Journal of Business Economics
https://doi.org/10.1007/s11573-024-01210-x

ORIGINAL PAPER

Robust design heuristics for product costing systems: a replication and extension using an ABC cost hierarchy

Matthias Meyer[1] · Mark Schmidt[1]

Accepted: 21 September 2024

Abstract

Accurately reported product costs are essential for many managerial decisions, such as product and resource planning, pricing, product mix decisions, and cost management. Balakrishnan et al. (Manag Sci 57(3):520–541, 2011) offered valuable managerial guidance on how costing system design heuristics affect the accuracy of reported product costs. This paper tests the internal and external validity of their numerical experiments and recommendations. Our replication results show that we are able to reproduce most of the findings, thereby confirming the results' internal validity. To test external validity and assess the robustness of their findings, we modify a key model element by implementing a resource consumption pattern that follows a four-tier Activity-Based Costing (ABC) cost hierarchy, with which we repeat the numerical experiments. Although the design heuristics are mostly robust in this modified environment, the benefits of improving the costing system design are less straightforward and less linear. For instance, single plant-wide cost drivers outperform more information-demanding costing systems with few (2–4) cost pools. Consequently, given a four-tier ABC cost hierarchy, refining the costing system incrementally can reduce accuracy, leaving costing system designers stuck in the middle. Overall, our study supports Balakrishnan et al. (2011) results' robustness but also identifies reasons why firms may still use simple costing systems.

Keywords Product costing · ABC cost hierarchy · Costing system design · Model replication · Internal validity · External validity · Robustness

JEL Classification M40 · M41 · M11 · M10

✉ Mark Schmidt
mark.schmidt@tuhh.de

Matthias Meyer
matthias.meyer@tuhh.de

[1] Institute of Management Accounting and Simulation, Hamburg University of Technology, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany

Published online: 04 December 2024

1 Introduction

Accurate product cost information is vital for decision-making in many domains, such as pricing (Banker and Hughes 1994), product mix decisions (Drury and Tayles 1994), and cost management (Anderson and Dekker 2009). Cost information is derived from measuring the resource consumption[1] of individual products. This, however, requires tracing all individual products' resource consumptions throughout the firms' operations and activities. As this is not only difficult but also often not possible (Balakrishnan and Sivaramakrishnan 2002) firms use costing systems as approximations. These group resources and activities into cost pools and select cost drivers representing the underlying activities (Labro 2019). When doing so, costing system design heuristics can indicate how to reduce the required effort while preserving sufficient accuracy. For example, Kaplan and Cooper (1998) suggested that firms should focus on high-cost resources when forming cost pools and selecting cost drivers ("the *Willie Sutton rule*"). In summary, costing system design heuristics are fundamental since they determine the accuracy with which resource consumption is measured and costs are allocated.

Balakrishnan et al. (2011; hereafter BHL) employed numerical experiments to provide essential insights into the effectiveness of different costing system design heuristics. Their study is the first to compare these design heuristics, which stem primarily from intuition, simplified textbook reasoning (Kaplan and Cooper 1998), and single observations of empirical costing systems (Hwang et al. 1993). The study is interesting from a theoretical perspective, and the results are essential for practitioners designing a costing system. Still, despite the usefulness of the results and recommendations that BHL provided, the simulation model in their paper uses a simple two-tier cost hierarchy distinguishing between volume-level and batch-level resources. However, theory and empirical observations indicate that resource consumption follows more complex and structured patterns.

More specifically, cost accounting research and Activity-Based Costing (ABC) literature often discuss the existence of a four-tier cost hierarchy (Anderson and Sedatole 2013; Cooper and Kaplan 1991). This four-tier cost hierarchy separates unit-level, batch-level, product-sustaining, and facility-sustaining-level costs (Anderson and Sedatole 2013; Banker et al. 2018, 2021; Labro 2004). Following the prior literature, we term this four-tier cost hierarchy the *ABC cost hierarchy*. This results in a resource consumption pattern in which single resources are consumed either more proportionally (i.e., with a more significant positive correlation) or more disproportionally (i.e., with a stronger negative correlation) along a cost hierarchy's different tiers (Noreen 1991). Consequently, resource consumption is not only more complex but also less random. While the complexity could be detrimental to the accuracy of costing systems, the reduced randomness could be utilized by costing system design heuristics to increase the accuracy of reported costs.

[1] In our understanding, resources are required to perform activities (e.g., for employees to do research) and are therefore consumed, which in turn leads to corresponding costs (e.g., salaries). We therefore use the terms resource consumption and usage interchangeably, as well as "activities performed," which we also use as a synonym.

Against this backdrop, the question arises whether the results and recommendations presented in BHL are robust toward such a substantial modification of the resource consumption pattern. This is relevant, as resource consumption is documented to be a key driver of costing system accuracy (Homburg 2005; Labro and Vanhoucke 2008; Schmidt et al. 2023), and the ABC cost hierarchy is typically considered the raison d'être for implementing an ABC system (Cooper and Kaplan 1991). Therefore, this paper aims to test the internal and external validity of BHL's results concerning the design heuristics of costing systems given an ABC cost hierarchy.

We proceed in two steps. First, we replicate the model closely and implement it in a different software environment based on the conceptual description provided by BHL to test internal validity. This establishes a comparable basis for extending the model, while we can also rule out programming and implementation errors (Brüggen et al. 2021). Second, we derive a plausible modeling approach to implement an ABC cost hierarchy in the replicated simulation model's resource consumption generation. This implementation extends BHL's original approach and is based on theoretical predictions and empirical observations from the literature. More precisely, we insert a four-tier ABC cost hierarchy into the modeled production environment's resource consumption, set the respective share of the costs and resources for each tier, and control the resource consumption correlations between the ABC cost hierarchy's different tiers. We, therefore, model a more complex and structured true resource consumption, which the costing system needs to reflect to compute the product costs accurately. By rerunning the numerical experiment in this modified production environment, we conduct a robustness analysis in the form suggested by Grimm and Berger (2016) and test the results' external validity regarding the costing system design heuristics (Libby et al. 2002).

Our close replication's results show that we can replicate most of the costing system design heuristics' selected results with a high alignment to those in the original paper. We found just one exception where the reproduced effects were less profound than those in the original paper. Nevertheless, we conclude that we successfully replicated the selected results from BHL, thereby demonstrating their internal validity.

Regarding the extension, we find that the main conclusions from BHL still hold when implementing the ABC cost hierarchy. However, we also observe that the ABC cost hierarchy affects the performance of different costing system design heuristics. First, we document heuristics that benefit from the more structured resource consumption in the ABC cost hierarchy. Second, we note that when an ABC cost hierarchy is present, very simple costing systems are more likely to outperform more refined costing systems compared to a setting without an ABC cost hierarchy. We suggest that in such a setting, the potential of errors offsetting one another increases. These offsetting effects in cost allocation refer to different resource consumptions depicted within a cost driver, deviating upward or downward in similar magnitudes, canceling one another out, and achieving reasonable accuracy in the final product costs (Datar and Gupta 1994). Improving just one aspect of the costing system (e.g., increasing the number of cost drivers) could result in reduced accuracy because the offsetting effects are diminished (e.g., due to the additional cost drivers introducing new measurements that reduce the canceling out of the upward and downward deviations within one cost driver) (Labro and Vanhoucke 2007). This effect could favor less information-demanding costing system designs.

The paper is structured as follows. Chapter 2 discusses our replication and extension approach and theoretical expectations regarding the ABC cost hierarchy. Chapter 3 describes the simulation model and experimental design. Chapter 4 presents the replication results. Chapter 5 details the ABC cost hierarchy modeling approach. Chapter 6 investigates the ABC cost hierarchy's impact on the system heuristics. Chapter 7 concludes with a summary and discussion of the findings.

2 Replication and extension

We test the internal validity of costing system design heuristics results by first closely replicating the original simulation model within a different software environment (Schmidt et al. 2023; Wilensky and Rand 2007). This helps rule out implementation or programming specifics that influence results (Edmonds and Hales 2003), and determines a comparable basis for the extension toward a four-tier ABC cost hierarchy (Brüggen et al. 2021).

We decided to implement an ABC cost hierarchy as an extension because this hierarchy is considered the raison d'être for implementing an ABC system (Cooper and Kaplan 1991). Additionally, there is a lengthy and ongoing discussion on the relevance, presence, and structure of an ABC cost hierarchy in firms' resource consumption (Anderson and Sedatole 2013; Banker et al. 2021). Although there is a general understanding that resource consumption differs for different resources and activities, it remains undecided which abstract cost hierarchy accurately represents actual resource consumption. More recent accounting theory suggests that resources are consumed in different patterns and are separated into several tiers (Banker et al. 2018), which correlate differently with a firm's activities (Banker et al. 2021). Although several studies support this view (Banker and Johnston 2006), the number, sizes, and correlations of different resource tiers in a cost hierarchy are still debated (Anderson and Sedatole 2013). The BHL model contains a simple two-tier cost hierarchy that separates the volume-level costs from the batch-level ones while assuming that these two tiers are of relatively equal size.[2] In contrast, the ABC cost hierarchy distinguishes between four tiers—unit-level, batch-level, product-sustaining-level, and facility-sustaining-level. Each tier accounts for different fractions of the costs and the resources and entails differently correlated resource consumption patterns.[3] Consequently, resource consumption in the ABC cost hierarchy is more differentiated than in the original model, which design heuristics potentially need to account for. Figure 1 compares the two different cost hierarchies conceptually.

We aim to investigate whether BHL's results of the costing system design heuristics still hold under an ABC cost hierarchy with empirically derived ratios regarding

[2] They model approximately 50% of the resources as batch-level that accounts for 20–50% of the total costs.

[3] The sizes of the tiers in Fig. 1 represent their respective fraction of the total costs (height) or of the resources consumed (width). We derive these different sizes from the empirical observations in Sect. 0.

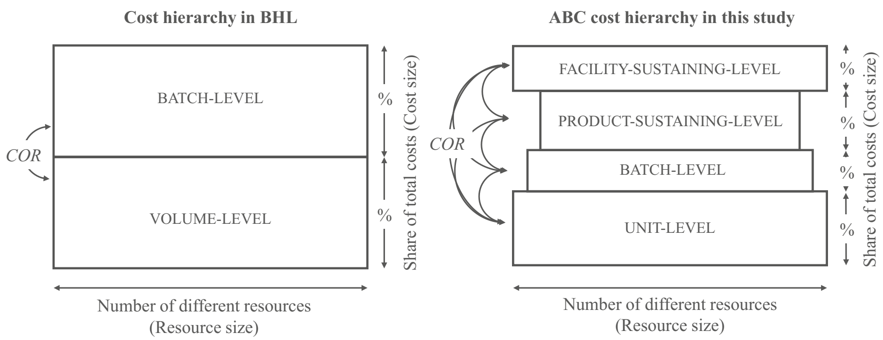

Fig. 1 Different cost hierarchies. *COR* refers to the correlation of resource consumption patterns between the cost hierarchy's different tiers. The number of different resources (Resource size) specifies how many different resources are consumed or activities performed in a particular tier. The share of total costs (Cost size) specifies the fraction of all costs evoked by the resources/activities within a tier. Please note that a high number of different resources does not necessarily imply a high share of total costs

the resource sizes, the cost sizes, and the resource consumption's correlations. We evaluate the results' internal and external validity by focusing on the key costing system design heuristics described in BHL. The overall objective is to examine how a four-tier ABC cost hierarchy affects the efficacy of different costing system design choices compared to a two-tier cost hierarchy.[4]

On average, costing system design literature maintains that more information-demanding costing systems should report product costs more accurately because more information about the true resource consumption is captured instead of approximated. This general rule entails, nevertheless, that more accurate product costs are given at the increasing cost of designing and maintaining a costing system (Labro 2019). The associated results from BHL guide managers to costing system design refinements that lead to the most significant accuracy pay-offs. Additionally, the results indicate exceptions to this rule. These exceptions are especially helpful since they inform designers when less information-demanding and cheaper costing systems provide more accurate product costs than more refined costing systems. By implementing an ABC cost hierarchy, we expect some of these exceptions to hold and increase in significance while others will diminish. We argue that the ABC cost hierarchy increases the true resource consumption's heterogeneity but gives it more structure. Heterogeneity in resource consumption[5] is considered one of the main drivers of errors in reported costs and, therefore, often discussed in the

[4] BHL provided a summary of the costing system design choices' key results in their paper's appendix, which we use as a focal point for our replication analyses. Table 6 provides overviews of these results and explains how the extension affects the different BHL recommendations. We use this approach to guide our analysis, because simulation experiments provide a vast output space (e.g., many variables and sub-experiments are possible), and the results are often too complex to report (Labro 2015). We therefore follow the general suggestion to use patterns to focus our reported results on (Grimm et al. 2005; Heine et al. 2005; Schmidt et al. 2023).

[5] Other studies frame the concept as diversity in resource consumption (Labro and Vanhoucke 2008) or product diversity (Abernethy et al. 2001).

costing system design literature (Christensen and Demski 1997; Gupta 1993; Hwang et al. 1993). Labro and Vanhoucke (2008) define this heterogeneity as a multifaceted construct reflected by "(1) differences in how resources are shared among activities and products across the whole of the costing system, (2) differences in proportional resource usage by activities and products at a particular cost pool, and (3) differences in the dollar size of different cost pools" (Labro and Vanhoucke 2008, p. 1716). We argue that implementing an ABC cost hierarchy addresses aspect (1) of Labro and Vanhoucke (2008)'s definition. The costing system design influences aspects (2) and (3). Since this paper aims to test different costing system designs and heuristics' validity, we focus on a treatment that addresses aspect (1)—heterogeneity in true resource consumption. Textbooks and prior studies define this type of heterogeneity as the degree of different products' differing consumption of resources and activities (Hilton 2011). Consequently, more significant differences in resources' consumption require more information to approximate all products' true consumption and, hence, to allocate costs accurately.

Per our understanding and modeling, an ABC cost hierarchy affects the differences in consumption directly by introducing resources that are consumed in structurally different patterns than other resources (e.g., product-sustaining resources vs. unit-level resources) (Banker et al. 2021). However, by doing so, we argue that it also introduces more structure to a firm's resource consumption. It separates different resources into four tiers and assumes that intra-tier resource consumption is strongly correlated while inter-tier resource consumption is unrelated or negatively correlated (Schmidt et al. 2023). We understand that this characteristic of an ABC cost hierarchy and its corresponding resource consumption evoke two effects that address the traditional intuition that more information-demanding costing systems generate more accurate product costs. First, we expect that design heuristics could exploit the more structured resource consumption and greater information content, which could, in turn, benefit their accuracy-related performance. Second, the increased heterogeneity in resource consumption could increase the probability of errors offsetting each other in simple costing systems (Datar and Gupta 1994). Both effects are relevant not only for the external validity of results concerning costing system design heuristics from BHL but also for providing important information for practitioners.

3 Model description and replication approach

3.1 Model description

The BHL model comprises two main elements: the *benchmark system* and the *costing system*. The *benchmark system* depicts a firm comprising resource costs, produced products, and resource consumption, linking the spent resource costs to individual products. The *costing system* is an aggregated version of the benchmark system aiming at accurately allocating resource costs to individual products. To achieve this, the costing system forms cost pools and selects cost drivers based on limited information retrieved from the benchmark system. The costing system design is based on heuristics resembling rules of thumb, textbook guidance, and intuition when designing a costing system (Balakrishnan et al. 2011). Figure 2 provides a conceptual overview of the model's main elements and functionality. A detailed description of the simulation model appears in the original paper (Balakrishnan et al. 2011).[6]

The model starts by generating *RCP*, the vector containing each resource's costs. First, the user determines the dissimilarity between the resource costs of all resources in *RCP* by defining the input variable *RC_VAR*, which represents the standard deviation of the normal distribution from which the random values in *RCP* are drawn. The greater the chosen value for *RC_VAR*, the greater the differences between all the resources' costs. Next, the model draws a random value from a uniform distribution with boundaries of 0.2 and 0.5. This value determines the share of the total costs assigned to the batch-level resources (e.g., machine set-ups) (*PER_BATCH*). Consequently, batch-level resources consume between 20 and 50% of the total costs. In total, the model generates 50 resources.

Next, the model generates the resource consumption matrix *RES_CONS_PAT*, which determines how individual products consume resources. Specific patterns within this matrix should, therefore, reflect different resource consumption patterns linked to different cost structures or production environments (e.g., flow-shop, work-shop) to generalize the findings across different industries (Balakrishnan et al. 2011). The model generates the *RES_CONS_PAT* matrix using the input parameters *DENS* and *COR*. *DENS* defines the matrix's density. High density resembles a high degree of resource sharing (few zero entries), meaning that the resources are more commonly shared between products, which might be the case in a mass-production line. *COR* defines the correlation of batch-level resource consumption with unit-level resource consumption. The model distinguishes between two tiers in the *RES_CONS_PAT* matrix, which correspond to the volume-level and batch-level resources in the *RCP* vector (see Fig. 1). Half of the modeled resources (i.e., 25) are volume- and batch-level resources, respectively. The greater the *COR* value, the more significant the resource consumption's correlation between the *RES_CONS_PAT* batch-level tier and unit-level tier. Accordingly, if the batch-level resources' consumption patterns correlate negatively (positively) with the volume-level consumption patterns, the true resource consumption would (not) distinguish between the two tiers. In other words, if *COR* has negative values, the modeled cost hierarchy comprises two tiers (see Fig. 1). A consumption pattern for a resource i (i.e., a column in the resource consumption matrix *RES_CONS_PAT*) RC_i is generated as:

$$RC_i = InputCor * BASE + \sqrt{1 - InputCor^2} * RandomValue$$

InputCor resembles the input variable *COR* of batch-level resources, *BASE*, and *RandomValue* are values drawn from a normal distribution for each product. *BASE* resembles a resource that each product consumes, regardless of *DENS*. In the original model, the value for *InputCor* of unit-level resource consumption is set to U[0.2;

[6] The original paper's authors also provided us with a detailed algorithm description. We deeply appreciate their support.

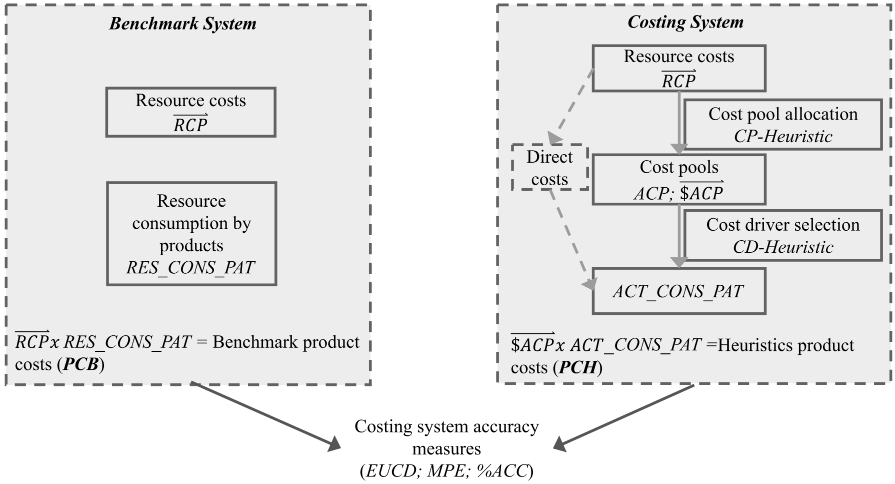

Fig. 2 Conceptual model. $\vec{RCP}$ = Resource costs; *RES_CONS_PAT* = Resource consumption pattern matrix; *PCB* = Benchmark product costs; *CP-Heuristic* = Heuristic used for allocating resource costs to cost pools; $\vec{ACP}$ = Activity cost pools; *CD-Heuristic* = Heuristic used for selecting cost drivers; *ACT_CONS_PAT* = Activity consumption pattern matrix; *PCH* = Heuristics' (reported) product costs; *EUCD* = Euclidean Distance; *MPE* = Mean percentage error; *%ACC* = Share of accurately costed products

0.8], regardless of *COR* and for all settings. Next, the model approximately sets the *DENS* share of entries in *RES_CONS_PAT* to zero and relativizes the remaining non-zero values to reflect the percentage of resource consumption. The model multiplies the resource cost vector *RCP* with the *RES_CONS_PAT* matrix to obtain the true benchmark product costs *PCB*. The sum of all product costs in *PCB* equals the sum of spent resource costs in *RCP*.

The model performs two tasks when generating the costing system: building cost pools and selecting cost drivers. The input variable *ACP* first defines the number of activity cost pools of the costing system. Next, the model employs six different heuristics with increasing degrees of information demand to build the cost pools listed below:

1. *Random*: Resources from *RCP* are randomly distributed among a set number of activity cost pools (*ACPs*), each containing roughly the same number of resources. No specific information is needed for this method.
2. *Size-Random*: Each *ACP* is initially seeded with one of the largest resources from *RCP*. The remaining resources are then randomly distributed across all *ACPs* to ensure each pool has a similar number of resources. This method requires identifying the cost-wise largest resources.
3. *Size-Misc*: The largest resources from *RCP* are placed into *ACP*-1, while the remaining resources go into a miscellaneous pool. This method also requires information about the largest resources.

4. *Correlation-Random*: Each *ACP* is seeded with a randomly selected resource from *RCP*. The remaining resources are allocated based on their correlation to the seed resource, ensuring each pool has a similar number of resources. This method requires information on the correlation between resources, which is information-demanding (Balakrishnan et al. 2011).
5. *Correlation-Size*: Each *ACP* is seeded with one of the largest resources from RCP, and the remaining resources are allocated based on their correlation to the seed resource. This method demands both correlation data and knowledge of the largest resources.
6. *Blended*: Resources are divided into unit-level and batch-level groups. Each group is assigned to half of the *ACPs*. Next, resources are allocated to cost pools using the *size-random* method in each group. This requires prior knowledge of the resource tiers and can be information-demanding due to the complexity of distinguishing between different resource consumption patterns. This differentiation may be easy for the costing system designer in some settings. Contrarily, research is still inconclusive about the relationship between different resources' consumption patterns (Foster and Gupta 1990) and argues that resource consumption is easily misinterpreted when related to production volume (Cooper and Kaplan 1991; Schmidt et al. 2023). Overall, we argue it may be information-demanding to distinguish between different tiers of resource consumption.

The model employs the following heuristics to select cost drivers for each activity cost pool, with increasing information demand, as the resource consumption patterns need to be measured if these are included in the cost driver:

1. *Big-Pool*: Only the cost-wise largest resource in a cost pool is set as the cost driver.
2. *Indexed*: This method averages a previously defined number of cost-wise largest activity cost pool resources as the driver. For instance, Num = 2 means that the average resource consumption of the two cost-wise largest resources in a cost pool is set as the cost driver of that activity cost pool. More resources are included in the cost driver if the value for Num is larger.
3. *Average*: The average resource consumption of all resources in an activity cost pool is set as the driver of that cost pool.

The model generates the activity consumption matrix *ACT_CONS_PAT* from the selected cost drivers. This matrix can be seen as an aggregated *RES_CONS_PAT*, which does not have full information about each resource consumption but only about consumption from cost driver resources. Additionally, the model assumes a measurement error in the cost drivers (Cardinaels and Labro 2008) of between ± 10% and ± 50% of the true resource consumption. Finally, by multiplying *ACT_CONS_PAT* with the dollar value of each *ACP ($ACP)*, the model calculates the heuristics product costs *PCH*, representing the product costs that the costing system reports. The sum of all products' costs *PCH* (or *PCB*) equals the sum of *RCP*.

Balakrishnan et al. (2011) computed three different accuracy measures to compare the true and the reported product costs (*PCB* and *PCH*).

1. *EUCD*: Euclidean Distance between *PCH* and *PCB*, calculated as $\sqrt{\sum_{i=1}^{CO} (PCB_i - PCH_i)^2}$
2. *%ACC*: The materiality measure measures the percentage share of products that are materially under- or overcosted (Labro and Vanhoucke 2007). Kaplan and Atkinson (1998) set this threshold at a 5% deviation from the true costs. *%ACC* is calculated as $\left(\frac{1}{CO}\right)\sum_{i=1}^{CO}\begin{cases} 1 \; if\, 0.95\,PCB_i < PCH_i < 1.05\,PCB_i; \\ 0 \; otherwise; \end{cases}$ (Labro and Vanhoucke 2007).
3. *MPE*: The mean percentage error between *PCH* and *PCB*, calculated as $\left(\frac{1}{CO}\right)\sum_{i=1}^{CO}\frac{|PCB_i - PCH_i|}{PCB_i}$

3.2 Design of numerical experiments

We aim to reproduce the original paper's selected results by rerunning the same experiments with our replicated model. More precisely, we reproduce Figures 1, 2, and 3 from BHL. In addition, we reproduce the original paper's Table 1, as it provides a general overview of the model's functionality and benchmark system and, therefore, serves as a baseline test of the replication's success or failure (results in Table 2).

Table 1 below illustrates the design of our replication approach's experiment by following the recommendations in Lorscheid et al. (2012). We reran the numerical experiment from BHL precisely, with 960 randomly generated benchmark systems and 300 unique costing systems, to obtain 14,400 design points (i.e., input parameter combinations) and 288,000 unique observations. We chose and varied the input variables' parameter values according to BHL. These variables are the leading independent variables of our numerical experiment, while the output variables are the primary dependent variables. We also control for further variables that affect the model's behavior but are not part of our primary analysis.

To evaluate the replication's success, we compare the replicated Table 1's descriptive statistics (Table 2 in our paper) and the replicated figures (Figs. 3, 4, 5 in our paper). We do so by primarily employing relational and distributional equivalence as the replication success criteria (Axtell et al. 1996), by comparing the curve progressions in the relevant figures as well as each observation point's means. We would need further quantitative results to apply other approaches to assess distributional equivalence in more detail (Fachada et al. 2017). However, we follow Schmidt et al. (2023) in the model comparison and argue that a qualitative assessment of replication results is sufficient concerning the design heuristics of costing systems because these provide general guidance instead of quantifications. Moreover, this approach follows the recommendation from Labro (2015) to focus on qualitative results in simulation research. To summarize, achieving relational and distributional equivalence in the replicated table and figures indicates high internal validity for the tested results.

Table 1 Design of experiments for replication following Lorscheid et al. (2012)

Input variables		Control variables		Output variables
Benchmark system		*PER_BATCH* Share of costs assigned to batch-level resources	U[0.2; 0.5]	*EUCD* Euclidean distance between true and reported product costs
RC_VAR Disparity of resource costs	0.25, 0.5, 0.75	*Runs* Number of runs of each input variable combination	20	*%ACC* Share of product costs within the materiality threshold (5%)
DENS Degree of resource sharing	−0.75, 0, 0.75, 1.5	*MSMT_ERR*[a] Measurement error in the cost driver	U[0.1; 0.5]	*MPE* Mean percentage error between the true and reported product costs
COR Correlation between batch-level and unit-level resource consumption	0.33, 0, −0.33, −0.66	*NUMB_PRO* Number of products in the product portfolio	50	
		NUMB_RES Number of resources in RCP	50	
Costing system				
ACP Number of activity cost pools in the costing system	1, 2, 4, 6, 8, 10			
CP-Heuristic Employed cost pool allocation heuristic	*Random*, *Size-Random*, *Size-Misc.*, *Correlation-Random*, *Correlation-Size*, *blended*			
CD-Heuristic Employed cost driver selection heuristic	*Big-Pool*, *Num* = *2*, *Num* = *4*, *Num* = *5*, *Average*			

Design points = 3*4*4*10*6*5 = 14,400; Observations = 14,400*20 = 288,000

[a]Similar to the original paper, we found that the MSMT_ERR's effect is trivial, and therefore omitted it from the replication analysis

4 Close replication results

The results from our close replication of the original BHL model show alignment with the original paper's results since we could generally reproduce the model's output. First, we recompute the original paper's Table 1 and compare the different metrics to our computations (see Table 2). Our replicated model computes near-similar values for all metrics in Table 2. By doing so, we show that (1) the provided explanations for calculating these metrics are sufficient for recalculation and (2) the replicated model's behavior is generally similar to that of the original model. However, numerical equivalence is nearly impossible in stochastic models (Belding 2000). Consequently, we find minor deviations between the two models in nearly all the metrics in Table 2.

These minor deviations do not necessarily affect the overall model's behavior and results but can eventually have larger effects. In particular, we observe deviations in the correlations between the resource consumption of the largest resource and all other resources, as well as between the largest resource consumption and that of batch resources (Table 2A). The replicated model produces weaker positive and stronger negative correlations than the original model. The model description for generating the resource consumption matrix is detailed and unambiguous. We therefore assume that these differences probably stem from using different software packages and random number generators. We contacted the authors from BHL, who gave the same explanation for these differences. We nevertheless conclude that Table 2's reported results are overall relationally and distributionally equivalent.

4.1 Forming cost pools

Next, we compare Figure 1A from BHL and observe a high similarity between the original model's results and our replicated model (see Fig. 3A). More precisely, the different cost pool allocation heuristics' performance order remains, while we observe minor differences in the absolute accuracy (*EUCD*) of certain heuristics. Additionally, the finding that more cost pools generally increase the accuracy holds.

Concerning Figure 1B from BHL, the replicated model does not fully replicate the results of all of the applied cost pool allocation heuristics. While the results of the correlation-based method with an *average* cost driver generally hold, the other methods' *EUCD* differs substantially from that of the original model (see Fig. 3B). More precisely, the *random* and *size-random* methods are more accurate than in the original model, while the method *size-misc* is notably less accurate than in the original model. To interpret this finding, it should be noted that to achieve this result, we had to conduct an additional experiment[7] to compute the required share of the costs in the top 20% of the resources because the experiment that BHL describe only produces more evenly distributed resource costs.[8] When we contacted the original paper's authors, they told us they had increased the input parameter RC_VAR even further than in the experiment in

[7] This additional experiment only affects the results in Figs. 3B and 7. All the other experiments were conducted as described in Table 1.

[8] See Table 2A. The top 20% of resources (i.e., the largest ten resources) have an average share of 34% of the total costs.

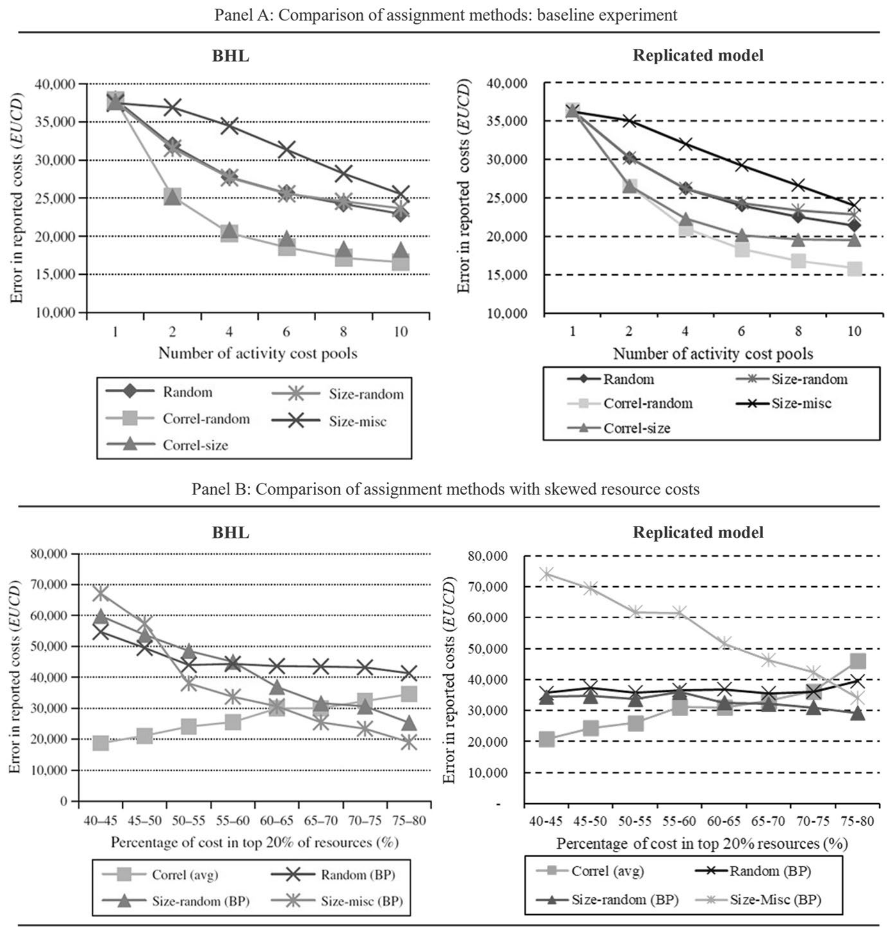

Fig. 3 Replication of Figure 1 from BHL

Table 1 to achieve the required skewness of resource costs. We followed their guidance and set the *RC_VAR* to 0.75; 1; 1.25; 1.5; 1.75; 2, while fixing ACP to 10 and leaving all the other parameters as in the baseline experiment (Table 1). Upon asking the original paper's authors for further clarification, they also assumed that the differences between the original and the replicated model probably stem from the deviations in the resource consumption matrix's correlations (Table 2A). Untabulated results substantiate this assumption further and show that increasing the number of cost pools to more than ten increases the two figures' alignment. Accordingly, the resource consumption matrix in the replicated model seems more heterogeneous than in the original model, which means a greater disaggregation of the costing system is required to achieve a similar accuracy. Overall, we only partially replicated Fig. 3B, concluding that when resource costs are disparate, size-based rules are less dominant over correlation-based ones.

Table 2 Descriptive statistics

(A) *Benchmark cost systems—characteristics systematically varied*

Variation resource costs (using parameter RC_VAR)	Units	Study	Global average	Low dispersion (RC_VAR = 0.25)	Med dispersion (RC_VAR = 0.50)	High dispersion (RC_VAR = 0.75)
			N = 960	N = 320	N = 320	N = 320
Percentage of cost in largest pool/percentage of cost in smallest pool	Ratio	BHL	6.78	3.2	5.75	11.39
		REP	6.88	3.29	5.93	11.42
Percentage of cost in top 10 resources	Percent	BHL	34	30	34	39
		REP	34	30	34	37

Density of consumption matrix (using parameter DENS)	Units	Study	Global average	Little sharing of resources (DENS = −0.75)	Medium sharing of resources (DENS = 0)	High sharing of resources (DENS = 0.75)	Very high sharing of resources (DENS = 1.50)
			N = 960	N = 240	N = 240	N = 240	N = 240
Percentage of zero entries in the consumption matrix	Percent	BHL	36.04	70.95	46.11	20.87	6.22
		REP	38.32	75.51	48.73	22.20	6.83
Average number of products consuming a resource	Number (max. = 50)	BHL	31.97	14.52	26.94	39.56	46.88
		REP	30.84	12.25	25.63	38.90	46.58
Average range in consumption of a resource across products (given positive use)	Percent	BHL	11.43	23.22	11.08	6.61	4.79
		REP	11.09	22.06	10.94	6.62	4.75

Table 2 (continued)

Importance of resources devoted to batch activities (using parameter COR)		Study	Global average	Similar consumption patterns (COR = 0.33)	Intermediate consumption patterns (COR = 0.0)	Intermediate consumption patterns (COR = −0.33)	Dissimilar consumption patterns (COR = −0.66)
			$N=960$	$N=240$	$N=240$	$N=240$	$N=240$
Correlation between largest pool and all resources	Number	BHL	0.264	0.376	0.298	0.232	0.149
		REP	0.102	0.202	0.124	0.069	0.013
Correlation between largest pool and batch resources	Number	BHL	−0.029	0.3	0	−0.061	−0.125
		REP	−0.045	0.138	0.000	−0.113	−0.207

(B) *Benchmark cost systems—characteristics not systematically varied*

Characteristic	Unit	Study	Average	Median	Interquartile range
Percentage of resources in pools devoted to batch activities	Percent	BHL	35.07	34.87	14.14
		REP	35.09	35.01	15.51
Correlation between largest pool and volume resources	Number	BHL	29.36	29.6	15.6
		REP	29.07	29.45	18.19

(C) *Error metrics—univariate statistics ($N=17,280$)*

	Study	Mean	Min	Quartile 1	Quartile 2	Quartile 3	Max
EUCD	BHL	28,448	2514	16,237	24,882	36,919	124,155
	REP	33,840	2280	17,928	29,288	45,120	284,562
%ACC	BHL	25.844	0	14	22	34	100
	REP	23.826	0	12	18	30	100
MPE	BHL	16.8	1.47	9.66	14.92	21.82	62.59
	REP	19.93	1.23	10.56	17.44	26.76	187.11

Table 2 (continued)

(D) *Correlation among error metrics (N = 17,280)*

	EUCD		%ACC		MPE	
	BHL	REP	BHL	REP	BHL	REP
EUCD	1	1	−0.745	−0.718	0.955	0.981
%ACC			1	1	−0.769	−0.730
MPE					1	1

N specifies the number of observations in the corresponding sub-sample. Columns or rows marked with REP contain values computed with the replicated model

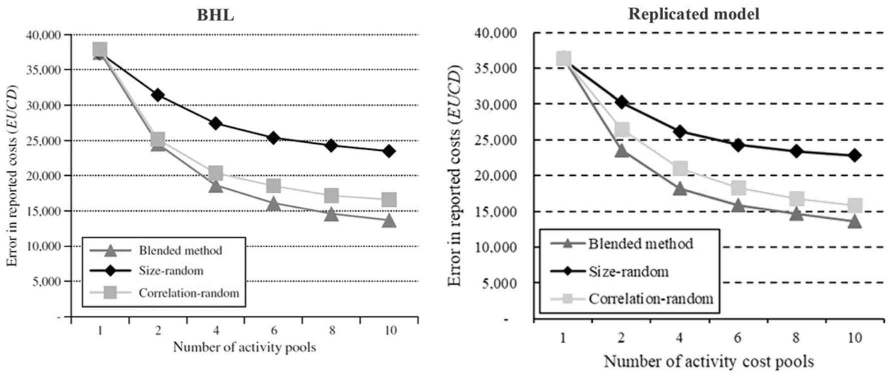

Fig. 4 Replication of Figure 2 from BHL

We can almost reproduce Figure 2 from BHL identically (see Fig. 4). We also find that the *blended* method performs best when grouping the resource costs into cost pools.

4.2 Selecting cost drivers

Figure 5 shows the replication results for Figure 3A and B in BHL. The figures from the replicated model differ slightly from those in BHL. More precisely, the replicated model produces more linear curve progressions for the different cost driver heuristics, and there are larger differences in accuracy between the heuristics when the costing system employs only one cost pool. Potential explanations for this discrepancy could be the different correlations in the resource consumption patterns or more minor deviations in the heuristics' actual algorithm (either the underlying *CP-Heuristic* or the respective *CD-Heuristic*). Although the reproduced results differ slightly from the original paper's results, we argue that the overall qualitative statements remain similar and are replicated.

Overall, the replicated model generally reproduces most of the costing system design heuristics results. The only major exception is Fig. 3B, where the results do not align for all heuristics, affecting the corresponding result in Fig. 7B. However, since this result requires an undescribed additional experiment, we argue that the other results are not dependent on this. Replication success is achieved through relational equivalence, meaning that the results' qualitative statements remain (Axtell et al. 1996). We also test for distributional equivalence by comparing the quantifications in Table 2, and the figures' observation points. In summary, we evaluate our replication as successful and the results as internally valid.

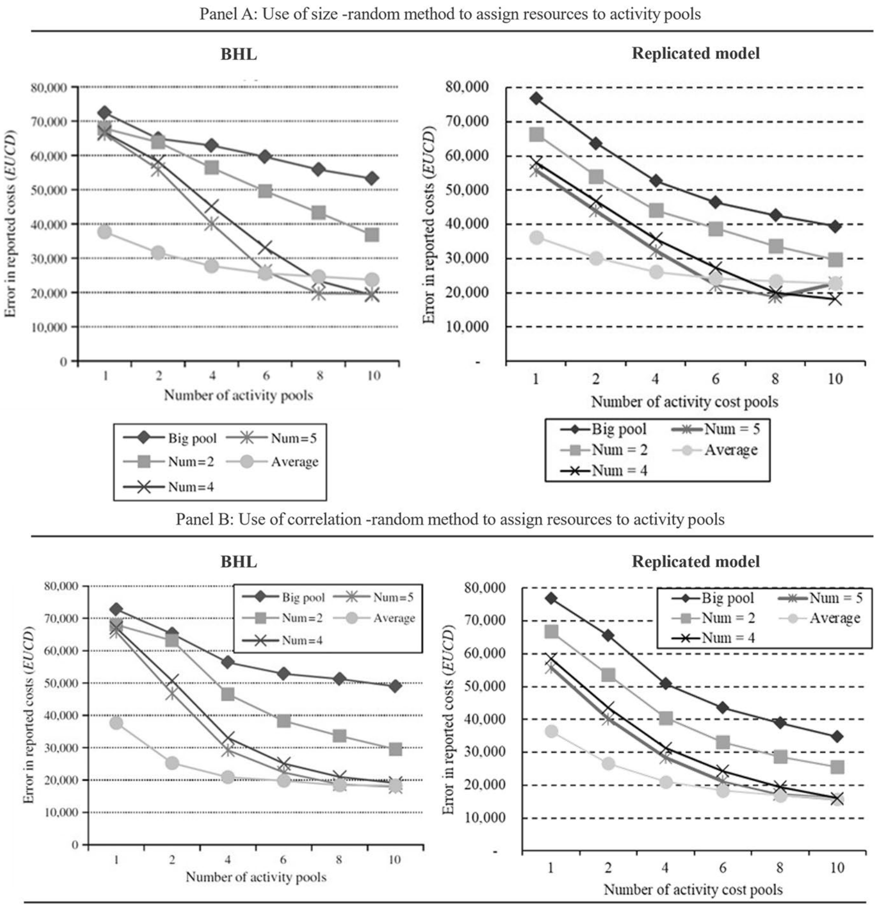

Fig. 5 Replication of Figure 3 from BHL

5 Implementation of an ABC cost hierarchy

A firm's cost hierarchy describes the pattern in which resources are consumed (Anderson and Sedatole 2013) and thus directly affects the costing system design and the errors in the reported costs. The original BHL model models a cost hierarchy with two resource tiers: the volume-level and the batch-level tier. Their resource consumptions are either positively or negatively correlated (see Fig. 1). ABC literature advocates for four tiers: the unit-level, batch-level, product-sustaining-level, and facility-sustaining-level tiers (Anderson and Sedatole 2013). Unit-level resource consumption (and the corresponding cost usage) varies with the number of units produced, for example, machining and assembly (Kaplan and Atkinson 1998). Batch-level resource consumption and costs, like batch set-up costs, vary with the number of batches (Labro 2004). Product-sustaining-level costs occur with activities that are decoupled from the direct production activities and, instead, relate to efforts to design and maintain a unique product (Banker et al. 2021). The number of unique products in a firm's portfolio could potentially be a product-sustaining-level cost driver. Lastly, facility-sustaining costs, such as directors' salaries, are decoupled from the production activities. Owing to its splitting of resources into four tiers, the ABC cost hierarchy assumes that firms' resource consumption is more heterogeneous than a two-tier cost hierarchy (Anderson and Sedatole 2013). The costing system design has to match this different structure to measure the true resource consumption accurately and, hence, to allocate costs accurately (Noreen 1991). Consequently, the ABC cost hierarchy is often proposed to motivate the implementation of more sophisticated ABC systems and to advocate their superiority (Shank and Govindarajan 1988).

Three dimensions determine a cost hierarchy's structure (see Fig. 1): the number of tiers, the correlations within and between the tiers, and the tiers' specific sizes in terms of their costs and resources. These dimensions distinguish the ABC cost hierarchy from the original BHL model's cost hierarchy.

First, the number of tiers defines how many different consumption patterns there are in a firm's cost hierarchy (Cooper and Kaplan 1991). A greater number of tiers probably represents a more diverse overall resource consumption pattern because the number of different activities that need to be captured by the costing system design also increases.

Second, the correlation of the resource consumption between two or more tiers describes whether there is a positive, neutral, or negative correlation between the resources consumed. Resource consumption between tiers often correlates negatively with one another (Cooper and Kaplan 1991). This not only increases the heterogeneity but also results in more structure since different resource consumption patterns can be separated more easily. For instance, several studies analyzing cost drivers investigate whether unit-level activities are proportional to the resource consumption of a firm's manufacturing overhead (MOH) (see Banker and Johnston (2006) for an overview). If this is given, unit-level and MOH resource consumption, which often describe product-sustaining-level and facility-sustaining-level resources (Mertens 2020), are highly correlated. Table 3 reviews the empirically observed and theorized directions of correlations (−; 0; +sign) between the four different tiers of the ABC cost hierarchy.[9]

Third, the size ratio between the ABC cost hierarchy's different tiers determines which tiers should be carefully considered when pooling resources and selecting cost drivers (Cooper and Kaplan 1999). More precisely, the size ratio determines the number of resources or performed activities per tier and its share of the total costs. Unit-level activities might, for instance, be responsible for high costs but actually involve few tasks (e.g., in automated flow-shop production). In contrast,

[9] Schmidt et al. (2023) provide a detailed derivation of Table 3. They also note that empirically observed correlations are potentially distorted, because (1) empirical studies rely on the employed cost drivers, which only approximate the different tiers' actual resource consumption (Labro 2006) and (2) because relating one tier's resource consumption with that of another could lead to the misconception of a positive relation over time (Cooper and Kaplan 1991).

Table 3 Observed correlations between resource consumption of different tiers of the ABC cost hierarchy based on Schmidt et al. (2023)

Tier/tier	Unit-level	Batch-level	Product-sustaining-level	Facility-sustaining-level
Unit-level	1	–	0; +	0
Batch-level	0.07 (Ittner et al. 1997) 0.02–0.82** (Banker et al. 2021) 0.05–0.10 (Ittner and Macduffie 1995)	1	0; –	0
Product-sustaining-level	0.19 (Ittner et al. 1997) 0.43**–0.78** (Banker et al. 2021)	−0.41** (Ittner et al. 1997) 0.30**0.93** (Banker et al. 2021) 0.12** (Datar et al. 1993)	1	0
Facility-sustaining-level	0.57* (Banker et al. 1995) 0.08–0.20 (Ittner and Macduffie 1995)	0.28** (Datar et al. 1993) −0.17 (Banker et al. 1995) −0.30** to 0.20 (Ittner and Macduffie 1995)	0.69** (Datar et al. 1993) 0.44* (Banker et al. 1995)	1

* indicates $p < .05$. ** indicates $p < .01$, as the original studies found. Lower diagonal cells contain empirical observations regarding correlations between different tiers. Some references observe multiple correlations, the absolute range of which is displayed. Upper diagonal cells contain theoretical predictions about the correlations between different tiers

facility-sustaining costs might cover many activities, but the total costs might be lower (Mevellec 2008). Several factors could shape these cost and size ratios. Engineering literature, for instance, often maintains that products' increased modularity tends to increase the product development efforts (i.e., the product-sustaining-level costs) (Mertens et al. 2023), and reduce the unit-level costs in the long term (Fixson 2006; Mertens et al. 2021; Thyssen et al. 2006). Table 4 lists the empirically observed cost, and resource size ratios of an ABC cost hierarchy's different tiers. Although only a few case studies and surveys report the precise ratios, it is evident that the cost and size ratios do not necessarily match (e.g., the batch-level cost share equals 12.4%, while the batch-level size share equals 43%). Consequently, the costing system design's accuracy could differ in tiers in which many resources and activities are combined with a low share of the total costs or vice versa.

We aim to model the ABC cost hierarchy to reflect the theoretical considerations and empirical observations of the ratios and correlations in Tables 3 and 4.

Table 4 Cost and resource size ratios of an ABC cost hierarchy's different tiers

Cost share

References	Total indirect costs				Additional information
	Unit (%)	Non-unit-level			
		Batch (%)	Product (%)	Facility (%)	
Cooper and Kaplan (1991)	52	22	19	7	Case study, $N=1$
Hundal (1997)	89.4	2.6	2.2	5.8	Case study, $N=1$
Al-Omiri and Drury (2007)[b]	41	59			Survey, $N=86$
Kallunki and Silvola (2008)[a,b]	39.1	60.9			Survey, $N=105$
Thyssen et al. (2006)[c]	81.6	9	9.4		Case study, $N=1$
Gunasekaran and Singh (1999)	23.4	15.1	27.2	34.2	Case study, $N=1$
Average	55.3	12.40	16.40	15.29	

Resource share

References	Total number of resources				
	Unit (%)	Non-unit-level			
		Batch (%)	Product (%)	Facility (%)	
Duh et al. (2009)	54	38	–	8	Case study, $N=1$
Ittner et al. (1997)[b]	42	58			Case study, $N=1$
Thyssen et al. (2006)[c]	28.5	28.5	43		Case study, $N=1$
Gunasekaran and Singh (1999)	21	58	7	14	Case study, $N=1$
Average	38	43	7.3	11.7	

The percentage shares of all the tiers are normalized to add up to 100% of the total indirect costs and total number of resources

[a]We exclude the direct costs as in Mertens (2020)

[b]Only unit-level and non-unit-level costs and resources are distinguished in the referenced study

[c]Product-sustaining and facility-sustaining costs and resources are grouped in the referenced study

Table 5 Modeled properties of the ABC cost hierarchy

Tier	Input correlation (*COR*)	Unit	Batch	Product	Facility
Original model (output correlations)					
Unit		1			
Batch	0.33; 0; −0.33; −0.66	−0.16 [0.60]	1		
Product	–	–	–	1	
Facility	–	–	–	–	1
Modeled cost share (%)	50–80	20–50	0	0	
Modeled resource share (%)	50	50	0	0	
ABC cost hierarchy extension (output correlations)					
Unit		1			
Batch	0; −0.33, −0.66	−0.73 [0.15]	1		
Product	0.33	0.48 [0.14]	−0.43 [0.13]	1	
Facility	0	0.00 [0.14]	0.00 [0.14]	0.00 [0.14]	1
Modeled cost share (%)	45	19	17	19	
Modeled resource share (%)	35	46	8.3	9.7	

The values resemble the averaged calculated output correlations between the resource consumption patterns of the two tiers in the respective row and column. The values in brackets resemble the corresponding standard deviation. Note that values are calculated based on the generated *RES_CONS_PAT*, whose computation process is described in Sect. 3

We adjust the resource consumption pattern in the benchmark system by modeling the resource consumption matrix (*RES_CONS_PAT*) and resource cost vector (*RCP*) based on the ABC cost hierarchy. The costing system heuristics remain unchanged when comparing the performance.[10]

We follow the original model's mechanics, splitting *RCP* into different tiers. BHL set about half of the resources as volume-level and batch-level, respectively, with the batch-level resources responsible for 20–50% of the costs (see Fig. 1). In contrast, we separate *RCP* into four tiers with regard to the ABC cost hierarchy (see Fig. 1), with each tier having the approximate share of resources and costs, which the different observations' average indicated in prior studies (see Table 4). Next, we employ a similar approach as in the original model to achieve the desired correlations between the different tiers' resource consumption (see Sect. 3 for the model description).[11] We continue to model the unit-level resource consumption exactly as the volume-level resource consumption in the original model. We reuse the variable *COR* to determine the batch-level resource consumption's correlation from the unit-level resource consumption but draw it from U[0; −0.66]. Consequently, the

[10] An exception is made for the *blended* method, which is slightly adjusted to account for the ABC cost hierarchy's four tiers compared to those of the original model's cost hierarchy. A comparison of the modeling overviews is given in Table 9 in the appendix.

[11] The chosen input values for the correlations and the sizes of costs and resources are based on the empirical observations and theoretical predictions from Tables 3 and 4.

correlation between the batch-level and unit-level resources is negative (Mertens 2020; Misra 1975). To determine the product-level resource consumption, we set *COR* as for unit-level resource consumption to U[0.2; 0.8] but multiply each product-level resource's drawn value with a random factor U[0; 0.5]. Therefore, product-level resources have a lower within and between-tier correlation but correlate positively with unit-level resource consumption. This, on average, approximately equals a value of 0.33 for *COR*, corresponding to the modeling in Schmidt et al. (2023). Finally, facility-level resource consumption is generated randomly to reflect the decoupled resource consumption from unit-level activities. Table 5 provides an overview of the achieved average cost shares, the resource size shares, the *calculated* averaged output correlations (standard deviations in brackets) of the closely replicated model, and the extension model with an implemented ABC cost hierarchy. Additionally, Fig. 6 compares the exemplary cost hierarchies of both modeling approaches, the original model and the ABC cost hierarchy extension.

We understand the ABC cost hierarchy as a more segregated and heterogeneous true resource consumption pattern and expect this modification to affect the costing system design heuristics and the associated results.

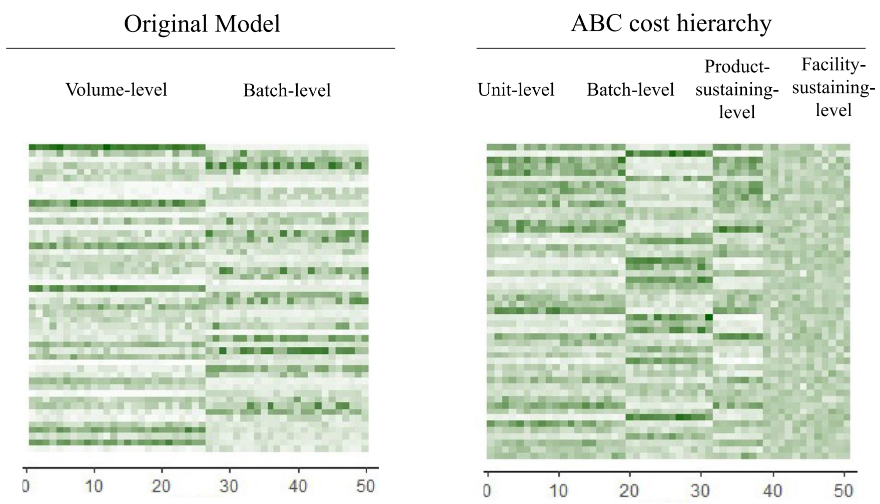

Fig. 6 Exemplary comparison of modeled cost hierarchies. Dark-shaded cells resemble higher resource consumption, and light-shaded cells indicate lower relative resource consumption of the corresponding resource (column) by the corresponding product (row). Consequently, the emerging color patterns imply differently correlated resource consumptions across the different tiers. The original model's respective relative resource sizes are volume-level: 50%; batch-level: 50%. The relative cost sizes are volume-level: 50–80%; batch-level 20–50%. The ABC cost hierarchy model's respective relative resource sizes are, on average, the unit-level: 35%; batch-level: 46%; product-sustaining-level: 8.3%; facility-sustaining-level: 9.7%. The respective relative cost sizes are the unit-level: 45%; the batch-level: 19%; product-sustaining-level: 17%; facility-sustaining-level: 19%. The shown cost hierarchies are exemplary. Cost and resource sizes vary due to random number generation

6 Robustness analysis results

6.1 Forming cost pools

To analyze the validity of the results toward an ABC cost hierarchy, we recompute the replicated figures of our close replication. Starting with Figure 1 in BHL, Fig. 7A shows that introducing an ABC cost hierarchy into the resource consumption pattern affects the overall performance of all cost pool allocation heuristics as the *EUCD* increases on average (note the different y-axes). Hence, more cost pools are required to achieve similar accuracy levels due to the increased heterogeneity in resource consumption.

For Figure 1B in BHL, Fig. 7B shows that the performance of the *size-misc* heuristic with a *Big-Pool* cost driver further decreases compared to the other heuristics. However, it is the only heuristic that profoundly increases accuracy with increasing resource cost disparity. Nevertheless, contrary to the original model, the heuristic *size-random* with a *Big-Pool* cost driver outperforms the *size-misc* heuristic. Arguably, as we could not reproduce the results precisely as in BHL, this result may also be driven by the close replication. However, Fig. 7B illustrates that the heuristic *size-random* outperforms the heuristic *size-misc* in the ABC cost hierarchy setting even more. In this case, distributing the low-cost resources across the cost pools leads to more accurate product costs than pooling them into one miscellaneous cost pool. Finally, the reported product costs' accuracy when using the heuristic *correlation-random* with an *average* cost driver decreases when the disparity in resource costs increases. This is in line with the original model.

We test the external validity of the results concerning the *blended* method (Figure 2 in BHL) by rerunning the same experiment with an ABC cost hierarchy and by providing an *updated blended* method. The original *blended* method distinguishes between volume-level and batch-level resources when grouping resources into cost pools. This differentiation does not apply to the ABC cost hierarchy, as it separates resources into four tiers instead of just two. We adjusted the *blended* method to this new structure by calculating which two tiers' resource consumption patterns have the highest correlation. We grouped these into one group and the other two tiers' resources into the second group. After that, we applied the *blended* method to these two groups, as this was done in the original model. Our results in Fig. 8 illustrate that this grouping (i.e., the *updated blended* method) shows the highest accuracy, as in the original model.[12] Conversely, Fig. 8 shows that all heuristics' overall performance decreases (note the different y-axes) since more cost pools are required to achieve similar accuracy. This again shows that the more heterogeneous resource

[12] We understand that a heuristic that distinguishes between four groups should perform even better, since it approximates true resource consumption most accurately by incorporating as many tiers as possible in the implemented ABC cost hierarchy. The modeling of this heuristic would, however, come with certain exceptions, especially when the number of cost pools is odd or not dividable by the number of groups (tiers). Since it is not this paper's objective to develop further costing system design heuristics, we focus on our simple, updated blended method, which echoes the same differentiation idea, and guide further research to test these more complex heuristics.

consumption pattern in the ABC cost hierarchy setting requires more information-demanding costing system designs to report accurate product costs. In comparison, Fig. 8 also plots the original model's *blended* method, which performs worse than the *correlation-random* method. We apply the original *blended* method to the four-tier cost hierarchy, treating unit-level costs as volume-level and considering all other tiers' costs as batch-level (i.e., non-unit-level). BHL argued that the *blended* method represents an ABC system more closely. In this regard, our results indicate that an ABC system's potential advantage (i.e., distinguishing between different tiers of resources) could become detrimental if the design does not match the underlying cost hierarchy and resource consumption. Additionally, the *correlation-random* method achieves accuracy similar to the *updated blended* method with ten cost pools and outperforms the original *blended* method. An explanation could be that introducing the ABC cost hierarchy also provides resource consumption with more structure, from which the correlation-based assignment benefits. More precisely, the improved structure leads to stronger positive or negative correlations between the different resources' consumption patterns. Grouping resources into cost pools based on their similarity, therefore, results in more homogeneous resource consumption within a cost pool and more heterogeneous resource consumption between the cost pools when a correlation-based assignment is selected.

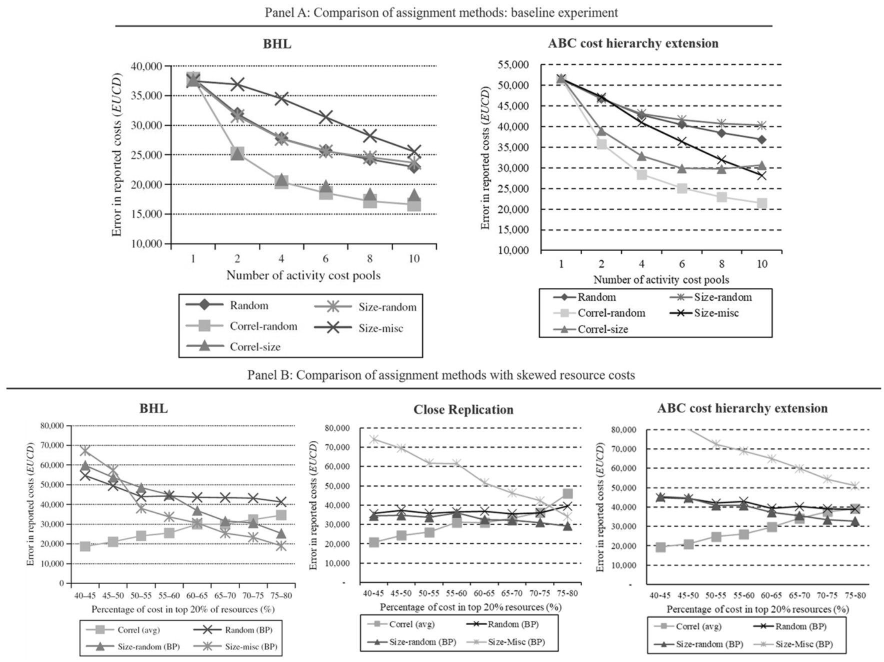

Fig. 7 Robustness analysis of Figure 1 from BHL

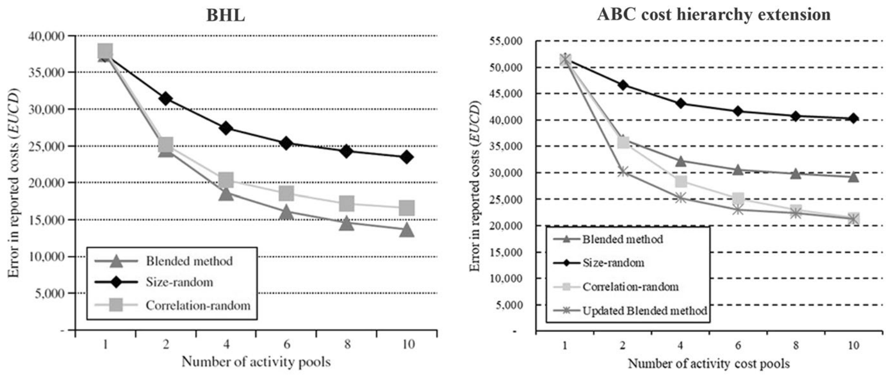

Fig. 8 Robustness analysis of Figure 2 from BHL

6.2 Selecting cost drivers

Figure 9A shows that the differences in accuracy between the cost driver selection heuristics are relatively small compared to those in the original model. Accordingly, selecting a more information-demanding cost driver is less relevant in an ABC cost hierarchy, given that the *size-random* method is used to assign resources to cost pools. When there is only one cost pool in the costing system, the *Big-Pool* method explicitly outperforms all other methods, indicating that a single, plant-wide cost driver could perform well in an ABC cost hierarchy-based resource consumption (Horngren et al. 2015).

Additionally, the accuracy of the *Num4* and *Num5* methods decreases again when the costing system employs eight or more cost pools. This is surprising, as it was generally believed that adding more cost pools increases the accuracy (Balakrishnan et al. 2011). We assume this effect is caused by the decreased offsetting effects of eight or more cost pools. We cannot find any alternative explanation for the methods *Num4* and *Num5*'s behavior, but reconcile our explanation by observing that *Num5* and *average* yield identical accuracy if the costing system has ten cost pools. The latter occurs because, in the modeled firm with 50 resources, a costing system with ten cost pools contains five resources in each cost pool. *Num5* and *average*, therefore, calculate identical cost drivers.

Regarding Fig. 9B, the *Big-Pool* method is again the most accurate concerning a single, cost pool costing system. However, both panel A and B show that the *Big-Pool* method's error increases and peaks regarding a costing system with two cost pools. As in the above results, this is due to offsetting effects when only one cost pool is used. In other words, allocating costs using a single resource's consumption pattern is simple and accurate because the differences in other resources' consumption offset one another. Consequently, adding a second resource's consumption to allocate the costs (e.g., by adding a second cost pool or increasing the information included in the cost driver) could lessen the costing system's accuracy.

Overall, incorporating an ABC cost hierarchy in our modeling approach has significant but different effects on the design heuristics concerning the reported product costs' accuracy. Most of BHL's results remain qualitatively valid in an

ABC cost hierarchy setting, reinforcing their external validity. Conversely, we highlight more exceptions to the general understanding that more information-demanding costing systems also result in fewer errors in the reported product costs. More precisely, we assume the presence of two different ABC cost hierarchy effects on the costing system design rules. First, the ABC cost hierarchy introduces better-structured resource consumption with greater information content, from which heuristics employing this information benefit (e.g., correlation-based heuristics; see Fig. 8). Second, despite being structured, the resource consumption also seems more heterogeneous, with lower or negative correlations between the different tiers' consumption patterns (Schmidt et al. 2023). This benefits errors' offsetting in very simple costing system designs, with these designs achieving higher accuracy (e.g., Fig. 9).

In summary, incorporating an ABC cost hierarchy in our modeling approach affects the design heuristics concerning the accuracy of reported product costs. Most results from BHL remain qualitatively valid in an ABC cost hierarchy setting, reinforcing their external validity. We summarize our findings in Table 6 and relate these to the core results from BHL, as reported in their appendix.

Overall, although BHL already noted exceptions to the general understanding that more information-demanding costing systems also result in lower errors in reported product costs, the ABC cost hierarchy model reports even more cases to which this applies. This challenges costing system design because the accuracy gained from incrementally refining a costing system becomes less linear and predictable.

We further illustrate this by ranking costing systems according to their information demand.[13] A simple system might use one cost pool and a basic cost-driver method, requiring minimal information (e.g., *Big-Pool*). A complex system with multiple cost pools, sophisticated cost pooling, and driver selection requires extensive data. Ideally, more information should improve its accuracy. Thereafter, we compare the relevant changes in the accuracy (i.e., *EUCD*) to the simplest costing system when refining the costing system incrementally toward more information-demanding design heuristics (e.g., adding more cost pools or increasing the number of measured resources for the cost driver). If more information increases accuracy consistently, we would also see a linear decrease in errors. However, very simple systems often benefit more from the ABC cost hierarchy, showing non-linear, and sometimes increased, errors with greater refinement. We test this by varying the costing systems' information-demand scores in two settings (Fig. 10).

In the first setting (Fig. 10A), we include all the design parameters, the number of cost pools, the cost pool allocation heuristics, and the selection of cost drivers in calculating the information demand score. Each additional cost pool requires more information, thereby increasing the information demand score by one point.

[13] Table 7 in our paper's appendix lists the respective scores of each costing system design heuristic. We used the explanations given in BHL for orientation to assign these scores and thereafter varied them to test our findings' robustness. We moreover note that the ranking is based on an ordinal scale.

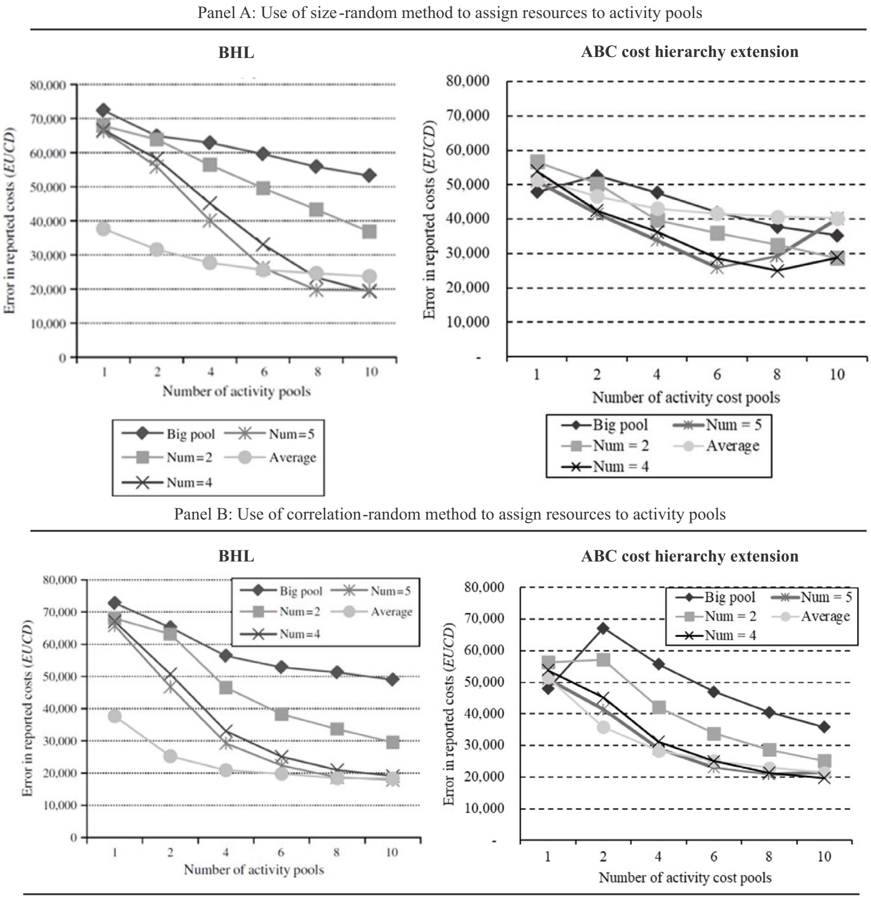

Fig. 9 Robustness analysis of Figure 3 from BHL

Similarly, more complex cost driver selections (*Big-Pool* vs. *Indexed* vs. *Average*) also increase the information demand score. It is less clear which heuristics require more information for the cost pool allocation heuristics. Therefore, Fig. 10B provides an alternative, more robust setting by focusing solely on the number of cost pools and the cost driver selection heuristics. Table 7, in our paper's appendix, explains the respective scores of each costing system design heuristic based on the explanations in BHL.

Figure 10A and B show the gains in accuracy compared to the simplest costing system for both models—the original model and the ABC cost hierarchy extension. First, we observe that the accuracy decreases across all types of costing systems, as the resource consumption in the ABC cost hierarchy model becomes more heterogeneous. Second, both panels show that the gain in accuracy due to the increasing information demand is more concave in the ABC cost hierarchy extension compared to the original model. Although this has been noted previously (Balakrishnan et al.

Table 6 Results overview and summary

Result/finding in BHL	Description (quoted from BHL pp. 540–541)	Results	
		Close replication	ABC cost hierarchy extension
Forming cost pools			
Result P1; Figure 1	*When the distribution of resource costs is moderately skewed (top 20% of costs account for less than 40% of total costs), correlation-based methods dominate size-based methods. When the distribution of resource costs is highly skewed (top 20% of costs account for greater than 75% of total costs), size-based methods dominate correlation-based methods*	Not entirely replicated	The *size-random* method outperforms correlation-based methods, but the *size-misc* method performs worse than correlation-based methods
Result P2; Figure 2	*A blended method that groups resources into tiers and uses a size-based rule within each tier results in an error that is comparable to the error obtained with more information-intensive methods*	Replicated	Pre-grouping similar resources before allocating them to cost pools is also beneficial in the ABC cost hierarchy model. However, the pre-grouping must lead to homogeneous groups. If not, no pre-grouping and correlation-based assignment perform better
Result P4; Figure 2	*For All methods assigning resources to cost pools, it is generally preferable to group the costs of low-cost resources into one pool rather than distribute them over the other pools*	Replicated	This result only holds when resource costs are moderately skewed. If so, a miscellaneous cost pool leads to even higher accuracy than the original model does. We argue that this is due to the increased offsetting effects
Result P5; Figure 1	*A moderate number of cost pools (10–20) seems enough, regardless of the method used to group resources into cost pools. For both size and correlation-based methods, the gain from adding more pools is concave in the number of pools formed*	Replicated	The increased heterogeneity in the ABC cost hierarchy model increases the demand for more cost pools. However, the reduced marginal utility of adding more cost pools does not change. In some settings, more cost pools could even result in worse accuracy
Selecting cost drivers			
Result D1	*In every environment and method for grouping resources into cost pools, an indexed driver is preferred to the "big pool" method of using the consumption pattern for the largest resource*	Replicated	The *Big-Pool* method is able to outperform more information-intensive cost driver selection methods, especially in aggregated costing systems with only one cost pool

Table 6 (continued)

Result/finding in BHL	Description (quoted from BHL pp. 540–541)	Results	
		Close replication	ABC cost hierarchy extension
Result D2; Figure 3	*Indexed drivers (using four or five resources) might even do better than the more information-intensive average driver with a moderate (8–12) number of cost pools*	Replicated	In general, the differences in accuracy between the different cost driver methods could be minor compared to the original model. The cost driver choice is therefore less critical. However, it could be detrimental to incrementally improve the cost driver method, as it can result in reduced accuracy

2011; Friedl et al. 2017), the marginal utility of refining the costing system could be even smaller.

In the original model, there is a relatively linear decrease in *EUCD*, which means incremental refinements in the costing system yield similar payoffs in accuracy. Conversely, in the ABC cost hierarchy model, there is a less linear pay-off for both information-demand scores. More precisely, we find that the accuracy decreases in the first third of the refinements from the simplest system. Consequently, there is a high probability that very simple costing systems outperform subsequent more information-demanding costing systems in the mid-range. Next, incremental refinements pay off for the second third of the information-demanding costing systems. Finally, with respect to the last third of the information-demanding costing systems, additional gains in accuracy decrease as the curves' decline decreases. This especially applies to Fig. 10B.

An explanation for these observations is that implementing the ABC cost hierarchy increases resource consumption's heterogeneity. True resource consumption is divided into four tiers (compared to the original model's two), each with different consumption patterns and contributing differently to the total costs. Overall, this has the potential for more offsetting than in a two-tier hierarchy. Very simple costing systems can benefit by leveraging these offsetting effects (Datar and Gupta 1994). This could enhance accuracy in very simple systems and lead to non-linear payoffs from subsequent refinements. In offsetting the effects of errors, overstated and understated measurements within a cost driver nullify each other (Labro and Vanhoucke 2007). Few studies have addressed these effects. Datar and Gupta (1994) and Gupta (1993) found that such effects are probably present in many costing systems. Labro and Vanhoucke (2007) noted that while refinements usually increase accuracy, the exceptions could be highly aggregated systems with few cost pools. Our findings support the latter presumption, showing that very simple systems might outperform subsequent, more information-demanding ones at times, making accuracy gains through refinements less predictable. From an accuracy perspective, firms

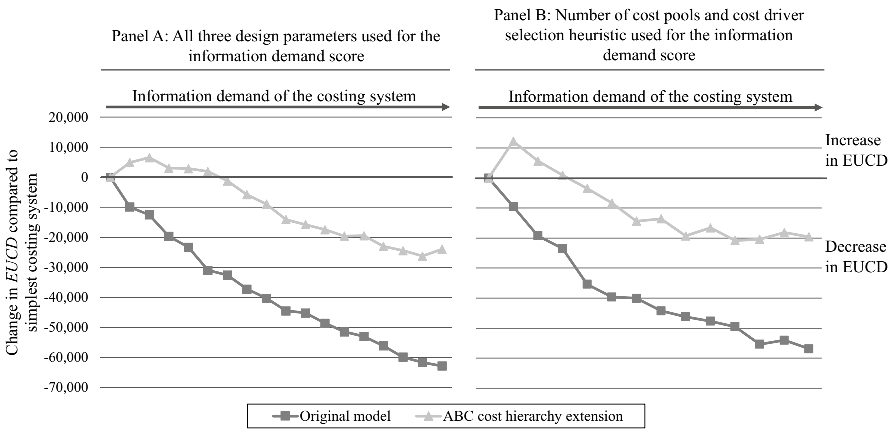

Fig. 10 Change in the cost allocation error (EUCD) when incrementally increasing the costing system's information demand

might, therefore, benefit from maintaining such simple costing systems and avoid being stuck in the middle. Achieving more accurate costing systems requires more than incremental refinements.

7 Conclusion and discussion

This paper aims to test the internal and external validity of costing system design heuristics concerning the accuracy of reported product costs, as reported by BHL. First, to test the results' internal validity and establish a comparable basis for our extension, we closely replicated the simulation model and the original paper's experiments. Our replication results show high internal validity in most findings. We only find one exception where the replicated results are less aligned with those of the original paper.

Second, we tested the external validity by conducting a robustness analysis in which we implemented a four-tier ABC cost hierarchy. This approach changed the resource consumption patterns systematically, which might have affected the costing system design heuristics. The results of our robustness analysis again indicate that many of the original paper's results remain valid. However, we observed recurring effects that challenged some of these results and the corresponding rules. Specifically, the offsetting effects within the costing system—when the ABC cost hierarchy is present—might be more relevant than assumed in the original model (Datar and Gupta 1994; Labro and Vanhoucke 2007). We find that very simple costing system designs could benefit from the ABC cost hierarchy. Our results generally indicate that incremental refinements of a costing system design with only a few cost pools seldom lead to significant gains in accuracy. In addition, certain heuristics such as *correlation-random* and *updated blended* benefit from the more structured resource consumption in the ABC cost hierarchy. Overall, these results frame the external validity of the costing system design heuristics' results.

Our study contributes to the discussion on replication's role in accounting research (Basu and Park 2014; Salterio 2014; Shields 2015). While the debate mainly centers on empirical methods, our study complements it with a computational social science perspective. More specifically, it aims to highlight the importance of testing the validity of simulation experiments' results. Testing internal validity helps establish a comparable foundation, while external validity could be scrutinized in subsequent extensions and robustness analyses (Grimm and Berger 2016; Schmidt et al. 2023).

Moreover, our results add to the discussion of an ABC cost hierarchy's presence (Anderson and Sedatole 2013), since we explain a cost hierarchy's effects on costing system design heuristics. In this regard, we show how costing system design refinements are far less linear than expected, which results in a higher probability of very simple costing systems being more accurate than increasingly information-demanding costing systems. We understand this as an additional potential explanation for the low adoption rate of complex costing systems, such as that of ABC systems (Gosselin 2006), and, in turn, the still high usage of very simple costing systems

(Al-Omiri and Drury 2007; Drury and Tayles 2005), which the literature frames as the "ABC-Paradox" (Cinquini et al. 2015; Gosselin 2006).

Our results also show that the benefits of incremental refinements toward a more complex costing system are less straightforward with an ABC cost hierarchy. Prior accounting research often compares very simple and very complex costing systems with one another (Banker and Potter 1993; Duh et al. 2009; Shank and Govindarajan 1988), neglecting in-between designs. We follow Labro and Vanhoucke (2007) by closing this gap and disentangling in-between costing system designs and incremental refinements' effect. We show how incremental refinements could also reduce accuracy, risking being stuck in the middle. By doing so, our study also contributes to education and practice and warns costing system designers to be cautious when making costly improvements to their costing system design. Finally, we provide an updated comparison between different design heuristics within an ABC cost hierarchy to support the practice of designing costing systems.

Our study also has limitations, which could provide fruitful avenues for future research. First, we emphasize the importance of code and data sharing. Although we could reproduce many findings, some discrepancies in the code and the results could not be fully explained. Second, although we derived our modeling approach from prior analytical and empirical studies, literature on and empirical data about costing systems and resource consumption in firms (including the ABC cost hierarchy) are sparse. Consequently, more empirical research into this avenue would further validate the simulation results. Third, we do not focus on several other ways of testing external validity. For example, engineering literature has observed that different product designs (Davila and Wouters 2004) and product variety management strategies (Erens and Verhulst 1997) shape resource consumption patterns and are therefore likely also to affect costing system design efficacy. Investigating the effects of these concepts on costing system design could, therefore, be a rewarding future research project.

Appendix 1

Compared to the *random* method, the *size-random* and *size-misc* methods require knowledge of the largest resources and their resource consumption. These methods are, therefore, more information-demanding than purely randomly assigned ones. Next, a *correlation-random* method requires correlation measures for different resources' consumption patterns. Additionally, the c*orrelation-size* method also requires knowledge about the largest cost-wise resources. Finally, the *blended* method requires knowledge about the different tiers and consumption patterns regarding the firm's actual resource consumption (Table 7).

With respect to *CD-Heuristics*, the indexed drivers *Num2*, *Num4*, and *Num5* measure an increasing number of resource consumption patterns for constructing the cost driver. Therefore, these patterns are increasingly demanding regarding information, with the *average* driver being the informationally most demanding cost driver heuristic.

Table 7 Information-demand score of each heuristic

CP-heuristic	Score	CD-heuristic	Score
Random	0	Big-Pool	0
Size-random	1	Num2	1
Size-misc	1	Num4	2
Correlation-random	2	Num5	3
Correlation-size	3	Average	4
Blended[a]	4		

[a]We employed the *blended* method as described in BHL's original model, and the updated *blended method* for the ABC cost hierarchy model

Appendix 2

To substantiate the results' robustness depicted in Fig. 10, we calculate two alternative information-demand scores, showing that the qualitative results remain valid regardless of this change. Figure 11 shows these two altered scores' results, while Table 8 reports each heuristic's scores regarding the two additional calculations. Regarding the first alternative calculation, we assume that grouping low-cost resources into a miscellaneous cost pool requires more information than distributing these resources' overall cost pools. In respect of the second alternative calculation, we furthermore assume that allocating resources to cost pools on the basis of correlating resources requires significantly more information than size-based heuristics require. Consequently, we increased the information demand score of these heuristics.

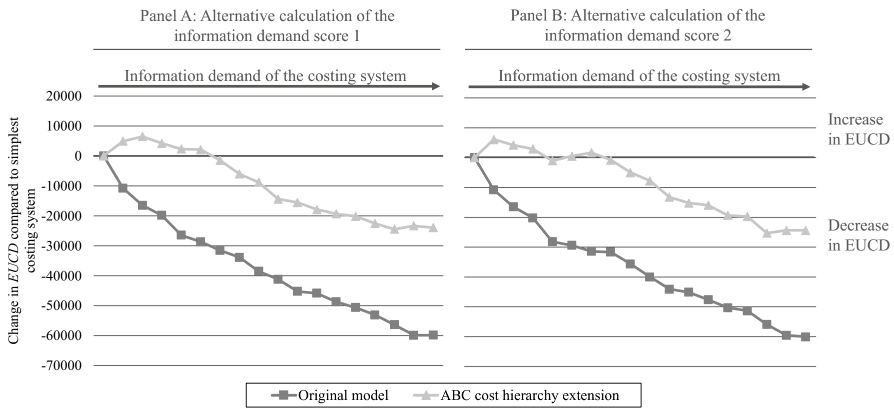

Fig. 11 The information-intensity score's alternative calculations

Table 8 Information-demand score of each heuristic's alternative calculations

CP-heuristic	Score	CD-heuristic	Score
Alternative calculation of information-demand score 1			
Random	0	Big-Pool	0
Size-random	1	Num2	1
Size-misc	2	Num4	2
Correlation-random	3	Num5	3
Correlation-size	4	Average	4
Blended[a]	5		
Alternative calculation of information-demand score 2			
Random	0	Big-Pool	0
Size-random	1	Num2	1
Size-misc	2	Num4	2
Correlation-random	4	Num5	3
Correlation-size	4	Average	4
Blended[a]	5		

[a]We employed the *blended* method as described in BHL's original model and in the *updated blended* method of the ABC cost hierarchy model

Appendix 3

See Table 9.

Table 9 Comparison between the original model and the ABC cost hierarchy extension

Parameter	Original Model/Close Replication	ABC cost hierarchy extension
Benchmark system		
Number of products	50	50
Number of resources	50	50
Number of tiers in the cost hierarchy	2 (volume-level and batch-level)	4 (unit-level, batch-level, product-sustaining-level, facility-sustaining-level)
Separation of the number of resources into tiers (number of entries in *RCP* and *RES_CONS_PAT* add up to a respective number of resources)	Volume-level: 50% Batch-level: 50%	Unit-level: 35% Batch-level: 46% Product-sustaining-level: 8.3% Facility-sustaining-level: 9.7%
Separation of costs into tiers (values in *RCP* add up to the respective share of the total costs [*TC]*)	Volume-level: 50-80% Batch-level: 20-50%	Unit-level: 45% Batch-level: 19% Product-sustaining-level: 17% Facility-sustaining-level: 19%
Costing System		
Cost pool allocation heuristics	*Random* *Size-random* *Size-misc* *Correl-random* *Correl-misc* *Blended*	*Random* *Size-random* *Size-misc* *Correl-random* *Correl-misc* *Updated-blended*
Cost driver selection heuristics	Big-Pool Num = 2 Num = 4 Num = 5 Average	Big-Pool Num = 2 Num = 4 Num = 5 Average

The changed parameters/heuristics between the original model and the ABC cost hierarchy extension are highlighted in grey-shade

Acknowledgements We thank Christian Hofmann and Wolfgang Breuer for their helpful comments on this paper's Extended Abstract and two anonymous reviewers for helpful suggestions on previous versions of this paper. Additionally, we are grateful for Eva Labro's assistance during the replication process.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Code and data availability Code and data for all conducted experiments can be found here: https://doi.org/10.5281/zenodo.13772864.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Abernethy MA, Lillis AM, Brownell P, Carter P (2001) Product diversity and costing system design choice: field study evidence. Manag Account Res 12(3):261–279. https://doi.org/10.1006/mare.2001.0168

Al-Omiri M, Drury C (2007) A survey of factors influencing the choice of product costing systems in UK organizations. Manag Account Res 18(4):399–424. https://doi.org/10.1016/j.mar.2007.02.002

Anderson SW, Dekker HC (2009) Strategic cost management in supply chains, part 1: structural cost management. Account Horiz 23(2):201–220. https://doi.org/10.2308/acch.2009.23.2.201

Anderson SW, Sedatole KL (2013) Evidence on the cost hierarchy: the association between resource consumption and production activities. J Manag Account Res 25(1):119–141. https://doi.org/10.2308/jmar-50293

Axtell R, Axelrod R, Epstein JM, Cohen MD (1996) Aligning simulation models: a case study and results. Comput Math Organ Theory 1(2):123–141. https://doi.org/10.1007/BF01299065

Balakrishnan R, Sivaramakrishnan K (2002) A critical overview of the use of full-cost data for planning and pricing. J Manag Account Res 14(1):3–31. https://doi.org/10.2308/jmar.2002.14.1.3

Balakrishnan R, Hansen S, Labro E (2011) Evaluating heuristics used when designing product costing systems. Manage Sci 57(3):520–541. https://doi.org/10.1287/mnsc.1100.1293

Banker RD, Hughes JS (1994) Product costing and pricing. Account Rev 69(3):479–494

Banker RD, Johnston HH (2006) Cost and profit driver research. In: Chapman CS, Hopwood AG, Shields MD (eds) Handbooks of management accounting research, vol 2. Elsevier, Amsterdam, pp 531–556

Banker RD, Potter G (1993) Economic implications of single cost driver systems. J Manag Account Res 5:15

Banker RD, Potter G, Schroeder RG (1995) An empirical analysis of manufacturing overhead cost drivers. J Account Econ 19(1):115–137. https://doi.org/10.1016/0165-4101(94)00372-C

Banker RD, Byzalov D, Fang S, Liang Y (2018) Cost management research. J Manag Account Res 30(3):187–209. https://doi.org/10.2308/jmar-51965

Banker RD, Gordon SP, Dhinu S (2021) Cost hierarchy: evidence and implications. Glob J Manag Bus Res 21(D2):1–16. https://doi.org/10.34257/GJMBRDVOL21IS2PG1

Basu S, Park H-U (2014) Publication bias in recent empirical accounting research. SSRN Electron J. https://doi.org/10.2139/ssrn.2379889

Belding TC (2000) Numerical replication of computer simulations: some pitfalls and how to avoid them. Center for the Study of Complex Systems. Retrieved from http://arxiv.org/abs/nlin/0001057

Brüggen A, Maas VS, Van den Abbeele A (2021) Replication in experimental management accounting research—an editorial. J Manag Control 32(2):155–160. https://doi.org/10.1007/s00187-021-00324-5

Cardinaels E, Labro E (2008) On the determinants of measurement error in time-driven costing. Account Rev 83(3):735–756. https://doi.org/10.2308/accr.2008.83.3.735

Christensen J, Demski JS (1997) Product costing in the presence of endogenous subcost functions. Rev Acc Stud 2(1):65–87. https://doi.org/10.1023/A:1018373627347

Cinquini L, Collini P, Marelli A, Tenucci A (2015) Change in the relevance of cost information and costing systems: evidence from two italian surveys. J Manage Governance 19(3):557–587. https://doi.org/10.1007/s10997-013-9275-4

Cooper R, Kaplan RS (1991) Profit priorities from activity-based costing. Harv Bus Rev 69(3):130–135

Cooper R, Kaplan RS (1999) Design of cost management systems, 2nd edn. Prentice-Hall, Upper Saddle River

Datar SM, Gupta M (1994) Aggregation, specification and measurement errors in product costing. Account Rev 69(4):567

Datar SM, Kekre S, Mukhopadhyay T, Srinivasan K (1993) Simultaneous estimation of cost drivers. Account Rev 68(3):602–614

Davila A, Wouters M (2004) Designing cost-competitive technology products through cost management. Account Horiz 18(1):13–26. https://doi.org/10.2308/acch.2004.18.1.13

Drury C, Tayles M (1994) Product costing in uk manufacturing organizations. European Accounting Review 3(3):443–470. https://doi.org/10.1080/09638189400000031

Drury C, Tayles M (2005) Explicating the design of overhead absorption procedures in uk organizations. Br Account Rev 37(1):47–84. https://doi.org/10.1016/j.bar.2004.08.003

Duh RR, Lin TW, Wang WY, Huang CH (2009) The design and implementation of activity-based costing. Int J Account Inf Manag. https://doi.org/10.1108/18347640910967726

Edmonds B, Hales D (2003) Replication, replication and replication: some hard lessons from model alignment. J Artif Soc Soc Simul 6(4):1–11

Erens F, Verhulst K (1997) Architectures for product families. Comput Ind 33(2):165–178. https://doi.org/10.1016/S0166-3615(97)00022-5

Fachada N, Lopes VV, Martins RC, Rosa AC (2017) Model-independent comparison of simulation output. Simul Model Pract Theory 72:131–149. https://doi.org/10.1016/j.simpat.2016.12.013

Fixson SK (2006) A roadmap for product architecture costing. In: Simpson TW, Siddique Z, Jiao JR (eds) Product platform and product family design: methods and applications. Springer, New York, pp 305–334

Foster G, Gupta M (1990) Manufacturing overhead cost driver analysis. J Account Econ 12(1):309–337. https://doi.org/10.1016/0165-4101(90)90052-6

Friedl G, Hofmann C, Pedell B (2017) Kostenrechnung: Eine entscheidungsorientierte einführung. Vahlen, Franz

Gosselin M (2006) A review of activity-based costing: technique, implementation, and consequences. In: Chapman CS, Hopwood AG, Shields MD (eds) Handbooks of management accounting research, vol 2. Elsevier, Amsterdam, pp 641–671

Grimm V, Berger U (2016) Robustness analysis: deconstructing computational models for ecological theory and applications. Ecol Model 326:162–167. https://doi.org/10.1016/j.ecolmodel.2015.07.018

Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF et al (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310(5750):987–991. https://doi.org/10.1126/science.1116681

Gunasekaran A, Singh D (1999) Design of activity-based costing in a small company: a case study. Comput Ind Eng 37(1):413–416. https://doi.org/10.1016/S0360-8352(99)00106-0

Gupta M (1993) Heterogeneity issues in aggregated costing systems. J Manag Account Res 5:180–212

Heine B-O, Meyer M, Strangfeld O (2005) Stylised facts and the contribution of simulation to the economic analysis of budgeting. J Artif Soc Soc Simul 8(4):4

Hilton RW (2011) Managerial accounting: creating value in a dynamic business environment, 9th edn. McGraw-Hill/Irwin, New York

Homburg C (2005) Using relative profits as an alternative to activity-based costing. Int J Prod Econ 95(3):387–397. https://doi.org/10.1016/j.ijpe.2004.02.002

Horngren CT, Datar SM, Rajan MV (2015) Cost accounting: a managerial emphasis. Prentice-Hall, Upper Saddle River

Hundal MS (1997) Product costing: a comparison of conventional and activity-based costing methods. J Eng Des 8(1):91–103. https://doi.org/10.1080/09544829708907954

Hwang Y, Evans IIIJH, Hegde VG (1993) Product cost bias and selection of an allocation base. J Manag Account Res 5:213–242

Ittner CD, Macduffie JP (1995) Explaining plant-level differences in manufacturing overhead: structural and executional cost drivers in the world auto industry. Prod Oper Manag 4(4):312–334. https://doi.org/10.1111/j.1937-5956.1995.tb00297.x

Ittner CD, Larcker DF, Randall T (1997) The activity-based cost hierarchy, production policies and firm profitability. J Manag Account Res 9:143–162

Kallunki J-P, Silvola H (2008) The effect of organizational life cycle stage on the use of activity-based costing. Manag Account Res 19(1):62–79. https://doi.org/10.1016/j.mar.2007.08.002

Kaplan RS, Atkinson AA (1998) Advanced management accounting, 3rd edn. Prentice-Hall, Upper Saddle River

Kaplan RS, Cooper R (1998) Cost and effect: using integrated cost systems to drive profitability and performance. Harvard Business School Press, Boston

Labro E (2004) The cost effects of component commonality: a literature review through a management-accounting lens. Manuf Serv Oper Manag 6(4):358–367. https://doi.org/10.1287/msom.1040.0047

Labro E (2006) Analytics of costing system design. Contemporary issues in management accounting. Oxford University Press, Oxford, pp 217–242

Labro E (2015) Using simulation methods in accounting research. J Manag Control 26(2):99–104. https://doi.org/10.1007/s00187-015-0203-4

Labro E (2019) Costing systems. Found Trends Account 13(3–4):267–404. https://doi.org/10.1561/1400000058

Labro E, Vanhoucke M (2007) A simulation analysis of interactions among errors in costing systems. Account Rev 82(4):939–962. https://doi.org/10.2308/accr.2007.82.4.939

Labro E, Vanhoucke M (2008) Diversity in resource consumption patterns and robustness of costing systems to errors. Manage Sci 54(10):1715–1730. https://doi.org/10.1287/mnsc.1080.0885

Libby R, Bloomfield R, Nelson MW (2002) Experimental research in financial accounting. Acc Organ Soc 27(8):775–810. https://doi.org/10.1016/S0361-3682(01)00011-3

Lorscheid I, Heine B-O, Meyer M (2012) Opening the 'black box' of simulations: increased transparency and effective communication through the systematic design of experiments. Comput Math Organ Theory 18(1):22–62. https://doi.org/10.1007/s10588-011-9097-3

Mertens KG (2020) Measure and manage your product costs right—development and use of an extended axiomatic design for cost modeling. Thesis, Technische Universität Hamburg. https://doi.org/10.15480/882.2888

Mertens KG, Schmidt M, Yildiz T, Meyer M (2021) Introducing a framework to generate and evaluate the cost effects of product (family) concepts. Proc Des Soc 1:1907–1916. https://doi.org/10.1017/pds.2021.452

Mertens KG, Rennpferdt C, Greve E, Krause D, Meyer M (2023) Reviewing the intellectual structure of product modularization: toward a common view and future research agenda. J Prod Innov Manag 40(1):86–119. https://doi.org/10.1111/jpim.12642

Mevellec P (2008) Cost systems design. Springer, Berlin

Misra RB (1975) Optimum production lot size model for a system with deteriorating inventory. Int J Prod Res 13(5):495–505. https://doi.org/10.1080/00207547508943019

Noreen E (1991) Conditions under which activity-based cost systems provide relevant costs. J Manag Account Res 3:11

Salterio SE (2014) We don't replicate accounting research—or do we? Contemp Account Res 31(4):1134–1142. https://doi.org/10.1111/1911-3846.12102

Schmidt M, Mertens KG, Meyer M (2023) Cost hierarchies and the pattern of product cost cross-subsidization: extending a computational model of costing system design. PLoS ONE 18(9):e0290370. https://doi.org/10.1371/journal.pone.0290370

Shank JK, Govindarajan V (1988) The perils of cost allocation based on production volumes. Account Horiz 2(4):71

Shields MD (2015) Established management accounting knowledge. J Manag Account Res 27(1):123–132. https://doi.org/10.2308/jmar-51057

Thyssen J, Israelsen P, Jørgensen B (2006) Activity-based costing as a method for assessing the economics of modularization—a case study and beyond. Int J Prod Econ 103(1):252–270. https://doi.org/10.1016/j.ijpe.2005.07.004

Wilensky U, Rand W (2007) Making models match: replicating an agent-based model. J Artif Soc Soc Simul 10(4):1–2