

Analyzing Influences on Pivotal ITO Contract Features: A Quantitative Multi-Study Design with Evidence from Western Europe

Completed Research

Michael Könnig

TU Dresden

michael.koenning@mailbox.tu-dresden.de

Kai Heinrich

TU Dresden

kai.heinrich@tu-dresden.de

Patrick Zschech

TU Dresden

patrick.zschech@tu-dresden.de

Christian Leyh

TU Dresden

christian.leyh@tu-dresden.de

Abstract

As information technology in private and public organizations continues to gain importance, IT outsourcing (ITO) has become a critical component of corporate strategy for many institutions. Consequently, substantial research has investigated topics around ITO decisions, outcomes, and contractual governance. Despite decades of academic research, however, quantitative analyses of key contract features are still scarce. By applying three statistical lenses on a dataset of more than 1,000 ITO deals from the ASG region, we shed light on three pivotal ITO contract characteristics. Our findings are twofold: Contextually, they point to declining contract lengths and run rates, identify significant influence factors on contract value, contract length, and pricing method, and show high prediction rates in binary classification of pricing method and contract value. Methodically, they serve as a proof of concept for IS theory building by applying quantitative mixed-method approaches to similar datasets with global scope and additional variables.

Keywords

IT Outsourcing; ITO; Sourcing Strategy; Explanatory Analysis; Predictive Modeling.

Introduction

The sourcing of information technology (IT) has developed into a critical component of corporate strategy. Several studies (e.g., Dibbern et al. 2004; Gonzalez et al. 2006; Lacity et al. 2016; Liang et al. 2015) provide an overview of the potential value that arises from outsourcing corporate IT services to an external party (IT outsourcing, ITO). Today, almost every Fortune 500 company and many large public institutions outsource significant parts of their IT services. Consequently, a considerable body of research and an entire global industry worth \$320 billion in 2015 have evolved around ITO (Faisal and Raza 2016).

ITO can take on many different forms along several dimensions. Whereas some organizations decide to almost completely replace their internal IT function with external service providers (full outsourcing), some choose to keep a significant part of competences in-house and only supplement them with external services (partial outsourcing). Moreover, while some organizations rely on a single main service provider (single-vendor sourcing), a growing share of companies decide to engage several main provider firms (multi-vendor sourcing). Finally, some firms outsource their IT services to an organization that operates from an offshore location (offshoring), while others prefer tasks to be fulfilled at the company location (onshoring) or at least from a nearby region (nearshoring) (Ruivo et al. 2015).

Strategic IT sourcing can be thought of as a six-step process, composed of data collection/spend analysis, market research, the go-to-market, negotiations, contracting, and implementation and continuous improvement (Payne and Dorn 2012). All decisions made during the sourcing process eventually manifest

in ITO contracts that govern the outsourcing relationship. These contracts exhibit various characteristics, including contract value, pricing method, contract length, or the annual run rate. Analyzing these deals as a lens on ITO can yield interesting insights concerning the IT sourcing strategy decisions that companies make, the client/provider relationship, and even the ITO market and its development in general. Against this background, in this paper we examine key characteristics of ITO contracts as observable manifestations of IT sourcing decisions. To this end, we analyze a rich dataset of characteristics of 1,016 ITO contracts from Austria, Switzerland, and Germany (the ASG region) that were signed between January 2006 and August 2017. This dataset is a subset of IDC's "BuyerPulse Deals Database", which in total contains data on more than 32,000 worldwide ITO and Business Process Outsourcing (BPO) deals. For each of the analyzed contracts, our dataset provides a multitude of characteristics. It includes information on the two contract partners (e.g., name, industry, revenue, number of employees), as well as on pivotal contract details (e.g., tasks to be provided, contract value, run rate, deal closure date, contract length, price method).

Considering the peculiarities of the given data in terms of numerous entries and high dimensionality, we want to examine to what extent such a dataset can be used towards the purpose of theory development in information systems (IS) research (see Gregor 2006). Therefore, the research in this paper has two goals: First, we seek to uncover latent or rarely-documented relationships between the ITO variables in our dataset. Second, in doing so, we also aim to evaluate the applicability of our empirical methods/models (see Shmueli and Koppius 2011) to such a dataset, in order to get a fundamental understanding of whether it can be used for future research projects (as "proof of concept"), i.e., to evaluate similar or even larger datasets. Therefore, the remainder of this paper is structured as follows: In the next section we provide an overview of the research context, before outlining our research method in detail and describing key characteristics of our dataset. We then present our findings from a descriptive-interpretive analysis, an explanatory design, and from predictive modeling. Finally, we conclude the paper by summarizing our main contributions, discussing limitations, and deriving an outline of future research opportunities.

Related Work

With its increasing importance in corporate practice, ITO has attracted substantial academic interest. During the last few decades, researchers have dealt with a multitude of research questions around ITO (for a full review of ITO literature see Dibbern et al. 2004; Gonzalez et al. 2006; Lacity et al. 2016; Liang et al. 2015). According to Lacity et al. (2016), the majority of these studies can be grouped into one of the following three categories (see their extensive literature review for specific papers in each category): First, a plethora of articles focuses on the most important ITO decisions, i.e., whether to outsource or why or which tasks to outsource. For example, cost savings have consistently been found to be the major motivator for ITO decisions. In addition, other reasons frequently cited for outsourcing include desires to increase business performance, increase flexibility, or access a provider's (technological) expertise or skills, as well as focus on core capabilities. When deciding the outsourcing partner with which to contract, clients are found to favor high levels of domain understanding and cultural proximity, among others. A second prominent category is composed of ITO outcomes, particularly the client's general perception of the success of an ITO endeavor and the effect of ITO on a client's business performance. Researchers found that especially high levels relationship quality with regard to knowledge sharing, communication, trust, and commitment consistently produced better ITO outcomes (Westner and Strahringer 2010). Moreover, higher levels of client expertise with outsourcing, the client's provider management capabilities, and technical capability of the service provider also yield higher levels of satisfaction. The third category has recently become "more prevalent" (Lacity et al. 2016) and is comprised of studies on contractual and relational governance as dependent variables, i.e., what influences contract design, pricing modes, contract duration, and different modes of collaboration, among others. But while a myriad of studies examines ITO decisions and outcomes, comparatively less research focuses on questions of pivotal contractual details such as pricing method, contract duration, or contract volume.

Qualitative case studies or survey-based evaluations provide insights into the ITO decision-making process and the contractual outcomes of individual companies. In terms of pricing method, for instance, these studies discuss characteristics and application scenarios for the two archetypal pricing mechanisms, fixed price and variable or mixed pricing (time-and-material contracts). Other studies examine costs and benefits of different contract durations (Goo et al. 2007; Ravindran et al. 2015b; Susarla 2012). However, these studies hardly provide a holistic view on the current ITO market situation or its development, as they lack

the required sample size for generalizability. Since ITO contract details are seldom published in sufficient quantity to justify general statements, quantitative studies around these questions are also very rare. Consequently, Ravindran et al. (2015b) determine that there is “comparatively limited guidance in the literature on how duration is awarded in contracts for [ITO]” (p. 380). The same argument holds for larger-size analyses of contract volume, i.e., the size of the contract measured in monetary terms (Gewald and Gellrich 2007) or the annual run rates, i.e., the annual cost.

One of the few quantitative publications on the topic is the study by Ravindran et al. (2015b). Drawing from a large IDC dataset of ITO deals, the researchers found that a previous client/provider relationship, the client’s prestige, longer duration of its other contracts, and the provider’s reputation positively affected contract duration. In a similar study, the authors concluded that “greater vendor experience in terms of breadth of industries is associated with a longer contract duration. Similarly, customers that have traded with a greater number of vendors are likely to enter into longer-term contracts” (Ravindran et al. 2015a). However, to the best of our knowledge, no other recent academic research articles describe the development of – or provide statistically significant determinants for – contract duration, contract size, and pricing method. Utilizing a large dataset of ITO deals and applying three different statistical lenses as described below, we seek to shed light on these three pivotal characteristics of ITO contracts.

Research Method

In terms of theory development in IS research, according to Gregor (2006) there are four primary goals of theory: analysis and description, explanation, prediction, and prescription. Within this paper, we focus on the first three goals. This results in the following three research questions (RQ):

- RQ1:** *Which characteristics does the recent development of the ITO market in the ASG region show in terms of contract value (CV), contract length (CL), and pricing method (PM)?*
- RQ2:** *Which explanatory variables have a statistically significant influence on those pivotal contract characteristics?*
- RQ3:** *To what extent can the high dimensional feature space be used to build empirical models with high predictive power?*

To answer our research questions, we applied a mix of analytical approaches to consider different data analysis tasks in relation to theory-driven research in IS (see Gregor 2006). As such, we analyzed our dataset through three lenses. We employed three different statistical approaches to analyze the dataset. First, we used a descriptive-interpretive approach to quantitatively analyze key characteristics of ITO contracts as aspects of IT sourcing strategy and their development over time (RQ1). Second, we designed an explanatory model to determine the main factors influencing those key characteristics (RQ2). Finally, we created predictive models designed to predict the outcome of upcoming ITO deals based on previous sourcing decisions in comparable settings (RQ3). By enhancing the descriptive-interpretive analysis with explanatory and predictive modeling approaches, we tried to determine whether such a data collection is suitable for theory development, in terms of testing causal hypotheses and creating empirical models with predictive power. For the distinction of the task within the explanatory and predictive methods, we followed the approach of Shmueli and Koppius (2011), which highlights the particularities of the two modeling approaches to the process of model development and respective evaluation criteria. Detailed explanations for the applied empirical methods are provided directly in combination with the respective results later on.

Dataset Characteristics

Our dataset was kindly provided by the International Data Corporation (IDC), a “global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets” (International Data Corporation 2018). The organization employs a variety of primary and secondary sources to compile its data. It constantly monitors press releases, public financial records, historic market data, and third-party media reports and complements these data points with interviews with IT service providers to yield more contextual information on specific deals.

We examined details of 1,016 observed ITO deals that were closed between Austrian, Swiss, or German client firms, with both national and international service providers, between January 2006 and August 2017. We triangulated the data by manually searching the web for publicly available information on ITO

deals, predominantly on IT news websites and in press releases. While IDC's dataset misses some of the smaller deals, we found that the dataset contains almost every large known deal of at least \$100 million. Moreover, the dataset also includes information on commercial terms raised from interviews and industry analyses - data that is hardly made public. Hence, we chose not to enrich the IDC dataset with our own findings and instead kept with IDC's database for the sake of consistency. The IDC dataset describes its 1,016 ITO deals along 74 variables. Ten of these variables represent IDC's internal IDs, keywords, or similar information which we did not consider relevant and therefore excluded from our analyses. In the following, we provide a brief overview of the nature of our dataset by describing it along: four client characteristics (see section a), the number of service providers (b), and three pivotal contract details (c). The full list of variables can be found in the Appendix of this paper in Table 6.

(a) In terms of client characteristics, we describe its country, industry sector, number of employees, and revenue. Of the 1,016 observed ITO contracts, 703 (69%) were closed with German client companies, 250 (25%) with Swiss clients, and 63 (6%) with Austrian companies. These clients are active in a multitude of different industry sectors, including discrete manufacturing (15% of deals), professional services (13%), public sector (11%), banking (11%), insurance (10%), process manufacturing (9%), transportation (5%), retail trade (4%), and communications and media (4%). The remaining 173 deals (18%) were closed with companies from eight other sectors. The dataset is similarly diverse regarding client size: 329 contracts (32%) were closed with clients with less than 1,000 employees, 313 contracts (31%) with clients that had 1,000-9,999 employees, and 306 contracts (30%) with clients employing at least 10,000 staff. The size of clients closing the remaining 68 deals (7%) was unknown. Finally, a similarly heterogeneous picture is presented in terms of client revenue: 425 deals (42%) were closed with clients that have revenues of up to \$1 billion, 245 deals (24%) with clients between \$1 billion and \$10 billion, and 229 deals (23%) with clients grossing \$10 billion and more. The revenue of 117 client companies (11%) was unknown.

(b) In terms of service providers, the dataset includes deals with 211 different service providers, from large multi-billion-dollar companies, such as Atos and Deutsche Telekom, to small and medium-sized firms with only one or a few deals totaling less than EUR 10 million of contract value.

(c) In terms of contract details, we focus on ITO engagement type, contract type, and bid type. The IDC database discerns three types of ITO engagement types. Application Outsourcing (AO) includes services designed to provide for the day-to-day operations, support, development, customization, implementation, and integration, as well as maintenance of enterprise applications and end-user support. IT Outsourcing (ITO) involves long-term, contractual arrangement[s] in which a service provider takes ownership of and responsibility for managing all or part of a client's IS infrastructure and operations. Network and Endpoint Outsourcing Services (NEOS) involves the support and management of one or more elements of the client/server and network communications infrastructure of an organization. 413 deals (41%) in our dataset were AO, 361 deals (35%) were ITO, and 242 deals (24%) concerned NEOS. In terms of contract type, we distinguish between new contracts, contract extensions, contract expansions, a combined extension and expansion, and contract renegotiation. In our dataset, we found 820 new deals (81%), 150 extensions (15%), and 38 expansions (4%). 8 contracts (<1%) were of a different type. In terms of bid type, 802 of our deals (79%) were closed after a competitive tender, whereas 125 deals (12%) were closed non-competitively. For 89 deals (9%), the bid type was undisclosed. Interestingly, over time, a decreasing share of contracts were closed non-competitively. While 19% to 26% of deals between 2008 and 2010 were non-competitive, their share declined to a mere 2 to 8% between 2014 and 2016.

Study Design and Results

Descriptive Analysis: Empirical Development of ITO Contract Characteristics

Before presenting our explanatory and predictive studies, we first outline our analysis of the behavior of three key contract characteristics, namely the average contract value, contract length, and pricing method. Of the 1,016 observed ITO deals, 436 deals (43%) have a value of less than \$10 million, and another 419 deals (41%) less than \$100 million. 86 deals (9%) are worth between \$100 and 250 million, and another 75 contracts (7%) more than \$250 million. Of those, 18 so-called "mega deals" have a value of more than \$1 billion. While only representing a mere 1% of deals, these "mega deals" account for 51% of the total contract value of all ITO contracts. The median contract value of all contracts is \$13.9 million. In terms of contract length, our data shows a declining development over the last ten years, from an average of 52.3 months for

contracts closed between 2006-08 to an average of 50.6 months between 2015-17 (after Winsoring at 5th and 95th percentile). Similarly, the average annual run rate, i.e., the total contract value divided by contract length in years, has also been in decline for the major part of the last ten years, from an absolute high of \$21.2 million in 2007 to \$10.8 million in 2017 (excluding mega deals). These two trends are in line with the findings of Bapna et al. (2010), who already identified recent outsourcing relationships to “involve several vendors and typically run over a shorter time span” (p. 786). Figure 1 illustrates our findings on contract value, contract length, and annual run rate. With regard to pricing, the market follows the typical distinction between fixed price and variable pricing, which includes consumption-based or on-demand pricing. 525 deals (52%) in our dataset were fixed-price contracts, and 446 deals (44%) were a combination of fixed and variable pricing (mixed pricing). The remaining 45 deals (4%) relate to other pricing methods. For a more exhaustive presentation/discussion of findings from our descriptive analysis, refer to Könning et al. (2018).

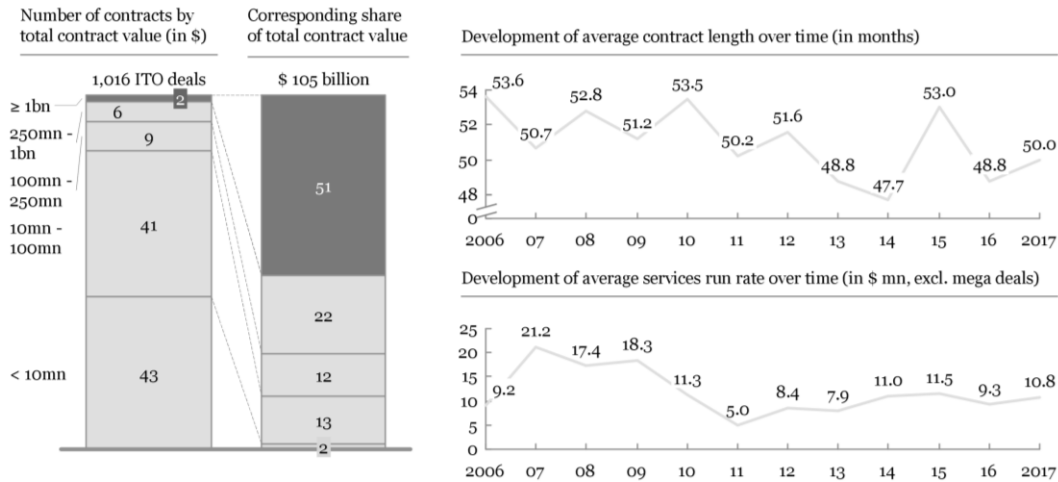


Figure 1. Findings of the Descriptive Analyses on Contract Value, Contract Length, and Annual Run Rate

Explanatory Study: Testing for Influences on ITO Contract Characteristics

In line with Shmueli and Koppius (2011), our first approach consisted of an explanatory statistical model and an evaluation method to be applied after model parameters are estimated. The analysis goal is defined by testing causal hypotheses on the ITO dataset using statistical inference in a Generalized Linear Model (GLM), a method frequently used in explanatory research (Fahrmeir and Tutz 2001; McCullagh 1984). As a first step, we defined our variables of interest. Using ITO theory, we selected the following pivotal contract characteristics as dependent variables: contract value (CV), contract length (CL), and pricing method (PM). A descriptive summary of those variables is given in Table 1. CV and CL are ratio scale attributes measured in dollar and months, respectively. The related models are estimated on all cases with n=1,016. The PM variable distinguishes only between two pricing levels: a fixed price and a mixed method, where the latter also includes dynamic pricing. We omitted the other minor categories of pricing methods from the dataset due to an insufficient subsample size (n<10) resulting in a sample of n=971 for the PM model.

Variable	Measurement	Mean/mode	Std Dev.	Min	Max
CV	Ratio Scale	\$103,798,600	\$563,407	\$75,000	\$11,657,096,541
CL	Ratio Scale	54.33 months	20 months	2 months	180 months
PM	Nominal Scale (binary)	Fixed (0.54)	-	Combination (0.46)	Fixed (0.54)

Table 1. Descriptive Statistics of Pivotal Contract Characteristics (Dependent Variables)

As a second step towards model estimation, we needed to prepare the data, in order to constrain the model to support statistical inference and hypothesis testing. This is particularly important as the dataset consists of 59 (out of 64) polynomial variables. Some of those variables contain a large number of levels (e.g., customer name, with 609 levels) that are subject to dummy encoding, resulting in a large number of attributes that involuntarily creates the curse of dimensionality that makes explanatory modeling cumbersome (Geenens 2011). Several preprocessing steps were taken for the final model. This included dealing with missing values through imputation or deletion, and removal of several nominal attributes with low variation in level frequencies or in high-level count. We also removed highly correlated features. After

preprocessing, our dataset contained $m=18$ explanatory variables. For each of the three dependent variables, the other two remained in the model as explanatory variables, so that we ended up with $m=17$ explanatory variables for each model. The variables are detailed in Appendix Table 6. For empirical hypotheses generation, we can assume that the dataset was put together by ITO experts and that it contains plausible candidates for explanatory variables. Thus, we can state the hypothesis “*we assume that variable j has an influence on CV (CL, PM respectively)*” for every j that is contained within the set of our 18 candidates. By applying dummy coding of polynomial features, the dataset expands to 63 variables. For modeling purposes, we used Gaussian-based GLMs for the ratio-scaled dependent variables (SCL, CL), and a Bernoulli-based GLM for the binary dependent variable (PM). The reference category for PM is “Combination Pricing”. We implemented an iterative Reweighted Least Squares Method (IRLSM), with a collinearity check for parameter estimation (Daubechies et al. 2010). The results of the model are summarized in Table 2. The prefix of the variable name (e.g., “M” for market) reflects the grouping depicted in Appendix Table 6. For reasons of clarity and comprehensibility, we only point out statistically significant dependencies. The table consists of the pool of all significant explanatory variables for all three models (CV, CL, PM). Thus, the table has some sparse fields denoted by x, which indicate that a certain variable was not significant for that particular dependent variable.

Explanatory variable	Dependent variable		
	CV (R ² =0.26)	CL (R ² =0.18)	PM (R ² =0.35)
<i>C_Number of Employees</i>	1.85*10 ⁹ ***	x	x
<i>CD_Contract Length (CL)</i>	1.12*10 ⁹ ***	x	x
<i>CD_Contract Value (CV)</i>	x	4.13***	x
<i>CD_AwardType = Master Service Agreement</i>	12.72*10 ⁹ ***	31.21**	x
<i>CD_Contract Status = Contract cancelled¹</i>	9.35*10 ⁹ ***	23.18*	x
<i>M_Number of submarkets</i>	0.64*10 ⁹ *	2.58*	-0.67***
<i>CD_Bid Type = Non-competitive</i>	1.80*10 ⁹ *	x	x
<i>CD_Bid Type = Competitive</i>	x	x	-1.77***
<i>C_Revenue</i>	x	-2.29***	x
<i>M_Discrete/Bundled = Discrete¹</i>	1.42*10 ⁹ *	x	x
<i>CD_Contract Type²</i>	x	-47.88***	x
<i>SG_Global Sourcing = Offshore / Nearshore¹</i>	x	6.94**	x
<i>GS_Geographic Scope = Latin America¹</i>	x	50.00*	x
<i>C_Macro Industry = Government¹</i>	x	-9.55*	-0.88**
<i>C_Market = Hosted Application Management¹</i>	x	11.81*	-2.88***
<i>SG_Global Sourcing = Onshore¹</i>	x	x	1.55***
<i>REL_Existing Relationship = No¹</i>	x	x	1.18*
<i>C_Macro Industry = Other¹</i>	x	x	-0.87*

Table 2. GLM Results of the Explanatory Study (sig. level: *=0.05, **=0.01, ***=0.001)

The results show that most of the significant influence factors are variables that reflect contract details (CD: 7 variables), followed by client (C: 5 var.), market (M: 2 var.), sourcing geography (SG: 2 var.), client/vendor relationship (REL: 1 var.), and geographic scope (GS: 1 var.). First, we examine the variables that influence the CV. We can see from Table 2 that, besides the obvious influence factor of *CL* and *number of submarkets*, *MSA contracts* have significantly greater underlying CVs. Furthermore, a *cancellation* of a contract seems to happen more often with a greater magnitude of the underlying CV. In addition, *discrete deals* and *competitive bids* have significantly higher CVs. From the variables that characterize the deal partners, only the *number of client employees* shows a significant influence towards higher contract values. Looking at CL, we find that, besides *CV* and *number of submarkets* (i.e. the number of different task categories the service provider fulfills), the *contract type* has an average negative effect on the contract length, regardless of the type of contract involved, resulting in only marginal differences throughout the different contract types. We find that *offshore sourcing* deals have significantly higher contract lengths. In addition, deals with the *geographic scope of Latin America* have higher contract lengths. The only significant *market* that drives the CL is *hosted application management*. Finally, contracts closed with public institutions (coded as *macro industry: government*) show significantly lower contract lengths. The results of the probability

¹ Table entry only reflects the significant level of nominal variables, e.g., the variable *award type* only has an effect on *CV* if the *award type* is *MSA*, other award types are not significant.

² Average effect over all levels (only where all levels are significant and have same direction of influence).

model for PM show that *onshore sourcing* has a significant positive influence on the probability for combined pricing. In contrast, *government-issued contracts* and contracts in the *hosted application management market* significantly tend towards a fixed price mechanism. Another important finding is that contracts built upon an *existing relationship* between client and vendor show a significantly higher probability of mixed pricing. Since we conducted the analysis over a 10 year period we checked for consistency by testing our hypotheses for the 2016/2017 period and found only one minor change that is reflected by significant influence of the *manufacturing industry* in the CL model.

Although we employed an explanatory approach to provide hypothesis testing for the dataset, we find that the overall model quality is rather poor. Therefore, we conducted a second study that concentrated on a predictive study design, which is concerned with overall model performance of predictions.

Predictive Study: Measuring Predictive Power of Feature Space

In contrast to explanatory modeling, which intends to identify causal relationships and to specify how and why certain empirical phenomena occur, predictive modeling aims to predict the outcome of future or unknown events. To assess a model's prediction ability, the accuracy is measured on out-of-sample data, which have not been used for model building, to ensure generalizability. For numeric outcomes (i.e., regression task), the predictive power is quantified via residuals that measure the difference between predicted and observed values; for nominal outcomes (i.e., classification task), metrics are used that assess the amount of correctly/incorrectly classified objects (Shmueli and Koppius 2011). Thus, the goal of this second study is to design empirical models with high predictive power – a goal that can be achieved by using the high dimensionality of the dataset and applying algorithms that are capable of identifying complex relationships in the data. Exemplary algorithms come from the field of Machine Learning, and include Decision Trees (DT) or Neural Networks (NN). Such methods are often non-parametric, and their flexibility enables them to capture non-linearities and higher-level interactions between variables without making statistical assumptions. Even though the price of this flexibility is lower transparency in terms of lacking model interpretability, such methods are predominantly used for forecasting due to their ability to accurately predict new observations. Moreover, it is recommended to apply diverse methods and then compare their accuracy, since learning algorithms like DT or NN use different modeling mechanisms to learn the inner relationships of data variables (Shmueli and Koppius 2011).

In our particular case, the predictive study was carried out for all three previously-considered target variables (CL, CV, and PM). Since we intended to increase dimensionality of the data, we modified the preprocessing steps from the explanatory study. Thus, we kept variables with unbalanced distributions, objectively small explanatory power, and a large amount of levels; we ended up with $m=32$ prediction variables (besides the three target variables, c.f. Appendix, Table 6), including only three numeric variables and 14 categorical variables with more than five levels. Furthermore, the data was split into training and test data in a ratio of 70:30, to avoid overfitting and to evaluate the generalizability between different prediction models on the same out-of-sample data. For CL and CV prediction, we applied four different regression algorithms due to the numeric scaling. These included a simple Linear Regression (LIN-R), a Boosted Decision Tree Regression (BDT-R), a Decision Forest Regression (DF-R), and a Neural Network Regression (NN-R). The predictive power was assessed via two commonly-used metrics that express the average model prediction error in units: the mean absolute error (MAE) and the root mean squared error (RMSE) (Chai and Draxler 2014). As can be seen in Table 3, the performance of the LIN-R was worst for both target variables, since it was not able to capture the non-linearities within the data. However, even if the other models performed much better, they over- or underestimated – for example the contract length on average by ~1-1.5 years, due to the strong impact of outliers. This effect was even higher in the case of CV prediction, since the few mega deals distort the results, and the models are not capable of correctly estimating the contract value based on the given variables.

Regression	LIN-R	BDT-R	DF-R	NN-R
MAE (CL)	34.97 months	13.18 months	12.56 months	14.62 months
RMSE (CL)	46.10 months	18.44 months	18.19 months	20.70 months
MAE (CV)	\$407,402,205	\$109,124,622	\$111,999,329	\$193,032,730
RMSE (CV)	\$833,444,361	\$602,920,453	\$684,233,613	\$728,631,033

Table 3. Regression Results for CL and CV Prediction

Moreover, it can be stated that the low number of metric predictors makes it difficult to correctly estimate a numeric outcome. For this reason, in a second run, we turned the CV prediction into a classification task to achieve higher predictive power. Thus, the target variable was converted to a nominal scaling using the four levels defined by IDC ($CV < \$10\text{mn}$; $\$10\text{mn} \leq CV < \100mn ; $\$100\text{mn} \leq CV < \250mn ; $CV \geq \$250\text{mn}$). Since the last two categories were strongly underrepresented, they were grouped into a single category, resulting in the following three classes: small (43%), medium (41%), and big (16%) deals. For model development, we applied four multiclass classifiers (MC), namely a Logistic Regression (MC-LGR), a Decision Forest (MC-DF), a Decision Jungle (MC-DJ), and a Neural Network (MC-NN). The predictive power was assessed via overall accuracy, which measures the proportion of correctly classified objects among the total number of cases over all three classes examined (c.f. Table 4, left). While both tree-based classifiers (DF and DJ) showed best model performance, they still lacked discriminatory power. This can also be illustrated within the confusion matrix of the DF classifier (see Table 4, right), where the model tends to classify big deals as medium ones (32.6%) and medium deals as small ones (34.8%).

MC classification	Overall accuracy	MC-DF			
MC-LGR	55.74%	Actual	Predicted		
MC-DF	66.23%	small	small	medium	big
MC-DJ	64.59%	small	78.5%	20.0%	1.5%
MC-NN	56.39%	medium	34.8%	55.3%	9.8%
		big	4.7%	32.6%	62.8%

Table 4. MC Classification Results for CV Prediction (left) and Confusion Matrix for MC DF (right)

As an alternative data binning approach, we applied an equal-height discretization method, using the median as a threshold for CV prediction, to overcome the issues of model training with unbalanced classes (Alfred 2009). Thus, the following two levels were established, turning the task into a binary classification: $CV < \$13.9\text{mn}$ and $CV \geq \$13.9\text{mn}$. For model development, we used four two-class (TC) classifiers (TC-LGR, TC-BDT, TC-DF, and TC-NN), and measured the model results again via accuracy (i.e. proportion of correctly classified objects); we also used another metric called AUC, which determines the area under a receiver operating characteristics (ROC) curve (Fawcett 2006). A similar design was applied to the PM prediction, where (like in the explanatory study) underrepresented classes (4%) were filtered to establish two main categories (i.e., fixed price and variable pricing). The results are depicted in Table 5, where it can be seen that for both tasks, acceptable prediction performances could be achieved across all models. Interestingly, this time the tree-based classifiers did not lead to the best model performances using this simplification of a binary class separation. Nevertheless, the results have to be regarded through a critical lens: while the models yield solid results, for the PM classification, the partitioning of the CV into two classes needs to be considered with caution, due to the highly-fragmented distribution of this metric variable. As such, the question “*Will the contract value be higher or lower than ~\$15 million?*” can be of practical interest; however, it strongly simplifies the scope, since it completely omits the explicit consideration of big deals up to \$100-250 million, or even mega deals beyond \$1 billion. This issue should be taken up in further research, where the aim of the investigation could be, for example, to develop dedicated models capable of predicting highly unbalanced characteristics (such as the occurrence of the rare mega deals).

TC classification	TC-LGR	TC-BDT	TC-DF	TC-NN
Accuracy (CV)	79.9%	77.0%	79.7%	78.3%
AUC (CV)	0.874	0.873	0.867	0.847
Accuracy (PM)	75.6%	78.0%	75.6%	76.3%
AUC (PM)	0.836	0.859	0.821	0.838

Table 5. TC Classification Results for CV and PM Prediction

Conclusion, Limitations, and Outlook

In this study, we sought to shed light on potential influences on contract duration, contract value, and pricing method as three pivotal characteristics of ITO contracts. To this end, we made use of three quantitative approaches, with different data analysis tasks in terms of theory-driven research in IS (as suggested by Gregor 2006), and applied them to a dataset with characteristics of 1,016 ITO deals. Our descriptive analysis provides a focused impression of the ITO market (RQ1). Our evidence found that both contract duration and annual services run rate are in decline for most of the last ten years, confirming previous findings by Bapna et al. (2010). The results are also in line with an observable trend towards multi-vendor sourcing (Herz et al. 2010; Könning et al. 2018; Oshri et al. 2009). Considering the results of our

explanatory study, we identified several significant influences on pivotal contract characteristics (RQ2). Beside obvious results, we found that certain combinations of industry and market yield significantly lower contract lengths and tend to use fixed pricing mechanisms. A downside to using explanatory approaches is the limited bandwidth of methods that come along with forced transformations or exclusions of possible explanatory variable candidates. However, by enriching the analysis with a higher feature space, and applying algorithms from the field of Machine Learning, it could be shown that - to a certain degree - the dataset contains characteristics with non-linear discriminatory power that can be used to build prediction models with solid accuracy (RQ3). Particularly for PM and CV prediction using a binary classification approach, we could achieve accuracies up to nearly 80%. On the other hand, the regression models for CV and CL prediction lacked predictive power due to a small number of numeric variables and a fragmented distribution of the target variables. Nevertheless, it can be stated that the dataset yields acceptable results, which deliver a basis for extending the analyses to cover larger areas of investigation. As such, we see several avenues for further research, and we already plan to extend our geographic scope by using a global dataset with a significantly higher number of data points. We also aspire to enrich our study with data from external sources (such as additional client- and vendor-specific characteristics, general market data like GDP, and other potential influence factors).

Besides the aforementioned methodical limitations, some dataset-induced limitations also exist in our research. While IDC's BuyerPulse service contracts database contains a multitude of data points on ITO contracts, the database is not complete. It relies on publicly available information on ITO contracts, compiled from analyst reports, media articles, or press releases issued by the contracting parties. However, these sources tend to report more on large-scale ITO projects than on the many small contracts that are also closed but not reported on. As a result, a majority of ITO deals cannot be recorded in IDC's database. Furthermore, ITO contracts of large institutions are more often published, which is why the database is subject to an unavoidable systematic sampling error and is biased towards larger ITO deals. In addition, many ITO deals from the ASG region (especially smaller ones) are only announced in German press releases. Hence, the chance is higher that they will be missed by an US-American market institute like IDC. When both these factors are taken into account, IDC suggests their database contains around 10% of existing outsourcing deals. Because the sample is biased towards larger deals, IDC estimates that they represent 20% of combined ITO contract value. Nonetheless, the IDC dataset represents the best perspective currently available on the ITO market, as our data triangulation effort indicated.

References

- Alfred, R. 2009. "Discretization Numerical Data for Relational Data with One-to-Many Relations," *Journal of Computer Science* (5:7), pp. 519–528.
- Bapna, R., Barua, A., Mani, D., and Mehra, A. 2010. "Research Commentary - Cooperation, Coordination, and Governance in Multisourcing: An Agenda for Analytical and Empirical Research," *Information Systems Research* (21:4), pp. 785–795.
- Chai, T., and Draxler, R. R. 2014. "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?: Arguments Against Avoiding RMSE in the Literature," *Geoscientific Model Development* (7:3), pp. 1247–1250.
- Daubechies, I., Devore, R., Fornasier, M., and Güntürk, C. S. 2010. "Iteratively Reweighted Least Squares Minimization for Sparse Recovery," *Communications on Pure and Applied Mathematics* (63:1), pp. 1–38.
- Dibbern, J., Goles, T., Hirschheim, R., Bandula, J., and Jayatilaka, B. 2004. "Information Systems Outsourcing: A Survey and Analysis of the Literature," *The DATA BASE for Advances in Information Systems* (35:4), pp. 6–102.
- Fahrmeir, L., and Tutz, G. 2001. *Multivariate Statistical Modelling Based on Generalized Linear Models*, New York City, NY, USA: Springer-Verlag New York.
- Faisal, M. N., and Raza, S. A. 2016. "IT Outsourcing Intent in Academic Institutions in GCC Countries: An Empirical Investigation and Multi-Criteria Decision Model for Vendor Selection," *Journal of Enterprise Information Management* (29:3), pp. 432–453.
- Fawcett, T. 2006. "An Introduction to ROC Analysis," *Pattern Recognition Letters* (27:8), pp. 861–874.
- Geenens, G. 2011. "Curse of Dimensionality and Related Issues in Nonparametric Functional Regression," *Statistics Surveys* (5:0), pp. 30–43.

- Gewald, H., and Gellrich, T. 2007. "The Impact of Perceived Risk on the Capital Market's Reaction to Outsourcing Announcements," *Information Technology and Management* (8:4), pp. 279–296.
- Gonzalez, R., Gasco, J., and Llopis, J. 2006. "Information Systems Outsourcing: A Literature Analysis," *Information & Management* (43:7), pp. 821–834.
- Goo, J., Kishore, R., Nam, K., Rao, H. R., and Song, Y. 2007. "An investigation of factors that influence the duration of IT outsourcing relationships," *Decision Support Systems* (42:4), pp. 2107–2125.
- Gregor, S. 2006. "The Nature of Theory in Information Systems," *MIS Quarterly* (30:3), pp. 611–642.
- Herz, T. P., Hamel, F., Uebernickel, F., and Brenner, W. 2010. "Deriving a Research Agenda for the Management of Multisourcing Relationships Based on a Literature Review," *Proceedings of the 16th Americas Conference on Information Systems*.
- International Data Corporation. 2018. About IDC. <https://www.idc.com/about>. Accessed 28 February 2018.
- Könning, M., Westner, M., and Strahringer, S. 2018. "Multisourcing on the Rise – Results from an Analysis of more than 1,000 IT Outsourcing Deals in the ASG Region," *Proceedings zur Multikonferenz Wirtschaftsinformatik 2018 (MKWI 2018)*, pp. 1813–1824.
- Lacity, M. C., Khan, S. A., and Yan, A. 2016. "Review of the Empirical Business Services Sourcing Literature: An Update and Future Directions," *Journal of Information Technology* (31:3), pp. 269–328.
- Liang, H., Wang, J.-j., Xue, Y., and Cui, X. 2015. "Information & Management IT Outsourcing Research from 1992 to 2013: A Literature Review Based on Main Path Analysis," *Information & Management*.
- McCullagh, P. 1984. "Generalized linear models," *European Journal of Operational Research* (16:3), pp. 285–292.
- Oshri, I., Kotlarsky, J., Rottman, J. W., and Willcocks, L. L. 2009. "Global Sourcing: Recent Trends and Issues," *Information Technology & People* (22:3), pp. 192–200.
- Payne, J., and Dorn, W. R. 2012. *Managing indirect spend: Enhancing profitability through strategic sourcing*, Hoboken, N.J: Wiley.
- Ravindran, K., Susarla, A., Krishnan, R., and Mani, D. 2015a. "The Mediating Effect of Formal Contractual Controls in the Relationship Between Experience and Contract Design," in *Achieving Success and Innovation in Global Sourcing: Perspectives and Practices*, I. Oshri, J. Kotlarsky and L. P. Willcocks (eds.), La Thuile, Italy. February 18-21, 2015, Cham: Springer International Publishing, pp. 52–61.
- Ravindran, K., Susarla, A., Mani, D., and Gurbaxani, V. 2015b. "Social Capital and Contract Duration in Buyer-Supplier Networks for Information Technology Outsourcing," *Information Systems Research* (26:2), pp. 379–397.
- Ruivo, P., Rodrigues, J., Neto, M., Oliveira, T., and Johansson, B. 2015. "Defining a Framework for the Development of ICT Services "Nearshoring" in Portugal," *Procedia Computer Science* (64), pp. 140–145.
- Shmueli, G., and Koppius, O. R. 2011. "Predictive Analytics in Information Systems Research," *MIS Quarterly* (35:3), pp. 553–572.
- Susarla, A. 2012. "Contractual Flexibility, Rent Seeking, and Renegotiation Design: An Empirical Analysis of Information Technology Outsourcing Contracts," *Management Science* (58:7), pp. 1388–1407.
- Westner, M., and Strahringer, S. 2010. "The Current State of IS Offshoring in Germany: Project Characteristics and Success Patterns," *Journal of Information Technology Management* (21:1), 49–70.

Appendix

Category	Examined Variables
Client (C)	Name, Parent*, Macro Industry**, Industry*, Sub Industry*, Government/Commercial**, SIC Code*, Number of Employees**, Number of Employees (Categorized)*, Revenue**, Revenue (Categorized)*
Vendor (V)	Name, Parent*, Role
C-V-Relation (REL)	Existing Relationship**, Distinct/New Combination**
Contract details (CD)	Award Type**, Bid Type**, Contract Status**, Contract Type**, Pricing Method**, Contract Length**, Contract Value**, Contract Value (Categorized), Contract Run Rate, Project Owner, Advisors, Contractors
Signing (SIGN)	Macro Region, Region, Sub Region, Country*, State/Province*, City*
Market (M)	Macro Market, Market Full Name**, Market*, Sub Market(s)*, Number of Sub Markets**, Discrete/Bundled Contract**, Engagement Type, Cloud (How)*, Cloud (What)*
Geographic scope (GS)	Scope**, Macro Region*, Region*, Country*
Dates (D)	Press Release, Signing Period, Signing Date, Signing Year/Quarter, Start Date, End Date, Signing Period ID, Signing Year, End Period, End Period ID, Last Edited Date, Contract Status Change Date
Sourcing geography (SG)	Global Sourcing**, Macro Region, Region, Country (* ** used for explanatory and predictive modeling, (*) used only for predictive modeling

Table 6. List of All Variables Examined during Descriptive, Explanatory and Predictive Analysis