A Quantitative Analysis of Faulty EPCs in the SAP Reference Model

J. Mendling[1], M. Moser[1], G. Neumann[1], H.M.W. Verbeek[2], B.F. van Dongen[2], and W.M.P. van der Aalst[2]

[1] Vienna University of Economics and Business Administration
Augasse 2-6, 1090 Vienna, Austria
{jan.mendling,h9950347,neumann}@wu-wien.ac.at
[2] Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{h.m.w.verbeek,b.f.v.dongen,w.m.p.v.d.aalst}@tm.tue.nl

Abstract. The *SAP reference model* contains more than 600 non-trivial process models expressed in terms of *Event-driven Process Chains* (EPCs). We have automatically translated these EPCs into YAWL models and analyzed these models using WofYAWL, a verification tool based on Petri nets. We discovered that *at least 34 of these EPCs contain errors* (i.e., at least 5.6% is flawed). We analyzed which parts of the SAP reference model contain most errors. Moreover, based on 15 characteristics (e.g., the size of the model), we used logistic regression to find possible predictors for these errors. This systematic analysis of the SAP reference model illustrates the need for verification tools such as WofYAWL.

1 Introduction

There has been extensive work on formal foundations of conceptual modeling and respective languages. However, little quantitative research has been reported on the actual use of conceptual modeling [5]. Moreover, literature typically discusses and analyses languages rather than evaluating enterprise models at a larger scale (i.e., beyond "toy examples"). A fundamental problem in this context is that large enterprise models are in general not accessible for research as they represent valuable company knowledge that enterprises do not want to reveal. In particular, this problem affects research on reference models, i.e., models that capture generic design that is meant to be reused as best practice recommendation in future modeling projects.

One case of a model that is, at least partially, publicly available is the SAP reference model. It has been described in [4, 14] and is referred to in many research papers (see e.g. [11, 17, 19, 22, 26]). The extensive database of this reference model contains almost 10,000 sub-models, most of them EPC business process models [4, 13, 14]. Fig. 1 shows the EPC model for "Certificate Creation" as an example of one of these models. The SAP reference model was meant to be used as a blueprint for the implementation of SAP's ERP system. It reflects

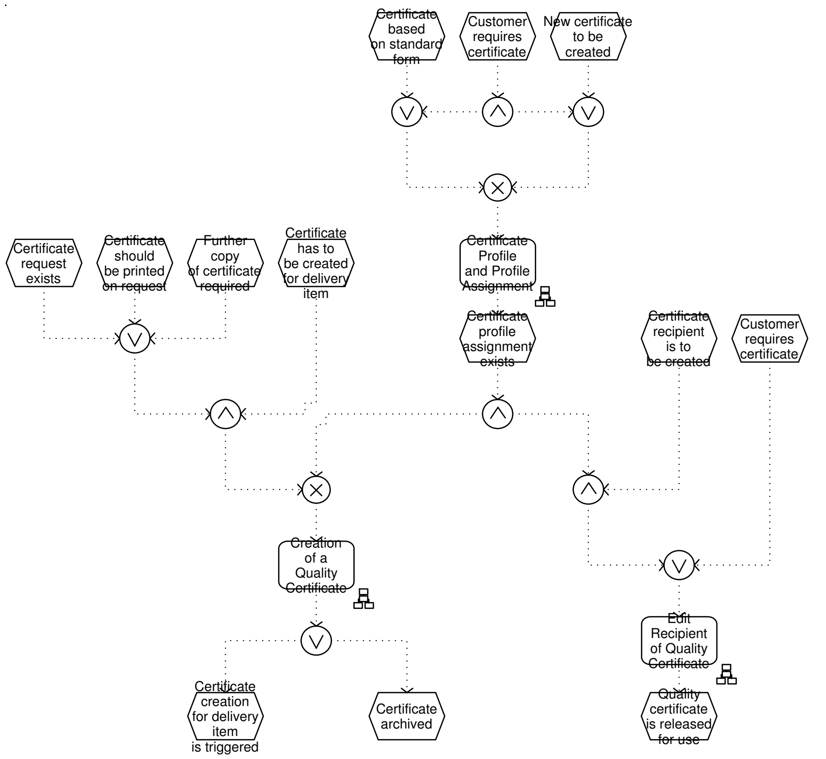

Fig. 1. One of the EPCs in the SAP reference model: the "Certificate Creation" process

Version 4.6 of SAP R/3 which was marketed in 2000. Building on recently developed techniques to verify the formal correctness of EPC models as reported in [27], we aim to acquire knowledge about how many formal modeling errors can be expected in a large repository of process models in practice, assuming that the SAP reference model can be regarded as a representative example. We will map all non-trivial EPCs in the SAP reference model onto YAWL models [1] and use the WofYAWL tool [27] as a means to verify the correctness of these EPC (using the relaxed-soundness criterion [6]). We have to stress that this analysis yields a lower bound for errors since some errors may not be discovered by this tool. Furthermore, wrong model content (wrong element labels, wrong order of elements) cannot be detected by WofYAWL. Therefore, it has to be expected that there are more errors than those that we actually identify.

The remainder of this paper is organized as follows. Section 2 describes the design of our quantitative study. In particular, we discuss the mapping of EPCs from the SAP reference model to YAWL models, the analysis techniques employed by WofYAWL, and the identification of how the models can be corrected. In Section 3 we focus on the analysis of the non-trivial EPCs in the SAP reference model. First, we calculate descriptive statistics that allow us to get a comprehensive inventory of errors in the SAP reference model. Secondly, we investigate the hypothesis that more complicated models have more errors. This hypothesis

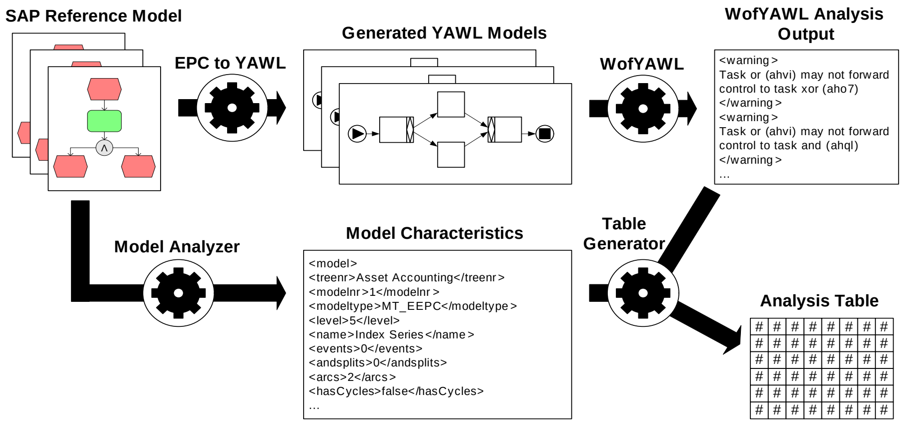

Fig. 2. Overview of the Evaluation Design

was suggested in [2] and we analyze it using different complexity measures and by testing whether they are able to explain the variance of errors. The results allow us to conclude which complexity metrics are well suited to explain error variance and that the impact of complexity on error probability is significant. Subsequently, we discuss our findings in the light of related research (Section 4) and conclude with a summary of our contribution and its limitations (Section 5).

2 Evaluation Design

In this section, we present the way we evaluated the SAP reference model. We use the ARIS XML export of the reference model as input to several transformation and analysis steps (see Fig. 2). In a first step, the EPC to YAWL transformation program generates a YAWL XML file for each EPC in the reference model (see Section 2.1). These YAWL models are then analyzed with WofYAWL that produces an XML error report highlighting the design flaws than have been discovered (see Section 2.2). Independent from these steps, the Model Analyzer extracts descriptive information such as the number of elements of a certain element type and whether there are cycles for each EPC model. An XML file of these model characteristics is then merged with the output of WofYAWL based on the ID of each EPC, and written to an analysis table in HTML format. Then, this table is imported in SPSS to do the statistical analysis. Additionally, Section 2.3 reports on how erroneous EPC models can be corrected.

2.1 Transformations of EPCs to YAWL

Several mappings from EPCs to Petri Nets have been proposed in order to verify formal properties, see e.g. [15] for an overview. In this paper, we use a transformation from EPCs to YAWL that has been recently defined in [18]. The advantage is that each EPC element can be directly mapped to a respective

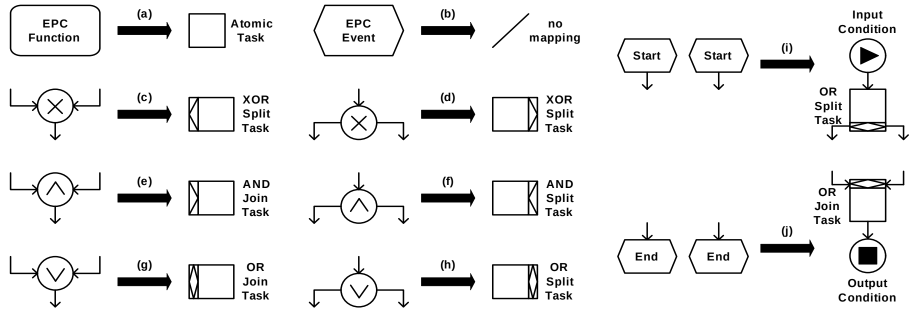

Fig. 3. Overview of the EPC to YAWL Mapping

YAWL element if a local interpretation of the XOR-join is assumed (see Fig. 3). Even though EPCs and YAWL are very similar in this sense, there are three differences that have to be considered in the transformation: state representation, connector chains, and multiple start and end events.

EPC functions can be mapped to YAWL tasks following mapping rule (a) of Fig. 3). The first difference between EPCs and YAWL is related to *state representation*. EPC events can be interpreted as states that define preconditions for the start of functions and postconditions after their completion. Though this definition might suggest a direct mapping of events to YAWL conditions (the YAWL equivalent to places in Petri nets), there is a problem of alternative event-function and function-event connectors: consider an event that leads to an AND split which is followed by two functions. Here we have one event, but there are actually two conditions needed to represent the preset of the subsequent functions. Accordingly, EPC events are related to states, but they do not directly match conditions in YAWL. Therefore, rule (b) defines that events are not mapped to YAWL taking advantage of the fact that arcs in YAWL represent implicit conditions if they connect two tasks. In EPCs connectors are independent elements. Therefore, it is allowed to build so-called *connector chains*, i.e. paths of two or more consecutive connectors (cf. Fig. 1). In YAWL there are no connector chains since splits and joins are part of tasks. The mapping rules (c) to (h) map every connector to a dummy task with the matching join or split condition (see Fig. 3). The third difference stems from *multiple start and end events*. An EPC is allowed to have more than one start event. Multiple end events represent implicit termination: the triggering of an end event does not terminate the process as long as there is another path still active. In YAWL there must be exactly one start condition and one end condition. Therefore, the mapping rules (i) and (j) generate an OR split for multiple starts and an OR join for multiple ends. Fig. 4 gives the result of applying the transformation to the "Certificate Creation" EPC of the first section. Note that connectors are mapped onto dummy tasks. To identify these tasks they are given a unique label extracted from the internal representation of the EPC, e.g., task "and (c8z0)" corresponds to the AND-split connector following event "Customer requires certificate".

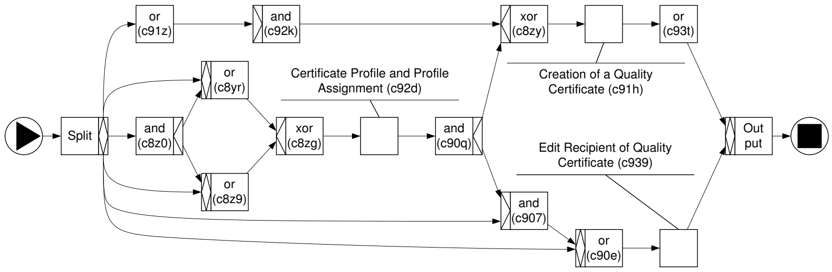

Fig. 4. YAWL model obtained by applying the mapping shown in Fig. 3 to the running example

2.2 WofYAWL Analysis

After mapping the EPC onto YAWL, we can use our verification tool WofYAWL [28]. WofYAWL is Petri-net based. Therefore, it first maps a YAWL model onto a Petri net [7, 20, 21]. Fig. 5 sketches a small fragment of the Petri net that results from mapping the YAWL model of Fig. 4. The fragment only considers the dummy tasks resulting from the mapping of the top four connectors in Fig. 1. Moreover, from the initial OR-split task “Split” in Fig. 4 we only consider the arcs connected to these four dummy tasks. Note that when mapping this OR-split onto transitions all possible interpretations are generated ($2^3-1 = 7$ transitions). Similarly, all other XOR/OR-splits/joins are unfolded.

The “happy smileys” in Fig. 5 are used to identify net elements that are involved in so-called “good execution paths”, that is, the execution paths in the Petri net that lead from the initial state to the *desired* final state. In Fig. 5, there exist two such paths, which join at the XOR-join named “xor (c8z9)”. The “sad smileys” visualize relevant parts in the Petri net that are not covered by some good execution path. As a result, these parts can in no way contribute to reaching the desired final state from the initial state. Since there is definitely something wrong with such parts, WofYAWL issues the following warnings for this fragment:

- `Task "or (c8yr)" may not receive control from task "and (c8z0)"`,
- `Task "or (c8z9)" may not receive control from task "and (c8z0)"`,
- `Task "or (c8yr)" may be an XOR-join instead of an OR-join`,
- `Task "or (c8z9)" may be an XOR-join instead of an OR-join`.

These warnings indicate that there is a problem involving the top four connectors in Fig. 1. Note that AND-split connector splits the flow into two paths that

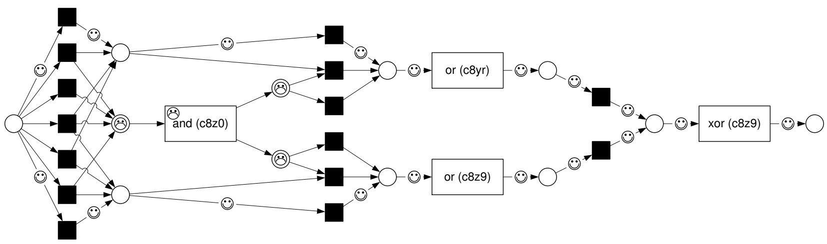

Fig. 5. Petri net fragment of the converted YAWL model

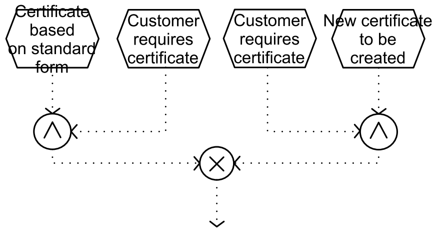

Fig. 6. Fragment of an alternative "Certificate Creation" EPC addressing the problems identified using WofYAWL

join with and XOR-join. Hence these two paths cannot be involved in a good execution as indicated by first two warnings. Moreover, if the AND-split connector is not allowed to occur, the two OR-joins could as well be XOR-joins. In Section 2.3 we will show how these diagnostics can be used to repair the problem.

In our analysis we use *transition invariants* to avoid constructing large or even infinite state spaces [27]. However, the mapping shown in Fig. 3 tends to generate very large models. For example, in the SAP reference model there are EPCs with 22 end events. Using the naive translation shown in Fig. 3 this results into 4 million transitions just to capture the final OR-join. Therefore, we have used a more refined mapping which scales much better. Moreover, we have used Petri-net-based reduction rules [20] to further reduce the complexity of the models without loosing any information. For additional details on this approach, we refer to our technical report [27].

2.3 Identification of Errors

Errors in EPCs can be identified in an automated way using WofYAWL. However, being able to detect problems is not enough. In practice, these problems should be repaired by the process owner. Take the EPC of Fig. 1 for example. In Section 2.2, we have shown that there were four error messages coming from WofYAWL. From this, it is rather trivial to conclude that the AND-split connector following the event "Customer requires certificate" can never occur, since it would always cause the following XOR-join to block[1]. To repair this mistake, the problem owner should decide whether to change the AND-split into an XOR-split, or to change the XOR-join into an AND-split. The decision cannot be made without explicit domain-knowledge of the process under consideration, and might even be different for each implementation of the process. In its current form however, the process model cannot be used.

In some cases, WofYAWL generates a message, suggesting that an OR-connector could be changed to an XOR. If such a message is generated for a connector in isolation (i.e. there are no other messages regarding the same connector) then this connector can indeed be changed without disturbing the model. However, if other messages relate to the same connector (which is the case in our example) special care has to be taken. In the "Certificate Creation" model

[1] For this conclusion, we followed the executable semantics of the ARIS-Simulation.

Table 1. Hierarchy Levels of the SAP Reference Model

Hierarchy Level	Models	eEPC	Function Allocation Diagram	Process Selection Diagram	Role Activity Diagram	EPC	Error
1	1	1	0	0	0	0	0
2	58	29	0	29	0	0	0
3	175	73	0	0	0	102	15
4	1226	724	0	0	0	502	19
5	8384	3035	3035	0	2014	0	0
All Levels	9844	3862	3035	29	2014	604	34

for example, the connectors can only be changed to an XOR-join under the assumption that the event "Customer requires certificate" cannot occur. Since this is not a valid assumption, we propose to repair the EPC as shown in Fig. 6.

3 Analysis of the SAP Reference Model

Using the approach depicted in Fig. 2 we analyze the SAP reference model. First of all, we analyze in which parts of the reference model most errors occur (Section 3.1). Second, in Section 3.2, we formulate hypotheses relating correctness to properties of the EPC (e.g., larger models are more likely to contain errors). Finally, we test these hypotheses using logistic regression (Section 3.3).

3.1 Descriptive Statistics

The sample of the SAP reference model that was available for this research is organized in two orthogonal dimensions: hierarchy levels and branches. Table 1 illustrates that five levels of abstraction are used to arrange the models. Each model at a lower level is a sub-model of a model on a higher level. On the top level there is one model which serves as the root for the model hierarchy. Most of the 9844 models are of model type extended EPC ("eEPC"), but only a fraction of them represent proper EPCs with at least one start event and one function. There are 604 of such process models as listed in the column "EPC". These EPCs have been the starting point of our analysis. Using the transformations and the WofYAWL tool described in Section 2, we discovered that at least 34 models have errors (5.6% of 604 analyzed EPCs).

Table 2 summarizes the SAP reference model subdivided into its 29 branches. It can be seen that the number of EPC models varies substantially (from none in Position Management to 76 in Sales & Distribution). Furthermore, the EPCs are of different size indicated by the mean number of events, functions, connectors, and arcs in columns $E_{av.}$, $F_{av.}$, $C_{av.}$, $A_{av.}$ respectively. The column "Cycle" states how many EPCs are cyclic, and "Error" for how many models WofYAWL reports an error. It is interesting to note that branches with more than 10% of faulty models tend to be larger. For example, refer to the Real Estate Management branch: 16.7% of the EPCs have errors and the mean number of events (12.7) per EPC is higher than the overall mean number of events (11.5). Similar observations can be made for functions (6.5 to 4.0), connectors (7.3 to 5.2), and arcs (27.0 to 20.8). In the following subsection, we test whether such characteristics of an EPC can be used to predict errors.

Table 2. Branches of the SAP Reference Model. The columns $E_{av.}$, $F_{av.}$, $C_{av.}$, $A_{av.}$ refer to the mean number of events, functions, connectors, and arcs.

Branch	Model	%	EPC	%	$E_{av.}$	$F_{av.}$	$C_{av.}$	$A_{av.}$	Cycle	Error	%
Asset Accounting	461	4.7%	43	7.1%	13.9	4.0	5.2	23.3	0	7	16.3%
Benefits Administration	50	0.5%	6	1.0%	9.5	3.3	5.8	19.7	3	0	0.0%
Compensation Management	122	1.2%	18	3.0%	7.6	3.4	3.3	13.7	3	1	5.6%
Customer Service	402	4.1%	41	6.8%	16.5	3.6	9.0	29.5	3	1	2.4%
Enterprise Controlling	599	6.1%	22	3.6%	14.3	10.1	6.1	32.1	0	3	13.6%
Environment, Health, Safety	102	1.0%	19	3.1%	3.5	2.7	1.2	7.0	0	0	0.0%
Financial Accounting	614	6.2%	54	8.9%	13.0	4.0	5.1	21.8	0	3	5.6%
Position Management	4	0.0%	0	0.0%	0.0	0.0	0.0	0.0	0	0	n.a.
Inventory Management	184	1.9%	3	0.5%	15.0	7.0	6.0	28.0	2	0	0.0%
Organizational Management	37	0.4%	5	0.8%	12.0	3.0	6.6	24.0	3	0	0.0%
Payroll	541	5.5%	7	1.2%	5.7	3.1	2.1	11.4	0	1	14.3%
Personnel Administration	15	0.2%	4	0.7%	7.3	1.5	4.0	12.3	0	0	0.0%
Personnel Development	60	0.6%	10	1.7%	8.7	2.5	4.4	15.6	3	1	10.0%
Personnel Time Management	87	0.9%	12	2.0%	10.8	3.0	5.3	19.5	1	2	16.7%
Plant Maintenance	399	4.1%	35	5.8%	20.5	4.2	11.4	37.8	9	1	2.9%
Procurement	444	4.5%	37	6.1%	6.7	3.5	2.7	12.4	0	2	5.4%
Product Data Management	366	3.7%	26	4.3%	4.5	5.4	2.2	13.7	0	0	0.0%
Production	296	3.0%	17	2.8%	8.8	3.0	2.9	13.7	0	1	5.9%
Production Planning	194	2.0%	17	2.8%	5.7	2.9	3.0	11.5	0	0	0.0%
Project Management	347	3.5%	36	6.0%	8.5	3.8	2.2	14.0	0	0	0.0%
Quality Management	209	2.1%	20	3.3%	20.5	3.8	11.7	37.8	1	1	5.0%
Real Estate Management	169	1.7%	6	1.0%	12.7	6.5	7.3	27.0	1	1	16.7%
Recruitment	56	0.6%	9	1.5%	7.4	2.6	4.1	13.8	3	0	0.0%
Retail	842	8.6%	1	0.2%	7.0	5.0	2.0	11.0	0	0	0.0%
Revenue & Cost Controlling	568	5.8%	19	3.1%	16.5	10.2	7.9	36.0	1	1	5.3%
Sales & Distribution	703	7.1%	76	12.6%	10.6	3.1	4.3	16.6	0	1	1.3%
Training & Event Management	95	1.0%	12	2.0%	13.0	2.7	6.2	22.2	0	1	8.3%
Travel Management	116	1.2%	1	0.2%	24.0	7.0	16.0	48.0	0	0	0.0%
Treasury	1761	17.9%	48	7.9%	10.5	3.5	4.5	18.1	0	6	12.5%
All 29 Branches	9844	100%	604	100%	11.5	4.0	5.2	20.8	33	34	5.6%

3.2 Hypotheses and Related Error Determinants

Determinants of errors in EPCs can be related to several aspects. In this subsection we discuss model size, model complexity, and typical error patterns.

Model Size: The size of the model can be considered as a potential error determinant if the model is produced by a human modeler. Simon [25] points to the limited cognitive capabilities and concludes that humans act only rational to a limited extent. In the context of modeling, this argument would imply that human modelers loose track of all interrelations of a large model due to their limited cognitive capabilities, and then introduce errors that they would not insert in a small model. Accordingly, we define the following hypotheses:

- S_1 : A higher number of events E increases the error probability.
- S_2 : A higher number of functions F increases the error probability.
- S_3 : A higher number of connectors C increases the error probability.
- S_4 : A higher number of arcs A increases the error probability.

Model Complexity: Recent work by Cardoso [2] discusses complexity as an error source. Similar to large models, the modeler is expected to introduce errors more likely in complex models due to limited cognitive capabilities. Yet, complexity may differ from size, e.g., a large sequence may be less demanding for a modeler than small model containing several joins and splits. In EPCs complexity is introduced by *connectors*. This supports S_3. Moreover, two EPCs can have the same number of connectors, but differ in complexity if the second model introduces additional *arcs* between the connectors. Therefore, S_4 is also backed from a complexity point of view. *Cycles* represent an additional aspect of complexity. Arbitrary cycles can lead to EPC models without clear semantics as shown in [16]. Cardoso introduces a *complexity metric* based on the observation that the three split connector types introduce a different degree of complexity. According to the number of potential post-states an AND-split is weighted with 1, an XOR-split with the number of successors n, and an OR-split with $2^n - 1$. We refer to the sum of all connector weights of an EPC as split-complexity SC (called Control-flow Complexity CFC in [2]). Analogously, we define the join-complexity JC as the sum of weighted join connectors based on the number of potential pre-states. Furthermore, we assume that a mismatch between potential post-states of splits and pre-states of joins can be modeled with the split-join-ratio $JSR = JC/SC$. Based on this we formulate the following hypotheses:

- C_1 : A higher number of connectors C increases the error probability.
- C_2 : A higher number of arcs A increases the error probability.
- C_3 : EPCs *with cycles* have a higher error probability than EPCs without.
- C_4 : A higher SC value of an EPC increases the error probability.
- C_5 : A higher JC value of an EPC increases the error probability.
- C_6 : A higher JSR value of an EPC increases the error probability.

Error Patterns: The last set of hypotheses is based on typical patterns that may point at potential problems. EPCs lack an explicit notion for the initial state, i.e., unlike a Petri net it is nor clear in which state the EPC starts because multiple start events may become triggered. This is reflected by the initial OR-split when translating an EPC to YAWL. Clearly, this may introduce errors and therefore the number of start events may influence the likelihood of errors being introduced. A similar observation may be made for the number of end events. A well-know source of errors are the so-called PT- and TP-handles in Petri nets [10]. A PT-handles starts with a place with multiple outgoing arcs joining later in a single transition. In terms of EPCs this means that an XOR-split connector corresponds to an AND-join connector. Clearly, this may indicate a deadlock problem: the process gets stuck just before AND-join. Similarly, an OR-split connector corresponding to an AND-join connector may be problematic. TP-handles are the reverse of PT-handles and start with a transition (AND-split) where outgoing arcs come together in a place (XOR-join). In terms of EPCs this corresponds to an AND-split or OR-split connector with a matching XOR-join connector. This establishes the following hypotheses:

- EP_1 : A higher number of start events increases the error probability.
- EP_2 : A higher number of end events increases the error probability.
- EP_3 : A higher number of XOR/OR-splits and AND-joins in an EPC increases the error probability.
- EP_4 : A higher number of AND/XOR-splits and XOR-joins in an EPC increases the error probability.

Table 3. Potential Determinants for Errors in the SAP Reference Model

Symbol	Definition	Motivation
A	Number of Arcs	S_4, C_2
E_{start}	Number of Start Events	S_1, EP_1
E_{end}	Number of End Events	S_1, EP_2
E_{int}	Number of Internal Events	S_1
F	Number of Functions	S_1
AND_j	Number of AND joins	S_1, C_1, EP_3
AND_s	Number of AND splits	S_1, C_1, EP_4
XOR_j	Number of XOR joins	S_1, C_1, EP_4
XOR_s	Number of XOR splits	S_1, C_1, EP_3
OR_j	Number of OR joins	S_1, C_1
OR_s	Number of OR splits	S_1, C_1, EP_3, EP_4,
Cycle	if the EPC has cycles	C_3
SC	Split Complexity	C_4
JC	Join Complexity	C_5
JSR	Join-Split-Ratio	C_6

Table 3 summarizes the input variables that we will investigate. The table also shows how these variables can be linked to the discussed hypotheses.

3.3 Testing of Error Determinants

We now utilize the analysis table of the SAP reference model (cf. Fig. 2) to test the significance of our hypotheses. The potential determinants listed in Table 3 serve as input variables to explain the variance of the dependent variable “hasError”. As the dependent variable is binary, we use a logistic regression (logit) model. The idea of a logit model is to model the probability of a binary event by its odds, i.e., the ratio of event probability divided by non-event probability. The relationship between input and dependent variables is represented by an S-shaped curve of the logistic function that converges to 0 for $-\infty$ and to 1 for ∞. The cut value of 0.5 defines whether event or non-event is predicted. $Exp(B)$ gives the change of the odds if the input variable is increased by one unit: $Exp(B) > 1$ increases and $Exp(B) < 1$ decreases error probability.

The significance of the overall model is assessed by the help of two statistics. First, the Hosmer&Lemeshow Test should be greater than 5% to indicate a good fit based on the difference between observed and predicted frequencies. Second, Nagelkerke's R^2 ranging from 0 to 1 serves as a coefficient of determination indicating which fraction of the variability is explained. Furthermore, each estimated coefficient of the logit model is tested using the Wald statistic for being significantly different from zero. The significance should be less than 5%. In Table 4 we also give the percentage of correct classifications and the number of wrong and correctly predicted faulty EPCs. As our sample includes only 5.6% error cases, a correct classification of 94.4% can easily be achieved by always predicting that the EPC is correct. Therefore, the number of correctly predicted errors is more interesting in this context. For more details on logistic regression see e.g. [12].

As a first step we calculated univariate logit models for each of the 15 input variables.[2] Each model for the 11 variables that indicate the number to elements of a specific type in the EPC had a Wald statistic at a significance level of 0.6%

[2] Due to space limitations, we do not give a table of the univariate results here.

Table 4. Multivariate Logit Models based on potential Error Determinants

	Complete Model		Without SC and JC		8-Step Model		5-Step Model	
Coefficient	Exp(B)	Wald Sig.	Exp(B)	Wald Sig.	Exp(B)	Wald Sig.	Exp(B)	Wald Sig.
Constant	0.023	0.0%	0.028	0.0%	0.024	0.0%	0.025	0.0%
A	1.097	39.0%	1.081	47.8%	-	-	-	-
E_{start}	0.641	0.2%	0.666	0.4%	0.719	0.2%	0.844	2.4%
E_{end}	1.151	24.3%	1.057	63.2%	1.128	6.1%	-	-
E_{int}	1.069	70.6%	1.045	80.8%	1.151	0.5%	1.162	0.3%
F	0.906	36.8%	0.903	35.8%	-	-	-	-
AND_j	1.065	81.8%	1.190	51.6%	1.321	10.9%	-	-
AND_s	0.786	35.7%	0.932	77.8%	-	-	-	-
XOR_j	1.705	3.8%	1.795	2.3%	2.010	0.0%	1.559	0.9%
XOR_s	0.493	0.6%	0.589	2.4%	0.654	2.2%	-	-
OR_j	2.209	0.3%	2.067	0.5%	2.233	0.0%	1.939	0.1%
OR_s	0.432	0.6%	0.426	0.6%	0.473	0.2%	0.639	0.9%
Cycle	0.951	94.1%	0.990	98.8%	-	-	-	-
SC	1.000	59.3%	-	-	-	-	-	-
JC	1.000	97.2%	-	-	-	-	-	-
JSR	1.032	45.6%	1.023	60.3%	-	-	-	-
Hosmer&Lem. Sig.		10.3%		89.5%		62.9%		52.0%
Nagelkerke R^2		0.326		0.304		0.300		0.266
Correct Classif.		95.2%		95.2%		94.7%		95.0%
Correct Error Pred.		8		8		6		5
Wrong Error Pred.		3		3		4		1

or better. The dichotomous variable for cycles showed a significance of 10.6% in the Wald test which not as good as the frequently used 5% significance level. The three complexity metrics all had a very poor Wald value with a significance between 70.8% to 78.1%. Accordingly, the null hypothesis that they have no impact on the odds of an error cannot be rejected. So based on the univariate logit models we can conclude that the various metrics related to the size of the model seem to be the best predictors for errors.

In a second step we tested multivariate logit models combining all input variables; Table 4 summarizes the results. We started with all 15 variables yielding the results given in the "Complete Model" column. Together they are able to predict 95.2% correctly. Note that Table 4 shows that the number of OR-joins is significant (Wald sig. is 0.3%) and has a considerable impact (Exp(B) is 2.209). As SC and JC were both estimated to be 1 (having no impact on the odds), we reduced the model to 13 variables. The result is given in column "Without SC and JC". The other two columns list the model with the maximum number of variables that all have Wald sig. better than 11% ("8-Step Model") and better than 5% ("5-Step Model"), respectively. The columns show that the estimated coefficients have a stable tendency and a relatively stable value. All Hosmer&Lemeshow and Nagelkerke R^2 values indicate good fit. The 8-Step model yields a prediction of 0.143 for our "Certificate Creation" EPC from the running example. This is below the 0.5 cut-off value and leads to an incorrect prediction of the model having no errors. The model with the highest prediction value (0.945) is a large EPC with 122 arcs, 24 connectors, 40 events, and 43 functions. This model includes an error which is correctly predicted.

The different multivariate logit models suggest the following conclusions. First, the *complexity metrics* proposed by [2] seem to have no impact on the odds of an error at all. The Wald test has both a bad significance and also predicts coefficients very close to zero. An explanation could be that OR connectors get a weight that depends exponentially on the connector cardinality. Consider the example of an AND-split-join block with 5 parallel threads. Both SC and JC would result in a complexity metric of 1. Changing the connector types from AND to OR changes both metrics to 32. This great change in the metric based on state complexity obviously does not reflect the perceived conceptual complexity by the modeler. As the modeler is the one who introduces errors, these metrics seem to be misleading when used for the prediction of errors. Furthermore, the fact that a model includes *cycles* is not significant in the Wald statistic. Moreover, the number of *arcs* does not seem to have a huge impact on the odds, maybe because size is also captured by the number of other model elements and complexity by the number of connectors. The number of *start events* has a coefficient that reduces the odds. This might be related to the way how start events are used in the SAP reference models. There are several EPC models with lots of start events that are directly joined for representing alternative start triggers. This leads to a very simplistic join structure that is unlikely to produce errors. The coefficients for number of *functions* is not significantly different from zero with a tendency to a negative impact on the error probability. In contrast to that, both the number of *end and internal events* increase error probability, but not very strong. Furthermore, it is interesting to see that all join *connectors* tend to have a positive impact on the odds of an error. The OR join has the highest coefficient of about 2. On the other hand, all split connectors have a negative impact. Interestingly, each pair of connectors has coefficients that have almost the same impact, but in a different direction. As an example, consider the coefficients for OR connectors of the 8-Step model. Introducing a pair of OR join and split connectors would have an impact on the odds of $0.473 * 2.233 = 1.056$. Finally, the very small constant of about 0.025 indicates that the probability of an error is very small. This coefficient might be higher if our evaluation design was able to detect more errors in the SAP reference model.

Beyond the significance of each individual coefficient, multivariate logistic regression appears to be a suitable tool to predict error probability in the SAP reference model. Based on only 5 coefficients we are able to classify 95% of the EPCs correctly with a Nagelkerke R^2 of above 0.25. Accordingly, complexity seems to be a major source of error probability, yet not in shape of complexity metrics but rather related to the number of join connectors in the EPC.

4 Related Research

This section discusses the work that is most related for the research areas verification (Section 4.1), execution of informal models (Section 4.2), and quantitative analysis in process modeling (Section 4.3).

4.1 Verification

Since the mid-nineties, a lot of work has been done on the verification of process models, and in particular workflow models. In 1996, Sadiq and Orlowska [23] were among the first to point out that modeling a business process (or workflow) can lead to problems like livelock and deadlock. In their paper, they present a way to overcome syntactical errors, but they ignore the semantical errors. Nowadays, most work that is conducted is focusing on semantical issues, i.e., "will the process specified always terminate" and similar questions. The work on verification that has been conducted in the last decade can roughly be put into three categories.

Verification of formal models, i.e. verification in the mathematical sense. The model with formal executable semantics is correct or not.

Verification of informal models, i.e. defining subclasses of informal models that are mapped onto formal models. Again, the model is correct or not.

Verification by design, i.e. the modeling language does not allow for syntactical errors. Examples are block structured models.

These three categories were presented before in detail in [8], where the authors give relevant literature for each of them.

Besides the three categories, there are some verification approaches that are more or less a combination of others. Consider for example the approach presented in [9], where EPCs are verified using a more or less formal verification approach. However instead of generating a subclass of EPCs for which the approach works, the process designer or process owner is made involved in the verification process by using his knowledge about the process, which is not made explicit in the model. The latter is the reason why this approach could not be used for the automatic verification of the entire SAP reference model, since we are not process owners.

The approach we use in this paper, i.e. the WofYAWL approach, has been introduced in [28]. Again, this approach is somewhat of a by-stander. The approach takes a model with a formal semantics (i.e. a YAWL model), but it isn't complete. The approach cannot decide whether the process is completely correct. It can however find errors in the YAWL model that should be corrected. By translating EPCs to YAWL models, we could use this approach.

4.2 Execution of informal models

It is interesting to note that verification is strongly related to the efficient execution of models. Especially the approaches presented in the previous paragraph, all rely on executable semantics of the process model under consideration. As an example, we mention YAWL models. YAWL models use an OR-join of which the intuitive idea is taken from EPCs. To obtain executable semantics for YAWL models, YAWL models are mapped onto reset nets to decide whether an OR-join is enabled or not in [29]. In the context of EPCs the possibility to provide executable semantics has been investigated in [16], where executable semantics are proven to exist for a large sub-class of all EPCs. In [3] an approach is presented to efficiently calculate the state space of an EPC, thereby providing executable semantics for the EPC. The authors mainly motivate this work from the viewpoint of simulation/execution although their approach can also be used for verification purposes. Because of the semantical problems in some EPCs [16] the algorithm does not always provide a result. Moreover, the authors also point out the need for "chain elimination" to reduce the state space of large models.

4.3 Quantitative Research on Process Modeling

In contrast to the rich set of work on formal aspects of process modeling, only little research has been dedicated to quantitative aspects. In [24] the understandability of join and split representation in EPCs is compared to Petri nets from a modeler perspective. According to this study, users seem to understand the EPC notation easier. A recent survey reported in [5] identifies the most popular conceptual modeling languages and tools in Australia. Furthermore, the authors identify a set of motivations why modeling is used in practice and summarize prior quantitative work on observed advantages and disadvantages of modeling. Beyond that, we are not aware of quantitative research that aims at identifying determinants for errors in process models. There has been some research on complexity metrics for process models motivated by the idea that complexity would increase probability of errors [2].

To summarize this overview of related work, we point out that this paper uniquely combines formal error identification with quantitative analysis of potential error determinants. This way, we have been able to provide a lower bound of 5.6% for the percentage of errors in the SAP reference model.

5 Contributions & Limitations

In this paper, we proposed an approach to automatically identify errors in the SAP reference model. This formal analysis builds on a mapping from EPCs to YAWL and on the utilization of the WofYAWL tool, and is one of the few studies using formal methods for quantitative research. We provided an in-depth analysis of errors in the SAP reference model which yields a lower bound for the number of errors (5.6% of the 604 non-trivial EPCs). As far as we know, this is the first systematic analysis of the EPCs in the SAP reference model.

Our findings demonstrate the need for formal analysis of process models in practice. Moreover, we used a multivariate logistic regression model to test whether certain model characteristics can serve as error determinants. Beyond the significance of each individual coefficient we can conclude that multivariate logistic regression appears to be a suitable tool to predict error probability in the SAP reference model. Based on only 5 coefficients we were able to classify 95% of the EPCs correctly with a Nagelkerke R^2 of above 0.25. Therefore, complexity seems to be a major source of error probability, yet not in shape of complexity metrics defined in [2] but rather related to the number of joins in the EPC.

Yet, our approach still has several limitations. It is a shortcoming for the estimation of a logit model that WofYAWL does not find all errors in the EPCs. Future research will have to investigate how those potential determinants that are not significant in the test perform in the context of other models. Better results could be possible if WofYAWL would be biased to detect only certain categories of errors, but others not. Therefore, we need further research on automatic identification of errors. Furthermore, we aim to reuse this research design for other large enterprise models in order to test whether the coefficients are stable. A systematic analysis of more large enterprise models could result in a theory explaining when human modelers are likely to introduce errors in a process model. Such a theory would offer valuable insights for the teaching of process modeling languages in companies and universities making people aware of situations where errors occur more frequently.

References

1. W.M.P. van der Aalst and A.H.M. ter Hofstede. YAWL: Yet Another Workflow Language. *Information Systems*, 30(4):245–275, 2005.
2. J. Cardoso. Control-flow Complexity Measurement of Processes and Weyuker's Properties. In *6th International Enformatika Conference*, Transactions on Enformatika, Systems Sciences and Engineering, Vol. 8, pages 213–218, 2005.
3. N. Cuntz, J. Freiheit, and E. Kindler. On the Semantics of EPCs: Faster Calculation for EPCs with Small State Spaces. In M. Nüttgens and F.J. Rump, editors, *Proceedings of Fourth Workshop on Event-Driven Process Chains (EPK 2005)*, pages 7–23, Hamburg, Germany, Dec. 2005. Gesellschaft für Informatik, Bonn.
4. T. Curran and G. Keller A. Ladd. *SAP R/3 Business Blueprint: Understanding the Business Process Reference Model.* Enterprise Resource Planning Series. Prentice Hall PTR, Upper Saddle River, 1997.
5. Islay Davies, Peter Green, Michael Rosemann, Marta Indulska, and Stan Gallo. How do practitioners use conceptual modeling in practice? *Data & Knowledge Engineering*, In Press, 2006.
6. J. Dehnert and P. Rittgen. Relaxed Soundness of Business Processes. In K.R. Dittrich, A. Geppert, and M.C. Norrie, editors, *Proceedings of CAiSE 2001*, volume 2068 of *LNCS*, pages 157–170. Springer-Verlag, Berlin, 2001.
7. J. Desel and J. Esparza. *Free Choice Petri Nets*, volume 40 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge Univ. Press, Cambridge, UK, 1995.
8. B.F. van Dongen and M.H. Jansen-Vullers. EPC Verification in the ARIS for MySAP reference model database. BETA Working Paper Series, WP 142, Eindhoven University of Technology, Eindhoven, 2005.
9. B.F. van Dongen, H.M.W. Verbeek, and W.M.P. van der Aalst. Verification of EPCs: Using reduction rules and Petri nets. In *Conference on Advanced Information Systems Engineering (CAiSE 2005)*, volume 3520 of *LNCS*, pages 372–386. Springer-Verlag, Berlin, 2005.
10. J. Esparza and M. Silva. Circuits, Handles, Bridges and Nets. In G. Rozenberg, editor, *Advances in Petri Nets 1990*, volume 483 of *LNCS*, pages 210–242. Springer-Verlag, Berlin, 1990.
11. P. Fettke and P. Loos. Classification of reference models - a methodology and its application. *Information Systems and e-Business Management*, 1(1):35–53, 2003.
12. J. F. Hair, jr., R. E. Anderson, R. L. Tatham, and W. C. Black. *Multivariate Data Analysis.* Prentice-Hall International, Inc., 5th edition edition, 1998.

13. G. Keller, M. Nüttgens, and A.W. Scheer. Semantische Processmodellierung auf der Grundlage Ereignisgesteuerter Processketten (EPK). Veröffentlichungen des Instituts für Wirtschaftsinformatik, Heft 89 (in German), University of Saarland, Saarbrücken, 1992.
14. G. Keller and T. Teufel. *SAP(R) R/3 Process Oriented Implementation: Iterative Process Prototyping*. Addison-Wesley, 1998.
15. E. Kindler. On the Semantics of EPCs: A Framework for Resolving the Vicious Circle. In J. Desel, B. Pernici, and M. Weske, editors, *International Conference on Business Process Management (BPM 2004)*, volume 3080 of *LNCS*, pages 82–97. Springer-Verlag, Berlin, 2004.
16. E. Kindler. On the Semantics of EPCs: Resolving the Vicious Circle. *Data and Knowledge Engineering*, 56(1):23–40, 2006.
17. K. R. Lang and M. Schmidt. Workflow-supported organizational memory systems: An industrial application. In *Proceedings of HICSS-35 2002*. IEEE Computer Society, 2002.
18. J. Mendling, M. Moser, and G. Neumann. Transformation of yEPC Business Process Models to YAWL. In *21st Annual ACM Symposium on Applied Computing*, Dijon, France, 2006. ACM.
19. Jan Mendling, Gustaf Neumann, and Markus Nüttgens. Yet Another Event-Driven Process Chain. In *Proceedings of BPM 2005*, LNCS 3649, pages 428–433, 2005.
20. T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, April 1989.
21. W. Reisig and G. Rozenberg, editors. *Lectures on Petri Nets I: Basic Models*, volume 1491 of *LNCS*. Springer-Verlag, Berlin, 1998.
22. Michael Rosemann and Wil van der Aalst. A Configurable Reference Modelling Language. *Information Systems*, In Press, 2006.
23. W. Sadiq and M.E. Orlowska. Modeling and verification of workflow graphs. Technical Report No. 386, Department of Computer Science, The University of Queensland, Australia, 1996.
24. K. Sarshar and P. Loos. Comparing the control-flow of epc and petri net from the end-user perspective. In W. v. d. Aalst, B. Benatallah, F. Casati, and F. Curbera, editors, *Business Process Management*, LNCS 3649, pages 434–439, 2005.
25. H. A. Simon. *Sciences of the Artificial*. The MIT Press, 3rd edition, 1996.
26. O. Thomas and A.-W. Scheer. Tool support for the collaborative design of reference models - a business engineering perspective. In *Proceedings of HICSS-39 2006*. IEEE Computer Society, 2006.
27. H.M.W. Verbeek and W.M.P. van der Aalst. On the verification of EPCs using T-invariants. BPM Center Report BPM-06-05, BPMcenter.org, 2006. http://www.BPMcenter.org/reports/2006/BPM-06-05.pdf.
28. H.M.W. Verbeek, W.M.P. van der Aalst, and A.H.M. ter Hofstede. Verifying workflows with cancellation regions and OR-joins: An approach based on invariants. BETA Working Paper Series, WP 156, Eindhoven University of Technology, Eindhoven, The Netherlands, 2006.
29. M.T. Wynn, D. Edmond, W.M.P. van der Aalst, and A.H.M. ter Hofstede. Achieving a General, Formal and Decidable Approach to the OR-join in Workflow using Reset nets. In G. Ciardo and P. Darondeau, editors, *Applications and Theory of Petri Nets 2005*, volume 3536 of *LNCS*, pages 423–443. Springer-Verlag, Berlin, 2005.

A.1.1 Univariate Logit Model for Number of Start Events

This section gives the results of a univariate logit model with **Number of Start Events** as the single input variable. The Wald test with a significance of 0.1% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 0	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-2,819	,177	255,030	1	,000	,060

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	NoofStartEvents	12,970	1	,000
	Overall Statistics		12,970	1	,000

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	252,229[a]	,016	,044

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6,343	5	,274

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofStartEvents	,108	,032	11,279	1	,001	1,114
	Constant	-3,334	,257	168,231	1	,000	,036

a. Variable(s) entered on step 1: NoofStartEvents.

A.1.2 Univariate Logit Model for Number of End Events

This section gives the results of a univariate logit model with **Number of End Events** as the single input variable. The Wald test with a significance of 0.5% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	255,229[a]	,011	,030

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	25,315	6	,000

Classification Table[a]

Observed			Predicted		
			hasError		Percentage Correct
			0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofEndEvents	,072	,026	7,781	1	,005	1,074
	Constant	-3,207	,243	174,708	1	,000	,040

a. Variable(s) entered on step 1: NoofEndEvents.

A.1.3 Univariate Logit Model for Number of Intermediate Events

This section gives the results of a univariate logit model with **Number of Intermediate Events** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	216,892[a]	,072	,203

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	28,251	6	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	567	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofIntermediateEvents	,194	,030	41,120	1	,000	1,214
	Constant	-3,804	,281	182,693	1	,000	,022

a. Variable(s) entered on step 1: NoofIntermediateEvents.

A.1.4 Univariate Logit Model for Number of Functions

This section gives the results of a univariate logit model with **Number of Functions** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	236,336[a]	,041	,117

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	13,559	5	,019

Classification Table[a]

Observed			Predicted		
			hasError		Percentage Correct
			0	1	
Step 1	hasError	0	569	1	99,8
		1	32	2	5,9
	Overall Percentage				94,5

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofFunctions	,170	,039	18,480	1	,000	1,185
	Constant	-3,692	,294	157,384	1	,000	,025

a. Variable(s) entered on step 1: NoofFunctions.

A.1.5 Univariate Logit Model for Number of AND-Joins

This section gives the results of a univariate logit model with **Number of AND-Joins** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	242,151[a]	,032	,091

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	10,529	2	,005

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	33	1	2,9
	Overall Percentage				94,5

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofANDjoins	,355	,074	22,814	1	,000	1,427
	Constant	-3,394	,244	193,448	1	,000	,034

a. Variable(s) entered on step 1: NoofANDjoins.

A.1.6 Univariate Logit Model for Number of AND-Splits

This section gives the results of a univariate logit model with **Number of AND-Splits** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	241,174[a]	,033	,095

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	16,564	2	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	568	2	99,6
		1	34	0	,0
	Overall Percentage				94,0

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofANDsplits	,397	,083	22,996	1	,000	1,487
	Constant	-3,443	,251	188,440	1	,000	,032

a. Variable(s) entered on step 1: NoofANDsplits.

A.1.7 Univariate Logit Model for Number of XOR-Joins

This section gives the results of a univariate logit model with **Number of XOR-Joins** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	236,840[a]	,040	,115

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4,478	2	,107

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofXORjoins	,433	,082	28,197	1	,000	1,542
	Constant	-3,504	,256	187,029	1	,000	,030

a. Variable(s) entered on step 1: NoofXORjoins.

A.1.8 Univariate Logit Model for Number of XOR-Splits

This section gives the results of a univariate logit model with **Number of XOR-Splits** as the single input variable. The Wald test with a significance of 0.6% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	255,357[a]	,010	,030

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	28,998	2	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofXORsplits	,220	,080	7,617	1	,006	1,246
	Constant	-3,083	,214	207,113	1	,000	,046

a. Variable(s) entered on step 1: NoofXORsplits.

A.1.9 Univariate Logit Model for Number of OR-Joins

This section gives the results of a univariate logit model with **Number of OR-Joins** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	239,654[a]	,036	,102

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4,499	1	,034

Classification Table[a]

	Observed		Predicted		
			hasError		Percentage Correct
			0	1	
Step 1	hasError	0	568	2	99,6
		1	33	1	2,9
	Overall Percentage				94,2

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofORjoins	,525	,111	22,232	1	,000	1,691
	Constant	-3,232	,218	219,684	1	,000	,039

a. Variable(s) entered on step 1: NoofORjoins.

A.1.10 Univariate Logit Model for Number of OR-Splits

This section gives the results of a univariate logit model with **Number of OR-Splits** as the single input variable. The Wald test with a significance of 0.1% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	252,975[a]	,014	,041

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4,309	1	,038

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofORsplits	,354	,109	10,496	1	,001	1,425
	Constant	-3,126	,219	204,117	1	,000	,044

a. Variable(s) entered on step 1: NoofORsplits.

A.1.11 Univariate Logit Model for Number of Arcs

This section gives the results of a univariate logit model with **Number of Arcs** as the single input variable. The Wald test with a significance of 0.0% indicates that the null hypothesis of the coefficient being zero is rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	226,548[a]	,057	,161

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	20,679	7	,004

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	568	2	99,6
		1	31	3	8,8
	Overall Percentage				94,5

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofArcs	,035	,006	37,166	1	,000	1,036
	Constant	-3,851	,288	178,970	1	,000	,021

a. Variable(s) entered on step 1: NoofArcs.

A.1.12 Univariate Logit Model for hasCycle

This section gives the results of a univariate logit model with **hasCycle** as the single input variable. The Wald test with a significance of 10.7% indicates that the null hypothesis of the coefficient being zero cannot be rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	259,543[a]	,004	,010

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	,000	0	.

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	hasCycle	,911	,565	2,597	1	,107	2,487
	Constant	-2,892	,188	237,764	1	,000	,055

a. Variable(s) entered on step 1: hasCycle.

A.1.13 Univariate Logit Model for CFC-Split

This section gives the results of a univariate logit model with **CFC-Split** as the single input variable. The Wald test with a significance of 78.1% indicates that the null hypothesis of the coefficient being zero cannot be rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	261,364[a]	,001	,002

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	54,081	6	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	CFCsplit	,000	,000	,077	1	,781	1,000
	Constant	-2,813	,177	253,418	1	,000	,060

a. Variable(s) entered on step 1: CFCsplit.

A.1.14 Univariate Logit Model for CFC-Join

This section gives the results of a univariate logit model with **CFC-Join** as the single input variable. The Wald test with a significance of 75.6% indicates that the null hypothesis of the coefficient being zero cannot be rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	604	100,0
	Missing Cases	0	,0
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	261,441[a]	,000	,001

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	55,769	7	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	570	0	100,0
		1	34	0	,0
	Overall Percentage				94,4

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	CFCjoin	,000	,000	,096	1	,756	1,000
	Constant	-2,813	,177	253,444	1	,000	,060

a. Variable(s) entered on step 1: CFCjoin.

A.1.15 Univariate Logit Model for Split-Join-Ratio

This section gives the results of a univariate logit model with **Split-Join-Ratio** as the single input variable. The Wald test with a significance of 70.8% indicates that the null hypothesis of the coefficient being zero cannot be rejected.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	601	99,5
	Missing Cases	3	,5
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 1: Method = Enter

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	261,232[a]	,000	,001

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	45,765	5	,000

Classification Table[a]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 1	hasError	0	567	0	100,0
		1	34	0	,0
	Overall Percentage				94,3

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	CFCquot	,010	,028	,140	1	,708	1,010
	Constant	-2,829	,182	241,913	1	,000	,059

a. Variable(s) entered on step 1: CFCquot.

A.2.1 Multivariate Logit Model including all 15 Input Variables

This section gives the results of a multivariate logit model with including **all 15 input variables**. The Hosmer & Lemeshow test has good significance from up step 6 (higher than 5%). Nagelkerke's R Square ranges from 0.204 to 0.326. The early inclusion of the CFC-Split variable leads to unsatisfactory Wald significance of the coefficient. As CFC-Split and CFC-Join are estimated to have no impact on the odds of an error, they are excluded resulting in a 13 input variable logit model (A.2.2).

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	601	99,5
	Missing Cases	3	,5
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 0	hasError	0	567	0	100,0
		1	34	0	,0
	Overall Percentage				94,3

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-2,814	,177	254,001	1	,000	,060

Variables not in the Equation[a]

			Score	df	Sig.
Step 0	Variables	CFCquot	,144	1	,704
		NoofStartEvents	13,070	1	,000
		NoofEndEvents	9,255	1	,002
		NoofIntermediateEvents	75,427	1	,000
		NoofFunctions	46,915	1	,000
		NoofANDjoins	28,632	1	,000
		NoofANDsplits	30,309	1	,000
		NoofXORjoins	36,291	1	,000
		NoofXORsplits	9,012	1	,003
		NoofORjoins	39,954	1	,000
		NoofORsplits	11,900	1	,001
		NoofArcs	57,713	1	,000
		CFCsplit	,097	1	,755
		CFCjoin	,144	1	,704
		hasCycle	2,734	1	,098

a. Residual Chi-Squares are not computed because of redundancies.

Block 1: Method = Forward Stepwise (Likelihood Ratio)

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	216,457[a]	,072	,204
2	213,292[a]	,077	,218
3	208,708[a]	,084	,238
4	204,742[a]	,090	,255
5	203,098[a]	,092	,262
6	199,694[a]	,098	,276
7	197,069[a]	,101	,288
8	190,349[b]	,111	,316
9	189,648[b]	,112	,319
10	189,312[b]	,113	,320
11	188,828[b]	,114	,322
12	188,661[b]	,114	,323
13	187,992[b]	,115	,326
14	187,986[b]	,115	,326
15	187,986[b]	,115	,326

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

b. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	27,948	6	,000
2	30,731	6	,000
3	23,116	8	,003
4	17,282	8	,027
5	19,677	8	,012
6	7,473	8	,487
7	10,669	8	,221
8	10,314	8	,244
9	9,063	8	,337
10	8,838	8	,356
11	6,560	8	,585
12	7,706	8	,463
13	11,792	8	,161
14	13,217	8	,105
15	13,261	8	,103

Classification Table[a]

Observed			Predicted hasError 0	Predicted hasError 1	Percentage Correct
Step 1	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 2	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 3	hasError	0	565	2	99,6
		1	30	4	11,8
	Overall Percentage				94,7
Step 4	hasError	0	565	2	99,6
		1	30	4	11,8
	Overall Percentage				94,7
Step 5	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 6	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 7	hasError	0	564	3	99,5
		1	29	5	14,7
	Overall Percentage				94,7
Step 8	hasError	0	564	3	99,5
		1	28	6	17,6
	Overall Percentage				94,8
Step 9	hasError	0	564	3	99,5
		1	28	6	17,6
	Overall Percentage				94,8
Step 10	hasError	0	564	3	99,5
		1	27	7	20,6
	Overall Percentage				95,0
Step 11	hasError	0	563	4	99,3
		1	26	8	23,5
	Overall Percentage				95,0
Step 12	hasError	0	563	4	99,3
		1	27	7	20,6
	Overall Percentage				94,8
Step 13	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2
Step 14	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2
Step 15	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofIntermediateEvents	,194	,030	41,152	1	,000	1,214
	Constant	-3,799	,281	182,354	1	,000	,022
Step 2[b]	NoofIntermediateEvents	,167	,033	25,203	1	,000	1,182
	NoofORjoins	,241	,126	3,658	1	,056	1,272
	Constant	-3,851	,286	181,424	1	,000	,021
Step 3[c]	NoofIntermediateEvents	,180	,034	27,533	1	,000	1,198
	NoofORjoins	,288	,124	5,398	1	,020	1,334
	CFCsplit	,000	,000	,227	1	,634	1,000
	Constant	-3,938	,297	176,023	1	,000	,019
Step 4[d]	NoofStartEvents	-,116	,062	3,570	1	,059	,890
	NoofIntermediateEvents	,205	,038	28,483	1	,000	1,228
	NoofORjoins	,509	,173	8,622	1	,003	1,664
	CFCsplit	,000	,000	,356	1	,551	1,000
	Constant	-3,726	,315	139,874	1	,000	,024
Step 5[e]	NoofStartEvents	-,169	,075	5,024	1	,025	,845
	NoofIntermediateEvents	,162	,050	10,612	1	,001	1,176
	NoofXORjoins	,196	,151	1,693	1	,193	1,217
	NoofORjoins	,603	,191	9,999	1	,002	1,828
	CFCsplit	,000	,000	,349	1	,555	1,000
	Constant	-3,665	,312	137,550	1	,000	,026
Step 6[f]	NoofStartEvents	-,189	,078	5,871	1	,015	,828
	NoofIntermediateEvents	,168	,052	10,400	1	,001	1,183
	NoofXORjoins	,365	,179	4,147	1	,042	1,441
	NoofORjoins	,702	,202	12,090	1	,001	2,019
	NoofORsplits	-,336	,186	3,258	1	,071	,714
	CFCsplit	,000	,000	,258	1	,611	1,000
	Constant	-3,665	,321	130,703	1	,000	,026
Step 7[g]	NoofStartEvents	-,182	,078	5,441	1	,020	,833
	NoofIntermediateEvents	,180	,053	11,354	1	,001	1,197
	NoofXORjoins	,513	,202	6,419	1	,011	1,670
	NoofXORsplits	-,259	,171	2,291	1	,130	,772
	NoofORjoins	,646	,208	9,655	1	,002	1,908
	NoofORsplits	-,295	,192	2,357	1	,125	,744
	CFCsplit	,000	,000	,351	1	,553	1,000
	Constant	-3,650	,324	126,994	1	,000	,026
Step 8[h]	NoofStartEvents	-,284	,086	10,870	1	,001	,753
	NoofEndEvents	,177	,067	7,069	1	,008	1,194
	NoofIntermediateEvents	,201	,051	15,784	1	,000	1,223
	NoofXORjoins	,599	,195	9,384	1	,002	1,820
	NoofXORsplits	-,567	,195	8,411	1	,004	,567
	NoofORjoins	,827	,215	14,804	1	,000	2,286
	NoofORsplits	-,602	,231	6,816	1	,009	,548
	CFCsplit	,000	,000	,346	1	,557	1,000
	Constant	-3,944	,350	126,893	1	,000	,019
Step 9[i]	NoofStartEvents	-,338	,109	9,614	1	,002	,713
	NoofEndEvents	,171	,067	6,397	1	,011	1,186
	NoofIntermediateEvents	,182	,055	10,859	1	,001	1,200
	NoofANDjoins	,155	,186	,701	1	,402	1,168
	NoofXORjoins	,626	,199	9,908	1	,002	1,870
	NoofXORsplits	-,547	,197	7,720	1	,005	,578
	NoofORjoins	,863	,220	15,396	1	,000	2,370
	NoofORsplits	-,643	,238	7,314	1	,007	,525
	CFCsplit	,000	,000	,291	1	,590	1,000
	Constant	-3,862	,358	116,094	1	,000	,021

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 10[j]	CFCquot	,030	,043	,507	1	,477	1,031
	NoofStartEvents	-,351	,110	10,190	1	,001	,704
	NoofEndEvents	,175	,068	6,671	1	,010	1,191
	NoofIntermediateEvents	,183	,055	11,016	1	,001	1,201
	NoofANDjoins	,165	,185	,796	1	,372	1,180
	NoofXORjoins	,621	,198	9,843	1	,002	1,861
	NoofXORsplits	-,546	,196	7,721	1	,005	,579
	NoofORjoins	,868	,219	15,740	1	,000	2,383
	NoofORsplits	-,635	,237	7,159	1	,007	,530
	CFCsplit	,000	,000	,296	1	,586	1,000
	Constant	-3,901	,366	113,693	1	,000	,020
Step 11[k]	CFCquot	,032	,043	,569	1	,451	1,033
	NoofStartEvents	-,370	,115	10,424	1	,001	,691
	NoofEndEvents	,214	,089	5,838	1	,016	1,238
	NoofIntermediateEvents	,194	,058	11,240	1	,001	1,214
	NoofANDjoins	,213	,200	1,132	1	,287	1,237
	NoofANDsplits	-,143	,207	,475	1	,490	,867
	NoofXORjoins	,659	,206	10,231	1	,001	1,934
	NoofXORsplits	-,573	,203	7,968	1	,005	,564
	NoofORjoins	,896	,225	15,816	1	,000	2,449
	NoofORsplits	-,708	,261	7,351	1	,007	,492
	CFCsplit	,000	,000	,305	1	,581	1,000
	Constant	-3,910	,367	113,619	1	,000	,020
Step 12[l]	CFCquot	,034	,043	,614	1	,433	1,034
	NoofStartEvents	-,376	,117	10,301	1	,001	,687
	NoofEndEvents	,207	,091	5,197	1	,023	1,230
	NoofIntermediateEvents	,210	,071	8,659	1	,003	1,234
	NoofFunctions	-,025	,060	,169	1	,681	,976
	NoofANDjoins	,220	,201	1,195	1	,274	1,246
	NoofANDsplits	-,119	,216	,300	1	,584	,888
	NoofXORjoins	,656	,208	9,927	1	,002	1,926
	NoofXORsplits	-,583	,207	7,955	1	,005	,558
	NoofORjoins	,904	,227	15,786	1	,000	2,469
	NoofORsplits	-,716	,262	7,470	1	,006	,489
	CFCsplit	,000	,000	,300	1	,584	1,000
	Constant	-3,836	,406	89,340	1	,000	,022
Step 13[m]	CFCquot	,032	,042	,565	1	,452	1,032
	NoofStartEvents	-,444	,143	9,687	1	,002	,641
	NoofEndEvents	,140	,120	1,358	1	,244	1,150
	NoofIntermediateEvents	,068	,177	,146	1	,702	1,070
	NoofFunctions	-,098	,109	,811	1	,368	,906
	NoofANDjoins	,062	,275	,051	1	,821	1,064
	NoofANDsplits	-,241	,262	,852	1	,356	,785
	NoofXORjoins	,531	,256	4,302	1	,038	1,701
	NoofXORsplits	-,704	,254	7,655	1	,006	,495
	NoofORjoins	,792	,263	9,066	1	,003	2,208
	NoofORsplits	-,838	,303	7,666	1	,006	,433
	NoofArcs	,093	,108	,736	1	,391	1,097
	CFCsplit	,000	,000	,285	1	,593	1,000
	Constant	-3,764	,414	82,521	1	,000	,023

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 14[n]	CFCquot	,031	,042	,557	1	,455	1,032
	NoofStartEvents	-,445	,143	9,661	1	,002	,641
	NoofEndEvents	,140	,120	1,363	1	,243	1,151
	NoofIntermediateEvents	,067	,177	,142	1	,706	1,069
	NoofFunctions	-,099	,110	,815	1	,367	,906
	NoofANDjoins	,063	,276	,052	1	,819	1,065
	NoofANDsplits	-,241	,262	,851	1	,356	,786
	NoofXORjoins	,533	,257	4,297	1	,038	1,704
	NoofXORsplits	-,707	,259	7,449	1	,006	,493
	NoofORjoins	,792	,263	9,063	1	,003	2,208
	NoofORsplits	-,840	,304	7,620	1	,006	,432
	NoofArcs	,093	,108	,743	1	,389	1,098
	CFCsplit	,000	,000	,286	1	,593	1,000
	hasCycle	-,051	,687	,005	1	,941	,951
	Constant	-3,762	,415	82,062	1	,000	,023
Step 15[o]	CFCquot	,031	,042	,556	1	,456	1,032
	NoofStartEvents	-,445	,143	9,647	1	,002	,641
	NoofEndEvents	,140	,120	1,362	1	,243	1,151
	NoofIntermediateEvents	,067	,177	,143	1	,706	1,069
	NoofFunctions	-,099	,110	,810	1	,368	,906
	NoofANDjoins	,063	,276	,053	1	,818	1,065
	NoofANDsplits	-,241	,262	,850	1	,357	,786
	NoofXORjoins	,533	,258	4,285	1	,038	1,705
	NoofXORsplits	-,707	,259	7,449	1	,006	,493
	NoofORjoins	,792	,263	9,043	1	,003	2,209
	NoofORsplits	-,840	,304	7,620	1	,006	,432
	NoofArcs	,093	,108	,739	1	,390	1,097
	CFCsplit	,000	,000	,286	1	,593	1,000
	CFCjoin	,000	,001	,001	1	,972	1,000
	hasCycle	-,051	,687	,005	1	,941	,951
	Constant	-3,763	,415	82,046	1	,000	,023

a. Variable(s) entered on step 1: NoofIntermediateEvents.

b. Variable(s) entered on step 2: NoofORjoins.

c. Variable(s) entered on step 3: CFCsplit.

d. Variable(s) entered on step 4: NoofStartEvents.

e. Variable(s) entered on step 5: NoofXORjoins.

f. Variable(s) entered on step 6: NoofORsplits.

g. Variable(s) entered on step 7: NoofXORsplits.

h. Variable(s) entered on step 8: NoofEndEvents.

i. Variable(s) entered on step 9: NoofANDjoins.

j. Variable(s) entered on step 10: CFCquot.

k. Variable(s) entered on step 11: NoofANDsplits.

l. Variable(s) entered on step 12: NoofFunctions.

m. Variable(s) entered on step 13: NoofArcs.

n. Variable(s) entered on step 14: hasCycle.

o. Variable(s) entered on step 15: CFCjoin.

Variables not in the Equation[a]

			Score	df	Sig.
Step 1	Variables	CFCquot	,218	1	,641
		NoofStartEvents	,006	1	,937
		NoofEndEvents	1,075	1	,300

Variables not in the Equation[a]

			Score	df	Sig.
Step 1	Variables	NoofFunctions	,074	1	,786
		NoofANDjoins	,071	1	,790
		NoofANDsplits	,146	1	,702
		NoofXORjoins	,223	1	,637
		NoofXORsplits	1,985	1	,159
		NoofORjoins	3,788	1	,052
		NoofORsplits	1,721	1	,190
		NoofArcs	,050	1	,822
		CFCsplit	2,138	1	,144
		CFCjoin	,082	1	,774
		hasCycle	,223	1	,637
Step 2	Variables	CFCquot	,024	1	,877
		NoofStartEvents	1,106	1	,293
		NoofEndEvents	1,048	1	,306
		NoofFunctions	,032	1	,858
		NoofANDjoins	,032	1	,858
		NoofANDsplits	,252	1	,616
		NoofXORjoins	,412	1	,521
		NoofXORsplits	1,277	1	,258
		NoofORsplits	3,214	1	,073
		NoofArcs	,505	1	,477
		CFCsplit	4,231	1	,040
		CFCjoin	,118	1	,731
		hasCycle	,253	1	,615
Step 3	Variables	CFCquot	,000	1	,994
		NoofStartEvents	3,624	1	,057
		NoofEndEvents	,668	1	,414
		NoofFunctions	,004	1	,949
		NoofANDjoins	1,035	1	,309
		NoofANDsplits	,055	1	,814
		NoofXORjoins	,003	1	,955
		NoofXORsplits	2,516	1	,113
		NoofORsplits	1,759	1	,185
		NoofArcs	1,534	1	,216
		CFCjoin	,124	1	,725
		hasCycle	,110	1	,740
	Overall Statistics		19,532	12	,076
Step 4	Variables	CFCquot	,194	1	,659
		NoofEndEvents	,128	1	,720
		NoofFunctions	,052	1	,819
		NoofANDjoins	,162	1	,687
		NoofANDsplits	,778	1	,378
		NoofXORjoins	1,687	1	,194
		NoofXORsplits	,782	1	,377
		NoofORsplits	,904	1	,342
		NoofArcs	,163	1	,686
		CFCjoin	,037	1	,847
		hasCycle	,348	1	,555
	Overall Statistics		17,456	11	,095

Variables not in the Equation[a]

			Score	df	Sig.
Step 5	Variables	CFCquot	,508	1	,476
		NoofEndEvents	,019	1	,892
		NoofFunctions	,078	1	,780
		NoofANDjoins	,313	1	,576
		NoofANDsplits	,374	1	,541
		NoofXORsplits	3,164	1	,075
		NoofORsplits	3,305	1	,069
		NoofArcs	,052	1	,819
		CFCjoin	,017	1	,895
		hasCycle	,227	1	,634
	Overall Statistics		15,226	10	,124
Step 6	Variables	CFCquot	,082	1	,774
		NoofEndEvents	,715	1	,398
		NoofFunctions	,048	1	,826
		NoofANDjoins	1,329	1	,249
		NoofANDsplits	,539	1	,463
		NoofXORsplits	2,365	1	,124
		NoofArcs	,346	1	,557
		CFCjoin	,012	1	,912
		hasCycle	,260	1	,610
	Overall Statistics		12,975	9	,164
Step 7	Variables	CFCquot	,008	1	,930
		NoofEndEvents	7,438	1	,006
		NoofFunctions	,293	1	,588
		NoofANDjoins	1,292	1	,256
		NoofANDsplits	1,800	1	,180
		NoofArcs	3,339	1	,068
		CFCjoin	,011	1	,916
		hasCycle	,188	1	,665
	Overall Statistics		9,920	8	,271
Step 8	Variables	CFCquot	,380	1	,538
		NoofFunctions	,163	1	,687
		NoofANDjoins	,702	1	,402
		NoofANDsplits	,129	1	,720
		NoofArcs	,087	1	,768
		CFCjoin	,001	1	,975
		hasCycle	,003	1	,956
	Overall Statistics		2,590	7	,920
Step 9	Variables	CFCquot	,570	1	,450
		NoofFunctions	,312	1	,577
		NoofANDsplits	,444	1	,505
		NoofArcs	,063	1	,802
		CFCjoin	,007	1	,932
		hasCycle	,004	1	,947
	Overall Statistics		1,919	6	,927
Step 10	Variables	NoofFunctions	,356	1	,551
		NoofANDsplits	,478	1	,489
		NoofArcs	,086	1	,769
		CFCjoin	,007	1	,933
		hasCycle	,002	1	,966
	Overall Statistics		1,302	5	,935
Step 11	Variables	NoofFunctions	,170	1	,680
		NoofArcs	,014	1	,907
		CFCjoin	,008	1	,927
		hasCycle	,000	1	,993
	Overall Statistics		,850	4	,932

Variables not in the Equation[a]

			Score	df	Sig.
Step 12	Variables	NoofArcs	,713	1	,398
		CFCjoin	,005	1	,945
		hasCycle	,000	1	,984
	Overall Statistics		,720	3	,868
Step 13	Variables	CFCjoin	,000	1	,987
		hasCycle	,005	1	,941
	Overall Statistics		,006	2	,997
Step 14	Variables	CFCjoin	,000	1	,988
	Overall Statistics		,000	1	,988

a. Residual Chi-Squares are not computed because of redundancies.

A.2.2 Multivariate Logit Model including 13 Input Variables without CFC-Join and CFC-Split

This section gives the results of a multivariate logit model with including **13 input variables without CFC-Join and CFC-Split**. The Hosmer & Lemeshow test has good significance apart from step 5 (higher than 5%). Nagelkerke's R Square ranges from 0.204 to 0.304. The 5-Step model is the last one with all coefficients having a Wald significance of lower than 5%. The 8-step model is the last one with all coefficeints having a Wald significance of lower than 11%.

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	601	99,5
	Missing Cases	3	,5
	Total	604	100,0
Unselected Cases		0	,0
Total		604	100,0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			hasError		Percentage Correct
	Observed		0	1	
Step 0	hasError	0	567	0	100,0
		1	34	0	,0
	Overall Percentage				94,3

a. Constant is included in the model.

b. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-2,814	,177	254,001	1	,000	,060

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	CFCquot	,144	1	,704
		NoofStartEvents	13,070	1	,000
		NoofEndEvents	9,255	1	,002
		NoofIntermediateEvents	75,427	1	,000
		NoofFunctions	46,915	1	,000
		NoofANDjoins	28,632	1	,000
		NoofANDsplits	30,309	1	,000
		NoofXORjoins	36,291	1	,000
		NoofXORsplits	9,012	1	,003
		NoofORjoins	39,954	1	,000
		NoofORsplits	11,900	1	,001
		NoofArcs	57,713	1	,000
		hasCycle	2,734	1	,098
	Overall Statistics		109,151	13	,000

Block 1: Method = Forward Stepwise (Likelihood Ratio)

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	216,457[a]	,072	,204
2	213,292[a]	,077	,218
3	209,890[a]	,082	,233
4	207,894[a]	,085	,241
5	202,256[a]	,094	,266
6	199,929[a]	,097	,275
7	197,656[a]	,101	,285
8	194,225[a]	,106	,300
9	193,863[a]	,106	,301
10	193,440[a]	,107	,303
11	193,232[a]	,107	,304
12	193,152[a]	,107	,304
13	193,152[a]	,107	,304

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	27,948	6	,000
2	30,731	6	,000
3	12,593	8	,127
4	12,911	7	,074
5	7,152	8	,520
6	17,972	8	,021
7	6,983	8	,538
8	6,161	8	,629
9	5,146	8	,742
10	4,793	8	,779
11	4,741	8	,785
12	3,560	8	,895
13	3,557	8	,895

Classification Table[a]

Observed			Predicted hasError 0	Predicted hasError 1	Percentage Correct
Step 1	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 2	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 3	hasError	0	564	3	99,5
		1	30	4	11,8
	Overall Percentage				94,5
Step 4	hasError	0	566	1	99,8
		1	30	4	11,8
	Overall Percentage				94,8
Step 5	hasError	0	566	1	99,8
		1	29	5	14,7
	Overall Percentage				95,0
Step 6	hasError	0	565	2	99,6
		1	28	6	17,6
	Overall Percentage				95,0
Step 7	hasError	0	564	3	99,5
		1	29	5	14,7
	Overall Percentage				94,7
Step 8	hasError	0	563	4	99,3
		1	28	6	17,6
	Overall Percentage				94,7
Step 9	hasError	0	563	4	99,3
		1	28	6	17,6
	Overall Percentage				94,7
Step 10	hasError	0	563	4	99,3
		1	27	7	20,6
	Overall Percentage				94,8
Step 11	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2
Step 12	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2
Step 13	hasError	0	564	3	99,5
		1	26	8	23,5
	Overall Percentage				95,2

a. The cut value is ,500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	NoofIntermediateEvents	,194	,030	41,152	1	,000	1,214
	Constant	-3,799	,281	182,354	1	,000	,022
Step 2[b]	NoofIntermediateEvents	,167	,033	25,203	1	,000	1,182
	NoofORjoins	,241	,126	3,658	1	,056	1,272
	Constant	-3,851	,286	181,424	1	,000	,021

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 3[c]	NoofIntermediateEvents	,213	,045	22,957	1	,000	1,238
	NoofORjoins	,302	,128	5,603	1	,018	1,352
	NoofORsplits	-,286	,160	3,176	1	,075	,751
	Constant	-3,880	,295	172,377	1	,000	,021
Step 4[d]	NoofIntermediateEvents	,180	,051	12,517	1	,000	1,197
	NoofXORjoins	,186	,132	1,968	1	,161	1,204
	NoofORjoins	,325	,127	6,522	1	,011	1,384
	NoofORsplits	-,366	,170	4,616	1	,032	,693
	Constant	-3,954	,304	169,052	1	,000	,019
Step 5[e]	NoofStartEvents	-,170	,076	5,079	1	,024	,844
	NoofIntermediateEvents	,151	,050	9,060	1	,003	1,162
	NoofXORjoins	,444	,170	6,807	1	,009	1,559
	NoofORjoins	,662	,200	10,956	1	,001	1,939
	NoofORsplits	-,448	,171	6,840	1	,009	,639
	Constant	-3,670	,320	131,885	1	,000	,025
Step 6[f]	NoofStartEvents	-,276	,106	6,789	1	,009	,759
	NoofIntermediateEvents	,126	,053	5,760	1	,016	1,135
	NoofANDjoins	,260	,171	2,313	1	,128	1,297
	NoofXORjoins	,495	,176	7,922	1	,005	1,640
	NoofORjoins	,743	,209	12,637	1	,000	2,102
	NoofORsplits	-,506	,180	7,893	1	,005	,603
	Constant	-3,573	,323	122,372	1	,000	,028
Step 7[g]	NoofStartEvents	-,274	,107	6,549	1	,010	,760
	NoofIntermediateEvents	,137	,052	6,825	1	,009	1,147
	NoofANDjoins	,270	,174	2,410	1	,121	1,310
	NoofXORjoins	,631	,198	10,110	1	,001	1,879
	NoofXORsplits	-,236	,166	2,029	1	,154	,790
	NoofORjoins	,699	,214	10,712	1	,001	2,011
	NoofORsplits	-,482	,184	6,851	1	,009	,618
	Constant	-3,557	,325	119,870	1	,000	,029
Step 8[h]	NoofStartEvents	-,330	,108	9,440	1	,002	,719
	NoofEndEvents	,121	,064	3,522	1	,061	1,128
	NoofIntermediateEvents	,141	,050	7,795	1	,005	1,151
	NoofANDjoins	,278	,173	2,573	1	,109	1,321
	NoofXORjoins	,698	,193	13,040	1	,000	2,010
	NoofXORsplits	-,424	,185	5,257	1	,022	,654
	NoofORjoins	,804	,213	14,179	1	,000	2,233
	NoofORsplits	-,748	,238	9,857	1	,002	,473
	Constant	-3,741	,342	119,908	1	,000	,024
Step 9[i]	NoofStartEvents	-,342	,112	9,314	1	,002	,711
	NoofEndEvents	,119	,065	3,332	1	,068	1,126
	NoofIntermediateEvents	,165	,067	6,027	1	,014	1,180
	NoofFunctions	-,035	,059	,365	1	,546	,965
	NoofANDjoins	,300	,179	2,815	1	,093	1,350
	NoofXORjoins	,699	,197	12,596	1	,000	2,011
	NoofXORsplits	-,440	,190	5,368	1	,021	,644
	NoofORjoins	,822	,218	14,192	1	,000	2,276
	NoofORsplits	-,772	,242	10,129	1	,001	,462
	Constant	-3,633	,381	90,785	1	,000	,026

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 10[j]	NoofStartEvents	-,378	,125	9,103	1	,003	,685
	NoofEndEvents	,055	,117	,216	1	,642	1,056
	NoofIntermediateEvents	,071	,158	,200	1	,654	1,073
	NoofFunctions	-,090	,104	,748	1	,387	,914
	NoofANDjoins	,170	,266	,408	1	,523	1,185
	NoofXORjoins	,590	,258	5,230	1	,022	1,805
	NoofXORsplits	-,509	,218	5,454	1	,020	,601
	NoofORjoins	,740	,251	8,681	1	,003	2,095
	NoofORsplits	-,806	,249	10,494	1	,001	,446
	NoofArcs	,061	,092	,436	1	,509	1,063
	Constant	-3,570	,393	82,588	1	,000	,028
Step 11[k]	CFCquot	,024	,044	,286	1	,593	1,024
	NoofStartEvents	-,389	,127	9,386	1	,002	,678
	NoofEndEvents	,058	,117	,249	1	,618	1,060
	NoofIntermediateEvents	,073	,158	,214	1	,644	1,076
	NoofFunctions	-,091	,104	,768	1	,381	,913
	NoofANDjoins	,184	,267	,473	1	,491	1,202
	NoofXORjoins	,592	,256	5,335	1	,021	1,807
	NoofXORsplits	-,506	,218	5,393	1	,020	,603
	NoofORjoins	,746	,250	8,915	1	,003	2,108
	NoofORsplits	-,803	,248	10,439	1	,001	,448
	NoofArcs	,060	,092	,421	1	,516	1,062
	Constant	-3,593	,396	82,236	1	,000	,028
Step 12[l]	CFCquot	,023	,044	,272	1	,602	1,023
	NoofStartEvents	-,406	,140	8,463	1	,004	,666
	NoofEndEvents	,055	,116	,229	1	,632	1,057
	NoofIntermediateEvents	,045	,182	,059	1	,807	1,046
	NoofFunctions	-,102	,111	,846	1	,358	,903
	NoofANDjoins	,174	,268	,422	1	,516	1,190
	NoofANDsplits	-,070	,248	,080	1	,778	,932
	NoofXORjoins	,584	,256	5,218	1	,022	1,794
	NoofXORsplits	-,530	,233	5,172	1	,023	,589
	NoofORjoins	,726	,257	7,969	1	,005	2,067
	NoofORsplits	-,853	,307	7,726	1	,005	,426
	NoofArcs	,078	,109	,503	1	,478	1,081
	Constant	-3,585	,397	81,692	1	,000	,028
Step 13[m]	CFCquot	,023	,044	,270	1	,603	1,023
	NoofStartEvents	-,406	,140	8,415	1	,004	,666
	NoofEndEvents	,056	,116	,230	1	,632	1,057
	NoofIntermediateEvents	,044	,183	,059	1	,808	1,045
	NoofFunctions	-,102	,111	,845	1	,358	,903
	NoofANDjoins	,174	,268	,422	1	,516	1,190
	NoofANDsplits	-,070	,248	,079	1	,778	,932
	NoofXORjoins	,585	,257	5,190	1	,023	1,795
	NoofXORsplits	-,530	,236	5,067	1	,024	,589
	NoofORjoins	,726	,257	7,967	1	,005	2,067
	NoofORsplits	-,853	,308	7,667	1	,006	,426
	NoofArcs	,078	,110	,504	1	,478	1,081
	hasCycle	-,010	,673	,000	1	,988	,990
	Constant	-3,585	,398	81,224	1	,000	,028

a. Variable(s) entered on step 1: NoofIntermediateEvents.

b. Variable(s) entered on step 2: NoofORjoins.

c. Variable(s) entered on step 3: NoofORsplits.

d. Variable(s) entered on step 4: NoofXORjoins.

e. Variable(s) entered on step 5: NoofStartEvents.

f. Variable(s) entered on step 6: NoofANDjoins.

Variables in the Equation

g. Variable(s) entered on step 7: NoofXORsplits.

h. Variable(s) entered on step 8: NoofEndEvents.

i. Variable(s) entered on step 9: NoofFunctions.

j. Variable(s) entered on step 10: NoofArcs.

k. Variable(s) entered on step 11: CFCquot.

l. Variable(s) entered on step 12: NoofANDsplits.

m. Variable(s) entered on step 13: hasCycle.

Variables not in the Equation

			Score	df	Sig.
Step 1	Variables	CFCquot	,218	1	,641
		NoofStartEvents	,006	1	,937
		NoofEndEvents	1,075	1	,300
		NoofFunctions	,074	1	,786
		NoofANDjoins	,071	1	,790
		NoofANDsplits	,146	1	,702
		NoofXORjoins	,223	1	,637
		NoofXORsplits	1,985	1	,159
		NoofORjoins	3,788	1	,052
		NoofORsplits	1,721	1	,190
		NoofArcs	,050	1	,822
		hasCycle	,223	1	,637
	Overall Statistics		22,862	12	,029
Step 2	Variables	CFCquot	,024	1	,877
		NoofStartEvents	1,106	1	,293
		NoofEndEvents	1,048	1	,306
		NoofFunctions	,032	1	,858
		NoofANDjoins	,032	1	,858
		NoofANDsplits	,252	1	,616
		NoofXORjoins	,412	1	,521
		NoofXORsplits	1,277	1	,258
		NoofORsplits	3,214	1	,073
		NoofArcs	,505	1	,477
		hasCycle	,253	1	,615
	Overall Statistics		17,876	11	,085
Step 3	Variables	CFCquot	,044	1	,835
		NoofStartEvents	,813	1	,367
		NoofEndEvents	,003	1	,954
		NoofFunctions	,505	1	,477
		NoofANDjoins	,015	1	,902
		NoofANDsplits	,629	1	,428
		NoofXORjoins	2,022	1	,155
		NoofXORsplits	,294	1	,588
		NoofArcs	,004	1	,953
		hasCycle	,449	1	,503
	Overall Statistics		15,216	10	,124
Step 4	Variables	CFCquot	,116	1	,734
		NoofStartEvents	4,954	1	,026
		NoofEndEvents	,290	1	,590
		NoofFunctions	,143	1	,706
		NoofANDjoins	,193	1	,660
		NoofANDsplits	,007	1	,933
		NoofXORsplits	2,468	1	,116
		NoofArcs	1,307	1	,253
		hasCycle	,204	1	,652
	Overall Statistics		13,040	9	,161
Step 5	Variables	CFCquot	,030	1	,862
		NoofEndEvents	,189	1	,664
		NoofFunctions	,020	1	,887
		NoofANDjoins	2,323	1	,127
		NoofANDsplits	,933	1	,334
		NoofXORsplits	2,003	1	,157
		NoofArcs	,385	1	,535
		hasCycle	,309	1	,578
	Overall Statistics		9,792	8	,280

Variables not in the Equation

			Score	df	Sig.
Step 6	Variables	CFCquot	,227	1	,634
		NoofEndEvents	,294	1	,588
		NoofFunctions	,242	1	,623
		NoofANDsplits	,202	1	,653
		NoofXORsplits	2,074	1	,150
		NoofArcs	,000	1	,987
		hasCycle	,087	1	,768
	Overall Statistics		7,171	7	,411
Step 7	Variables	CFCquot	,071	1	,790
		NoofEndEvents	3,641	1	,056
		NoofFunctions	,581	1	,446
		NoofANDsplits	1,020	1	,313
		NoofArcs	1,288	1	,256
		hasCycle	,060	1	,807
	Overall Statistics		4,665	6	,587
Step 8	Variables	CFCquot	,261	1	,609
		NoofFunctions	,366	1	,545
		NoofANDsplits	,000	1	,986
		NoofArcs	,011	1	,916
		hasCycle	,000	1	,996
	Overall Statistics		1,150	5	,950
Step 9	Variables	CFCquot	,336	1	,562
		NoofANDsplits	,019	1	,890
		NoofArcs	,434	1	,510
		hasCycle	,000	1	,996
	Overall Statistics		,875	4	,928
Step 10	Variables	CFCquot	,309	1	,578
		NoofANDsplits	,088	1	,766
		hasCycle	,002	1	,966
	Overall Statistics		,392	3	,942
Step 11	Variables	NoofANDsplits	,080	1	,778
		hasCycle	,000	1	,983
	Overall Statistics		,080	2	,961
Step 12	Variables	hasCycle	,000	1	,988
	Overall Statistics		,000	1	,988