



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Performance effects of multiple control forms in a Lean organization

A quantitative case study in a systems fit approach

Kristensen, Thomas Borup; Israelsen, Poul

Published in:
Management Accounting Research

DOI (link to publication from Publisher):
[10.1016/j.mar.2013.10.001](https://doi.org/10.1016/j.mar.2013.10.001)

Creative Commons License
Unspecified

Publication date:
2014

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Kristensen, T. B., & Israelsen, P. (2014). Performance effects of multiple control forms in a Lean organization: A quantitative case study in a systems fit approach. *Management Accounting Research*, 25(1), 45-62.
<https://doi.org/10.1016/j.mar.2013.10.001>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Performance effects of multiple control forms in a Lean organization: A quantitative case study in a systems fit approach

Thomas Borup Kristensen, PhD, Aalborg University
Poul Israelsen, Aalborg University

Published in "Management Accounting Research" 2014.
<https://doi.org/10.1016/j.mar.2013.10.001>

Abstract

The primary objective of this paper is to research and test how control forms function and perform in a Lean organization. In the present quantitative case study, we provide statistical support that Lean is a set of multiple control forms (output, behavioral, and social controls) that complement each other to enhance performance, i.e., it is a control package. Therefore, performance is increased if the average level of control forms is increased, and performance is further increased if the control forms are balanced at the same level representing a complementary effect between them. Moreover, we provide a refinement to the statistical approach in testing systems fit models like ours by supplementing the Euclidian distance with the city-block distance. In this way, we are able to show that the control forms in Lean have a balanced complementary effect on performance, which is distinct from a solely additive effect or no effect. The refined understanding of complementary effect between control forms, the notion of balance, in a Lean organization can be utilized in understanding and testing more general control package theory in other contexts. Our data are archival data spanning multiple years in a dedicated Lean organization. This Scandinavian organization has around 2,000 employees and produces small electronic components that are sold to business customers.

Keywords: Lean, Control package, Systems fit, Performance, Statistical method, Control forms, Configuration fit

1. Introduction

Since the introduction of Lean (Womack et al., 1991) in organizations in Europe and the USA, many companies have adopted the methods in question. In their book, Womack et al. (1991) published statistical results that showed how, according to a series of core business performance measures, Japanese production methods (especially that of Toyota) were superior to production methods in the contemporary American car industry. Toyota's production method has become widely dispersed throughout many parts of the world, including Scandinavia, under the label of Lean. Even though the techniques, methods, and mindsets of Lean have been well documented (Liker, 2004; Liker and Meier, 2006; Bicheno, 2004; Monden, 2010), research on Lean as a control package is scarce.

The main objective of this paper is to research and empirically test whether the performance effects from Lean are related to complementary effects between the control forms in a Lean organization. In this

endeavor, we pursue a deeper understanding of complementarities among control forms in Lean by introducing the notion of “balance between control forms,” which we find important to enhancing performance. We also use this insight on “balance” to derive consequences for the more general control package literature.

Kennedy and Widener (2008) find that Lean is a control package consisting of many variables that constitute an integrated system of multiple control forms. Their case study shows that management accounting research needs to understand the operations management technology, as this implies the use of several control forms in certain ways. Also, Chenhall (2003) stresses that management accounting systems cannot be studied in a vacuum without taking other control forms into consideration. This could cause the same problems as with omitted variables, where the real explanatory factor is left out of the model. This also implies that management accounting should be researched as part of packages that include other control forms. Expanding research to encompass the many variables in a control package, for example a Lean package, emphasizes the need to study the phenomenon using a systems fit approach (Drazin and Van de Ven, 1985).

The widespread use of Lean in companies begs the questions of if and how Lean as a control package enhances performance. Published research on statistical tests of how Lean works as a control package and relates to performance is almost nonexistent. One previous attempt to test the effect of Lean in a systems fit model did not provide any significant results (Selto et al., 1995). With Kennedy and Widener’s (2008) qualitative case study, research has developed a step in the right direction towards building theory that answers the question of how Lean works as a control package. However, their study provides no quantitative evidence of performance effects. Moreover, their study identifies a number of intervening and bidirectional relations in the controls of the Lean package, and they consequently argue that the framework they construct is suggestive of a configuration fit. Perceiving Lean through the lens of systems fit begs the additional question of whether performance effects are solely based on complete mutual dependency among control forms or whether improvements to the individual control forms can enhance performance on their own continuously. This includes understanding how the right balance between control forms is constituted, and thus, how off-balance can affect performance.

The study of control packages is not new, as previous empirical papers have studied the application of multiple control forms such as personnel controls, action/behavior controls, output controls, and clan controls, with various labels attached to them (Kren and Kerr, 1993; Eisenhardt, 1985; Kihn, 2007; Kihn, 2010; Abernethy and Stoelwinder, 1995; Abernethy and Brownell, 1997; Otley and Berry, 1980; Gerdin, 2005; Sandelin, 2008; Chenhall and Langfield-Smith, 1998). However, there is still much to understand about how this works in a modern Lean context. Abernethy and Brownell (1997) call for further research on how organizations rely on combinations of control forms in given settings, and they state that virtually nothing is known about how the effects of any one control are governed by the level of simultaneous reliance on other forms. They conclude that this understanding remains piecemeal until research explores these complex relations. Kihn (2010) adds to this by stating that little research exists on how managers combine accounting information and various other types of information, and it is not clear how managers emphasize multiple forms of controls. Both Kihn (2007) and Kihn (2010) welcome further research on various types of

controls. Our study also responds to Abernethy and Brownell's (1997) call by examining how multiple control forms may enhance performance in Lean organizations, and to Malmi and Brown's (2008) request for research on management control systems as a package.

In the general control package literature, it is not found how Lean organizations pick and balance control forms to create complementary performance effects. Nevertheless, Kennedy and Widener (2008) use the general control package literature to address which control forms are used in a Lean organization; however, they do not research in depth what happens to the complementarity effects on performance if control forms are balanced or off-balance. They argue, though, that removing one control form will cause the whole system to roll back to another control system, but they do not look into situations where control forms are present at different levels of implementation. Hence, they do not present an understanding of how performance is affected when there are different balance and off-balance situations between the multiple control forms.

Understanding balance and off-balance in a complementarity perspective in a Lean context also has potential within general control packages theory. In the general control package literature, complementary effects are described, or at least tested, as additive relations between control forms in a profile deviation analysis using a distance measure from an ideal state – or they are tested as interaction effects, which is a pure multiplication that does not necessarily represent or portray how control forms should be balanced. The latter merely states that one control form is dependent on another, and increasing one of them is just as beneficial for performance as increasing the other (i.e., doubling one of them is just as good as doubling the other), even though the off-balance between them is further increased.

Within the theory of equifinality, Sandelin (2008) uses the term “internal consistency between control forms” to describe how control forms can be balanced in different ways and can generate the same outcome/performance, even though the external context is the same for two or more ideal combinations of control forms. However, internal consistency only describes balance for multiple ideal combinations, and does not describe how off-balance can affect the performance stemming from complementary relations amongst multiple control forms.

We perceive the use of multiple control forms in Lean organizations to be quite tightly coupled in a system, as opposed to loosely coupled (Roberts, 2004). Thus, having the right balance is important, as off-balance will thwart the complementary performance effects more in a tightly coupled system than in a loosely coupled system. Off-balance occurs when one or more control forms is deviating from the ideal distance to other control forms. Hence, the level of tightness is important to understand for control packages other than the one in Lean organizations, as it influences how off-balance between control forms can affect performance.

Furthermore, statistical analysis techniques of systems fit models need to be further developed. Previous research papers have thoroughly mapped and described statistical techniques within a systems fit approach (Boyd et al., 2012; Gerdin and Greve, 2008; Venkatraman, 1989). However, new techniques are needed in order to better understand how the control mechanisms function in a Lean organization, i.e., how they co-

function either as mutually dependent elements or as independent elements without synergy (i.e., just additive). These two different ways of functioning cannot be distinguished in the Euclidian distance technique used by Drazin and Van de Ven (1985) and propagated by Gerdin and Greve (2004) for systems fit model testing, even though the systems fit model is founded upon a synergetic way of functioning. The statistical technique of Euclidian distance cannot stand alone in testing for the existence of synergy. However, research has not moved beyond the use of Euclidian distance in testing systems fit models. By enhancing the conventional profile deviation analyses techniques, we are able to test for performance effects of the Lean control forms package and to assess whether these are based on complementarities among the individual controls. This is an additional purpose of the paper. Thereby, we respond to Gerdin (2005), who sees an interesting potential in exploring holistic approaches.

The remainder of the paper is structured in the following way: Section 2 contains the theoretical background and hypothesis development. In Section 3, the research method is described, and, in Section 4, the analysis is presented. Section 5 presents the conclusion. Finally, we present the limitations of our study and point to future research directions in Section 6.

2. Background and hypothesis development

Recent research points to the fact that Lean implementation involves the implementation of an entire package of different control mechanisms. According to Kennedy and Widener (2008), these control mechanisms can be divided into three different control forms: behavioral control, output control, and social control. In addition, the Lean organization uses a value stream organizational layout where resources are dedicated to product groups with similar production flows.

Behavioral control encompasses standard operating procedures (SOPs). These SOPs are aligned with output control measures and hence act as an aid for employees to reach the required levels of output. Standard operating procedures are not strict instructions from management to the value stream cells, but they describe the standard way to operate a machine or how products should flow in the cell. They describe the current best practice and help the employees create continuous improvements (Adler and Borys, 1996) through systematic description of value-added and non-value added activities. Stated differently, SOPs are put in place to visualize (e.g., using markings on the walls and floors) and ensure that operators work similarly and according to one best practice. However, the SOPs can be updated and circumvented if contingencies arise, which they often do (Chapman and Ahrens, 2004). In most cases, visualized problems should be handled by empowered workers (Liker, 2004).

In output controls, the Lean organization relies mostly on non-financial measures (Fullerton and McWatters, 2002), typically lead time, delivery performance, and first time through (Maskell and Baggaley, 2004). Output controls are employed to influence and motivate, and they are not intended to affix blame (Hopper and Jazayeri, 1999). Just like SOPs, output controls are used to enhance visual control and to serve as decision-making aids for empowered workers. Financial measures are also put in place for lower-level

management and shop-floor employees (Lind, 2001). In these cases, financial measures are typically oriented towards the value streams. Labor tracking systems based on IT are often perceived to be dysfunctional to Lean, or at least redundant, and cost variance analyses of responsibility units/value streams are not performed (Kaplan and Cooper, 1998; Kennedy and Widener, 2008).

Training, visualization, empowerment, and peer pressure are elements of social control (Kennedy and Widener, 2008). Training enables employees to carry out their operations better, and, through training, they are subject to internalization of what is best practice in the company. This is part of ensuring goal congruence between the employees' actions and the Lean philosophy; increased training generates a multi-skilled workforce and reduces performance errors. Peer pressure is strongly related to training. Employees' level of training is typically visualized on an open board on the shop floor for everybody to see and improve on. Members of the cells pressure each other to increase their skill levels.

Visualization works together with both output controls and SOPs. Whiteboards on the shop floor post numbers generated through output controls (Carreira, 2005), such as the ability to deliver on time, so that cell workers know how well they are performing. This becomes input that helps determine whether or not corrective actions are required. Visual control also involves the ability to look into the production cells by removing walls and other things that block the visual lines between cells. This provides an opportunity for cells to help each other, which is necessary, as they are quite interdependent due to the low buffer stocks in between cells. Moreover, visualization functions in close conjunction with SOPs because without best practices, or SOPs, it is difficult to benefit fully from visualization. For example, yellow lines on the floor are used to ensure that the work-in-progress levels are under control, and standard levels of work-in-progress can only be pre-calculated if SOPs exist.

As mentioned above, a value stream layout is typical in a Lean organization (Kennedy and Widener, 2008). Hence, the functional layout is abandoned (Lind, 2001). The value stream layout is characterized by high flow, low changeover time, and short lead time, because buffer stocks are continuously cut as much as possible (Liker, 2004). It is the nature of this high flow that creates the need for empowered workers, as there are no buffer stocks to address problems between responsibility centers. Furthermore, in this layout, the value stream is given the responsibility for finishing a whole product (Bicheno, 2004). As the flow in the value streams increases, more pressure is put on the control forms to function well and in balance with one another.

2.1. Testing effects of Lean

A main lesson to be learned from Kennedy and Widener (2008) is their breakdown of the practice-defined variable, Lean, into multiple theoretical variables, i.e., the control forms described in the previous section. This is the approach recommended by Luft and Shields (2003) for investigating a practical relevant contemporary phenomenon in a theory-consistent manner.

Kennedy and Widener (2008) find that the individual control elements cannot work without the other control elements, i.e., complementarities exist among them. Hence, in a Lean environment, none of the control forms can provide substantial performance effects without support from the other control forms. Social controls and output controls are dependent on each other to produce performance effects, and social controls are dependent on bureaucratic/behavioral (SOP) control and output control. So, in their case study, Kennedy and Widener (2008) conclude that Lean is a control package with synergies between a number of control forms. The proposition that Lean brings about synergy between its elements is supported by Kobayashi (1995) and by the founder of Toyota, Toyoda (Liker, 2004). Taiichi Ohno, the inventor of the Toyota production system, argues that the key to the success of the Toyota production system is not found in any of the individual elements, but in the system made up of all the elements together (Liker, 2004). Liker (2006) warns against perceiving Lean as piecemeal technical projects. This will only provide the organization with an incentive to pick the low-hanging fruit and to thereby miss out on the chance to acquire a long-term sustainable system. In Kobayashi's (1995) book, he argues that Lean can only lead to good performance if all the pillars of the system are functioning and if a cross-functional synergy has evolved between the pillars. Finally, Lind's (2001) case study shows that after implementing World Class Manufacturing (similar to Lean), the case company introduced changes to many of the elements in its control package.

Viewing Lean as a control package impacts how the effects of Lean can be tested. Kennedy and Widener (2008) propose—in the words of Drazin and Van de Ven (1985)—a systems fit between the various control forms. The systems fit approach is labeled a “configuration fit” by Gerdin and Greve (2004). Within this approach, it is not adequate to investigate the performance effects of individual parts of Lean, as this would be cutting away the interactions of the control mechanisms within Lean; hence, we need to tease out the complementary effects among the control forms statistically, instead of letting them remain vague.

Despite this, Selto et al. (1995) have provided the only case study that tests the systems fit approach (in a Just-in-Time setting). However, they find no significant effect when they apply the statistical technique of Euclidian distance. The reason for the lack of significance is not ascribed to the statistical technique but to problems with degrees of freedom causing their test results to be non-significant. The Euclidian distance test procedure is supported by Gerdin and Greve (2004), who refer to it as profile deviation analysis. Therefore, even though qualitative case studies have suggested that Lean is a synergetic control package, no empirical statistical study has been able to support this proposition.

Previous cross-sectional surveys of Lean, Just-in-Time, or World Class Manufacturing have presented mixed results in terms of profitability, stemming from the advanced manufacturing philosophies. Kinney and Wempe (2002), Huson and Nanda (1995), and Balakrishnan et al. (1996) all find no sustainable effect on profitability. Contrary to this, Shah and Ward (2003) do find significant effects on cost per unit. With the recent findings of Kennedy and Widener (2008), we believe that previous research has shown mixed results as a function of researching only a few variables in a selection or interaction fit approach, with minor

attention being paid to Lean as a package of synergetic elements of multiple control forms.¹ Therefore, the main purpose of this paper is to research and empirically test whether the performance effects from Lean are related to synergetic, i.e. complementary, effects between its control forms.

2.2. Development of hypotheses

Our paper is inspired by the case study of Kennedy and Widener (2008) and their proposition of Lean as a systems fit model with multiple control forms. Their proposition is based on in-depth analyses of interviews, archival data, and on-site visits; however, no quantitative data, data analysis, or explicit performance measures were included. Hence, their study gives little direction as to how performance deviations occur between organizations or value streams. Therefore, in our study of the performance effects of Lean as a control package influenced by complementarity, we need to propose a refinement/modification of how we perceive the relation between control forms and performance.

The basic assumption of configuration (systems) fit, which says that changes only take place and only have effect in quantum jumps (Gerdin and Greve, 2004), can be reconciled with the continuous improvement philosophy in Lean (Womack and Jones, 2003), where the organization is constantly fine-tuned. The Lean approach is seen to be one configuration, and within this configuration, companies/value streams are trying to come as close as possible to the ideal state. Hence, when a value stream moves away from the ideal state, the larger is the misfit, and performance declines. Therefore, the continuous improvement approach is to be seen as closing the gap between the current state of a value stream and the ideal state, and not as a continuous process of changing to another configuration (like mass production, for instance). In our research, we are interested in the performance effect of Lean, as this has been questioned in the studies presented above. Thus, in our research, we have adopted the contingency version of systems fit, contrary to the congruence fit, where no performance variables are included.

In order to understand the expected synergy (configuration fit) between the control forms, we use Milgrom and Roberts' (1995) notion of complementary effects on performance. They state that activities are complementary if doing more of one of them increases the returns from doing more of the others. In our case, the activities are the control forms used in Lean, where the performance (return) gained by one control form is dependent on the level of the other control forms, thus representing our study's complementarity, i.e., a situation with control forms at a similar level is expected to lead to better performance than a situation where one control form is really good and another control form is weak—even if the average of the two situations is the same. However, the effect of the latter situation (with both good and bad control levels) would create better performance than a situation with only bad control form levels. Therefore, balancing the control forms is expected to improve performance because of synergy effects, but a high level is also required. This notion of balance in understanding the complementary effects is our refined approach to it, and we discuss this further below.

¹Selection and interaction fit approaches are the terminology of Drazin and Van de Ven (1985). In Gerdin and Greve (2004), these two approaches are part of the Cartesian fit approach.

Kihn (2010) presents survey results supporting a somewhat similar perspective. The author concludes that managers perceive each control form as relevant, but none of them separately as complete information. Hence, managers combine the controls in such a way that they complement, as opposed to substitute, each other. The author's key findings were that managers differed significantly in the way they emphasize controls and six principal styles could be identified, and all styles were not equally effective in enhancing performance. Otley and Berry (1980) suggest concentrating on observing the whole range of control activities undertaken by specific organizations, as this is most fruitful in the long term. Thus, not only financial controls should be observed, but also other control forms. Moreover, they state that these other control forms should not replace financial controls but should be complementary. Both Kihn (2010) and Otley and Berry (1980) are somewhat in line with the notion of complementarities in this study.

Kennedy and Widener (2008) provide an example of how they perceive synergy between control forms in a Lean context. They argue that removing peer pressure as a control form would cause a reversion of the entire control system to the previous state, i.e., to a completely different control form system. Hence, they argue that the Lean control form system is a package that is tightly coupled. A tightly coupled system, as opposed to a loosely coupled system, is characterized by a situation in which changes in one element will severely compromise performance unless numerous other elements are also adjusted (Roberts, 2004). Roberts (2004) states that a standard assembly line is the ultimate example of a tightly coupled system, as nobody can reduce the pace without affecting the overall performance. With respect to the notion of the complementary effect stated above, this leads us to expect that having one or more control forms at a very low level will thwart some of the performance effects of having other control forms on a high level. Therefore, we need to study the balance between the control forms to understand and test for this type of complementary performance effects. However, increasing the level of a decent level control form is expected to have at least a minimum performance effect, despite one of the other control forms being at a very low level; we have labeled this as the additive effect. This additive effect can be thwarted by having off-balance between the control forms, and this thwarting represents the complementary effect on performance. Thus, we do not expect the Lean control form system to be completely and extremely tightly coupled, as the natural consequence thereof would be to correlate just the lowest level control form with performance. The other control form would not be of relevance in this extreme view, as only the weakest link in the chain would set the performance level. Additionally, the extreme approach to tightness would violate our approach to complementary effects. Roberts (2004) states that hidden in the definition of complementary effects are the fact that doing more of one activity (control form) does not automatically demand doing more of another activity simultaneously. Hence, the returns of increasing one activity are not fully dependent on the levels of the other activities. Our refined notion of complementary effects is not a conventional multiplicative crossover interaction effect (Gerdin and Greve, 2008), where two or more variables are multiplied as a representation of the complementarity effect on performance. In the multiplicative interaction notion of complementary effects, one control form can be increased to double, and we would expect performance to be doubled as well. However, this notion of complementarity does not take into account whether the control form that is doubled is increased from a low or high level compared to the other control forms, and consequently it does not reveal whether the control forms are being brought more into balance or more into off-balance. It could be

that a current high level control form is doubled and the low level is not, and the multiplicative interaction would suggest doubled performance, whereas our refined notion of the balance versus off-balance complementary effect suggests there would be a lower than doubled increase in performance due to increased off-balance. Hence, in our refined notion of the complementarity effect on performance, it does matter whether or not off-balance is increased compared to the ideal balance. Therefore, our notion of off-balance in understanding the complementarity effect on performance is distinct from the multiplicative-interaction notion of the complementarity effect. This is especially important when considering tightly coupled systems, as they rely more on complementary effects. We further elaborate on this notion of complementary effects in Section 4.3 with the statistical approach and with a numeric example.

A potential threat to the Lean performance could be if one of the control forms is not improvable or is very costly to improve. This would cause drastically lower performance if we assume a high synergy effect. In some cases, it might be impossible for an organization to get every element of Lean working right. In the case of Ezzamel and Willmott (1998), for example, the company struggled to convince the employees that empowerment was right. In this sense, the Lean organization could potentially be quite fragile. If only one element cannot be fully implemented, then the entire system will miss out on performance benefits stemming from the complementarities, i.e., the synergy from balancing the control forms. However, a cross-sectional study by Shah and Ward (2003) show that individual bundles of Lean practices increase performance irrespective of synergy effects (interactions) between these bundles. This supports the proposition that there might also be an additive effect on performance from Lean elements. However, they did not test for balance effects with their interaction approach. Hence, in their study, our balance approach to complementary effects is not researched. The question of whether or not improvements in single control elements can create improved performance is built into our hypotheses. We propose that both hypotheses are supported.

This conceptualization leads us to the following hypothesis:

In the Lean organization, the output, behavioral controls, and social controls affect performance. The higher the average level of the control forms, the higher the level of performance.

The three control forms complement each other. Hence, they create the highest performance effect when they are *balanced* (symmetrical) at the highest *level* possible.

This general hypothesis can be divided into two separate hypotheses to be tested:

1. The higher the level (closer to the ideal state) of the three control forms (output, behavioral, and social), the higher the performance in a Lean context.
2. The more balanced (symmetrical) the three control forms, the higher the performance in a Lean context.

Hypotheses 1 and 2 are not mutually exclusive. In fact, Hypothesis 2 is an extension of Hypothesis 1. If Hypothesis 1 is confirmed and Hypothesis 2 is not, we could describe the performance effects from the Lean control forms to be additive, and thus not a synergetic configuration fit model. If Hypotheses 1 and 2 are both confirmed, we would describe the performance effects as complementary with respect to our balance approach. In this situation, the additive effect of increasing the average level of the control forms is either reduced or increased, with the effect coming from the control forms being more or less symmetrical. If we have a simple situation with only two control forms, and one is improved from level 4 to level 5 (while the other is kept at level 2), then there will be an increase in performance caused by the additive effect (higher average level of control forms), but the effect will be somewhat reduced by lower complementarity, i.e., balance, as the gap between the control forms is increased. Therefore, one could say that the full performance potential coming from the increase from level 4 to 5 is not exploited. In the converse situation, where a level 2 control form is increased to level 3 (while the other remains at level 4) there is an additive effect that is further increased by a complementary effect due to the control forms now further complementing each other. We expect to confirm both hypotheses. Specifically, Hypothesis 2 is essential, as it confirms the configuration fit nature of the Lean control forms. However, it would be odd if only Hypothesis 2 is confirmed, and not Hypothesis 1, as this would reflect a fit where the control forms only have to be balanced, regardless of whether this is at a low or high level.

As a final note on hypothesis development, we do not include variables on “value stream layout” or “accounting system” (unlike Kennedy and Widener, 2008) due to the non-variance of these in our empirical dataset, cf. the next sections.

3. Research method

We study the performance effects of multiple control forms in a single case company using longitudinal Lean audit measures as approximations of the control forms utilized in Lean. These measures are tested via an enhanced profile deviation analysis against income statements and other non-financial performance measures of the value streams in the case company. All measures used are archival data stored by the case company.

In the following sections, we present the case company, the constructs, and the statistical techniques applied to our dataset.

3.1. Case company

The company employs about 2,000 blue- and white-collar workers in total, and the annual revenue is a little more than € 250 million. The company is headquartered in Scandinavia, and its main production site, the site of our investigation, is situated just next to the HQ.

The company produces a variety of small electronic equipment for business customers around the world. It is a family-owned company that has expanded to its current size since the present CEO took over in the 1970s.

In 2003, several Lean tools, such as JIT, Kanban, 5S, Lean office, Kaizen, etc. were implemented, and the company has been dedicated to the Lean philosophy ever since. The company also changed its organizational structure and production layout at this point in time. The new structure involved a product line for each product group, very much like value streams, although all product groups still have some functions in common, e.g., raw materials inventory, and packaging. Nevertheless, the vast majority of activities and resources are located within each value stream, as recommended in the Lean approach (Maskell and Baggaley, 2004). Since 2003, the company has won several national awards for its substantial Lean program, and it is recognized by Lean experts in its country to be at the frontier of Lean implementation.

3.2. Quantitative single case study

All data are from this single company, and our unit of analysis is the company's seven product groups/value streams. All value streams sell only to other businesses as a supplier of goods that are part of an assembly of a final product for consumers. Even though there are minor variations between the value streams, these are not considered influential. A main advantage of this is that external control variables are held constant (Young and Selto, 1993), e.g., the level of environmental turbulence and uncertainty. As we investigate the impact of control forms as a systems package, we can exclude the contingency factors of the environment. In addition, the company has chosen to employ the same Lean practices in each value stream, which implies that slightly different market contingencies do not create enough variance to create a need for different Lean practices. It is also worth noting that all value streams have used the same Lean practices since they initiated their work with Lean. Hence, performance differences as a function of differences in Lean tools used are nonexistent. Variance in performance is likely to be related to how well the control forms are implemented and working. Likewise, the same accounting model is used for each value stream.

Another reason for concentrating on only one company is access to very rich data. In this study, we have access to quarterly contribution margin accounts for the seven separate value streams for multiple years. We also have access to quarterly Lean audit reports for each of these value streams. Last but not least, the organization has provided us with a number of non-financial indicators per value stream for the same period of years. Combined with our visits and interviews with many company representatives over three years, these data have provided rich insights into the performances of the value streams and their drivers. Getting access to confidential disaggregated accounting figures and Lean reports gives a unique insight into an organization that has employed Lean extensively for many years. This is a main advantage of using active site selection (Anderson and Widener, 2007). Furthermore, we agree with Chapman (1998) that qualitative and quantitative studies ought to respond to each other to build up knowledge, and a multiple methods approach is desirable (Birnberg et al., 1990). Our response is to draw on the findings of Kennedy and Widener's (2008) qualitative case study and to further develop propositions and test these statistically. Additionally, we can draw on information gained during the many site visits and interviews, both formal and informal.

3.3. Measurement of constructs

There are three categories of data in this paper. For each of the seven value streams, there are (i) internal income statements, (ii) non-financial performance measures, and (iii) Lean audits as measures of control forms. The latter category contains the independent variables, while the first two are the dependent variables in our hypotheses.

3.3.1 Financial and non-financial measures of performance: dependent variables

As performance measures, we employ two categories for each value stream: quarterly data from actual internal accounting reports, and various non-financial performance data. These data cover the period from 2005 to 2009. The company generates accounting reports monthly. The non-financial figures are available in almost real time for the production managers and team leaders and are discussed at weekly whiteboard meetings in each production cell. For our purpose, we have aggregated both financial and non-financial data to a quarterly level. This aggregation is done to relate these financial and non-financial performance data to the measures of control forms in the Lean audits, which are captured quarterly by the case company.

The first set of performance data, the accounting reports, is set up in a tiered contribution margin format per value stream. A monthly income statement is created for each of the seven product groups. Each group has its own revenues from the sales units around the world, and revenues reflect actual prices achieved by the sales units. Hence, sales numbers are market-based and are not based on more or less dubious internal transfer prices. The contribution margin sheet is divided into four main groups, namely, Contribution Margins 1, 2, and 3, and capacity costs. In Contribution Margin 1, materials and direct labor are deducted from revenues. Contribution Margin 1 less indirect variable costs (e.g., costs of small tools, repair of small tools) and additional night shift payments gives Contribution Margin 2. Costs of customer claims, cost of scrapped poor-quality products and write-down on stocks constitute the difference between Contribution Margins 2 and 3. Finally, capacity costs, mainly white-collar salaries and depreciation, are deducted as separate line items. Contribution Ratio 3 (Contribution Margin 3 divided by revenue) is of particular importance, as this includes all costs that vary in the period under scrutiny, thus representing the variation in accounting performance. Quality costs, which make up the cost difference between Contribution Margins 2 and 3, are treated as a separate performance variable in addition to being included in Contribution Ratio 3. For this purpose, we divide quality costs with revenue to express a quality cost percentage. The case company uses a wage cost ratio as a performance indicator, and we will include this in our analysis as a performance measure, calculated by dividing the direct wages with revenues. Hence, this is a component of Contribution Ratio 3, which only focuses on wage fluctuations. To avoid effects stemming from currency exchange rate fluctuations when selling to the American market from the Euro-zone, revenues are corrected for fluctuations in the dollar/Euro exchange rate.

The management accounting system of the case company is quite similar to the one described in Kennedy and Widener's (2008) paper. They describe a management accounting system based on actual costs with no allocation of overhead costs. These elements are similar to our case company, which uses a tiered contribution margin model with no allocations. Unlike Kennedy and Widener, however, our case company

also uses standards in the management accounting system. However, the cases are similar in the sense that there is no detailed labor tracking system per job order or product unit.

The second set of performance data is the non-financial measures per value stream. These data are generated in almost real time by the organization. We have received weekly aggregated data and have aggregated these further to match those of the quarterly Lean audit reports. The organization measures three Lean non-financial performance indicators: delivery performance and two different lead times. Delivery performance is measured as the percentage of customer deliveries made within the promised time frame. The first lead-time measure is the time between when an order is released to operations and when it arrives at packing. The second lead-time measure reflects the time between when an order is started and when it is finished.

These three non-financial performance measures are central to Lean. Lead time is an especially important indicator of Lean performance, and it is one of the first indicators to be reduced through the introduction of Lean (Liker, 2004). Obviously, delivery performance is critical in Lean (Just-in-Time), as operating with small buffer stocks would otherwise make stability highly fragile (Kalagnanam and Lindsay, 1998). These three performance indicators are therefore critical to the Lean implementation success of the case organization and are thus also important when measuring how the Lean package performs as a control package.

Using both objective financial and non-financial performance measures in the value streams enhances the reliability of measures. Selto et al. (1995) criticize Govindarajan (1988) for using self-reported performance measures, and, using these self-reported measures, Govindarajan (1988) also acknowledges a limitation of his study, as he believes that objective data would yield more powerful results. Configuration fit is used in both these studies and they both support the use of objective performance measures, not self-reported measures. Hence, the use of actual objective performance measures is perceived as a crucial benefit of our study. Moreover, Venkatraman and Ramanujam (1986) recommend the use of both financial and operational performance effect variables that come from the primary source, as this will reduce the number of potential threats to the validity of the research design. According to these authors, the advantages of this approach are that it provides a comprehensive operationalization of business performance and it can be adopted at the SBU level. Furthermore, they recommend not aggregating multiple performance measures into one variable, as this can cover up underlying types of performance that may be in conflict. Thus, we analyze the performance variables in our tests separately, and not aggregated across multiple performance indicators.

3.3.2. Lean audit measures as measures of control forms: independent variables

The third set of data is internal Lean audit reports. The Lean audit measures in this paper are deployed to approximate the control forms used in a Lean setting according to Kennedy and Widener (2008). First, we describe how the Lean audits are measured and then how these are used as measures of the control form constructs.

The Lean audit reports are made quarterly for each value stream, and they consist of 14 areas of interest. The inspiration for the Lean audit in the case company is the seminal book by Kobayashi (1995). The Lean audits were implemented by a consultancy company, and only small changes have been made to the audits over the years. Within our area of interest, no changes have been made except for a change in the scoring scale in 2008. Before 2008, the scale used was 1–5, and thereafter it was changed to 1–10.

Each quarter, the value stream is expected to ask for a Lean audit at the company's central Lean department. The Lean department then sends out a specialist to do an assessment using the Lean audit tool. Over the years, only two persons have made these assessments. In our opinion, this enhances the reliability of the measurements. Furthermore, this provided the opportunity to engage in a dialog with the persons who fill out these audits. From these interviews, we conclude that the Lean audit instrument has been consistent over the years. We had the opportunity to follow a Lean audit session within one value stream. This lasted three hours, which is typical. The areas in the Lean audit are filled out while walking through the value stream shop floor and in the offices of the value stream. After acquiring all the scores, the Lean auditor presents the results to the relevant persons responsible for the value stream. This often leads to a discussion of how to improve the scores, and the Lean auditor—an expert in Lean—provides recommendations on how to do this. The value stream managers do not see this as a traditional top-down control mechanism, but merely as support to improve the Lean value stream.

The Lean audit instrument consists of descriptions of how each score should be achieved for each of the measurement areas. Therefore, as the guidelines for achieving a certain score must be followed, subjectivity is reduced. An example of this is “continuous team improvements: *the teams are not working independently, as the teams are not providing solutions to the current problems in the team area but just carrying out the instructions given by the superior. However, small improvements are seen, such as an opportunity to provide improvement suggestions to the superiors,*” which will lead to a score of 2. To raise the score to 3, the value stream is supposed to “*react towards problems and try to solve them autonomously.*” Similar descriptions are found in Kobayashi's (1995) framework, which is the main foundation of the company's Lean audit and which also stresses the importance of balance between improvement areas.

Both in Kobayashi's (1995) framework and in the Lean audit used in the case company, the scales of measurements for each Lean audit measure are constructed in such a way that maximizing them will reflect an ideal state of performance. This is in line with the framework of Kennedy and Widener (2008). This maximization of the control forms is not to be understood as Lean companies maximizing all aspects (all possible techniques/methods) of the multiple control forms in the sense that Ouchi (1979) and Ouchi and Maguire (1975) describe them. For instance, the output control forms in Kennedy and Widener (2008) and our case company do not include detailed labor variance tracking per employee, as would be the case in a traditional mass production setting. Instead, the output control form in a Lean setting would be more focused on non-financial performance measures that are collected by the cell workers themselves. In general, with fragile production systems like Lean and with the experimentation that is expected of work groups, no single control form would be enough.

Table 1 shows the general description of the areas in the Lean audit report, grouped according to the dimensions of controls. The dimensions of controls are taken from Kennedy and Widener (2008). Hence, we have related the Lean audit measures to the control forms. Kobayashi's (1995) framework² is well known among organizations that want to measure the implementation level of Lean (Bicheno, 2004), or at least in consulting companies. Hence, it may not be a coincidence that Kennedy and Widener's (2008) model has many similarities to the Kobayashi framework, as this framework has inspired Lean implementations worldwide. Moreover, the framework is constructed such that it reflects the main elements of Lean, and some of these elements are the control forms as structured by Kennedy and Widener (2008).

Table 1 is structured so that the left-hand column shows the label from the Lean audit used in the case company, the middle column is the control forms from Kennedy and Widener (2008), and the right hand column is a summarized description of the Lean audit measure taken from the case company's own Lean audit reports. The "goal and result control" Lean audit measure is used in our data model to approximate Kennedy and Widener's first control form, "output control." They describe output control as generated from the bottom up, which is also the case for this Lean audit measure, as numbers are to be "maintained" in the cells. Moreover, many of the performance measures used in both our case company and that of Kennedy and Widener are very similar—cf. on-time-delivery, units per hour (day by the hour), lead time (order to invoice), and financial measures such as cost of poor quality.

Kennedy and Widener's second control form, "behavior control," is mainly implemented through the use of Standard Operating Procedures in their case. These are used as roadmaps for workers to assist them in completing their tasks. This is similar to our case company's Lean audit, where this is measured as "standardized work" and "mapping," in which standards are updated continuously, ensuring that behavior is sustainable (that is, not reverting to old habits).

The third control form, "social control," Kennedy and Widener (2008) describe as consisting of employee empowerment, peer pressure, visualization, and training. This involves bulletin boards showing the training level of each employee, self-selection of training needed in groups, empowerment to make decisions in cells, peer pressure within groups to make sure everybody is well-trained and performs highly, influence on hire and lay-off decisions, and color-coding to manage work. The same content is found in our case company under the labels of "competences," "continuous team improvements," "neat and clean," and "coupled production." Our case company also has bulletin boards for training levels and self-initiated training under "competences." Under the Lean audit label, "continuous team improvements," includes team involvement in hiring and firing, together with the team's own responsibility to make improvements by being able to make decisions. Kennedy and Widener (2008) include peer pressure and empowerment as two different sublevel constructs, whereas the Lean audit treats this as only one measure (continuous team improvement). However, we consider that the content is very similar, so this will not affect the conclusions of this paper. Kennedy and Widener's (2008) visualization (a sublevel to "social control") is measured by two measures in the Lean audit, i.e., "neat and clean" and "coupled production," but the content is basically the same. Therefore, in

²Known as the "PPORF" program in Japan.

general, the control forms as described by Kennedy and Widener are very similar to the Lean audit measures in our case company. Hence, the Lean audit measures provide a solid approximation of the control forms.

<Insert Table 1 about here>

Table 1: Measures in the Lean audit³

The Lean audit reports also score the goal congruence of the value streams with the overall company goals. We take this measurement as a proxy of goal congruence performance. This creates a dependent performance variable. The controls of the organization are implemented to create goal congruence between the parts and the whole, thus creating the desired actions in line with the company's strategy. The goal congruence is therefore a "softer" form of measuring whether the control package achieves the purpose intended, i.e., goal congruence. The measure offers a different angle on performance that is more tightly coupled with the immediate purpose of control and is less sensitive to "unintended effects" that might affect the financial and non-financial performance measures.

3.4. Statistical test method

Our hypotheses are formulated under the regime of systems fit, where we have many variables interacting (Lean audit variables) and where we expect companies to be able to survive (with different performances, however), even though they have not attained the ideal state. Additionally, we have performance data. In such a situation, Gerdin and Greve (2004) suggest using profile deviation analysis. This model originates from a paper by Drazin and Van de Ven, in which they argue that pattern analysis (deviation analysis), is an appropriate technique for analyzing the "gestalt" characteristics of the organization in a systems fit approach (Drazin and Van de Ven, 1985).

Within the profile deviation analysis, the Euclidian distance is used to measure the distance between the ideal state and the measurement scores for the constructs under scrutiny. The result of the deviation analysis is an overall "distance score" and is a measure of the misfit to the ideal state. This distance score is regressed against performance and a negative correlation is expected, meaning that the greater the distance from the ideal state, the lower the performance (Hult et al., 2007). For the purpose of our study, we need to address the complementarities between the control forms. In the Lean audits, the ideal state would have maximum scores on all elements. However, with the Euclidian distance score only, we cannot distinguish between Hypothesis 1 (improvement of control without balance) and Hypothesis 2 (balanced improvements of controls), so we are not able to investigate whether complementary effects exist.

A high score on one control measure cannot fully compensate for a low score on another. For example, a value stream that scores "3" on each of the three control forms generally performs better than a value stream that scores "2" on one control form, "3" on another, and "4" on the third control form. The Euclidian

³ The company's Lean audit scheme is in the local language, not in English. Translation has been done by the authors.

distance measure needs to be supplemented to enable us to test whether the improvement comes from complementarities between control forms or just from an increase in the average level of the control forms. We call this supplementary distance measure the “city-block” score and discuss it further in the analysis section. With this new statistical approach, we are able to address whether complementary effects on performance from the control forms are thwarted if the off-balance between the control forms is increased.

According to Miller and Friesen (1984), the ideal state of a configuration can be derived empirically or conceptually. Earlier, Selto et al. (1995) derived the ideal configuration empirically in the aforementioned Just-in-Time setting, using the best performing work-group as the ideal state. In our case, we used the conceptual approach. Three reasons exist for this choice. The first reason has to do with the case company’s Lean audits, where the measures/constructs are made in such a way that scoring the maximum value (5 on the 1–5 scale, 10 on the 1–10 scale) is perceived to be the best level. Hence, the Lean audits have a built-in conceptual best level. The second and third reasons are based on Govindarajan (1988). He argues that the empirical approach demands a benchmark group of the best performers, but it is somewhat arbitrary how many performers should constitute this benchmark group. We would further argue that using only one observation to create this empirical benchmark would entail the risk of choosing an outlier (nonetheless, this is what Selto et al. (1995) did). On the other hand, including more observations to create the benchmark would lead to a loss of observations in the test procedure, as the benchmark performers cannot be included in the calculation of the Euclidean distance. This will reduce the degrees of freedom in the statistical test (Govindarajan, 1988), which is the third reason for choosing the conceptual approach in setting the ideal state.

4. Data presentation and analysis of test results

In this section, we first present our data and the necessary indexing that was done, followed by test results that provide statistical support for Hypotheses 1 and 2. First, we present support for Hypothesis 1 using the traditional distance measure, Euclidean distance. Next, support for Hypothesis 1 is presented using city-block distance. Finally, support for Hypothesis 2 is revealed using a novel method of disentangling the Euclidean and city-block measures.

4.1. Data presentation

We need to index the archival data from the company for a number of reasons.

All eight performance measures—financial and non-financial—are indexed using the same points in time as the baseline measure. Especially for the financial performance measures, this is done to rule out the effect of value streams with slightly different market conditions. Some of the streams make products that have higher contribution margins due to more value-adding product features in these segments; the amount of value-adding features vary between value streams, but not within streams. Therefore, Contribution Margin 3 is indexed per value stream. As a result, even though the value streams may have significantly different contribution margins, this effect will not cause any disturbance to our results.

Because of the change of scales in the Lean audit reports, we create two Lean audit-indexes per value stream. As mentioned, the scale in the Lean audit reports expanded from 1–5 to 1–10. Unfortunately, this transformation is not linear and cannot be unified simply by dividing the latter by two. Fortunately, however, this only leads to the loss of the seven observations that now constitute the base for the second index of each value stream; this is the first Lean audit of 2008 per value stream. To match the performance observations, the eight performance measures are also transformed into two indexes per value stream.

After indexing both the Lean audit data and the performance data, we have (approximately) 86 cases for further analysis (i.e., 86 data points for each regression analysis). The total number of useable Lean audit reports for further analysis was 100. These 100 cases represent slightly less than four years, lasting from around 2005 to 2009, with quarterly Lean audit reports. We lost 14 cases due to the calculation of the baseline-indexes (seven value streams times two base numbers per stream). Eighty-six data points is an adequate sample size, as the regression analysis made below has significant test results, and because analysis of statistical power⁴ revealed no problems related to the regression analysis presented in table 5 below.

The descriptive statistics and correlations of the indexed Lean audit scores are presented in Table 2. Pearson correlation coefficients indicate several associations between the control form variables. They are all positively associated, except for one relation. Furthermore, many of the coefficients are significant.

<Insert Table 2 about here>

Table 2: Descriptive statistics and correlations based on indexing

We have also calculated a Euclidian distance index. This calculation is based on all seven areas of interest (control forms) in the Lean audits presented in Table 1. For each Lean audit measurement, e.g., second quarter 2005 in Value Stream A, the absolute Euclidian distance is calculated against the ideal case. In the ideal case, the maximum score is 5 (but is later 10) on each of the seven items assessed. The Euclidian distance represents a multidimensional or seven-dimension calculation of the difference between each Lean audit and the ideal case. These absolute Euclidian distances are then indexed for each value stream. This means that the Euclidian distance index shows the change in control forms compared to the first time it was measured for each particular value stream. For example, in the fourth quarter of 2004, the first Lean audit report was made for Value Chain B, and this measure is used as the index starting point. If the first quarter of 2005 has a Euclidian distance index of 98 (.98), it means that the distance to the ideal has dropped 2% and has thus come closer to the ideal state. According to our hypothesis, a lower distance should indicate improved performance.

In general, the average Euclidian distances dropped by around 15%, while the average Contribution Margin 3 Ratio index increased by 15% over the whole period. Over the same period, the data fluctuate up and down. In the middle of the period, the Contribution Margin 3 Ratio fluctuates up and down by more than

⁴ We used GPower 3.1 to compute a post-hoc statistical power analysis test.

20% compared to the average for the contribution margin ratio. The Lean audit scores also fluctuate up to 20% more than average in the middle of the period, and, a year later, are down by 20% below average. Therefore, it cannot be said that both the Euclidean distance and Contribution Margin 3 improved linearly from 2005 to 2009, as in between they fluctuated quite substantially up and down. Consequently, we do not suspect the overall increase in Contribution Margin 3 to be just a common learning curve effect.

We also index the city-block distance. Basically, we do the same transformations as for the Euclidean distance.

4.2 Euclidean distance regression tests of Hypothesis 1

We have performed tests to determine whether the correlations between the Euclidian distance and each of the eight performance measures can be found as expected. The results are shown in Table 3.

<Insert Table 3 about here>

Table 3: Linear regression (OLS) of Euclidian distance on performance

Table 3 presents support for our hypothesis that an improvement of the seven control areas of the Lean audit (seen by a decrease in the Euclidian distance index) leads to an increase in the Contribution Margin 3 Ratio (cf. CM-Ratio 3-index). Therefore, as anticipated, there is a negative correlation between them. In other words, the closer the value stream is to the ideal state, the better the Contribution Margin 3 Ratio. In our case, the adjusted R square is .12, which is generally acceptable for this kind of study. For example, in a study of World Class Manufacturing and activity-based costing, an adjusted R square of about .06 was found to be acceptable (Banker et al., 2008). Their research context is somewhat similar to our context, with both studies researching advanced manufacturing and organizational control.

The performance effect is also significant for both the wage ratio (direct wages divided by revenue) and the quality-cost-ratio (quality costs divided by revenue), with adjusted R square of .082 and .053, respectively, and in the expected direction, positive beta coefficients. The latter is in line with Fullerton and McWatters (2001), who found reduced scrap and rework as a consequence of Just-in-Time. However, there is no significant effect for the materials ratio (direct materials divided by revenue). This indicates that, statistically, improvements to the control forms in the Lean approach lead to reductions in direct wages and quality costs. The quality-cost ratio measures the cost of complaints and scrapped materials. Therefore, it is to be expected that the materials-cost ratio will not be improved much by improving the Lean control forms, as the effect it has on the materials costs is somewhat transferred to the quality-cost ratio. Even though some of the relations between financial and non-financial performance measures are statistically significantly related to the Euclidean distance (Lean audit scores), this cannot be claimed to be causal. Nevertheless, it is a statistical pattern that is supported by ex-ante theory.

Table 3 also presents a significant relation between the Euclidian distance and goal congruence, and the beta coefficient for the goal congruence measure is negative, as expected. Therefore, when the distance from the ideal state is increased, the goal congruence is reduced, and vice versa. As mentioned above, the goal congruence construct is part of the quarterly Lean audit, and it assesses (and scores) the extent to which employees have the corporate goal anchored in their actions. This is important, as the control forms are supposed to make the employees act in congruence with corporate goals. Fortunately, the financial indicators and the goal congruence results are aligned with each other. Thus, there is no incoherence in terms of employees acting goal-incongruently but still being able to provide improved financial results.

The non-financial performance indicators have less correlation with the Euclidian distance than do the financial measures. The delivery reliability performance has not been affected by the control forms. Even though the organization in general has improved (shortened) the Euclidian distance, it has not improved the delivery performance in the same manner. Some of the organizational members explained this by pointing to increased sourcing from China, with unreliable deliveries hampering the delivery performance measure. For the same reason, the measure of lead time between release of the order to production and packing of the order (Leadtime-release2pack-index) is subject to this “noise.” The other lead-time measure applied by the company (Leadtime-start2finish-index) was less sensitive to problems caused by the sourcing companies. The lead-time measure indicates the lead time from when the order is started in the production lines within a value stream to when the order is finished. Often, the order is packed when it is finished. The lead time from start to finish is positively correlated with the distance measure. This was also expected. When the value stream is far from the ideal state of the control forms, the lead time is also high. This lead-time measure thus supports the results of the financial performance measures.

All results in Table 3 have been created using linear regression models and they thereby assume linearity between the performance measures and the independent distance construct. Potential violation of this assumption has been checked by Pearson correlations and scatterplots. When CM3-ratio, wage-ratio, quality-cost-ratio, and goal congruence are used as the dependent performance measures, the Pearson correlations with the distance construct are significant at a p-level of .05. The lead time start2finish is a borderline case with a p-level of .089 and is thus significant at a p-level of .1. Scatterplots basically represent the same linearity, though with a slight tendency to be S-shaped, but as the Pearson correlations are significant, we stick to the linearity in our analysis above in order to keep it parsimonious. Scatterplots of the standardized residuals and predicted values show little or no indication of heteroscedasticity for all the regressions in Table 3 that are significant, i.e., the assumption of homoscedasticity in the simple OLS regression models is not violated. The Kolmogorov-Smirnov (KSL) test statistic of normality is .066, which is higher than the critical value of .04301 when the CM3-ratio is the dependent variable, and is thus non-significant at a p-level of .05. When wage-ratio is the dependent variable, the KSL statistic is .063, i.e., also non-significant at a p-level of .05, meaning low indication of non-normality of the residuals. This has also been assessed using plots. Consequently, the normality of residuals assumptions of the OLS regressions is not violated in these cases. However, for the three regression models with dependent variables—goal congruence, quality cost ratio, and lead time start2finish—there are significant KLS test statistics, indicating a violation of normality

of residuals for the models with these three variables. Therefore, these three latter models should be interpreted with some caution.

4.3 City-Block distance regression tests of Hypothesis 1

In this section, we analyze Hypothesis 1 based on the city-block distance. The advantage of using the city-block distance to test the hypothesis is the absence of balancing effects. The Euclidean distance could include some balance effects, which are not included in the city-block distance. These balance effects are described and analyzed in the next section, Section 4.4. Therefore, in order to fully accept Hypothesis 1, we supplement the traditional profile deviation analysis (which uses Euclidean distance) with a profile deviation analysis based on city-block distance.

The city-block distance is analyzed in the same manner as we analyze the Euclidean distance in the previous section.

< Insert Table 4 about here >

Table 4: Linear regression (OLS) of city-block distance on performance

Table 4 presents the results using the city-block distance index as the independent predictor variable and the same eight performance variables used in the previous section. The results are produced by linear OLS regressions. The conclusions made using the city-block distance index are similar to the conclusions made in the previous section. Hence, when Contribution Margin 3 Ratio is the dependent performance variable, the regression produces a significant (p-level of .0014) model. The equivalent adjusted R square is .059. This same similarity is found when wage ratio, quality-cost ratio, and goal congruence are used as the performance variable. These are all significant using the city-block distance, as when using Euclidean distance. Only lead time (start2finish) is not significant using city-block distance, whereas it was significant using Euclidean distance.

Because of the similarities to the conclusions found in the previous section, we come to the same conclusion in this section, which is that Hypothesis 1 is supported.

We have assessed whether there are any violations of the assumptions of the OLS regressions in Table 4. Pearson correlations are all significant at a p-level of .05 for relations between the city-block distance index and the four performance variables, with significant statistics in Table 4. This, together with scatterplot relations, indicates linearity. Furthermore, scatterplots of the standardized residuals and predicted values show little or no indication of heteroscedasticity for all the regressions in Table 4 that are significant, except for a couple of outliers when quality-cost ratio is the dependent variable. The Kolmogorov-Smirnov (KSL) test statistic of normality is .076, which is higher than the critical value of .04301 when the CM3 ratio is the dependent variable and is thus non-significant at a p-level of .05. When the wage-ratio is the dependent variable, the KSL statistic is .058, i.e., also non-significant at a p-level of .05, meaning low indication of

non-normality of the residuals. This has also been assessed using plots. However, for the two regression models with dependent variables of goal congruence and quality-cost, there are significant KLS test statistics, indicating a violation of normality of residuals for these models. Therefore, these two latter models should be interpreted with some caution.

There is one remarkable difference when comparing Tables 3 and 4, which is the substantially lower adjusted R square for the Contribution Margin 3 Ratio when city-block distance is the predictor compared to Euclidean distance as the predictor. In the first case, it is .120 (see Table 3) and in the latter, .059 (see Table n4). This could represent the absence of balancing effects in the city-block measure. For this, we provide a section of test results below in order to test Hypothesis 2. Section 4.4 will also be used to illustrate a refinement of the profile deviation analysis, compared to the more traditional use in Sections 4.2 and 4.3. This refined technique enables us to better assess whether there is a complementary effect on performance among the multiple Lean control forms with respect to our balancing of control forms.

4.4 New approach to distance measures testing Hypothesis 2

The Euclidian distance used in Section 4.2 integrates the balance between the control forms and the average (absolute) distance from the ideal state. This can be illustrated by a simple example. Say we only have two measures of control forms included, and they both have a score of 3 in the Lean audit. At a different point in time, the Lean scores are 2 and 4. In both cases, the average is 3. The ideal state is two scores of 5. In both cases, the distance to the ideal state is 4 (2 + 2 and 3 + 1). To reflect this distance, we require what we call “city-block” distance.⁵ Contrary to this, the Euclidian distance is not the same in both cases. The Euclidian distance to the ideal state is higher in the case with the scores of 4 and 2 than in the case with the scores of 3 and 3. The reason for this is that the Euclidian distance is equivalent to the hypotenuse, which is the shortest distance between two points in a co-ordinate system (Pythagoras: $a^2 + b^2 = c^2$). In the case with scores of 3 and 3, the Euclidian distance is 2.83, whereas with scores of 2 and 4, the Euclidian distance is 3.16. Therefore, even though the Lean audit scores at two different points in time have the same average, they may have different Euclidian distances to the ideal state. Thus, in the city-block distance measure, the degree of balance between control forms is left out and the degree of balance is only incorporated in the Euclidian measure. We need to separate these two effects to test Hypotheses 1 and 2 separately. The results are shown in Table 5.

Before we assess the results in Table 5, we need to develop an understanding of this refined statistical technique, providing an example operating with more than just the two control forms used above. Let us assume that we have a situation where a value stream is measuring output control (“goals and result control” in the Lean audit) to be at a high level, a score of 5. Likewise, we have a high level for behavioral control

⁵ Selto et al. (1995) calculate both the Euclidean distance and an absolute distance when testing their systems/configuration model. We think that their absolute distance is similar to what we label “city-block” distance. However, Selto et al. (1995) test Euclidean distance and absolute distance in two separate models. In our testing, we use city-block distance and the difference between the city-block distance and the Euclidean distance in the same model. Selto et al. (1995) do not use this *combination* of the distances, and they keep the distances separate in the two different models.

(“standardization” – SOP), say a score of 4. Training (“competences” – a social control form) is, however, very poor, with a score of 1. Thus, we have scores of 5, 4, and 1. This would produce a distance score of 5 for city-block distance and 4.123 for Euclidean distance, and this equals a difference of $-.877$, and when we divide this by the city-block distance, it equals $-.175$.⁶ In another situation, let’s assume the scores to be 5, 3, and 2 (output, behavior/SOP, and training). The latter situation would produce a city-block distance of 5, like the first, and a Euclidean distance of 3.605, which equals a difference of -1.395 , and when we divide this by city-block distance, we get $-.279$. Hence, the latter situation has a better balance, as represented in the divided difference between the Euclidean and city-block distances of $-.175$ and $-.278$, respectively. Hence, the lower the number, the higher the performance expected, i.e., beta in the regression analysis needs to be negative. The performance is expected to be higher in the latter case due to complementary effects, as the medium level of training would not thwart the performance effect of the other two control forms. This, however, will be the case in the first situation, where the synergy between the two high-level control forms is thwarted by the low level of training, due to the tight couplings between them, cf. the discussion in Section 2.

The high level of the output control form in the first situation would mean that the output numbers are updated frequently and that employees use them when trying to improve the company’s performance. As the SOP (behavioral control) is also at a high level, the standards are based on reliable, noise-free data, and everybody is aware of the current best practice. The high level of output and the SOP control forms will both affect performance positively, but they will also reinforce each other, representing a complementary effect on performance. The output numbers can, for example, make variations from the standards visible. These variations can be crucial to a Lean organization, as it works with low levels of inventory to buffer variation problems. Moreover, if the standards are very reliable, these variation numbers will be more reliable, and updating the standards can be used to document and sustain the new level of standards generated from improvement ideas. The ideas could stem from the output numbers. However, the “competence” level (training—a social control form) is low in the first situation. This low-level control form thwarts some of the performance effects coming from the two high-level control forms. The presence of low-level competences would reduce the possibility of generating valuable improvement ideas from the output numbers, and it would reduce the possibility of working according to standards. Furthermore, even if a few ideas are generated, the employees would be unable to make a skilled judgment about whether the ideas would actually affect future performance positively. If the competence level is at a more decent level, the employees would use their skills to take advantage of the output numbers, even if the SOP control form is reduced from a high level to a middle level. In that case, the standards would have some noise and would not be updated that frequently, but there would at least be some competences for utilizing this level.

The example shows how the control forms are quite tightly coupled - but are not completely tight. Having a completely tight coupled system would result in a low-level control form completely thwarting the performance effects. The result would equal assessing/predicting performance just using the control form (Lean audit area) with the lowest score, regardless of the level of the other control forms. However, in

⁶ We do not use the difference between the Euclidean distance and city-block distance in the tests carried out; we used the difference divided by city-block distance. A more detailed mathematical explanation is provided later in this section.

general, the more tightly coupled the control forms are, the higher the complementary effect, which we label “balancing the control forms.” As the Lean control package is quite tightly coupled, when the control forms are off-balance, the performance effects from complementarity between the control forms is more thwarted than if the control package was more loosely coupled. Off-balance refers to deviations from the ideal distance *between* the control forms. In our Lean context, off-balance is when the control forms are not on the same level on the 1–5 scale of the Lean audits. However, in other contexts and other configurations of control packages, off-balance may be defined differently. These other, or more general, contexts may benefit from our understanding of off-balance in conjunction with complementarities and tightly/loosely coupled control systems.

Kobayashi’s (1995) framework also recommended focusing on improvements to the low-level areas, as his framework is also based on the idea of synergy between the areas. Hence, this framework has the same foundation as we expect in Hypothesis 2 for a Lean organization.

To keep the numeric example above somewhat simple, we included only three control forms, but the effects are similar for more control forms.

<Insert Table 5 about here>

Table 5: Regressions (OLS) of city-block distance and Euclidian-block difference/city-block distance

The main conclusion, based on Table 5, is a confirmation of Hypothesis 2. This is based upon Contribution Margin 3 Ratio as a dependent performance variable in the OLS regression analysis. Table 5 is an extension of Table 4. In Table 5, we include not only the city-block distance as a predictor variable, but also the difference between Euclidean distance (index) and city-block distance (index), divided by the city-block distance (labeled “eucli-block difference/city-block index”). This new variable indicates the performance effect stemming from the control forms being in balance. Hence, it indicates the complementary effects on performance from multiple control forms in the Lean organization. This added variable increases the adjusted R square from .059 (see Table 4) to .232 (see Table 5), and the Beta score is as predicted. Thus, as Hypotheses 1 and 2 are both supported by the statistical results, it appears that the theoretically developed hypotheses are empirically sound. The tests show that the Lean control form package is tightly coupled, but is not completely tight, as we also support Hypothesis 1.

We base our conclusions on the model in which Contribution Margin 3 Ratio is used as a dependent variable. We consider this a “core” performance measure of the organization, as it directly reflects the goal of being profitable. Hence, it reflects goal congruence between the value streams and the owner of the company. Furthermore, this is a “hard” measure of performance, and we agree with Kihn (2005) that use of financial measures is generally preferred.⁷

⁷ We also tested for time-lagged models to see whether a lagged model could increase the adjusted R square. However, with the Contribution Margin 3 Ratio index as a dependent variable and with the city-block distance index and “eucli-block difference/city-block index” as predictors, neither of the time-lagged models ($t + 1$ and $t - 1$) improved the R square.

We have tested models with other dependent performance indicators, i.e., wage-cost ratio, quality-cost-ratio, and lead time “start2finish,” as these were significant using only the Euclidean distance index or the city-block distance index as the predictor variable. However, they do not yield significant results for eucli-block difference/city-block index as the predictor, and adjusted R square was not increased. This may be due to the lower adjusted R square (see Tables 3 and 4) before adding this extra predictor of the dependent variables, compared to the Contribution Margin 3 Ratio.

The reason we do not just use the Euclidean-block difference (Euclidean distance minus city-block distance) in Table 5 is mathematical. To illustrate, if we have three different situations with two control form variables, scoring 1 and 1; 4 and 4; and 4 and 2, the difference alone (without division) would be misleading. In the first situation (1, 1), the difference equals -2.34; in the second situation (4, 4), the difference equals -.59; and in the last situation (4, 2), it equals -.84. However, we want Situations 1 and 2 to be equal and Situation 3 to be lower than Situations 1 and 2, which is not the case. To obtain this result, a division of the difference between the Euclidean distances and city-block distance is required. Divided by the city-block distance, the results for Situations 1 and 2 are both -.29, and the result for Situation 3 is -.21. This new variable, “eucli-block difference/city-block index,” represents the effect of the control forms being symmetrical and in balance.

There is no sign of multicollinearity in the model in Table 5. The variance inflation factor (VIF) is very close to 1, meaning that there is no sign of multicollinearity. Furthermore, the condition number is below 10, which is generally considered acceptable as an indication of low multicollinearity. The studentized residuals plot shows a close to normal distribution (normality) of the residuals, as assumed in linear OLS regressions. Scatterplots of the standardized residuals and prediction values show low indications of heteroscedasticity, meaning that the assumption of homoscedasticity in the linear regression model is not violated. As the dataset is longitudinal, a test of autocorrelation is needed. To serve this purpose, we applied a Durbin-Watson test, which compares the level of the independent variables at points t and $t - 1$ in time. This is done per value stream. The tests show no significant measure of autocorrelation.

5. Conclusion

The findings of this paper are of both theoretical and practical importance, and we address both issues below.

Little research has been carried out on the topic of control forms in a Lean organization, despite the importance of Lean in many organizations today. To our knowledge, no study has quantitatively tested the impact and complementarities between various control forms in a Lean organization. Kennedy and Widener’s (2008) case study on this topic is the first to propose systems fit of control forms with synergetic complementarities (Gerdin and Greve, 2004). Their case study, however, did not provide quantitative analysis of the performance effects. Earlier case studies had difficulties in proving the effect of Lean in a quantitative manner (Selto et al., 1995). Our findings add to this modest base of research. Additionally, we

found evidence that the control forms do not have a purely additive (incremental), but rather a balanced complementary, synergetic effect on performance. Thus, increasing the level of a control form enhances performance, but if off-balance to other control forms is increased, some of the effects from the increased control forms will be thwarted, and in total this shows the complementary effects on performance. Hence, it would be better if both the level of the control form is closer to the ideal state, but it would also be better if the balance between them was the same as the balance in the ideal state. Consequently, our findings indicate that the controls of Lean should be seen as a control package. We consider this empirical result our first contribution.

An important addition to this first contribution is the effect this can have on theorizing and testing more general control form packages in the future. Analyzing and understanding the control package in the Lean organization is a more useful way of understanding how control forms can be balanced, and how off-balance can thwart the performance effects from complementarities. The more tightly coupled the control forms in the package, the more off-balance reduces the performance effects. The level of off-balance can be addressed as the deviations from the ideal distance between the control forms. Also, off-balance between the control forms will have a greater negative effect on performance in a tightly coupled system than in a loosely coupled system.

Our second contribution is the refinement of the statistical test approach. In tests of systems fit models, it has previously been recommended to use the Euclidian distance method (Gerdin and Greve, 2004) only. However, this technique does not allow us to distinguish the difference between enhanced performance that comes from independent variables that are in balance, or from an increase in their average level only. The Euclidian distance includes both phenomena. By using the city-block distance as well as the Euclidian distance simultaneously, it becomes possible to isolate the performance effects that stem from independent variables in balance and/or “just” at a higher average level. However, it should be noted that in our approach, the ideal state is at the maximum of scores for the control forms. The ability to use this distinction between the city-block distance and the Euclidean distance, where the ideal state is not at the maximum, minimum, or in the middle for all constructs, is beyond the scope of this paper.

A third contribution of this paper is the confirmation of management accounting model measurements being correlated to the “right” Lean behavior. For several years, researchers have debated whether management accounting models can capture the performance effects gained from Lean (Kaplan and Cooper, 1998; Maskell and Baggaley, 2004; Maskell and Kennedy, 2007). This is an important discussion for organizations working with Lean because they need a management accounting model that somehow reflects the results of what they believe to be the right behavioral patterns. Our case company has applied a (somewhat advanced) tiered contribution margin model, where the responsible units make up the value streams. The tiered, multi-level contributions margin helps to understand the variability of costs but also helps to understand the cost of quality, in our case the difference between Contribution Margins 2 and 3. In practice, it is difficult to trace the effects in the management accounting model from improvements in the Lean control forms, as it is difficult to trace the effects from complementarities between the control forms in, for example, trend curves

of costs. Our findings suggest that over a longer period of time, and with the right statistical techniques, the effects of Lean are traceable in traditional management accounting models.

Finally, we address how practitioners should interpret our findings, as recommended by Merchant (1984). Practitioners are concerned about how practice-defined variables can add to their companies' performance. Therefore, from practitioners' point of view, researchers should be better at researching practice-defined variables (in our case, Lean) and their performance effects. Our statistical findings suggest that managers should optimize their Lean organization by focusing on balance between control forms and by increasing the right use of them. Therefore, if a manager is in a position to decide whether to increase a low-level control form or increase a high-level control form, he should pick the former. Thereby, *ceteris paribus*, the manager would increase financial performance by making the control forms more balanced and not by just bringing about a higher average level of control forms. This complements the perspective of Otley and Berry (1980), who state that accountants should not only measure cost and profit, but should also use other variables in consequence prognosis. As such, the Lean measurements in our study (Lean audits) may be part of these other variables.

6. Limitations and future research

There are both strengths and weakness in basing our study on only one case company. A weakness is, of course, that we cannot claim that our case is empirically representative. On the other hand, sampling a sufficiently large number of Lean companies would most likely have forced us to use cross-sectional survey methods and, consequently, to rely on respondents' memories. The benefit of our approach is the access to a very rich amount of archival data extending across many years, collected using (almost) the same measurement instruments over the years (Lean audits and tiered contribution margins), especially instruments not created for research purposes. Consequently, the reliability is very high. Nonetheless, a cross-sectional survey addressing some of the same hypotheses could be interesting from an external validity point of view. Building a bridge between these types of studies would certainly improve validity, if they provide consistent results.

Using the Euclidian distance in a regression model with performance indicators as the dependent variables has certain limitations. Despite being recommended as the best statistical technique for addressing a systems fit model (Gerdin and Greve, 2004), its ability to measure the impact of individual variables is somewhat limited (Drazin and Van de Ven, 1985). However, this method is intended for systems fit models, and these models focus on the impact of a package (group) of independent variables on a dependent variable, rather than the impact of single independent variables.

A limitation of our study is that we cannot claim to have found causal relations between the control forms and performance. We can only claim to have found significant statistical relationships between the datasets. However, we have validated the results with the actors of the organization in order to test whether they match their perceptions of how their Lean program intends to affect performance. The actors believed in

synergy effects coming from the package of Lean audit scores on performance, but they have never applied sophisticated statistical techniques to test this.

7 Acknowledgements

The authors gratefully acknowledge the partial funding of the research project by a small group of Scandinavian companies and one of these is the case company in this paper. The case companies did not affect any choices made by the authors during the research process including data collection, analysis etc. Therefore, this research can be regarded as independent research. Moreover, the authors would like to thank Jan Greve for valuable comments on an earlier draft of the paper presented at the Manufacturing Accounting Research Conference, Helsinki, 13-15 June, 2012. We are also thankful for the comments and suggestions made by the two anonymous reviewers and the editor-in-chief.

References

- Abernethy, M.A., Stoelwinder, J.U., 1995. The role of professional control in the management of complex organizations. *Accounting, Organizations and Society*, 20 (1), 1–17.
- Abernethy, M.A., Brownell, P., 1997. Management control systems in research and development organizations: The role of accounting, behavior and personnel controls. *Accounting, Organizations and Society*, 22 (3/4), 233–248.
- Adler, P.S., Borys, B., 1996. Two types of bureaucracy: Enabling or coercive. *Administrative Science Quarterly*, 41, 61–89.
- Anderson, S., Widener, S., 2007. Doing quantitative field studies. In: Chapman, C.S., Hopwood, A., Shields, M.D. (Eds.). *Handbook of Management Accounting Research*. Oxford: Elsevier, 319–341.
- Balakrishnan, R., Linsmeier, T.J., Venkatachalam, M., 1996. Financial benefits from JIT adoption: Effects of customer concentration and cost structure. *The Accounting Review*, 71 (2), 183–205.
- Banker, R.D., Bardhan, I.R., Chen, T-Y, 2008. The role of manufacturing practicing in mediating the impact of activity-based-costing. *Accounting, Organizations and Society*, 33 (1), 1–19.
- Bicheno, J., 2004. *The new Lean toolbox: Towards fast, flexible flow*. Picsie books.
- Birnberg, J.G., Shields, M.D., Young, M.S., 1990. The case for multiple methods approach in empirical management accounting research (with an illustration from budget setting). *Journal of Management Accounting Research*, 2 (fall), 33–66.
- Boyd, B.D., Haynes, K.T., Hitt, M.A., Bergh, D.D., Ketchen, D.J., 2012. Contingency hypotheses in strategic management research: use, disuse, or misuse? *Journal of Management*. 38 (1), 278–313.
- Carreira, B., 2005. *Lean manufacturing that works: Powerful tools for dramatically reducing waste and maximizing profits*, USA: Amacon.
- Chapman, C., 1998. Reflections on a contingent view of accounting. *Accounting, Organizations and Society*, 22 (2), 189–205.
- Chapman, C.S., Ahrens, T., 2004. Accounting for flexibility and efficiency: A field study of management control systems in a restaurant chain. *Contemporary Accounting Research*, 21 (2), 271–301.
- Chenhall, R.H., 2003. Management control systems design within its organizational context: Findings from contingency-based research and direction for the future. *Accounting Organizations and Society*, 28 (2-3), 127–168.
- Chenhall, R.H., Langfield-Smith, K., 1998. The relationship between strategic priorities, management techniques and management accounting: an empirical investigation using a systems approach. *Accounting, Organizations and Society*, 23 (3), 243–264.
- Eisenhardt, K.M., 1985. Control: Organizational and economic approaches, *Management Science*,

31 (2), 134-149.

- Ezzamel, M., Willmott, H., 1998. Accounting for teamwork: A critical study of group-based systems of organizational control. *Administrative Science Quarterly*, 43, 358–396.
- Drazin, R., Van de Ven, A., 1985. Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30 (4), 514–539.
- Fullerton, R.R., McWatters, C.S., 2001. The production performance benefits from JIT implementation. *Journal of Operations Management*, 19 (1), 81–96.
- Fullerton, R.R., McWatters, C.S., 2002. The role of performance measures and incentives systems in relation to the degree of JIT implementation. *Accounting, Organizations and Society*, 27 (8), 711–735.
- Gerdin, J., Greve, J., 2004. Forms of contingency fit in management accounting research: A critical review. *Accounting Organizations and Society*, 29 (3-4), 303–326.
- Gerdin, J., 2005. Management accounting system design in manufacturing departments: an empirical investigation using a multiple contingencies approach. *Accounting, Organizations and Society*, 30, 99–126
- Gerdin, J., Greve, J. 2008. The appropriateness of statistical methods for testing contingency hypotheses in management accounting research. *Accounting, Organizations and Society*, 33, 995–1009.
- Govindarajan, V., 1988. A contingency approach to strategy implementation at the business-unit level: Integrating administrative mechanism with strategy. *Academy of Management Journal*, 31 (4), 828–853.
- Hopper, T., Jazayeri, M., 1999. Management accounting within world class manufacturing: a case study. *Management Accounting Research*, 10 (3), 263–301.
- Hult, T., Boyer, K.K., Ketchen D.J., 2007. Quality, operational logistics strategy, and repurchase intentions: A profile deviation analysis. *Journal of Business Logistics*, 28 (2), 105–132.
- Huson, M., Nanda, D., 1995. The impact of just-in-time manufacturing on firm performance in the U.S. *Journal of Operations Management*, 12 (3-4), 297–310.
- Kalagnanam, S.S., Lindsay, R.M., 1998. The use of organic models of control in JIT firms: Generalizing Woodward's findings to modern manufacturing practices. *Accounting Organizations and Society*, 24 (1), 1–30.
- Kaplan, R.S., Cooper, R., 1998. *Cost and effect: Using integrated cost systems to drive profitability and performance*. Boston: Harvard Business School Press.
- Kennedy, F.A., Widener, S.K., 2008. A control framework: Insights from evidence on Lean accounting. *Management Accounting Research*, 19 (4), 301–323.
- Kihn, L.A., 2005. Comparing performance measurement approaches in accounting research. *Finnish Journal of Business Economics (LTA)*, 2, 143–184.
- Kihn, L.A., 2007. Financial consequences in foreign subsidiary manager performance evaluations.

- European Accounting Review, 16 (3), 531–554.
- Kihn, L.A., 2010. An empirical investigation of the principal top management styles in the emphasis of multiple forms of controls. *International Journal of Accounting, Auditing and Performance Evaluation*, 6 (1), 108–127.
- Kinney, M.R., Wempe, W.F., 2002. Further evidence on the extent and origins of JIT's profitability effects. *The Accounting Review*, 77 (1), 203–225.
- Kobayashi, I., 1995. 20 keys to workplace improvement. Productivity Press.
- Kren, L., Kerr, J., 1993. The effect of behavior monitoring and uncertainty on the use of performance-contingent compensation. *Accounting and Business Research*, 23 (9), pp.159–168.
- Liker, J., 2004. *The Toyota Way*. New York: McGraw-Hill.
- Liker, J., Meier, D., 2006. *The Toyota Way Fieldbook*. New York: McGraw-Hill.
- Lind, J., 2001. Control in world class manufacturing: A longitudinal case study. *Management Accounting Research*, 12 (1), 41–74.
- Luft, J., Shields, M.D., 2003. Mapping management accounting: Graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations and Society*, 28 (2-3), 169–249.
- Malmi, T., Brown, D., 2008. Management control systems as a package: Opportunities, challenges and research directions. *Management Accounting Research*, 19 (4), 287–300.
- Maskell, B., Baggaley, B., 2004. *Practical Lean accounting*. Productivity Press.
- Maskell, B., Kennedy, F., 2007. Why do we need lean accounting and how does it work? *Journal of Corporate Accounting & Finance*, 18 (3), 59–73.
- Merchant, K., 1984. Influences on departmental budgeting: An empirical examination of a contingency model. *Accounting, Organizations and Society*, 9 (3/4), 291–307.
- Milgrom, P., Roberts, J., 1995. Complementarities and fit: Strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics*, 19, 179–208.
- Miller, D., Friesen, P.H., 1984. *Organizations a quantum view*. Prentice-Hall.
- Monden, Y., 2010. *Toyota production system*. Productivity Press.
- Otley, D., Berry, A., 1980. Control, organization and accounting. *Accounting, Organizations and Society*, 5 (2), 231–244.
- Ouchi, W., Maguire, M.A., 1975. Organizational control: Two functions. *Administrative Science Quarterly*, 20 (4), 559–569.
- Ouchi, W., 1979. A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25 (9), 833–848.
- Roberts, J., 2004. *The modern firm: Organizational design for performance and growth*. Oxford University Press.
- Sandelin, M., 2008. Operation of management control practices as a package – A case study on control system variety in a growth firm context. *Management Accounting Research*, 19, 324–343.

- Selto, F.H., Renner, C.J., Young, S. M., 1995. Assessing the organizational fit of a just-in-time manufacturing system: Testing selection, interaction, and systems model of contingency theory. *Accounting Organizations and Society*, 20 (7-8), 665–684.
- Shah, R., Ward, P.T., 2003. Lean manufacturing: Context, practice bundles, and performance. *Journal of Operations Management*, 21 (2), 129–149.
- Womack, J.P., Jones, D.T., Roos, D., 1991. *The machine that changed the world*. HarperCollins Publishers.
- Womack, J.P.; Jones, D.T., 2003. *Lean Thinking, Banish Waste and Create Wealth in Your Corporation*. New York: Free Press Simon & Schuster.
- Young, S.M., Selto, F.H., 1993. Explaining cross-sectional workgroup performance differences in a JIT facility: A critical appraisal of a field based study. *Journal of Management Accounting Research*, 5, 300–326.
- Venkatraman, N., Ramanujam, V., 1986. Measurement of Business Performance in Strategy Research: A Comparison of Approaches. *Academy of Management Review*, 11 (4), 801–814.
- Venkatraman, N., 1989. The concept of fit in strategy research: toward verbal and statistical correspondence. *Academy of Management Review*, 14 (3), 423–444.

Lean audit measures		
Label	Control form	How to improve it and be good at it..
Goal and result control	Output	<ul style="list-style-type: none"> - better goal definition in the cells. - employees know what is expected for the numbers to be at a good level. - non-financial scores (numbers) are measured and “maintained” in the cells every day such as units per hour, on-time delivery, lead time, downtime, absenteeism. - employees think of improvements based on the numbers.
Standardized work	Behavioral (SOP)	<ul style="list-style-type: none"> - work studies are made and diagrams are made. - they are continuously updated and reliable. - employees know what is value added and not.
Mapping	Behavioral (SOP)	<ul style="list-style-type: none"> - maps that show current state – current best practice flow of operations. - action plans to control how to get from current state to future state. - mapping ensures that actions plan are followed and behavior is not falling back to the “usual”.
Competences	Social (training)	<ul style="list-style-type: none"> - time is assigned for training. - cell-teams are initiating training themselves. - training is based on standard descriptions of work. - bulletin board with training level of employees
Continuous team improvements	Social (peer pressure and empowerment)	<ul style="list-style-type: none"> - teams should fix problems – not superiors. - teams are defining their own internal roles. - teams are active in hiring and firing. - teams try to combine own interest with the agenda of the company. - top-management only have a supporting role and act as coaches, i.e. improvements are driven by the employees.
Neat and clean	Social (visual)	<ul style="list-style-type: none"> - color marking (or text) of all things in the cells. - whiteboards are neat and outdated info removed. - no excess materials in the cells.
Coupled production	Social (visual)	<ul style="list-style-type: none"> - WIP between cells is located in color controlled areas. - whiteboards between the cells help coordination. - no walls or visual blocks between cells.
Management systems	Goal congruence – not a control form – but a performance construct	<ul style="list-style-type: none"> - better understanding of corporate goals. - goals are anchored at the employees. - employees and top-management believe they work together to reach the goals.

Table 1: Measures in the Lean audit

	CIT	Stan	CP	Map	Com	GRC		Mean	Std. Dev.	N
NaC	0.128	.289***	.056	.303***	.537***	.405***		1.351	.484	86
CIT		.284***	.499***	.228**	.426***	.459***		1.880	1.016	86
Stan			.192*	.747***	.271**	.081		1.472	.721	86
CP				.069	.137	.460***		1.892	1.002	86
Map					.122	-.068		1.818	.836	86
Com						.480***		1.330	.473	86
GRC								1.573	.636	86
***Correlation is significant at the 0.01 level (2-tailed).										
**Correlation is significant at the 0.05 level (2-tailed).										
*Correlation is significant at the 0.1 level (2-										
Labels:	NaC=Neat and Clean									
	CIT=Continuous team improvements									
	Stan=Standardized work									
	CP=Coupled production									
	Map=Mapping									
	Com=Competences									
	GRC=Goal and result control									

Table 2: Descriptive statistics and correlations based on indexing