Research Note

Accounting Research for the Management Accounting Profession
Tan Boon Seng*

Abstract

This paper is motivated by the desire to educate practitioners, who are sponsors and consumers of management accounting research, on the fundamental of research. We review the scientific method with a hypothetical example to investigate the drivers of budget gaming using Vroom’s expectancy theory as the theoretical basis.

This review contributes to the debate of the state of management accounting research arising from the Zimmerman paper in the Journal of Accounting and Economics in 2001, and several responses published in the European Accounting Review in 2002. This review is also useful for practitioners and novice researchers, to understand the principles of research and judge the validity of research output.

Keywords
Managerial Accounting Research
Scientific Method
Threats to Validity
Budget Gaming

Introduction

Two papers in the 2001 issue of the Journal of Accounting and Economics sparked a lively debate about the state of managerial accounting research. Ittner and Larcker (2001) reviewed the managerial accounting literature using the Value-Based Management Accounting Framework (Copeland et al, 1996) derived from McKinsey’s consulting practice. Generally, the authors thought that managerial accounting research was practice-oriented and relevant, but the diversity in samples, research methods, and theories makes the results less generalizable. Zimmerman (2001) critiqued the progress in managerial accounting research and conjectured six reasons for the slow progress compared with other areas such as financial accounting. The six reasons are: the lack of reliable and consistent data; the lack of theory in many field-based research projects; incentives of researchers to conduct more practical research and less theoretical research; the literature’s failure to embrace economics that have been productive in other research areas, such as financial accounting, in developing a body of knowledge; few empirically testable theories; the focus on decision making and not on control which is characteristic of accounting systems.

Zimmerman’s paper received several rebuttals in the 2002 issue of the European Accounting Review. Hopwood (2002) recognised much of Zimmerman’s critiques, but argued that a critique on Ittner and Larcker (2001) is not a critique on managerial accounting research. Furthermore, economics may be less appropriate for some research questions such as studying cultural differences. Ittner and Larcker (2002) argued that accounting is an applied discipline and require a practice perspective with a view of improving practice. This is not independent from theory development. Lukka and Mouritsen (2002) and Luft and Shields (2002) both recognised that economic theory is valuable but not superior to other sources such, as sociology or psychology, and the heterogeneity is preferable. The three volumes of the Handbook of Managerial Accounting Research (Chapman, Hopwood and Shields, 2006a; 2006b; 2008) published a few years later allude to a serious effort to link research with a rich source of theoretical frameworks.

* Institute of Singapore Chartered Accountants
The historical development in financial accounting research in the past few decades may predict the future trajectory of managerial accounting research. Financial accounting research before the 1960s mostly argued what “correct” accounting treatment “should” be (i.e., normative theory), but it has shifted over time to examine what accounting treatment “is” in the real or analytical world (i.e., positive theory). Positive accounting theory (Gaffikin, 2007) therefore seeks to establish the empirical and analytical relationships between variables by using the scientific method, and the established relationships form the basis for decisions and policies. Managerial accounting research appears to be following this trend from normative research to positive research backed by empirical findings.

Among this rich source of theoretical frameworks, we argue that economic theory is the most promising – although not the only source – to advance the body of knowledge in managerial accounting in informing practice. The use of economic reasoning in other social sciences – such as using economic tools to study crime, marriage, discrimination, personnel policy and countless other “non-economic” topics – is given the name “economic imperialism” (Lazear, 2000). Lazear explains that the power of economics lies in its rigor in following the scientific method of stating a formal refutable theory, testing theory, and revising the theory based on the evidence, which has been consistently applied over the past half century. Furthermore, there is a trade-off, not generally seen in other social sciences, to allow a high level of abstraction in economic theory. Abstractions strip away complexity and focus economic analysis to produce a rigorous answer, but they also strip away the descriptive richness of the phenomena that may miss out important issues1. Finally, economic theory has a long tradition of positivism which is the norm in financial accounting research today,

and could be the future trajectory for management accounting research.

This paper provides a basic background of the scientific method – which we argue should be the foundation of managerial accounting research independent of the theoretical framework –and the concept of conclusion validity which will be useful in evaluating research. This paper is not a comprehensive review, but we hope it is educational to consumers of research for examining the validity of a research, and useful to sponsors of research for evaluating the merits of a research proposal. We use a hypothetical example to give some concreteness to the discussion.

The Scientific Method

The scientific method refers to the systematic search for cause and effect – i.e., analytical relationships – by engaging the following steps iteratively (i.e., moving back and forth between the steps):

**Observation:** Begin by making an observation, identifying a potential pattern, or considering how an action or event may affect a future action or event. Rich description of a phenomenon is often found in professional journal giving a good context for the observation.

**Theory Development:** Use observation to develop a theory, i.e. an explanation, of what might be causing these actions or events.

**Background Research:** Review past, current, and alternative theories, previously reported evidence, and established tests. Background research is essential to guide the formulation of a useful research question to examine in the investigation phase via an analytical or empirical approach. Background research usually reveals why people are interested in the question and what the debate is about. Multiple competing theories can be discovered. Background research, which is also called a literature review, can generate review papers summarising the current state of the knowledge useful for other researchers.

**Hypothesis Development:** A theory should provide predictions that can be proved wrong with data (i.e., falsified)—a falsifiable prediction is a hypothesis. Hypotheses are

---

1 In the words of Lazear (2000): “Our narrowness allows us to provide concrete solutions, but sometimes prevents us from thinking about the larger features of the problem. This specialisation is not a flaw; much can be learned from other social scientists who observe phenomena that we often overlook. But the parsimony of our method and ability to provide specific, well-reasoned answers gives us a major advantage in analysis”
developed by using logical arguments or a formal economic model. Analytical research, as oppose to empirical research, uses formal modelling to develop hypotheses but does not proceed to empirical testing. Research using logical arguments for the analytical relationship usually proceeds to empirical testing of the analytical relationship.

**Hypothesis Testing:** While the theory developed in step (b) may plausibly explain the observation in step (a), there is no way to tell if the theory is better than competing theories, or which one is the best, until an empirical analysis is done on the predictions derived from the theories (i.e., data are analysed to render a conclusion). The outcome of empirical research is a conclusion, which is claimed to be valid within certain limitations, and is claimed to be useful for answering an important question.

The conclusion validity of empirical research is dependent on three components: construct validity, internal validity, and external validity. Each empirical method entails trade-offs between these three components. We elaborate on the threats to conclusion validity to build the foundation for discussing empirical methods.

**Construct Validity:** Constructs are conceptual ideals of the variables in the analytical relationship. Constructs have to be operationalised, i.e. defined and measured, before data collection can begin. Construct validity refers to how well the data measures the construct. The design of a valid measure for a theoretical construct is a pillar of the foundation for conclusion validity in empirical research.

**Internal Validity:** The conclusions of empirical research can arise from the hypothesised analytical relationship between constructs, or from other factors proposed in alternative theories. The source of theory need not be derived from an economic model built with advanced mathematics as long as the relationships between constructs have been logically argued. The strength of the internal validity depends on the rigor of the theory. It is common in management literature (e.g. strategic management literature) to argue logically, citing and interpreting results from other studies to support the argument, and then theorise a hypothesised relationship between constructs. To claim that a conclusion is valid, an empirical analysis must control for the effects of other factors from alternative theories.

**External Validity:** The usefulness of an empirical result increases when it can be generalised to a larger population. However, external validity decreases when the result is generalised beyond the sample. To reduce threats to external validity, data should be collected from representative samples that are selected using a suitable sampling method. The considerations for proper sampling is quite lengthy to discuss but is well understood.

**Classification of Management Accounting Research**

Following the classification used by the BYU Marriot School for ranking universities, we classify management accounting research into four types based on research methodology: analytical, archival, experimental, and others. Analytical research deals with conceptual problems and is non-empirical. Archival research uses data that are collected from third-party repositories such as the Compustat database, or more commonly in management accounting by surveys. Experimental research randomly divides subjects into two groups, administers a “treatment” to one group but not the other (control) group, and then compares the outcomes in both groups. In a variation of experimental research, called a natural experiment, data collected from a group affected by an event are compared with data from another (control) group that was not affected by the event. Other research method includes a wide variety - the most common ones in financial accounting research are survey and case research. In descending order,

---

4 See http://www.byuaccounting.net/rankings/univrank/rankings.php
the most common methods used in published financial accounting research are archival, experimental, and analytical research. Case research is more common, archival and experiments are less common, in management accounting research than financial accounting research due to historical and data availability reasons.

Strictly speaking, survey is a data collection method and not a research method. A survey uses a questionnaire to collect data consistently for a fresh research question. The data – especially if the survey is a census repeated on a regular basis – then become an archival database useful for answering other research questions in archival research. Experiment can be conducted using surveys and other methods. Furthermore, survey may be used to obtain data just to describe a situation (e.g. survey of preference) and is not research; or for testing a theory-driven hypothesis (e.g. perceived budget manipulation is higher when performance bonus is tied to the budget) and is research.

The key strength of case research is its flexibility to handle rich qualitative data, especially when the theoretical basis of the research question is not well established. Case research can be used in the exploratory phase of research – i.e. the observation, theory building, and background research phase – or on its own as a research method. The main problem with case research is the difficulty in generalising the result. The other problems are the difficulties in interpreting rich data to form constructs, and the limitations of statistical tools to rigorously analyse the data from a series of cases.

Why are there so many methods, and which one is the best for empirical research? In addition to cost and feasibility, research methods are evaluated based on their conclusion validity.

Archival research has become the mainstream method in financial accounting research because of the availability of large financial databases, cheap computing power, and positive network externality. Census data are often available and therefore offer high external validity. However, the generalisation of research results to another country, or another period, can be contested and depends on the context of the study. As the literature using archival research is large, many constructs have already been designed, tested and improved. There is usually strong construct validity. The weakness of archival research is that archival data are not collected for the research, i.e., they are secondary data. Therefore, compared with experimental research, archival research has lower internal validity on controlling for other factors that may produce the result. The use of regression analysis with proper control variables addresses the weaknesses of internal validity in archival research to some extent. In summary, archival research is the mainstream method in scientific research because it is characterised by low data access costs, high construct validity, high external validity, and reasonable internal validity.

Experimental research is associated with high internal validity because of the ability to manipulate controlled conditions\(^5\) within the limits of feasibility (for example, an experiment that induces subjects to commit fraud in a real setting, even if after the experiment, is ethically infeasible). However, the experimental results may not be generalizable beyond the subjects in experimental research because the carefully controlled conditions are often difficult to replicate in a natural environment. The choice between archival research and experimental research thus involves trade-offs.

As there is no one best method given the threats to conclusion validity, replication by using different methods (i.e., triangulation) is ultimately the best research strategy. Practical considerations such as availability of data, mastery of theory and methodology often overwhelm the threat to validity in selecting the method.

**Hypothetical Example**

**Research Topic versus Research Question**

We begin the hypothetical example by first making a distinction between research topic (also known as research area) and research

---

question. It is necessary for a researcher to broadly understand the research topic and possibly even produce a review paper. However, a research topic is too broad for empirical work – a clearly scoped research question is needed. For example, it is impossible to do empirical research on the research topic of budget gaming, but it is possible to review the current debate on the topic to produce a review paper. A research question—what are the drivers of budget gaming—can immediately relate to the analytical literature and produce falsifiable hypotheses, which is what the empirical research aims to test.

A subtlety is worth noting here: empirical result cannot prove that a theory is true when the empirical evidence shows that the predicted relationship is statistically significant; it can only falsify the theory if the empirical result shows otherwise. A theory that survives the falsification process is the accepted wisdom of the day. Therefore, in stating a hypothesis, we begin by stating that there is no relationship between auditor independence and audit tenure (i.e., the null hypothesis is: the theory is “wrong”), and we then conduct an empirical test to falsify the null hypothesis.

If the empirical evidence rejects the null hypothesis, we say that the result is significant (meaning that the theory is “not wrong” about the hypothesised relationship). If the empirical evidence does not reject the null hypothesis, there are two possibilities. First, the theory that produces the hypothesis is wrong. Second, the effect is too small to be detected given the required confidence level and sample size, i.e., we may obtain a significant result with more data. An appropriate method should be used to determine the proper sample size to avoid using some arbitrary sample size.

The Example

This is sometimes called Popper’s cannon of falsification, attributed to the philosopher Karl Popper.

“Not wrong” does not mean “right”.

The minimum sample size to detect the difference is calculated from the statistical power of testing the hypothesis; see http://en.wikipedia.org/wiki/Statistical_power

Firms prepare operating budgets for two reasons. First, budgeting is an integral part of business strategy to help allocate resources, monitor progress, and provide feedback on how well the strategy is working, i.e. strategic control and coordination. Second, the budget is an input to performance measure to motivate behaviour that supports the strategy, i.e. performance measurement. These objectives are related but distinct. However, using the budget for performance measurement has a side effect of budget gaming because the incentive works too well. The consequence of budget gaming is a broken budgeting process for strategic control and coordination because the budget is no longer reliable. An extreme view of addressing the budget gaming problem is to forego using the budget in setting performance target (e.g. the Beyond Budgeting Round Table). Therefore, research that provide insight on the drivers of budget gaming has application for fixing the broken budgeting process and is therefore a significant research problem.

While our description of the scientific method in section 2 suggests that researchers progress from (a) to (e) in a sequential, linear fashion, in reality, the research process involves following the general sequence but frequently moving back and forth in a seemingly chaotic manner as described in the Garbage Can Model (Martin, 1981). However, for clarity, we follow the sequence (a) to (e) below:

Observation: Budget gaming occurs at the points of setting the target and meeting the target. Budget preparation requires forecasting that involve judgments and estimates that produce a range of feasible targets. The budget preparers may prefer to have a conservative number in the budget that is easily achievable – e.g., a low profit or earnings before interest and tax. This practice is known as sandbagging (or paddling) the budget. A more conservative number increases the chance of exceeding the target to receive the incentive. The budget reviewer tries to negotiate to a less conservative number that is deemed fair and feasible, and the dynamics of the negotiation strongly influence the final target. Sandbagging can also occur for expense line in the budget and cost centres are not spared. The budget reviewer is at an informational disadvantage but has the actual result for the previous year. This allows the reviewer to set
the previous year actual plus an adjustment as
target. A strong performer last year was
rewarded then, but is punished with a higher
target this year. The effect of the target
creeping up if it is consistently hit, but not
lowered if the target is missed, is known as
ratcheting. Ratcheting produces a perverse
incentive that the preparer’s best strategy is to
meet the budget and not optimise performance,
giving budget the bad reputation of a fixed
performance contract. Ratcheting is also seen
in the expense line – unnecessary spending
towards the end of the financial year so that the
next year’s expense is maintained.

Budget gaming can produce serious
consequent and make budget useless for
strategic control, and even corrupt a culture of
honesty (Jensen, 2003). Therefore, the question
of what factors drive budget gaming is an
important question.

Theory Development: Vroom’s Expectancy
Theory of motivation (Vroom, 1964) can be
modified to examine the drivers of
sandbagging the budget (here after budget
gaming9). The theory posits that behaviour is
rational, and motivation is strong when effort
leads to performance (expectancy),
performance leads to reward (instrumentality),
and reward leads to satisfaction of
psychological needs (valence). The theory
therefore focuses specifically on what
motivates preparers to sandbag the budget. We
can then look for explanatory variables for the
dependent variable of budget gaming.

Background Research: The theory
development indicates that an empirical
analysis may answer the question regarding the
relationship between budget gaming and the
motivation factors. However, several issues
need to be resolved before proceeding to
collecting data.

First, the data source should be investigated
and the research method should be decided to
ensure feasibility. Given the researcher’s
expertise and resource, we use archival
research and collect data through a survey in
this example.

Second, although we have used Vroom’s
Expectancy Theory as the basis to investigate
budget gaming, there are alternative theories
that can also be plausible. For example,
experienced preparers are more able, and also
more willing, in gaming the budget. However,
more experienced preparers are also better at
forecasting, and have less need for budget
gaming. This means that the relationship
between experience and budget gaming is
likely to be an inverted U shape. The variables
from alternative theories are control variable.

Third, the theory provides the constructs –
budget gaming, expectancy, instrumentality,
valence and experience – and that the level of
budget gaming is the dependent variable,
which is predicted by the other variables as
explanatory variables (according to the theory)
and control variables (according to the
alternative theory). The research problem can
be analysed with regression analysis if we can
measure these constructs from the survey data
collected.

Measuring budget gaming is rather
challenging. It is common to set target for
revenue or profit. Budget gaming for profit is
more complicated because it can occur in
revenue and all the costs. Therefore, we focus
on revenue budget and administer the survey to
sales managers whose performance is
measured by meeting revenue targets. We ask
the question “How do you rate the seriousness
that the revenue budget is sandbagged
(padded)?” with the possible responses as
insignificant (1); significant but the target is
still reliable; (3) very significant and the target
is not reliable. Hence, the dependent variable is
measured using an ordinal scale (ranking
scale), which complicate the regression
analysis.

Expectancy – the construct that effort leads to
performance – is interpreted as the chance that
sandbagging lead to meeting the revenue
target. A significant and unexpected change in
the external environment is usually the root
cause that a sandbagged revenue target is still
missed. The variance of the actual revenue in
the last five years, i.e. volatility, is a proxy of
how likely the external shock can occur
(alternatively, volatility can be measured by a

---

9 This nomenclature is not exactly correct because it
does not include budget ratcheting, but budget
sandbagging is a very clumsy term. Furthermore,
the interaction between ratcheting and sandbagging
can be quite complicated, and is best examine by
methods such as case study.
rating scale if revealing actual revenue is sensitive). Volatility is negatively correlated to expectancy.

Instrumentality – the construct that performance leads to reward – is interpreted as the link that budget performance is considered for promotion and the bonus payment. Therefore, there are several variables for instrumentality. We can ask the question “What is the subjective weight, between 0 and 1, given to budget performance when considering promotion for sales staff?” to obtain the explanatory variable for instrumentality for promotion as the reward. Jensen (2003) discusses the pay-for-performance profile that motivate budget gaming. We investigate if features of the pay-for-performance profile – minimum revenue to receive bonus, bonus cap, and bonus payoff increasing faster at higher revenue (i.e. stretched targets) – lead to budget gaming argued by Jensen.

Valence – the construct that reward leads to satisfaction of psychological needs – is measured by survey questions such as “What is the subjective weight, between 0 and 1, that monetary reward is important to you?” Another valence measure is to replace ‘monetary reward’ by ‘personal satisfaction without monetary reward’ as it is observed people can be motivated by personal satisfaction without monetary reward to surpass set targets.

The construct for experience is proxied by the number of years the survey respondents have prepared revenue budget. This construct is quite straightforward although the theoretical relationship with budget gaming can be complicated – it can be linear (positive or negative correlation) or nonlinear (rising to a maximum then falling, or falling to a minimum then rising). We can accommodate these relations using a quadratic term in the regression equation and use the data to check which relation is correct.

**Hypothesis Development:** The information from the background research can be summarised in the following equation\(^{10}\):

\[
\text{Budget Gaming} = a + b_1(\text{Explanatory Variables}) + c_1(\text{Experience}) + c_2(\text{Experience}^2) + \text{error}
\]

Equation 1: Relationship between Budget Gaming, Explanatory Variables and Control Variables

The intercept is \(a\), the parameters of interest are \(b_i\) and the parameters for the control variables are denoted by \(c_i\). If the expectancy theory is correct, we should find a correlation between budget gaming (the independent variable) and each of the explanatory variables, holding the other variables constant, with the sign predicted by the theory. This (conditional) correlation is given by \(b_i\) which can be tested if it is significantly different from zero (further details are discussed in the Hypothesis Testing section). For example, \(b_1\) is the correlation between budget gaming and expectancy measured by variance of the sales revenue, which the theory predict is negative. The correlation between budget gaming and instrumentality of promotion is given by \(b_2\), which the theory predict is positive.

**Hypothesis Testing:** As the procedure for hypothesis testing is similar for all the variables, we illustrate the procedure for the correlation between budget gaming and instrumentality of promotion, which is the parameter \(b_2\).

We start with the null hypothesis that \(b_2\) is zero and the alternate hypothesis that \(b_2 > 0\). We obtain data to run the regression for equation 1. There are two possible results:

First, we may find that \(b_2\) is indeed not statistically different from zero. We can see from equation 1 that there is no relation between budget gaming and instrumentality of promotion when \(b_2\) is zero. This scenario tends to occur when \(b_2\) is very small, the variation is large, and we want to be very strict in accepting the conclusion the theory is not wrong (i.e., if we say that the theory is not wrong, the chance that it turns out to be wrong is very small). For this case, we cannot reject the hypothesis instrumentality of promotion has no effect on budget gaming, hence there is no support for the expectancy theory.

regression requires using a model such as the multinomial probit or multinomial logit.

\(^{10}\) As budget gaming is measured using an ordinal (ranked) scale, estimating the equation by
Second, we may find that \( b_2 \) is significantly different from zero; thus, we can reject the null hypothesis, and we have a significant result. This scenario tends to occur when \( b_2 \) is very large, when the variation is small, and when we want to be less strict in accepting the conclusion the theory is not wrong. If the empirical results also show that \( b_2 > 0 \), the instrumentality part of the theory cannot be rejected. If the result is statistically significant but the parameter is of the wrong sign \( (b_2 < 0) \), we conclude that the data do not support the theory. Hence, empirical support for the theory requires significant result of the correct sign.

We can also test whether \( b_i \) are jointly statistically different from zero by using the same approach. The procedure is the same although the statistical distribution is different – the single parameter test uses a t-distribution and the joint test uses an F-distribution. We can also test the control variables. If the relationship is linear, \( c_1 \) is not significantly different from zero and \( c_2 \) determines the linear correlation. If the relationship is non-linear, \( c_1 \) is significantly different from zero: \( c_1 \) is positive if there is a minimum point, and \( c_1 \) is negative if there is a maximum point.

**The Myth that Academic Research is not Useful for Practice**

The scientific method is the foundation for academic research, and a common myth is that academic research is not applied research that professional practice need. It is not entirely correct, and this paper contributes to bridging the gap. The academic-professional divide is well documented by Evans, Burritt and Guthrie (2011): academic research is criticized to be inaccessible to professional, and too focused on publishing in academic journal rather than improving accounting practice (Ratnatunga, 2012).

While there is some truth in the criticism, our review shows that accounting research is motivated to solve a significant problem, and the scientific method is to ensure the rigor of the research. Publishing in a peer-reviewed academic journal guarantees that rigor, but has made the research inaccessible to professionals. Therefore, the problem is not the usefulness, but the accessibility, of such research to the accounting profession.

Identifying the correct problem is the first step of solving the problem.

This paper contributes to making research more accessible by equipping the professionals with the basic understanding of the principles of research. The other solution to improve accessibility is to “translate and synthesize” academic research for the benefit of the profession. Evans, Burritt and Guthrie (2011) suggest that professional accounting organisations have an important role to transmit the findings to professionals.

**Concluding Remarks**

Management accounting research, especially those done by academics, is getting more sophisticated to inform practice. This paper can contribute to bridging the academic-professional divide by demystify accounting research to practicing professionals – who may be sponsors and consumers of accounting research – and to students who will become practicing professionals. Furthermore, novice researchers such as graduate students may find this article useful as a review, and educators may consider using this paper as teaching material.

The scientific method is also a useful framework for examining accounting research, although the research process does not proceed sequentially from (a) to (e), as outlined in this paper. The scientific method ensures the rigor of the investigation and is consistent with positive accounting theory. The considerations for conclusion validity – internal validity (arising from theory), construct validity and external validity – are applicable to all research methods including case research. The scientific method has become the mainstream in financial accounting research in recent decades, and we expect the same to happen to management accounting research in future. Understanding these ideas is an important first step for the accounting profession to evaluate accounting research.

**References**


