates that the simulation predicts 30 million Euro is required to execute the project proposals.

The vertical axis depicts the cumulative chance a particular outcome occurred in the simulations. For instance, $x = 20$ has a value of $y = 0.67$. This means that in the simulations the project cost in 67% of the cases were lower or equal to 20 million Euro.

The vertical solid line represents the sum of the forecasted project cost of the new project proposals. Recall, that the forecasted project cost for the combined project proposals was 22.0 million Euro. The vertical dashed line is the actual required investment that was used when the projects were actually executed. This turned out to be 30.4 million Euro.

Using this figure, it becomes possible for an IT governor to make a risk assessment for the decision of which projects to invest in. The simulation shows that if all project proposals would be executed, there is a 67% chance the projects will cost less than the available budget of 20 million Euro. However, there is also a 33% chance more money is required.

IT governors are able to use the quantified information of the figure to enhance their decision making in many ways. For instance, depending on the risk averseness or appetite of the organization, the IT governor can decide to take the risk and commence with all projects. Or, for example, the executive can determine that a selection needs to be made among the proposals. For each selection, it is possible to create a new simulation to reassess the risk involved when executed that selection.

7.2. Asset business values

In the second simulation example, we illustrate how to enhance decision information when considering project proposals for their asset business value. Suppose, we have the same project proposals as in the previous example. That is, there are 5 Cost Reduction proposals and 7 New Product Development proposals. The predicted asset business value in terms of NPV for the project proposals is 201 million Euro. An IT governor may wonder what the chance is that the predicted value will actually be generated. It is possible to answer this question using another simulation.

7.2.0.4. Creating the simulation. As in the previous example, we use the historic forecasting quality of completed projects to make a prediction of the asset business value of the new project proposals. For the NPV, we need to use the $(f - r)/s$ ratio instead of the f/a ratio. In this case, we use the $(f - r)/s$ ratio with $s = f_A$, the forecasted asset cost.

For each Cost Reduction proposal, we randomly draw a $(f - r)/s$ ratio from the completed Cost Reduction projects. However, the $(f - r)/s$ ratios show a large variation and there are only a limited number of data points. This may cause the extremes of the ratios to have a higher probability of occurring than they have in reality. If we would randomly draw ratios including these extreme ratios, we will find our simulation to produce many unlikely outcomes. Therefore, to avoid this, we remove the upper and lower ten-percent quantiles from the completed projects. This leaves 80% of the completed projects.

With the randomly drawn ratios, we simulate the value generated by each project proposal. This is done for project proposal i using a randomly chosen completed project j, by calculating $u_i = f_i - s_i \cdot (f_j - r_j)/s_j$. In this formula, f_i is the originally forecasted NPV value, s_i the original forecasted asset cost, $(f_j - r_j)/s_j$ the $(f - r)/s$ ratio of a completed project and u_i the simulated actual NPV. Summing these simulated outcomes computes to an aggregated simulated NPV for the combined new project proposals. This procedure is iterated 1000 times.

7.2.0.5. Simulation of organization Z. The results of the simulations are shown in Fig. 19. The figure shows a cumulative distribution function of the outcomes of the simulation. The horizontal axis depicts the simulated total asset business value of the project proposals. The vertical axis shows the cumulative chance of the outcomes of the simulation. For instance, $x = 0$ corresponds with the value $y = 0.18$, which means that in 18% of the cases the simulation had an adjusted business value smaller or equal to 0.

The figure shows that even with an originally forecasted NPV of 201 million Euro, there is a 18% chance that this selection of proposals will be loss generating. Of course, it is unlikely to actually occur, as the management will undertake actions to prevent the projects to become loss-making. For instance, losses can be reduced by canceling a project, when it becomes clear the project will not yield the desired results.

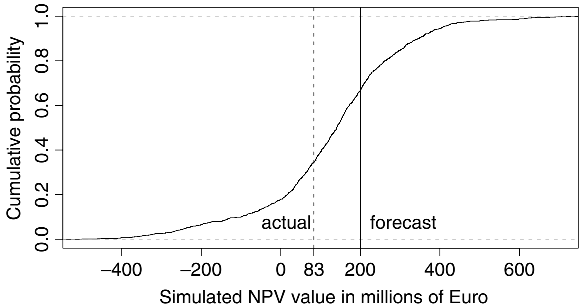

Fig. 19. Simulation results in the form of a cumulative distribution function of simulated NPVs of 5 Cost Reduction project proposals and 7 New Product Development project proposals.

The figure also shows that there is a 67% chance that the proposals will yield less than the predicted 201 million Euro. In this case, the actual business value generated by the proposals turned out to be 83 million Euro.

7.3. Summary

The two simulation examples illustrate that it is possible to enhance decision information using the quantified forecasting information. In the basic examples, we showed how, for instance, the information of the project cost and asset business value can be used. It is also possible to incorporate asset usage cost and benefit information in the simulations. These examples show that it is possible to assist IT governors in their decision making.

8. Discussion

In the previous sections, we discussed how to quantify the forecasting quality of IT business value and use it to enhance decision information. We covered many aspects in detail. The generalized method is not a new statistical way to assess forecasting quality. However, it summarizes, visualizes and quantifies the forecasting quality in such a way that it is accessible to executives. In this section, we finalize with a number of general remarks about the subject.

In this article we addressed quantifying the forecasting quality from the perspective of accuracy. For instance, we argued that increasing the accuracy of the forecasts is beneficial to the organization. Namely, more accurate forecast information will allow for better decision making. Therefore, removing a bias to increase the forecasting accuracy is beneficial.

However, political biases can have their purpose. For instance, suppose we would ask participants of a tennis tournament up front to raise their hand if they think they are going to win. It is good for the play in the tournament if more than 1 participant will raise their hand. Of course, we know that these forecasts are inaccurate and biased; only one player can actually win. The forecasts show the ambition of the players and their motivation.

Business cases can also be used to communicate ambition. These ambitions should not be killed at the expense of increased forecasting accuracy. IT governors should avoid that, by increasing the forecasting accuracy, the organization becomes risk-averse. However, it remains important for executives to quantify the forecasting accuracy and potential bias to account for the quality in their decision making.

Also, in the article we discussed the quality of *forecasts*, not of the estimators per se. Forecasts can be made by several estimators. For instance, one estimator predicts the business domain and another forecasts the IT domain. If a forecast turns out to be less accurate than desired, this may, for instance, be caused by differences in the business domain. In this case, the IT estimator may have made accurate forecasts, yet the overall forecast is still inaccurate. In these situations, caution is needed when assessing the quality of the forecasts. The forecasting accuracy does not necessarily reflect on a particular estimator.

Finally, in this article, we discussed it is important to search for heterogeneity in the data. The number of factors at our disposal in the data sets were limited. In the case study of organization Z, we found the data to be heterogeneous with respect to the asset classification. Apart from that, we found no conclusive evidence for factors that influence the accuracy of the forecasts.

However, this does not mean other factors do not exist. Other research [33] is performed to qualitatively or quantitatively determine what influences the accuracy of the forecasts. Although these factors may differ from organization to organization, we stress the importance of investigating them. These factors will assist the estimators in determining how they are able to further improve their forecasting.

9. Conclusions

This article discussed how to quantify the forecasting quality of IT business value. To support decision making, IT business value is captured in numerous economic indicators that summarize the expected business value of project proposals.

Although not without its limitations, theory advocates the use of the Net Present Value (NPV) method. Recall that the Net Present Value is a summation of the predicted monetary benefits and cost of a project discounted to current value.

We showed that the NPV is composed of components, such as the benefits, asset usage cost and project cost. Each of the components, needs to be predicted. Therefore, in practice many organizations use multiple economic indicators to gain insight in the value of an asset proposal. Like the NPV, any economic indicator is highly dependent on the accurate forecasting of its elements. In this article, we assessed the forecasting quality of the NPV, and its components: the benefits, asset usage cost and project cost.

To quantify the forecasting quality of the NPV, we generalized a method from another article [14]. This method consists of the f/a ratio, the Estimating Quality Factor (EQF) and the reference cone. Recall that the EQF is a measure of the deviation between forecast and actual. For that method, three problems arise when we use it to quantify the forecasting quality of IT business value.

First, the f/a ratio has visual complications in asymptotic behavior. For instance, if the actual a is zero, the ratio becomes infinite and cannot be drawn in a normal way.

Second, the f/a ratio compares the forecast with the actual that is objectively measurable. However, when we want to use the forecasting quality to enhance decision information, the actual is not necessarily the most useful reference point. For instance, the actual may become known after many years when the original forecasting method may have already changed. Or, in case of the NPV, the actual may not be computed at all. Then, we need an alternative: a reference point that is not the actual, but still a better approximation than when the forecasting started. For instance, it is possible to re-estimate the actual two or three years after the start of the asset.

Third, the f/a ratio is only applicable for non-negative valued entities. However, the NPV can take negative values as well.

To remove these problems, we extended and generalized the method. This generalized method is applicable for any entity. The tools to be used depend on the entity in question.

We assessed the differences between the existing method for positively valued indicators with our generalized method. It turns out that overall the same conclusions are drawn.

Before using the method to quantify the forecast quality of the IT business value for a real-world case study, we performed a prerequisite data analysis. We found the real-world data to be heterogeneous with respect to the asset classification. For Cost Reduction and New Product Development projects, the forecasting quality of the NPVs and its components, the benefits and asset usage cost are significantly different. The investigation of other variables did not result in further heterogeneity.

With this knowledge, we performed the analysis of the forecasting quality of the NPV, benefits, asset usage cost and project cost for 102 IT assets of a telecommunication organization. The real-world case study represents an NPV value of 1812 million Euro, with discounted benefits of 4714 million Euro and an investment value of 173 million Euro.

We found that Cost Reduction assets have no biases for the benefit, asset usage cost and project cost forecasts. However, the New Product Development forecasts showed optimistic biases for the benefits and asset usage cost forecasts. Moreover, the New Product Development forecasts were of lower quality in terms of EQF than the Cost Reduction forecasts. Compared to benchmarks from the literature, the forecasting quality of the project cost is reasonable.

Also, we performed a sensitivity analysis to determine the impact of the different components on the forecasting quality of the NPV. The analyses showed that an improvement of the benefit forecasts results in the highest improvement of the NPV forecasting quality. Interestingly, in our real-world case study, if the quality of the asset usage cost forecasts improves, it results in a decreased forecasting accuracy of the NPV. These results support the idea that CIO's must manage IT assets to maximize business value, instead of controlling their cost.

Finally, we discussed how to enhance decision information using the quantified forecasting information. We described two simulation examples to show how to use the quantified information of the project cost and asset business value. These examples assists with questions, such as: if a certain number of project proposals would be executed, what is the likelihood that the available budget is sufficient to finance them? Or, given a set of proposals, what is a likely range of business values that will be generated? In these cases, the simulations will provide for additional information to assist the IT governors in their decision making.

This article showed how to assess the forecasting quality of IT business value. Forecasts are a crucial aspect for IT governors as they determine the potential value and risk related to IT assets. Therefore, knowing the forecast quality and potential bias is essential in the decision making process.

Acknowledgements

This research received partial support by the Dutch *Joint Academic and Commercial Quality Research & Development (Jacquard)* program on Software Engineering Research via contract 638.004.405 *EQUITY: Exploring Quantifiable Information Technology Yields* and via contract 638.003.611 *Symbiosis: Synergy of managing business-IT-alignment, IT-sourcing and off shoring success in society*. Furthermore, we would like to thank a number of organizations that will remain anonymous and Todd Little of Landmark Graphics Corporation for kindly sharing their forecasting data with us. Also, we thank our colleague Rob Peters for his help, comments and remarks in reviewing this article numerous times. Finally, we are grateful to Capers Jones and the reviewers for their comments, which allowed us to further improve this article.

References

[1] Fred Arditti, The weighted average cost of capital: some questions on its definition, interpretation, and use, The Journal of Finance 28 (4) (1973) 1001–1007.
[2] Glen Arnold, Panos Hatzopoulos, The theory-practice gap in capital budgeting: evidence from the United Kingdom, Journal of Business Finance & Accounting 27 (5) (2000) 603–626.
[3] François Bergeron, Jean-Yves St-Arnaud, Estimation of information systems development efforts, Information & Management 22 (1992) 239–254.
[4] Barry Boehm, Software Engineering Economics, Prentice Hall PTR, 1981.
[5] R.A. Brealey, S.C. Myers, Principles of Corporate Finance, McGraw-Hill, Inc., 1988.
[6] Lionel Briand, Khaled El Emam, Dagmar Surmann, Isabella Wieczorek, Katrina D. Maxwell, An assessment and comparison of common software cost estimation modeling techniques, in: Proceedings of the 21st international conference on Software engineering, 1999, pp. 313–322.
[7] Robert Bruner, Kenneth Eades, Robert Harris, Robert Higgins, Best practices in estimating the cost of capital: survey and synthesis, Financial Practice and Education (1998) 13–28.
[8] Manfred Bundschuh, Carol Dekkers, The IT Measurement Compendium, Springer, 2010.
[9] Murali Chemuturi, Software Estimation Best Practices, Tools & Techniques, J. Ross Publishing, 2009.
[10] S.D. Conte, H.E. Dunsmore, V.Y. Shen, Software Engineering Metrics and Models, The Benjamin/Cummings Publishing Company, 1986.
[11] Tom DeMarco, Controlling Software Projects, Prentice Hall PTR, 1982.
[12] Sanjeev Dewan, Charles Shi, Vijay Gurbaxani, Investigating the risk-return relationship of information technology investment: firm-level empirical analysis, Management Science 53 (12) (2007) 1829–1842.
[13] J.B. Dreger, Function Point Analysis, Prentice Hall, 1989.
[14] J.L. Eveleens, C. Verhoef, Quantifying IT forecast quality, Science of Computer Programming 74 (2009) 934–988.
[15] J.L. Eveleens, C. Verhoef, The rise and fall of the chaos report figures, IEEE Software 27 (1) (2010) 30–36.
[16] Marek Fisz, Probability Theory and Mathematical Statistics, John Wiley & Sons, 1963.
[17] James M. Fremgen, Capital budgeting practices: a survey, Management Accounting 54 (11) (1973) 19–25.
[18] D. Garmus, D. Herron, Function Point Analysis — Measurement Practices for Successful Software Projects, Addison-Wesley, 2001.
[19] Lawrence J. Gitman, John R. Forrester, A survey of capital budgeting techniques used by major US firms, Financial Management 6 (3) (1977) 66–71.
[20] META Group, The Business of IT Portfolio Management: Balancing Risk, Innovation, and ROI, Technical report, META Group, Stamford, CT, USA, January 2002.
[21] Harold Bierman Harold, Seymour Smidt, The Capital Budgeting Decision: Economic Analysis of Investment Projects, 8th ed., Macmillan Publishing Company, 1993.
[22] Peter R. Hill, Practical Software Project Estimation, McGraw-Hill, 2011. International software benchmarking standards group.
[23] Jeremy Hope, Robin Fraser, Beyond budgeting: building a new management model for the information age, Management Accounting 77 (1) (1999) 16–21.
[24] Rob J. Hyndman, Anne B. Koehler, Another look at measures of forecast accuracy, International Journal of Forecasting 22 (2007) 679–688.
[25] Jonathan E. Ingersoll, Stephen A. Ross, Waiting to invest: investment and uncertainty, The Journal of Business 65 (1) (1992) 1–29.
[26] Capers Jones, Conflict and litigation between software clients and developers, Technical note, April 2001. Version 10.
[27] Capers Jones, Estimating Software Costs: Bringing Realism to Estimating, second ed., McGraw-Hill, Inc., 2007.
[28] Capers Jones, Applied Software Measurement: Global Analysis of Productivity and Quality, third ed., McGraw-Hill, Inc., 2008.
[29] Capers Jones, Personal communication, March 2011.
[30] Magne Jørgensen, Experience with the accuracy of software maintenance task effort prediction models, IEEE Transactions on Software Engineering 21 (1995) 674–681.
[31] P. Kampstra, Beanplot: a boxplot alternative for visual comparison of distributions, Journal of Statistical Software 28 (2008).
[32] Chris Kemerer, An empirical validation of software cost estimation models, Communications of the ACM 30 (5) (1987) 416–429.
[33] G.P. Kulk, C. Verhoef, Quantifying IT estimation risks, Science of Computer Programming 74 (2009) 900–933.
[34] Linda M. Laird, M. Carol Brennan, Software Measurement and Estimation: A Practical Approach, John Wiley & Sons, 2006.
[35] Joseph L. Bower, Managing the resource allocation process, Harvard University, 1970.
[36] I. Leliveld, M. Jeffery, IT portfolio management — Challenges and best practices, Technical report, Kellogg School of Management and Diamond Cluster International Inc., 2003.
[37] Tim Lister, Becoming a better estimator — An Introduction to using the EQF Metric, 2002. Available via www.stickyminds.com.
[38] Todd Little, Context-adaptive agility: managing complexity and uncertainty, IEEE Software 22 (3) (2005) 28–35.
[39] Todd Little, Schedule estimation and uncertainty surrounding the cone of uncertainty, IEEE Software 23 (3) (2006) 48–54.
[40] James H. Lorie, Leonard J. Savage, Three problems in rationing capital, The Journal of Business 28 (4) (1955) 229–239.
[41] Maplesoft. www.maplesoft.com/Products/Maple/.
[42] Steve McConnell, Software Estimation: Demystifying the Black Art, Microsoft Press, 2006.
[43] Andy Neely, Mike Bourne, Chris Adams, Better budgeting or beyond budgeting? Measuring Business Excellence 7 (3) (2003) 22–28.
[44] R.J. Peters, C. Verhoef, Quantifying the yield of risk-bearing IT-portfolios, Science of Computer Programming 71 (2008) 17–56.
[45] J. Petty, D. Scott, M. Bird, The capital expenditure decision-making process of large corporations, The Engineering Economist 20 (3) (1975) 159–172.
[46] Dien Phan, Douglas Vogel, Jay Nunamaker, The search for perfect project management, Computerworld 26 (1988) 95–100.
[47] Richard H. Pike, Capital budgeting in the 1980s: a major survey of the investment practices in large companies, The Institute of Cost and Management Accountants, 1982.
[48] T. Pisello, P.A. Strassmann, IT Value Chain Management – Maximizing the ROI from IT Investments, in: Digital Publications from The Information Economics Press, vol. D10, The Information Economics Press, New Canaan, Connecticut, USA, 2003.
[49] L.H. Putnam, A general empirical solution to the macro software sizing and estimating problem, IEEE Transactions of Software Engineering 4 (4) (1978) 345–381.
[50] Stephen Ross, Randolph Westerfield, Bradford Jordan, Fundamentals of Corporate Finance, McGraw-Hill, Inc., 2000.
[51] Stephen A. Ross, Uses, abuses, and alternatives to the net-present-value rule, Financial Management 24 (3) (1995) 96–102.
[52] Patricia A. Ryan, Glenn P. Ryan, Capital budgeting practices of the fortune 1000: how have things changed? Journal of Business and Management 8 (4) (2002) 355–364.
[53] Lawrence D. Schall, Gary L. Sundem Jr., William R. Geijsbeek, Survey and analysis of capital budgeting methods, The Journal of Finance 33 (1) (1978) 281–287.
[54] Martin Shepperd, Chris Schofield, Estimating software project effort using analogies, IEEE Transactions of Software Engineering 23 (11) (1997) 736–743.
[55] Oxford University Press, Concise Oxford English Dictionary, 10th ed., 1999.
[56] C. Verhoef, Quantifying the value of IT-investments, Science of Computer Programming 56 (2005) 281–287.