MIS Quarterly

SPECIAL ISSUE: IT AND INNOVATION

EXTERNAL KNOWLEDGE AND INFORMATION TECHNOLOGY: IMPLICATIONS FOR PROCESS INNOVATION PERFORMANCE

Konstantinos Trantopoulos
IMD, Chemin de Bellerive 23, 1001 Lausanne, SWITZERLAND, and
ETH Zurich, Weinbergstrasse 56/58, 8092 Zurich, SWITZERLAND {kontranto@gmail.com}

Georg von Krogh
ETH Zurich, Weinbergstrasse 56/58, 8092 Zurich, SWITZERLAND {gvkrogh@ethz.ch}

Martin W. Wallin
Chalmers University of Technology, SE-412 96 Gothenburg, SWEDEN {martin.wallin@chalmers.se}

Martin Woerter
ETH Zurich, Leonhardstrasse 21, 8092 Zurich, SWITZERLAND {woerter@kof.ethz.ch}

Appendix A

Sample Characteristics

Table A1. Industry Composition of the Sample

	Manufacturing Industry	2005		2008		2011	
		Freq.	%	Freq.	%	Freq.	%
1	Food/Beverages/Tobacco	107	8.48	83	7.74	95	8.22
2	Textile	31	2.46	21	1.96	23	1.99
3	Clothing	10	0.79	8	0.75	10	0.87
4	Wood	41	3.25	36	3.36	41	3.55
5	Paper & Paper Products	29	2.30	21	1.96	22	1.90
6	Printing	70	5.55	58	5.41	56	4.84
7	Chemicals & Chemical Products	95	7.53	87	8.12	102	8.82
8	Rubber/Plastics	48	3.80	38	3.54	54	4.67
9	Non-metallic Minerals	44	3.49	36	3.36	35	3.03
10	Manufacture of Basic Metals	30	2.38	26	2.43	28	2.42
11	Fabricated Metals	152	12.04	143	13.34	164	14.19
12	Machinery & Equipment	230	18.23	190	17.72	189	16.35
13	Electrical Equipment	70	5.55	60	5.60	54	4.67
14	Electronics	142	11.25	129	12.03	129	11.16
15	Watches	46	3.65	42	3.92	44	3.81
16	Vehicles	27	2.14	22	2.05	22	1.90
17	Other Manufacturing	39	3.09	28	2.61	32	2.77
18	Energy	51	4.04	44	4.10	56	4.84
Total		**1,262**	**100.00**	**1,072**	**100.00**	**1,156**	**100.00**

Note: Industry distribution for the Swiss manufacturing firms in the sample (two-digit NACE codes).

Appendix B

Variables and Descriptive Statistics

Table B1. Utilization of External Knowledge Sources

Sources		Mean Score
1	Customers	3.24
2	Suppliers (materials)	3.22
3	Suppliers (investment goods)	2.59
4	Suppliers (software)	2.36
5	Competitors	2.78
6	Own Enterprise Group	2.12
7	Universities	2.48
8	Public or Private Research Institutes	2.19
9	Consultants	2.02
10	Technology Transfer Offices	1.81
11	Patent Disclosures	2.01
12	Fairs/Exhibitions	3.12
13	Conferences/Scientific Literature	2.99
14	Databases (information networks)	2.48

Note: Mean values on survey responses (five-point Likert scale, 1 = no usage; 5 = high usage) over the nine-year sample period (N = 3,490).

Table B2. Variable Definitions

Variable Name	Variable Construction/Definition
Cost Reduction	Continuous variable: Cost reduction achieved by process innovation (log).
External Search Depth	Number of external knowledge sources utilized intensively: the questionnaire contains 14 external knowledge sources for innovation activities. The variable counts the knowledge sources which has been assessed as very important (value 4 or 5 on a 5-point Likert scale).
External Search Breadth	Number of external knowledge sources used: the variable counts the number of external knowledge sources (maximum 14) that are of some importance to the innovation activities of the focal firm (value 2, 3, 4, or 5 on the 5-point Likert scale).
Data Access Systems	Sum of three binary variables: Adoption of systems for enterprise resource planning (ERP) (0/1), supply chain management (SCM) (0/1), and customer relationship management (CRM) (0/1).
Network Connectivity	Binary variable: Adoption of local area network (LAN) (0/1).
IT Investments	Continuous variable: Investments in IT (log).
R&D Intensity	Fractional variable: R&D expenditures to total sales.
% Employees Academic Degrees	Fractional variable: Percentage of employees with academic degrees.
Product Innovation Objective	Binary variable: value 1 if at least one out of four product innovation goals (improve product quality, replace outdated products, expand product portfolio, keep or increase market share) is assessed as very important by the focal firm (value 5 on a 5-point Likert scale); 0 otherwise.
Process Innovation Objective	Binary variable: value 1 if at least one out of five process innovation goals (increase flexibility of production, reduce labor costs, reduce material cost, reduce energy cost) is assessed as very important by the focal firm (value 5 on a 5-point Likert scale); 0 otherwise.
Outsourcing Production (intermediate)	Binary variable: value 1 if the focal firm outsourced the production of intermediate products; 0 otherwise.
Outsourcing Production (all)	Binary variable: value 1 if the focal firm outsourced the whole production process; 0 otherwise.
Mergers	Binary variable: value 1 if the focal firm merged with other firms; 0 otherwise.
Acquisitions	Binary variable: value 1 if the focal firm acquired other firms/parts of other firms; 0 otherwise.

Table B2. Variable Definitions (Continued)

Variable Name	Variable Construction/Definition
Firm Concentration on Core Business	Binary variable: value 1 if the focal firm made steps towards a stronger concentration on its core business; 0 otherwise.
% Employees Further Education	Fractional variable: share of employees that received further education.
% Costs of Further Education	Fractional variable: share of costs for further education covered by the focal firm.
Employees' Change of Responsibilities	Binary variable: value 1 if there has been a change in the responsibilities of employees; 0 otherwise.
% Employees Switch Function and/or Department	Fractional variable: share of employees that switched function and/or department.
Size	Continuous variable: number of employees (log).
Industry	18 binary variables: industry dummies defined at the NACE two-digit level.
Time	3 binary variables: time dummies to capture the three waves of the survey.

Data source: KOF (Swiss Economic Institute), ETH Zurich.

Table B3. Descriptive Statistics and Pair-Wise Correlation Matrix of Major Variables

	Variable	Mean	S.D.	1	2	3	4	5	6	7	8	9	10
1	*Cost Reduction*	1.92	7.36	1									
2	*External Search Depth*	3.29	2.65	0.21	1								
3	*Data Access Systems*	1.18	1.01	0.12	0.18	1							
4	*Network Connectivity*	0.85	0.36	0.18	0.23	0.37	1						
5	*IT Investments*	10.65	3.05	0.21	0.18	0.20	0.33	1					
6	*External Search Breadth*	10.36	3.28	0.17	0.45	0.18	0.32	0.24	1				
7	*R&D Intensity*	0.02	0.04	0.21	0.15	0.08	0.14	0.11	0.12	1			
8	*Product Innovation Objective*	0.31	0.46	0.28	0.27	0.11	0.17	0.17	0.16	0.24	1		
9	*Process Innovation Objective*	0.15	0.36	0.38	0.27	0.10	0.13	0.16	0.14	0.13	0.40	1	
10	*Size*	4.26	1.41	0.28	0.22	0.32	0.49	0.62	0.31	0.10	0.21	0.18	1

Note: Pair-wise Pearson correlations are reported based on the IV GMM estimations (N = 1,057). All correlations are significant at $p < .01$.

Appendix C

Endogeneity Tests

Test of Underidentification

A rejection of the null indicates that the selected set of excluded instruments is correlated with the endogenous variables. Since we present cluster-robust statistics, this test refers to the Kleibergen-Papp test, which is a generalization of the Anderson canonical correlation rank statistic in the non-i.i.d case (Kleibergen and Paap 2006). This test results in an LM statistic of 20.10 (significant at the 1% level) for Model 1, 30.45 (significant at the 1% level) for Model 2, 9.989 (significant at the 1% level) for Model 3, and 9.367 (significant at the 1% level) for Model 4 (see Table 1 in the main results). Therefore, all instruments pass the underidentification test (i.e., they are sufficiently correlated with the endogenous variables).

Test of Overidentifying Restrictions

Since the number of excluded instruments exceeds the number of our endogenous regressors (overidentification), we can test for the instrumental exclusion restriction, that is, test the joint null hypothesis that the group of instrumental variables is valid (i.e., uncorrelated with the error terms) (Baum et al. 2007). A rejection of the null would cast suspicion on the validity of the instruments. As we present

heteroskedastic-robust covariance estimator, the Hansen's J statistic is reported. The Hansen J statistic reports values of 1.610 for Model 1, 1.585 for Model 2, 0.375 for Model 3, and 0.535 for Model 4 (see Table 1 in the main results). These values indicate that the null hypothesis is not rejected; thus, our instruments are valid.

Endogeneity Test

Finally, we test the null hypothesis that the specified endogenous regressors can actually be treated as exogenous (Baum et al. 2007). The Wald test of exogeneity reports the values 7.196 (significant at the 5% level) for Model 1, 12.58 (significant at the 5% level) for Model 2, 14.27 (significant at the 5% level) for Model 3, and 17.83 (significant at the 1% level) for Model 4 (see Table 1 in the main results). Thus, we confirm that the specified endogenous regressors cannot be treated as exogenous.[1]

References

Baum C. F., Schaffer, M. E., and Stillman, S. 2007. "Enhanced Routines for Instrumental Variables/GMM Estimation and Testing," *The Stata Journal* (7:4), pp. 465-506.

Kleibergen, F., and Paap, R. 2006. "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics* (133:1), pp. 97-126.

Appendix D

Robustness Tests

Alternative IT Proxies

Table D1 examines the introduction of a binary variable for the presence of *ERP* in the focal firm as an alternative proxy for *Data Access Systems* (for a similar approach, see Bloom et al. 2015). We keep the same model specification as in our main regressions. The interaction term *Data Access Systems (ERP) * External Search Depth* in the first column of Table D1 takes a positive and significant coefficient (coeffficient = 1.535, $p < .05$). This is in line with the finding of our main model depicted in column (2) of Table 1. Next, we introduce the *Network Connectivity* variable. The second column of Table D1 qualitatively reproduces the main results depicted in column (4) of Table 1. However, the focus on ERP as a measure of *Data Access Systems* causes the interaction term *Network Connectivity * External Search Depth* to lose its significance ($p = .184$). This ambiguity of the *Network Connectivity* interaction term may well be rooted in the fact that ERP—although a system that primarily affects access to information—may also to some modest degree be used for internal communication (for a similar disussion, see Bloom et al. 2015, p. 2878). Thus, ERP might cover some of the *Network Connectivity* (LAN) effect, causing the *Network Connectivity * External Search Depth* term to become insignificant. To examine this issue in more depth, we reran our regressions by substituting *ERP*, first, with a binary for *SCM* and second, with a binary for *CRM*. In contrast to ERP, SCM and CRM are enterprise systems with no internal communication capabilities, thus we should expect those systems to have less or no overlap with the *Network Connectivity* effect. Indeed, the results of these models reproduce our main findings in column (4) of Table 1.[2]

[1]We implemented the endogeneity tests using the "endogtest" option for the "ivreg2" command in Stata 13. Notice that the test on exogeneity is performed after the overidentification restrictions test, as the first is not valid if the latter rejects the validity of the instruments.

[2]In particular, the coefficients (standard errors) for *Network Connectivity * External Research Depth* become 3.506 (1.603) and 2.967 (1.331) for the models including *SCM* and *CRM*, respectively. Full results available upon request.

Table D1. Robustness Tests, IV GMM Regressions, 2003–2011

Regressors	Dependent Variable (*Cost Reduction*)	
	Alternative *Data Access Systems* Measure (ERP)	Alternative *Data Access Systems* Measure (ERP)
	Coefficient (Std. Err.)	Coefficient (Std. Err.)
External Search Depth	-0.453 (0.657)	-1.838* (1.046)
Data Access Systems (ERP)	-7.379*** (2.606)	-5.753 (3.549)
*Data Access Systems (ERP) * External Search Depth*	1.535* (0.760)	0.846 (1.052)
Network Connectivity		-3.632 (4.253)
*Network Connectivity * External Search Depth*		2.251 (1.696)
IT Investments	0.732** (0.317)	0.787** (0.348)
External Search Breadth	-0.044 (0.128)	-0.087 (0.139)
Product Innovation Objective	0.708 (0.661)	0.542 (0.706)
Process Innovation Objective	4.600*** (0.940)	4.298*** (1.022)

Notes: Values are unstandardized regression coefficients. Huber-White robust standard errors are clustered at the firm level (i.e., robust to arbitrary heteroskedasticity and autocorrelation). All columns include the full set of control variables as in the main regressions (see Table 1). *$p < .10$; **$p < .05$; ***$p < .01$.

Tobit Estimator

The continuous dependent variable *Cost Reduction* includes a set of zero values since some firms did not introduce cost-reducing process innovations during the observed period. Consequently, the distribution of the dependent variable shows a pileup at the value zero (corner solution). Hence we conducted robustness tests by using an instrumental-variables Tobit (nonlinear) estimator (see Wooldridge 2002). Table D2 presents the results of the IV Tobit estimation and compares them with the ones obtained from our main regressions (IV GMM; see also Table 1).

While in linear models the interpretation of the coefficients of interaction terms is straightforward, this does not extend to nonlinear models like the Tobit. Unfortunately, inference and model testing of interaction effects cannot be conducted simply via the magnitude, statistical significance, or sign of the coefficients of the interaction terms. For example, the sign of the coefficient does not necessarily indicate the sign of the interaction effect (Ai and Norton 2003). Therefore, Tobit coefficient estimates for interaction terms should not be directly compared with the ones of linear regressions. As a consequence, we compare IV GMM with IV Tobit estimates, present Tobit estimates only for Model 1, and will not attempt to utilize Tobit estimates as robustness tests for the models with interaction effects (Models 2–4 in Table 1).

The comparison of the results between IV GMM and IV Tobit (Table D2) confirms the robustness of our main estimations shown in Table 1. In the IV Tobit estimation, all coefficients display the same sign and similar statistical significance as the corresponding GMM estimates. In particular, note that our main independent variable *External Search Depth* remains positive and even gains in significance (the magnitudes of the coefficients of the Tobit and GMM estimations are not directly comparable).

Table D2. Comparison Between IV GMM and IV Tobit Regressions, 2003-2011		
	Dependent Variable (*Cost Reduction*)	
	IV GMM	IV Tobit
Regressors	Coefficient (Std. Err.)	Coefficient (Std. Err.)
External Search Depth	0.660* (0.335)	3.383** (1.686)
IT Investments	0.561** (0.281)	5.285** (2.188)
External Search Breadth	-0.161 (0.116)	-0.995 (0.620)
Product Innovation Objective	0.710 (0.639)	2.507 (2.740)
Process Innovation Objective	4.858*** (0.913)	15.67*** (3.630)
% Employees Cont. Education	0.019** (0.009)	0.086* (0.046)

Notes: All columns include the full set of control variables as in the main regressions (see Table 1).
*p < .10; **p < .05; ***p < .01.

References

Ai, C., and Norton, E. C. 2003. "Interaction Terms in Logit and Probit Models," *Economics Letters 18* (80:1), pp. 123-129.

Bloom, N., Garicano, L., Sadun, R., Van Reenen, J. 2014. "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science* (60:12), pp. 2859-2885.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

Appendix E

Interaction Effect Between External Search Depth and IT Investments

The information systems literature makes wide use of IT investments in order to study how the IT artifact affects various types of performance (e.g. Bardhan et al. 2013; Kleis et al. 2012; Mithas et al. 2012; for a review, see Melville et al. 2004). Thus, we also run a supplementary analysis with the interaction effect between *External Search Depth* and *IT Investments* in a model specification similar to the one in our main regressions. Analogous models including interaction effects have appeared in recent work on IT and innovation (Bardhan et al. 2013; Joshi et al. 2010; Kleis et al. 2012; Tambe et at. 2012).

Table E1 presents the results. Column 1 shows the baseline regression, which is also presented in column (1) of Table 1. The coefficients of *External Search Depth* and *IT Investments* are positive and significant. In column (2) we introduce the interaction between *External Search Depth* and *IT Investments.* The coefficient of *External Search Depth* * *IT Investments* is positive and significant (coefficient = 0.381, p < .01). Importantly, this result suggests that intense usage of multiple external sources interacts with IT investments to positively affect process innovation performance. However, while the coefficient estimates for the model with the interaction effect are informative, they have limited explanatory value with regard to the marginal effect of *External Search Depth* on process innovation performance. For example, the coefficient estimate of *External Search Depth* conveys little information about the marginal effect of *External Search Depth* on *Cost Reduction* conditional on the value of *IT Investments*. Notice that in this model, the marginal effect of *External Search Depth* is dependent on *IT Investments*, and therefore the coefficients of the constitutive term of *External Search Depth* should not be interpreted as the average effect of a change in the independent variable on the dependent variable. (This coefficient only captures this effect correctly when *IT Investments* is zero.) Moreover, the coefficient estimates fail to indicate whether *IT Investments* has a statistically significant impact on the aforementioned marginal effect of the *External Search Depth* variables within the sample range of observed *IT Investments* values.

Table E1. Interaction Effect (IT Investments)

Regressors	Dependent Variable (*Cost Reduction*)	
	(1)	(2)
	Coefficient (Std. Err.)	Coefficient (Std. Err.)
External Search Depth	0.660* (0.335)	-3.726** (1.520)
IT Investments	0.561** (0.281)	-0.032 (0.388)
*External Search Depth * IT Investments*		0.381*** (0.128)
External Search Breadth	-0.161 (0.116)	-0.034 (0.136)
Product Innovation Objective	0.710 (0.639)	0.915 (0.700)
Process Innovation Objective	4.858*** (0.913)	4.091*** (1.074)
% Employees Cont. Education	0.019** (0.009)	0.007 0.011)

Notes: All columns include the full set of control variables as in the main regressions (see Table 3).
*p < .10; **p < .05; ***p < .01.

In order to provide a substantively meaningful description of the marginal effect of *External Search Depth* while accounting for the interaction effect, we go beyond the traditional results table and graphically illustrate the marginal effect of the *External Search Depth* variable on process innovation performance conditional on *IT Investments*, along with the corresponding standard errors (Brambor et al. 2006). The graphical results correspond to the instrumental-variables GMM estimator (column (2) of Table E1) with all other covariates being set to their mean values. Figure E1 illustrates the marginal effect of *External Search Depth* on *Cost Reduction*. The horizontal axis of the plot extends from the minimum observed value of *IT Investments* in the sample (0) to the maximum (19.1). The solid line in the figure indicates how the marginal effect of *External Search Depth* changes as IT investments increase. The significance of the marginal effect is depicted by the 95% confidence intervals around the sloping line: the marginal effect is significant when the upper and lower bounds of the confidence interval are either above or below the horizontal zero line (Brambor et al. 2006). We also overlay the frequency distribution of the *IT Investments* variable in the sample over the marginal effect plot. Figure E1 depicts that the marginal effect of *External Search Depth* is negative and significant when *IT Investments* is zero. This value of the marginal effect corresponds to the coefficient estimate of *External Search Depth* in column (2). As *IT Investments* increases, the marginal effect of *External Search Depth* increases (ascending line), a consequence of the positive coefficient of the interaction effect *External Search Depth * IT Investments*. The marginal effect becomes positive but still insignificant for *IT Investments* = 9.5, and converts to significant for *IT Investments* > 11.5.[3] This makes clear that *External Search Depth* has a positive impact on process innovation performance when the IT investments of the firm are moderate to high. Overall, the baseline model predicts a positive and significant relationship between *External Search Depth* and *Cost Reduction*, while the extended model predicts a positive and significant relationship for moderate to high IT investments.

To summarize, Figure E1 sheds light on the impact of the intensive usage of multiple external sources on process innovation performance, for different levels of IT investments. It also offers insights into the competitive significance of IT investments, while painting a nuanced and comprehensive picture of their strategic implications for process innovation performance. The graph illustrates the conditions under which the intense use of external knowledge sources is strategically beneficial to the firm, and shows that this use needs to be carefully orchestrated with a firm's IT investments for it to pay off in terms of production-related cost reductions.

[3]In our sample, 38.7% of the observations with *IT Investments* > 11; see histogram in Figure E1.

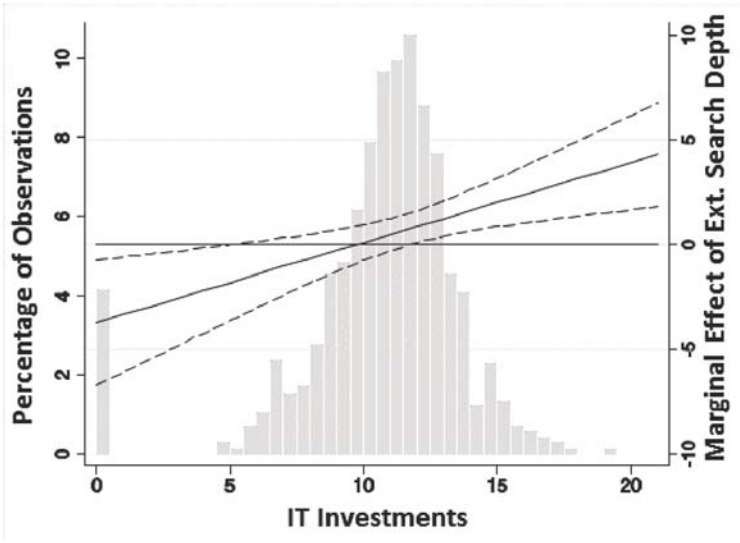

Figure E1. Marginal Effect of *External Search Depth* on *Cost Reduction*, Contingent on *IT Investments* (IV GMM, Column (2) of Table E1)

References

Bardhan, I., Krishnan, V., and Lin, S. 2013. "Business Value of Information Technology: Testing the Interaction Effect of IT and R&D on Tobin's Q," *Information Systems Researc*h (24:4), pp. 1147-1161.

Brambor, T., Clark, W. R., and Golder, M. 2006. "Understanding Interaction Models: Improving Empirical Analyses," *Political Analysis* (14), pp. 63-82.

Joshi, K. D., Chi, L., Datta, A., and Han, S. 2010. "Changing the Competitive Landscape: Continuous Innovation Through IT-Enabled Knowledge Capabilities," *Information Systems Research* (21:3), pp. 472-495.

Kleis, L., Chwelos, P., Ramirez, R. V., and Cockburn, I. 2012. "Information Technology and Intangible Output: The Impact of IT Investment on Innovation Productivity," *Information Systems Research* (23:1), pp. 42-59.

Melville, N., Kraemer, K., and Gurbaxani, V. 2004. "*Review*: Information Technology and Organizational Performance: An Integrative Model of IT Business Value," *MIS Quarterly* (28:2), pp. 283-322.

Mithas, S., Tafti, A., Bardhan, I., and Goh, J. M. 2012. "Information Technology and Firm Profitability: Mechanisms and Empirical Evidence," *MIS Quarterly* (36:1), pp. 205-224.

Tambe, P., Hitt, L. M., and Brynjolfsson, E. 2012. "The Extroverted Firm: How External Information Practices Affect Innovation and Productivity," *Management Science* (58:5), pp. 843-859.