

Diversity in resource consumption patterns and costing system robustness to errors

Eva Labro

Department of Accounting and Finance,
London School of Economics.

Mario Vanhoucke

Faculty of Economics and Business Administration, Ghent University
and Vlerick Leuven Gent Management School.

Version 30 October 2006.

Corresponding author: Eva Labro, Department of Accounting and Finance, London School of Economics, Houghton Street, London WC2A 2AE, U.K. e.labro@lse.ac.uk. Comments and suggestions from Eddy Cardinaels, Matt Hall, Thomas Hemmer and participants at workshops at Free University Amsterdam, HEC Paris, Rotterdam School of Management, LSE and Tilburg University are gratefully acknowledged.

Diversity in resource consumption patterns and costing system robustness to errors

Abstract

When decisions are made on the basis of reported product cost, even modest distortions in product cost may create significant distortions in decision-making. But costing systems are unlikely to be error-free. On the other hand, firms' budgets to enhance costing accuracy are typically constrained, and should be used where they are most effective. There is little guidance available, however, on how to (1) assess costing system quality, (2) improve costing system robustness to unwanted errors and (3) identify situations where costing system refinement efforts (such as introducing an ABC system intended to better reflect causal relationships) are likely to pay off most in terms of increased accuracy. Using a simulation method, this paper addresses these issues. We model various aspects and degrees of diversity in the resource consumption patterns to be reflected by the costing system and find that the hypothesis in academic and practitioner literature that more diversity in resource consumption patterns may lead to increased costing system sensitivity to errors holds only for some aspects of this diversity. Also, we identify situations in which following the high diversity rule of thumb and allocate costing system refinement resources to cases where there is a lot of diversity in resource consumption patterns is detrimental to improved accuracy.

Keywords: costing system design, costing errors, high diversity rule.

Data availability: the simulated datasets are available from the corresponding author on request.

Note: tables that include graphs are best printed in color to enhance readability. After publication, color plots can be provided on the author's website.

1. Introduction

When decisions are made on the basis of reported product cost, even modest distortions in product cost may create significant distortions in decision-making. When a firm's costing system does not reflect resource consumption patterns with sufficient accuracy, managerial decisions may result in wrong pricing policies, wrong product planning and ultimately in sub-optimal profits. It has long been accepted, however, that costing systems in general, and Activity Based Costing (ABC) in particular, provide accurate costs only under very specific circumstances (Noreen 1991; Datar and Gupta 1994; Christensen and Demski 1997). As a consequence, costing systems are unlikely to be error-free. Firms' budgets to enhance costing accuracy, however, are typically constrained, and accountants working in an era of increasing demand for their services (e.g. due to Sarbanes-Oxley) are stressed for time. Time and resources should therefore be used where they can be expected to be most effective. It is therefore surprising that little academic guidance is available on how to (1) assess the quality of the costing systems, (2) improve the robustness of costing systems to unwanted errors and (3) identify situations where costing system refinement efforts (such as introducing an ABC system intended to better reflect causal relationships) are likely to pay off most in terms of increased accuracy.

Gupta (1993) and Hwang et al (1993) hypothesize that more diversity¹ in resource consumption patterns may lead to increased costing system sensitivity to unwanted errors. Christensen and Demski (1997) also stress the complex nature of interactions between production technology and costing procedure.² Practitioners and management accounting textbooks propose the high diversity rule of thumb: efforts to refine costing systems should be focused on those cases where there is a lot of diversity in resource consumption patterns (e.g. Cooper 1988; Estrin et al 1994; Cooper and Kaplan 1998; Horngren et al 2005). To our knowledge, the performance of this rule of thumb has never been empirically tested.

Interestingly, however, Gupta (1993) points out that diversity in resource consumption patterns not only adds to costing error, but can also contribute to reductions in costing error through offsetting effects. Recognition of such offsetting effects in cost allocation systems makes it impossible to formally demonstrate simple propositions such as “diversity in resource consumption patterns leads to more errors in product costs” (Gupta 1993). We therefore resort to simulation to study the effects of various degrees of diversity in resource consumption patterns on the costing system’s robustness to errors. While not common (Balakrishnan and Sivaramakrishnan 2002), there exists management accounting literature that uses simulation methods.³

We build on the simulation method used by Labro and Vanhoucke (2007), in the remainder of the paper abbreviated to LV. They simulate a large variety of true costing systems to cover the range of potential true cost benchmarks, and for each of those a large set of false costing system approximations where they vary aggregation, specification and measurement errors in a controlled way. They present results on the interactions between these errors. In this paper, we not only vary these errors, but also model diversity in resource consumption patterns to be reflected by the costing system designs. This diversity is multi-faceted: it can relate to (1) the way in which resources are shared across activities and products in the whole of the costing system, (2) differences in proportional resource usage by activities and products at one particular cost pool, and (3) differences in dollar size of the cost pools. We develop a distribution measure to vary these aspects of diversity in the costing system in a controlled way. This approach allows us to provide insights on the interaction between errors and aspects of diversity in the costing system and speak to whether our results generalize or only hold for a sub-set of costing systems. Also, we find that the practitioner intuition that more diversity in resource consumption patterns leads to increased sensitivity to errors only holds for some aspects of diversity.

In doing so, we explore parameters that affect the design of effective costing systems, as called for by, amongst others, Datar and Gupta (1994) and Balakrishnan and Sivaramakrishnan (2002). We develop ex-ante indicators to help assess whether a costing system exhibits characteristics that are likely to make it more robust or sensitive to errors. The decision maker can then bear in mind the degree of robustness to errors when using the costing information. If needed, she can take appropriate accuracy enhancing action. Two options are available. First, she can improve the costing systems' robustness to aggregation, specification and measurement errors by working on the robustness indicators. Alternatively, she can also invest resources in reducing aggregation and specification error (e.g. by the introduction of an ABC system) or measurement error (e.g. by the introduction of a more precise measurement system). These resources can be allocated to those cases where the likely pay off in terms of enhanced accuracy is highest.

Importantly, we identify situations in which following the high diversity rule of thumb (e.g. Cooper 1988; Estrin et al 1994; Cooper and Kaplan 1998; Horngren et al 2005) and focus such costing system refinement efforts on cases where there is a lot of diversity in resource consumption patterns is detrimental to improved accuracy. Scarce resources will be allocated where they are not most effective in reducing overall error. We identify cases where less diverse resource consumption patterns benefit more from reductions in certain errors and therefore a "low diversity rule" holds. Further, we also find that the impact of some other errors is non-monotonic in increases in diversity in the sharing of resources across the whole of the costing system. Sometimes, a less diverse system will benefit more from a reduction in errors than a more diverse system.

The remainder of the paper is organized as follows. The next section develops the modeling of the various aspects of diversity in resource consumption patterns to be reflected by the costing system design. Building on Datar and Gupta (1994), we explain the variety of

errors in the design of costing systems that we are studying in section 3. Section 4 explains the workings of our simulation method. Section 5 presents our results on indicators that allow decision makers to assess and improve costing systems' robustness to errors. Section 6 discusses in which cases investment of resources in costing system refinements are likely to pay off most and tests the performance of the high diversity rule of thumb in identifying these. Finally, we conclude.

2. Modeling diversity in resource consumption patterns in costing systems

In order to capture the multi-faceted nature of diversity in resource consumption patterns to be reflected by a costing system design, we model parameters that relate to (1) the way in which resources are shared across activities and products in the whole of the costing system, (2) differences in proportional resource usage by activities and products at one particular cost pool, and (3) differences in dollar size of the cost pools. Where previous literature typically focused on the second stage of the allocation mechanism (e.g. Noreen 1991⁴; Datar and Gupta 1994; Babad and Balachandran 1993), we follow LV by modeling a two-staged allocation process, since they identify offsetting effects between the two stages of the cost allocation system. The basic setup of such two-staged costing system consists of a number of resource cost pools (RCP) that provide resources to a number of activity cost pools (ACP), who in their turn are used by the cost objects or products (CO). Figure 1 graphically depicts the modeled diversity indicators and introduces the notation used.⁵

- Insert Figure 1 about here-

To model *diversity in the way in which resources are shared across activities and products in the whole of the costing system*, we introduce the following parameters. First, as suggested by Gupta (1993) we introduce the number of activity drivers AD. The lower the number of activity driver links, the more diverse the resource usage of products across

activities is, as each activity only serves few cost objects. The higher the number of activity driver links, on the other hand, the lower the diversity in the use of activities by cost objects, as many cost objects will use many activities. AD reaches its maximum when each activity cost pool is needed for each cost object ($AD^{\max} = ACP * CO$) and its minimum when there is a one-to-one relationship between activity cost pools and cost objects ($AD^{\min} = \max(ACP, CO)$). The latter is equivalent with a SHARE-value of zero as defined by Balachandran et al. (1997) where there is no resource sharing among products. The number of resource drivers RD models the equivalent parameter in the first stage of the costing system.

Second, having many activity drivers AD is not sufficient to guarantee low diversity in the sharing of resources in the costing system; the activity drivers must also be distributed evenly over all activity cost pools. ADD captures the variance in the distribution of the number of outgoing activity drivers over the activity cost pools. The lower ADD, the smaller the diversity in the resource sharing in the costing system, as the activity drivers will be spread more evenly over the activity cost pools. For example, in a costing system with 10 activity cost pools and 50 activity driver links, ADD is at its lowest when each activity cost pool has 5 outgoing activity drivers. This even spread of activity drivers over the activity cost pools will contribute towards low diversity in resource sharing. On the other hand, ADD is at its highest when 9 of the activity cost pools each have one outgoing activity driver, whereas the remaining 10th cost pool has the 41 others. Such high ADD will contribute towards high diversity in resource sharing. RDD parameterizes the equivalent in the first stage: the variance in the distribution of the number of outgoing resource drivers over the resource cost pools.

Where the previous measures parameterized diversity in resource consumption patterns across the whole of the costing system, we follow Hwang et al (1993) and Gupta (1993), and also model *diversity in proportional resource usage by activities and products at the level of the individual cost pools*. We capture this aspect of diversity by the variance in the

distribution of percentages of activity cost allocated per outgoing set of activity drivers at each activity cost pool (APD).⁶ If APD is zero, we have products that all consume the activity in equal proportions. For example, when an activity is used by 4 products, all products will use 25% of the activity. When APD approaches 1, we are faced with products that exhibit high diversity in their resource consumption, as some products use a lot of the activity, whilst others only use little bits. For example, one of the 4 products uses 97% of the activity, while the other 3 only use 1% each. RPD parameterizes the equivalent in the first stage: the variance in the distribution of percentages of resource allocated per outgoing set of resource drivers at each resource cost pool.

Finally, we introduce the variance in the distribution of total resources over the resource cost pools in Stage I (TRD) to model *diversity in the size of the cost pools*. If TRD is zero, all resource cost pools are of equal dollar size. Diversity in resource consumption patterns is low, as all resource types are of equal monetary importance. However, if TRD is very high we have a system with many very small resource cost pools (in dollar terms) and a few very large. In this case, diversity in resource consumption patterns is high as the monetary importance of the various resource types exhibits big differences.

In sum, situations with low TRD, high AD, low ADD and low APD (and equivalent in stage I) exhibit less diverse resource consumption patterns, whereas high TRD, low AD, high ADD and high APD (and equivalent in stage I) result in more diverse resource consumption patterns to be reflected by the costing system design. When $TRD = 0$, $RD = RD^{\max}$, $RDD = 0$, $RPD = 0$, $AD = AD^{\max}$, $ADD = 0$ and $APD = 0$ we have the lowest possible diversity in the resource consumption patterns as all activities consume all resources, all products consume all activities, and this in equal proportions. The SHARE-metric as defined by Balachandran et al (1997) is equal to 1 in this case.

We define a general distribution measure to model and vary variance in these diversity indicators in a controlled way. This measure captures the deviation from the mean for the parameter under consideration (between minimum and maximum allowable levels of deviation). We return to our earlier example on the measure of the spread of activity drivers over activity cost pools ADD in a costing system with 10 cost pools and 50 activity drivers. The mean number of activity drivers per cost pool is $50/10 = 5$. When the distribution measure is set to zero, no deviation from this mean is allowed and every activity cost pool will be simulated with 5 outgoing activity drivers. At the other extreme, when ADD is set to its maximum, the deviation from the mean number of activity drivers per cost pool is ensured to be as high as possible by letting 9 of the cost pools have the minimum number of one cost driver each and allocating all 41 remaining cost drivers to the 10th cost pool. We simulate 10 different settings in between these two extremes. Appendix A explains in detail the intuition behind the formula of the distribution measure, shows how it is applied to model each of the costing system diversity indicators and presents a more elaborate example.

3. Aggregation, specification and measurement errors in costing systems

Datar and Gupta (1994) classify costing errors in aggregation error, specification error and measurement error.⁷ We use their definitions as modified by LV to reflect the two-staged nature of the costing allocation, as summarized in Figure 2.

- Insert Figure 2 about here -

Aggregation error can occur at two different places in the costing system. First, prior studies (e.g. Noreen 1991; Babad and Balachandran 1993; Datar and Gupta 1994) have typically defined aggregation error on the *activity cost pool* (AE-ACP) level, which is the link between the first and second stage of the costing system. Heterogeneous activities such as an expensive engineering setup and a cheap laborer setup are pooled incorrectly in one activity

cost pool, for example, when only one setup activity is defined. Second, LV also acknowledge that aggregation error can occur at the first stage of the costing system. Such errors arise when heterogeneous resources are aggregated in the same *resource cost pool* (AE-RCP). For example, rent and supervision costs are pooled into “administration costs”.

Measurement error can occur at three different places in the costing system. First, a particular *resource cost pool* might be over- or undercharged (ME-RCP). For example, when the cost of a marketing campaign is accidentally charged to the administrative expenses resource cost pool, the latter resource cost pool will be overcharged and the marketing expenses resource cost pool will be undercharged. Second, there can be measurement error on the *resource driver* (ME-RD) where an inaccurate number of allocation base is employed in the allocation of resource cost to activity. For example, such errors arise when an administrative assistant erroneously estimates the time spent on handling invoices, and this estimate is used to allocate the administrative salaries among activities. Third, measurement error can arise at the level of *activity drivers* (ME-AD) when, for example, the cost object is estimated to consume 10 incoming inspections, whereas it consumes only 8.

Specification error relates to the selection of the wrong cost driver unit, such as volume-based drivers rather than transaction-based drivers. As in Babad and Balachandran (1993, 565) and LV, our analysis uses the relative, rather than absolute, frequency (or volume of use) of a cost driver associated with a cost object or process. Therefore, here, specification error as defined in Datar and Gupta (1994) coincides with measurement error on the cost driver, as both result in a wrong percentage of the resource cost being allocated to the activity (in a first stage example). We therefore use only one construct to model both specification error and measurement error on the cost driver.⁸

4. Simulation experiments

We simulate a wide variety of true costing systems in order to cover the range of potential true cost benchmarks in Visual C++ version 6.0 under Windows XP, where we vary in a controlled way the diversity indicator under study. For each of these true costing systems, we simulate, in turn, a wide variety of (false) costing system approximations, whereby we vary either the aggregation, specification or measurement error that could have been introduced when the system was developed. We run simulations of all combinations of an error and a diversity indicator in a controlled experimental design.⁹ The errors are simulated using 10 settings of a disturbance factor (df) that represents the magnitude of the deviation from the true costing system by the introduction of aggregation, specification or measurement error. df ranges between 0 % and 90%¹⁰ with increments of 10%. Errors not under study are set to 0% to exclude confounding effects.

The diversity indicators are simulated using 11 settings of the general distribution measure as described in section 2 and appendix A, ranging from 0 % to 100% in increments of 10%. To get stable simulations, we run 15 and 40 simulation runs for each error and diversity indicator setting, respectively. For example, when simulating the data to study the combined effect of diversity indicator RPD and error AE-RCP, we use 11 settings for RPD ($u = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ and 1) and 10 for AE-RCP ($df = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9). For RPD = 0.1, we simulate 15 cost systems with AE-RCP = 0.1, 15 with AE-RCP = 0.2 and so on. This results in $15 \times 10 = 150$ “false” cost systems (i.e. cost systems in which error is introduced) for one true cost system with an RPD of 0.1. We simulate 40 true cost systems that have this value of RPD = 0.1. This results in $40 \times 150 = 6,000$ observations for costing systems that combine an RPD of 0.1 with some level of AE-RCP. We repeat this for all other settings of RPD, resulting in a total of $11 \times 6,000 = 66,000$ observations.

The diversity indicators that are not studied in a particular effect are put to their average value of 50% in the results reported on in the body of the paper. We recognize that there may be higher order interaction effects of the diversity indicator under study with the settings chosen for all other diversity indicators, but computational restrictions prevent us from studying more than one diversity indicator (together with one error) at the time. As explained in Appendix B, we therefore run two robustness checks for the level of diversity in the other diversity indicators. A “less diverse design” has values for the unvaried diversity indicators that tend towards less diversity in resource consumption patterns, whereas a more “more diverse design” has values tending towards more diversity. This results in $66,000 \times 3 = 198,000$ observations for one particular effect. Accordingly, the results in this paper are based on over 7,000,000 observed data points. In general we find qualitatively similar results for these robustness checks; the odd exceptions are reported in footnotes to the paper.

Measuring the overall error in the costing system

The introduction of aggregation, specification and measurement errors in the design of the costing system will result in inaccurate costing of cost objects. Following Babad and Balachandran (1993), Hwang et al (1993), Homburg (2001) and LV we use the square root of a symmetric quadratic error measure EUCD as a measure of the costing system’s overall error, which we compute for each simulated data point. This dependent variable is defined as follows (in dollar terms):

$$EUCD = \sqrt{\sum_{k=1}^{CO} (tc_k - fc_k)^2}$$

with k indexing the number of cost objects ($1, \dots, CO$)

tc_k : true cost accruing to cost object k in the true benchmark scenario

fc_k : (false) cost allocated to cost object k by the costing system approximation.

This measure is symmetric and, given we keep total resources TR constant (at 1 million) across all simulations, also captures the magnitude of the resulting overall error in the costing system.

The desired accuracy target and the relevant measure of accuracy necessarily depends on the context faced by the firm, the objective for which costing figures are provided (e.g. product costing, pricing, product line decisions, capacity planning, capacity allocation, control, performance measurement, benchmarking and project selection) and the capability of human decision makers to work with error-prone data. More accuracy may not always be needed or desired (e.g. Merchant and Shields 1993; Banker and Potter 1993; Alles and Datar 1998; Callahan and Gabriel 1998 and Kanodia et al. 2005) and various measures of accuracy may be relevant in a variety of contexts. Appendix B reports on our robustness checks for several other costing accuracy measures that may be relevant in a variety of contexts in which the costing figures can be used.

Data analysis methodology

For each set of simulated data we perform ANOVA analyses and report R^2 , F-values, and significance levels. Given our large sample sizes (typically 66,000 observations per data set that looks at the effect of an error and a diversity indicator), all effects easily become significant at very high levels. Therefore, we also report partial h^2 values that provide a better picture of effect size as they reflect the proportion of total variability attributable to a factor.

We graphically show marginal estimated means plots whereby the X-axis shows the diversity indicator under study ranging from low to high diversity, the Y-axis measures ?EUCD and the lines in different colors represent various sizes of the error under study. Further, we run regressions of the form

$$EUCD = a + b(X_1 - \bar{X}_1) + c(X_2 - \bar{X}_2) + d(X_1 - \bar{X}_1)(X_2 - \bar{X}_2) + u ,$$

where X_1 is the error and X_2 the diversity indicator under study. These independent variables in the regressions are centered on their respective means to remove multi-collinearity and to scale the interaction effect (Aiken and West 1993).¹¹ We report the standardized coefficients, t-values and R^2 . Running the regressions allows us to sign the main and interaction effects. For those data sets where the sign of the effects changes for different parameter values we additionally perform split sample analyses.

5. Assessing and improving costing systems' robustness to errors.

We capture the multi-faceted nature of diversity in the resource consumption patterns to be reflected in the costing system design by the diversity indicators developed in section 2. Studying the interaction between these indicators and the aggregation, specification and measurement errors that can be made in the development of a costing system, allows us to (1) assess how robust costing systems are to specific errors and (2) improve their robustness by working on these diversity indicators.

Diversity in the size of the resource cost pools.

Our results suggest that decreasing diversity in the size of the resource cost pools increases the robustness of the costing system to errors. Costing systems with resource cost pools that are very similar in size are likely to be more robust to errors, whereas costing systems with high variance in the sizes of their resource cost pools are highly sensitive to errors. This also means that when a firm operates a costing system with very different sizes in its resource cost pools, it can improve its robustness to errors by redefining the cost pools so that they exhibit less size variance. Note, however, that the firm's ability to make this robustness improvement is restricted by the underlying production technology the costing

system is meant to reflect; and as a consequence there will be contextual limits to the level of robustness that can be achieved.

Table 1 summarizes these results for the interaction between the variance in the distribution of resources over resource cost pools (TRD) and measurement / specification error on the resource drivers (ME-RD). This table is representative of our (untabulated) results on all other combinations of the TRD diversity indicator with other types of measurement and aggregation errors, which allows us to generalize our results on this specific diversity indicator. The analyses and plot in panels A and B suggest there is a strong positive main effect of TRD: decreasing diversity in the sizes of the resource cost pools (reducing TRD) improves overall accuracy (as measured by reductions in the dependent variable EUCD on the Y-axis). There also seems to be a positive interaction effect between ME-RD and TRD (as shown by the positive sign on the standardized regression coefficient) in that, higher values of ME-RD and TRD interact to increase overall error even more.¹²

The decision maker can calculate the value of TRD for the existing costing system (as is shown in the example in appendix A), use our marginal estimated means plots of the effect of TRD and errors (such as provided in Table 1 for ME-RD) to evaluate how robust to errors her system is likely to be, and take appropriate action. She has three alternatives. First, she can decide that costing figures are robust enough to error for their particular use, so that no improved costing system is needed. Second, she can improve robustness of the costing system by trying to achieve a more equal spread of resources over resource cost pools. Alternatively, she can invest resources in reducing error in the system by, for this example, introducing a better measurement system. In section 6 we discuss where such costing system refinements are likely to pay of most in terms of enhanced accuracy.

- Insert Table 1 about here -

Diversity in proportional resource consumption at each cost pool.

Our results suggest that decreasing diversity in the proportional resource consumption at each cost pool increases costing systems' robustness to errors. Costing systems in which the proportional resource consumption at the cost pools is equal are more robust to errors, whereas costing systems with high variance in the proportional usage of cost pools by activities and cost objects are very sensitive to errors. Table 2 shows this result for the interaction of the variance in the distribution of percentages allocated by activity drivers at activity cost pools (APD) and measurement error on the activity drivers (ME-AD). Reducing APD improves overall accuracy. This table is representative of our results on all other combinations of APD with other measurement and aggregation errors, and of the equivalent diversity indicator in the first stage of the costing system (RPD).

- Insert Table 2 about here -

Diversity in resource sharing across the whole costing system.

So far, our results support the common intuition that decreasing diversity in resource consumption patterns results in increased costing system robustness to errors: reducing diversity in the size of the resource cost pools and diversity in proportional resource consumption at each cost pool leads to enhanced accuracy. The results we present in this section, however, indicate that decreasing diversity in (what is arguably the most important aspect of resource consumption patterns) resource sharing across the whole costing system does not necessarily go hand in hand with increased robustness to errors. Diversity in resource consumption patterns can also reduce costing error (Gupta 1993).

As explained in section 2, one aspect of diversity in resource sharing across the whole costing system is the distribution of cost driver links over the cost pools. Untabulated results suggest that less diverse resource consumption patterns where cost drivers are spread more

evenly over cost pools usually do result in increased robustness to errors. Systems with lower RDD and ADD, resulting in a similar number of cost driver links at each cost pool, are more robust to errors, as hypothesized.

There is, however, an important exception to this. Table 3 suggests that in combination with measurement / specification error on the cost drivers the *opposite* pattern emerges. Spreading activity drivers more unevenly over the activity cost pools (ADD high) results in lower overall error when combined with measurement / specification error on those activity drivers (ME-AD). When the diversity in the spread of activity drivers over the activity cost pools increases, it becomes more and more likely that the measurement / specification error on the activity drivers occurs on drivers from pools that are only providing allocations to a very limited number of cost objects and that therefore the impact of these errors is smaller than when the activity drivers are more evenly spread over all activity cost pools. Untabulated results indicate the same effect in the first stage of the costing system.¹³

- Insert Table 3 about here -

A second aspect of diversity in resource sharing across the whole costing system is the number of cost driver links. Costing systems with a high number of cost drivers are more likely to exhibit a low level of diversity in resource consumption patterns as many resources will be used by many activities in Stage I and many cost objects will use many activities in Stage II. Cooper (1989, 35) asserts that as the number of cost drivers used increases, the accuracy of reported costs rises. We do, however, not find such a monotonically decreasing effect on EUCD of increasing the number of cost drivers. Rather, we document interesting interactions between the various types of errors and the number of cost drivers. The effect of increasing or reducing the number of cost drivers on costing system robustness will depend heavily on the aggregation, specification and measurement error levels and on how many cost driver links are already used.

In contrast to common intuition, the overall picture that emerges from Table 4 is that increases in diversity in resource sharing across the whole costing system in terms of a reduction in the number of cost driver links result in *decreased* overall error when there is *measurement or specification error on the cost drivers* (ME-AD). This error on the drivers often outweighs the positive effects of the decreasing diversity. Untabulated results for this effect in the first stage are similar. Increasing the number of cost drivers is not the panacea it sometimes is portrayed to be (eg.g Cooper 1989), when these cost drivers are badly specified or their units inaccurately measured.

In the marginal estimated means plot we see that at the minimum number of activity drivers ($AD^{\min} = \max(\text{ACP}, \text{CO})$, the setting where $AD = 0$) which reflects the most diverse setting, the impact of measurement error on these drivers is low, as there are only few drivers in the system. Moving towards the left on the X-axis in the plot by adding drivers initially results in an increased impact of measurement and specification error on these few drivers, which outweighs the effect of decreased diversity in resource consumption patterns. Subsequently, however, over the range of $\%AD \in [0.2, 0.4]$ the effect of decreased diversity outweighs the measurement error and overall error decreases. Then measurement error outweighs again until we reach very high levels of AD, getting close to the maximum level of AD (whereby each activity cost pool provides resources for each cost object). When nearly every activity cost pool is linked to nearly every cost object (AD very high), the probability that the allocation process will cancel out measurement and specification error on the drivers will increase. The sub-sample correlation analyses in panel C show indeed that the sign of the AD effect changes between positive and negative moving over the various ranges of AD.¹⁴

- Insert Table 4 about here -

Further, although the overall trend in Tables 5 through 7 confirms the common intuition that higher diversity leads to increased sensitivity to errors, these tables also suggest

that this effect is non-monotonic. First, Table 5 indeed suggests that increasing the number of outgoing resource drivers (RD) (i.e. moving from right to left on the X-axis) in the presence of *measurement error on the resource cost pools* (ME-RCP) usually results in increased robustness to this error, as the decreased diversity in the resource consumption patterns is able to offset the measurement error. However, for the range of $\%RD \in [0.2, 0.4]$, the effect of decreased diversity is outweighed by the effect of measurement error on the resource cost pools, and over this range the overall error increases with an increase in the number of resource drivers. For this range of an intermediate number of resource drivers, allocation processes do not dampen the effect of measurement error on the resource cost pools; a costing system with an intermediate number of resource drivers is quite sensitive to measurement error on the resource cost pools. Sub-sample correlation analyses show again the effect of RD changing sign.

- Insert Table 5 about here -

Second, we observe a similar pattern when the number of *outgoing* cost drivers is increased in combination with *aggregation error on the cost pools*. Table 6 shows the results for the second stage, studying aggregation error on the activity cost pools in combination with changes in the number of outgoing activity drivers. Usually there is a beneficial effect on overall accuracy of increasing the number of cost drivers. This effect of decreased diversity is, however, outweighed by the effect of aggregation error on the activity cost pools over a range of an intermediate number of activity drivers ($\%AD \in [0.2, 0.4]$).¹⁵ Untabulated results for this effect in the first stage are similar.

- Insert Table 6 about here -

Third, if the number of *incoming* resource drivers is increased in combination with *aggregation error on the activity cost pools*, Table 7 suggests that the beneficial effect of adding resource drivers is outweighed by this aggregation error for a low to intermediate

range for the number of resource drivers ($\%RD \in [0.1, 0.3]$) as these few drivers are likely to be affected strongly by this error. Once there are more resource drivers in the costing system, these are not affected so strongly by aggregation error and the dampening effect of decreased diversity in resource consumption patterns dominates.

- Insert Table 7 about here -

6. Testing the high diversity rule of thumb.

ABC advocates propose the high diversity rule of thumb that recommends to focus costing system refinement efforts (such as the introduction of an ABC system to reduce aggregation and specification error or an on-line measurement system to reduce measurement error) in those cases where there is a lot of diversity in resource consumption patterns (Cooper 1988, Estrin et al 1994, Cooper and Kaplan 1998).

In these cases, cost system refinements are considered most cost effective and likely to yield most benefits for the company (Horngren et al 2005). To our knowledge, the performance of this rule in identifying those cases where costing systems benefit most from reductions in aggregation, specification and measurement error has never been tested. In this section, we identify cases where the rule holds, but more importantly also single out situations where following the high diversity rule is detrimental to accuracy improvements. Costing accuracy enhancing budgets will be spend on cases where there is little to gain from the refinements, whereas important improvement areas will remain unexplored. Further, our results also suggest that there is a non-monotonic relation between the impact of costing system refinements on overall accuracy and the degree of diversity in the resource consumption patterns to be reflected by the costing system. On several occasions a less diverse resource consumption pattern will benefit more from reductions in these errors than a more diverse pattern.

In order to test the performance of the high diversity rule of thumb in identifying cases where the benefit of costing system refinements on increased accuracy is the highest, we split each sample in halves according to the aspect of diversity under study. Next, we perform regression analyses (as specified in section 4) on both the “high diversity” and “low diversity” sub-samples and calculate the 99% confidence interval on the difference between the unstandardized regression coefficients on the aggregation, specification or measurement error in both samples. If this regression coefficient for the “high diversity” sub-sample is significantly higher than that for the “low diversity” sub-sample, we can conclude that the high diversity rule correctly identifies that resource consumption patterns that are highly diverse in this particular aspect indeed benefit more from reductions in aggregation, specification and measurement error than the less diverse patterns. On the other hand, when the regression coefficient for the error in the “low diversity” sub-sample is significantly higher than that for the “high diversity” sub-sample, we have identified a situation where following the high diversity rule is detrimental to improved accuracy, as resources for costing system refinements would be more effectively spend on “low diversity” cases.

Panels C in Tables 1 and 2 suggest that the high diversity rule of thumb holds with respect to two aspects of diversity in resource consumption patterns: differences in dollar size of the cost pools (TRD) and differences in proportional resource usage by activities and products at one particular cost pool (APD and RPD).¹⁶ Costing system refinements pay off most in terms of improved accuracy in cases where the sizes of the resource cost pools are very different and when there are big differences in the proportional resource usage at each cost pool.

In relation to the third aspect of diversity in resource consumption patterns -the way in which resources are shared across activities and products in the whole of the costing system- however, we find two important exceptions where a “low diversity rule” holds. First, panel D

in Table 3 suggests that cases where cost drivers are more evenly spread over the cost pools (ADD and RDD low) benefit more from reductions in measurement error and specification error on the cost drivers than in cases where there is more diversity in this distribution. The regression coefficient on ME-AD is significantly higher in the “low diversity” sub-sample than in the “high diversity” sub-sample. These measurement and specification improvements will have an impact on a larger number of cost objects under a less diverse pattern of resource sharing amongst activities and products.

Second, panel D in Table 4 suggests that systems with many cost drivers (that are therefore likely to have lower diversity in resource consumption patterns as resources are shared more amongst activities and products) benefit more from reductions in measurement and specification error on the cost driver than more diverse systems with few cost drivers. The impact of errors in the measurement of the allocation bases and of wrongly specifying the drivers is bigger when there are more drivers in the costing system. Together, panel C in Tables 3 and 4 provide strong counterevidence against the use of the high diversity rule when considering costing system refinements that reduce measurement or specification error on the cost drivers, as the opposite rule, a “low diversity rule” seems to hold.

Additionally, Tables 5 through 7 present weak counterevidence against the high diversity rule, suggesting that the impact of aggregation and measurement error on the resource cost pools exhibits a non-monotonic relation to increases in diversity in the way resources are shared among activities and products in the costing system, as modeled by AD and RD. The plots in these tables provide counter examples where a less diverse system will benefit more from a reduction in the errors than a more diverse system. For example, untabulated results related to the relation between aggregation error on the activity cost pools (AE-ACP) and the number of activity drivers (AD) as depicted in Table 6 show that, overall the high diversity rule holds for this sample. However, the unstandardized regression

coefficient on AE-ACP for the sub-sample where $30 \leq AD \leq 40$ is significantly greater than that for the more diverse sub-sample with $AD = 20$. This means that this particular less diverse sub-sample (with $AD = 20$) benefits more from reductions in aggregation error than the specific more diverse sub-sample (with $30 \leq AD \leq 40$). There is not a monotonic increase of the impact of the aggregation error with increased diversity in resource consumption patterns as reflected by a lower number of cost drivers.

7. Conclusions

When decisions are made on the basis of reported product cost, even modest distortions in product cost may create significant distortions in decision-making. But costing systems are unlikely to be error-free. On the other hand, firms' budgets to enhance costing accuracy are typically constrained, and should be used where they are most effective. This paper provides guidance on how to (1) assess costing system quality, (2) improve costing system robustness to unwanted errors and (3) identify situations where costing system refinement efforts (such as introducing an ABC system intended to better reflect causal relationships) are likely to pay off most in terms of increased accuracy.

Academic and practitioner literature hypothesize that more diversity in resource consumption patterns may lead to increased costing system sensitivity to errors. We model various aspects and degrees of diversity in the resource consumption patterns to be reflected by the costing system design, relating to (1) the way in which resources are shared across activities and products in the whole of the costing system, (2) differences in proportional resource usage by activities and products at one particular cost pool, and (3) differences in dollar size of the cost pools. We study the impact of each of these diversity indicators on the costing system's robustness to aggregation, specification and measurement errors using a simulation method. In doing so, we provide guidance for assessing whether a costing system

exhibits characteristics that likely make it more or less robust to these errors; and thus to improve costing system robustness.

We find that for two aspects of diversity in resource consumption patterns the intuition that increased diversity leads to decreased robustness to errors holds. Decreasing diversity in the dollar amount of the resource cost pools and in the proportional resource consumption by cost drivers at each cost pool increases robustness to errors. When all resource cost pools are of equal dollar magnitude and when the percentages allocated by cost drivers at cost pools are equal, costing systems are more robust to aggregation, specification and measurement error.

On the other hand, decreasing diversity in the sharing of resources across the whole of the costing system - which is arguably the most important aspect of diversity - does not have a monotonic effect on costing system robustness to errors. Increasing the number of cost drivers decreases the diversity in the resource allocation pattern and is therefore expected to make the costing system more robust to errors. Whilst we find this effect on occasions in our results, other effects of the interaction with the various errors also regularly outweigh it, so that increasing the number of cost drivers may lead to decreased robustness to these errors.

Also, decreasing diversity in the spread of cost drivers over cost pools often, but not always, increases robustness. The main exception to this occurs when there is measurement error or specification error on the cost drivers. When the diversity in the spread of cost drivers over the cost pools increases, it becomes more and more likely that this error occurs on drivers that are only providing allocations to a very small number of activity cost pools (in first stage) or cost objects (in second stage) and that therefore the impact of these errors is smaller than when the cost drivers are more evenly spread over all cost pools.

Further, we test the performance of the high diversity rule of thumb that says that costing system refinements should be focused on situations where there is high diversity in resource consumption patterns. Our results suggest that the high diversity rule of thumb holds

with respect to diversity in the dollar size of resource cost pools and in the proportional resource consumption by cost drivers at each cost pool in that the more diverse resource consumption situations benefit most from reductions in aggregation, specification and measurement error. We do however find that the rule does not hold with respect to diversity in resource consumption patterns in terms of the sharing of resources across the whole of the costing system.

We identify cases where less diverse resource consumption patterns benefit more from reductions in measurement and specification error on the cost driver and therefore a “low diversity rule” holds. Further, we also find the impact of aggregation error and measurement error on the resource cost pools to be non-monotonic in increases in diversity in the sharing of resources across the whole of the costing system. This means that there also exist counter examples of the high diversity rule of thumb for the latter errors where a less diverse system will benefit more from a reduction in these errors than a more diverse system.

8. Bibliography

- Alles, M., and S. Datar. 1998. Strategic transfer pricing. *Management Science* 44(4): 451-461.
- Aiken, L.S., and S.G. West. 1993. Multiple regression: testing and interpreting interactions. Sage Publications, Newbury Park, CA.
- Babad, Y.M., and B.V. Balachandran. 1993. Cost driver optimization in activity-based costing. *The Accounting Review* 68(3): 563-575.
- Balachandran, B.V., Balakrishnan, R., and K. Sivaramakrishnan. 1997. On the efficiency of cost-based decision rules for capacity planning. *The Accounting Review* 72(4): 599-619.
- Balakrishnan, R., and K. Sivaramakrishnan. 2001. Sequential solutions to capacity-planning and pricing decisions. *Contemporary Accounting Research* 18 (Spring): 1-27.

- Balakrishnan, R., and K. Sivaramakrishnan. 2002. A critical overview of the use of full-cost data for planning and pricing. *Journal of Management Accounting Research* 14: 3-31.
- Banker, R.D., and S.C. Hansen. 2002. The adequacy of full-cost-based pricing heuristics. *Journal of Management Accounting Research* 14: 33-58.
- Banker, R.D., and G. Potter. 1993. Economic implications of single cost driver systems. *Journal of Management Accounting Research* 5: 15-33.
- Callahan, C., and E. Gabriel. 1998. The differential impact of accurate product cost information in imperfectly competitive markets. *Contemporary accounting research* 15(4): 419-456.
- Christensen, J., and J. Demski. 1997. Product costing in the presence of endogenous subcost functions. *Review of Accounting Studies* 2(1): 65-87.
- Cooper, R. 1988. The rise of Activity-Based Costing – part two: when do I need an Activity-Based Costing system? *Journal of Cost Management* Fall: 41-48.
- Cooper, R. 1989. The rise of Activity-Based Costing – part three: how many cost drivers do you need, and how do you select them? *Journal of Cost Management* Winter: 34-46.
- Cooper, R., and R.S. Kaplan. 1998. The design of cost management systems. Prentice Hall, New Jersey.
- Datar, S., and M. Gupta. 1994. Aggregation, specification and measurement errors in product costing. *The Accounting Review* 69(4): 567-591.
- Demeulemeester, E., M. Vanhoucke, and W. Herroelen. 2003. A random network generator for activity-on-the-node networks. *Journal of Scheduling* 6: 13-34.
- Dhavale, D.G. 2005. Cost considerations in optimal capacity acquisition: an option pricing approach. *Journal of Management Accounting Research* 17: 75-93.
- Estrin, T.L., J. Kantor, and D. Albers. 1994. Is ABC suitable for your company? *Management Accounting* April: 40-45.

- Friedrich, R.J. 1982. In defense of multiplicative terms in multiple regression equations. *American Journal of Political Science* 26: 797-833.
- Gokins, G. 2001. Activity-Based Cost management: an executive's guide. John Wiley & Sons, Canada.
- Gupta, M. 1993. Heterogeneity issues in aggregated costing systems. *Journal of Management Accounting Research* 5 (fall): 181-212.
- Homburg, C. 2001. A note on optimal cost driver selection in ABC. *Management Accounting Research* 12: 197-205.
- Horngren, C., A. Bhimani, G. Foster, and S. Datar. 2005. Management and cost accounting. FT Prentice Hall.
- Hwang, Y., J.H. Evans III, and V.G. Hegde. 1993. Product cost bias and selection of an allocation base. *Journal of Management Accounting Research* 5 (fall): 213-242.
- Kanodia, C., R. Singh, and A. Spero. 2005. Imprecision in accounting measurement: Can it be value enhancing? *Journal of Accounting Research* 43(3): 487-519.
- Labro, E., and M. Vanhoucke. 2007. A simulation analysis of interactions among errors in costing systems. *The Accounting Review*, forthcoming.
- Leitch, R.A., Philipoom, P.A., and T.D. Fry. 2005. Opportunity costing decision heuristics for product acceptance decisions. *Journal of Management Accounting Research* 17: 95-117.
- Merchant, K. and M.D. Shields. 1993. When and why to measure cost less accurately to improve decision making. *Accounting Horizons* 7(2), 76-81.
- Noreen, E. 1991. Conditions under which activity-based costing systems provide relevant costs. *Journal of Management Accounting Research* 3 (fall): 159-168.

FIGURE 1

Modeling diversity in resource consumption patterns in costing systems

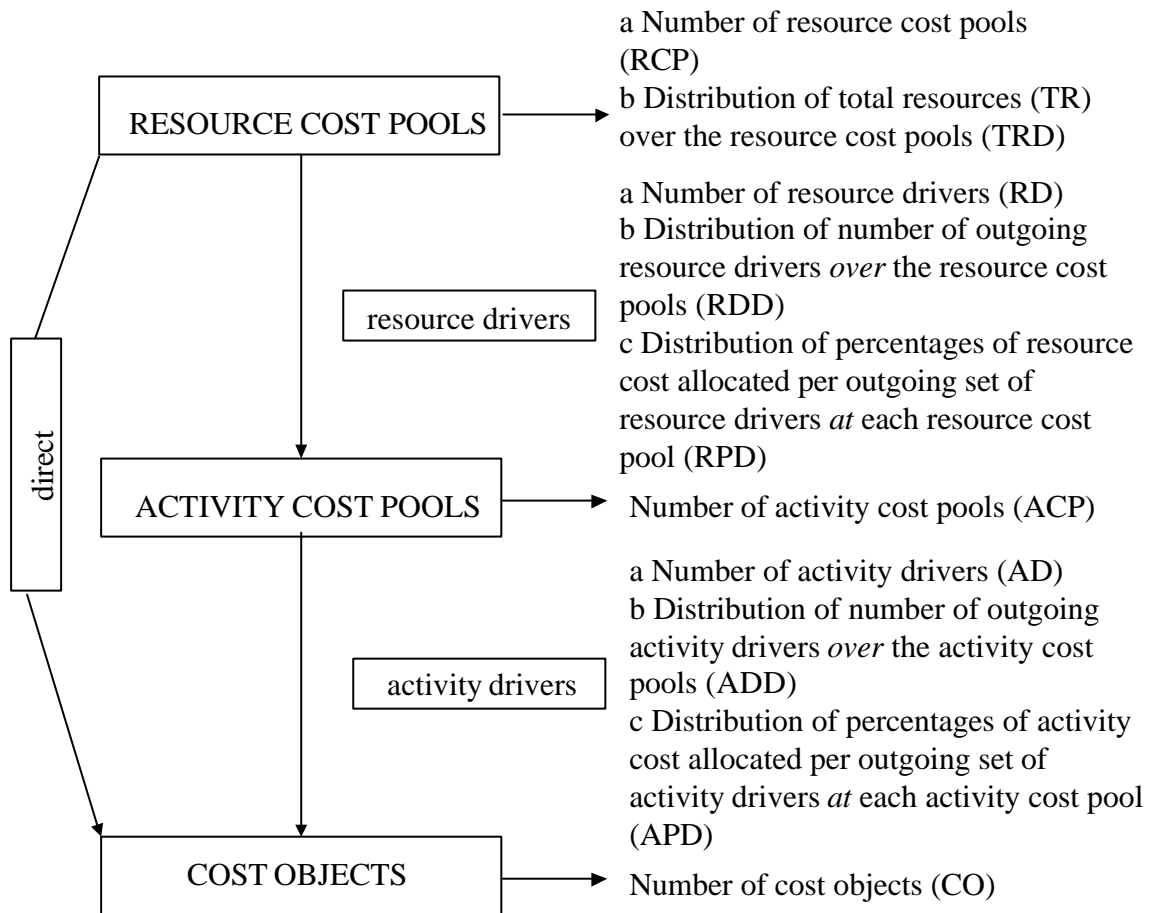
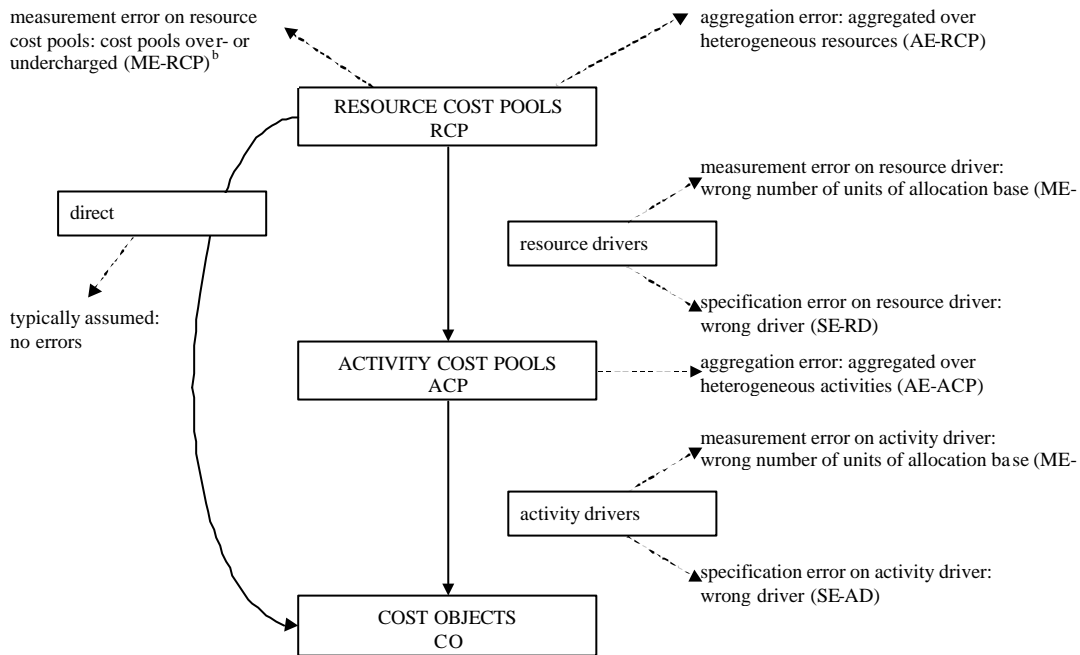


FIGURE 2
Errors in costing system design ^a



^a Source: LV.

^b ME-RCP is modeled by two parameters. ME1-RCP models the number of recourse cost pools subject to measurement error, whilst ME2-RCP models the size of the measurement error on these resource cost pools.

^c Empirically, specification error on the driver collapses with measurement error on the driver as both result in a wrong percentage of the resource to be allocated to the activity (in a first stage example) (LV). We will therefore only use one construct to model both in our simulation design.

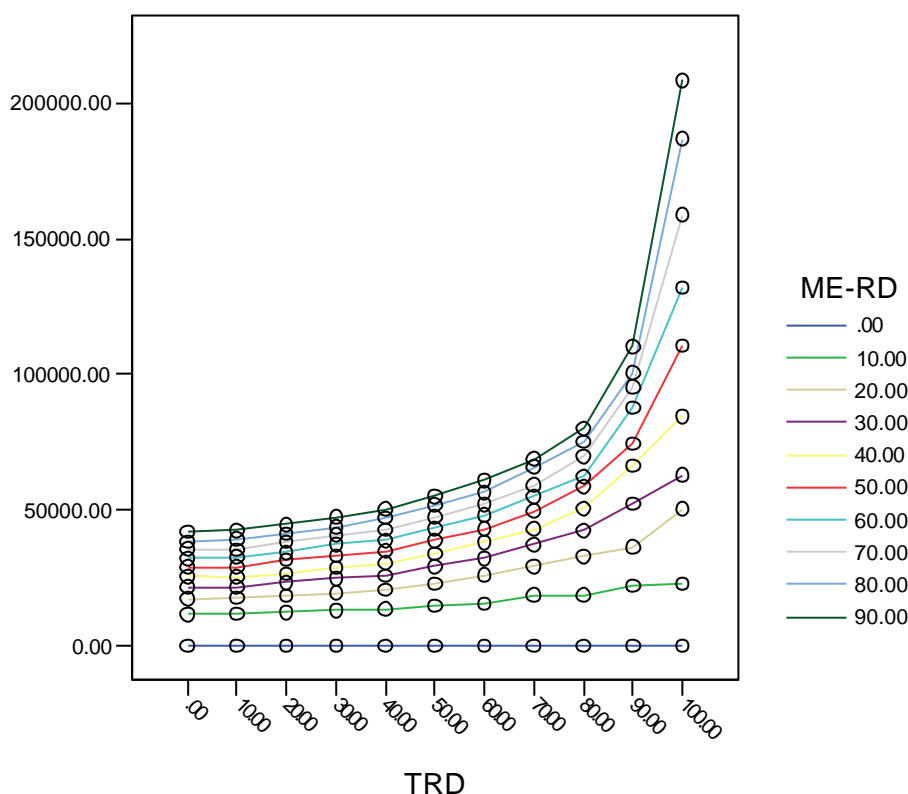
TABLE 1

Analysis of the distribution of resources over resource cost pools – measurement error on the resource driver effects ^a

Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression coefficient β ^b	t-value
TRD	1680	0.000*	0.203	0.329	102*
ME-RD	1947	0.000*	0.210	0.397	123*
TRD*ME-RD	92	0.000*	0.112	0.210	65*
R-Squared	0.393			0.310	

Panel B: Marginal Estimated Means Plot of EUCD



Panel C: High diversity rule test

	Less diverse Sub-sample	More diverse Sub-sample	99% confidence interval on difference between B's
Sub-sample with TRD B (unstandardized regression coefficient on ME-RD)	≤ 50 473.788	> 50 1126.918	653.13 ± 36.71

* indicates significance at the 1% level.

^a EUCD = square root of mean squared error, TRD = variance in the distribution of resources over resource cost pools, ME-RD = measurement error on resource drivers. Note that, as explained in section 3, measurement error on the cost driver can in this empirical setting also be interpreted as specification error.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

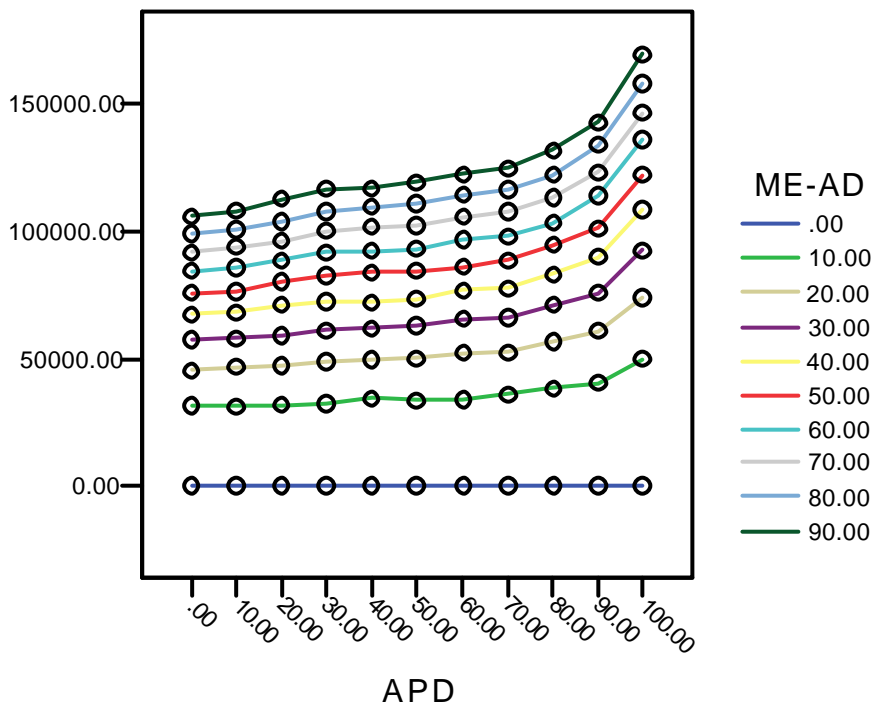
TABLE 2

Analysis of the distribution of percentages of activity cost allocated by activity drivers at the activity cost pools – measurement / specification error on the activity driver effects ^a

Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression coefficient β ^b	t-value
APD	3284	0.000*	0.333	0.230	131*
ME-AD	41374	0.000*	0.850	0.857	490*
APD*ME-AD	81	0.000*	0.099	0.105	60*
R-Squared	0.862			0.798	

Panel B: Marginal Estimated Means Plot of EUCD



Panel C: High diversity rule test

	Less diverse Sub-sample	More diverse Sub-sample	99% confidence interval on difference between B's
Sub-sample with APD	≤ 50	> 50	
B (unstandardized regression coefficient on ME-AD)	1133.212	1381.547	248.34 ± 14.34

* indicates significance at the 1% level.

^a EUCD = square root of mean squared error, APD = variance in the distribution of percentages allocated by activity drivers at activity cost pools, ME-AD = measurement / specification error on activity drivers.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

TABLE 3

Analysis of the distribution of activity drivers *over* activity cost pools – measurement / specification error on the activity driver effects ^a

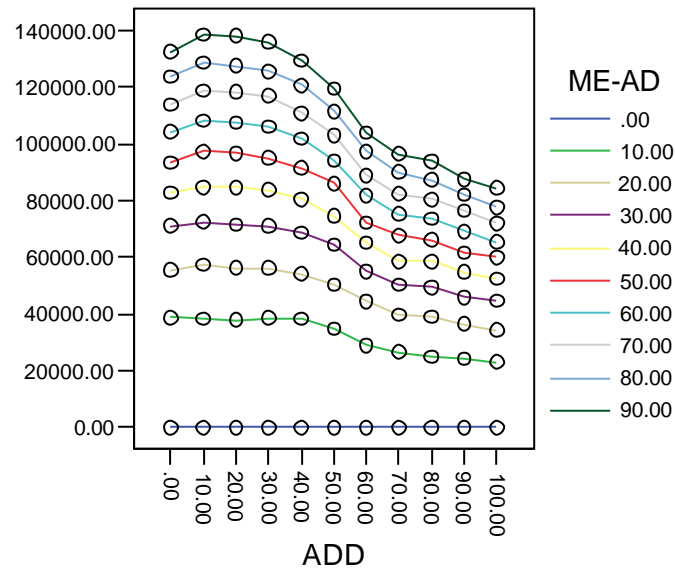
Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression coefficient β ^b	t-value
ADD	5626	0.000*	0.461	-0.306	-190*
ME-AD	45627	0.000*	0.862	0.845	525*
ADD*ME-AD	145	0.000*	0.165	-0.143	-89*
R-Squared	0.879			0.829	

Panel C: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with ADD = 10	Sub-sample with ADD = 10
ADD	0.033 (0.000*)	-0.318 (0.000*)
ME-AD	0.918 (0.000*)	0.842 (0.000*)

Panel B: Marginal Estimated Means Plot of EUCD



Panel D: High diversity rule test

	Less diverse sub-sample	More diverse sub-sample	99% confidence interval on difference between B's
Sub-sample with ADD = 50			
B (unstandardized regression coefficient on ME-AD)	1329.63	942.179	-387.45 ± 11.30

* indicates significance at the 1% level. ^a EUCD = square root of mean squared error, ADD = variance in the distribution of activity drivers *over* activity cost pools, ME-AD = measurement / specification error on activity drivers.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

^c Spearman correlations are consistent with these.

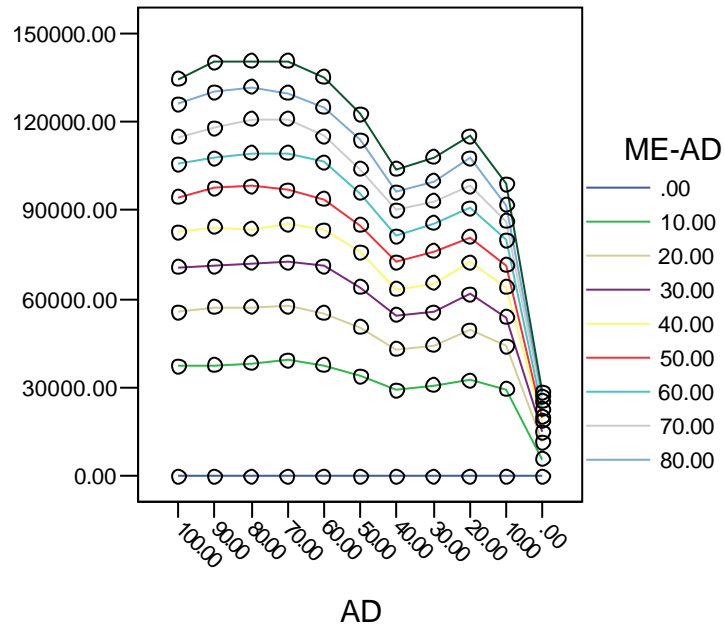
TABLE 4

Analysis of the number of activity drivers – measurement / specification error on the activity driver effects^a

Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression coefficient β ^b	t-value
AD	11874	0.000*	0.643	0.347	186*
ME-AD	44002	0.000*	0.857	0.787	420*
AD*ME-AD	317	0.000*	0.302	0.171	91*
R-Squared	0.892			0.769	

Panel B: Marginal Estimated Means Plot of EUCD



Panel C: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with AD = 80 and ME-AD = 40	Sub-sample with AD = 80 and ME-AD = 50	Sub-sample with AD = 40 and ME-AD = 80	Sub-sample with AD = 20 and ME-AD = 40	Sub-sample with AD = 20 and ME-AD = 20
AD	-0.017 (0.178)	-0.089 (0.000*)	0.184 (0.000*)	-0.086 (0.000*)	0.582 (0.000*)
ME-AD	0.897 (0.000*)	0.638 (0.000*)	0.892 (0.000*)	0.893 (0.000*)	0.624 (0.000*)

Panel D: High diversity rule test

	Less diverse sub-sample	More diverse sub-sample	99% confidence interval on difference between B's
Sub-sample with AD = 50 B (unstandardized regression coefficient on ME-AD)	1367.879	915.676	-452.20 ± 15.79

* indicates significance at the 1% level.

^a EUCD = square root of mean squared error, AD = number of activity drivers, ME-AD = measurement / specification error on activity drivers.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

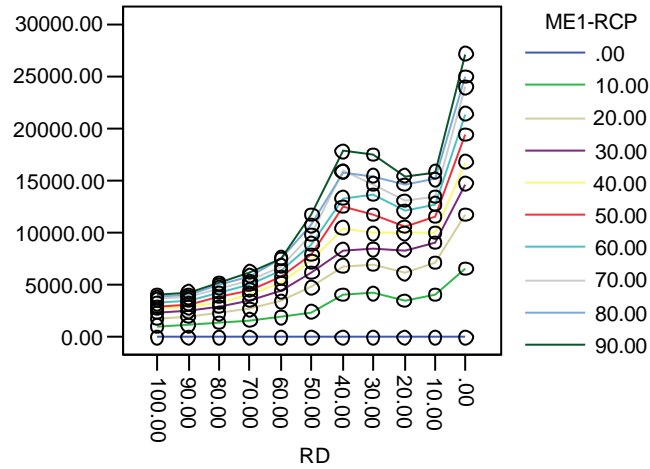
^c Spearman correlations are consistent with these.

TABLE 5: Analysis of the number of resource drivers – measurement error on resource cost pools effects^a

Panel A: Analysis of the EUCD error measure for effects of RD and ME1-RCP for ME2-RCP = 50%

Source of Variation	F	Sig.	Partial h^2	Standardized regression ^b coefficient β	t-value
RD	3829	0.000*	0.368	-0.481	-171*
ME1-RCP	3199	0.000*	0.304	0.438	156*
RD*ME1-RCP	107	0.000*	0.128	-0.233	-83*
R-Squared	0.538			0.477	

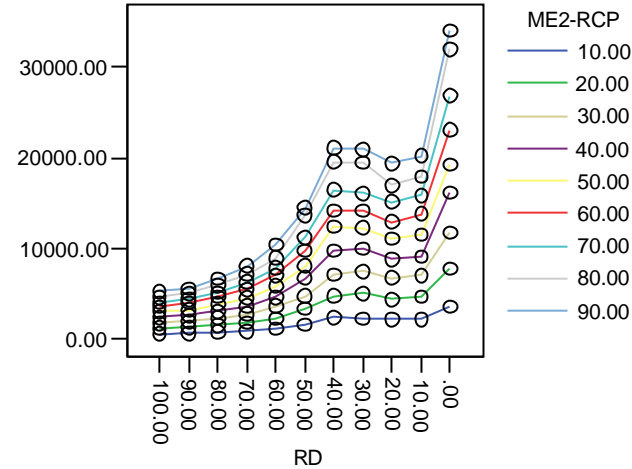
Panel B: Marginal Estimated Means Plot of EUCD



Panel D: Analysis of the EUCD error measure for effects of RD and ME2-RCP for ME1-RCP = 50%

Source of Variation	F	Sig.	Partial h^2	Standardized regression ^b coefficient β	t-value
RD	3467	0.000*	0.369	-0.478	-164*
ME2-RCP	3452	0.000*	0.318	0.459	158*
RD*ME2-RCP	115	0.000*	0.134	-0.244	-84*
R-Squared	0.546			0.499	

Panel E: Marginal Estimated Means Plot of EUCD



Panel C: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with RD = 40	Sub-sample with 20 = RD = 40	Sub-sample with RD = 20
RD	-0.437 (0.000*)	0.050 (0.000*)	-0.293 (0.000*)
ME1-RCP	0.434 (0.000*)	0.549 (0.000*)	0.564 (0.000*)

Panel F: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with RD = 40	Sub-sample with 20 = RD = 40	Sub-sample with RD = 20
RD	-0.437 (0.000*)	0.044 (0.000*)	-0.292 (0.000*)
ME2-RCP	0.461 (0.000*)	0.574 (0.000*)	0.593 (0.000*)

* indicates significance at the 1% level. ^a EUCD = square root of mean squared error, RD = number of resource drivers, ME1-RCP = number of resource cost pools subject to measurement error, ME2-RCP = size of measurement error on resource cost pools. ^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term. ^c Spearman correlations are consistent with these.

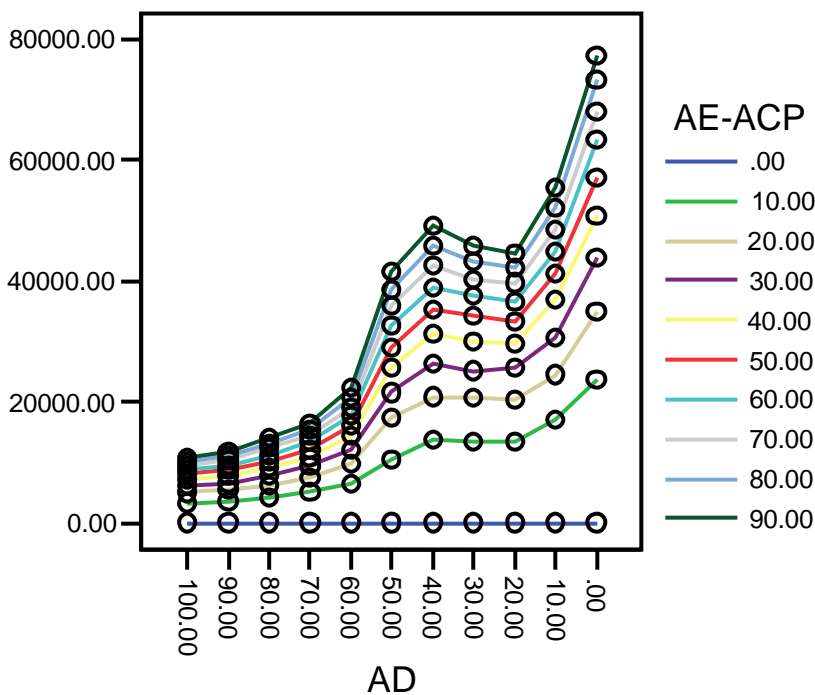
TABLE 6

Analysis of the number of activity drivers – aggregation error on the activity cost pools effects^a

Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression ^b coefficient β	t-value
AD	22344	0.000*	0.772	-0.648	-358*
AE-ACP	15917	0.000*	0.685	0.527	291*
AD*AE-ACP	551	0.000*	0.430	-0.293	-162*
R-Squared	0.863			0.783	

Panel B: Marginal Estimated Means Plot of EUCD



Panel C: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with AD = 40	Sub-sample with 20 = AD = 40	Sub-sample with AD = 20
AD	-0.609 (0.000*)	0.044 (0.000*)	-0.379 (0.000*)
AE-ACP	0.524 (0.000*)	0.815 (0.000*)	0.760 (0.000*)

* indicates significance at the 1% level.

^a EUCD = square root of mean squared error, AD = number of activity drivers, AEACP = aggregation error on activity cost pools.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

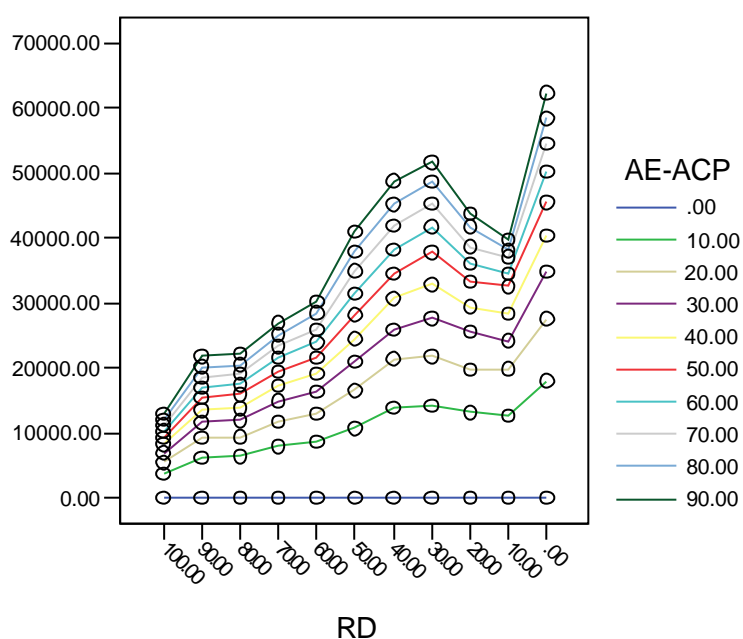
^c Spearman correlations are consistent with these.

TABLE 7
Analysis of the number of resource drivers –
aggregation error on the activity cost pools effects ^a

Panel A: Analysis of the EUCD error measure

Source of Variation	F	Sig.	Partial h^2	Standardized regression coefficient β ^b	t-value
RD	8971	0.000*	0.577	-0.507	-240*
AE-ACP	14426	0.000*	0.663	0.631	299*
RD*AE-ACP	226	0.000*	0.236	-0.225	-107*
R-Squared	0.785			0.706	

Panel B: Marginal Estimated Means Plot of EUCD



Panel C: Correlation analysis for the split samples

Pearson ^c correlation of EUCD with	Sub-sample with RD = 30	Sub-sample with 10 = RD = 30	Sub-sample with RD = 10
RD	-0.531 (0.000*)	0.133 (0.000*)	-0.322 (0.000*)
AE-ACP	0.625 (0.000*)	0.774 (0.000*)	0.758 (0.000*)

* indicates significance at the 1% level.

^a EUCD = square root of mean squared error, RD = number of resource drivers, AE-ACP = aggregation error on activity cost pools.

^b The independent variables in the regression are mean-centered to remove multicollinearity and to scale the interaction term.

^c Spearman correlations are consistent with these.

Appendix A: Use of a distribution measure in our modeling of diversity in resource consumption patterns in costing system design¹⁷

The distribution measure is defined as $u = \frac{\mathbf{a}_w}{\mathbf{a}_{\max}} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{\mathbf{a}_{\max}}$ with $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, the average value of all x_i 's. Consequently, it measures the distribution of all x_i values ($i = 1, \dots, n$) by calculating the total absolute deviations \mathbf{a}_w and \mathbf{a}_{\max} . \mathbf{a}_w measures the total absolute deviation of all x_i values (i.e. (x_1, x_2, \dots, x_n)) from the average deviation $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ by $\mathbf{a}_w = \sum_{i=1}^n |x_i - \bar{x}|$. In this formula \mathbf{a}_{\max} is used to denote the maximal possible value of \mathbf{a}_w . By dividing \mathbf{a}_w by \mathbf{a}_{\max} , we make sure that our distribution measure lies between zero and one inclusive. The maximal deviation \mathbf{a}_{\max} depends on the minimal allowable value (l) and maximal allowable value (u) of each variable x_i and can be shown¹⁸ to be equal to

$$\mathbf{a}_{\max} = \left(u - \bar{x} \right) \left\lfloor \frac{\sum_{i=1}^n x_i - n \cdot l}{u - l} \right\rfloor + \left| l + \left(\sum_{i=1}^n x_i - n \cdot l \right) \bmod (u - l) - \bar{x} \right| + (\bar{x} - l) \left(n - 1 - \left\lfloor \frac{\sum_{i=1}^n x_i - n \cdot l}{u - l} \right\rfloor \right).$$

If no constraining minimal x_i values (i.e. $l = 0$) and no constraining maximal x_i values (i.e. $u = \sum_{i=1}^n x_i$) are imposed, this formula collapses to $\mathbf{a}_{\max} = \left(\sum_{i=1}^n x_i - \bar{x} \right) + (n - 1) \cdot \bar{x}$. This maximal deviation occurs in a situation where one of the x_i 's is at its maximum value of $u = \sum_{i=1}^n x_i$ (first term) and all the other $(n - 1)$ terms are at their minimal value of $l = 0$. From here, also the intuition behind the general formula for \mathbf{a}_{\max} becomes clear. The first term denotes the number of x_i 's that can be put at their maximum value of u . The second term is greater than zero if $\sum_{i=1}^n x_i$ is not fully allocated after putting each x_i to at least the minimal value of l and as many as possible x_i 's to the maximum value of u . The second term then helps to create the maximum variance by allocating the remainder of $\sum_{i=1}^n x_i$ to just one x_i . The third term sets the remainder of the x_i 's to their minimal value of l .

We use this distribution measure to model the diversity indicators related to the differences in the dollar size of the resource cost pools (TRD), to the spread of cost drivers over cost pools across the whole costing system (RDD and ADD) and to the differences in proportional resource consumption at each cost pool (RPD and APD). As in LV, we assume a fixed input value TR to refer to the total resources that need to be allocated to ensure comparability of our simulation results across the various settings. Table A1 summarizes the notation used. Capitals indicate the number of a parameter and small italics with a subscript refer to a particular parameter. As an example, we refer to RCP to denote the number of resource cost pools and we use rcp_i to denote the i^{th} resource cost pool (with $i = 1, \dots, \text{RCP}$). In Table A1, we display the parameter values of our general distribution measure for the calculation of the various costing system diversity indicators.

- Insert Table A1 about here -

The **Total Resource Distribution (TRD)** measures the distribution of the total resources TR over the total number of resource cost pools (RCP). To that purpose, we rely on the distribution measure with $n = \text{RCP}$, $\sum_{i=1}^n x_i = \text{TR}$ and $\bar{x} = \frac{\text{TR}}{\text{RCP}}$. Consequently, \bar{x} denotes the average amount of resources in a resource cost pool and x_i denotes the resources in rcp_i . If no minimal (l) or maximal (u) values are imposed, then $\text{TRD} = 0$ if $x_i = \bar{x}$ for all resource cost pools i . Resources are spread evenly over resource cost pools. If $\text{TRD} = 1$ then $x_i = 0$ for all but one resource cost pool. This remaining resource cost pool has a $x_i = \text{TR}$. In order to guarantee that every resource cost pool contains some resources and the costing system does not collapse into one with fewer resource cost pools (a setting that will be covered by varying RCP in the simulations), we will impose minimal values l for each resource cost pool.

The **number of Resource Drivers (RD)** defines the way in which resource cost pools are linked to activity cost pools. We use RD^{\min} to denote the minimal number of links and RD^{\max} to denote the maximum number of links between resource cost pools and activity cost pools, with

$$\text{RD}^{\min} = \max(\text{RCP}, \text{ACP})$$

$$\text{RD}^{\max} = \text{RCP} * \text{ACP}$$

RD is at its maximum when each resource cost pool provides resources for each activity cost pool. RD is at its minimum when there is a one-to-one relation between resource cost pools

and activity cost pools. If we express the number of links as a percentage ($\%RD \in [0,1]$, i.e. between RD^{\min} and RD^{\max}) then $RD = \lfloor RD^{\min} + \%RD * (RD^{\max} - RD^{\min}) \rfloor$.

The **Resource Driver Distribution (RDD)** over the resource cost pools measures the distribution of number of outgoing resource drivers over the resource cost pools. For the application of the general distribution measure, \bar{x} denotes the average number of resource drivers allocating the resources of a resource cost pool to activities, x_i denotes the number of resource drivers allocating the resources of rcp_i to activities, and $\sum_{i=1}^n x_i = RD$. The maximal value u equals ACP when rcp_i is linked with every activity cost pool. Similarly, l equals 1, denoting that each resource cost pool has to be linked with at least one activity cost pool. We ensure that every resource cost pool has at least one outgoing resource driver and every activity cost pool has at least one incoming resource driver. Applying RDD results in a number of resource drivers rd_i for every rcp_i and hence, $\sum_{i=1}^{RCP} rd_i = RD$.

The **Resource driver Percentage Distribution (RPD)** measures the distribution of percentages of resource allocated per outgoing set of resource drivers at each resource cost pool. Earlier, we determined for each rcp_i a number of resource drivers rd_i . Now we have to assign percentages to these resource driver links such that the total sum equals 100%, as we assume that arithmetical errors arising when one “forgets” to allocate some of the do not occur.¹⁹ Again, we rely on our distribution measure for each rcp_i with $n = RD_i$, $\sum_{i=1}^n x_i = 100$, $\bar{x} = \frac{100}{n}$ and $l = 0.1$.

The **number of Activity Drivers (AD)** defines the way in which activities are linked to cost objects. We use AD^{\min} to denote the minimal number of links and AD^{\max} to denote the maximum number of links between activities and cost objects, with

$$AD^{\min} = \max (ACP, CO)$$

$$AD^{\max} = ACP * CO$$

AD is at its maximum when each activity is needed for each cost object. AD is at its minimum when there is a one-to-one relation between activities and cost objects. Similar as for RD, we express the number of links as a percentage ($\%AD \in [0,1]$, i.e. between AD^{\min} and AD^{\max}) then $AD = \lfloor AD^{\min} + \%AD * (AD^{\max} - AD^{\min}) \rfloor$.

The **Activity Driver Distribution (ADD)** measures the distribution of the activity drivers over the activity pools in terms of how many outgoing links per activity cost pool. We use our distribution measure where \bar{x} denotes the average number of activity drivers allocating activity costs to cost objects per activity cost pool, x_i denotes the number of activity drivers allocating the cost of acp_i to cost objects and $\sum_{i=1}^n x_i = AD$. The maximal value u equals CO when acp_i is linked with every cost object. Similarly, l equals 1, denoting that each activity cost pool has to be linked with at least one cost object. In doing so, we ensure that every activity cost pool has at least one outgoing activity driver and every cost object has at least one incoming activity driver. Applying ADD results in a number of activity drivers ad_i for every acp_i .

The **Activity driver Percentage Distribution (APD)** measures the distribution of percentages of activity cost allocated per outgoing set of activity drivers at each activity cost pool. Earlier, we determined for each acp_i a number of activity drivers ad_i . Here we have to assign percentages to these activity driver links. We again assume, for the same reasons as explained earlier, that the total sum equals 100%. Again, we rely on our distribution measure for each acp_i with $n = ad_i$, $\sum_{i=1}^n x_i = 100$, $\bar{x} = \frac{100}{n}$ and $l = 0.1$.

Table A2 shows the working of this parameterization in a few simple examples of a first stage allocation. Second stage allocations are equivalent. Example 2 has resource cost pools that are of an equal dollar size (TRD = 0), while example 3 has the most diversity in the size of resource cost pools (TRD = 1, with the minimum amount of resources allocated to a resource cost pool $l = 1,000$). For a given number of resource drivers, example 1 has the lowest diversity in the distribution of number of outgoing resource drivers over the resource cost pools (RDD = 0; each resource cost pool has two drivers) and of percentage of resource allocated per outgoing set of resource drivers at each resource cost pool (RPD = 0; each of those two drivers consumes 50% of the resource cost pool). RDD is at the highest level of diversity in example 2 where all but one resource cost pools provide resources for just one activity cost pool. RPD reaches the highest level of diversity in example 3 where, for a particular resource cost pool, most of its resources serve one activity cost pool. Only 1% of the resources in each resource cost pool are allocated to each of the other activity cost pools this resource cost pool serves, as this is set as the lower bound l used in the distribution measure for RPD.

- Insert Table A2 about here -

Appendix B: Robustness checks

Other resulting error measures

The body of the paper reports on a symmetric Euclidian distance measure between the true cost and the false costing approximation as the dependent variable, similar to error measures used in Babad and Balachandran (1993), Hwang *et al.* (1993), Homburg (2001) and LV. We run robustness checks with a percent error measure, similar to the one used by Christensen and Demski (1997), a mean absolute error metric, a Mean Squared Error measure as used by Datar and Gupta (1994), and a “materiality” measures as used by LV as alternative dependent variables. The latter tracks the number of cost objects measured without substantial error (i.e. resulting error smaller than 10%). These costing accuracy measures are defined and summarized in panel A of Table B1.

Other costing system shapes

LV choose an “hourglass” shape for their design of a costing system where there are many resource cost pools, fewer activity cost pools and again many cost objects, based on survey and case evidence showing that this is the most common shape in practice.²⁰ We follow LV in choosing this “hourglass” shape for most hypotheses, but have explicitly studied the impact of shape on the robustness of the costing system (details are available from authors on request).

Other costing system designs

Although computational restrictions prevent us from studying more than one diversity indicators (together with one error) at the time, there may be an interaction of the diversity indicator under study with the settings chosen for all other diversity indicators. We therefore run robustness checks for level of diversity in the other diversity indicators. The body of the paper reports on an “average design”, where average values of 0.5 for all diversity indicators not under study in the particular hypothesis are imposed. A “less diverse design” has values for the diversity indicators that tend towards less diversity in resource consumption patterns, whereas a more “more diverse design” has values tending towards more diversity. The parameter values used are summarized in panel B of Table B1.

- Insert Table B1 about here -

In general we find similar results in terms of the sign of the effects and the shape of the marginal estimated means plots for all these robustness checks. Higher resulting error levels are achieved for the more diverse designs, and lower resulting error levels for the less diverse design. The F-values for the diversity indicators in the ANOVA can sometimes vary substantially, but the effects remain significant at the same levels. We report the cases where these robustness checks did not lead to qualitatively similar results in footnotes to the paper. Tables with marginal estimated means plots, ANOVA and regression results for the more and less diverse designs for each effect are available from the authors on request.

TABLE A1

Use of a general distribution measure to parameterize diversity in resource consumption patterns in costing systems ^a

$$u = \frac{a_w}{a_{\max}} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{(u - \bar{x}) \cdot \left[\frac{\sum_{i=1}^n x_i - n \cdot l}{u - l} \right] + \left| l + \left(\sum_{i=1}^n x_i - n \cdot l \right) \bmod (u - l) - \bar{x} \right| + (\bar{x} - l) \cdot \left(n - 1 - \left[\frac{\sum_{i=1}^n x_i - n \cdot l}{u - l} \right] \right)}$$

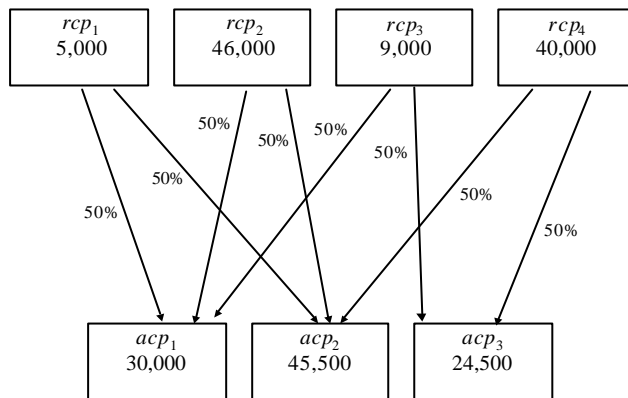
Diversity indicator to model: u	<i>n</i>	$\sum_{i=1}^n x_i$	\bar{x}	<i>l</i>	<i>u</i>
TRD	RCP	TR	TR/RCP	100	TR
RDD	RCP	RD	RD/RCP	1	ACP
RPD	rd _i	100	100/rd _i	0.1	100
ADD	ACP	AD	AD/ACP	1	CO
APD	ad _i	100	100/ad _i	0.1	100

^a Capitals indicate the number of a parameter, small italics with a subscript refer to a particular parameter. RCP = number of resource cost pools, ACP = number of activity cost pools, TRD = distribution of total resources over the resource cost pools, RD = number of resource drivers, RDD = distribution of number of outgoing resource drivers *over* the resource cost pools, RPD = distribution of percentages of resource allocated per outgoing set of resource drivers *at* each resource cost pool, AD = number of activity drivers, ADD = distribution of number of outgoing activity drivers *over* the activity cost pools, APD = distribution of percentages of activity cost allocated per outgoing set of activity drivers *at* each activity cost pool, TR = total resources, *l* = lower bound, *u* = upper bound.

TABLE A2

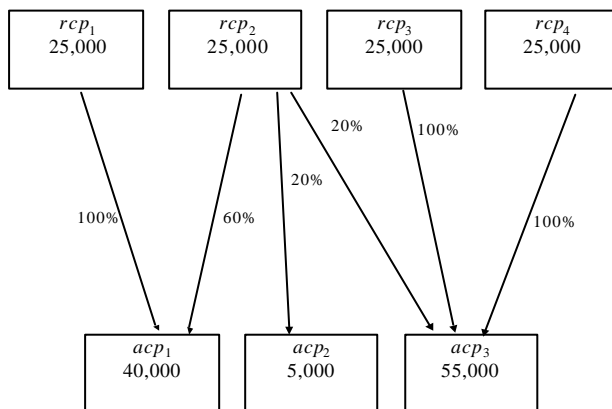
Examples of the use of a general distribution measure to parameterize costing system designs ^a

Panel A: Example 1



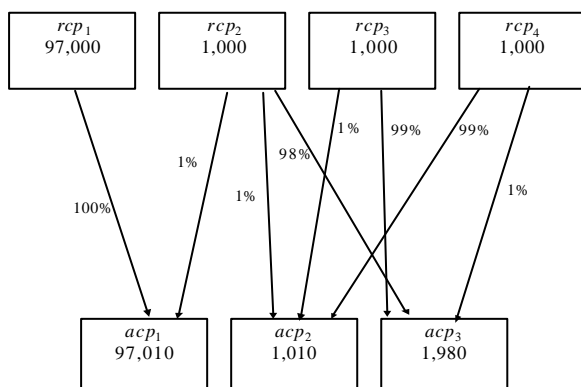
TRD	0.5
RD	8
RDD	0
RPD	0

Panel B: Example 2



TRD	0
RD	6
RDD	1
RPD	0.75

Panel C: Example 3



TRD	1
<i>l</i> for TRD	1,000
RD	8
RDD	0.33
RPD	1
<i>l</i> for RPD	1%

^a Note that these examples focus on the first stage of the costing system. The second stage is modeled in a similar way. $rcp_i = i^{\text{th}}$ resource cost pool, $acp_i = i^{\text{th}}$ activity cost pools, TRD = distribution of total resources over the resource cost pools, RD = number of resource drivers, RDD = distribution of number of outgoing resource drivers over the resource cost pools, RPD = distribution of percentages of resource cost allocated per outgoing set of resource drivers at each resource cost pool, l = lower bound.

TABLE B1: Robustness checks for resulting error measurement and diversity indicator settings

Panel A: Resulting error measurement^a

Measure	Description	Definition	References where used
EUCD	Euclidian Distance	$\sqrt{\sum_{k=1}^{CO} (tc_k - fc_k)^2}$	Babad and Balachandran (1993) Hwang <i>et al.</i> (1993) Homburg (2001) LV
MAE	Mean Absolute Error	$\frac{1}{CO} \sum_{k=1}^{CO} tc_k - fc_k $	n.a.
MPE	Mean Percent Error	$\frac{1}{CO} \sum_{k=1}^{CO} \frac{ tc_k - fc_k }{tc_k}$	Christensen and Demski (1997) LV
MSE	Mean Squared Error	$\frac{1}{CO} \sum_{k=1}^{CO} (tc_k - fc_k)^2$	Datar and Gupta (1994) LV
ACC	Percent of accurate cost objects (10% error allowed)	$\frac{1}{CO} \sum_{k=1}^{CO} \{1 0.95 * tc_k < fc_k < 1.05 * tc_k; 0 \text{ otherwise} \}$	LV

Panel B: Diversity indicator settings for Davg, Dhom and Dhet^b

Diversity indicator ^b	an “average design”	a “less diverse design”	a “more diverse design”
TRD	0.5	0.3	0.7
RD	0.5	0.7	0.3
RDD	0.5	0.3	0.7
RPD	0.5	0.3	0.7
AD	0.5	0.7	0.3
ADD	0.5	0.3	0.7
APD	0.5	0.3	0.7

^a k indexes the number of cost objects ($1..co$), CO is the total number of cost objects, tc_k is true cost accruing to cost object k in the true benchmark scenario, fc_k is the cost allocated to cost object k by the costing system approximation.

^b These settings are used for all diversity indicators other than for the one varied in a controlled way for the study of a particular effect.

¹ Gupta (1993) and Hwang et al (1993) use the term “heterogeneity” which is defined as “diversity or dissimilarity of variable values or characteristics under consideration” (Gupta 1993, 181).

² Gupta (1993) and Hwang et al (1993) have studied this issue at the “product”- and “all products at *one* cost pool”-level respectively, whereas we will take the view of the whole costing system and therefore extend our measurement to all products costed by the costing system. Christensen and Demski (1997) also consider the product portfolio as a whole, but focus on where in the product portfolio large or small errors occur.

³ e.g. Balachandran et al 1997, Balakrishnan and Sivaramakrishnan 2001, Banker and Hansen 2002, Christensen and Demski 1997, Dhavale 2005, LV, Leitch et al 2005.

⁴ Noreen (1991) assumes a one-to-one relationship between RCP and ACP.

⁵ Since results in LV suggest that diversity in the dollar size of the cost objects leads to increased sensitivity to errors, we do not consider that aspect here again.

⁶ Hwang et al (1993) measure diversity in the underlying production technology by the magnitude of the differences in input ratios at cost pools: how does the consumption by product *j* at cost pool *i* relate to the average consumption at cost pool *i* by all products. Gupta (1993) also uses a measure based on the difference between the proportional resource usage of a product across activities and the average proportional resource usage by that product to measure intra-product diversity.

⁷ As a caveat, note that we do not consider other issues that could result in errors in product cost such as jointness or non-linearity (Noreen, 1991).

⁸ Of course, it remains important to continue differentiating between specification error and measurement error on the driver as both have different causes in a case-based setting (wrong drivers vs. a measurement problem with units of allocation bases, respectively) and will therefore require different treatment.

⁹ Computational restrictions (both in CPU times for the running of the simulations as well as for the data analysis capabilities of the statistical software) and interpretational restrictions prohibit us from varying more factors at the same time.

¹⁰ Hypothetically, a 100% error is the maximum error.

¹¹ Because X_1 and X_2 are simulated to be uncorrelated, the mean of the cross product term of their z-scores will also be zero and therefore be itself standardized. As a result, the regression analysis used here is appropriate (Friedrich 1982).

¹² Untabulated results show an exception for an overall error metric that measures the number of cost objects that are accurately costed (allowing for a 10% symmetrical error) ($ACC =$

$$\frac{1}{CO} \sum_{k=1}^{CO} \{1 | 0.95 * tc_k < fc_k < 1.05 * tc_k; 0 \text{ otherwise} \}. \text{ In this case, accuracy improves again when moving to very high}$$

levels of TRD ($\geq 80\%$). Intuitively, when TRD is very high, the probability that the measurement error is introduced on resource drivers that allocate resources from a very small resource cost pool is very high. This will lead to many small errors on cost objects that are not picked up in ACC, which allows for a 10% error before it counts cost objects as materially miscosted.

¹³ The marginal estimated means plot in panel B and the correlation analyses in panel C of Table 3 suggest that only in the extreme case where each cost pool has the same number of outgoing cost drivers ($ADD=0$), this effect is outweighed by the effect of ultra low diversity.

¹⁴ The first column also shows that, for AD very high and ME-AD only low, the occurrence of this offsetting effect is not frequent and sizable enough for the negative correlation between AD and EUCD to become significant.

¹⁵ For a percent error metric ($MPE = \frac{1}{CO} \sum_{k=1}^{CO} \frac{|tc_k - fc_k|}{tc_k}$) this effect disappears and the effect of decreasing

diversity caused by increasing the number of cost drivers always outweighs.

¹⁶ Note again that these tables are representative for all combinations of an error with TRD and APD / RPD, so the high diversity rule of thumb result for these aspects of diversity generalizes across all aggregation, measurement and specification errors.

¹⁷ This distribution measure is closely related to the variability measure developed by Demeulemeester et al (2003) in the context of the generation of project scheduling networks.

¹⁸ A proof is available from the authors on request.

¹⁹ This assumption is less stringent than the one Datar and Gupta (1994, p. 583) make at this stage, where they assume zero measurement error in the total *units* of the cost allocation base at the cost pool level.

²⁰ They report robustness checks for 8 other design “shapes”.