

Integrated Modeling Environments in Organizations: An Empirical Study

Gordon P. Wright • Alok R. Chaturvedi • Radha V. Mookerjee • Susan Garrod

Krannert Graduate School of Management, Purdue University, West Lafayette, Indiana 47907

gordy@mgmt.purdue.edu

Krannert Graduate School of Management, Purdue University, West Lafayette, Indiana 47907

alok@mgmt.purdue.edu

Information Systems Architecture, Boeing Commercial Group, Seattle, Washington 98124

radha.v.mookerjee@boeing.com

School of Technology, Purdue University, West Lafayette, Indiana 47907

sgarrod@purdue.edu

Considerable attention in the information systems and management science literature has focused on computer-based modeling environments, sometimes called integrated modeling environments or model management systems. This research has been primarily concerned with suggesting features/components of modeling environments such as improved executable modeling languages for model creation, integration, and data representation; specialized database systems for managing model data; and customized model-solver software. However, there has been little (if any) empirical guidance offered in the literature about the specific needs of business and industry for computer-based integrated modeling environments. Using a data set compiled from a national survey of modelers (analysts) and model users (decision makers), we empirically investigate the validity of several of the key assumptions of modeling environment research reported in the literature, and examine the relationships between the modeling factors: *data complexity*, *model complexity*, *modeling intensity*, *modeler/user requirements*, and *need for computer-based integrated modeling environments* in organizations.

Our empirical analysis of the data set shows that practitioners rank automated access to model data and automated error checking (e.g., model syntax and semantics checking) high as desirable components in modeling environments. We find that users prefer to have modeling environments linked to their current modeling and modeling-support software systems. Our findings further suggest that a high percentage of modelers and users are dissatisfied with the software systems they are currently using to support their modeling activities. Finally, a covariance structure analysis of the modeling environment factors clearly shows that: (a) model complexity has a direct positive effect on modeling intensity; (b) data complexity has an insignificant direct effect on modeling intensity, but has a negative effect on modeler/user requirements; and (c) modeler/user requirements have a direct positive effect on need for computer-based integrated modeling environments in organizations.

(Model Management Systems; Integrated Modeling Environments; Structured Modeling; Decision Support Systems)

1. Introduction

An active area in information systems research is the development of computer-based modeling environments, sometimes called *integrated modeling environments* or *model management systems* (e.g., see Geoffrion 1989a, 1991, and 1994; Blanning 1993; and Wright et al. 1997). By an integrated modeling environment we mean a software system with facilities for model creation and data representation, and the capability for interfacing with database systems and model-solving tools.

A key factor motivating most of the published research on modeling environments is the assumption that models, like data, are an important organizational resource and need to be managed well.¹ Further, much of the published research can be characterized as proposing components of modeling environments such as: improved executable modeling languages for model creation and data representation; specialized database management systems for managing model data; and customized model-solver software. However, *the literature offers little empirical guidance as to the needs in business and industry for integrated modeling environments.*

Using a data set compiled from a national survey of modelers and users, this study focuses on empirically describing several of the key assumptions of modeling environment research reported in the literature. In addition, we attempt to empirically describe and compare the important factors and trends that are relevant to organizations evaluating their need for (and use of) modeling environments, as suggested by the data, to those features of modeling environments suggested in the literature. This study also examines the relationships between the modeling constructs: *data complexity, model complexity, modeling intensity, modeler/user requirements, and the need for integrated modeling environments in organizations.*

¹For the purpose of this research, the term "model" refers to a representation (idealized) that uses variables and equations or constraints to describe a business problem/system. Examples include statistical (e.g., econometric, time series, decision analysis, simulation), optimization/mathematical programming, spreadsheet, and network (e.g., scheduling, distribution, production planning, PERT) representations of business problems that can be manipulated using solver software systems.

Research Assumptions—Integrated Modeling Environments

Most of the published research on modeling environments is theoretical in content and based on assumptions RA1-RA4 as follows.

RA1: MODEL USAGE. *Models are widely used by firms for the planning and control of business functions; usage is increasing due to improved user-friendly hardware, modeling languages, and solver software* (Will 1975; Dolk and Konsynski 1984; Applegate et al. 1985, 1986; Geoffrion 1987, 1989a, 1989b, 1994).

RA2: MODEL COMPLEXITY. *Models used in organizations differ widely in complexity, application domain, and modeling paradigm* (Elam et al. 1980; Geoffrion 1987; Chang et al. 1993; Lenard 1993).

RA3: MODELING ACTIVITIES. *There exists considerable interaction of models in organizations; as a result, there exists a need for formal methods to manage models from different modeling paradigms and application domains* (Dolk and Konsynski 1984, 1985; Bradley and Clemence 1988; Geoffrion 1989b; Krishnan 1993; Dolk and Kottermann 1993, 1994).

RA4: MODEL-DATA ACTIVITIES. *Organizations need capable and versatile database management systems to support their modeling efforts* (Dolk and Konsynski 1984, 1985; Wright et al. 1997; Kang et al. 1997).

Desirable Features of Modeling Environments

The following is a comprehensive list of the desired features of computer-based modeling environments suggested in the literature.

DF1. *Have facilities for model formulation, model representation, model verification, data management, model integration, model selection, and model solution* (Dolk and Konsynski 1984; Shaw et al. 1988; Ramirez et al. 1993; Geoffrion 1987, 1991; Neustadter et al. 1992; Krishnan 1993).

DF2. *Serve as a buffer between models and data, enabling users from many functional areas to use a single modeling paradigm (to perform task-specific analysis) and a common database* (Dolk 1988; Geoffrion, 1987, 1989; Neustadter et al. 1992).

DF3. *Support the entire life-cycle of the model (formulation, documentation, software development, evaluation,*

validation, training, installation, implementation and maintenance) (Will 1975; Dolk 1988; Fourier et al. 1990, Gass 1983, 1987; Ramirez et al. 1993; Geoffrion 1991).

DF4. *Be hospitable to users of models as well as modeling professionals* (Gass 1983, 1987; Geoffrion 1987, 1989a, 1991).

DF5. *Support a single easily understood language for multiple users of the system* (Geoffrion 1987, 1989c, 1992a, 1992b; Krishnan 1993; Wright et al. 1997).

Examples of computer-based modeling systems for optimization that have some of the features given previously are AMPL (Fourier et al. 1990), GAMS (Brooke et al. 1988), SAS®-OR (1989) and LINGO (Cunningham and Schrage 1991). Some of the features can also be found in SAS® (1989) and SPSS® (Norusis 1993), two popular computer-based systems for data analysis and econometric modeling, and SIMSCRIPT® (1983), a well-known simulation system.

An excellent survey of the model management literature is given in Blanning (1993). Two notable prototype implementations of modeling environments that possess most of the preceding features are FW/SM developed at UCLA (Geoffrion 1991) and OR/SM developed at Purdue (Wright et al. 1997, Kang et al. 1997). FW/SM and OR/SM are built on Framework III® (FW) (1991) and ORACLE® Tools and Database (OR) (ORACLE 1989), respectively. Both use Structured Modeling (SM) (Geoffrion 1987) as their conceptual framework and the Structured Modeling Language (Geoffrion 1992a and 1992b) as their executable modeling language. Briefly stated, the advantages of Structured Modeling as a conceptual framework for integrated modeling environments are: (1) applicability across several problem domains; (2) an executable, functional, and mathematically rigorous language for model and data representation; (3) the ability to manage different versions and families of models; and (4) a natural capability for interfacing with database systems and model-solving tools. The work of Geoffrion and his colleagues (e.g., see Geoffrion 1987, 1989c, 1991, 1992a, 1992b; Neustadter et al. 1992) on FW/SM strongly influenced the design and implementation of OR/SM.

Since this study is believed to be the first attempt to

gather and analyze empirical data from practitioners regarding their use of integrated modeling environments, we chose to investigate the data set from an exploratory perspective. This was done by analyzing several summary statistics on modeling and model-related activities in organizations, and raising questions (hypotheses) and testing relationships that are related to several of the key assumptions of the modeling environment research reported in the literature. Therefore, this paper proceeds as follows. The next section describes the data set collected from a national survey of modeling professionals and users of models. This is followed by a somewhat detailed empirical description of several modeling-related activity variables that are directly related to the research assumptions RA1-RA4 given previously. We then empirically describe several of the desired features of modeling environments suggested by the data. This is followed by more exploratory analysis where we use covariance structure analysis (Jöreskog and Sörborn 1986) to examine the relationships between the modeling constructs: data complexity, model complexity, modeling intensity, modeler/user requirements, and the need for modeling environments (labeled "modeling environment-need" in this study) in organizations. We conclude with some brief comments of what we believe are the contributions of this study to the literature.

2. Data

Data Collection

Our empirical analysis uses a data set obtained from a national survey of modeling practitioners in business and industry. Coded surveys were mailed to modelers and users of models in several areas of organizations including: market research, financial planning, production and operations planning, strategic management, human resource management, and management information systems. Sent along with each questionnaire was a cover letter from the researchers explaining the coding systems, the purpose of the study, and the confidentiality of the responses. Informants were also offered a brief statement describing the importance of the survey to encourage their participation. Questionnaires were returned directly to the researchers in pre-addressed and postage-paid envelopes to emphasize

the academic control of the information. One follow-up mailing with a duplicate survey and cover letters were sent to nonrespondents; also, one follow-up telephone call was made to randomly selected nonrespondents to encourage them to complete and return the questionnaire.²

Respondents (Modelers and Users) and Key Terms

The Purdue Technical Practitioners' Database was used to select the informants for the questionnaires. This database contains information on several modeling professionals and users of models including: (a) name, address, and job description; (b) type of industry employed; (c) years employed; and (d) education. Questionnaires were mailed to 825 randomly selected members of the database. Questionnaires were returned by 192 of the potential 825 respondents covering 90 different organizations. This gives a response rate of 23% that is above average for such surveys and suggests that the results are probably robust and somewhat generalizable (Dillman 1978). The respondents are a diverse group in terms of position and experience. For example, over 75% of the respondents belong to Fortune 500 companies; approximately 60% spend at least 20% of their time on modeling activities; and 67% have over 5 years of modeling experience. Table 1 gives some additional characteristics of the respondents.

The first page of the questionnaire gives the definition of several key terms and phrases frequently used in the survey, which are reported in Table 2. This is followed by 15 pages containing 71 questions on modeling and modeling-related activities. Since several questions have multiple parts (i.e., variables), the resulting data set has 192 responses on 153 variables. Nearly all the variables are categorical on a five-point scale. A few responses are observations on continuous variables such as percentages, while other responses indicate choices such as types of software systems being used to support modeling activities.

The questions in the survey were: (a) designed using the research assumptions and desirable features of

Table 1 Characteristics of the Respondents—Modelers and Users

	Frequency	Percentage
Industry		
Manufacturing	109	54.8
Finance, Banking, and Insurance	30	15.1
Wholesale and Retail Trade	19	9.5
Transportation, Communication, and Utilities	12	6.0
Other	29	14.6
Position		
Senior Management	30	15.1
Middle Management	83	41.7
Professional/Staff	70	35.1
Other	16	8.1
Modeling and Model-Use Experience		
<5 years	64	32.2
5–10 years	66	33.2
10–15 years	22	11.1
>15 years	47	23.6

Table 2 Definitions of Key Terms/Phrases in the Questionnaire

Model*

A model is a statement of a problem, characterized by a set of inputs, a set of outputs and the relations between them. A model may contain one or more of the following: an objective function, a set of constraints, or a set of assumptions.

Solver

A procedure (e.g., a statistical estimation method or an algorithm) used to execute or solve a model; it is assumed to be independent from the form of model representation.

Model Integration

The process of combining two or more models to form one composite model.

Model-Data Interface

A computer-based link between a model and the data needed for model solution.

Model-Solver Interface

A computer-based link between a model and a procedure needed to solve the model.

Solver Selection

The selection of a solver (assuming more than one solver is available for a particular model) best suited for the problem (i.e., model instance) at hand.

*See footnote 1.

²The design and implementation of the survey followed the generally accepted guidelines (e.g., format, follow-up procedures, etc.) given in Dillman (1978).

modeling environments discussed earlier (in the introduction); and (b) stated in simple, concrete terms using terminology understood and employed by modelers and model users. Some field interviews were used to pretest and improve the wording of questions. Also, exploratory factor analysis and an examination of the question intercorrelations, means, and standard deviations were used to refine the scales. A few questions were not considered because of low variance in responses and the possibility of multiple interpretations. The questions are organized into six distinct sections: Models, Data, Solvers, Software Systems, Model Complexity, and User Characteristics. These sections are briefly described as follows.

Models Section. Seventeen questions were used to assess model usage within and across departments. The questions address the multiple facets of modeling incorporated in our definitions of model, model integration, and model-solver interface, as listed in Table 2 including: (a) time spent creating models, (b) functional areas using models, (c) future use (planned for) versus present use of models, (d) frequency of model usage in direct support of decision making in the firm, (d) interaction of models within and across application domains, and (e) how firms manage models.

Data Section. This section contains eight questions that focus on measures of the complexity of model-data, including data accuracy and data volatility, and how data (e.g., databases) needed to manipulate (e.g., solve) models are managed (e.g., located, stored, retrieved, and reused) in organizations.

Solver Section. Types and satisfaction levels of computer-based solvers were assessed with three questions, including a question asking the respondent to list the reasons, if any, why he/she is dissatisfied with his/her current solver software systems.

Software Systems. Thirty-three questions were used to assess the need for: (a) automated support for model creation, selection, and integration; (b) automated support for solver selection; and (c) model-data interfaces. Several of the questions also were used to assess the necessity for a computer-based integrated modeling environment in the respondent's department.

Model Complexity. The complexity of the models

in each respondent's department or organization was assessed by six questions. For example, each respondent was asked to identify the factors that best indicate (i.e., measure) the complexity of the models being used in their department. Examples of factors measuring model complexity include: (a) number of variables, (b) number of constraints, and (c) expected computing time.

User Characteristics. The last section of the survey gives some descriptive information about each respondent including statistics concerning his/her work experience and level of computer literacy. In addition, each respondent is asked to list the computer hardware most frequently used to support his/her modeling efforts.

Table 3 summarizes the content of the questionnaire in more detail.

3. Some Empirical Results

First, we describe several key summary statistics (Tables 4–7) from our analysis of the practitioners' data set that are directly related to the research assumptions RA1–RA4, as shown in Figure 1. This is followed by a discussion (Table 8) of the need for modeling environments in business, organized around the desired features DF1–DF4 (given in the introduction) of integrated modeling environments suggested in the literature.

RA1: Model Usage—Descriptive Findings

Tables 4A and 4B—Model Hardware and Software. High-end PCs and mainframes are the most common computers being used in organizations to support modeling and modeling-related activities. Workstations and Low-end PCs rank last and are only marginally different from each other in frequency of use. On the average, the data indicates that modelers use three to four different software systems (e.g., solvers, database management systems) to support their modeling efforts.

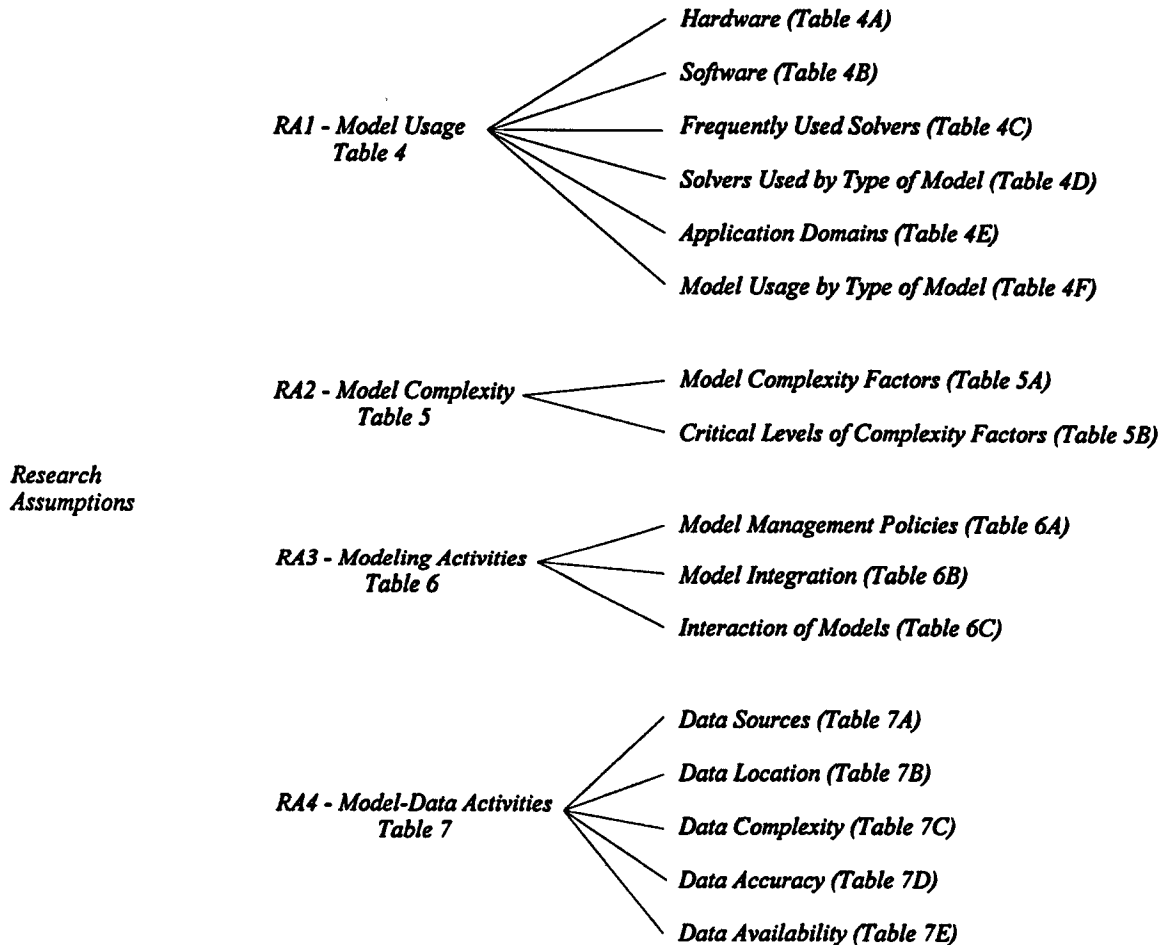
Tables 4C and 4D—Model Solvers. Not surprising, statistical methods and spreadsheet-based procedures/algorithms are the most frequently used solvers in business applications. SAS ranks first as the most frequently used solver, followed by spreadsheets and project management software ranking second and

WRIGHT, CHATURVEDI, MOOKERJEE, AND GARROD
Integrated Modeling Environments in Organizations

Table 3 **Survey Design**

Section	Variables Measured	Related to Assumptions/Features
Models (17 questions)	Time spent on model-related activities, size and variability of model library, and level of computerization	RA1, RA2
	Application domains, types of problems solved using models, and frequency of use of different model types	RA1, RA2/DF1
	Model interactions: within and across departments by type of model	RA1-RA4/DF1
	How/why models are stored	DF1
	Demand for models to support decision making: within and across departments	RA1-RA3/DF1
Data (8 questions)	Source of data	RA4/DF1, DF2, DF4
	Data location—centralized or decentralize	RA3, RA4/DF1, DF2, DF4
	Data availability—hard copy, automated	RA4/DF1, DF2, DF4
	Data volatility—(ranking)	RA4/DF1, DF2, DF4
	Data accuracy—(ranking)	RA4/DF1, DF2, DF4
	Ease of data measurement	RA4/DF1, DF2, DF4
	Ease of acquiring data	RA4/DF1, DF2, DF4
	Proportion of models using real-time data	RA3, RA4/DF1, DF2
Solvers (3 questions)	Names of solvers being used for each of eleven different model types	RA1, RA2/DF1, DF2, DF4
	Satisfaction levels for the five most frequently used solvers	RA1/DF1, DF2, DF4
	Reasons for dissatisfaction with current solvers	RA1-RA2/DF1, DF2, DF4
Software Systems (33 questions)	Number of software systems for model-related activities	RA1, RA2/DF1
	Names of software systems for model-related activities	RA1, RA2/DF1, DF4
	Need for model formulation languages/software	RA1-RA3/DF1, DF3, DF5
		RA1, RA2/DF1, DF3, DF4, DF5
	Satisfaction levels—model formulation software	RA1, RA2/DF1, DF3, DF4
	Need for model selection software	RA2, RA3/DF1
	Similarity of variables across models	RA2, RA3/DF1, DF3, DF4
	Need for model integration software	RA1, RA2/DF1, DF3, DF4
	Need for solver selection software	RA3, RA4/DF1-DF5
	Need for automated model-data interfaces	RA1-RA4/DF1-DF5
	Need for computer-based modeling environment systems	RA1-RA4/DF1-DF5
	Desired characteristics of computer-based modeling environments	RA1-RA4/DF1-DF5
Model Complexity (6 questions)	Indicators of model complexity	RA1, RA3/DF1, DF3-DF5
	Measuring model complexity	RA1, RA3/DF1, DF3-DF5
	Proportion of complex models in use	RA1, RA3/DF1, DF3-DF5
	Adequacy of current software to manage complex models	RA1, RA4/DF1, DF3-CF5
User Characteristics (4 questions)	Years of modeling experience	RA1, RA2
	Job title	RA1
	Computer literacy	RA1, RA2
	Hardware systems being used	

Figure 1 Research Assumptions and Summary Statistics



third, respectively. The data further suggest that simulation and optimization solvers (e.g., MPS and SIMSCRIPT) rank the lowest in frequency of use. A total of 150 different systems were named by the respondents as frequently used solvers; however, only nine were mentioned as frequently used solvers by five or more respondents, which probably suggests that many commercial solver software products are either similar in terms of purpose or performance, or they are highly specialized and meet the requirements of only a small number of modelers.

Tables 4E and 4F—Application Domains and Model Usage. As shown in Table 4E, manufacturing, finance, strategy, and accounting are the functional ar-

reas where models are used most frequently in organizations. Human resource management ranks noticeably the lowest in model usage. Statistical models (e.g., regression, forecasting) are the most frequently used types of models in organizations, followed by scheduling and project management models. The least used models are mathematical programming models.

Other Statistics on Model Usage. The data indicate that the average number of models used across departments within an organization is approximately equal to twenty. Further, the average number of models across organizations is estimated to be about 200. The data also indicate that models are used to support approximately 50% of departmental decisions. Sixty

WRIGHT, CHATURVEDI, MOOKERJEE, AND GARROD
Integrated Modeling Environments in Organizations

Table 4 Model Usage

(A) Hardware Used to Support Modeling Environments

Type	Frequency	Percentage
High-End PCs	167	87.0
Mainframes	115	59.9
Workstations	48	25.0
Low-End PCs	40	20.8

(B) Number of Software Systems Used to Support Modeling Environments

Range	Frequency	Percentage
0-5	159	82.8
6-10	13	6.8
11-15	11	5.8
>15	9	4.7

(C) Most Frequently Used Solvers

Solver Software	Frequency	Percentage
SAS	116	60.4
Spreadsheets	78	40.6
Project Management	23	12.0
SIMAN	15	7.8
MPS	15	7.8
SPSS	10	5.2
LINDO	9	4.7
SIMSCRIPT	7	3.7
SLAM	6	3.1

(D) Solver Software Used by Model Type

Solver Software	Model Type
SAS	Statistical, Forecasting, Math, Programming
Spreadsheets	Statistical, Forecasting, Project Management
Project Management	Forecasting, Scheduling, Project Management
SIMAN	Simulation
MPS	Scheduling, Distribution, Math, Programming
SPSS	Statistical, Forecasting
LINDO	Distribution, Scheduling, Math Programming
SIMSCRIPT	Simulation
SLAM	Simulation

(Table 4 continued)

(E) Application Domains* in Rank Order

Domain/Area	Frequency	Percentage
Manufacturing	139	72.4
Finance	137	71.4
Strategic Mgmt.	119	62.0
Marketing	112	58.3
Accounting	94	49.0
Human Resources	57	29.7

*Areas where models are being used.

(F) Model Usage by Model Type*

Type	Frequency	Percentage
Forecasting	158	82.3
Simulation	151	78.6
Statistical	134	69.8
Scheduling	109	56.8
Project Management	109	56.8
Decision Analysis	99	51.6
Math. Programming	84	43.8

*Each individual response is a rank ordering of model types based on frequency of use.

percent of the practitioners indicated that they use models primarily for solving unstructured problems. Finally, less than 5% indicated that the modeling software they use is developed in-house. The later supports the claim that most modeling software systems are: (a) purchased from software vendors; and (b) standard across organizations.

RA2: Model Complexity—Descriptive Findings

Tables 5A and 5B—Model Complexity. It is perhaps not surprising that the key factors that modelers and users suggested as contributors to model complexity are the number of variables, equations, and constraints. However, it is surprising to learn of the "very low" critical levels for the complexity factors measured, as listed in Table 5B. Models with an average of 26 equations, 21 constraints, or between 13 and 20 variables were at the critical level of complexity as reported by the practitioners. Table 5B also shows that there is considerable variability in the critical levels of

these factors. It is interesting to note that computer time (i.e., model execution time) ranks the lowest as a contributing factor to model complexity. Finally, the data indicates (not shown in Table 5) that over 55% of the practitioners are dissatisfied with the capabilities of the software systems they use to support their modeling activities.

RA3: Modeling Activities—Descriptive Findings

Table 6A, 6B, and 6C—Model Integration and Interactions. Models are most likely to be managed at the departmental level rather than centrally in an organization, and surprisingly, the results strongly suggest that model integration (i.e., linking two or more existing models) is a frequently occurring task (88%) in organizations. As expected, model integration is the highest among models of the same type. The respondents were asked to indicate the degree of interaction between models, within and across departments, with special emphasis on the sharing of data between models. As expected, within department interactions (e.g., model sharing) are the most frequent and interactions across departments are the least frequent.

Table 5 Model Complexity

(A) Factors That Contribute to (Measure) Model Complexity

Variable	Frequency	Percentage
Number of equations	162	84.4
Number of constraints	128	66.7
Number of input variables	125	65.1
Number of assumptions	110	57.3
Number of output variables	81	42.2
Execution time (minutes)	47	24.5

(B) Critical Levels of Model Complexity Factors

Variable	Average Critical Level	Standard Deviation
Number of equations	26.0	20.3
Number of constraints	21.1	24.8
Number of input variables	20.3	32.1
Number of assumptions	28.7	19.5
Number of output variables	13.2	12.6
Execution time (minutes)	29.2	34.3

Other Statistics on Modeling Activities. Over 95% of the respondents indicated that they reuse models frequently. Furthermore, 86% indicated that models are located in several model bases in their organization. The survey shows that 82% of the interactions between models are data sharing activities. Other model interactions indicated by the respondents included: (a) using results from one or more *area* (e.g., manufacturing, marketing, finance, etc.) models to develop a strategic planning model for the firm; and (b) using the results from one model to set goals and constraints for another model. Twenty-five percent of the respondents indicated in their qualitative responses that most interactions between models in their organization are not automated but rather are manual operations. Finally, respondents indicated a strong need for automated support of model integration. This finding supports the assumption of a need for computer-based integrated modeling environments in many organizations.

RA4: Data Activities—Descriptive Findings

Tables 7A and 7B—Data Sources and Locations. The results indicate that data are most often collected when needed for model execution (i.e., model solution); whereas 36% of model data are from routine automated transactions. Only a small percentage (17%) of the respondents indicated their model data are centrally located in their organization; and storing data occurs most often at the department level in organizations.

Tables 7C, 7D, and 7E—Data Complexity, Accuracy, and Availability. The respondents were asked to rate (on a scale of 1 to 5) four measures of data complexity. The complexity measures included: (1) volatility, accuracy, and ease of measuring data; and (2) availability (accessibility) of data for use by models. The data availability measure distinguishes between data collected during electronic transactions as compared to data that is manually collected and compiled. Data accuracy and availability ranked the highest, whereas data volatility ranked the lowest as measures of data complexity.³ Further, 25% of the respondents indicated

³The interpretation of data complexity in this study did not refer to the file structure of the data. This is discussed to a limited degree in the previous section.

WRIGHT, CHATURVEDI, MOOKERJEE, AND GARROD
Integrated Modeling Environments in Organizations

Table 6 Model Integration

(A) Model Management Policies

Models are Managed:	Frequency	Percentage
Centrally	25	14.1
Type of model	44	25.0
Department	63	35.8
Individual user	44	25.0

(B) Model Integration

Integration, If Any, Among Models of:	Frequency	Percentage
Same and different types	62	35.0
Usually the same type	56	32.0
Usually different types	37	21.0
No integration	21	12.0

(C) Frequency of Model Interaction

Ranking	Output-Input Model Interactions		Other Interactions	
	Within Departments	Across Departments	Within Departments	Across Departments
1 (= high)	13 (6.8)	8 (4.2)	16 (8.5)	5 (2.6)
2	64 (33.7)	40 (21.1)	39 (20.6)	20 (10.6)
3	71 (37.4)	65 (34.2)	53 (28.0)	54 (28.6)
4	37 (19.5)	65 (34.2)	49 (25.9)	51 (27.0)
5 (= low)	5 (2.6)	12 (6.3)	33 (16.9)	60 (31.2)

that model data are made available to users only on an ad-hoc noncomputerized basis (e.g., manually collected), and 36% indicated that model data are made available using database management software.

Other Statistics—Model-Data Activities. Almost 35% of the respondents indicated that their models, data, and solution procedures are managed at the departmental level. Further, 42% indicated that models are stored separate from their solution procedures. Perhaps an interesting finding is that 35% of the respondents indicated that model users, rather the model developers, are responsible for the design and implementation of model-data interfaces in their organization.

**Desirable Features of Modeling Environments—
Descriptive Findings**

Tables 8A and 8E—Model-Data Interfaces (see features DF1–DF3 in the Introduction). The results clearly show that there is a significant need (i.e., many respondents rank high) for automated access to model data (i.e., databases), and automated error and consistency checking as components of model-data interfaces. This is further supported at the area level, as shown in Table 8E, with the possible exception of error checking in market planning/research departments. It is clear that practitioners rank type conversion as the least desired feature of model-data interfaces.

Tables 8B and 8F—Model Management Systems

(see DF1, DF3–DF5). Model integration and model formulation software systems (e.g., an executable general purpose modeling language) are ranked by users as nearly equivalent desirable features in modeling environments. Further, the results suggest that automated model selection and solver selection are the least desired features of modeling environments. These results are consistent across functional areas in organizations with the possible exception of human resource management.

Tables 8C and 8G—Current Software Systems (DF1–DF3). The results in Table 8C show that modeling environment software currently used in organizations is about equally split between decision support systems and model management systems. Further, as shown in Table 8G, this result is consistent across functional areas, with the exception of human resource management.

Tables 8D and 8H—Usefulness of Integrated Modeling Environments (DF1–DF5). The results strongly suggest that computer-based modeling environments would not be considered useful to modelers and users as separate (nonintegrated) systems. As shown in Table 8H, this is strongly supported in the manufacturing, marketing, and accounting areas, and to a lesser extent in finance, strategic management, and human resource management.

Other Statistics—Modeling Environments. More than 75% of our respondents indicated that they support the use of computer-based integrated modeling environments in their organization, but only if they are linked to an existing modeling support system (such as an existing decision support system). The respondents were asked to rate (on a scale of 1 = not important at all to 5 = very important) the importance of having one modeling system capable of managing all their models. Almost a third of the respondents (30.3%) answered with a score of 1 or 2. Other qualitative responses to this question strongly suggest that although modelers and users believe integrated modeling environments are desirable, they recognize that it is probably difficult to implement a single comprehensive modeling system that meets most of the needs of their organization.

Table 7 Model—Data Activities

(A) Data Sources		
Source	Frequency	Percentage
Collected whenever needed	62	32.5
Computer based transactions	36	18.8
Hard copy	8	4.2
All of the above	85	44.5
(B) Data Locations		
Location	Frequency	Percentage
Centrally	33	17.3
Department	72	37.7
Other	86	45.0
(C) Data Complexity		
Factor	Average Ranking	Standard Deviation
Accuracy	3.6	0.8
Availability	3.3	1.1
Measurability	2.9	1.0
Volatility	2.6	1.5
(D) Data Accuracy		
How accurate are the data?	Frequency	Percentage
Inaccurate	121	63.0
Accurate	21	11.0
Neither	35	18.0
(E) Data Availability		
How are the data made available?	Frequency	Percentage
Manually	65	34.0
Automatically	50	26.0
Both	77	40.0

4. User Requirements, Modeling Intensity, and Data Complexity

The goal of this part of our study is to provide researchers and practitioners more information about

WRIGHT, CHATURVEDI, MOOKERJEE, AND GARROD
Integrated Modeling Environments in Organizations

Table 8 Features/Components of Modeling Environments

(A) Model-Data Interfaces

Desired Features	Average Ranking*	Standard Deviation
Automatic access to model data	4.2	1.1
Error Checking	4.0	1.0
Consistency Checking	3.8	1.1
Type Conversion	3.3	1.2

*(5 = high; 1 = low).

(B) Components of MMS*

Desired Features	Average Ranking**	Standard Deviation
Model Integration	3.1	1.3
Model Formulation	3.0	1.4
Model Selection	2.7	1.5
Solver Selection	2.4	1.3

*MMS = Model Management Systems. ** (5 = high; 1 = low)

(C) Current Software Systems*

System Used	Frequency	Percentage
Decision Support	50	26.0
Model Management	48	25.0
Both of the above	17	19.0
Neither of the above	58	30.0

*Being used to support modeling and modeling-related activities.

(D) Modeling Environments

Response	Frequency	Percentage
Not very useful	84	44.0
Useless	50	26.0
Somewhat useful	35	18.0
Useful	23	12.0

the salient relationships between the five modeling environment constructs: *modeler/user requirements, modeling environment-need, modeling intensity, model complexity, and data complexity*. By a modeler/user

(Table 8 continued)

(E) Model-Data Interfaces

Functional Area	Average Rankings (5 = high; 1 = low)			
	AAMD	EC	CC	TC
Manufacturing	4.2	4.2	3.9	3.6
Marketing	4.6	3.7	4.0	3.5
Accounting	4.1	4.3	3.7	3.3
Finance	4.3	4.2	3.9	3.4
Strategic Mgmt.	3.9	3.8	3.9	3.0
Human Resource Mgmt.	4.1	4.1	3.6	2.9

AAMD = Automatic access to model data; EC = Error Checking; CC = Consistency Checking; and TC = Type Conversion.

(F) Modeling Environments

Functional Area	MI	MF	MS	SS
Manufacturing	3.6	3.7	2.6	2.5
Marketing	3.4	3.2	2.7	2.1
Account	3.0	2.9	2.4	2.5
Finance	3.0	3.1	2.4	2.4
Strategic Mgmt.	3.3	3.1	2.2	2.2
Human Resource Mgmt.	2.7	2.8	2.0	2.5

MI = Model Integration; MF = Model Formulation; MS = Model Selection; and SS = Solver Selection.

(G) Current Software Systems (Modeling Environments)

Functional Area	DSS	MMS	Both	Neither
Manufacturing	28%	25%	21%	26%
Marketing	28	27	17	29
Accounting	27	25	20	28
Finance	27	25	18	30
Strategic Mgmt.	22	26	22	29
Human Resource Mgmt.	14	31	31	24

DSS = Decision Support Systems; MMS = Model Management Systems.

(H) Modeling Environments*

Functional Area	NU	U	SU	U
Manufacturing	40%	22%	10%	28%
Marketing	37	21	15	27
Accounting	43	21	19	18
Finance	44	20	24	12
Strategic Mgmt.	46	27	23	4
Human Resource Mgmt.	45	28	19	8

*As a separate nonintegrated system. NU = Not very useful; U = Useless; SU = Somewhat useful; and U = Useful.

requirement we mean a specific modeling or software system needed by a modeler or user to support his/her modeling efforts such as an executable modeling language, an automated model-data interface, automated support for model creation and integration, or a customized model-solver system. By a modeling environment need we mean the need for a computer-based integrated modeling environment (see Geoffrion 1991 and Wright et al. 1997). Modeling intensity refers to the level of effort spent by modelers and users in their modeling activities. Recall that model complexity refers to the number of "parameters" of a model such as number of variables, equations, or constraints. Finally, data complexity refers to the accuracy, availability, and volatility of the data required by a model.

We seek to answer specific questions about modeling activities in organizations such as:

Modeling Intensity. Do complex models (model complexity) require a significant effort (modeling intensity) from modelers and users?

Do complex and automated model databases (data complexity) require more or less effort (modeling intensity) from modelers and users?

Modeler/User Requirements. What is the relationship between the complexity of a model (model complexity) and the number of different software systems/components (modeler/user requirements) needed or used by modelers and users?

What is the relationship between complex and automated model databases (data complexity) and the number of different software systems/components (modeler/user requirements) needed or used by modelers and users?

Does an increase in modeling effort (modeling intensity) by modelers and users usually cause an increase in the number of different software systems/components (modeler/user requirements) desired by modelers and users?

Modeling Environment Need. Does an increase in the number of complex models (model complexity) usually cause an increase in the need for an integrated modeling environment (modeling environment need) in organizations?

Does an increase in the number of complex automated model databases (data complexity) cause an increase in the need for an integrated modeling environment (modeling environment need) in organizations?

Does an increase in modeling effort (modeling intensity) by modelers and users cause an increase in the need for an integrated modeling environment (modeling environment need) in organizations?

Does an increase in the number of different software systems/components (modeler/user requirements) needed or used by modelers and users cause an increase in the need for an integrated modeling environment (modeling environment need) in organizations?

We now turn to a more detailed discussion of the five constructs including their measurements/indicators.

Modeler/User Requirements and Modeling Environment Need

Modeling often involves an intensive set of activities including model formulation, model integration (i.e., combining two or more models into a single model), model solution, and/or model selection (see DF1; Murphy and Stohr 1986; Liang 1993, 1988b; Bradley and Clemence 1988; and Dolk and Kotterman 1993). As discussed earlier, most of the research in the modeling environment literature focuses on proposing features of modeling environments, such as improved modeling languages or customized model-solver software, without empirical support from modelers and users of models. In addition (and without empirical support), a key assumption in the literature is that usage of models in organizations is increasing due to improved user friendly hardware, modeling languages, and solver software (see research assumption RA1). This discussion leads to the following choices of measurement variables for the Modeler/User Requirements and Modeling Environment Need constructs.

Measuring Modeler/User Requirements and Modeling Environment Need. *Modeler/User Requirements* is measured by five possible respondent ratings (calibrated on a five-point scale) elicited by the question: "Given your needs, how necessary is automated support for each of the following features in a computer-based modeling environment?" The features are: (1) model formulation, (2) model integration, (3) model

selection, (4) model extraction, (5) version management, and (6) model-data interface.

The *Modeling Environment Need* construct is measured by two questions. The first question is: "How useful would a Model Management System be in your department?" The options for this question are: "Very Useful," "Somewhat Useful," "Neither Useful nor Useless," "Not very Useful," and "Useless." The second question is: "What is your opinion regarding the use of models outside a decision support system?" The options for this question (calibrated on a five-point scale) are: "necessary," "somewhat necessary," "neither necessary nor unnecessary," "somewhat unnecessary," or "unnecessary."

Modeling Intensity

The model management/modeling environment literature assumes that the demand for models is increasing in most organizations; and further, this increase is caused by several factors including the widespread availability of user friendly modeling and solver-software systems (e.g., see Dolk and Konsynski 1984, Geoffrion 1991, and Kang et al. 1997). The literature further suggests that the increase in modeling efforts in organizations leads "naturally" to the need for automated support for model formulation and other model-related activities such as, for example, model integration and model-data interfaces (e.g., see Applegate et al. 1985, Geoffrion 1991, and Wright et al. 1997).

Measuring Modeling Intensity from the Data. *Modeling Intensity* is measured by two questions concerning individual and departmental modeling efforts in organizations. The first question is: "What proportion of your time is spent on modeling and using models?" The options for this question are: "0% to 20%," "21% to 40%," "41% to 60%," "61% to 80%," or "81% to 100%." The second question is: "How many different models are used in your department (computerized or otherwise)?" The options for the second question are: "0 to 10," "10 to 20," "20 to 30," "30 to 50," or "greater than 50."

Model Complexity and Data Complexity

Two important assumptions common to most (if not all) of the published model management/modeling

environment research are that: (a) models in organizations differ widely in complexity, application domain, and modeling paradigm (see RA2); and (b) organizations need capable and versatile database management systems to support their modeling efforts (see RA4). The literature further assumes that model and data complexity increases modeler/user requirements and the need for integrated modeling environments (e.g., see Wright et al. 1997, Geoffrion 1991, Dolk and Konsynski 1984, and Elam et al. 1980).

Measuring Model Complexity and Data Complexity from the Data. The *complexity* of a model is measured by asking respondents to give a critical level to each of three attributes of a model such that if a given model exceeds the critical level in one or more of the attributes, then he/she would consider the model to be complex. Using the definition of a model given in Table 2,⁴ the three attributes this study uses to measure model complexity are: "number of variables," "number of equations/constraints," and "number of assumptions." The options for each attribute are: "less than 5," "5-10," "11-20," "21-30," or "greater than 30." We note that over 80% of the respondents selected all of the three attributes as measures/contributors to model complexity.

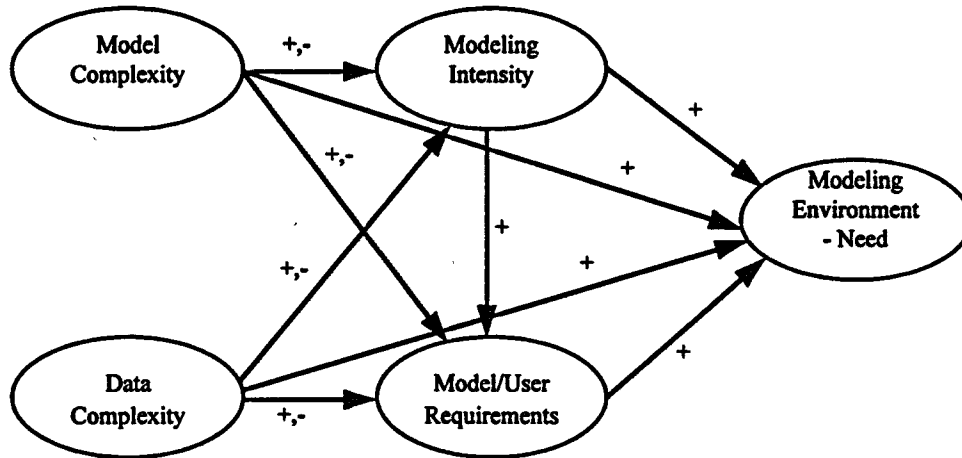
Data Complexity is measured by three questions concerning data accuracy, data volatility, and data measurability. The first question is: "How accurate are the data required by your models?" The options are: "Very accurate," "fairly accurate," "neither accurate or inaccurate," "fairly inaccurate," or "very inaccurate." The second question is: "How volatile are the data required by your models?" The options are: "very volatile," "fairly volatile," "neither volatile, nor stable," "fairly stable," or "very stable." The third question is: "How easy is it to measure the data required by your models?" The options are: "very easy," "fairly easy," "neither very easy, nor very difficult," "fairly difficult," or "very difficult."

4. Hypotheses and Results

We propose the relationships (model) given in Figure 2 as plausible (and important) hypotheses to be tested.

⁴See footnote 1.

Figure 2 Proposed Hypotheses*



*Here "+" and "-" indicate significant positive and negative effects, respectively.

The model in Figure 2 depicts modeling intensity and modeler/user requirements (consequences of model complexity and data complexity) as antecedents of modeling environment-need. The estimated parameters of the structural equations model are given in Table 9. Appendix A contains a discussion of some goodness of fit, reliability, and validity statistics, and a summary of the results from a rigorous simulation analysis of the estimated parameters. The parameters, test statistics, and some of the goodness of fit statistics are estimated using LISREL (Jöreskog and Sörbom, 1986).⁵

It is perhaps appropriate at this point in our discussion to emphasize what we believe are two important aspects of this study which are the data set and our exploratory approach. As discussed earlier, there are no published empirical bases for most of the model management/modeling environment studies reported in the literature. Our analysis is based on a data set compiled from a national survey of modelers and users. Because the data set is somewhat unique and timely, we have chosen to investigate it from an exploratory perspective, giving summary statistics, raising questions and testing hypotheses as the analysis

progressed.⁶ Given our exploratory approach and limited data on some variables, we refrain from adopting a truly causal view of any of the hypothesized relationships between the latent constructs; we merely attempt to demonstrate and interpret "consistent themes" that are present in the data. We now give a brief discussion of the hypotheses and test results implied in Figure 3.

Modeling Intensity

HYPOTHESIS 1A. *Model complexity has a direct effect on modeling intensity.*

HYPOTHESIS 1B. *Data complexity has a direct effect on modeling intensity.*

Hypothesis 1A is strongly supported by the data. This supports what is perhaps obvious and anticipated: that modelers/users tend to spend a significant effort manipulating (e.g., revising, integrating, etc.) complex models (as measured by the number of variables, number of constraints, and the number assumptions). The magnitude of the t-statistic for Hypothesis 1B given in Table 9 is 0.76, which suggests that data

⁵We use the maximum likelihood estimation procedure given in the CALIS procedure, as implemented by SAS™.

⁶In general, this type of inductive approach is common (and endorsed by several researchers) in several areas of management research including the marketing science literature.

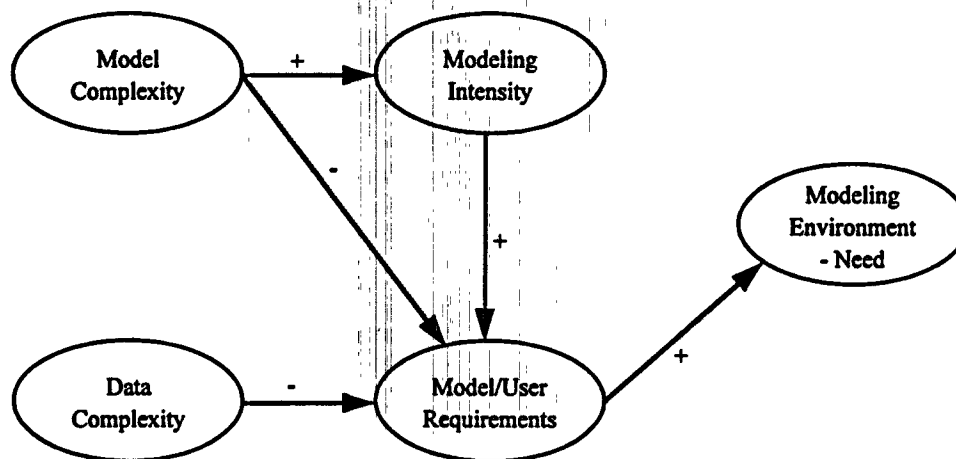
WRIGHT, CHATURVEDI, MOOKERJEE, AND GARROD
Integrated Modeling Environments in Organizations

Table 9 Parameter Estimates and Goodness of Fit Statistics* (T-Statistic in Parentheses)

Path From	To	Original Model	Revised Model
Model Complexity	Modeling Intensity	0.06 (2.80)	0.18 (2.97)
Data Complexity	Modeling Intensity	-0.08 (-0.76)	—
Model Complexity	Modeler/User Requirements	-0.08 (-1.32)	-0.08 (-1.49)
Data Complexity	Modeler/User Requirements	ms0.69 (-1.97)	-0.69 (-1.79)
Model Intensity	Modeler/User Requirements	0.17 (1.33)	0.21 (1.70)
Model Complexity	Modeling Environment-Need	-0.01 (-0.58)	—
Data Complexity	Modeling Environment-Need	-0.03 (-0.50)	—
Model Intensity	Modeling Environment-Need	0.03 (1.16)	—
Modeler/User Requirements	Modeling Environment-Need	0.16 (8.28)	0.21 (9.61)
<i>Goodness-of-Fit Indicators:</i>			
Chi-Square Statistic		141	159
degrees of freedom		100	104
Goodness-of-Fit Index		0.92	0.91
Adjusted Goodness-of-Fit Index		0.87	0.87
Root Mean Square Residuals		0.10	0.10
Bentler's Comparative Fit Index		0.93	0.91

*Here "—" indicates that the corresponding path coefficient is constrained to be equal to zero in the revised structural equations model.

Figure 3 Results*



*Here "+" and "-" indicate significant positive and negative effects, respectively.

complexity has an insignificant effect on modeling intensity. This supports the view that modelers (and sometime users) are separated (e.g., may not be trained on how to use automated database systems) from the important task of obtaining or managing model data.

User Requirements

HYPOTHESIS 2A. *Model complexity has a direct effect on the number of different software systems/components (modeler/user requirements) needed or used by modelers and users.*

HYPOTHESIS 2B. *Data complexity has a direct effect on the number of different software systems/components (modeler/user requirements) needed or used by modelers and users.*

HYPOTHESIS 2C. *Modeling intensity has a direct and positive effect on the number of different software systems/components (modeler/user requirements) needed or used by modelers and users.*

Hypothesis 2A is at best weakly supported by the data as indicated by the magnitude of its t-statistic ($= 1.32$). We see that model complexity has a "weak" negative effect on modeler/user requirements. Perhaps a plausible explanation for this result is that modelers and users usually do not manipulate (e.g., revise and integrate) models that have high complexity ratings (i.e., ratings as measured by the number of variables, number of equations/constraints, and number of assumptions), and therefore they do not need additional or sophisticated modeling software systems for this purpose. In our analysis we split the data set into two sets: one set consisting of observations where one or more of the model complexity variables are, on the average, low, and the other containing observations where one or more of the model complexity variables are high. Next, we performed a correlation analysis between the model complexity variables and the measurement variables of the modeler/user requirements construct. Again, here the approach is exploratory. The two sets of variables (model complexity and model/user requirements) showed a positive correlation in the first data set where model complexity is low

and a negative correlation in the second data set where model complexity is high. These correlations indicate that for only "small" increases in model complexity does modeler/user requirements increase, and for large increases, modeler/user requirements are likely to stay constant or decrease.

Hypotheses 2B and 2C are supported by the data. We see that data complexity has a negative direct effect on modeler/user requirements. This supports the view that modelers (and sometime users) choose to become more separated from the task of managing models that require complex data, and therefore do not need additional or new software systems to support managing of such models. We also note that when modeling intensity increases so does the need for new and specialized software systems. As anticipated, we find that modeling intensity positively affects modeler/user requirements.

Modeling Environment Need

HYPOTHESIS 3A. *Model complexity has a positive effect on the need for a computer-based integrated modeling environment.*

HYPOTHESIS 3B. *Data complexity has a positive effect on the need for a computer-based integrated modeling environment.*

HYPOTHESIS 3C. *Modeling intensity has a positive effect on the need for a computer-based integrated modeling environment.*

HYPOTHESIS 3D. *Modeler/user requirements has a positive effect on the need for a computer-based integrated modeling environment.*

It is interesting that Hypothesis 3D is supported while Hypotheses 3A, 3B, and 3C are completely rejected. These results indicate that modeler/user requirements has a positive direct effect on the need for integrated modeling environments. We also conclude that data and model complexity, and modeling intensity have only an indirect effect, through modeler/user requirements, on the need for integrated modeling environments in organizations.

5. Some Comments and Conclusions

Understanding the practice and needs of modelers and model-users in business is an important issue for designers of computer-based modeling environments. This study attempts to describe some of the specific needs in business and industry for modeling environments using a data set compiled from a national survey of practitioners. As mentioned previously, since this study is believed to be the first attempt to gather and analyze empirical data from practitioners regarding their modeling activities, we investigated the data set from an exploratory perspective. This was done by analyzing several summary statistics on modeling and model-related activities in organizations, and raising questions (hypotheses) and testing relationships about several of the key assumptions of the modeling environment research reported in the literature. Specifically, we examined the relationships between the following factors: model complexity, data complexity, model intensity, model/user requirements, and the need for integrated modeling environments.

The findings of this study largely support the literature that computer-based integrated modeling environments are needed in organizations. What is interesting is that a large proportion (over 70%) of the reporting practitioners indicated that modeling environments should be integrated with their currently existing modeling-support (e.g., decision support) software systems. Other perhaps significant findings were that only a very small percentage (5%) of modeling software is developed "in-house," and a very high percentage of modeling efforts in organizations are directed toward solving highly complex unstructured management problems. An analysis of the comments and opinion section in the questionnaire indicates that many modelers believe they lack the necessary expertise to effectively design or use database management systems for model-based work. Finally, we found that the majority of the reporting practitioners are dissatisfied with their current modeling-support software systems.

Surprisingly, a high proportion of the respondents (88%) indicated that model integration is a frequently

Table 10 Measurement Results

Construct	Number of Questions	Range of Reliabilities	Average Reliability	Composite
Model Complexity	3	0.66 to 0.92	0.80	0.90
Data Complexity	3	0.96 to 0.98	0.98	0.98
Modeling Intensity	3	0.80 to 0.89	0.85	0.89
User Requirements/Needs	6	0.94 to 0.98	0.96	0.99
System Needs	2	0.87 to 0.99	0.93	0.99
	17			

occurring task in their organizations; this strongly supports the view that automated support for model integration is a highly desired feature in their modeling software. In addition, practitioners reported that statistical models are the most frequently types of models used to support decision making, whereas mathematical programming models are the least frequently used in organizations.

The quantitative and qualitative findings, the later from the comments and opinion section of the questionnaire, show that: (a) the availability of microcomputers, executable modeling languages, and general purpose spreadsheets has led to an increase in the use of models to support decision making in many organizations; (b) available sophisticated and user-friendly hardware and software have made it easy for even the novice modeler/user to create and use models; (c) there exists a wide application domain for models (e.g., manufacturing, finance, marketing, and human resource management); (d) there is considerable variability in the types of models and model-solvers being used in firms (e.g., statistical, econometric, simulation, scheduling, decision analysis, mathematical programming); and (e) modelers and users of models believe there is a lack of user-friendly commercial software to support the modeling and solution of unstructured problems.

Using covariance structure analysis we empirically tested several hypotheses, some of which have been advanced in the literature, regarding the relationships between the modeling factors: data complexity, model complexity, modeling intensity, modeler/user requirements, and need for an integrated modeling environment. The analysis clearly suggests that: (a) model

Table 11 Reliability and Significance Tests

		Revised Model (t-statistic in parenthesis)	Sampling Estimates (n = 200 estimates/data sets)	
Path From	To	Parameter Estimate	Average Parameter Estimates	Test Results
Model Complexity	Modeling Intensity	0.18 (2.97)	0.19	accept 0.18 ^a
Model Complexity	Modeler/User Requirements	-0.08 (1.49)	-0.13	accept -0.08 ^b
Data Complexity	Modeler/User Requirements	-0.59 (-1.79)	-0.65	accept -0.59 ^a
Model Intensity	Modeler/User Requirements	0.21 (1.70)	0.16	accept 0.21 ^b
Modeler/User Requirements	Modeling Environment-Need	0.21 (9.61)	0.20	accept 0.21 ^a

^aRisk level equals 0.05.

^bRisk level equals 0.10.

complexity has a direct positive effect on modeling intensity; (b) data complexity has an insignificant effect on modeling intensity; (c) data complexity has a direct negative effect on modeler/user requirements; and (d) only modeler/user requirements directly affects the need for computer-based integrated modeling environments.

We emphasize that this empirical study is best viewed as exploratory (i.e., descriptive and raising and testing hypotheses as the analysis progressed) and is not without its shortcomings. Because of the limited sample size, we were unable to examine the cross-validity of the parameter estimates from the structural equations analysis. However, we did perform a rigorous simulation analysis on the estimated parameters which is reported in Appendix A. Also, given our exploratory objectives we refrained from adopting a causal view of any of the hypothesized relationships between the latent constructs. Instead we attempted to: (a) present and interpret consistent themes that relate to the key assumptions of the modeling environment research reported in the literature; and (b) offer some guidance concerning the needs in business and industry for integrated computer-based modeling environments. An important goal of this paper is to persuade

readers to think about their current research (e.g., validity of their research assumptions) in terms of what is being suggested and reported by professional modelers and users in Tables 4–8. Finally, we hope that the results from the structural equations analysis will be of benefit to practitioners and designers of integrated modeling environments.⁷

Appendix A Reliability and Validity of the LISREL Model

Measurement Model

Before examining the structural model, the measurement model was evaluated by examining both individual item reliability and construct reliability. Table 10 shows that the item reliability of the measures are high ranging from 0.66 to 0.99. The average variance extracted (AVE) should be at least 50% in order to avoid having more variance due to error than valid measurement (Fornell and Larcker 1981). The AVE of all the constructs in the revised model exceeds the acceptable limit of 50%.

⁷Partial support by grants from the NCR Corporation (Cooperative Systems Division), NATO Scientific Affairs Division, and the AT&T Foundation is gratefully acknowledged. However, the views contained in this paper are those of the authors and not of the sponsors. The authors thank the associate editor and referees for many constructive suggestions. We also thank Professor Peg Williams, who helped with the design of the Purdue Technical Database questionnaire.

Overall Model

Although the chi-square statistic reported in Table 9 indicates that the revised model should be rejected, this test has several well documented problems/limitations (e.g., see Mulaik et al. 1989, Marsh et al. 1988, Anderson and Gerbing 1984, Bearden et al. 1982). Other measures of the model's fit reported in Table 9 include the goodness of fit index (Jöreskog and Sörbom 1986), the root mean square residual (Jöreskog and Sörbom 1986), and the comparative fit index (Bentler 1990). The goodness of fit index of 0.91 indicates that the revised model accounts for ninety-one percent of the sum of squares of the sample covariances and provides an acceptable fit (Cuttance 1987). Similarly, the comparative fit index of 0.91 indicates that the revised model provides a substantially better fit to the observed covariance matrix than the null model with no parameters. The root mean square residual shows that the average of the residuals between the fitted and observed covariances in 0.10.

Simulation Analysis of the Parameter Estimates

Our sample size of 170 is well above the suggested minimum of 100 required for estimating the parameters of the structural equations models of Figures 2 and 3; however, it is not large enough to cross-validate the results using a holdout sample. As an alternative, we perform a simulation analysis of the parameter estimates to assess their stability in repeated random sampling. This procedure allows us to estimate the bias and covariance matrix of the estimators (see McCarthy et al. 1992), and detect any parameters whose estimated values are strongly affected by seemingly insignificant variations (see Bard 1975, p. 176-183) in the data set. Using the covariance matrix from the original sample, we randomly generate 200 data sets, each consisting of 170 observations on the 17 measurement variables. Next we obtain the maximum likelihood estimates of the five structural parameters for each of the 200 simulated data sets. These estimates are then used to test the null hypothesis that the estimated parameters from the revised model are the true estimates.⁸ As shown in Table 11, we could not reject the null hypothesis for any of the estimates in the revised model at the 5% level of significance. The results of a correlation analysis performed on the 17 measurement variables further support the conclusion of the test reported in Table 11.

⁸The interested reader can write to the authors for a detailed description of our simulation and test procedures. The test procedure used in this study is similar to the one used by Chandrasekharan et al. (1994). It probably suffices to say that the procedure is simply a sequence of steps involving standard statistical techniques (for generating additional data samples and performing Binomial tests) that are often used in applied econometric modeling for testing the precision of parameter estimates.

References

- Anderson, J. C. and D. W. Gerbing, "The Effect of Sampling Error on Convergence, Improper Solutions, and Goodness-of-Fit Indices for Maximum Likelihood Confirmatory Factor Analysis," *Psychometrika*, 49 (1984), 155-173.
- Applegate, L. M., G. Klein, B. R. Konsynski, and J. F. Nunamaker, "Model Management Systems: Proposed Model Representations and Future Design," *Proc. Sixth International Conf. on Information Systems*, (1985), 1-16.
- , B. R. Konsynski, and J. F. Nunamaker, "Model Management Systems: Design for Decision Support," *Decision Support Systems*, 2 (1986), 81-91.
- Ashton-Tate, *Framework III*, Ashton-Tate, Torrance, CA, 1988.
- Bard, Y., *Nonlinear Parameter Estimation*, Academic Press, New York, 1975.
- Blanning, R. W., "Model Management Systems: An Overview," *Decision Support Systems*, 9 (1993), 9-18.
- Bentler, P. M., "Comparative Fit Indices in Structural Models," *Psychological Bulletin*, 107 (1990), 238-246.
- Bollen, K. A., *Structural Equations with Latent Variables*, Wiley, New York, 1989.
- Bradley, G. H. and R. D. Clemence, "Model Integration with a Typed Executable Modeling Language," *Proc. Twenty-First Annual Hawaii International on System Sciences III: Decision Support and Knowledge Based Systems Track*, (1988), 403-410.
- Brooke, A., D. Kendrick, and A. Meeraus, *GAMS: A User's Guide*, Scientific Press, Redwood City, CA, 1988.
- C.A.C.I., *SIMSCRIPT II.5, Programming Language*, C.A.C.I., Los Angeles, CA, 1983.
- Chandrasekharan, R., M. Moriarity, and G. P. Wright, "Testing for Unreliable Estimators and Insignificant Forecasts in Combined Forecasts," *J. Forecasting*, 13 (1994), 611-624.
- Chang, A. M., C. W. Holsapple, and A. B. Whinston, "Model Management Issues and Directions," *Decision Support Systems*, 9 (1993), 19-37.
- Cunningham, K. and L. Schrage, *LINGO Optimization Modeling Language*, LINDO Systems, Chicago, IL, 1991.
- Cuttance, P., "Issues and Problems in the Application of Covariance Structure Models," in P. Cuttance and R. Ecob (Eds.), *Structural Modeling by Example*, Cambridge University Press, New York, 1987.
- Dillman, D. A., *Mail and Telephone Surveys*, Wiley, New York, 1978.
- Dolk, D. R., "The Role of an Information Resource Dictionary System," *Comm. ACM*, 31, 6 (1988), 704-718.
- and J. Kottermann, "Model Integration and a Theory of Models," *Decision Support Systems*, 9, 1 (1993), 51-63.
- and B. Konsynski, "Model Management in Organizations," *Information and Management*, 9, 1 (1985), 35-47.
- and ———, "Knowledge Representation for Model Management Systems," *IEEE Trans. on Software Engineering*, 6 (1984), 71-91.
- Elam, J. J., J. C. Henderson, and L. W. Miller, "Model Management Systems: An Approach to Decision Support in Complex Organizations," *Proc. First International Conf. on Information Systems*, (1980), 98-110.
- Fornell, C. and D. F. Larcker, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *J. Marketing Res.*, 18, 1, (1981), 39-50.
- Fourier, R., D. M. Gay, and B. W. Kernighan, "A Modeling Language for Mathematical Programming," *Management Sci.*, 36 (1990), 519-554.

- Gass, S. I., "Managing the Modeling Process: A Personal Reflection," *European J. Operational Res.*, 31 (1987), 1-8.
- , "Decision-Aiding Models: Validation, Assessment, and Related Issues for Policy Analysis," *Oper. Res.*, 31, 4 (1983), 603-631.
- Geoffrion, A. M., "An Introduction to Structured Modeling," *Management Sci.*, 33, (1987), 547-588.
- , "Computer-Based Modeling Environments," *European J. Oper. Res.*, 41 (1989a), 33-43.
- , "Integrated Modeling Environments," *Computer Sci. in Economics and Management*, 21 (1989b), 3-15.
- , "The Formal Aspects of Structured Modeling," *Oper. Res.*, 37 (1989b), 30-51.
- , "FW/SM: A Prototype Structured Modeling Environment," *Management Sci.*, 37 (1991), 1513-1538.
- , "The SML Language for Structured Modeling: Levels 1 and 2," *Oper. Res.*, 40 (1992a), 38-57.
- , "The SML Language for Structured Modeling: Levels 3 and 4," *Oper. Res.*, 40 (1992b), 58-75.
- , "Structured Modeling: Survey and Future Research Directions," *ORSA CSTS Newsletter*, 15, (1994), 10-20.
- Jöreskog, K. G., "A General Method for Estimating a Linear Structural Equation System," *Structural Equations in the Social Sciences*, A. S. Goldberger and O. D. Duncan (eds.), Princeton University Press, Princeton, NJ, 1973.
- and Dag Sörbom, "LISREL: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood," *User's Guide*, Version VI, Scientific Software, Mooresville, IN, 1986.
- Kang, M., G. P. Wright, R. Chandrasekharan, R. Mookerjee, and N. D. Worobetz, "The Design and Implementation of OR/SM: A Prototype Integrated Modeling Environment," 1997 *Ann. Oper. Res.* (special issue on the IS/OR interface).
- Kottermann, J. E. and D. R. Dolk, "Model Integration and Modeling Languages: A Process Perspective," *Information Systems Res.*, 3, 1, (1992), 1-16.
- Krishnan, R., "Model Management: Survey, Future Directions and a Bibliography," *ORSA CSTS Newsletter* 14, 1 (1993).
- Lenard, M. L., "An Object Oriented Approach to Model Management," *Decision Support Systems*, 9 (1993), 67-73.
- Liang, T. P., "Integrating Model Management with Data Management in Decision Support Systems," *Decision Support Systems* 1 (1985), 221-232.
- , "Development of a Knowledge-Based Model Management System," *Oper. Res.*, 36, 6 (1988), 849-863.
- Marsh, H. W., J. R. Balla, and R. P. MacDonald, "Goodness of Fit Indices in Confirmatory Factor Analysis: The Effect of Sample Size," *Psychological Bulletin*, 103 (1988), 391-410.
- McCarthy, P. S., P. K. Kannan, R. Chandrasekharan, and G. P. Wright, "Estimating Loyalty and Switching with an Application to the Automobile Market," *Management Sci.*, 38 (1992), 1371-1393.
- Medsker, G. J., L. J. Williams, and P. J. Holahan, "A Review of Current Practices for Evaluating Causal Models in Organizational Behavior and Human Resources Management Research," Working Paper, Krannert Graduate School of Management, Purdue University, West Lafayette, IN, 1990.
- Mulaik, S. A., J. Van Alstine, N. Bennet, and C. D. Stillwell, "An Evaluation of Goodness of Fit Indices for Structural Equation Models," *Psychological Bulletin*, 105 (1989), 430-445.
- Neustadter, L., A. Geoffrion, S. Maturana, Y. Tsai, and F. Vicuna, "The Design and Implementation of a Prototype Structured Modeling Environment," *Ann. Oper. Res.*, 38 (1992), 453-484.
- Norusis, M., *SPSS Base System User's Guide*, Release 6.0, SPSS Inc., Chicago, IL, 1987.
- ORACLE, *ORACLE Version 5.1.c*, ORACLE, Belmont, CA, 1989.
- Ramirez, R., "The ISUMMS Project and Bibliography," ISUMMS Project Informal Note No. 2, Department of Management, Iowa State University, 1993.
- SAS Institute Inc., *SAS/STAT User's Guide*, Version 6, SAS Institute, Cary, NC, 1989.
- SAS Institute Inc., *SAS/OR User's Guide*, Version 6, SAS Institute, Cary, NC, 1989.
- Shaw, M. J., J. P. Tu, and P. De, "Applying Machine Learning to Model Management in Decision Support Systems," *Decision Support Systems*, 4, 3 (1988), 285-305.
- Will, H. J., "Model Management Systems," *Information Systems and Organization Structure*, W. Gruyter, Berlin, 468-482, 1975.
- Wright, G. P., N. D. Worobetz, M. Kang, R. Mookerjee, and R. Chandrasekharan, "OR/SM: A Prototype Integrated Modeling Environment Based On Structured Modeling," *INFORMS J. Computing*, 9, 2 (1997) 134-153.

Michael J. Shaw, Associate Editor. This paper was received on November 17, 1995 and has been with the authors 7 months for 1 revision.