I.T. INVESTMENT AND COMPLEMENTS IN VALUE CREATION: A META-ANALYTIC REPLICATION

Pankaj Nagpal, Case Western Reserve University, Cleveland, Ohio, USA, pxn18@case.edu
Kalle Lyytinen, Case Western Reserve University, Cleveland, Ohio, USA, kalle@case.edu
Susan Helper, Case Western Reserve University, Cleveland, Ohio, USA, sxh23@case.edu

Abstract

We execute four theoretically motivated replications of earlier analyses on business value of information technology (IT). We classify firm level IT value research and select alternative studies to understand and the core logics of current research and its sensitivity to used measures. Simultaneous replication of several models enables generalization across a variety of constructs, and abstracts their findings as is commonly done through meta-analyses. An access to a common dataset and availability of well matched analogs allow execution of these replications. Our results show robustness of some constructs across studies in predicting aspects of IT value. The models also offer theoretical insights of plausible interactions between the technology artifact, organizational characteristics and productivity outcomes. These results are preliminary due to a limited dataset, which was not customized to the proposed replications. Extensions to current theories are suggested which would enable better explanation and prediction.

1 INTRODUCTION

Although research in IS economics has theorized on the productivity effects of IT investment successfully over a decade, comparisons across studies are difficult at best. In this paper, we review and replicate firm level analyses to validate the robustness of findings across studies that used different predictors. We classify analyses into different model types, and discuss the value of meta-analyses and replications. To overcome current limitations of meta-analyses, we explore the following questions: what constructs can be used to assess whether IS adds value to a firm? How can economic impacts of IS be theorized by extending ‘unique’ factors? What aspects of theoretical explanation are replicable across ‘types’ of IT artifacts? The approach uses an extensive dataset with a large number of constructs to draw conclusions about the robustness of findings across studies. The paper is organized as follows. We first review implications of the firm-level literature on IT value. This includes a critical review of research approaches, which is developed into a discussion of key constructs. The role of meta-analyses and replication in theory validation is discussed, and the logic for selection of productivity models is given. The dataset and survey sample are introduced subsequently, followed by a replication of productivity analyses that used alternative models. We conclude with implications of our findings for future research.

2 RELATED RESEARCH

The ‘productivity paradox’ has been resolved with several studies (Brynjolfsson & Hitt, 1993; Lichtenberg, 1995) that used comprehensive datasets and powerful econometric techniques. As a result, available empirical evidence shows positive and significant productivity effects associated with IT investment. Since then, research interest has shifted to understanding the role of complementary assets that work with IT investment. These factors had initially been studied (Barua, Kriebel, & Mukhopadhyay, 1995; Kelley, 1994) only at relatively finer levels of organizations. As of now, they...
still represent the key questions for researchers. The following table summarizes what we know, and do not know.

<table>
<thead>
<tr>
<th>Topic</th>
<th>What we know…</th>
<th>And do not know…</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Returns’ to IT investment</td>
<td>Returns from IT investment are generally positive.</td>
<td>How is IT investment conceptualized?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What is/should be included in cost?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>At what level of analysis, and time horizon, is return estimated?</td>
</tr>
<tr>
<td>Complementary characteristics</td>
<td>Organizational practices show interactions with IT investment.</td>
<td>Which organizational characteristics or practices impact the effects?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>How do they work?</td>
</tr>
<tr>
<td>Industry structure</td>
<td>There are strong environmental and industry effects.</td>
<td>How does industry/environment affect the returns?</td>
</tr>
</tbody>
</table>

Table 1. What do we know and do not know of IT productivity

Organizational characteristics are complementary to IT investment, and have been the main focus of recent research. In general, these studies identify multiple complementary capabilities by considering factors that interact with value addition processes. These include training that helps employees use computers in their work (Hempell, 2002), and the general profile of employees, e.g., education and skills. Managerial intensity has been proposed (Francalanci & Galal, 1998) to differentiate more productive service firms. Workplace practices in manufacturing firms (Black & Lynch, 1997; Kelley, 1994) such as TQM and benchmarking, have been claimed to improve productivity along with IT investment. Work organization (Bresnahan, Brynjolfsson, & Hitt, 2002) as determinants of productivity show significant interactions with IT investment. Yet, we do not know the typical set of complementarities a priori, and how they work individually, or in combination with other investments.

In the paper, we consider only those studies which examine the firm as a unit of analysis. So we do not cover studies that have used establishments or production plants as unit of analysis (Black & Lynch, 1997). The major distinctions that we handle are cross sectional versus longitudinal research designs, type of dependent variable, the use of complements, and the treatment of industry specificity.

2.1 Research Design

Longitudinal designs have been used to control for the ‘unobserved’ effects across years, and tracing the lag effects. Although simpler cross sectional designs have recognized positive effects of IT investment, there is more consistency in longitudinal studies, as shown by the sign and magnitude of regression coefficients. In contrast, taking a snapshot of the phenomenon at a given point has resulted in conflicting results, sometimes even within the same study. For example, Byrd and Marshall (1997) concluded that the ratio of PCs to employees had a positive effect, while the processor value of IT (as % of revenue) and the share of IT budget spent on IT staff training had negative effects. These conflicting results were seen even with a common dependent variable, which suggests a lack of robust constructs. This suggests that different variables and ratios have different statistical effects even in the same model. We can see somewhat higher consistency in other cross sectional studies. As these are relatively ‘objective’ econometric measures, not susceptible to raters’ biases or perceptual measurement challenges, the results are all the more surprising.

2.2 Dependent Variable

There are many ways of measuring the effect of technology and complementary assets, over and above productivity and profitability, in which there is higher clarity on productivity effects (Kohli & Devaraj, 2003). The dependent variable is an overall measure in some studies, using value added or sales, either at the firm or SBU level. Other researchers have conceptualized the result as a ‘per capita’ effect, adjusted for the number of employees. Using financial ratios is yet another dependent variable used,
which are difficult to categorize into productivity or profitability. Moreover, the ratios tend to be specific to the industry (Francalanci & Galal, 1998) or the sample. Hence the industry specific ratios are difficult to compare, and therefore to generalize. A majority of cross sectional studies use per capita effects. Overall, there is some evidence that cross sectional surveys, using per capita DVs, can overcome the limitations of being a snapshot in time. In a logarithmic form, using per capita DV is equivalent to adding an additional control in the regression equation, which brings these (longitudinal and cross sectional) designs to common ground.

2.3 Independent Variables

There is greater variation among complementary measures as compared to those in technology constructs. The large variety of organizational capabilities makes it difficult to conceptualize the nature of causal relationships of independent factors that cause productivity effects. In selecting models for replication, we adapt Black and Lynch (1997) way to classify independent variables into sub-groups. In addition to technology characteristics, the sub-groups are: 1) characteristics of workers, 2) existence of work systems, 3) existence of employee organizations, and 4) profit sharing schemes. In the selected studies, the majority include only technology-related IVs, making their results more comparable. These less complex studies show the positive effects of IT investment, with a few exceptions (Byrd & Marshall, 1997; Francalanci & Galal, 1998; Loveman, 1994). Hence these exceptions are interesting from the point of view of research, as to whether these reflect only idiosyncrasies of the constructs or data, or truly different results that will have stronger implications for theory. Additional analyses will therefore have more value, if they help test robustness of the exceptional findings. Depending on these results, we can draw stronger implications for the role of technology investments. Among worker characteristics; skill, education, and training are seen to increase productivity with IT investment. The exception is the study at a very fine level of analysis (Kelley, 1994) in which unit production hours are seen to increase with wages and technical education. Existence of ‘high performance’ work systems and sharing schemes were hypothesized in Black and Lynch (1997), and not been used elsewhere. Employee organizations such as work groups (Black & Lynch, 1997) and labor management committees (Kelley, 1994) show conflicting results. Work organization, a survey-based measure of the use of teamwork and decision authority among workers (Bresnahan et al., 2002) also showed positive effects. Hence the organizational complements are candidates for rigorous testing and replication as to whether the results are valid under other conditions. There is a high level of sensitivity of the results to the raw complementary measures used in the study, regardless of its type.

2.4 Diversity of sample

Some studies cover a single industry (Francalanci & Galal, 1998; Mukhopadhyay, Kekre, & Kalathur, 1995; Prasad & Harker, 1997; Weill, 1992) which makes them relatively free of confounding influences of inter-industry variation. On the other hand, these results could be due to the peculiarities of the selected industry. At the other extreme, there are studies (Bresnahan et al., 2002; Byrd & Marshall, 1997; Dewan & Min, 1997; Rai & Patnayakuni, 1997) with a broad coverage, which sample manufacturing as well as service sectors. In this respect, their results are most generalizable at the firm level. We observe clearly positive effects of IT investment, with the exception of some ratios (Byrd & Marshall, 1997) that relate IT to economic variables. Model specification appears to be a problem in some studies, in that the variables directly reflect the results, regardless of the existence of causal effects. For example, in the financial firms, not all types of IT are significant in that only transactional IT has an effect on return on assets (Weill, 1992). Likewise in the automotive industry, EDI was important (Mukhopadhyay et al., 1995) to improved inventory turnover. The literature is dominated by large firms, such as US Fortune 500. Yet, this limits generalizability for smaller firms, and IT productivity in general.
3 RESEARCH METHOD

We use replication as the key research method. A number of reviews (Hubbard & Armstrong, 1994; Hubbard & Vetter, 1991; Hubbard, Vetter, & Little, 1998) have concluded that a high proportion of replication studies contradict original findings. This trend holds across a variety of disciplines including finance (Hubbard & Vetter, 1991), economics, strategic management (Hubbard et al., 1998), and marketing (Hubbard & Armstrong, 1994). While low incidence of replications is expected, the surprising result is a lack of support for the original research. It is seen even in more positivistic disciplines as finance and economics where only 20% of replications support the original results. The low replicability generates the so-called ‘file drawer problem’ (Rosenthal, 1979) in which the 5% of studies that are Type I errors (falsely rejecting the null hypothesis) are favored by reviewers, and get disproportionately large space in journals. The majority of papers which do not demonstrate ‘significant’ results are stored in drawers. This has consequences for the reliability, validity, and generalizability of findings (Hubbard et al., 1998). While we do not know of reviews that have studied the extent of replications in IS research, the few which have been published are somewhat controversial, in that the original authors ( Kettinger, Lee, & Lee, 1995) questioned whether their work was ‘replicable’ when replication (Van Dyke, Prybutok, & Kappelman, 1999) found low support for the original model.

Meta-analyses have been seen as more acceptable research methodology vis-à-vis replication (Eden, 2002). In meta-analyses, the researcher uses a sample of statistical findings across studies as raw data. The approach results in a grand survey with few variables, and also reduces sampling error (Lindsay & Ehrenberg, 1993). Replication is conventionally almost the reverse in which a researcher selects a single empirical study, and subjects it to an additional test to increase generalizability. Meta-analytic research (Eden, 2002) is seen to have higher chances in discovering new knowledge. However, the key assumption is that the research findings being used as discrete data points are relatively ‘proven’, ironically true only if the studies are themselves replicable. In this sense, replication is a necessary requirement for high quality meta-analyses.

The main assumption in meta-analyses (Figure 1) is that the constructs being used for testing need to be ‘very’ similar, before entering into a regression. In new disciplines such as Information Systems, there has been high growth in new knowledge, which differs in methods and theoretical rationale. Hence it is unlikely that the constructs will be adequately similar to use in meta-analyses. In this case, replications help collate many studies, and thus develop a cumulative body of knowledge. Deliberate variation in experimental conditions of research (Lindsay & Ehrenberg, 1993) extends the external

![Figure 1. Meta-analyses](image1.png)

![Figure 2. Meta-analytic Replication](image2.png)
validity of the research. The existence of statistically significant results in the original research (Lindsay & Ehrenberg, 1993) mean that the results are unlikely to result from a sampling error. However, we are less sure of the ‘law like’ nature of theories in positivistic research, and whether theories are replicable outside their original scope.

We augment the conventional view of replication by proposing a replication through competing models, based on a review of firm level literature, and a meta-analytic approach to replication (Figure 2). In this approach, we replicate a number of studies simultaneously by using a single dataset as the common denominator. In this respect, we bring alternative models to a common level as is done in meta-analytic research. The approach is similar, but also different in many ways from meta-analyses. Meta analyses have to make certain assumptions about individual studies when abstracting them into discrete ‘data’. These assumptions include the comparability of variables across studies, degree of support, and type of study. In contrast, we take each individual study ‘as is’, abstracting the variables to a replicated/revised model. This abstraction shares greater similarity to the original study, as compared to reducing each model to ‘data’. Our approach can handle a number of constructs, while meta-analyses are limited in that respect, using ‘similar’ constructs (figure 1). We generalize theoretically across constructs and their robustness, whereas meta-analyses generalize across populations. In this respect, our approach is highly applicable to the current state of research in IT value, where a large number of disparate constructs have been deployed in theory building and theory testing. Essentially, we recognize that constructs are different (Figure 2), and hence do not try to ‘force’ similarity across diverse research strands but to discover it by careful analyses. Essentially, our analyses have a greater ‘fidelity’ to the original work, by selection of analogs.

We will use the section discussing the classification of studies to guide our selection. An additional criterion is the availability of data, as our survey did not collect detailed data to match with all the IT related variables used in studies. These difficult to match variables include details on IT budget (Byrd & Marshall, 1997), staffing (Francalanci & Galal, 1998) and different ‘types’ of IT expenses (Weill, 1992). These criteria limited the choice of studies that were replicable, in cross-sectional studies. We selected cross sectional (Black & Lynch, 1997; Prasad & Harker, 1997; Rai & Patnayakuni, 1997) as well as longitudinal (Bresnahan et al., 2002; Brynjolfsson & Hitt, 1993; Hempell, 2002) studies for replication.

3.1 Data Source and Models

Given the research coverage skewed towards larger firms, we deploy a new sample of firms. These include small firms that serve automotive manufacturers and their suppliers, and account for more than 10 percent of U.S. manufacturing employment1. In this replication study, analogs have a key role in ensuring the validity of the results. The controversies on the quality and applicability of replication research have been heavily influenced by the choice of analogs used by researchers. So we identify the analogs along with our results, for each replication. We select studies across different types in research, i.e., vary in design, independent and dependent variables. Starting with relatively simpler research which study only the IT artifact (Brynjolfsson & Hitt, 1993; Prasad & Harker, 1997), we replicate more complex research (Hempell, 2002) which includes complementary factors on employee characteristics and organization structure. The latter would be more difficult to replicate, given the interactions of IT and organizational factors.

In the following section, we show the results of the original study, and those from its replication. The analogs in our dataset are clearly shown, so that their similarity with the variables (or constructs) is known for each replication. Each replication is presented as a set of parallel regressions, with the coefficients, and significance levels shown in the table. We show the dependent variable, independent

---

1 Details of sample are available on http://www.mmtc.org
variables, sample size, and the adjusted R square for each regression. The first two replications do not include any complementary factors. Later replications have different complementary factors, which are matched with analogs in our dataset.

4 PRELIMINARY RESULTS

The purpose of replication is to verify the results that could be observed using existing models, and new dataset. The tables show the original study in the left, and the replication results in the right panel. In general, we have smaller sample sizes, limiting the statistical power.

4.1 Replications focusing on IT capital and labor

Prasad and Harker (1997) based their work on data from ongoing project on retail banks. IT data from a survey was matched with secondary sources. The significance of labor is replicable (Table 2), although we have the ‘overall’ effect. The additional effect of non IT capital is a surprise. The question is whether it is responsible for higher R square? This replication shows the conventional effects of labor and capital. The results suggest the limitations of highly industry specific variables such as loans and deposits, which were unrelated to either type of capital. In the replication (right hand side of Table 2), non IT capital has a significant effect.

<table>
<thead>
<tr>
<th>Prasad and Harker (1997) results</th>
<th>Replication results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original variable</td>
<td>Coefficient (SE)</td>
</tr>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
</tr>
<tr>
<td>Loans and deposits</td>
<td>.259 (.031)**</td>
</tr>
<tr>
<td>Non IT labor</td>
<td>.533 (.059)**</td>
</tr>
<tr>
<td>Non IT capital</td>
<td>-.0207 (.026)</td>
</tr>
<tr>
<td>IT capital</td>
<td>.0011 (.013)</td>
</tr>
<tr>
<td>Adj. R square</td>
<td><strong>0.41</strong></td>
</tr>
<tr>
<td></td>
<td>141</td>
</tr>
</tbody>
</table>

* ** *** denote sig. levels of .1, .05, .01 respectively. Est. equations also include a constant.

Table 2. Variables, analogs, and results

Brynjolfsson & Hitt (1993) used sophisticated deflators to correct the commonly used dependent variable of sales. Their results were key to resolving the productivity paradox, and an elegant formulation of the problem. The results are easily replicated, although with a lower R square value in our smaller sample (Table 3). As we are using cross sectional data, we do not have elaborate controls (year and sector). The results are remarkably robust\(^2\) when value added is used as DV, instead of sales. Notice that the significance levels for the coefficients across the models, on the left and right hand sides of the table, are remarkably similar. In the replication, coefficient for IT capital is much larger than that for conventional (non IT) capital.

These results show that the effects of IT investment are replicable. This suggests the role of IT capital as a key factor of production, regardless of the dependent variable. The study (Brynjolfsson & Hitt, 1993) which was an important landmark in resolving the productivity paradox, appears to use a remarkably robust model.

---

\(^2\) R squares are not strictly comparable, as the DVs are different.

\(^3\) not shown here
Replications focusing on IT, innovation and work organization

Hempell (2002) used longitudinal data collected using a survey, along with econometric variables. Selected firms were classified as innovators, based on perceptual data from respondent surveys. Large sample size, and longitudinal data were used. Significant effects were observed for innovation, and its interaction with IT capital. In our results (Table 4) we have the conventional effects of labor and capital. IT capital is an additional effect clearly seen in the replication. We observe the direct effects of innovation, but not the interaction with IT. It is possible that we have insufficient power to capture the effect. With IT capital excluded from the regression, the interaction term becomes significant. While this hints at the original results, the statistical interpretation of the interaction term is difficult when the original variable is excluded in the regression.

Bresnahan et al (2002) studied large US firms, from manufacturing (55%) and services (45%) sectors. HR managers responded on the firm and its “most typical” establishment. Work organization was assessed through manager ratings on the use of teams and decentralized decision making. The study combined surveys and secondary data in new ways. Elaborate controls were used in longitudinal data to tease out the effects of work organization. In the replication (table below), we have similar results for labor and capital. However, there is no evidence that use of work teams adds to the explanation over and above the conventional inputs. The significance levels are not close to 10% levels seen in the original study. The effects of ‘work organization’ (a composite measure) are clearly difficult to replicate. There are clearly additional analogs which add variance over and above the extent of teamwork, such as autonomy (Bresnahan et al, 2002).
Table 5. Variables, analogs, and results

<table>
<thead>
<tr>
<th>Bresnahan et al (2002) results</th>
<th>Replication results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
</tr>
<tr>
<td>Value added</td>
<td>Value added</td>
</tr>
<tr>
<td><strong>Independent variables</strong></td>
<td></td>
</tr>
<tr>
<td>labor</td>
<td>.748*** (.0374)</td>
</tr>
<tr>
<td>non IT capital</td>
<td>.355*** (.0246)</td>
</tr>
<tr>
<td>IT stock</td>
<td>.0358** (.0147)</td>
</tr>
<tr>
<td>Work organization</td>
<td>.0218* (.0129)</td>
</tr>
<tr>
<td>IT stock*work organization</td>
<td>.0162* (.0084)</td>
</tr>
<tr>
<td>Adj. R square</td>
<td>0.93</td>
</tr>
<tr>
<td>Sample size</td>
<td>2228</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>125</td>
</tr>
</tbody>
</table>

Table 5. Variables, analogs, and results

5 DISCUSSION

Replications suggest that IT capital measures are comparatively robust, and findings can be replicated even with smaller sample sizes we have used. IT capital is a well understood economic variable, and is comparable across different models. Value added, or sales, are robust DVs that help replication and comparability. Complementary capabilities, however, are difficult to replicate. These have different measures in different contexts, and are not easy to compare across models. Organizational capabilities, specifically, have been operationalized as highly context-dependent in literature. In replication, it is difficult to select analogs for these specific contexts. Findings of studies which use extremely fine levels of analysis, viz. machining production “jobs” at manufacturing plants (Kelley, 1994) are difficult to replicate at aggregated levels. The selection of these micro level units of analysis affects their replicability in later research. Given the difficulties, there is a clear relationship between use of well understood econometric variables, and generalizability of research. The lack of statistical significance among the ‘organizational’ analogs is therefore, not so surprising. It is more difficult to guess whether low replicability is due to data or use of unrelated constructs (analogs). Using additional analogs that are more similar to those used in the original study will provide some hints. The results indicate that there are higher impacts from use, as compared to work practices. The results show differing levels of robustness of the constructs, used in IT value research.

A key limitation of the data is that IT investment is aggregated, i.e., we do not have the details of labor and capital. Hence studies which used such detailed components of IT investment were not replicable, or replicable to a limited extent. The current dataset might have an impact on these results. The biggest threat to validity is the similarity of results across a variety of large firms and a specific sample.

We plan to collect additional data in the following surveys to replicate additional models, through capturing detailed analogs. This will be done in line with the annual surveys in the next few months. The large number of models will be later integrated into a framework, as we draw implications from the results.

References


