Economies of scale in ICT: how to balance infrastructure and applications for economies of scale in ICT and business

Woudstra, U.A.

Citation for published version (APA):
6 DISCUSSION, LIMITATIONS AND CONCLUSION

6.1 Introduction

In this section we first discuss the research findings. Then the limitations of this research are analyzed in terms of different validity aspects. Finally we address the implications of the research findings in a concluding section.

6.2 Discussion

In summary, our findings substantially support our research model. They provide evidence which is consistent with the hypotheses $H_1$ and $H_2$ about (dis)economies of scale in ICT respectively business processes; this holds especially for Housing Corporations in combination with the Infrastructure Factor. In this research we performed a global and a detailed analysis concerning the validity of the hypotheses, all based on least squares. The global analysis consists of Partial Least Square (PLS) regression and linear regression to obtain a global view regarding the validity of the hypotheses; we assumed linear relations between the inputs and outputs of the ICT respectively business processes. In the detailed analysis we assumed functions of the form $y = a.x^b$ with $y =$ output and $x =$ input of the ICT respectively business processes. We determined the absolute productivity ($y/x$) and the relative productivity using Data Envelopment Analysis. The discussion below is based on the results of the detailed analysis.

The measurement of ICT expenditure has only for Municipalities and Hospitals resulted in higher validation values for FTE Operations than for Total ICT costs (see Table 5.21), which partly confirms proxy assumption $P_1$. However, for Housing Corporations the opposite holds; a possible explanation could be that the measurement of the FTE for operations and maintenance is insufficiently accurate, and that part of the workforce is spent on projects rather than on operations and regular maintenance. This is always a difficult distinction: for example, what about maintenance projects? Housing Corporations are relatively small organizations, using on average only 7 FTE for operations and maintenance, so there is in general no clear separation of functions.

We supposed in proxy assumption $P_2$ that the number of ICT assets would be a better approximation of the scale of ICT assets than the cost of hardware/software. Furthermore we expected in $P_{2a}$ that the number of workstations gives higher validation values than the combined scale measurement. In general both $P_2$ and $P_{2a}$ are validated, with two exceptions: in the situation of Housing Corporations ($P_2$) and Hospitals ($P_2$ and $P_{2a}$), with the Maturity Factor as efficacy criterion (see Table 5.21). Below we will discuss that the Maturity Factor has a low $H_1$-validation value compared with the Infrastructure Factor, which makes the rejection of $P_2$ and $P_{2a}$ in this case more explainable and less important.

The COBIT Maturity Factor as an ICT management policies criterion has an average $H_1$-validation value of just about 11 % of the average Infrastructure Factor validation value (see Table 5.19). The validation of $H_1$ could not be improved by using subsets of the 17 maturity
criteria (Figure 4.3). These results are in line with proxy assumption P3, which states that the maturity of the ICT organization is less important for the productivity of the ICT organization. As the COBIT maturity factor leads to very low H1-validation values, we can state that the results in the “MF” columns in Table 5.21 are less relevant than those in the “IF” columns.

Why are the Housing Corporations’ validation values higher than those of Municipalities and Hospitals? Note that only 16 of the 70 entries (23%) in Table 5.11 (Municipalities) show a statistically significant (p < 0.05) difference between the high and the low efficacy ICT management policies groups of Municipalities. For Hospitals this value is even lower: only 9 of the 60 entries (15%) in Table 5.15 show a statistically significant (p < 0.05) difference. These values are lower than the results for the Housing Corporations in Table 5.7, where 50 of the 100 entries (50%) show a statistically significant difference. The distinction between Housing Corporations, Municipalities and Hospitals is also demonstrated in Table 5.19, which shows an overall H1-validation value of 0.16 for the Housing Corporations versus 0.06 for the Municipalities and 0.05 for the Hospitals. Another difference between Housing Corporations, Municipalities and Hospitals can be found in Appendix 4, where the Housing Corporations could better be divided in groups based on the cybernetic view on ICT management policies (according to Figure 2.17) than the Municipalities and the Hospitals.

The explanation can probably be found in the following differences, which will be analyzed in more detail in section 6.3 (Limitations):

1) The Housing Corporations hardly show any variety in their business processes and are more mutually comparable than the Municipalities and the Hospitals, whose business processes show a larger variety.

2) The size of the benchmark: a number of 196 measurements (Housing Corporations) provides better statistical results than 55 (Municipalities) or 37 (Hospitals).

In Table 5.19 we can see that the H1-validation values calculated with Data Envelopment Analysis (DEA) are on average the same as those calculated with the absolute productivity analysis. For Housing Corporations the average DEA value is lower than the absolute productivity value (0.15 versus 0.18). For Municipalities on the contrary the average DEA value is somewhat higher than the absolute productivity value (0.07 versus 0.06). For Hospitals the average DEA value is the same as the absolute productivity value (0.05). A possible explanation for the differences between both techniques could be that DEA is sensible for measurement errors (Kitchenham 2002).

In this research a positive relation between ICT expenditure and the maturity of the ICT organization could hardly be validated. It is however likely that in organizations with a higher ICT budget the maturity of the ICT organization has more influence on the productivity of the ICT organization. In general can be stated that the higher the ICT budget, the higher the level of maturity of the ICT organization, the better the ICT organization is under control, the less errors are made and, the higher the level of productivity of the ICT labour is (Juran 1979). On the other hand, a higher level of COBIT maturity usually involves costs which could be equivalent to or even supersede possible savings (see Figure 2.16 for these opposing views). Further research is therefore required to analyze the effects at higher COBIT maturity levels.

In this research we have primarily analyzed the productivity of the ICT conversion process and secondarily the productivity of the business processes. In order to realize economies of scale in ICT and in business, the spending in infrastructure and applications must both be
sufficient. In Figure 3.7 the relation between the average Infrastructure Factor (IF) and Costs of ICT processes is represented: we assume that these costs decrease with increasing IF. On the other hand the Costs of Business Processes increase with increasing IF (and decreasing application spending, thus decreasing application functionality). Therefore there must be a minimum for the total costs (the sum of business and ICT) at a certain value of IF. We suppose there is a minimum value (Min) for IF below which the ICT expenditures rise more steeply and a maximum (Max) above which the business expenditures rise more steeply. Then there must be an area between Min and Max where total costs are more or less the same. We found for Housing Corporations that Min=0.42 and Max=0.53. For Hospitals these values are Min=0.54 and Max=0.58. In this research we could not determine a percentage for Municipalities, as the available data did not permit us to draw a conclusion. The above mentioned percentage should however be considered with care, as the individual situation of an organization may require specific investments in necessary functionalities. Besides, these percentages will change over time. Further research is required, in particular in the contribution of ICT to the performance of business processes (Brynjolfsson and Hitt 2003).

In this research we did not analyze the relation between the Maturity factor and the productivity of business processes, as the ICT/Business alignment is outside the scope of this research. As stated above the focus is primarily on the productivity of the ICT conversion process. We analyzed the business productivity only from the point of view of the upper limit of the infrastructure factor, to be sure that the organizations with high IF values that contribute to the high ICT productivity curve (Figure 4.7 graph 1), are the same organizations that contribute to the low business productivity curve (Figure 4.7 graph 2). We also found however that that the organizations with low IF values that contribute to the low ICT productivity curve, are the same organizations that contribute to the high business productivity curve. As the simultaneous analysis of H1 and H2 gives the same results as the separate analysis of H1 and H2, we can state that we reach a higher level of internal validity: we exclude the possibility that high/low ICT productivity is reached by other organizations than those that realize low/high business productivity. Validity aspects will be analyzed in more detail in the next chapter.

6.3 Limitations

In this section the validity of this research will be elaborated. First we will define the different validity aspects. Then we treat the question concerning the independency of the measurement variables of the constructs. Afterwards the influence of the quality of the empirical dataset on the different validity aspects will be investigated. Finally we will apply these ideas to this study, to determine the limitations of this PhD research.

6.3.1 Definitions validity aspects

In Figure 6.1 the different elements in the research process are represented. The starting point is made up be the phenomena in reality which are studied in the research, which determines the scope of the research. This is in terms of Popper (1972) part of World I. The Research model consists of the definitions of the constructs (efficacy of ICT management policies, ICT expenditure, scale ICT assets, organization performance), the hypotheses H1–H2 and proxy assumptions P1–P3. In Figure 3.10, Research model, we can see that the approximation of
constructs by variables is determined by the proxies P1-P3. The research model is in terms of Popper World II, as these are descriptions of the phenomena in reality. These constructs and hypotheses are based upon theories, e.g. Cobb and Douglas for economies of scale, which can be considered as a point of view to “look” at reality (see also Figure 2.12). These theories are in terms of Popper World III. In the Measurement model the constructs are represented by variables that are measured in reality with certain reliability. The Measurement model can be considered as a specification of the Research model. In the measurement model the measured variables fill in the (“specified”) hypotheses and proxies. On the other hand, by the use of the M&I dataset as secondary data source the measurement model imposes a restriction on the researchmodel. The measurement model is in terms of Popper also part of World II.

---

**Figure 6.1 Relation between reality, theory and models**

*Content validity* is the degree to which items in an instrument reflect the content universe to which the instrument will be generalized (Cronbach 1971; Rogers 1995; Boudreau et al 2001). This validity is generally established through literature reviews and expert judges or panels. The content validity can in terms of Figure 6.1 be defined as the degree until which the research model (consisting of definitions, hypotheses and proxies) is a good representation of the (perceived) reality (Leeuw 1996; Baarda and de Goede 2001). This type of validity can only be made plausible by the usage of general theories. For example the laws of gravitation cannot be proved to be valid, but every day again everybody is taking them into account as a well corroborated theory (Popper 1934). This is for hypotheses H1 and H2 determined by the question whether the theory of economies of scale according to the Cobb-Douglas function is an adequate representation of the reality in this case. For proxy assumptions P1-P3 there is no general theory to support the validity question. However, the cited references in this research should support the content validity of the proxies.
Construct validity is the extent to which an operationalization measures the concepts that it purports to measure (Straub 1989; Boudreau et al 2001). It asks whether the measures chosen are true constructs describing the event or merely artifacts of the methodology itself (Campbell and Fiske 1959; Cronbach 1971). The construct validity can in terms of Figure 6.1 be defined as the degree until which the measurement model (consisting of measured variables) is a good representation of the research model, given the measurement conditions (Leeuw 1996; Baarda and de Goede 2001). Within the scope of the measurement model the analysis of the empirical data is performed, which leads to the determination of the degree of validity of the hypotheses and of the proxy assumptions. For example hypothesis H1 has in Table 5.19 an average validation value of 0.16 for Housing Corporations. That means that the construct validity of this hypothesis has a value of 0.16, given the measurement conditions. The validation values as such depend on the definition of the validation value (which depends upon the values of b, Rsq and the number of years, see for example Table 5.8). The following remarks can be made concerning the construct validity:

- In section 4.3.1 we explained that the measurement according to the M&I cost model is based on a different definition of “Infrastructure” as a collection hardware components instead of a collection services including the integrative software. We concluded however that the M&I definition is better at the current level of data granularity. Only if more detailed data is available about the implementation of integrative infrastructural software, then it is worthwhile to measure the infrastructural software separately.

- The measurement model is “loaded” with data which is the result of measurement of reality with a certain reliability (see Figure 6.1). The validation values in Table 5.19 are based on the validation values in other Tables which all can be considered as construct validation values. High construct validation values are dependant on high reliability values in the Mann Whitney Tables; these values are ultimately based upon the reliability of the measurements of reality, which will be treated below.

- An important question regarding construct validity is (Straub 1989): “are the data a reflection of true scores or artefacts of the kind of instrument chosen”? In section 6.3.2 this will be elaborated in more detail.

Reliability is essentially an evaluation of measurement accuracy, for example the extent to which the respondent can answer the same or approximately the same questions the same way each time (Cronbach 1951). The measurement of the data in this research is already explained in Section 1.4.3: the measurements that are based on maturity data and on application availability data have the lowest reliability. The measurements with the highest reliability are based upon the cost data. This difference is also reflected in the validation value of the hypotheses: for example the IF-based analyses lead to higher validation values than the MF-based analyses. So we can state that the quality of an IF-based measurement model is higher than an MF-based model: an IF-based model is a better representation of reality than a MF-based model.

Internal validity raises the question of whether the observed effects could have been caused by or correlated with a set of un hypothesized and/or unmeasured variables (Straub 1989). In short, are there viable, rival explanations for the findings other than the single explanation offered by the researcher's hypotheses? The internal validity can in terms of Figure 6.1 be defined as the degree until which the measurement model is a good representation of the reality. Indeed, there could be an alternative explanation for high construct validation values, which are based upon a solid theoretical foundation that implies a high content validity. For example high investments in infrastructure can lead to neglecting investments in applications,
which can lead to insufficient functionality for users. We found that the simultaneous analysis of H1 and H2 gives the same results as the separate analysis of H1 and H2, so the possibility is excluded that high ICT productivity is reached by other organizations than those that realize low business productivity (and vice versa). By analyzing H1 and H2 simultaneously we realized a higher internal validity than by the separate analysis of H1 and H2.

External validity refers to generalizing across times, settings, and individuals (Cook and Campbell 1976; Sackett and Larson 1990). Sackett and Larson state that “external validity is the type of validity closest to our definition of generalizability”. The external validity can in terms of Figure 6.1 be defined (just as the internal validity) as the degree until which the measurement model is a good representation of the reality. In this case the focus is extending the scope of the investigated reality. We think that the results of this research are applicable to organizations with a low ICT budget, with a “utility” function of ICT (Weill and Broadbent 1998), that run the risk of neglecting investments in ICT infrastructure (caused by the rapidity of technological aging in ICT). Therefore the results of hypotheses H1 and H2 might be generalized for low ICT budget organizations. According to Lee and Baskerville (2003) generalization is however never possible without empirical verification, so we have to be very careful with this statement. The results of the other hypotheses (see Table 5.21) depend too much on the type of organization and therefore generalization is impossible.

Content validity, construct validity and reliability are necessary conditions for internal and external validity. In terms of Figure 6.1 we can state that the product of content validity (derivation of the research model from reality) and construct validity (derivation of the measurement model from the research model) and reliability (measurement of data in reality) determines the quality of the measurement model as a representation of reality. However, this is not a sufficient condition; in the concept of internal validity there is also an alternative modelling possible (see Figure 2.12, based on a different representation of reality); in the case of external validity we are dependant of the extension of the scope of the reality that is covered by the model (in terms of Figure 2.12 we could also speak of the scope of the representation of the reality).

6.3.2 Construct validity and the measurements of the constructs

In this section we will analyze some aspects of construct validity in more detail. The measurement model as defined in the foregoing section should ideally consist of variables that are orthogonal between different constructs. In Figure 6.2 the dimensions of the measurement variables of the 3 constructs of hypothesis H1 are represented on orthogonal axes X, Y and Z. Theoretically the correlation values of the inter-construct measures should be zero. However, in practice inter-construct correlation values up to 0,4 are considered acceptable (Bollen and Lennox 1991). There are in Figure 6.2 some combinations of inter-construct measurement variables that have to be considered with care:
- If Y is measured by Total ICT cost in year N and Z is measured by the average Infrastructure Factor, then Z contains cost factors of year N, that are also part of Y. However the correlation values between Y (ICT cost) and Z(IF), see Table 5.1, are -0,28 (HC), 0,18 (M) and 0,57 (Hosp). Only in the case of Hospitals this value is higher than 0,4 but the H1-validation values of Hospitals are lower than those of Housing Corporations and Municipalities, see Table 5.21. We discussed already in section 5.2 the possible
reasons why there is in the case of Hospitals a high correlation between IF and scale related variables.

- If X is measured by Cost of hardware/software in year N and Z is measured by the average Infrastructure Factor, then Z contains cost factors of year N, that are also part of X. In this case the correlation values between X(Cost HW/SW) and Z(IF), are -0.23 (HC), 0.19 (M) and 0.59 (Hosp), which is comparable with the situation described above.

- Note that we did already exclude the combination of Y(Total ICT cost year N) and X(Cost of hardware/software year N) in the research model, see Figure 3.10. In this case the correlation values between both variables are between 0.97 and 0.99.

- In section 4.4.1 we mentioned for Housing Corporations a correlation value of 0.51 between Z(MF) and X(Applications), which was explained by a subjective measurement procedure. However, the correlation between Z(MF) and X(Combined scale) has the value 0.25 (see Table 5.1) which is lower than 0.4.

The intra-construct variables should ideally be highly correlated in the case of “reflective” variables: in section 4.4.1 we already discussed this point and calculated a value of Cronbach’s alpha of 0.94 for Y(ICT cost) and Y(FTE Operations). The other intra-construct variables (in case of more than one variable per construct) are “formative” which does not demand a high level of correlation. Indeed, in Figure 6.2 there are 2 alternative measurement variables for each dimension, which are treated independently. Only in the situation of the composite variable (see section 4.3.2) scale ICT assets = f(workstations, FTE, applications), the correlation between these 3 variables is important. For Housing Corporations the value of Cronbach alpha appears to be 0.75 (N=196, N of items is 3).

Figure 6.2 Dimensions constructs measurement hypothesis H1
Note that in the case of hypothesis H2 we do not have these problems with cost factors that are part of measurement variables of different constructs. In Table 5.1 we can see that the correlation values of inter-construct measures between Turnover and the other variables are comparable with the correlation values of inter-construct measures between ICT cost and the other variables, which was already discussed above.

### 6.3.3 Statistical power and validity aspects

In this section we will relate the different validity aspects with the concept of statistical power. The power of any statistical test of a null hypothesis is a function of the following three parameters (Baroudi and Orlikowski 1989):

1. The significance criterion \( a \), which is the chosen risk of incurring a Type I error, and whether the test is directional (one-sided) or non-directional (two-sided). Power increases with larger \( a \) and with directional hypothesis tests. In this research Mann Whitney was used since ANOVA starts from a normal distribution and is therefore not always correct. Because Mann Whitney does not have this requirement, it was more suited for our research study. A statistically significant difference between the productivities in two groups was assumed when Mann Whitney \( p \) (two sided) < 0.05.

2. The precision of sample estimates, which is primarily influenced by the sample size \( n \). The larger the \( n \), all else being equal, the smaller the error, and the greater the precision, which increases the probability of rejecting the false null hypothesis. In this research the biggest sample size was realized by the dataset of the Housing Corporations. However, the datasets of Municipalities and Hospitals achieve additional statistical power.

3. The effect size, which represents the magnitude or strength of the relationship among the variables in the population. If other factors are controlled, the larger the effect size, the greater the degree that a phenomenon manifests itself, and the greater the probability that it will be detected and the null hypothesis rejected. In this research we have a limited number of variables and relations to describe a complex phenomenon in reality. We will analyze this problem in more detail below and relate this to the different validity aspects.

The quality of the measurement model as a representation of reality depends on the quality of the empirical dataset. In an ideal situation the dataset would contain data of all relevant variables (\( Z_2 \) in Figure 6.3) of some phenomenon, for example economies of scale in low ICT budget organizations. Then all organizations of this type would be for example the organizations with low ICT cost (for example < 10% of total cost). We would be able to construct an isomorphic model of the relevant part of reality that is able to explain completely this part of reality. The dataset would contain data of enough organizations (\( X_2 \) in Figure 6.3), measured with a sufficient level of reliability (\( Y_2 \) in Figure 6.3). However, we are limited to a dataset that contains data of a limited number of organizations (\( X_1 \)) concerning a limited number of variables (\( Z_1 \)), measured with a limited level of reliability (\( Y_1 \)).

In the “ideal” situation (\( X_2, Y_2, Z_2 \)) the content validity is 100%, as the research model is an isomorphic model of the relevant part of reality (there is no relevant ignorance at a conceptual level). The construct validity is 100%, as the measurement model is an isomorphic model of the relevant part of reality (all variables are mutual independant and there is no relevant
ignorance at the data level). The internal validity is 100%, as there are no competing theories that can explain the phenomenon in reality, because the measurement model is perfect (there is no relevant ignorance about economies of scale in ICT in low ICT budget organizations). The external validity is 100%, as the measurement model can explain perfectly the behaviour of the phenomenon in all relevant organizations (organizations with low ICT cost). Therefore we can define the quality of the ideal dataset as being 100%. In Figure 6.3 this can be visualized by stating that the quality of 100% belongs to point (X2, Y2, Z2).

In situation 1 (limited dataset) the content validity is less than 100%, as the research model is an incomplete model of the relevant part of reality. There is an amount of relevant ignorance at a conceptual level: in our situation we have only a Cobb-Douglas function, which is an incomplete model of the phenomenon of economies of scale in ICT. The construct validity is less than 100%, as the measurement model is an incomplete model of the relevant part of reality. The measured variables are determined by the definitions of the variables in the secondary dataset of M&I/Partners, that are just approximations of the defined constructs. These constructs cannot be completely measured, for example there is no definition of the level of knowledge of the users in the M&I dataset. The internal validity is less than 100%, as there are possibly competing theories that can partly explain the phenomenon in reality, because the measurement model is incomplete (there is an amount of relevant ignorance about economies of scale in ICT in low ICT budget organizations). The external validity is less than 100%, as the measurement model can only partly explain the behaviour of the phenomenon in a limited number of organizations (a number of Housing Corporations, Municipalities and Hospitals with ICT cost less than 10% of total cost). We can determine the quality of the limited dataset in Figure 6.3 the point (X1, Y1, Z1). A value of the quality of the limited dataset might for example be defined as the quotient 100%*(X1*Y1*Z1)/(X2*Y2*Z2).

An example will elucidate these concepts. The efficacy of ICT management policies as construct is defined by the variables Infrastructure Factor (IF) and Maturity Factor (MF), as part of the research model. These variables are measured in a certain way as part of the measurement model.

Suppose firstly that we find an alternative way to measure MF. This leaves the research model unchanged, as the definition of the construct and the variables is the same. The measurement model is changed, as there is an alternative way of measurement of a variable. In this example the content validity remains the same, as the constructs and variables are unchanged. The construct validity is higher when this appears to be a better way to measure this variable. The internal validity is higher, as there is less room for alternative models. The external validity is higher if the behaviour of more organizations can be explained correctly.

Suppose secondly there would be another variable to define efficacy of ICT management policies, besides IF and MF, for example an independent Skill Factor (SF), which enriches the research model. Also the measurement model would be more complete, provided that SF can be measured correctly. In this example the content validity is higher, as the definition of the constructs is changed. The construct validity is higher, as there are more variables to measure this construct. The internal validity is higher, as there is less room for alternative models. The external validity is higher if the behaviour of more organizations can be correctly explained.
X = number of organizations; Y = measurement reliability; Z = number of variables
X2 = sufficient organizations; Y2 = sufficient reliability; Z2 = all relevant variables

Figure 6.3 Quality of the empirical dataset

Summarizing we can state that the content validity depends upon the validity of the research model. The more variables (representing constructs) are defined in the research model, the better the research model should represent reality (provided that these variables are relevant). So content validity = f (Z).

The construct validity depends upon the validity of the measurement model. The more variables are defined in the measurement model and the higher the reliability of the measurement, the better the constructs can be measured (provided that the variables are relevant and independent). Furthermore, the more organizations are available in the dataset, the higher the probability to attain statistical significance and the better the validity of the measurement model. Therefore construct validity = f (X,Y,Z).

The internal validity depends both on the validity of the construct model and the measurement model. Therefore internal validity = f (X,Y,Z).

The external validity depends both on the validity of the construct model and the measurement model. Therefore external validity = f (X,Y,Z).

6.3.4 Validity aspects of this PhD research

We will apply the model in Figure 6.3 on this research. First we will evaluate the definition of constructs and the measurement of variables related to these constructs (corresponding with the dimensions Y and Z in section 6.3.3):

The efficacy of ICT management policies is defined in two ways:
- The infrastructure policy is measured by just one financial indicator, the Infrastructure Factor (IF). More infrastructure variables would raise construct validity.
The maturity of the organization is measured in a subjective way by a compound measurement, the *Maturity Factor (MF)*. A more objective measurement of this variable would raise reliability.

The ICT expenditure, the input of the ICT conversion process, is measured in two ways:
- The number of ICT *personnel* for operations and maintenance. Determination of this number can be executed objectively, provided that also the decentralized personnel is counted and that innovation hours are not counted as operations. Solving this problem in small organizations like Housing Corporations (as discussed in section 6.2) would raise reliability.
- The *total ICT costs* can be determined objectively and in a reliable way, provided that the accounting procedures are correctly executed by the accounting department.

The *Turnover*, the output of the Business processes, can be determined objectively and in a reliable way, provided that the accounting procedures are correctly executed by the accounting department.

The scale of ICT assets, the output of the ICT conversion process, is measured in three ways:
- The *number of workstations* as a representation of the scale of the infrastructure and the scale of the users, based on the assumption that the application variety is equal for all organizations of the same type:
  \[
  \text{Scale ICT assets} = \text{number of workstations}
  \]
  This measurement gives on average the highest hypotheses validation values.
- *Combined scale factor*: A combination of the number of workstations (scale infrastructure), the weighted number of application types (the variety of applications in the test environment determines the scale of applications on the desktop), and the number of FTE of the organization (scale users). Formally defined, according to the distance-based approach a compound measurement is conducted as follows (Poels and Dedene 2000):
  \[
  \text{Scale ICT assets} = \sqrt{\text{workstations}^2 + \text{FTE}^2 + \text{applications}^2}
  \]
  The statistical significance appears to be lower than in the foregoing approach (with only the number of workstations as scale of ICT assets). A better measurement of the variety and scale of applications and the number of users with knowledge would raise construct validity.
- The *TIR cost* (the cost of hardware / software) as a measurement of the scale of ICT assets. This approach gives lower levels of statistical significance in the analyses, compared with the number approach of ICT assets. The difference between the “ontological” (based on numbers) and the “descriptive” (based on cost HW/SW) approach is discussed in considerable detail in section 3.4.3: this explains the lower level of validity of the descriptive approach.

The number of variables in the research model is the same for Housing Corporations (HC), Municipalities (M) and Hospitals (Hosp). However, HC are less complex organizations than M and Hosp. Therefore the “ideal” description of HC needs fewer variables than M and Hosp. Thus the value of Z for HC is greater than the value of Z for M or Hosp. Formally written: \( Z(\text{HC}) > Z(\text{M}) \) and \( Z(\text{HC}) > Z(\text{Hosp}) \). In Table 6.1 \( Z(\text{HC}) = \text{“low”} \) and \( Z(\text{M}) = \text{“very low”} \), which also holds for \( Z(\text{Hosp}) \).

The measurement reliability is the same for HC, M and Hosp. Formally written:
\[
Y(\text{HC}) = Y(\text{M}) = Y(\text{Hosp})
\]
In Table 6.1 this is indicated as “comparable”.

The number of measurements of Housing Corporations \( X(\text{HC}) = 196 \). For the other organization types: \( X(\text{M}) = 55 \) and \( X(\text{Hosp}) = 37 \). These numbers are represented in Table
6.1. The value of X(HC)\*Y(HC)\*Z(HC) is indicated as “high”, to differentiate it from the value for M (“low”) and Hosp (“very low”).

The value “p<0.05” in Table 6.1 is an indication of the percentile of measurements with p<0.05 in the Tables 5.7 (HC), 5.11 (M) and 5.15 (Hosp) concerning the statistically significant difference between the high and low efficacy ICT management policies groups. The relation in Table 6.1 between “X\*Y\*Z” and “P<0.05” is straightforward, given the explication of the statistical power model in the foregoing section. The same holds for the ultimate H1 validation “H1 val” in Table 6.1.

Table 6.1 Relation between statistical power model and H1 validation

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>X*Y*Z</th>
<th>p &lt; 0.05</th>
<th>H1 val</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>196</td>
<td>low</td>
<td>comparable</td>
<td>high</td>
<td>0.50</td>
<td>0.16</td>
</tr>
<tr>
<td>M</td>
<td>55</td>
<td>very low</td>
<td>comparable</td>
<td>low</td>
<td>0.23</td>
<td>0.06</td>
</tr>
<tr>
<td>Hosp</td>
<td>37</td>
<td>very low</td>
<td>comparable</td>
<td>very low</td>
<td>0.15</td>
<td>0.05</td>
</tr>
</tbody>
</table>

We can consider “H1 val” as a measure of the internal validity of this research, which can be described as “low” for HC and “very low” for M and Hosp, as there are very much alternative explanations possible caused by variables that are not measured in this research. Therefore the external validity is limited to Housing Corporations, Municipalities and Hospitals.

6.4 Conclusion

This study offers new evidence for and insights into the economies of scale of ICT departments with low ICT budgets and a “utility view” on ICT (Weill and Broadbent 1998). We have concluded that these lower ICT spending organizations should spend a certain minimum of their ICT expenditure on infrastructure in order to attain economies of scale in the ICT conversion process. The measured economies between low and high infrastructure spending organizations are on average more than 20% for operational ICT labour and for Total ICT costs.

In this research the relationships between the efficacy of ICT management policies, ICT expenditure, and the scale of the ICT assets have been analyzed. The efficacy of ICT management policies has been defined as the ability of ICT management to govern an efficient and effective conversion from operational ICT expenditure to ICT assets. This implies that effective policies create the conditions for an efficient and effective ICT conversion process. Efficacy therefore refers to the optimal utilization of both technological resources and labour (or skills), which is increasingly important in order to leverage commodities. Our basic hypothesis is that a higher efficacy of organizations’ ICT management policies results in lower levels of ICT expenditure, given a similar scale of ICT assets. Using empirical data on 74 Housing Corporations, 22 Municipalities and 36 Hospitals, support has been found to demonstrate economies of scale with respect to ICT assets for (comparable) organizations with a high efficacy of ICT management policies. Furthermore, we have demonstrated diseconomies of scale for organizations with a low efficacy of ICT management policies. The theoretical contribution of this research lies in the presentation of a definition of the efficacy of ICT management policies and the measurement of this efficacy.
In addition, a new methodology has been introduced to analyze the relation between the efficacy of ICT management policies and cost savings in ICT by economies of scale. In this research study infrastructure expenditure appears to be the most important ICT management policies criterion. The organizations investigated in this research (Housing Corporations, Municipalities and Hospitals) have low ICT budgets and have a “utility” view on ICT. According to Weill and Broadbent these organizations should concentrate on infrastructure and transactional applications to attain economies of scale. This confirms the general notion that a certain level of ICT infrastructure is a necessary condition for an efficient ICT management. This study demonstrated that ICT expenses grow at a higher rate compared to the scale of the ICT assets whenever the level of infrastructure expenses is insufficient.

The maturity of the ICT organization formed the second ICT management policies criterion investigated in this research. However, a positive relation between the productivity of the ICT organization and the maturity of the ICT organization, as measured according to COBIT 4.0 (2005), could hardly be validated. This means that for the organizations investigated, the COBIT maturity concept cannot serve as a real differentiator between low or high efficacy of ICT management policies in the sense of the level of ICT expenses at a certain scale of ICT assets. The conclusion of this study is that (for organizations with a low level of ICT spending) a minimum level of ICT infrastructure is more important than COBIT maturity. It is suggested that further research is conducted to find out the percentage of ICT spending as part of the total costs above which a higher COBIT maturity would result in lower ICT expenses, given a similar scale of ICT assets.

We suggest that managers of ICT departments in lower ICT spending organizations start considering ICT-infrastructure investments as a prerequisite for achieving economies of scale in ICT management. There should be equilibrium between the business demand for more applications and the necessary infrastructure for the integration of these applications into the whole ICT architecture. In a situation where ICT is hardly more than just a utility, and where the authority of the ICT manager is generally only limited, the findings of this study may help the ICT manager convince the other managers in the organization to start focusing on infrastructure and transactional applications (Enns et al 2001; Enns et al 2003).

In this research we have primarily analyzed the productivity of the ICT conversion process and secondarily the productivity of the business processes. In order to realize economies of scale in ICT and in business, the spending in infrastructure and applications must both be sufficient. Thus there must be not only a lower limit but also an upper limit for the spending in ICT infrastructure. We found for Housing Corporations a lower limit of 42% of total ICT costs and an upper limit of 53%. For Hospitals these values are 54% and 58%. For Municipalities the available data did not permit us to determine these percentages. These values should be considered with care, as they change over time and for individual organizations it can on a certain moment be necessary to invest in infrastructure or applications.

Weill and Broadbent (1998) argue that lower ICT spending organizations with a “utility view” on ICT should spend relatively more on their ICT infrastructure. This study confirms their view and quantifies their ICT infrastructure findings. Furthermore, this is the first study that applies in-company firm data, which are superior to stock data, to this type of analysis.
Moreover, our study contradicts Carr’s notion of ICT as a commodity: even in our lower ICT expenditure level organizations there appear to be significant differences in the application and competitive use of ICT. It would be worthwhile to benchmark organizations in other sectors based on the Infrastructure Factor. Besides it would be interesting to compare the (results of) Infrastructure Factor benchmarking with other benchmark methods.

The degree of validity of hypothesis H1 is higher in the case of the Infrastructure Factor than in the case of the Maturity Factor based on 17 COBIT/ITIL aspects. Furthermore, we have seen that a subset of these aspects does not lead to a higher degree of validity. It would be worthwhile to investigate how another definition and measurement of the Maturity Factor could lead to a stronger validity.

In this research we demonstrated that standardization of ICT components and ICT processes is from a theoretical point of view the most important management policy to realize economies of scale in ICT conversion process. We did not analyze the effect of standardization of ICT assets on the productivity of business processes. We believe however that business process redesign, based on standardized applications has a positive effect on economies of scale in business processes. It would be worthwhile to investigate the combined effect of standardized ICT assets on economies of scale in the ICT conversion process and in the business processes.

Although in many respects ICT has proven to be a commodity to the organizations investigated in this research, its effective deployment is by no means straightforward. In this study the efficacy of ICT management policies appeared to be highly dependent on the percentage of ICT infrastructure investments. In other words, large scale infrastructure investments are a prerequisite for the efficient use of ICT (and not vice versa). The analysis presented in this study has clearly produced new and valuable insights into the realization of effective ICT management policies.