ABSTRACT

Title of Document:	DEVELOPMENT AND VALIDATION OF METHODOLOGY FOR FIX EFFECTIVENESS PROJECTION DURING PRODUCT DEVELOPMENT
	Stephen Mark Brown, Ph.D. Reliability Engineering, 2009
Directed By:	Ph.D., Ali Mosleh, Reliability Engineering

One of the challenges that design and reliability engineers face is how to accurately project fix effectiveness during reliability planning of a product development project. All reliability projection methods currently in use require estimates of the fix effectiveness factors (FEF) in their mathematical formulation. Obviously, required test results from multiple test phases are unavailable at the onset of a project and therefore practice is to rely on engineers' subjective assessment FEFs. Such estimates are often inaccurate and mostly optimist, resulting in potentiality significant project risks in the form of delays, additional development costs, and costs associated with field failures, returns, and market position. This dissertation provides a methodology that significantly improves the accuracy of FEF estimates and also the resulting reliability metrics such as projected failures rates and MTBFs. The methodology identifies key "performance shaping factors" (PSF) that enhances or impedes an engineer's ability to "fix" a problem, and puts that information into a "causal model" via Bayesian Belief Networks (BBN) to predict FEFs. Tests and confirmation of the methodology for various products and diverse industries show a systematic error reduction in FEF estimates over the current use of unstructured subjective estimates. A second major contribution of the research is an

investigation of the effect of interdependencies among various FEFs in projecting the reliability of the same product or several different products by the same organization. Independence is currently assumed by all reliability projection methods. The research (i) shows that FEFs are indeed dependent, (ii) provides a composite BBN model showing the level of dependency among two different fix activities, and (iii) quantifies the impact that fix effectiveness factors have on MTBF projections. The research therefore presents an important augmentation to the current IEC standard for reliability growth, Crow-AMSAA model, showing how to include dependent FEFs in the calculation of failure intensity.

DEVELOPMENT AND VALIDATION OF METHODOLOGY FOR FIX EFFECTIVENESS PROJECTION DURING PRODUCT DEVELOPMENT

By

Stephen Mark Brown

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Ph.D. Reliability
Engineering
2009

Advisory Committee:
Ali Mosleh, Ph.D., Chair
Mohammad Modarres, Ph.D.
Shapour Azarm, Ph.D.
Michel Cukier, Ph.D.
Elise Miller-Hooks, Ph.D.
Paul Ellner, Ph.D.

Preface

Purpose

This research defines the process required to develop a group of subject matter experts and develop generic fix effectiveness performance shaping factors, mapping them within Bayesian Belief Network structures to project fix effectiveness during early stages of product development prior to incurring test results. Gradient Ascent algorithms provide methodology to update or learn subject matter expert conditional probabilities. Methods to shorten model development and eliminate group think are explored. Last, we prove dependency among fix activities and provide an augmentation to the IEC reliability growth model.

Significance

The significance of this research is two-fold. First no structured method of projecting fix effectiveness factors during product development planning exists. Second, no reliability growth projection model accounts for dependencies among fix effectiveness factors. We have defined the process to account for both of these short comings and provided test cases for confirmation the process works. Both of these concepts, structured FEF projection and FEF dependencies have been overlooked since the beginning of reliability growth projection modeling.

Outline of Chapters

Chapter 1: Introduction

Chapter I discusses the motivation for research and the significance fix effectiveness error has on reliability growth projection. This chapter provides an overview of deficiencies current reliability growth projection models have in determining fix effectiveness during early stages of product development. This chapter also provides a synopsis of fix effectiveness variability impact on AMPM-Stein and Crow-AMSAA reliability growth projection models.

Research questions to be answered are: (a) can *BBN* model structures provide more accurate fix effectiveness estimates than estimates made by subject matter experts as projected at the onset of a project, (b) can generic performance shaping factors and *BBN* structures provide accurate fix effectiveness estimates across diverse industries, (c) can implementation of a learning algorithm reduce model error, (d) are FEF dependent and if so, how, and (e) how does one account for FEF dependencies in reliability projection?

Chapter 2: Background

Chapter 2 provides an overview of the most popular reliability growth projection and planning models. It also provides insight into Bayesian methodologies which serve as the framework to propagate soft data, such as subject matter expert judgment, through Bayesian Belief models ultimately projecting fix effectiveness in the absence of test results.

Chapter 3: Model Building Methodology

As its title implies, this chapter provides the methodology to collect performance shaping factors that enhance or impede an engineer's ability to "fix" a problem, put that information into a respective structure, and collect subject matter expert judgment in terms of conditional probabilities and project fix effectiveness. Gradient ascent methods are used to allow models to learn conditional probabilities and reduce fix effectiveness error.

Chapter 4: Testing and Confirmation

Two test cases are presented whereby methods developed in Chapter 3 are applied to two diverse industries, HVAC and Automotive. Model structures and conditional probability tables were developed by subject matter experts within the HVAC industry. In a test of generalization, the same model structures and conditional probability tables were used to project fix effectiveness within the automotive industry. Gradient ascent methods are used to allow CPT to learn and reduce model error.

Chapter 5: Simulated Subject Matter Expert Judgment

Numerous challenges presented themselves during this research. Teams were overwhelmed by the amount of time required to build conditional probability tables for each model structure. In addition, group think became prevalent as dominant subject matter experts pushed their ideas on other team members. This chapter addresses both issues by development of methodology to simulate conditional probability tables for given model structures thus reducing the time required to build the models and eliminate group think in the process.

Chapter 6: Fix Effectiveness Dependency

This chapter provides four scenarios of two failure modes being "fixed" by two design teams chosen from one resource pool. Common factors such as management commitment, resource availability, and test facilities allow multiple subject matter Bayesian Belief Networks to be joined. Interdependency of fix effectiveness factors is evident as performance shaping factor evidence propagates from one side of the model to the other via common nodes. This evidence leads to the augmentation Crow-AMSAA Reliability Growth model to include FEF dependency.

Chapter 7: Conclusion

Test cases among HVAC and Automotive indicate that generic performance shaping factors and Bayesian Belief Networks can be used to project fix effectiveness both with a given industry and between industries if FEF dependencies are included in the model dynamics. The method developed in this research provided fix effectiveness projections that were much more accurate than those made by subject matter experts, 64% and 53% less within HVAC and Automotive industries respectively. Gradient Ascent methods were successful in learning subject matter expert knowledge for better representation of reality. Lastly, methods proving fix effectiveness dependency via common performance shaping factors led to model error reduction. The output of this research is a new failure intensity function that we recommend replace the current IEC standard for reliability growth projection.

Dedication

To my family

Mom and Dad

Who have spent much time in prayer asking
GOD to guide and help their children
Without their support and support from above
none of this would have been possible

My wife, Jacqueline, and

children, Kristi, Stephanie, David, and Jeremy

For your continued patience, understanding, and
support of my educational journey

Acknowledgements

This paper could not have been written without the support of the subject matter experts at Trane Commercial Systems who volunteered their time and energy to the success of this research. They are Rick Bishop, Michelle Cloyd, Moataz Elheddeny, Bert Feinberg, Joe Gresham, Bill Hansen, Bob Rowlands, Patrick Sowada, Bob Schult, Jay Johnson, Wayne Gorski, and John Wold.

I am grateful for the insightful suggestions provided by the following reviewers: Lesa Ross, Glen Arnold, Greg Ruhlander, Chris Etter, Melissa Lilley, and Kathy Wheeler. I also want to thank Dr. Paul Ellner and Dr. Brian Hall for numerous sessions of constructive feedback that elevated my understanding of the topic. Their interaction proved crucial to the success of this research.

Numerous professors at colleges and universities supported my educational efforts through the years. Professors at Jackson State Community College include Professors Chumney, Kemp, Hadley, and Frost. Professors at the University of Tennessee at Martin TN include Professor Charles Callis and Dr. Robert Drake. Professors at the University of Maryland include my advisor Dr. Ali Mosleh and Dr. Modarres. I thank each professor for their mentoring and support of my educational journey.

Lastly I want to thank Prakash Bedapudi, Chief Technical Officer at Lennox International. Prakash has supported the reliability cause in my career for many years. He has kept my career on tract, challenged me to grow technically, and think outside the box. He has proved himself to be leader among leaders, a great mentor, and a true friend.

Table of Contents

List of Acronyms

A	
A-mode	Failure Mode That Will Not Receive a Corrective Action
AMPM	AMSAA Maturity Projection Model
AMSAA	Army Material Systems Analysis Activity
ANOVA	Analysis of Variance
B	
BBN	Bayesian Belief Network
B-Mode	Failure Mode That Will Receive a Corrective Action
BC	B-mode That Will Receive a Corrective Action Upon Failure
BD	B-mode That Will Receive a Delayed Corrective Action
C	
CF	Critical Feature
CGAM	Global Air Cooled Chiller
CPT	Conditional Probability Table
CI	Confidence Interval
F	
FEF	Fix Effectiveness Factor
FMEA	Failure Mode and Effects Analysis
FRACAS	Failure Reporting and Corrective Action System
H	
HDBK	Hand book
HVAC	Heating-Ventilation-Air Conditioning
I	
IRIS	Integrated Risk Information System
IOM	Installation-Operation-Maintenance
IPAK1	Type of Roof-Top HVAC Unit

M

M-review	Marketing Review
MFG	Manufacturing
MGTP	Modified Growth-Tracking-Projection
MIL	Military
MLE	Maximum Likelihood Estimate
MME	Method of Moments
MTBF	Mean Time between Failures

N

NPI	New Product Introduction

O

ODYSSEY	Type of Roof Top HVAC Unit
OEM	Original Equipment Manufacturer

P

P-review	Program Review
Pdf	Probability Density Function
PSF	Performance Shaping Factor

R

RG	Reliability Growth
RGPC	Reliability Growth Planning Curve
RRX	Rank Regression on variable x
RTWD	Water Cooled Chiller

S

SME	Subject Matter Expert
S-SME	Simulated Subject Matter Expert
SPLAN	System Plan

T

T-review	Technical Review
TGP	Tracking-Growth-Projection

V

VAR	Variance
VOC	Voice of the Customer

List of Tables

Table

List of Figures

Figure

Chapter 1: Introduction

Motivation for Research

New product development involves an iterative process of failure mode mitigation and testing. Product development engineers may benchmark an existing design, determine all respective failure modes that surfaced during the warranty period and develop plans to mitigate those failure modes to an acceptable level. Testing validates the effectiveness of the "fix" or failure mode mitigation. In many cases, this process is repeated again and again due to an inability to correctly project the effectiveness of failure mode mitigation.

Existing models for reliability growth projection require test results in order to make the projection (Gibson & Crow, 1995). The new product introduction process (NPI) of Figure 1 will not produce test results until the second technical review (T2), which may be years after the required projection. The reliability growth planning models must have the capability of assessing specified business parameters such that accurate fix effectiveness projections can be in the absence of test results. Current reliability growth models are inept in meeting this requirement. In addition, reliability growth planning curve (RGPC), described in the next section, uses assumed reliability growth rates, estimates for fix effectiveness factors (FEF) based on previous history, or expert judgment with little mention of performance shaping factors (PSF) and their

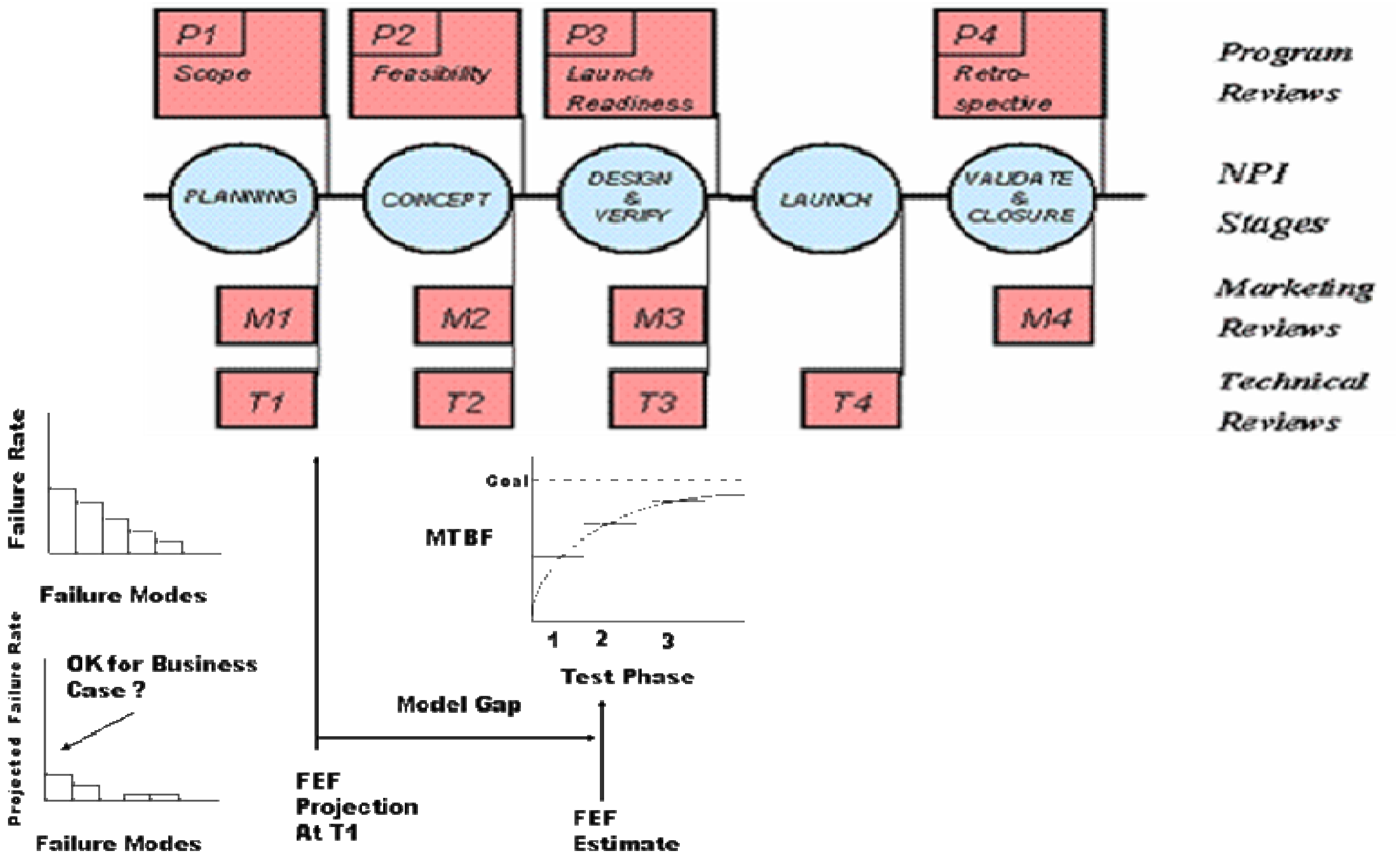


Figure 1. New product introduction process

interdependencies. New FEF projection methods are needed to allow reliability growth projections during project planning.

Significance of Fix Effectiveness Error on Reliability Growth Projections

A portion of this research has been dedicated to exploration of the impact that fix effectiveness variation has on reliability growth projection error. The original intent of the research was to determine the more robust reliability growth model, Crow-AMSAA or AMPM-Stein. Mode failure rates were held constant while each model experienced random realizations of increasing fix effectiveness variation. Reliability growth projection error was noted for each model. Results show that test of equal variances indicate one must fail to reject the null hypothesis of no difference in variances between each model for a given fix effectiveness variation. Two sample T tests however, indicate one must reject the null of no difference in mean error. In every instance, when a difference in mean error was detected, the AMPM-Stein mean error was lower. Statistically, the AMPM-Stein model is more robust against the effects of fix effectiveness variability than the Crow-AMSAA reliability projection model.

While AMPM-Stein proved to be statistically more robust, the distinction was made at 1031 samples, detecting a difference of 1/7 standard deviation, with a constant power value of 0.9 rendering the findings insignificant from a practical perspective. Most product development programs seldom see sample sizes above 50. At low sample sizes, fix effectiveness variation will not impact one model more than the other.

One major observation during experimentation was the overall error in reliability growth projection with each incorrect estimate for fix effectiveness. Both AMPM-Stein and Crow-AMSAA models experienced up to 20% error in reliability growth projections as fix effectiveness variation approached 0.05. Thus research efforts shifted toward development of a method for a more accurate fix effectiveness projection, specifically during the planning process in the absence of test results. But what elements influence fix effectiveness projections? Is it management commitment, test facilities, complexity of the failure mode, availability of resources, etc? How would one build such a structure? Is one method of projecting fix effectiveness better than another?

Reliability growth projections developed during project planning incorporate use of an estimate for fix effectiveness. Subject Matter Experts (SME) provide this estimate based on years of experience within a particular industry or past experiences. As discussed earlier, minimizing fix effectiveness projection error is crucial for accurate reliability growth projections. Over confident SME may project high fix effectiveness providing an overestimate of improvements (i.e., elimination and or reduction of critical failure modes, much improved MTBF [mean time between failure], more positive return on investment, etc.). Low fix effectiveness estimates will overly reduce return on investment estimates or MTBF projections such that a project may be inadvertently eliminated from an engineer's scope of annual projects.

The methodology presented in this paper established a framework for SME to determine performance shaping factors (PSF) that enhance or reduce an engineering team's ability to "fix" a failure mode. In addition those same SME arranged PSF direction of influence and derived an estimate for fix effectiveness (FEF) via Bayesian Belief Network (BBN) methodology. Three methods were used to build a BBN and project FEF: (a) M-1 expert aggregate, (b) M-2 fixed structure, and (c) M-3 consensus models. A fourth method, (d) M-4 aggregate BBN was simply an aggregate of each of the three structures.

Within M-1, expert aggregate model, SME were allowed to build their own BBN using the PSF previously defined, assign the direction of node influence, and establish parent-child relationships and node conditional probabilities. This method allowed SME to project FEF for their respective BBN. An aggregate FEF was quantified from the output of individual SME FEF.

Within the second model, fixed structure (M-2), SME reached consensus on the model structure, and entered their respective judgment on parent-child conditional probabilities for that BBN and project FEF. For the consensus model (M-3), SME reached consensus on the conditional probability tables (CPT) for the previously agreed upon structure of M-2. In addition we explored an M-4 model as an aggregate of M-1, M-2 and M-3. This provided an estimate for FEF while addressing model uncertainty. FEF projections from each method were evaluated against both known FEF from past projects and SME FEF projections.

Research questions to be answered are

1. Can *BBN* model structures M-1, M-2, M-3 or M-4 provide a more accurate *FEF* estimate than current industry *FEF* projection methods?
2. Can generic *PSF* and *BBN* structures provide accurate *FEF* estimates across diverse industries?
3. Can implementation of a learning algorithm such as Expectation Maximization reduce model error?

4. Are *FEF* dependent and if so, how?

5. How does one account for *FEF* interdependencies in reliability projection?

The significance of this research is two-fold. First no structured method of projecting fix effectiveness factors during planning exists. Second, no reliability growth projection model accounts for dependencies among fix effectiveness factors. The researcher has defined the process to account for both of these short comings and provided test cases for confirmation the process works. Both of these concepts, structured FEF projection and FEF dependencies have been overlooked since the beginning of reliability growth projection modeling. One can see in Figure 2 and Figure 3, unstructured subject matter estimates for fix effectiveness led to overestimates and excessive error when compared to actual results for fix effectiveness. Structured methodology described within this dissertation, provides specific steps subject matter experts can follow such that fix effectiveness error is substantially reduced, (M1-M4).

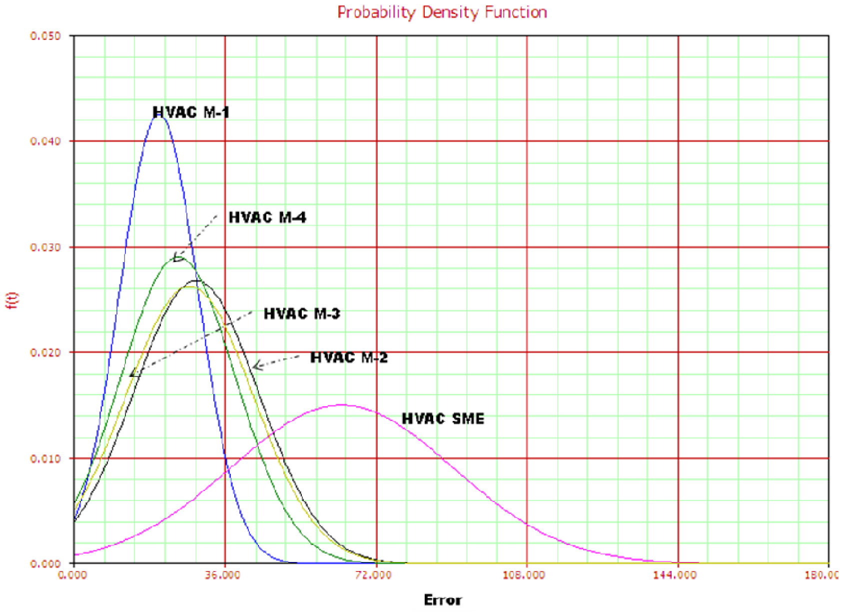


Figure 2. HVAC FEF projection error by model type

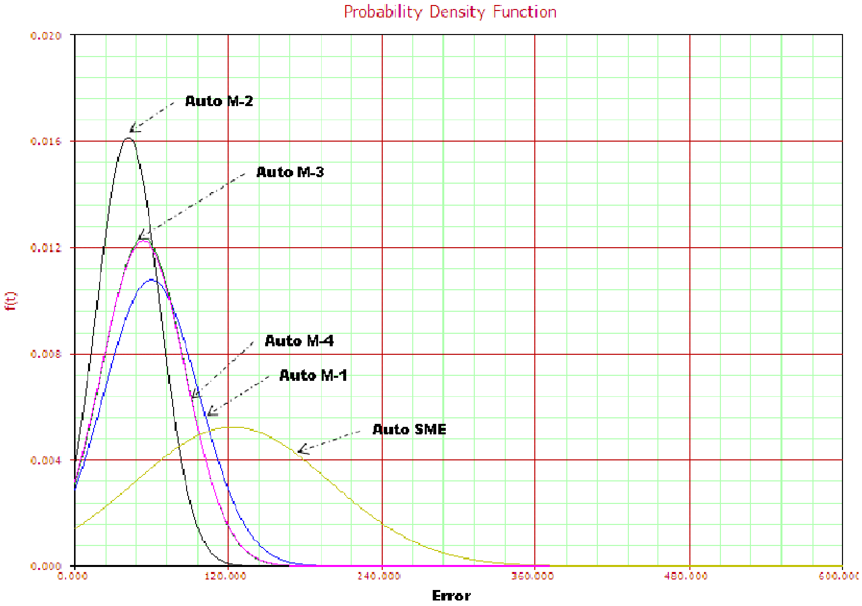


Figure 3. Automotive FEF projection error by model type

Chapter 2: Background

This section provided background information for reliability planning, tracking and projection models commonly used in industry. It was not an attempt to cover all models, but merely show how estimates of fix effectiveness impact reliability growth.

Reliability Planning Models

Duane (1964) recognized that a logarithmic relationship exists between cumulative failure rate and cumulative time on test. The original intent of the model was to track reliability improvement based on growth rates demonstrated during testing. Selby and Miller (1970) as well as the U.S Department of Defense (1981) expanded the Duane postulate into a planning tool to predict future MTBF based on assumed growth rates. MIL-HDBK-189 (Figure 4) provides a detailed discussion of reliability growth planning. Reliability growth planning involves evaluation of schedule, testing requirements, and technical resource needs and availability of those resources to construct a planned reliability growth curve. Previous programs, past lessons learned, etc. may be evaluated to estimate initial MTBF, time on test requirements, growth rate assumptions, and final MTBF. The curve contains interim reliability goals such that stage gate reviews can compare program reliability progress against the curve and flag potential reliability growth concerns. Should slip from target occur, management can reallocate resources or adjust other variables within their control to put the program back on track. Two approaches are used to build a reliability growth planning curve (RGPC). The first approach involves combining expert judgment from similar projects to develop an idealized RGPC representing an "expectation" of growth. The second method is a planned RGPC based on program milestones. Management establishes the growth target and time the project must be completed, thus dictating the growth rate. During stage gate reviews the realized growth curve is compared to program milestones allowing for resource reallocation to meet program metrics.

SPLAN (system plan) is a derivative of MIL-HDBK-189 in that it provides options for obtaining planning parameters (Ellner, McCarthy, Mortin, & Querido, 1995). For example, during the planning process management may enter initial MTBF, goal MTBF, assumed growth rates, and total length of a test program. SPLAN then can calculate the required time on test such that program objectives can be met.

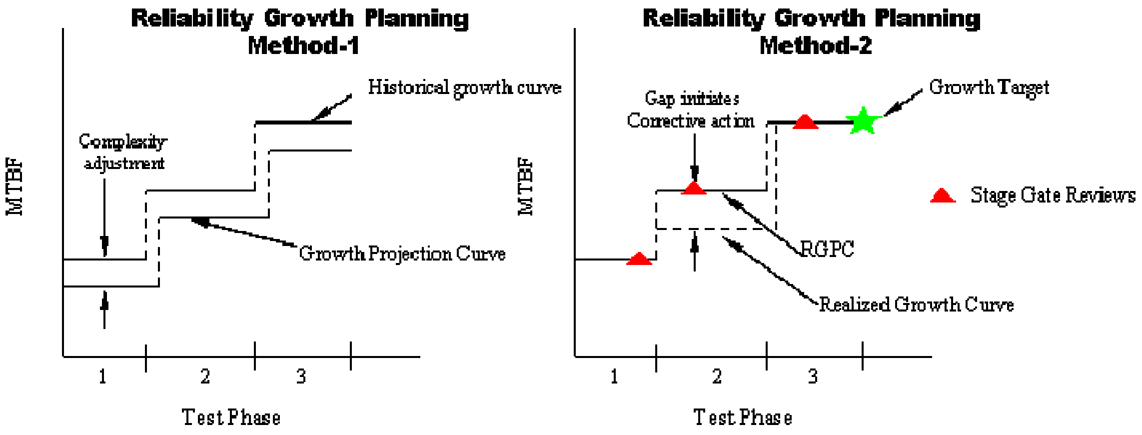


Figure 4. Mil-HDBK-189 reliability growth planning models

Victor Pellicione developed the first Tracking-Growth-Projection (TGP) model based on the logistic function. Farquhar and Mosleh (1995) modified TGP (MTGP) by adding a growth rate restricting factor ρ, (13) as a function of the business culture.

$$\lambda(t) = (\lambda_o - \lambda_p)e^{-\rho K t} + \lambda_p \tag{13}$$

Quantification of the business culture involves normalizing subjective input from the following categories:

1. Management
2. Reliability Engineer's Experience
3. Reliability Growth Test Plan
4. Growth Test Controls
5. Specification Requirements
6. FRACAS
7. Schedule
8. Starting Point
9. Reputation

The overall failure rate for the system is quantified during testing by characterization of the business culture restricting factors ρ, the growth rate K and test time t.

Crow-AMSAA projection model is used as a planning model (Figure 5). Engineers review warranty data and determine field failure modes that will not be fixed (A-modes) and those that will receive corrective action (B-modes). A failure rate projection can be estimated during planning based on assumed values of d_i.

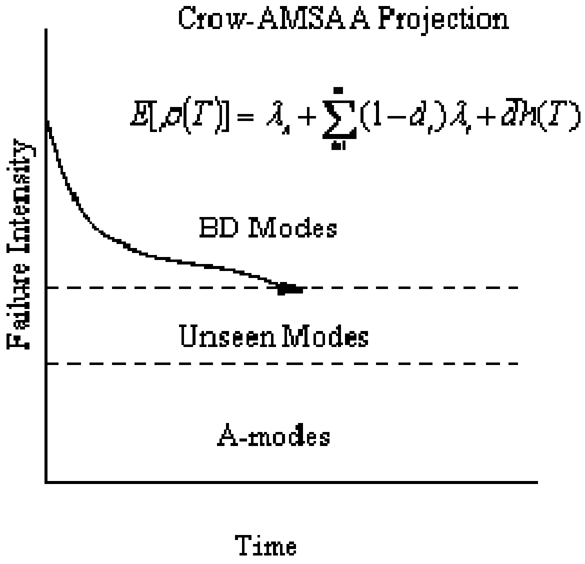


Figure 5. Crow-AMSAA projection model

Figure 6 shows an analysis method used by numerous industries whereby they evaluate the reliability improvement potential associated with assumed values of fix effectiveness. Based on the Pareto principle, one can see the curves begin to flatten as improvement opportunities diminish. The output of the analysis is a list potential B-modes upon which an engineering manager can optimize resource allocation for maximum reliability improvement. A short coming of the method is that it is based on assumed values of fix effectiveness. Once the number of “fixes are quantified, engineering can make an estimate of life improvement for the population in question (Figure 7).

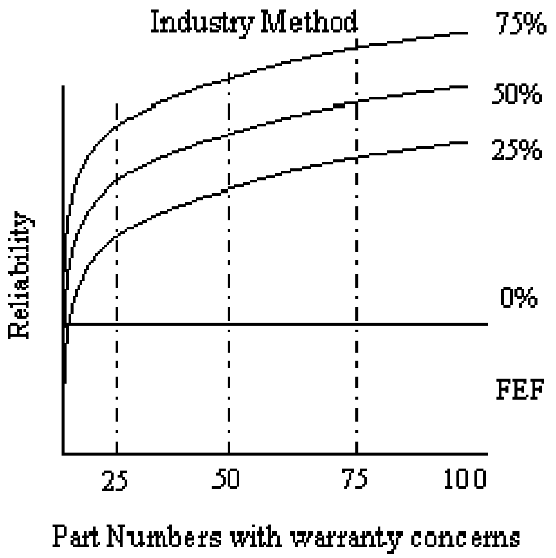


Figure 6. Optimized resource allocation

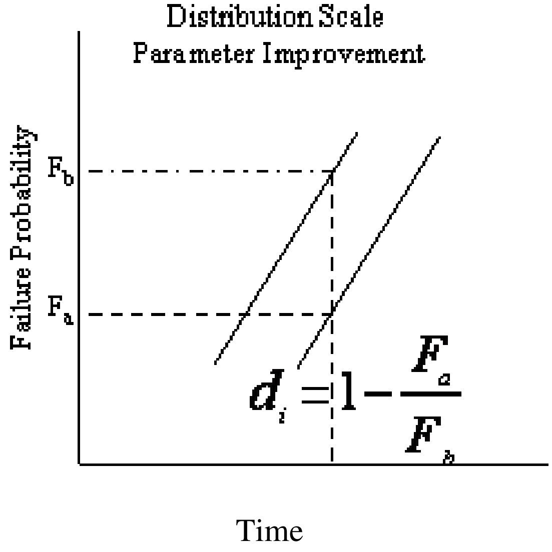


Figure 7. Improved scale based on assumed fix effectiveness

<u>*Reliability Tracking Models*</u>

The Weiss (1956) model, modeled guided missile systems with Poisson type failures. MTTF is assumed to change over successive trials given fix effectiveness of surfaced failure modes. The model results in a logistic reliability growth curve. This model represents one of the earliest models that track reliability as a result of fix effectiveness.

Aroef's model (1957) tracks reliability growth for continuous systems. For a given point in time, Aroef believed the rate of reliability improvement of a system is directly proportional to the growth achieved and inversely proportional to the test duration squared. Rosner (1961) also modeled system reliability via the failure intensity function. However he believed the rate of occurrence of failure at a particular time is proportional to the number of non-random defects that are still in the system at time t. His model was able to estimate the required test duration for a system as a function of the fraction of original failures corrected.

Lloyd-Lipow (1962) developed a reliability growth model for a system with one failure mode. Testing is conducted in multiple phases, with corrective actions implemented as failures surface. Their model, $R_k = R_\infty - \frac{\alpha}{k}$ provides an estimate of system reliability within a given test phase. They present MLE for model parameters, R_∞ and α resulting in an estimate of long term reliability given the current growth rate between test phases. Chernoff and Woods (1962) developed a similar model estimating the probability a system will be successful after a given number of failures have occurred. The model is of the form $P_r = 1 - e^{-(\alpha + \beta * r)}$ where α and β are models parameters developed by least squares methods.

Wolman's model (1963), assumed all assignable cause failures are of equal probability within each trial and are completely eliminated upon initial observation, thus the model assumes fix effectiveness is 100% for observed assignable cause failure modes. Therefore reliability is improved over a sequence k trial given by $R_k = 1 - q_o - (M + 1 - k)q$ where q_o represents the probability of a non assignable cause failure mode, M is the initial number of assignable cause failure modes and q is the probability of occurrence of a single assignable cause failure mode. Barlow-Scheuer (1966) is also a k stage reliability growth model. They debate that exactly one of three outcomes can occur in a given stage, success, inherent failure, or an assignable cause failure whereby the reliability for the i^th^ stage is given $r_i = 1 - q_o - q_i$, where q_o is the probability of inherent failure and q_i is the probability of an assignable cause failure.

Virent's Gompertz model (1968) is based on the trinomial Gompertz equation, $R = ab^{c^t}$ where $b, c \in (0,1)$. As $t \to \infty$ the parameter a is defined as the upper limit on reliability.

Pollock's model (1968) utilized Bayesian methodologies to model parameters as random variables with associated prior distributions such that one could project system reliability with or without test data. Pollock's model may represent the earliest Bayesian reliability growth model.

Within Crow's Continuous tracking model (1974), the instantaneous failure rate for reliability growth, given by Duane's model is reparameterized and as a Weibull hazard rate function for a repairable system. The model is given by $r(t) = \lambda\beta t^{\beta-1}$, where λ and β are model parameters. Crow goes on to develop goodness of fit for reliability growth for both time and failure truncated data. Numerous application examples are noted.

Singpurwalla (1978) developed a discrete reliability growth model to: determine if the binomial parameter p_i is increasing after k design phase modifications. In addition Singpurwalla provides for future projections of p_i beyond the k^{th} design modification.

Reliability Growth Projection Models

Product design engineers establish new product specifications as a function of customer requirements for performance, efficiency and reliability. The voice of the customer (VOC) may be described in a broad sense such as “easy to use”, “quiet operation”, “safe”, “lasts a long time” or “easy to maintain.” The engineering community must transpose VOC into technical requirements. For example, “quiet operation” may consist of technical requirements defined as sound levels not to exceed X decibels, or “safe” is defined as an operator reach to an interface panel is to not exceed a distance of Y, and “lasts a long time” may be defined in terms of $MTBF \geq Z$. These global level technical requirements must be driven down into critical features (CF) at the component and part level. The objective of this process is to control CF to satisfy technical requirements, thus satisfying the voice of the customer.

Concept or prototype units are built utilizing previously defined CF. Design, assembly, and manufacturing concerns are noted as the prototype units are tested against a battery of specified conditions. The intent is to determine failure mode existence and the time of occurrence, thus exposing deficient design concepts, manufacturing processes, or supplier variation. Surfaced failure modes receive corrective actions, and tests are repeated to determine fix effectiveness and system reliability.

Reliability growth models are used to quantify MTBF at the end of the initial test phase and project MTBF based on assumed fix effectiveness factors (FEF) of corrective actions. Various reliability growth models exist such as the Duane (1964) model, Crow-AMSAA, Crow-AMSAA Projection (Crow, 2004, 2006), and the AMPM-Stein (Ellner & Hall, 2005). Duane recognized a logarithmic relationship between cumulative MTBF and cumulative time on test. Crow-AMSAA stochastically represented the Duane model as a Weibull process defining reliability growth within a test phase. The reliability growth models evaluated in this research are those used to predict reliability growth across test phases where testing is continuous, and corrective actions are delayed until the end of test. The two models under evaluation include the Crow-AMSAA Projection and AMPM-Stein reliability projection.

The Crow-AMSAA Projection model classifies failure modes into A and B modes. A-modes will not receive corrective actions whereas B-modes will. Fix effectiveness factor (FEF) is defined as the percent reduction in the failure intensity for the i^{th} B-mode as a result of permanent corrective actions to the product design and/or manufacturing processes (Crow, 2004, 2006). AMPM-Stein uses a similar method with the exception that A and B-modes are defined by zero and positive realizations of *FEF* respectively, and estimates for the true failure intensity for the i^{th} failure mode are based on the Stein shrinkage estimator (Ellner & Hall, 2005; Ellner & Wald, 1995). Both models define failure intensity function contribution for unsurfaced failure modes, but utilize different methods for quantification. Crow-AMSAA Projection utilizes an average fix effectiveness factor multiplied by a Poisson intensity function that quantifies the rate at which new failure modes are being introduced. Conversely, the AMPM-Stein shrinkage estimator naturally allows for estimation of unsurfaced modes without assumption of an underlying distribution.

Crow-AMSAA

The Crow-AMSAA Projection model classifies failure modes into two broad categories, A-modes and B-modes (Crow, 2004, 2006). A-failure modes are those that will have no corrective actions. B-failure modes are further categorized into BC and BD modes. BC failure modes are B-modes that will be "corrected" during the testing phase whereas BD failure mode corrective actions will be delayed until the end of the test. It was previously stated the reliability growth models under consideration are continuous, and corrective actions are delayed until the end of the test, therefore BC failure modes will not be considered in this research.

The Crow-AMSAA Projection model assumes all BD modes are in series and fail according to an exponential distribution. A-mode occurrences also follow an exponential distribution with failure intensity $\lambda_A = \frac{N_A}{T}$. Since corrective actions are delayed until the end of the test phase, *MTBF* remains constant through the test and then jumps to a higher value pending effectiveness of fixes. Let k indicate the total number of BD modes and λ_i the failure intensity for the i$^{\text{th}}$ BD mode where $\lambda_{BD} = \sum_{i=1}^{k} \lambda_i$ such that at t=0, $\rho(0) = \lambda_A + \lambda_{BD}$. *FEF* are denoted by d_i, representing the fraction decrease in λ_i due to corrective action on the i$^{\text{th}}$ mode with (1-d_i) representing the remaining portion after fix. If during a test phase *m* of *k* BD modes surface, corrective actions are implemented on the *m* surfaced modes with a *FEF* of d_i, thus at time T, $\rho(0)$ becomes $\rho(T)$.

$$\rho(T) = \lambda_A + \sum_{i=1}^{m}(1-d_i)\lambda_i + \left(\lambda_{BD} - \sum_{i=1}^{m}\lambda_i\right) \text{ (Reduces to equation 2)} \tag{1}$$

$$\rho(T) = \lambda_A + \lambda_{BD} - \sum_{i=1}^{m} d_i \lambda_i \tag{2}$$

$$\sum_{i=1}^{m}(1-d_i)\lambda_i \text{ remaining portion of BD modes after corrective action} \tag{3}$$

$$\lambda_{BD} - \sum_{i=1}^{m}\lambda_i \text{ contribution to failure intensity due to all unseen failure modes} \tag{4}$$

Note the failure intensity; $\rho(T)$ of equation 1 has failure contribution from three areas; (a) A-mode failure intensity, (b) the remaining portion of corrected BD-modes, and (c) the bias correction term. The bias correction terms is estimated using average FEF multiplied by the instantaneous rate h(t) at which first occurrence of new BD modes are occurring at time T with the MLE for $h(t)$ defined as

$$\hat{h}(T) = \hat{\lambda}_{BD}\,\hat{\beta}_{BD}\,t^{\hat{\beta}_{BD}-1} \tag{5}$$

Thus, the expected value of ρ(T) is defined as

$$E[\rho(T)] = \lambda_A + \sum_{i=1}^{m}(1-d_i)\lambda_i + \bar{d}\,\hat{h}(T) \tag{6}$$

AMPM-Stein

The AMPM-Stein model also assumes failure modes fail independently according to an exponential distribution. FEF represents the fraction decrease in failure intensity due to implementation of a corrective action with all corrective actions delayed until the end of the test phase. The AMPM-Stein model does not label failure modes as A or B-modes but distinguishes between them by zero or positive FEF for surfaced A and B-modes respectively. One of the unique characteristics of the AMPM Stein model is the estimation procedure is based on the Stein shrinkage estimator (Ellner & Hall, 2005; Ellner & Wald, 1995).

Again, assume a system has $k > 1$ potential failure modes with initial failure rates $\lambda_1,....\lambda_k$, and N_i represents the number of failures for mode *I* during the test phase with the *MLE* of $\lambda = \hat{\lambda}_i = \frac{N_i}{T}$. In order to more accurately estimate λ_i, Ellner and Hall (2005) utilizes the Stein estimator $\tilde{\lambda}$ given by:

$$\tilde{\lambda}_i = \theta \hat{\lambda}_i + (1-\theta)avg\left(\hat{\lambda}_i\right) \tag{7}$$

The value that optimizes θ is θ_s, which is chosen to minimize the expected sum of squared errors between $\tilde{\lambda}_i$ and λ_i such that,

$$\theta_s = \frac{Var[\lambda_i]}{\left(\frac{\lambda}{kt}\right)\left(1-\frac{1}{k}\right) + Var[\lambda_i]} \tag{8}$$

$$\text{Where } \lambda = \sum_{i=1}^{k} \lambda_i \qquad \bar{\lambda} = \frac{\lambda}{k} \quad Var[\lambda_i] = \frac{\sum_{i=1}^{k}\left(\lambda_i - \bar{\lambda}\right)^2}{k} \tag{9}$$

Given that $N_i = 0$ for failure modes that have not occurred, Ellner and Hall use equation (7) to show the failure intensity for failure modes not surfaced by t, as:

$$\sum_{i \in \overline{obs}} \tilde{\lambda}_i = \frac{\left(1-\frac{1}{k}\right)\left(1-\frac{m}{k}\right)\left(\frac{N}{t}\right)}{\left(1-\frac{1}{k}\right) + \left(\frac{\sum_{i=1}^{k}\lambda^2}{\lambda} - \frac{\lambda}{k}\right)t} \tag{10}$$

The resulting Stein failure intensity is denoted as $\rho_s(T)$ where,

$$\rho_s(T) = \sum_{i \in obs} (1 - d_i\,^*)\tilde{\lambda}_i + \sum_{i \in \overline{obs}} \tilde{\lambda}_i \text{ and } d_i^* \text{ is a realization of } d_i. \tag{11}$$

Ellner and Hall (2005) note the exact solution of (11) is a function of unknown constants, k, Var[λ_i] and λ. Approximation models are developed using N_i and m to determine Gamma Distribution parameters, α and β based on both MLE and MME. This technique led to approximations for θ_s, λ and Var[λ_i]. Equations are then developed for large k and $k \to \infty$ resulting in an estimate for failure intensity based on the Stein estimator as $k \to \infty$ defined as $\rho_{s,\infty}(T)$.

$$\rho_{s,\infty}(T) = \lim_{k \to \infty} \hat{\rho}_{k,\infty}(T) = \sum_{i \in obs} (1 - d_i\,^*)\hat{\theta}_{s,\infty}\hat{\lambda}_i + \left(1 - \hat{\theta}_{s,\infty}\right)\frac{N}{T} \tag{12}$$

<u>Hierarchial Bayesian Framework</u>

A more recent development in the field of Reliability Growth modeling is a Hierarchical Bayesian framework developed by (Droguett & Mosleh, 2008). This methodology allows use of various sources of information such as the historical data on earlier system designs, in-house test data under both accelerated and nominal test conditions, engineering judgments about the impact of design modifications and failure mode fixes on the product’s failure intensity, and finally the observed performance in the field by the system.

The methodology implements an analysis procedure which breaks down the problem of assessing the reliability of future systems into a number of analysis steps that are part of different stages in the system’s design evolution. Each analysis step consists of a Bayesian analysis in a particular stage of the projected design evolution. Therefore, using the evidence from the sources described in the previous section, a different reliability function is estimated at each step in the analysis. The result of the estimation at each step consists of uncertainty distributions over the reliability as a function of time.

Figure 8 shows the various types of data that might be available for assessing the reliability of a system during its development phase. In the beginning, such data includes heritage data and results of reliability modeling for the new system based on heritage data for its components. During the design and development process, data on tests, impact of design modifications, and failure mode removal become available. After the development and release of the system, field data were accumulated, constituting the most relevant data for assessing the actual (observed) reliability of the system. These types of data are shown on a time line of system development and release in Figure 8. The analysis steps in each of these stages are also detailed in the figure.

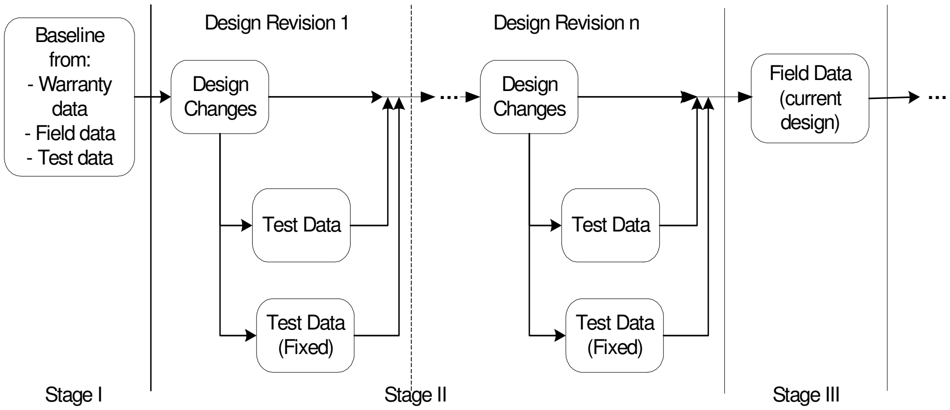


Figure 8. Overview of the methodology analysis steps

(1) The first step in this analysis flow is to establish a reliability assessment of the baseline comparator. A comparator is a previous design that is usually the newest, most

relevant system in the market. To do this assessment, multiple data sets of the comparator can be used. These data sets are based on heritage data and usually the available data types are in the form of warranty data, field data, and test data. These heritage data were considered relevant to the baseline comparator. In order to be able to scale the impact of the data on the baseline estimate, a relevance factor, ranging between 0 and 1, is assigned to data originating from the comparators.

(2) Following the baseline analysis, as shown in Figure 8 the development stage is comprised of several "design revision" programs. For each design program, three analysis steps are possible. The 'Design Changes' step modifies the result of the previous design step corresponding to the anticipated impact of the design changes. This step therefore is not a Bayesian update in the conventional sense, where data were added to update the estimate of a given quantity, but rather it transforms the results from earlier steps in order to estimate new values of the product reliability by a "design credit" factor (which is often an uncertain quantity assessed by SMEs).

(3) The 'Test Data' and 'Test Data (Fixed)' steps are used to validate the above results based on the modified design. These analyses steps include a check to see whether the test data indicated reliability metric significantly different from values estimated based on the anticipated impact of design changes. The difference between the 'Test Data' and 'Test Data (Fixed)' is whether the *FEF* and design credits were taken into account or not. Together with the 'Design Changes' step, they form the three analysis steps that can be carried out for each design round. Depending on which of the steps have been performed, one of the three analysis steps is used as the baseline point for the analysis of the next design round.

When a given reliability target has been reached, the product is deployed to the market. During this stage, operational experience is accumulated for the current system design. This field data can, in a later point in the design evolution, be used as a baseline for the reliability assessment of a new product design, thus restarting the cycle illustrated in Figure 8.

The methodology assumes an underlying reliability such as a Weibull with the following hazard rate, with two parameters:

$$h(t) = \frac{\beta}{\alpha^{\beta}} \cdot t^{\beta - 1}$$

Basic analysis procedure consists of a hierarchical Bayesian estimation procedure where data were applied to find joint posterior probability distributions for the Weibull parameters at each of the stages shown in Figure 8. At each stage the posterior distribution from the previous stage plays the role of the prior distribution for the next round on Bayesian updating. At each stage the reliability metrics (e.g, hazard rate) are found using the updated distribution of the Weibull parameters. The last round of calculations prior to product release are the prediction of the reliability of the product in its intended field application and environment.

The inclusion of the “design credit” into the likelihood functions is done through “proportional hazard” model (adjustment of failure rate). The inclusion of FEF is done either by the proportional hazard model or proportion reduction of failure counts (used for instance in a Poisson likelihood function) for the failure modes affected. The FEF values can be an uncertain value specified by, for instance, a beta probability distribution.

In this research, the Bayesian methodology provided a philosophically consistent framework for reliability growth projection as it recognizes the subjective and uncertain nature of FEF assessment. It provides the mathematical formalism for inclusion of such uncertainties and explicit account of their impact on the reliability metrics of interest.

One can see from all of the models noted above, the term di if determined during planning, is the result of unstructured subject matter expert judgment. Left unstructured, their judgment is subject to bias, resulting in under estimates or over estimates. This leads to reliability growth projection uncertainty that until this research has never been considered as a function of FEF variability (Corcoran, Weingarten, & Zehna, 1964; Crow, 1982, 1983, 1989, 2004, 2006; Ellner & Hall, 2005; Ellner, McCarthy, Mortin, & Querido, 1994; Hall, 2007, 2008; Selby & Miller, 1970).

Chapter 3: Model Building Methodology

Causal Model Building Process

The process defined below provides the steps necessary to build a causal model via Bayesian Belief Networks such that one can develop a structured SME FEF projection model. Possibly performance shaping factors, model structures, conditional probability tables, direction of influence among PSF etc. within a given organization are different than those defined in this research. However; the process to build the causal model and project fix effectiveness will be the same for your organization as was used for this research.

Walls and Quigley (2001) define five steps for elicitation and organization of expert judgment: (a) select the subject matter experts, (b) brief experts, (c) elicitation of judgment, (d) aggregation of judgment and (5) feedback for calibration. We expanded their process to include more detail as shown in Figure 9. Our process consists of (a) selection of subject matter experts, (b) defining PSF, (c) assigning the direction of influence among PSF (d) building CPT, (e) projecting FEF via BBN (structured SME), (f) collecting unstructured SME FEF projections from past projects, (g) obtain actual FEF from past projects, (h) compare difference from causal model FEF projections to unstructured SME projections and actual FEF, and (i) initiate learning algorithm and repeat step 8.

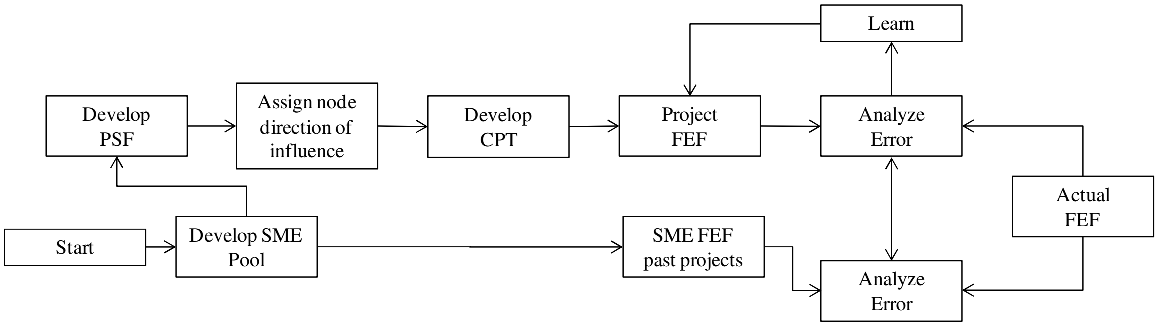


Figure 9. Causal model building process

As shown in Figure-8, the first step in the causal model process is the selection of subject matter experts. Subject matter experts need to have proven strong technical expertise within a given functional area and be familiar with the areas of interest for this study. They must be willing to act as an impartial evaluator, commit the needed time and effort for this research, participate in debates, be prepared for discussions, strong communication and interpersonal skills, and impartiality (Ashton, 1986; Ayyub, 2001; Mosleh & Apostolakis, 1986; Mosleh, Bier, & Apostolakis, 1988; Mosleh, 2002). Lastly, subject matter experts must be able to generalize and simplify.

The number of subject matter experts required is similar to a sample size representing a specific attribute within a population. The larger the sample size, the more representative of the true mean of a population for the attribute in question. However as sample size increases, we reach a point of diminishing returns as to what "new information" the additional data provides. Hogarth's normative model (1978) suggests that 6-10 subject matter experts provide the most accuracy. Ashton's work (Ashton & Ashton, 1985; Ashton, 1986) implies that eliciting 3-6 experts lead to high accuracy levels, whereas Calvin Shirazi and Mosleh (2009) conclude 6-7 experts are adequate.

Few researchers suggest that gains in accuracy are attributed to the inter-correlations of the experts, and minimal gain in judgment accuracy is achieved from redundancy in experts (Budescu & Rantilla, 2000).

Ten subject matter experts were chosen to develop performance shaping factors. Experts are briefed in methods of Bayesian Belief Networks by reviewing single and multiple parent nodes along with providing them with an understanding of marginal probabilities and observed evidence propagation. SME are asked the question, what impedes or enhances their ability to fix a failure mode, whereby they ultimately reach consensus on a list of PSF. The next step in the process is to use the PSF to build one of four model structures to project FEF. The first model is the expert aggregate model, M-1. Each SME is free to determine the direction of influence among the PSF and develop their CPT. The second model, M-2 fixed structure method, SME reach consensus as to the model structure and direction of influence of all PSF. Each SME is left to determine their respective CPT for the agreed upon "fixed" structure. Within the third model, M-3 consensus model, SME reach CPT consensus for the fixed structure of M-2. Lastly, the BBN aggregate, M-4, addresses model uncertainty by aggregation of M1, M-2 and M-3.

As shown in Figure 8 data, unstructured subject matter expert fix effectiveness factor projections, actual FEF and PSF states are collected from previous projects. PSF states are entered into each model structure, M-1, M-2, M-3 and M-4 allowing for a comparison between unstructured SME FEF projection, actual FEF and the causal model "structured" FEF projection. Various iterations of the learning algorithm are used to reduce the causal model BBN FEF projection.

The information in the remainder of this dissertation is test cases used to confirm the structured causal model BBN FEF projection process and associated implementation, e.g. fix effectiveness dependency.

Performance Shaping Factors

Sources of Fix Effectiveness Variation

Numerous areas may impact fix effectiveness as product moves from the drawing board to the field (Figure 10). For example, a design team may have all the best intentions of solving a failure mode, but may lack understanding of the complexity of the problem whereby their efforts prove unsuccessful. Manufacturing and supplier processes may be incapable of long term control of critical features (CF) post implementation of a "fix", whereby FEF results diminishes. In addition, FEF variability may be associated with shipping, installation, service, sales, and the customer. An original equipment manufacturer (OEM) may develop installation, operation, and maintenance (IOM) guidelines that instruct contractors, maintenance personnel, and end-users in the do's and do-not's of equipment IOM. For this study, it is assumed that shipping, installation, service, sales, and the customer operate within the IOM guidelines, whereby FEF variation is considered a function of design, manufacturing and supplier only (Figure 11).

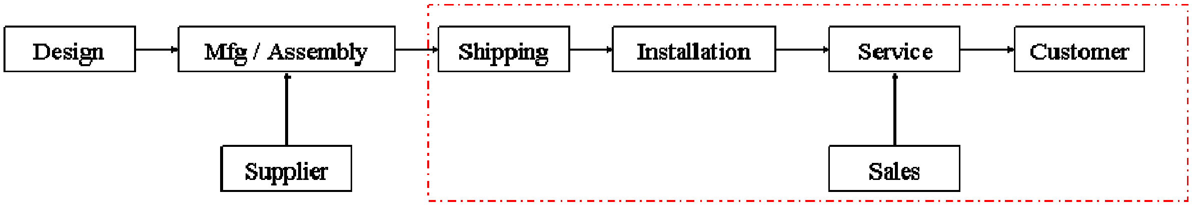


Figure 10. Sources of FEF variation

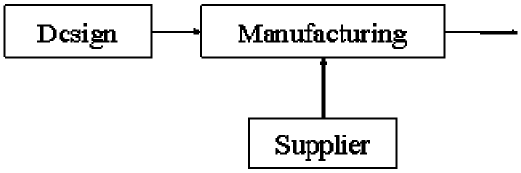


Figure 11. FEF variation reduction model

Assumptions for FEF variability include:

1. Assembly is included as part of the manufacturing process.
2. Shipping methods/route do not damage the product.
3. Installation/service is within the IOM guidelines.
4. Sales or aftermarket teams operate within IOM guidelines.
5. Customer uses the product within the IOM guidelines.

Identification of Performance Shaping Factors

The sample frames for SME included in this research are those with expertise in the numerous fields of engineering, manufacturing, and program management. SME have careers that span over 15 diverse industries including pipeline construction, HVAC, ice machines, communications, business/finance, automotive, aluminum forging, electric motor, furniture, power tools, heavy truck, metrology equipment, boat manufacturing, electronic controls, and process control. In addition SME has demonstrated expertise in one or more of the following disciplines with an average of 27 years in their respective field.

1. Heat transfer
2. Electrical/Mechanical/Reliability Engineering
3. Compressor design
4. Motor design

5. Aerodynamics
6. Fan technology
7. Quality
8. Manufacturing
9. Program management

Team diversity was chosen to generalize PSF across a broad range of engineering communities. For example, choosing engineers from pipeline construction, HVAC, business/finance, automotive, etc. provided a broad sample of expert knowledge, increasing the likelihood of generating the vital few PSF that influence FEF. SME are asked to categorize what areas positively or negatively impact an engineer's ability to "fix" a failure mode. They openly debate PSF reaching consensus on PSF they feel significantly impact fix effectiveness (Smith, 1989). SME reached consensus on a list of dominant PSF, states of each PSF, and operational definitions. The operational definitions provide a means of using surrogate data to populate PSF states. For example when attempting to determine the states of the PSF management commitment, it may be politically incorrect for an SME to openly state management is not committed to a "fix" activity. However, by using the operational definition of management commitment, the SME can ask the questions, does leadership review the project each month, quarter (i.e. is it on their radar) and are they providing monies to support the project? If the answers are no review and no monies, the SME can use this surrogate to determine management commitment is low. Conversely, if management does review the project in a timely manner and provides monies, management commitment can be determined to be high.

1. *Management commitment* – [low, high] does leadership support this project, do they review in a timely manner, i.e., monthly, quarterly, will they support monetarily.
2. *Time of project* – [adequate, inadequate] time from start to launch…if time interval is short, it decreases success probability
3. *Complexity of the failure mode* – [low, high] phenomenon not understood, field conditions not understood, multiple mechanisms driving failure mode, etc.
4. *Technical expertise* – [adequate, inadequate] assembled team is a well rounded, technically competent, pool of engineering talent relative to the failure mode in question.
5. *Availability of resources* – [adequate, inadequate] are the required resources available to dedicate enough time to make the project successful.
6. *Design complexity* – [low, high] numerous design iterations make it impossible to verify all combinations
7. *Test facilities* – [adequate, inadequate] are facilities identified (OEM or supplier) to generate the failure mode in question, i.e. failure modes can be turned on and off.
8. *Quality system maturity also denoted as new product introduction (NPI) maturity* – [low, high] team has NPI knowledge and execution skills, includes identification of CF by design and understanding of CF capability by manufacturing and supplier.

9. *FEF* – [10%, 30%, 50%, 70%, 90%] percent reduction in an initial mode failure probability due to implementation of corrective action. NOTE: Initial FEF states were 5% to 95% at 10% intervals. Due to the amount of time and effort required to build a 10 state FEF CPT, SME all agreed to change to a 5 state FEF.

Model Structure Development

Bayesian Belief Network Overview

Bayesian Belief Networks (BBN) provide a framework to collect soft information and organize the data in such a way to show logical relationships between the variables whereby conditional probabilities capture the uncertainty in the dependency between the variables (Bedford & Cook, 2001; Hall, 2007; Sigurdsson, Walls, & Quigley, 2001; Walls & Quigley, 2001). BBN serves as a graphical tool representing random variables (nodes) and associated conditional dependence or independence (edges) (Smith, 1989) among the nodes. This allows one to specify a joint distribution over a set of nodes in terms of conditional distributions (Howard, 1989; Nyberg, Marcot, & Sulyma, 2006; Wilson & Huzurbazar, 2007). The nomenclature of a BBN is shown in Figure 12. The nodes are identified as A, B, and C with the edges represented by the connecting lines between the nodes. In Figure 12, node C represents the parent whereas node B is a descendant or child of C and A is a child of B. A has no descendants. Each parent has a direct influence on the child. Figure 12 shows the three probabilistic relationships that a BBN can model. Typical applications of BBN would be to collect information about B and C to make inferences about A or inferences about all three nodes given system knowledge, etc.

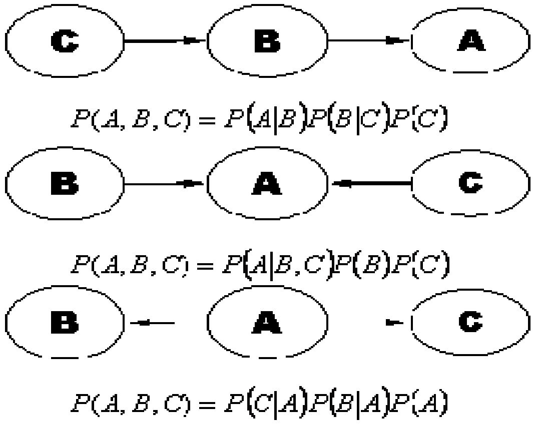


Figure 12. BBN nomenclature

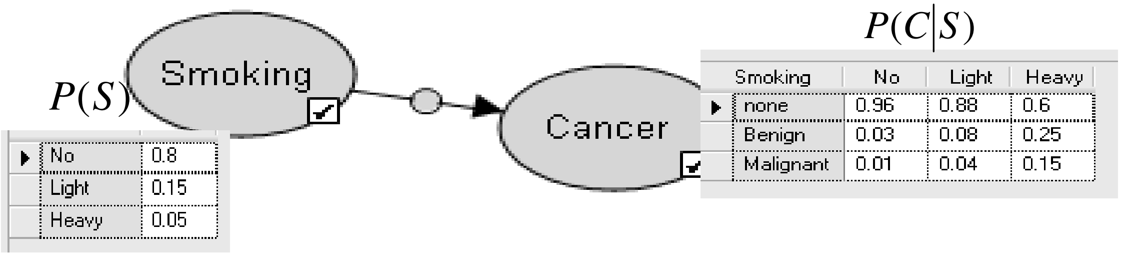
BBN Structure Development

With the identification of FEF PSF, the focus now turns to the challenge of structuring the PSF in such a way as to provide a more accurate FEF projection. Figure 13 and Figure 14 contain an overview of the training material used to expose SME to BBN methods. SME are shown how one PSF can influence another for both single and multiple parent structures. Smoking states are shown as no smoking, light and heavy with cancer states noted as none, benign and malignant. One can see in Figure 13, cancer is shown to be conditioned upon smoking. SME are shown the concepts of joint probabilities, and Bayesian updates by use of conditional probabilities.

Three methods are used to construct the BBN, M-1 expert aggregate model, M-2 fixed structure and M-3 consensus model (Hodge, Evans, Marshall, Quigley, & Walls, 2001; Tang & McCabe, 2007; Trucco, 2007). Within M-1, each SME is allowed to construct their respective BBN, organizing PSF as they wish and establish the direction of influence between nodes.

Figure 15 shows a model structure developed by one SME and the relationship among the aforementioned variables. For example, one can see FEF is conditioned on five variables, (1) technical aptitude, (2) NPI maturity, (3) resource availability, (4) management commitment, and (5) test facilities. Resource availability is conditioned on four variables, (1) failure mode complexity, (2) management commitment, (3) design complexity, and (4) test facilities. One can review Figure 15 for other conditional relationships and Appendix A for other SME M-1 models.

After each SME builds their respective structure, they must define the parent node, state probabilities, and associated child conditional probabilities. For example, the engineer that developed Figure 15 indicates the probability of the node; time of project (child) being adequate is conditioned upon the parents, design complexity, and management commitment. This SME believed the probability of an adequate time of project given design complexity is high and management commitment is low, only 20%. The process of developing CPT for PSF conditioned on other PSF allows one to capture SME knowledge as a function of a complex web of interacting performance shaping factors. This process is repeated for each SME BBN.



No	0.8
Light	0.15
Heavy	0.05

Smoking	No	Light	Heavy
none	0.96	0.88	0.6
Benign	0.03	0.08	0.25
Malignant	0.01	0.04	0.15

1. Experts determine variables of concern (smoking, cancer)
2. Establish the direction of influence among the variables (cancer is influenced by smoking)
3. Develop probability tables from expert judgment (probability of smoking and the probability of cancer given smoking)
4. Update marginal probabilities with new evidence as it becomes available (the probability of being a smoker given cancer)

$$P(C,S) = P(C|S)P(S)$$

Cancer	None	Benign	Malignant	Total
P(smoking)				
No	0.768	0.024	0.008	0.8
Light	0.132	0.012	0.006	0.15
Heavy	0.03	0.0125	0.0075	0.05
Total	0.93	0.0485	0.0215	

P(smoking)

P(cancer)

Given Evidence
Update Using Bayes Theorem

$$P(S|C) = \frac{P(C|S)P(S)}{P(C)}$$

Cancer	None	Benign	Malignant
P(smoking)			
No	0.768 / 0.930=0.826	0.024 / 0.0485=0.495	0.008 / 0.0215=0.372
Light	0.132 / 0.930=0.142	0.012 / 0.0485=0.247	0.006 / 0.0215=0.279
Heavy	0.030 / 0.930=0.032	0.0125 / 0.0485=0.258	0.0075 / 0.0215=0.349

Figure 13. BBN Training with one parent

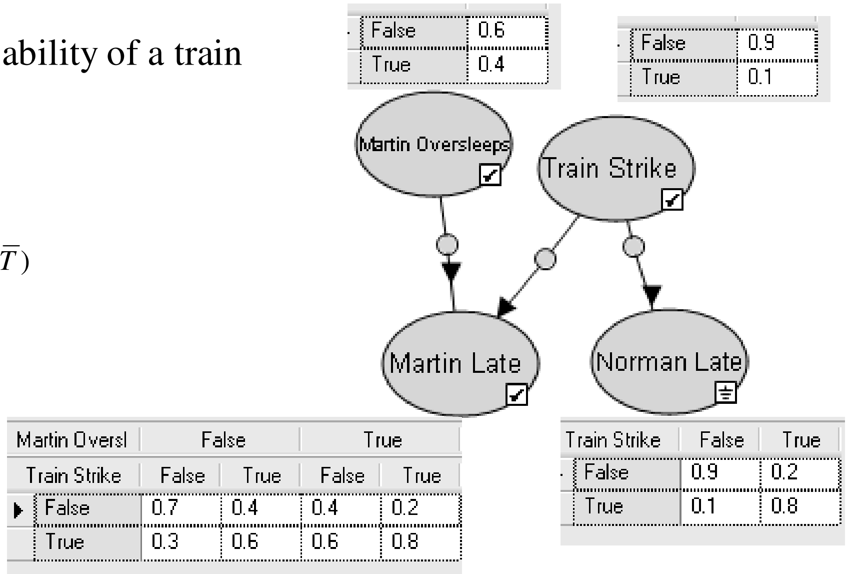
Suppose Norman is late, one may then feel the probability of a train strike has increased, but by how much?

Let N=Norman late T=Train strike

$$P(T|N) = \frac{P(N|T)P(T)}{P(N)} \text{ where } P(N) = P(N|T)P(T) + P(N|\overline{T})P(\overline{T})$$

$$P(N) = (0.8)(0.1) + (0.1)(0.9) = 0.17$$

$$P(T|N) = \frac{(0.8)(0.1)}{0.17} = 0.471$$



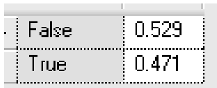
False	0.6
True	0.4

False	0.9
True	0.1

Martin Oversl	False		True	
Train Strike	False	True	False	True
False	0.7	0.4	0.4	0.2
True	0.3	0.6	0.6	0.8

Train Strike	False	True
False	0.9	0.2
True	0.1	0.8

Train Strike updated with evidence of Norman late: increased from 0.1 to 0.471



False	0.529
True	0.471

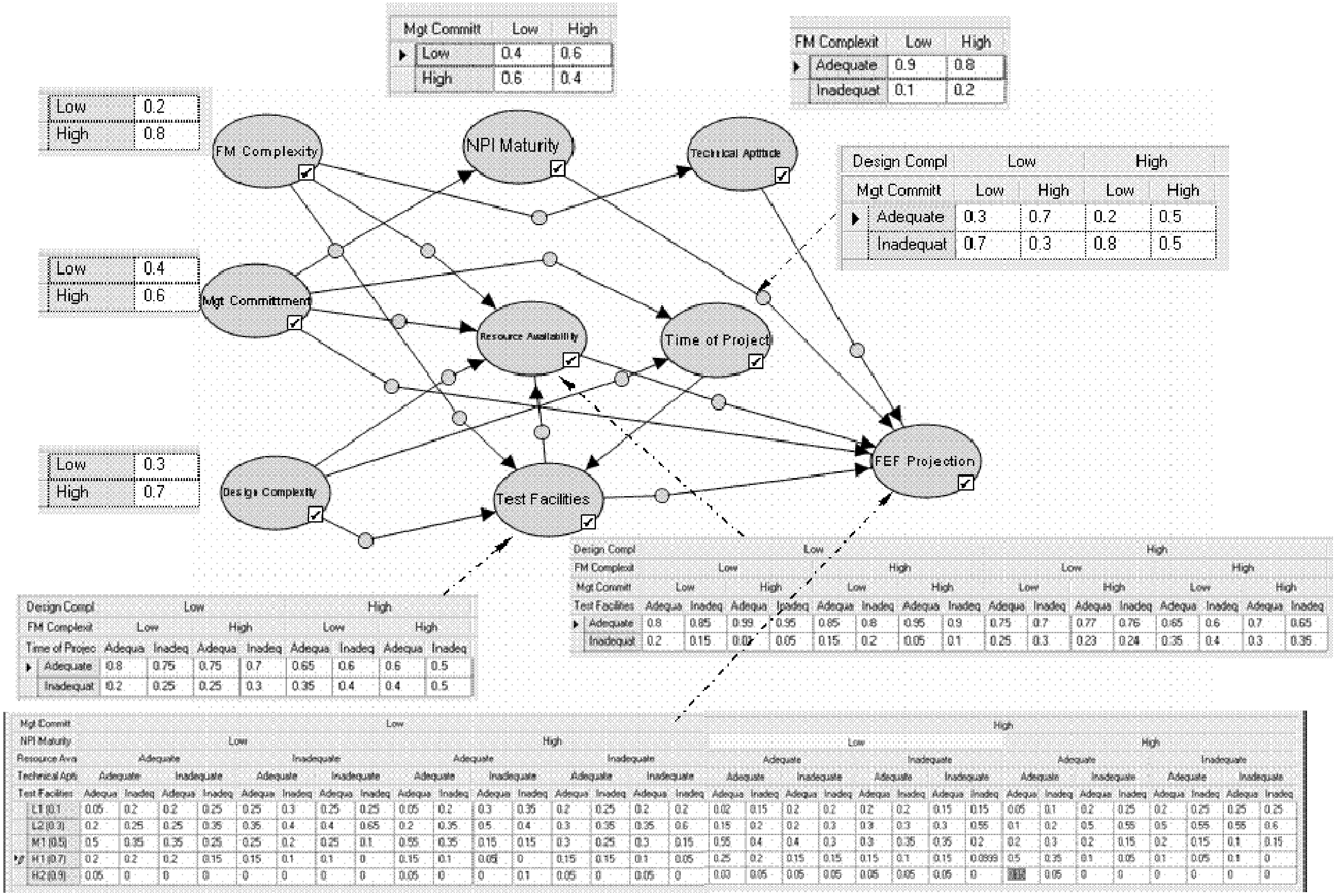
Given the evidence of Norman late and the updated likelihood of a train strike, the probability of Martin late can be determined. Let P(M)=Martin Late and P(O)=oversleeps

$$P(M) = P(M|T,O)P(T)P(O) + P(M|T,\overline{O})P(T)P(\overline{O}) + P(M|\overline{T},O)P(\overline{T})P(O) + P(M|\overline{T},\overline{O})P(\overline{T})P(\overline{O})$$

$$P(M) = (0.8)(0.471)(0.4) + (0.6)(0.471)(0.6) + (0.6)(0.529)(0.4) + (0.3)(0.529)(0.6)$$

$$P(M) = 0.542$$

Figure 14. BBN training with two parents



Low	0.2
High	0.8

Mgt Committ	Low	High
Low	0.4	0.6
High	0.6	0.4

FM Complexit	Low	High
Adequate	0.9	0.8
Inadequat	0.1	0.2

Low	0.4
High	0.6

Design Compl	Low		High	
Mgt Committ	Low	High	Low	High
Adequate	0.3	0.7	0.2	0.5
Inadequat	0.7	0.3	0.8	0.5

Low	0.3
High	0.7

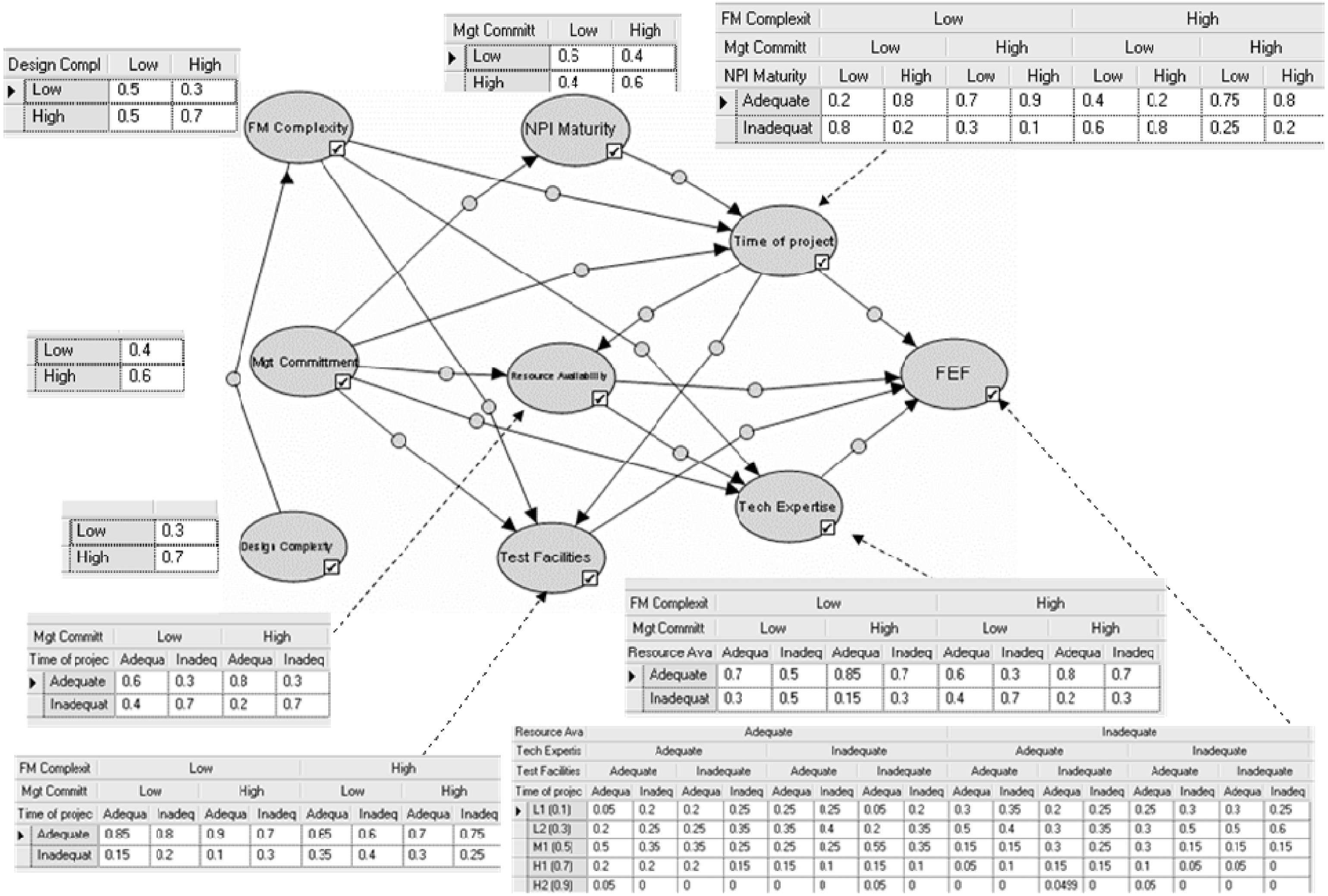
Design Compl	Low								High							
FM Complexit	Low				High				Low				High			
Mgt Commit	Low		High		Low		High		Low		High		Low		High	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.8	0.85	0.99	0.95	0.85	0.8	0.95	0.9	0.75	0.7	0.77	0.76	0.65	0.6	0.7	0.65
Inadequat	0.2	0.15	[illegible]	0.05	0.15	0.2	0.05	0.1	0.25	0.3	0.23	0.24	0.35	0.4	0.3	0.35

Design Compl	Low				High			
FM Complexit	Low		High		Low		High	
Time of Projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.8	0.75	0.75	0.7	0.65	0.6	0.6	0.5
Inadequat	0.2	0.25	0.25	0.3	0.35	0.4	0.4	0.5

Mgt Commit	Low																High															
NPI Maturity	Low								High								Low								High							
Resource Ava	Adequate				Inadequate				Adequate				Inadequate				Adequate				Inadequate				Adequate				Inadequate			
Technical Apti	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1	0.05	0.2	0.2	0.25	0.25	0.3	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.2	0.2	0.02	0.15	0.2	0.2	0.2	0.2	0.15	0.15	0.05	0.1	0.2	0.25	0.2	0.25	0.25	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.4	0.65	0.2	0.35	0.5	0.4	0.3	0.35	0.35	0.6	0.15	0.2	0.2	0.3	0.3	0.3	0.3	0.55	0.1	0.2	0.5	0.55	0.5	0.55	0.55	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.2	0.25	0.1	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.55	0.4	0.4	0.3	0.3	0.35	0.35	0.2	0.2	0.3	0.2	0.15	0.2	0.15	0.1	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.1	0	0.15	0.1	0.05	0	0.15	0.15	0.1	0.05	0.25	0.2	0.15	0.15	0.15	0.1	0.15	0.0999	0.5	0.35	0.1	0.05	0.1	0.05	0.1	0
H2 (0.9)	0.05	0	0	0	0	0	0	0	0.05	0	0	0.1	0.05	0	0.05	0	0.03	0.05	0.05	0.05	0.05	0.05	0.05	0	[illegible]	0.05	0	0	0	0	0	0

Figure 15. M-1 SME-3 FEF projection model

Within M-2 fixed structure method, SME reach consensus as to the model structure and direction of influence of all PSF. Each SME is left to determine their respective CPT for the agreed upon structure. For example, in Figure 16 one can see the team reached consensus that FEF is conditioned on four PSF, (a) time of project, (b) resource availability, (c) adequacy of test facilities, and (d) technical expertise. This same structure is used for M-3, however the SME reach consensus for the CPT given the M-2 structure. Additionally an M-4 model serves as an aggregate for M-1, M-2 and M-3, providing an estimate for FEF while addressing model uncertainty.



Design Compl	Low	High
Low	0.5	0.3
High	0.5	0.7

Mgt Committ	Low	High
Low	0.6	0.4
High	0.4	0.6

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.2	0.8	0.7	0.9	0.4	0.2	0.75	0.8
Inadequat	0.8	0.2	0.3	0.1	0.6	0.8	0.25	0.2

Low	0.4
High	0.6

Low	0.3
High	0.7

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.3	0.8	0.3
Inadequat	0.4	0.7	0.2	0.7

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.5	0.85	0.7	0.6	0.3	0.8	0.7
Inadequat	0.3	0.5	0.15	0.3	0.4	0.7	0.2	0.3

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.85	0.8	0.9	0.7	0.65	0.6	0.7	0.75
Inadequat	0.15	0.2	0.1	0.3	0.35	0.4	0.3	0.25

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.05	0.2	0.2	0.25	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.25	0.3	0.3	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.2	0.35	0.5	0.4	0.3	0.35	0.3	0.5	0.5	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.25	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.15	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.15	0.1	0.05	0.1	0.15	0.15	0.1	0.05	0.05	0
H2 (0.9)	0.05	0	0	0	0	0	0.05	0	0	0	0.0499	0	0.05	0	0	0

Figure 16. M-2 SME-5 FEF projection

Weighted Posterior Development Methodology

Collecting information from multiple SME tends to compensate the effects of heuristic faults, thus providing better predictive capability than individuals (Stiber, Small, & Pantazidou, 2004). The challenge to the analyst however is how to model the beliefs of each SME and provide an aggregate such that judgment weights correctly represent SME prior knowledge. An aggregation method is provided by Stiber et al. whereby each SME judgment is weighted by the posterior probability, given the current evidence, making the model correct for a specific problem. As evidence is collected, SME that are more consistent with reality, obtain a heavier posterior weight. Stiber et al. provides a method to extract the likelihood function from the BBN structure. For example, assume J SME are used for a study, where M_j represents the SME model j. The probability weighted average for an event probability is given by:

$$P(E) = \sum_{j=1}^{J} P(E|M_j)P(M_j) \tag{14}$$

where $P(E|M_j)$ is the probability of the event in question given the SME model *j* is correct. $P(M_j)$ represents the probability SME model j is correct. If all J SME are of equal weight in the prior, then $P^0(M_j) = \frac{1}{J}$. When evidence x is observed, one can update the probability that each SME judgment is correct given by:

$$P(M_j|x) = \frac{P(x|M_j)P^0(M_j)}{\sum_{h=1}^{J} P(x|M_h)P^0(M_h)} \tag{15}$$

The likelihood function, $P(x|M_j)$ is defined as the probability that evidence x would occur under model j. The likelihood function shown below must be evaluated N times, first for the first evidence under prior model, and again evaluated for subsequent evidence.

$$P(x|M_j) = P^0{}_j[(E_1 = e_1)]P_j[E_2 = e_2|(E_1 = e_1)]\prod_{n=1}^{N} P_j[E_n = e_n|E_1 = e_1 \cap \ldots \cap (E_{n-1} = e_{n-1})] \quad (16)$$

For example, imagine two SME, SME-3 and SME-5 noted fix effectiveness PSF, established parent and child nodes (Appendix A) with associated CPT.

Table 1 shows node marginal (true) probabilities with each SME given equal weight in the prior, 0.5 for each. Given new evidence of management commitment high, (denoted as 1), each SME judgment receives a new weight based on how close their prior model represents the actual state given the new evidence. One can see SME-3 is closer to reality, thereby the weight is adjusted from 0.5 to 0.537 and SME-5 is adjusted from 0.5 to 0.463 per equation (16). The aggregate can then be calculated using the adjusted weights.

The probability of M_j is determined per equation (15)

$$P(M_3) = \frac{(0.695)(0.5)}{(0.695)(0.5) + (0.6)(0.5)} = 0.537 \qquad P(M_5) = \frac{(0.6)(0.5)}{(0.6)(0.5) + (0.695)(0.5)} = 0.463$$

Judgment weights are used to proportion each SME contribution to the aggregate. For example consider the aggregate for quality system maturity.

$$Aggregate_{QualitySystemMaturity} = 0.537(0.66) + 0.463(0.4) = 0.540$$

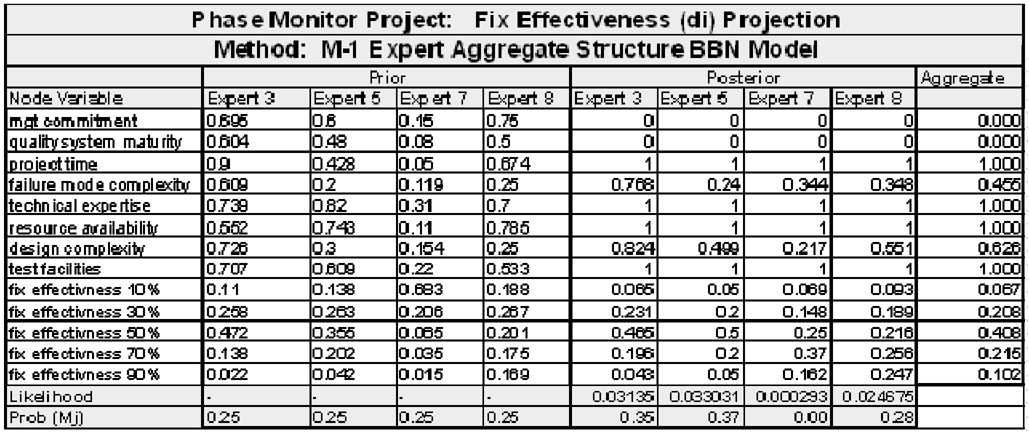
Repeat this process for all nodes. The aforementioned process is repeated for both M-1 and M-2 models with results shown in Figures 17 and 18.

M-4 *BBN* aggregate is shown in Figure 19. At this time M-3 was not available, therefore the resulting aggregate is for M-1 and M-2 only.

Table 1

Marginal Probability Event (Adequate)

	Prior		Posterior (MgtC=1)		
Node Variable	Expert 3	Expert 5	Expert 3	Expert 5	Aggregate
Mgt. commitment	0.695	0.6	1	1	1.000
Quality system maturity	0.604	0.48	0.66	0.4	0.540
Project time	0.9	0.428	0.97	0.56	0.780
Failure mode complexity	0.609	0.2	0.624	0.2	0.428
Technical expertise	0.739	0.82	0.8	0.82	0.809
Resource availability	0.552	0.74	0.75	0.77	0.759
Design complexity	0.726	0.3	0.72	0.3	0.525
Test facilities	0.707	0.609	0.71	0.619	0.668
Fix effectiveness 10%	0.11	0.139	0.109	0.118	0.113
Fix effectiveness 30%	0.258	0.263	0.247	0.239	0.243
Fix effectiveness 50%	0.472	0.355	0.483	0.347	0.420
Fix effectiveness 70%	0.138	0.202	0.014	0.242	0.187
Fix effectiveness 90%	0.022	0.042	0.021	0.054	0.036
Likelihood $P(x\mid M_j)$	-	-	0.695	0.6	
Prob (M_j)	0.5	0.5	0.537	0.463	

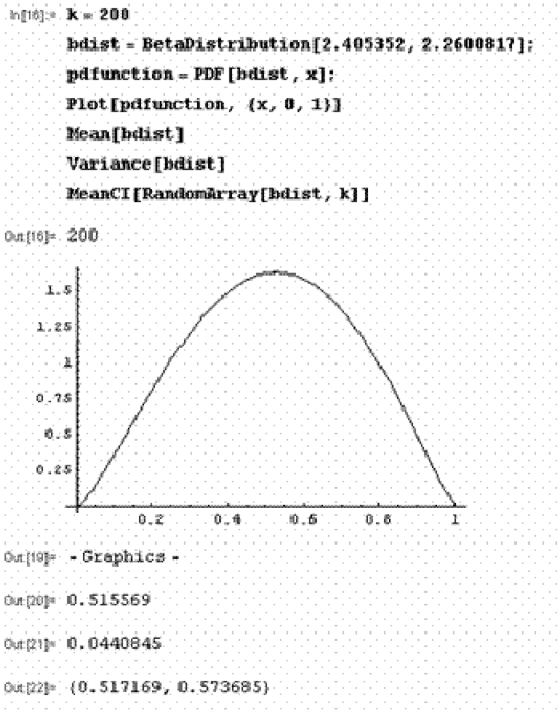


Phase Monitor Project: Fix Effectiveness (di) Projection									
Method: M-1 Expert Aggregate Structure BBN Model									
	Prior				Posterior				Aggregate
Node Variable	Expert 3	Expert 5	Expert 7	Expert 8	Expert 3	Expert 5	Expert 7	Expert 8	
mgt commitment	0.695	0.6	0.15	0.75	0	0	0	0	0.000
quality system maturity	0.604	0.48	0.08	0.5	0	0	0	0	0.000
project time	0.9	0.428	0.05	0.674	1	1	1	1	1.000
failure mode complexity	0.609	0.2	0.119	0.25	0.768	0.24	0.344	0.348	0.455
technical expertise	0.739	0.82	0.31	0.7	1	1	1	1	1.000
resource availability	0.552	0.743	0.11	0.785	1	1	1	1	1.000
design complexity	0.726	0.3	0.154	0.25	0.824	0.499	0.217	0.551	0.626
test facilities	0.707	0.609	0.22	0.533	1	1	1	1	1.000
fix effectivness 10%	0.11	0.138	0.683	0.188	0.065	0.05	0.069	0.093	0.067
fix effectivness 30%	0.258	0.263	0.206	0.267	0.231	0.2	0.148	0.189	0.208
fix effectivness 50%	0.472	0.355	0.065	0.201	0.465	0.5	0.25	0.216	0.408
fix effectivness 70%	0.138	0.202	0.035	0.175	0.196	0.2	0.37	0.256	0.215
fix effectivness 90%	0.022	0.042	0.015	0.169	0.043	0.05	0.162	0.247	0.102
Likelihood	-	-	-	-	0.03135	0.033031	0.000293	0.024675	
Prob (Mj)	0.25	0.25	0.25	0.25	0.35	0.37	0.00	0.28	

Use fix effectiveness distribution to determine mean and variance of di
Then calculate alpha and beta such that random realization for 95% CI of di
Mean $d_i = 0.516$
Variance = 0.044
95% CI for di = 0.517-0.573
Actual di for Phase monitor project = 0.58
Projection vs actual error (1% - 11%)

$$\alpha = \left(\frac{1-d_i}{FEF_{Var}}\right)\left(d_i^2\right) - d_i = 2.405$$

$$\beta = \left(\frac{d_i(1-d_i)^2}{FEF_{Var}}\right) - (1-d_i) = 2.260$$



```
In[16]:= k = 200
         bdist = BetaDistribution[2.405352, 2.2600817];
         pdfunction = PDF[bdist, x];
         Plot[pdfunction, {x, 0, 1}]
         Mean[bdist]
         Variance[bdist]
         MeanCI[RandomArray[bdist, k]]

Out[16]= 200
```

```
Out[19]= - Graphics -

Out[20]= 0.515569

Out[21]= 0.0440845

Out[22]= {0.517169, 0.573685}
```

Figure 17. M-1 expert aggregate example

Phase Monitor Project Fix Effectiveness (di) Projection

Method: M-2 Fixed Structure BBN Model

	Prior				Posterior				Aggregate
Node Variable	Expert 3	Expert 5	Expert 7	Expert 8	Expert 3	Expert 5	Expert 7	Expert 8	
mgt commitment	0.5	0.6	0.4	0.6	0	0	0	0	0.000
quality system maturity	0.45	0.52	0.46	0.64	0	0	0	0	0.000
project time	0.418	0.622	0.364	0.348	1	1	1	1	1.000
failure mode complexity	0.68	0.38	0.24	0.32	0.920	0.3	0.917	0.95	0.636
technical expertise	0.549	0.66	0.43	0.277	1	1	1	1	1.000
resource availability	0.51	0.582	0.36	0.535	1	1	1	1	1.000
design complexity	0.4	0.3	0.2	0.3	0.506	0.289	0.306	0.504	0.362
test facilities	0.541	0.736	0.365	0.344	1	1	1	1	1.000
fix effectivness 10%	0.204	0.187	0.413	0.389	0	0.05	0.0125	0	0.028
fix effectivness 30%	0.241	0.33	0.191	0.190	0.1	0.2	0.0375	0.05	0.147
fix effectivness 50%	0.256	0.316	0.133	0.093	0.15	0.5	0.05	0.1	0.333
fix effectivness 70%	0.164	0.138	0.102	0.085	0.2	0.2	0.1	0.1	0.185
fix effectivness 90%	0.136	0.019	0.16	0.105	0.56	0.05	0.8	0.75	0.307
Likelihood	-		-		0.017412	0.033763	0.006664	0.002566	
Prob (Mj)	0.25	0.25	0.25	0.25	0.29	0.53	0.11	0.04	

Use fix effectiveness distribution to determine mean and variance of di

Then calculate alpha and beta such that random realization for 95% CI of di

Mean $d_i = 0.618$

Variance = 0.068

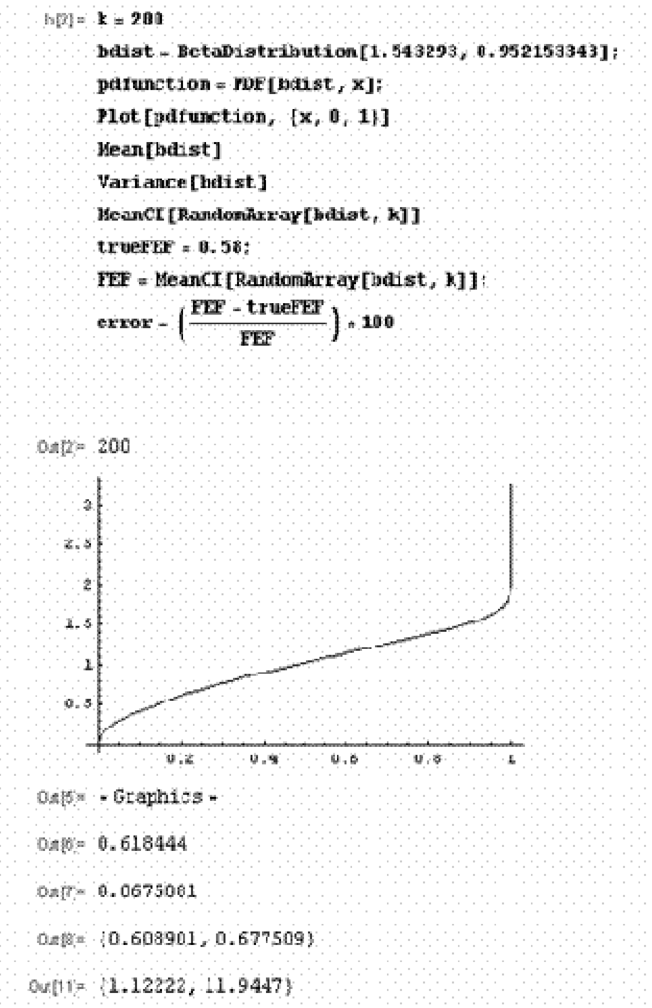
95% CI for di = 0.609-0.678

Actual di for Phase monitor project = 0.58

Projection vs actual error (4.8% - 14.5%)

$$\alpha = \left(\frac{1-d_i}{FEF_{VAR}}\right)(d_i^2) - d_i = 1.543$$

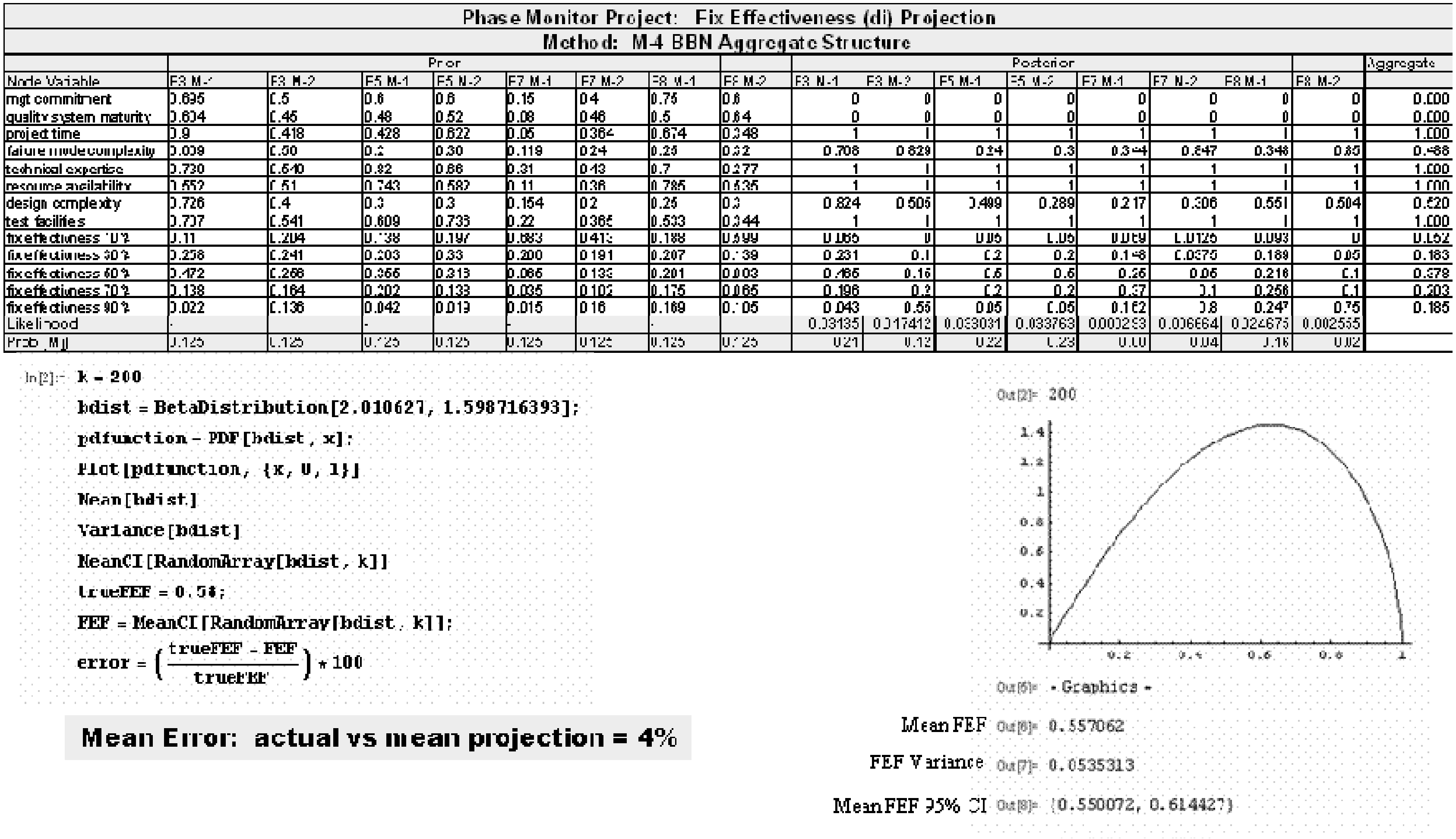
$$\beta = \left(\frac{d_i(1-d_i)^2}{FEF_{Var}}\right) - (1-d_i) = 0.952$$



```
In[2]:= k = 200
        bdist = BetaDistribution[1.543293, 0.952153343];
        pdfunction = PDF[bdist, x];
        Plot[pdfunction, {x, 0, 1}]
        Mean[bdist]
        Variance[bdist]
        MeanCI[RandomArray[bdist, k]]
        trueFEF = 0.58;
        FEF = MeanCI[RandomArray[bdist, k]];
        error = ((FEF - trueFEF)/FEF) * 100

Out[2]= 200
```

```
Out[5]= - Graphics -

Out[6]= 0.618444

Out[7]= 0.0675061

Out[8]= {0.608901, 0.677509}

Out[11]= {1.12222, 11.9447}
```

Figure 18. M-2 fixed structure example



Phase Monitor Project: Fix Effectiveness (di) Projection

Method: M-4 BBN Aggregate Structure

	Prior								Posterior								Aggregate
Node Variable	F3 M-1	F3 M-2	F5 M-1	F5 M-2	F7 M-1	F7 M-2	F8 M-1	F8 M-2	F3 M-1	F3 M-2	F5 M-1	F5 M-2	F7 M-1	F7 M-2	F8 M-1	F8 M-2	
mgt commitment	0.695	0.5	0.6	0.6	0.15	0.4	0.75	0.6	0	0	0	0	0	0	0	0	0.000
quality system maturity	0.604	0.45	0.48	0.52	0.08	0.46	0.5	0.64	0	0	0	0	0	0	0	0	0.000
project time	0.9	0.418	0.428	0.622	0.05	0.364	0.674	0.348	1	1	1	1	1	1	1	1	1.000
failure mode complexity	0.039	0.50	0.2	0.30	0.119	0.24	0.25	0.32	0.708	0.629	0.24	0.3	0.344	0.647	0.348	0.65	0.488
technical expertise	0.730	0.640	0.82	0.66	0.31	0.43	0.7	0.277	1	1	1	1	1	1	1	1	1.000
resource availability	0.552	0.51	0.743	0.582	0.11	0.36	0.785	0.635	1	1	1	1	1	1	1	1	1.000
design complexity	0.726	0.4	0.3	0.3	0.154	0.2	0.25	0.3	0.824	0.505	0.409	0.289	0.217	0.306	0.551	0.504	0.520
test facilities	0.707	0.541	0.609	0.733	0.22	0.365	0.533	0.344	1	1	1	1	1	1	1	1	1.000
fix effectiveness 10%	0.11	0.204	0.138	0.197	0.683	0.413	0.188	0.599	0.165	0	0.05	0.05	0.069	0.0125	0.093	0	0.152
fix effectiveness 30%	0.258	0.241	0.203	0.33	0.200	0.191	0.207	0.139	0.231	0.1	0.2	0.2	0.148	0.0375	0.189	0.05	0.163
fix effectiveness 50%	0.472	0.268	0.366	0.313	0.066	0.133	0.201	0.003	0.466	0.15	0.5	0.5	0.26	0.05	0.216	0.1	0.378
fix effectiveness 70%	0.138	0.164	0.202	0.133	0.035	0.102	0.175	0.065	0.196	0.2	0.2	0.2	0.37	0.1	0.256	0.1	0.203
fix effectiveness 90%	0.022	0.136	0.042	0.013	0.015	0.16	0.169	0.105	0.043	0.55	0.05	0.05	0.162	0.8	0.247	0.75	0.185
Likelihood	-		-		-		-		0.03135	0.017412	0.033031	0.033763	0.000253	0.006664	0.024675	0.002555	
Prob_Mij	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.21	0.12	0.22	0.23	0.00	0.04	0.16	0.02	

```
In[2]:= k = 200
        bdist = BetaDistribution[2.010627, 1.598716393];
        pdfunction = PDF[bdist, x];
        Plot[pdfunction, {x, 0, 1}]
        Mean[bdist]
        Variance[bdist]
        MeanCI[RandomArray[bdist, k]]
        trueFEF = 0.54;
        FEF = MeanCI[RandomArray[bdist, k]];
        error = ((trueFEF - FEF) / trueFEF) * 100
```

Mean Error: actual vs mean projection = 4%

```
Out[2]= 200
```

```
                  Out[5]= - Graphics -
Mean FEF          Out[6]= 0.557062
FEF Variance      Out[7]= 0.0535313
Mean FEF 95% CI   Out[8]= {0.550072, 0.614427}
```

Figure 19. M-4 BBN aggregate

Data Collection Methodology

Five improvement projects were chosen at random for two diverse industries, HVAC and Automotive. Warranty data were used to determine the percentage fail pre and post implementation of corrective action. BBN FEF projections were compared against actual FEF for past "fixes". BBN FEF projections were defined using four methods, (a) M-1 expert aggregate, (b) M-2 fixed structure models, (c) M-3 consensus model, and (d) M-4 BBN aggregate. Bayesian analysis was used to quantify posterior FEF model parameters based on evidence from SME.

For M-1, expert aggregate model, performed the following:

1. Using the predetermined PSF, allow each engineer to map the variables, establish direction of influence (parent-child relationships), and build CPT
2. Each SME are to enter evidence within their respective BBN model for a defined failure mode
3. Output FEF
4. Repeat steps 1-4 for SME 2….n
5. Determine FEF (d_i) mean and variance using Stiber et al. (2004)
6. Develop posterior parameters using Martz and Waller (1982)

$$\alpha = \left(\frac{1-d_i}{FEF_{VAR}}\right)\left(d_i^{\,2}\right) - d_i \tag{17}$$

$$\beta = \left(\frac{d_i(1-d_i)^2}{FEF_{Var}}\right) - (1-d_i) \tag{18}$$

7. Plot posterior Beta distribution for FEF
8. Determine 95% CI for Mean if required

For M-2, perform the following:

9. Using the predetermined PSF, allow the team of SME to reach consensus on the BBN structure
10. Each SME will build CPT within the agreed upon BBN
11. Each engineer to enter available evidence within the agreed upon BBN
12. Output FEF
13. Determine FEF (d_i) mean and variance using Stiber et al. (2004)
14. Repeat steps 9-13 for SME 2….n
15. Develop posterior parameters per equations 17-18
16. Plot posterior Beta distribution for FEF
17. Determine 95% CI for Mean FEF if required

For M-3, perform the following:

18. Using the predetermined PSF and the structure agreed upon in M-2, allow the team of SME to reach consensus on the CPT
19. Allow the team to reach consensus on the available node evidence
20. Output FEF
21. Determine FEF (d_i) mean and variance using Stiber et al. (2004)
22. Develop posterior parameters per equations 17-18
23. Plot posterior Beta distribution for FEF
24. Determine 95% CI for Mean FEF if required

For M-4, perform the following

25. Determine positive node state marginal probabilities for each SME model. Assign equal weight to all models and perform an aggregate of the models as though each were a different SME. Reference M-4 detail in Figure 19.
26. Determine FEF (d_i) mean and variance using Stiber et al. (2004)
27. Develop posterior parameters per equations 17-18
28. Plot posterior Beta distribution for FEF
29. Determine 95% CI for Mean FEF if required

Analysis Methods

Fix effectiveness projections were accomplished by entering evidence into node PSF and allowing that respective BBN to project FEF. BBN models consist of M-1 expert aggregate model, M-2 fixed structure, M-3 consensus model, M-4 BBN aggregate structure, and SME industry methods. Actual FEF was determined from warranty data for pre/post fix failure probabilities where $d_i = 1 - \frac{F_{a_i}}{F_{bi}}$. Error was calculated per

$$\breve{E} = \frac{FEF_{actual} - FEF_{projection}}{FEF_{actual}}.$$

Statistical significance of models was determined by project for each case study.

1. Power value for each experiment was held constant at 0.9. M-1 BBN model standard deviation was used as the baseline reference for sample size calculation. The difference in FEF projection we wish to detect is considered 0.05.
2. Perform random realizations per power and sample size calculations of previous step.

3. Perform test of equal variances.
4. Perform ANOVA to determine statistical significance among model means, with the null hypothesis equating to no difference in means. Perform Tukey pairwise comparisons as necessary.
5. Perform a one sample T-test comparing each model mean against the actual FEF obtained from field data.
6. Repeat analysis for the next project.

Model Qualification

Four SME were used in an initial review of M-1, expert aggregate and M-2, fixed structure FEF projection methods (Tables 2 and 3). The project of concern is called, "phase monitors." Design teams reviewed tear down data of a specific type of compressor and agreed to add phase monitors to eliminate a respective field issue. The four SME used in the analysis were on the phase monitor team, therefore they have knowledge of the team's activities, and ultimately BBN node states at the onset of the phase monitor project. Positive node state marginal probabilities are calculated for each expert. The SME agreed that management commitment was low and the "fix" team was a relatively young team with limited or low NPI (quality system) maturity. Relative to the required fix, the SME agreed that project time, technical expertise, resource availability, and test facilities were adequate. At project onset, the SME had full confidence they would solve the field issue and declared a FEF projection of 100%. After project launch, warranty data indicated the actual FEF ended up at 58%.

Table 2

Phase Monitor Project: Fix Effectiveness (di) Projection
Method: M-1 Expert Aggregate Structure BBN Model

	Prior				Posterior				
	Experts								
Node Variable	3	5	7	8	3	5	7	8	Aggregate
Mgt. commitment	0.695	0.6	0.15	0.75	0	0	0	0	0.000
Quality system maturity	0.604	0.48	0.08	0.5	0	0	0	0	0.000
Project time	0.9	0.428	0.05	0.674	1	1	1	1	1.000
Failure mode complexity	0.609	0.2	0.119	0.25	0.768	0.24	0.3744	0.348	0.455
Technical expertise	0.739	0.82	0.31	0.7	1	1	1	1	1.000
Resource availability	0.552	0.743	0.11	0.785	1	1	1	1	1.000
Design complexity	0.726	0.3	0.154	0.25	0.824	0.499	0.217	0.551	0.626
Test facilities	0.707	0.609	0.22	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.11	0.138	0.683	0.188	0.085	0.05	0.069	0.093	0.067
Fix effectiveness 30%	0.258	0.263	0.208	0.267	0.231	0.2	0.148	0.189	0.208
Fix effectiveness 50%	0.472	0.355	0.065	0.201	0.465	0.5	0.25	0.216	0.408
Fix effectiveness 70%	0.138	0.202	0.035	0.175	0.196	0.2	0.37	0.256	0.215
Fix effectiveness 90%	0.022	0.042	0.015	0.169	0.043	0.05	0.162	0.247	0.102
Likelihood $P(x\ M_j)$	-	-	-	-	0.03135	0.033031	0.000293	0.024675	
Prob (M_j)	0.25	0.25	0.25	0.25	0.35	0.37	0.00	0.28	

Table 3

Phase Monitor Project: Fix Effectiveness (di) Projection
Method: M-2 Expert Aggregate Structure BBN Model

	Prior				Posterior				
	Experts								
Node Variable	3	5	7	8	3	5	7	8	Aggre-gate
Mgt. commitment	0.5	0.6	0.4	0.6	0	0	0	0	0.000
Quality system maturity	0.45	0.52	0.46	0.64	0	0	0	0	0.000
Project time	0.418	0.622	0.364	0.348	1	1	1	1	1.000
Failure mode complexity	0.56	0.36	0.24	0.32	0.829	0.3	0.847	0.85	0.536
Technical expertise	0.549	0.66	0.43	0.277	1	1	1	1	1.000
Resource availability	0.51	0.582	0.38	0.535	1	1	1	1	1.000
Design complexity	0.4	0.3	0.2	0.3	0.505	0.289	0.306	0.504	0.382
Test facilities	0.541	0.736	0.385	0.344	1	1	1	1	1.000
Fix effective-ness 10%	0.204	0.197	0.413	0.599	0	0.05	0.0125	0	0.029
Fix effective-ness 30%	0.241	0.33	0.191	0.139	0.1	0.2	0.0375	0.05	0.147
Fix effective-ness 50%	0.256	0.316	0.133	0.093	0.15	0.5	0.05	0.1	0.333
Fix effective-ness 70%	0.164	0.138	0.102	0.085	0.2	0.2	0.1	0.1	0.185
Fix effective-ness 90%	0.136	0.019	0.16	0.105	0.55	0.05	0.8	0.75	0.307
Likelihood	-	-	-	-	0.017412	0.033763	0.006664	0.002555	
Prob (M_j)	0.25	0.25	0.25	0.25	0.29	0.56	0.11	0.04	

Let's review the process for determining likelihood, probability that SME-3's model is correct and the method to obtain an aggregate of all SME. Note in Table 1, SME-3 indicated management commitment is low, thus P(0)=1-P(1)=1-0.695=0.305. Quality management system is low = 1-0.604=0.396. The probability of project time adequate = 0.9, probability of technical expertise adequate is 0.739, probability of resource availability adequate is 0.552, and the probability of test facilities being adequate is 0.707. Using equation 16, the likelihood of the evidence is the product of the aforementioned data, thus the likelihood is:

$$P(x|M_3) = (1-0.695)(1-0.604)(0.9)(0.739)(0.552)(0.707) = 0.03135$$

This process is repeated for each *SME* yielding:

$$P(x|M_5) = (1-0.6)(1-0.48)(0.428)(0.82)(0.743)(0.609) = 0.033031$$

$$P(x|M_7) = (1-0.15)(1-0.08)(0.05)(0.31)(0.11)(0.22) = 0.000293$$

$$P(x|M_8) = (1-0.75)(1-0.5)(0.674)(0.7)(0.785)(0.533) = 0.024675$$

Equation 15 is used to determine the probability that *SME-3* model is correct.

$$P(M_3|x) = \frac{P(x|M_3)P^0(M_3)}{\sum_{h=1}^{J} P(x|M_h)P^0(M_h)} =$$

$$= \frac{(0.03135)(0.25)}{(0.03135)(0.25) + (0.033031)(0.25) + (0.000293)(0.25) + (0.024675)(0.25)} = 0.35$$

This process is repeated for each *SME* resulting in the following:

$$P(M_5|x) = 0.37 \quad P(M_7|x) = 0.0 \quad P(M_8|x) = 0.28$$

Marginal probabilities are calculated for each node given management commitment is low, quality system maturity low, project time, technical expertise, resource availability, and test facilities are all adequate. The aggregate is the sum of the marginal probability times the weight for each respective SME. For example, reference the node failure mode complexity, noting the marginal probability of 0.768, 0.24, 0.344 and 0.348 for SME 3, 5, 7, and 8 respectively.

$$Aggregate_{FMComplexity} = (0.35)(0.768) + (0.37)(0.24) + (0.0)(0.344) + (0.28)(0.348) = 0.455$$

The aggregate process is repeated for all nodes.

Model Calibration

During the development of PSF, SME collected a list of areas they feel impact, either positively or negatively, an engineering community's ability to fix a failure mode. SME have developed model structures and associated conditional probabilities that provide fix effectiveness prediction. How close SME are to reality is dependent on numerous factors, but in the end, their FEF prediction is representative of their life's experiences. These experiences may or may not be the same as other SME within their team. These differences ultimately lead to differing fix effectiveness predictions; therefore, it becomes imperative to develop a method to adjust SME perspective to that of reality. Gradient ascent methodology was chosen to tweak or update CPT in the presence of PSF evidence.

Gradient ascent is an optimization algorithm that allows one to find a local minimum or maximum (Friedman & Goldszmidt, 1998; Gueston, 2007; Hsu, 1999). Consider the joint distribution of a *BBN* over all variables by:

$$P(y_1....y_n) = \prod_{i=1}^{n} P(y_i | Parents(Y_i)) \tag{19}$$

where Y_i denotes the immediate predecessor (parent) of yi. Conditional gradient ascent can be used to learn *CPT* by converging the network such that the probability of the data given the network is maximized. Let w_{ijk} denote one cell in a CPT for variable y_i in the network under evaluation. According to Gueston, CPT updates can be performed per equation 20 whereby a specific cell is increased or decreased based on parent and child conditional evidence.

$$w_{ijk} \leftarrow w_{ijk} + \eta \sum_{d \in D} \frac{P\left(y_{ij}, u_{ik} \middle| data\right)}{w_{ijk}} \tag{20}$$

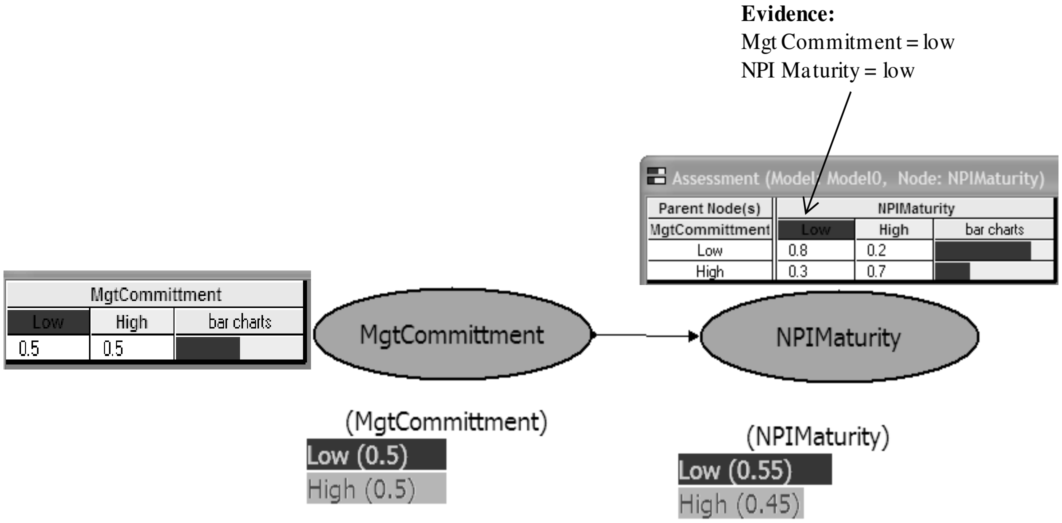
$$\sum w_{ijk} = 1 \qquad 0 \le w_{ijk} \le 1 \qquad \eta = \text{learning rate}$$

Since this research attempted to reduce FEF error prediction based on field data, we do not require an optimization of the data given the structure. The algorithm was used only as an error reduction algorithm regardless as to whether PSF evidence was positive or negative; therefore a slight modification of equation 20 was required. The modification (equation 21) involved using the child marginal probability instead of the child probability given the evidence. This allows CPT to be adjusted such that PSF marginal's move up or down depending on evidence being positive or negative otherwise no downward adjustments could be made in the presence of negative PSF.

$$w_{ijk} \leftarrow w_{ijk} + \eta \sum_{d \in D} \frac{P\left(y_{ij}\right) P\left(u_{ik} \middle| data\right)}{w_{ijk}} \tag{21}$$

Figure 20 provides a simple two node *BBN. NPI* maturity was the child node conditioned on the parent node management commitment. Positive state marginal probabilities are 0.5

and 0.45 for management commitment and NPI maturity respectively and we have evidence that management commitment and NPI maturity are both low. We can update that CPT given the evidence, i.e. that cell contains an SME value of 0.8. The marginal probability NPI maturity was low – equals 0.55, whereas the probability management commitment was low given NPI maturity was low – equals 0.727273 (Figure 21).



Evidence:
Mgt Commitment = low
NPI Maturity = low

Assessment (Model: Model0, Node: NPIMaturity)

Parent Node(s)	NPIMaturity		
MgtCommittment	Low	High	bar charts
Low	0.8	0.2	
High	0.3	0.7	

MgtCommittment		
Low	High	bar charts
0.5	0.5	

Figure 20. Gradient ascent methodologies

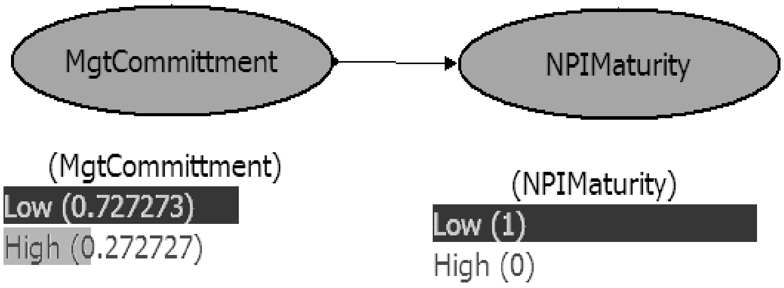
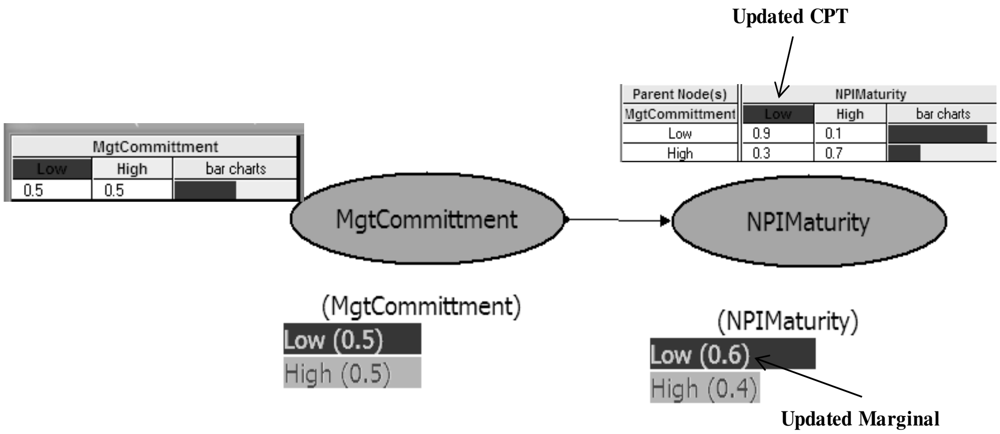


Figure 21. Parent marginal given evidence

Using equation 21, the CPT was updated from 0.8 to 0.9. The resulting marginal probability moved up from 0.55 to 0.6 with the likelihood more closely representing reality. This process was repeated for each PSF node with evidence throughout the BBN (Figure 22).

$$w_{ijk} = 0.8 + 0.2\frac{(0.55)(0.727273)}{0.8} = 0.9$$



Updated CPT

MgtCommittment		
Low	High	bar charts
0.5	0.5	

Parent Node(s)	NPIMaturity		
MgtCommittment	Low	High	bar charts
Low	0.9	0.1	
High	0.3	0.7	

Updated Marginal

Figure 22. Updated marginal probability

Chapter 4: Testing and Confirmation

Chapter Summary

Four different Bayesian Belief Network models were developed by subject matter experts at a major HVAC organization. The first BBN developed was the M-1Expert Aggregate Model. SME were allowed to assign direction of influence among PSF and assign their respective node CPT. The second BBN developed was M-2 Fixed Structure. SME reached consensus on the model structure and then developed their respective CPT. M-3, consensus model, involved SME reaching consensus on CPT for the fixed structure developed during M-2. M-4, BBN aggregate was simply an aggregate of M-1, M-2, and M-3.

SME collected PSF evidence from past projects and entered that information into each of the aforementioned BBN and predicted fix effectiveness. BBN projections were then compared against known FEF and SME projections of FEF at the onset of each project. FEF error was calculated in two ways, first, BBN FEF projection versus actual and second, SME FEF projection versus actual. The same BBN and associated CPT were used in an automotive organization to once again determine BBN FEF projection versus actual and SME FEF projection versus actual.

Table 4 shown below, indicates that BBN methodology provide less overall FEF projection error than that of SME as predicted at the onset of a product development project. M-4, BBN aggregate provide the least overall error at 32%, 59% less than SME.

M-1 expert aggregate and M-3 consensus BBN, exhibit 34% overall error, 57% less than SME. M-2, fixed structure BBN, exhibits 35% overall error, 56% less than SME.

Table 4

Percent Error by Model Type

Variable	% Error					
	M-1	M-2	M-3	M-3	BBN Avg	SME
HVAC	20	22	24	22	22	62
Auto	47	48	44	43	45	96
Overall	34	35	34	32	34	79

The net results of all this is the fact that HVAC BBN FEF projections reduced model error by 65% while Automotive BBN FEF projections reduced projection error by 53%, thus providing empirical evidence that Generic PSF developed by one industry can be used in another to provide substantially less FEF projection error than projections made by SME within that industry. Error reduction methodology will be further expanded in Chapter 6, by exploring the dependency among fix activities and their influence on FEF projection, failure intensity and ultimately MTBF projections.

The following pages in this chapter contain explicit detail and analysis of each model within each project.

HVAC Test Case

HVAC SME reviewed past projects and collected PSF evidence and FEF that would have been known or “perceived” to be known at the onset of each project (Table 5). For example, evidence for project-1 indicates management commitment and quality system maturity are low, project time, technical expertise, resource availability and test facilities are adequate. Unknown at project onset are how complex the failure mode may be and how many design iterations it may encompass. SME feel confident in their ability to solve this particular failure mode and assign a FEF of 100% to the project.

Table 5

HVAC Project Evidence

	HVAC Project Data				
Variable	Project-1	Project-2	Project-3	Project-4	Project-5
Mgt. commitment	0	1	0	1	1
Quality system maturity	0	1	0	1	0
Project time	1	?	0	0	0
Failure mode complexity	?	0	?	1	1
Technical expertise	1	?	?	?	1
Resource Availability	1	1	?	0	0
Design complexity	?	0	1	1	?
Test facilities	1	1	0	0	?
SME Proposed FEF	100%	90%	60%	99.8%	72%
Actual FEF	58%	70%	30%	62%	49%

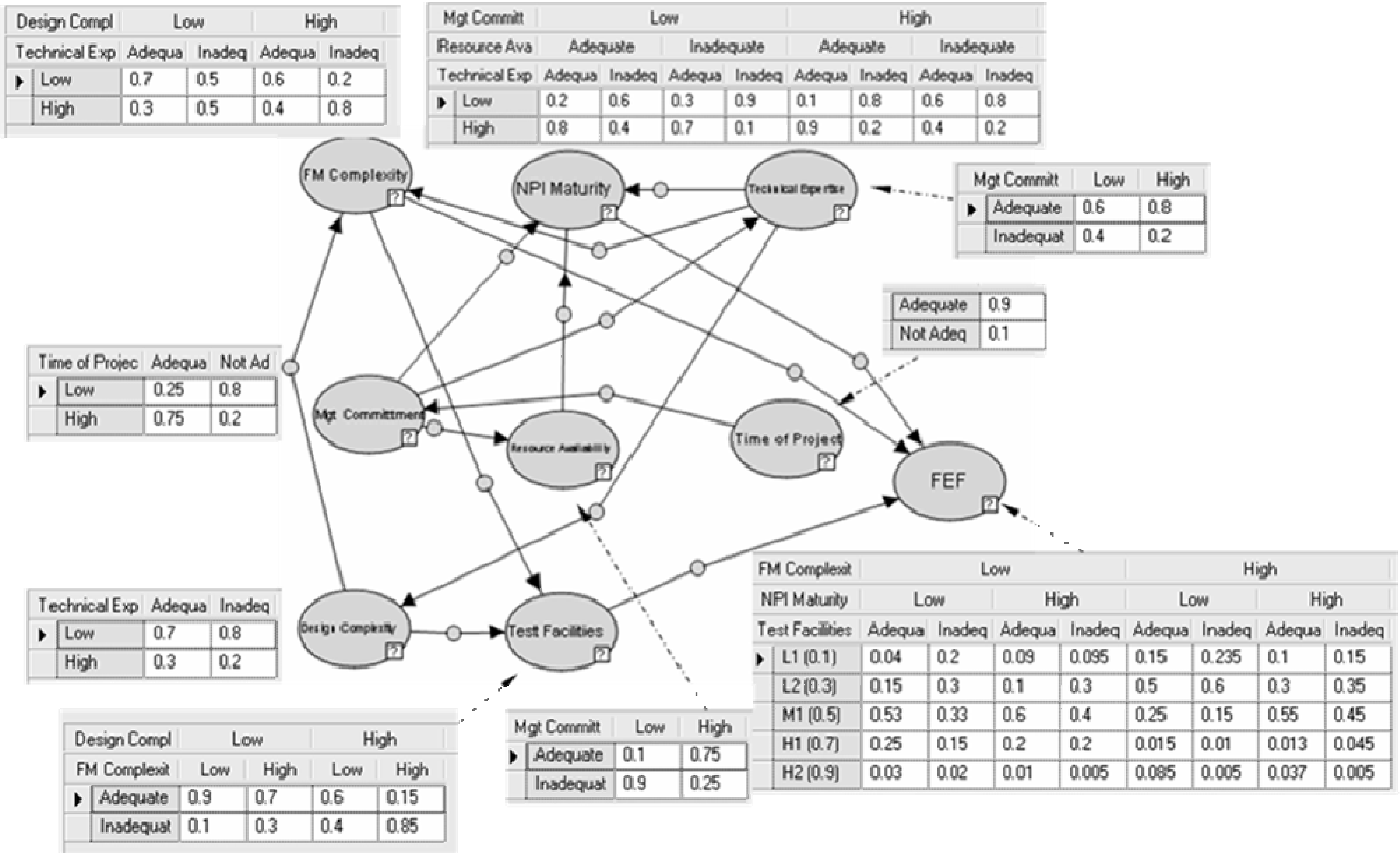
Note: ? – unobserved; 0 – negative evidence; 1 – positive evidence

HVAC M-1 Expert Aggregate

The first BBN developed is the M-1Expert Aggregate Model. SME were allowed to assign direction of influence among PSF and assign their respective node CPT. Figures 23 and 24 are examples of SME-3 and SME-7 M-1 models. Note how SME-3 believed FEF was a function of FM complexity, NPI maturity, and test facilities whereas SME-7 believed FEF was a function of FM complexity, resource availability, technical expertise, and test facilities (Appendix A).

PSF evidence was entered into each SME M-1 model allowing for an FEF projection. In addition an aggregate model was developed by calculating how close the marginal probability of each SME was to the observed evidence allowing a posterior SME judgment weight to be established (Stiber et al., 2004). For project-1 noted earlier, one can see in Table 6 the SME weighted posterior calculated to be 35% for SME-3, 37% for SME-5, 0% for SME-7 and 28% for SME-8. The M-1 BBN model provided a more accurate FEF projection than that of SME at project onset and projects FEF of 52.5% vs actual SME projections of 100% with projection error of 9% vs actual SME FEF projection error of 72%. Subsequent FEF projections are found in Table 7 with errors noted in Table 8.

In four of five FEF projections, the M-1 aggregate model proved to be more accurate than SME initial projections. Overall the M-1 aggregate BBN exhibited an average error of 19.6% vs an average SME error of 61.8%. The M-1 aggregate model exhibits 66.4% less error. Details can be found in Appendix B.



Design Compl	Low		High	
Technical Exp	Adequa	Inadeq	Adequa	Inadeq
Low	0.7	0.5	0.6	0.2
High	0.3	0.5	0.4	0.8

Mgt Committ	Low				High			
Resource Ava	Adequate		Inadequate		Adequate		Inadequate	
Technical Exp	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Low	0.2	0.6	0.3	0.9	0.1	0.8	0.6	0.8
High	0.8	0.4	0.7	0.1	0.9	0.2	0.4	0.2

Mgt Committ	Low	High
Adequate	0.6	0.8
Inadequat	0.4	0.2

Adequate	0.9
Not Adeq	0.1

Time of Projec	Adequa	Not Ad
Low	0.25	0.8
High	0.75	0.2

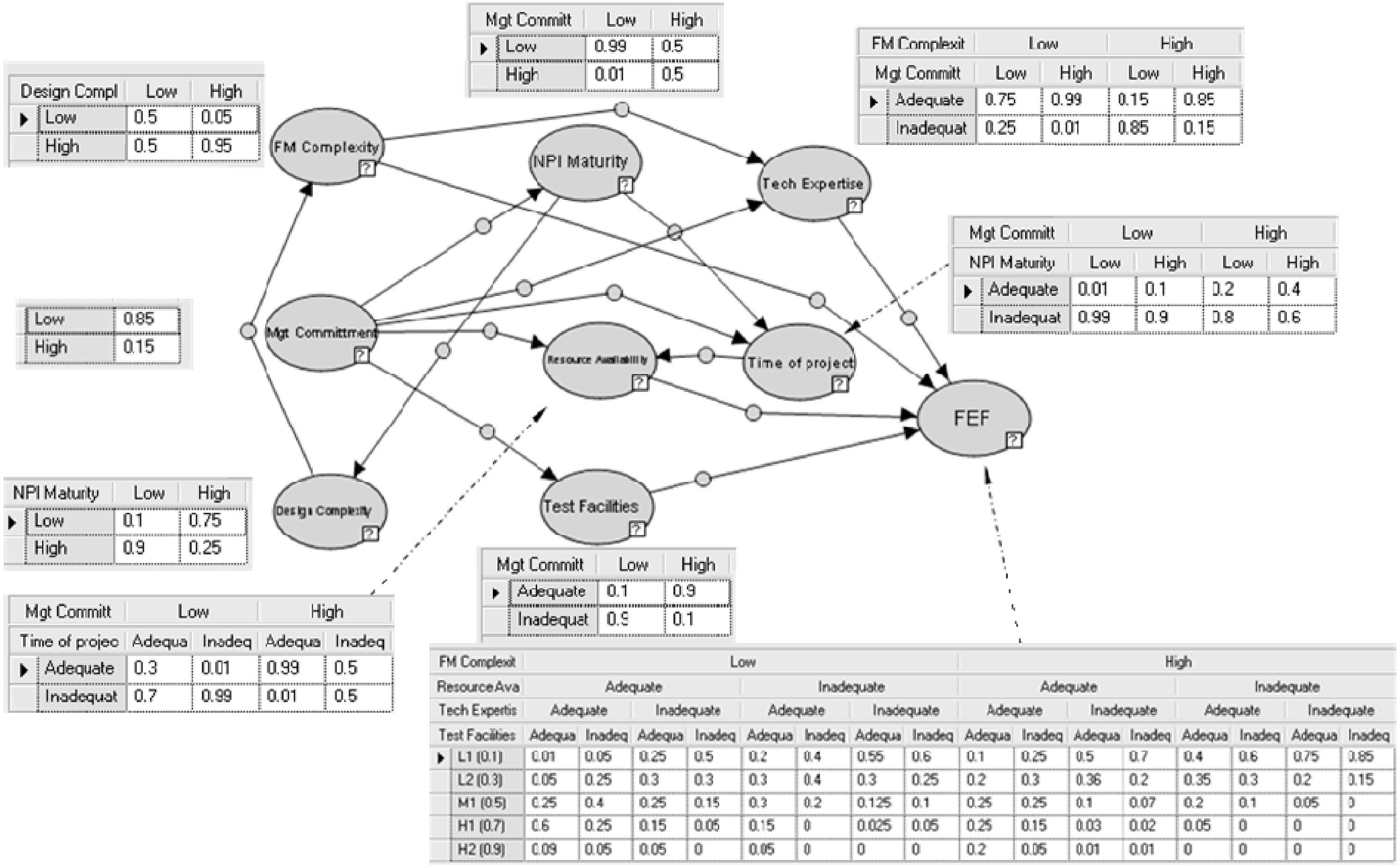
Technical Exp	Adequa	Inadeq
Low	0.7	0.8
High	0.3	0.2

FM Complexit	Low				High			
NPI Maturity	Low		High		Low		High	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.04	0.2	0.09	0.095	0.15	0.235	0.1	0.15
L2 (0.3)	0.15	0.3	0.1	0.3	0.5	0.6	0.3	0.35
M1 (0.5)	0.53	0.33	0.6	0.4	0.25	0.15	0.55	0.45
H1 (0.7)	0.25	0.15	0.2	0.2	0.015	0.01	0.013	0.045
H2 (0.9)	0.03	0.02	0.01	0.005	0.085	0.005	0.037	0.005

Design Compl	Low		High	
FM Complexit	Low	High	Low	High
Adequate	0.9	0.7	0.6	0.15
Inadequat	0.1	0.3	0.4	0.85

Mgt Committ	Low	High
Adequate	0.1	0.75
Inadequat	0.9	0.25

Figure 23. M-1 SME-3 BBN



Design Compl	Low	High
Low	0.5	0.05
High	0.5	0.95

Mgt Commit	Low	High
Low	0.99	0.5
High	0.01	0.5

FM Complexit	Low		High	
Mgt Committ	Low	High	Low	High
Adequate	0.75	0.99	0.15	0.85
Inadequat	0.25	0.01	0.85	0.15

Mgt Committ	Low		High	
NPI Maturity	Low	High	Low	High
Adequate	0.01	0.1	0.2	0.4
Inadequat	0.99	0.9	0.8	0.6

Low	0.85
High	0.15

NPI Maturity	Low	High
Low	0.1	0.75
High	0.9	0.25

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.3	0.01	0.99	0.5
Inadequat	0.7	0.99	0.01	0.5

Mgt Committ	Low	High
Adequate	0.1	0.9
Inadequat	0.9	0.1

FM Complexit	Low								High							
ResourceAva	Adequate				Inadequate				Adequate				Inadequate			
Tech Expertis	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.01	0.05	0.25	0.5	0.2	0.4	0.55	0.6	0.1	0.25	0.5	0.7	0.4	0.6	0.75	0.85
L2 (0.3)	0.05	0.25	0.3	0.3	0.3	0.4	0.3	0.25	0.2	0.3	0.36	0.2	0.35	0.3	0.2	0.15
M1 (0.5)	0.25	0.4	0.25	0.15	0.3	0.2	0.125	0.1	0.25	0.25	0.1	0.07	0.2	0.1	0.05	0
H1 (0.7)	0.6	0.25	0.15	0.05	0.15	0	0.025	0.05	0.25	0.15	0.03	0.02	0.05	0	0	0
H2 (0.9)	0.09	0.05	0.05	0	0.05	0	0	0	0.2	0.05	0.01	0.01	0	0	0	0

Figure 24. M-1 SME-7 BBN

Table 6

Project-1 Results: HVAC Fix Effectiveness (di) Projection – Method: M-1 SME Models

	Prior				Posterior				Aggregate				
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8					
Mgt. commitment	0.695	0.600	0.150	0.750	0	0	0	0	0.000				
Quality system maturity	0.604	0.480	0.084	0.500	0	0	0	0	0.000				
Project time	0.900	.0428	0.054	0.674	1	1	1	1	1.000				
Failure mode complexity	0.610	0.200	0.119	0.250	0.769	0.240	0.344	0.407	0.472				
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000				
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000				
Design complexity	0.726	0.300	0.154	0.250	0.824	0.499	0.217	0.555	0.627				
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000				
Fix effective-ness 10%	0.110	0.139	0.683	0.188	0.065	0.050	0.069	0.057	0.057				
Fix effective-ness 30%	0.258	0.267	0.202	0.268	0.231	0.200	0.148	0.162	0.200				
Fix effective-ness 50%	0.472	0.345	0.066	0.201	0.485	0.500	0.250	0.227	0.412				
Fix effective-ness 70%	0.138	0.205	0.035	0.175	0.196	0.200	0.370	0.277	0.220				
Fix effective-ness 90%	0.023	0.043	0.014	0.169	0.043	0.050	0.162	0.277	0.111				
Likelihood	-	-	-		3.13E-02	3.30E-02	3.12E-04	2.47E-02		Mean	Variance	Alpha	Beta
Prob (M_j)	0.25	0.25	0.25	0.25	0.35	0.37	0.00	0.28		0.525	0.043	2.517	2.275

Table 7

HVAC M-1 FEF Projections by Model Type

	FEF	M-1				M-1	SME
Project	Observed	SME-3	SME-5	SME-7	SME-8	Aggregate	
1	58	48	50	58	61	52	100
2	70	42	56	50	36	45	90
3	30	32	38	20	54	24	60
4	62	46	51	57	64	52	100
5	49	44	30	26	53	42	72

Table 8

HVAC M-1 Projection Error by Model Type

	M-1				M-1	SME
Project	SME-3	SME-5	SME-7	SME-8	Aggregate	
1	17%	14%	0%	-5%	9%	-72%
2	40%	19%	28%	49%	36%	-29%
3	-7%	-28%	34%	-79%	21%	-100%
4	25%	18%	8%	-3%	18%	-61%
5	9%	39%	47%	-8%	14%	-47%

HVAC M-2 Fixed Structure

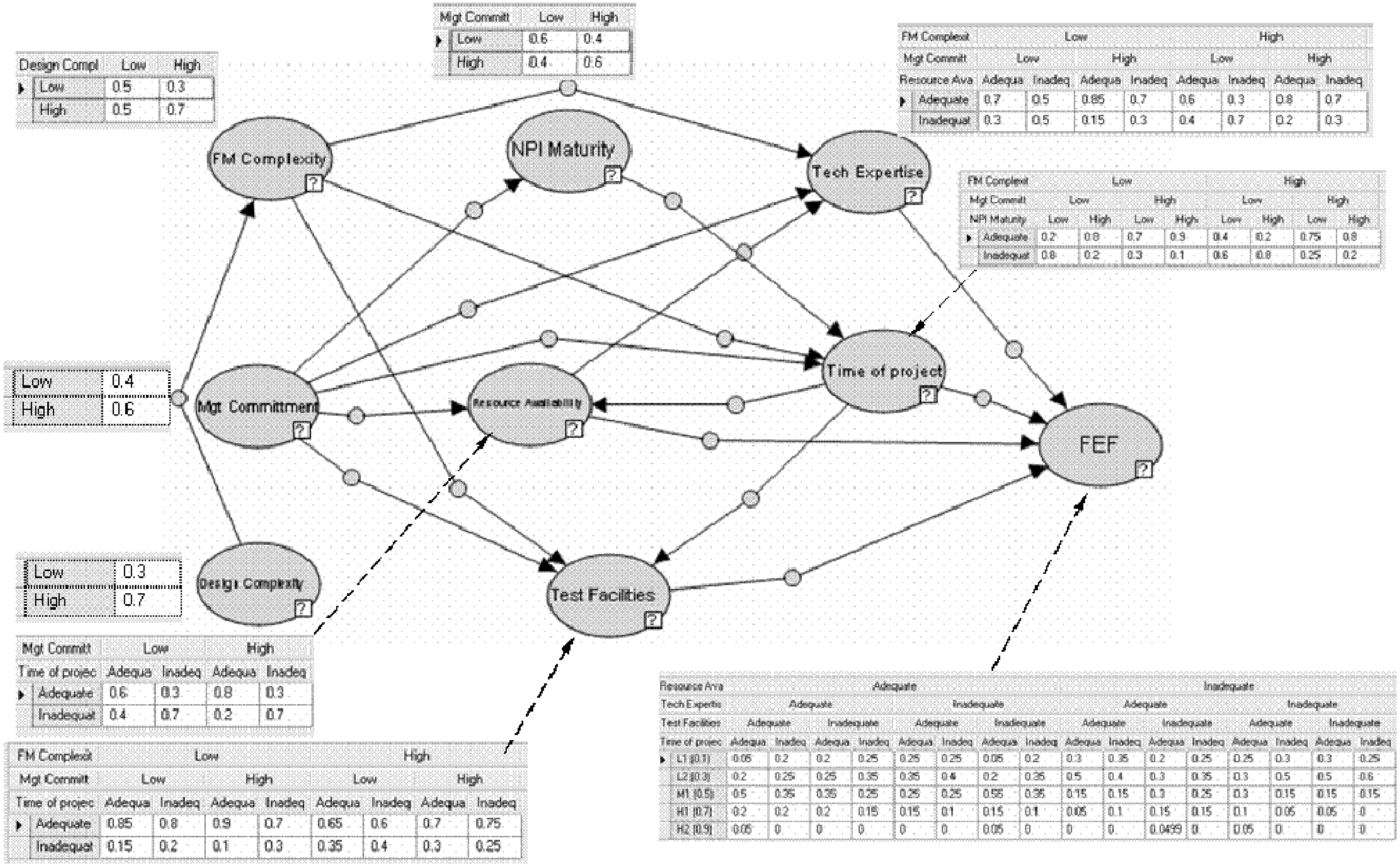
The second BBN developed was M-2 Fixed Structure. SME reached consensus on the model structure and then developed their respective CPT. After numerous spirited debates the team agreed to the model structure (Figure 25). Each SME used their judgment to populate CPT for each PSF (Appendix C). This structure noted FEF to be a function of resource availability, technical expertise, test facilities, and project time.

The same PSF evidence (Table 4) used in M-1 were input into the fixed structure M-2 models. Weighted posterior SME judgments were determined via Stiber et al. (2004). For project-1 SME-3 judgment weight was 28%, SME-5 57%, SME-7 11%, and SME-8 weight was 4% (Table 9). The M-2 BBN model provided a more accurate FEF projection than that of SME at project onset. The M-2 aggregate model projected a FEF of 61.5% vs actual SME projection of 100% with projection error of 6% vs actual SME FEF projection error of 72%. Subsequent project FEF projections can be found in Table 10 with errors noted in Table 11.

In all five FEF projections, the M-2 aggregate model proved to have equal or better FEF projection accuracy than the SME initial projections. Overall the M-2 aggregate BBN exhibited an average error of 21.8% vs an SME projection error of 61.8%. The M-2 fixed structure exhibited 48.6% less error (Table 11). Project details can be found in Appendix D.

HVAC M-3 Consensus Model

The third BBN developed was M-3. SME use the structure developed in M-2 and reach consensus on node CPT (Appendix E, Appendix F). Again, numerous spirited debates erupted among the SME. Ultimately they agreed to create one CPT by averaging their respective node M-2 CPT (Figure 26).



Design Compl	Low	High
Low	0.5	0.3
High	0.5	0.7

Mgt Commit	Low	High
Low	0.6	0.4
High	0.4	0.6

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.5	0.85	0.7	0.6	0.3	0.8	0.7
Inadequat	0.3	0.5	0.15	0.3	0.4	0.7	0.2	0.3

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.2	0.8	0.7	0.9	0.4	0.2	0.75	0.8
Inadequat	0.8	0.2	0.3	0.1	0.6	0.8	0.25	0.2

Low	0.4
High	0.6

Low	0.3
High	0.7

Mgt Commit	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.3	0.8	0.3
Inadequat	0.4	0.7	0.2	0.7

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.85	0.8	0.9	0.7	0.65	0.6	0.7	0.75
Inadequat	0.15	0.2	0.1	0.3	0.35	0.4	0.3	0.25

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.05	0.2	0.2	0.25	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.25	0.3	0.3	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.2	0.35	0.5	0.4	0.3	0.35	0.3	0.5	0.5	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.25	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.15	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.15	0.1	0.05	0.1	0.15	0.15	0.1	0.05	0.05	0
H2 (0.9)	0.05	0	0	0	0	0	0.05	0	0	0	0.0499	0	0.05	0	0	0

Figure 25. M-2 SME-5 BBN

Table 9

Project-1 Results: HVAC Fix Effectiveness (di) Projection
Method: M-2 SME Models

	Prior				Posterior				Aggregate			
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8				
Mgt. commitment	0.500	0.600	0.400	0.600	0	0	0	0	0.000			
Quality system maturity	0.450	0.520	0.460	0.640	0	0	0	0	0.000			
Project time	0.418	0.629	0.364	0.348	1	1	1	1	1.000			
Failure mode complexity	0.560	0.460	0.240	0.320	0.829	0.394	0.847	0.850	0.583			
Technical expertise	0.549	0.670	0.430	0.277	1	1	1	1	1.000			
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000			
Design complexity	0.400	0.300	0.200	0.300	0.505	0.289	0.306	0.504	0.360			
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000			
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.000	0.050	0.013	0.000	0.030			
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.100	0.200	0.038	0.050	0.148			
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.150	0.500	0.050	0.100	0.337			
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.200	0.200	0.100	0.100	0.185			
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.550	0.050	0.800	0.750	0.299			
Likelihood	-	-	-		1.74E-02	3.56E-02	6.70E-03	2.56E-03	Mean	Var.	Alpha	Beta
Prob (M_j)	0.25	0.25	0.25	0.25	0.28	0.57	0.11	0.04	0.6151	0.0528	2.1447	1.34213

Table 10

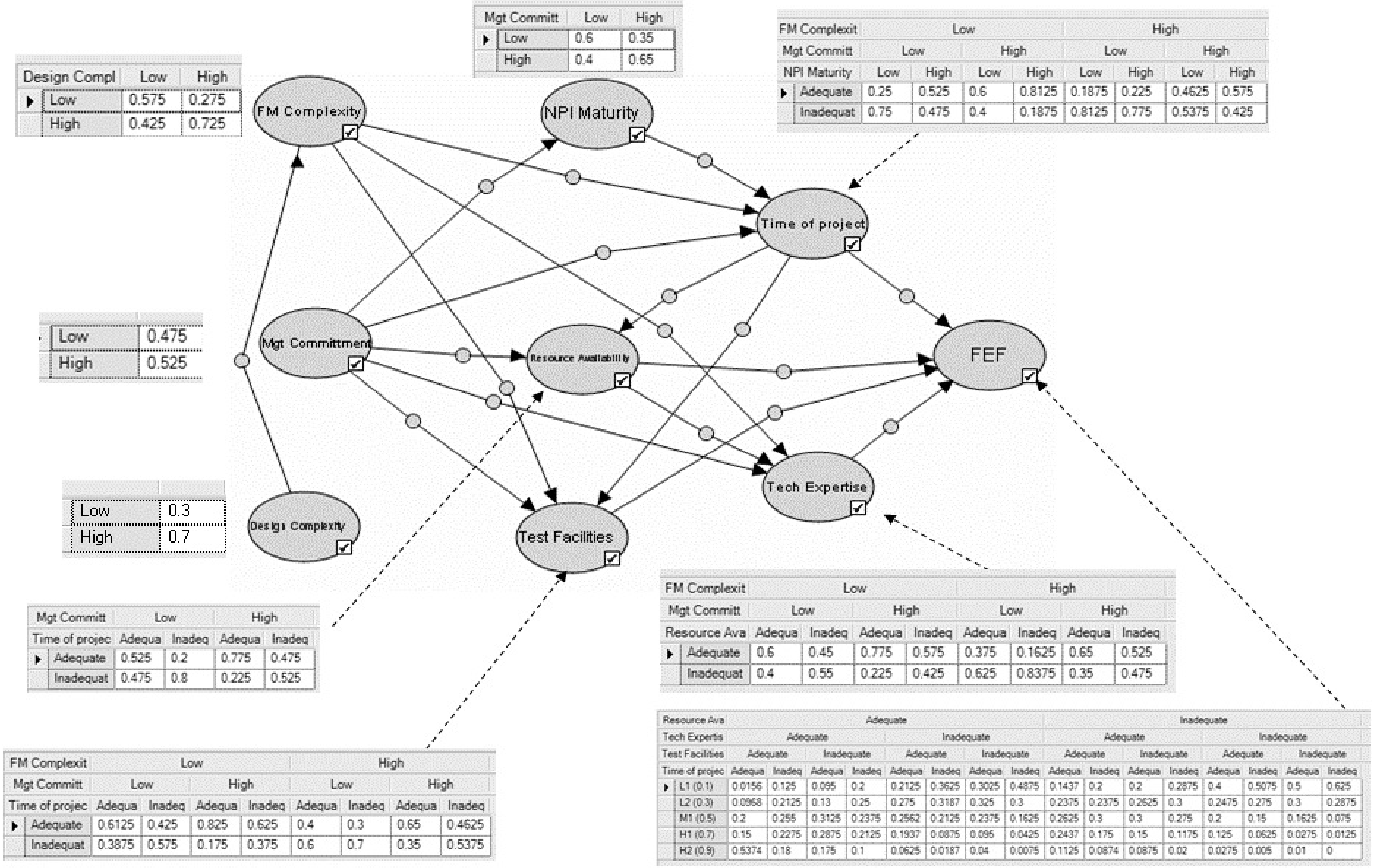
M-2 HVAC FEF Projections by Model Type

	FEF	M-2				M-2	SME
Project	Observed	SME-3	SME-5	SME-7	SME-8	Aggregate	
1	58	74	50	83	81	62	100
2	70	54	47	75	45	51	90
3	30	30	33	20	18	24	60
4	62	69	48	79	52	56	100
5	49	30	34	33	18	26	72

Table 11

HVAC M-2 FEF Projection Error by Model Type

	M-1				M-1	SME
Project	SME-3	SME-5	SME-7	SME-8	Aggregate	
1	-28%	14%	-43%	-40%	-6%	-72%
2	22%	34%	-8%	36%	27%	-29%
3	-1%	-9%	33%	41%	19%	-100%
4	-11%	23%	-28%	14%	10%	-61%
5	38%	31%	33%	62%	47%	-47%



Design Compl	Low	High
Low	0.575	0.275
High	0.425	0.725

Mgt Committ	Low	High
Low	0.6	0.35
High	0.4	0.65

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.25	0.525	0.6	0.8125	0.1875	0.225	0.4625	0.575
Inadequat	0.75	0.475	0.4	0.1875	0.8125	0.775	0.5375	0.425

Low	0.475
High	0.525

Low	0.3
High	0.7

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.525	0.2	0.775	0.475
Inadequat	0.475	0.8	0.225	0.525

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.45	0.775	0.575	0.375	0.1625	0.65	0.525
Inadequat	0.4	0.55	0.225	0.425	0.625	0.8375	0.35	0.475

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6125	0.425	0.825	0.625	0.4	0.3	0.65	0.4625
Inadequat	0.3875	0.575	0.175	0.375	0.6	0.7	0.35	0.5375

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.0156	0.125	0.095	0.2	0.2125	0.3625	0.3025	0.4875	0.1437	0.2	0.2	0.2875	0.4	0.5075	0.5	0.625
L2 (0.3)	0.0968	0.2125	0.13	0.25	0.275	0.3187	0.325	0.3	0.2375	0.2375	0.2625	0.3	0.2475	0.275	0.3	0.2875
M1 (0.5)	0.2	0.255	0.3125	0.2375	0.2562	0.2125	0.2375	0.1625	0.2625	0.3	0.3	0.275	0.2	0.15	0.1625	0.075
H1 (0.7)	0.15	0.2275	0.2875	0.2125	0.1937	0.0875	0.095	0.0425	0.2437	0.175	0.15	0.1175	0.125	0.0625	0.0275	0.0125
H2 (0.9)	0.5374	0.18	0.175	0.1	0.0625	0.0187	0.04	0.0075	0.1125	0.0874	0.0875	0.02	0.0275	0.005	0.01	0

Figure 26. HVAC M-3 Model

The same PSF evidence (Table 4) used in M-1 and M-2 were input into the fixed structure-consensus CPT M-3 model (Table 12). Since there was only one model, weighted posterior methodology was not used. For project-1 M-3 projects an FEF of 72% vs actual SME projection of 100% with projection error of 24% vs actual SME FEF projection error of 72%. Subsequent FEF projections can be found in Table 13 with errors noted in Table 14.

Table 12

Project Results – HVAC Fix Effectiveness (di) Projection Project 1
Method: M-3 SME Models

	Prior	Posterior				
Node Variable	M-3	Project 1	Project 2	Project 3	Project 4	Project 5
Mgt. commitment	0.525	0	1	0	1	1
Quality system maturity	0.531	0	1	0	0	1
Project time	0.444	1	0.756	0	1	0
Failure mode complexity	0.365	0.652	0	0.506	0.690	1
Technical expertise	0.491	1	0.650	0.344	0.736	0.575
Resource availability	0.481	1	1	0.200	1	0
Design complexity	0.300	0.378	0	1	1	1
Test facilities	0.518	1	1	0	1	0
Fix effectiveness 10%	0.299	0.016	0.073	0.549	0.044	0.553
Fix effectiveness 30%	0.245	0.097	0.146	0.296	0.127	0.295
Fix effectiveness 50%	0.209	0.200	0.236	0.121	0.215	0.125
Fix effectiveness 70%	0.123	0.150	0.198	0.027	0.170	0.021
Fix effectiveness 90%	0.124	0.538	0.348	0.007	0.443	0.006

Table 13

HVAC M-3 FEF Projections

Project	FEF observed	M-3	SME
1	58	72	100
2	70	62	90
3	30	23	60
4	62	67	100
5	49	23	72

Table 14

HVAC M-3 FEF Projection Error

Project	M-3	SME
1	-24%	-72%
2	11%	-29%
3	23%	-100%
4	-8%	-61%
5	54%	-47%

In four of five FEF projections, the M-3 model proved to be more accurate than SME initial projections. Overall the M-3 BBN exhibited a 24% FEF projection error vs. an SME error of 61.8%, 61% less than the SME projection (Table 14).

HVAC M-4 BBN Aggregate

The forth *BBN* developed was *M-4* which is simply an aggregate of models *M-1, M-2* and *M-3* (Table 15). Weighted posterior methodology is used to weight each *SME* judgment within a given model. *M-4 FEF* projection for project-1 was 57.4% vs *SME*

Table 15

HVAC M-4 Model- Fix Effectiveness (di) Projection Project 1 – BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	0	0	0	0	0	0	0	0	0	0.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	0	0	0	0	0	0	0	0	0	0.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	.769	.240	.344	.407	.829	.394	.847	.850	.652	0.528
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	.824	.499	.217	.555	.505	.289	.306	.504	.378	.507
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table 15 (*continued*).

Node Variable	Prior									Posterior									*Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.065	.050	.069	.057	.000	.050	.013	.000	.016	.044
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.231	.200	.148	.162	.100	.200	.038	.050	.097	.173
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.465	.500	.250	.227	.150	.500	.050	.100	.200	.368
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.196	.200	.370	.277	.200	.200	.100	.100	.150	.202
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.043	.050	.162	.277	.550	.050	.800	.750	.538	.214
Likelihood	-		-		-		-			3.13 E-02	3.30 E-02	3.12 E-04	2.47 E-02	1.74 E-02	3.56 E-02	6.70 E-02	2.50 E-03	1.21 E-02	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	0.19	0.20	0.00	0.15	0.11	0.22	0.04	0.02	0.070	

Note: *Aggregate

projection of 100% with projection error of 1% vs actual *SME FEF* projection error of 72%. Project *FEF* projections can be found in Table 16 with errors noted in Table 17.

Table 16

HVAC M-4 FEF Projections

Project	FEF observed	M-4	SME
1	58	57	100
2	70	49	90
3	30	24	60
4	62	53	100
5	49	28	72

Table 17

HVAC M-4 FEF Projection Error

Project	M-4	SME
1	1%	-72%
2	31%	-29%
3	20%	-100%
4	14%	-61%
5	43%	-47%

In four of five FEF projections, the M-4 model proved to be more accurate than the subject matter expert's initial projections. Overall the M-4 BBN exhibited an FEF projection error of 21.8% whereas SME projection error is 61.8%. The M-4 BBN aggregate exhibits 64.7% less error than SME projections. Empirical evidence indicates that all four BBN exhibited less overall FEF projection error than projections made by SME during project planning (Table 18). Project details can be found in Appendix G.

Table 18

HVAC FEF Projection Error by Model Type

Project	M-1	M-2	M-3	M-4	SME
1	10%	-6%	-24%	1%	-72%
2	42%	27%	11%	34%	-29%
3	26%	19%	23%	23%	-100%
4	19%	10%	-8%	15%	-61%
5	15%	47%	54%	42%	-47%

HVAC Case Study Analysis Overview

The process for analysis involved the following:

1. Power value for each experiment was held constant at 0.9. M-1 BBN model standard deviation was used as the baseline reference for sample size calculation. The difference in FEF projection we wished to detect was considered 0.05.
2. Perform random realizations per power and sample size calculations of previous step.

3. Perform test of equal variances.
4. Perform ANOVA to determine statistical significance among model means, with the null hypothesis equating to no difference in means. Perform Tukey pairwise comparisons as necessary.
5. Perform a one sample T-test comparing each model mean against the actual FEF obtained from field data.
6. Repeat analysis for the next project

Results for HVAC Project-1

Power and sample size calculations (Table 19) indicate 455 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Levene's test indicated one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure 27).

One way ANOVA analysis indicated one must reject the null hypothesis of no difference in the models (Table 20). The Tukey pairwise comparisons indicate which models are statistically significant (Table 21), should a zero crossing occur for any of the paired model combinations. The data indicates no significance among any model combinations. Finally, the one sample T (Table 22) looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject H_o of no difference for all models except M-4. M-4 proved to be statistically significance whereby one must fail to reject the null hypothesis of no difference. Details for the remaining HVAC projects can be found in Appendix (H).

Table 19

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	455	0.9000	0.9004	0.05

Note: Sigma = 0.2, Alpha = 0.05, Number of Levels = 4

Table 20

One-way ANOVA: HVAC Project-1: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	3	9.0942	3.0314	62.81	0.000
Error	1816	87.6515	0.0483		
Total	1819	96.7457			

Individual 95% Cis for Mean Based on Pooled StDev

Level	N	Mean	StDev	---------+---------+---------+-------
M1	455	0.5209	0.2055	(-*--)
M2	455	0.6270	0.2267	(--*-)
M3	455	0.7145	0.2261	(--*--)
M4	455	0.5810	0.2198	(--*--)
				---------+---------+---------+-------
Pooled StDev =		0.2197		0.560 0.630 0.700

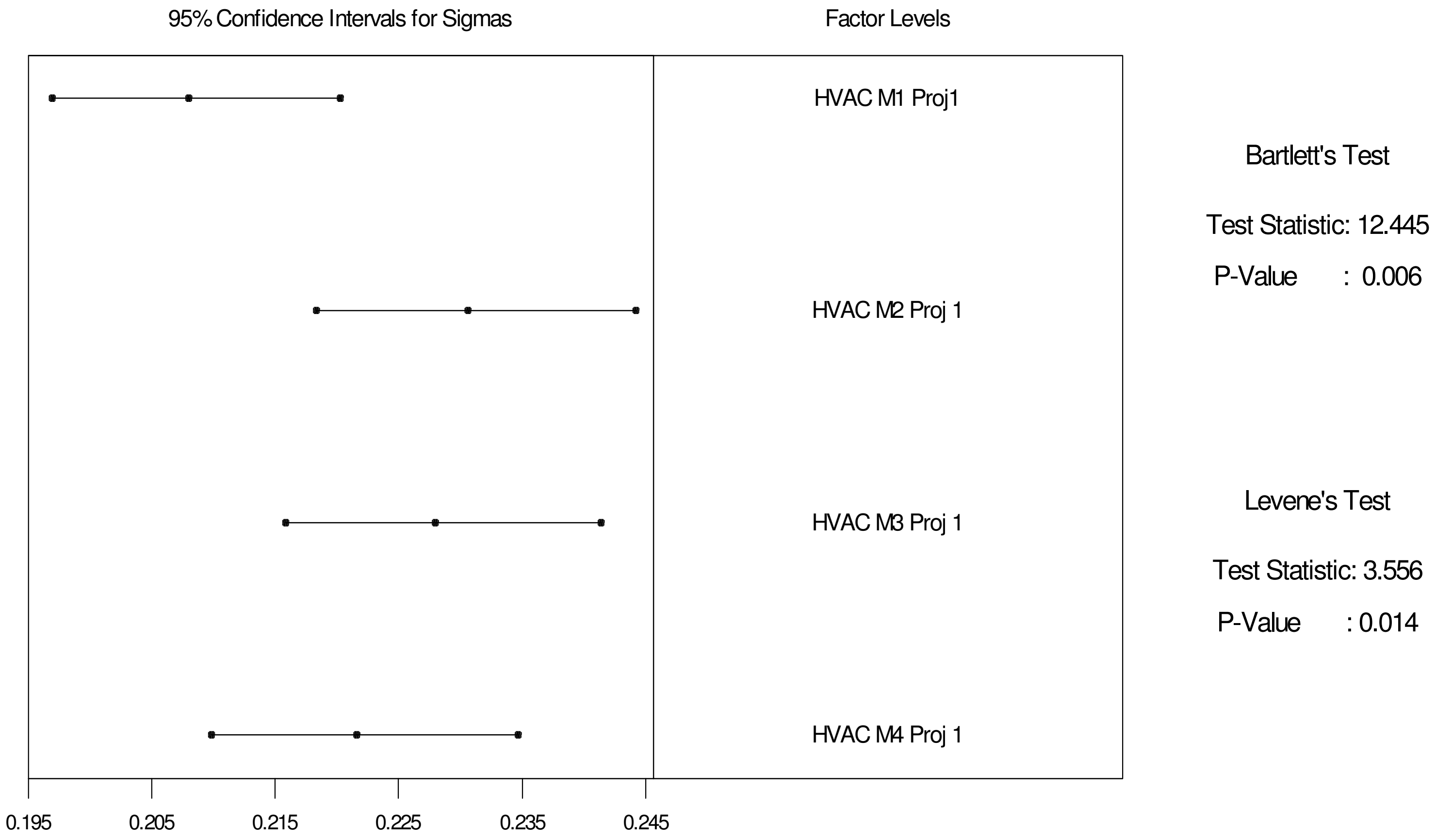


Figure 27. Test for equal variances for HVAC project 1

Table 21

Tukey's Pairwise Comparisons

	HVAC M1	HVAC M2	HVAC M3
HVAC M2	-0.1131		
	-0.0621		
HVAC M3	-0.2173	-0.1298	
	-0.1663	-0.0787	
HVAC M4	-0.0812	0.0064	0.1107
	-0.0301	0.0574	0.1617

Table 22

One-Sample T: HVAC Project-1: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1	455	0.52088	0.20553	0.00964
M2	455	0.6270	0.2267	0.0106
M3	455	0.7145	0.2261	0.0106
M4	455	0.5810	0.2198	0.0103

Variable	95.0% CI	T	P
M1	(0.50194, 0.53981)	-6.14	0.000
M2	(0.6061, 0.6479)	4.42	0.000
M3	(0.6937, 0.7353)	12.69	0.000
M4	(0.5607, 0.6012)	0.10	0.923

Note: *Note*: Test of mu = 0.58 vs mu not = 0.58

A P-value > 0.05 indicates one must fail to reject the null of no difference between actual FEF and M-4 FEF projections. Details for the remaining HVAC projects can be found in Appendix H.

Automotive Test Case

One question to be answered by this research was; can generic PSF and BBN structures provide more accurate FEF estimates across diverse industries than projections made by SME? Models M-1 expert aggregate, M-2 fixed structure, M-3 consensus model, and M-4 BBN aggregate, developed by HVAC SME, were used to project fix effectiveness for the automotive industry.

Table 23 indicates that BBN methodology provides less overall FEF projection error than that of SME as predicted at the onset of a product development project. Reviewing each auto BBN, one can see the M-4, BBN aggregate provides the least overall error at 43%, 55% less than SME. M-3 consensus model, exhibits 44% overall error, 54% less than SME. M-1, expert aggregate BBN, exhibits 47% error, 51% less than SME and M-2 exhibits 48% error, 50% less than SME. Average error from all automotive BBN models is 45%, 53% less than subject matter experts.

Table 23

Percent Error by Model Type

Variable	% Error					
	M-1	M-2	M-3	M-4	BBN Avg	SME
HVAC	20	22	24	22	22	62
Auto	47	48	44	43	45	96
Overall	34	35	34	32	34	79

Automotive SME reviewed past projects and collected PSF evidence and FEF that would have been known or “perceived” to be known at the onset of each project (Table 24). For example, evidence for project-1 indicated management commitment, quality system maturity, technical expertise, resource availability, design complexity, and test facilities were adequate, whereas project time and failure mode complexity were inadequate. *SME* felt confident in their ability to solve the failure mode so they assigned a fix effectiveness of 90%.

Table 24

Automotive Project Evidence

Variable	Project 1	Project 2	Project 3	Project 4	Project 5
Mgt. commitment	1	1	1	1	1
Quality system maturity	1	1	0	1	1
Project time	0	1	1	1	1
Failure mode complexity	0	0	1	1	0

Table 24 (*continued*).

Variable	Project 1	Project 2	Project 3	Project 4	Project 5
Technical expertise	1	1	?	1	0
Resource availability	1	1	1	1	?
Design complexity	1	1	?	1	1
Test facilities	1	?	1	1	1
SME Projected FEF	90%	90%	90%	90%	90%
Actual FEF	26.5%	50%	99.8%	85.3%	56.9%

Automotive M-1 Expert Aggregate

PSF evidence (Table 18) was entered into each SME M-1 model, (developed by HVAC SME) allowing for an FEF projection. In addition an aggregate model was developed by calculating how close the marginal probability of each SME was to the observed evidence allowing a posterior SME judgment weight to be established (Stiber et al., 2004). For project-1 noted earlier, one can see in Table 25, the SME weighted posterior calculates to be 14% for SME-3, 59% for SME-5, 0% for SME-7, and 27% for SME-8.

The M-1 aggregate model projects FEF of 60.1% (Table 26) vs actual SME projections of 90%. Field data suggests the resulting fix effectiveness was 26.5%, thus M-1 aggregate projection error is 56% vs actual SME FEF projection error of 240%. Subsequent FEF projections can be found in Table 26 with errors noted in Table 27.

Table 25

Automotive M-1 Project Results

	Prior				Posterior				Aggregate			
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8				
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000			
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000			
Project time	0.900	0.428	0.054	0.674	0	0	0	0	0.000			
Failure mode complexity	0.610	0.200	0.119	0.250	0	0	0	0	0.000			
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000			
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000			
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000			
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000			
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.10	0.05	0.10	0.04	0.054			
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.30	0.10	0.20	0.12	0.133			
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.55	0.20	0.25	0.20	0.248			
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.01	0.50	0.25	0.32	0.384			
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.04	0.15	0.20	0.32	0.180			
Likelihood	-	-	-		3.43E-03	1.47E-02	1.19E-05	6.72E-03	Mean	Var.	Alpha	Beta
Prob (M_j)	0.25	0.25	0.25	0.25	0.14	0.59	0.00	0.27	0.601	0.048	2.395	1.593

Appendix I references other SME M-1 automotive models. In three of five *FEF* projections, the *M-1* aggregate model demonstrated improved *FEF* projection accuracy. Overall M-1, expert aggregate BBN, exhibits 47% error, 51% less than SME.

Table 26

Automotive M-1 FEF Projections by Model Type

		M-1					
Project	FEF Observed	SME-3	SME-5	SME-7	SME-8	Aggregate	SME
1	26.5	41.7	62.0	55.0	65.2	60.1	90
2	50.0	40.7	59.4	53.5	62.8	51.1	90
3	99.8	51.6	51.5	64.0	60.8	54.1	90
4	85.3	48.8	50.0	64.2	69.0	51.1	90
5	56.9	41.7	36.6	23.7	55.7	44.3	90

Table 27

Automotive M-1 FEF Projection Error

	M-1				M-1	
Project	SME-3	SME-5	SME-7	SME-8	Aggregate	SME
1	-134%	-108%	56%	-146%	56%	-240%
2	19%	-19%	-7%	-26%	2%	-80%
3	48%	48%	36%	39%	-84%	10%
4	43%	41%	25%	19%	-67%	-5%
5	-14%	0%	35%	-52%	-28%	-146%

Automotive M-2 Fixed Structure

The same PSF evidence (Table 18) used in M-1 are input into the fixed structure M-2 models (Table 28). Weighted posterior SME judgments were determined via Stiber (2004). For project-1 SME-3 judgment weight is 28%, SME-5 is 44%, SME-7 is 8% and SME-8 weight is 21%. The M-2 aggregate projects a FEF of 47.7% vs actual SME projection of 90% with projection error of 44% vs actual SME FEF projection error of 240%. Subsequent project FEF projections can be found in Table 29 with errors noted in Table 30. M-2 project details can be found in Appendix J.

Table 28

Automotive M-2 Project-1 Results –Fix Effectiveness (di) Projection
Method: M-2 SME Models

	Prior				Posterior				Aggregate			
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8				
Mgt. commitment	0.500	0.600	.3400	0.600	1	1	1	1	1.000			
Quality system maturity	0.450	0.520	.0460	0.640	1	1	1	1	1.000			
Project time	0.418	0.629	0.364	0.348	0	0	0	0	0.000			
Failure mode complexity	0.560	0.460	0.240	0.320	0	0	0	0	0.000			
Technical expertise	0.549	0.670	0.430	0.277	1	1	1	1	1.000			
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000			
Design complexity	0.400	0.300	0.200	0.300	1	1	1	1	1.000			
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000			
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.200	0.200	0.050	0.050	0.157			
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.450	0.250	0.100	0.050	0.252			
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.270	0.350	0.250	0.150	0.279			
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.060	0.200	0.500	0.150	0.175			
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.020	0.000	0.100	0.600	0.137			
Likelihood	-	-	-		3.50E-03	5.53E-03	1.01E-03	2.61E-03	Mean	Var.	Alpha	Beta
Prob (M_j)	0.25	0.25	0.25	0.25	0.28	0.44	0.08	0.21	0.477	0.064	1.392	1.528

Table 29

Automotive M-2 FEF Projections by Model Type

		M-2				M-2	
Project	FEF Observed	SME-3	SME-5	SME-7	SME-8	Aggregate	SME
1	26.5	55.0	41.0	60.0	74.0	47.7	90
2	50.0	68.6	47.3	77.4	74.5	58.6	90
3	99.8	68.8	47.9	81.4	55.5	55.9	90
4	85.3	74.0	50.0	82.8	81.0	58.6	90
5	56.9	47.7	36.6	44.3	28.7	36.9	90

Within three of five *FEF* projections, the *M-2* aggregate model proved to have equal or better *FEF* projection accuracy. Overall the *M-2* aggregate *BBN* exhibited an average projection error of 47.6%, 50% less than SME.

Table 30

Automotive M-2 FEF Projection Error by Model Type

	M-2				M-2	
Project	SME-3	SME-5	SME-7	SME-8	Aggregate	SME
1	-32%	-55%	-126%	-179%	44%	-240%
2	-37%	5%	-55%	-49%	15%	-80%
3	31%	52%	18%	44%	-79%	10%
4	13%	41%	3%	5%	-46%	-5%
5	-30%	0%	-21%	22%	-54%	-146%

Automotive M-3 Consensus Model

The M-3 consensus model is a fixed structure consensus CPT developed by HVAC SME (Figure 28). The same PSF evidence (Table 18) used in M-1 and M-2 were input into the fixed structure-consensus CPT M-3 model. Since there was only one model, weighted posterior methodology was not used. For project-1 (Table 31), M-3 projects an FEF of 56% vs actual SME projection of 90% with projection error of 113% vs actual SME FEF projection error of 240%. Subsequent FEF projections can be found in Table 32 with errors noted in Table 33.

In three of five FEF projections, the M-3 model proved to be more accurate than the subject matter expert's initial projections. Overall the M-3 aggregate BBN exhibited an average projection error of 44.2%, 54% less that SME projections. M-3 project details can be found in Appendix K.

Table 31

Automotive M-3 Project Results

	Prior	Posterior				
Node Variable	M-3	Project 1	Project 2	Project 3	Project 4	Project 5
Mgt. commitment	0.525	1	1	1	1	1
Quality system maturity	0.531	1	1	0	1	1
Project time	0.444	0	1	1	1	1
Failure mode complexity	0.365	0	0	1	1	0
Technical expertise	0.491	1	1	0.775	1	0
Resource availability	0.481	1	1	1	1	0.717

Table 31 (*continued*).

	Prior	Posterior				
Node Variable	M-3	Project 1	Project 2	Project 3	Project 4	Project 5
Design complexity	0.300	1	1	0.473	1	1
Test facilities	0.518	1	0.650	1	1	1
Fix effectiveness 10%	0.299	0.095	0.085	0.040	0.016	0.146
Fix effectiveness 30%	0.245	0.130	0.159	0.123	0.097	0.220
Fix effectiveness 50%	0.209	0.313	0.220	0.212	0.200	0.268
Fix effectiveness 70%	0.123	0.288	0.165	0.167	0.150	0.213
Fix effectiveness 90%	0.124	0.175	0.371	0.457	0.538	0.154
FEF Projection		56%	62%	68%	72%	50%

Table 32

Automotive M-3 FEF Projections

Project	FEF Observed	M-3	SME
1	26.5	56	90
2	50.0	62	90
3	99.8	68	90
4	85.3	72	90
5	56.9	50	90

Table 33

Automotive M-3 FEF Projection Error

Project	M-3	SME
1	-113%	-240%
2	-23%	-80%
3	32%	10%
4	16%	-5%
5	-37%	-146%

Automotive M-4 BBN Aggregate

The automotive M-4 is an aggregate of models M-1, M-2 and M-3 developed by HVAC SME (Table 34) incorporating PSF evidence from automotive projects. Weighted posterior methodology is used weight each SME judgment within a given model. M-4 FEF projection for project-1 is 55.9% vs SME FEF projection of 90% with projection error of 111% vs actual SME FEF projection error of 240%. Project FEF projections can be found in Table 35 with errors noted in Table 36. Additional project details can be found in Appendix L.

Table 34

Automotive M-4 Results – Fix Effectiveness (di) Project 1 – BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	0	0	0	0	0	0	0	0	0	0.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	0	0	0	0	0	0	0	0	0	0.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table 34 (*continued*).

Node Variable	Prior									Posterior									*Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.100	.05	.100	.040	.200	.200	.050	.050	.095	.089
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.300	.10	.200	.120	.450	.250	.100	.050	.130	.169
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.550	.20	.2550	.200	.270	.350	.250	.150	.313	.263
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.013	.50	.250	.320	.060	.200	.500	.150	.288	.311
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.037	.15	.200	.320	.020	.000	.100	.600	.175	.167
Likelihood	-		-		-		-			3.43 E-03	1.47 E-02	1.19 E-05	6.72 E-03	3.50 E-03	5.53 E-03	1.01 E-03	2.61 E-03	3.61 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	0.08	0.36	0.00	0.16	0.09	0.13	0.02	0.06	0.09	
																Mean	Var.	Alpha	Beta
																0.559	0.057	1.872	1.476

Note: *Aggregate

Table 35

Automotive M-4 FEF Projections

Project	FEF Observed	M-4	SME
1	26.5	56	90
2	50.0	53	90
3	99.8	56	90
4	85.3	53	90
5	56.9	42	90

Table 36

Automotive M-4 FEF Projection Error

Project	M-4	SME
1	-111%	-240%
2	-6%	-80%
3	44%	10%
4	38%	-5%
5	-15%	-146%

In three of five FEF projections, the M-4 model proved to be more accurate than the subject matter expert's initial projections. Overall, the M-4 aggregate BBN exhibited an average projection error of 42.8%, 56% less FEF projection error than that of SME.

<u>Automotive BBN Analysis</u>

The same analysis methodology used for the HVAC case study was repeated for the automotive case study. Power and sample size calculations (Table 37) indicate 541 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Levine's test indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure 28). The remaining Automotive model analysis can be found in Appendix M.

Table 37

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	541	0.9000	0.9003	0.05

Note: Sigma = 0.218161, Alpha = 0.05, Number of Levels = 4

One way ANOVA indicated one must reject the null hypothesis of no difference in the models (Table 38) with Tukey pairwise comparisons (Table 39) indicating M3 FEF projection mean is statistically equal to M4. One sample T (Table 40) looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference. Overall summary is the models did not perform well in predicting actual FEF on this project. One would expect these results though given most node evidence were very positive however actual FEF was low at 26.5%.

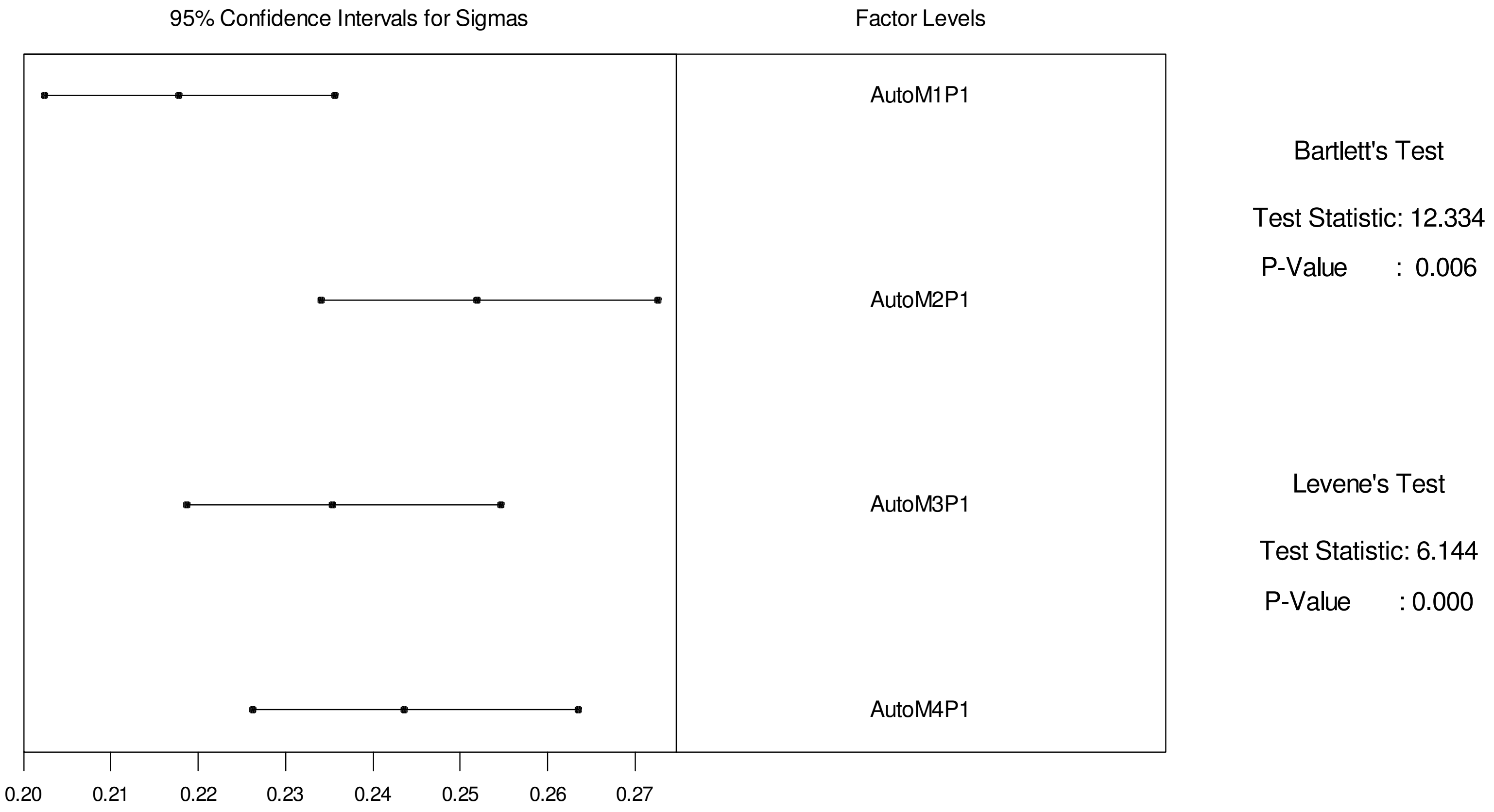


Figure 28. Test for equal variances for automotive project-1

Table 38

One-way ANOVA: Auto Project-1: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP1St	3	4.1682	1.3894	24.63	0.000
Error	2160	121.8382	0.0564		
Total	2163	126.0064			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level        N      Mean     StDev      ---------+---------+---------+-------
M1         541    0.6048    0.2178                                 (---*---)
M2         541    0.4829    0.2520         (---*---)
M3         541    0.5496    0.2354                     (---*---)
M4         541    0.5640    0.2435                        (---*---)
Pooled StDev =    0.2375                ---------+---------+---------+-------
                                               0.500      0.550     0.600
```

Table 39

Tukey's Pairwise Comparisons

	Auto M1	Auto M2	Auto M3
Auto M2	0.0849		
	0.1590		
AUTO M3	0.0182	-0.1038	
	0.0923	-0.0296	
Auto M4	0.0038	-0.1181	-0.0514
	0.0779	-0.440	0.0227

Table 40

One-Sample T: Auto Project-1: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1	541	0.60484	0.21780	0.00936
M2	541	0.4829	0.2520	0.0108
M3	541	0.5496	0.2354	0.0101
M4	541	0.5640	0.2435	0.0105

Variable	95.0% CI	T	P
M1	(0.58654, 0.62324)	36.29	0.000
M2	(0.4616, 0.5042)	20.12	0.000
M3	(0.5297, 0.5695)	28.12	0.000
M4	(0.5434, 0.5845)	28.55	0.000

Note: Test of mu = 0.265 vs mu not = 0.265

Model Calibration Summary

Gradient ascent methodologies are used to calibrate or tweak CPT, such that FEF projection error is reduced (Chapter-3, Model Calibration). Two approaches are used to update CPT, the first involved repetitive updating of a single project. This method allowed FEF projection error to diminish with each learning iteration, however the CPT for other parent nodes maxed out prior to FEF reaching the desired FEF value, therefore this method was scrapped. The second method involved a single update iteration of each project with the updated model becoming the baseline for the next project update. This process is repeated for all five projects within a respective industry (Table 41).

Table 41

HVAC Error Post Learning

Project	Actual Observed	Prior to Learn Projection Expected	Post Learn Projection Expected	Prior to Learn Error	Post Learn Error
1	58	71.9	71.4	23.97%	23.10%
2	70	62.0	61.7	11.43%	11.86%
3	30	23.0	23.3	23.33%	22.33%
4	62	66.8	66.2	7.74%	6.77%
5	49	22.6	23.0	53.88%	53.06%
Average Error				24.07%	23.42
Difference 2.67%					

HVAC project-1 evidence was used to update M-3 CPT. The resulting model was again updated using project-2 evidence and so on until all known evidence was propagated through the model, tweaking CPT with each learning iteration. After all learning iterations were complete, the resulting M-3 was again used to project fix effectiveness.

One can see in Table 41 how fix effectiveness values tweaked toward an improved projection, and after 5 learning iterations, (one for each project), FEF projection error reduced by 2.67%. While a 2.7 % reduction in error is appreciated, it hardly represents the grandiose expectations had at the onset of this research. Numerous issues presented themselves. For example, the narrow band of FEF projection capability of the model, caused the CPT to max out at a value of 1.0, prior to FEF projections reaching the actual field measured FEF. Thus, the error reduction capability of the gradient ascent algorithm was capped as well.

A second concern was PSF evidence that contradicted itself confused the model. One would think as PSF evidence became more negative fix effectiveness would follow and shrink to a lower value. Conversely; as PSF evidence improved so would FEF. Project-4 had more positive PSF than Project-3, but its actual FEF was almost 15% less. CPT would yo-yo back and forth as contradictory evidence propagated through the model. Previous CPT updates were at times reversed with subsequent updates, whereby error reduction capability of gradient ascent was reduced. Both of these concerns were influencing factors that led to development of Simulated SME (S-SME) models. Though outside the scope of the approved proposal, the magnitude of these issues warranted exploration of a proposed solution. S-SME methodology is discussed in Chapter 5.

Model Calibration Detail

The remainder of this chapter contains the detail of the calibration of various SME models. Shown below are repetitive updates for an M-2 fixed structure model for project-1 of the HVAC industry. Four update iterations are performed. Given the standard deviation of the M-2 baseline (prior to updates), we require a random realization of 626 samples such that the actual power of the experiment is 0.9, allowing a maximum difference detection of 0.5 (Table 42). One must reject the null of no difference in variance, with update-4 variance observed as the lowest (Figure 29). One way ANOVA results show one must reject the null of no difference in means among the baseline and updates (Table 43). The Tukey pairwise comparisons indicate statistical significance is achieved by the second learning iteration (Table 44). T-tests confirm statistical significance of the model as compared to actual FEF by the second and third learning iterations (Table 45). Thus in this case, gradient ascent was successful in adjusting CPT to reduce FEF projection error, resulting in a statistically significant FEF projection.

Table 42

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	626	0.9000	0.9001	0.05

Note: Sigma = 0.225167, Alpha = 0.05, Number of Levels = 5

Table 43

One-way ANOVA: Problem-1: HVACP M2 Baseline vs Repetitive M2 Updates

Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	4	0.5544	0.1386	3.05	0.016
Error	3125	142.2179	0.0455		
Total	3129	142.7723			

Individual 95% Cis for Mean Based on Pooled StDev				
Level	N	Mean	StDev	----+------------+------------+------------+--
M2P1Base	626	0.6273	0.2284	(--------*--------)
M2P1U1	626	06.071	0.2202	(--------*--------)
M2P1U2	626	0.5922	0.2130	(--------*--------)
M2P1U3	626	0.5905	0.2099	(--------*--------)
M2P1U4	626	0.6009	0.1936	(--------*--------)
Pooled StDev =		0.2133		----+------------+------------+------------+-- 0.580 0.600 0.620 0.640

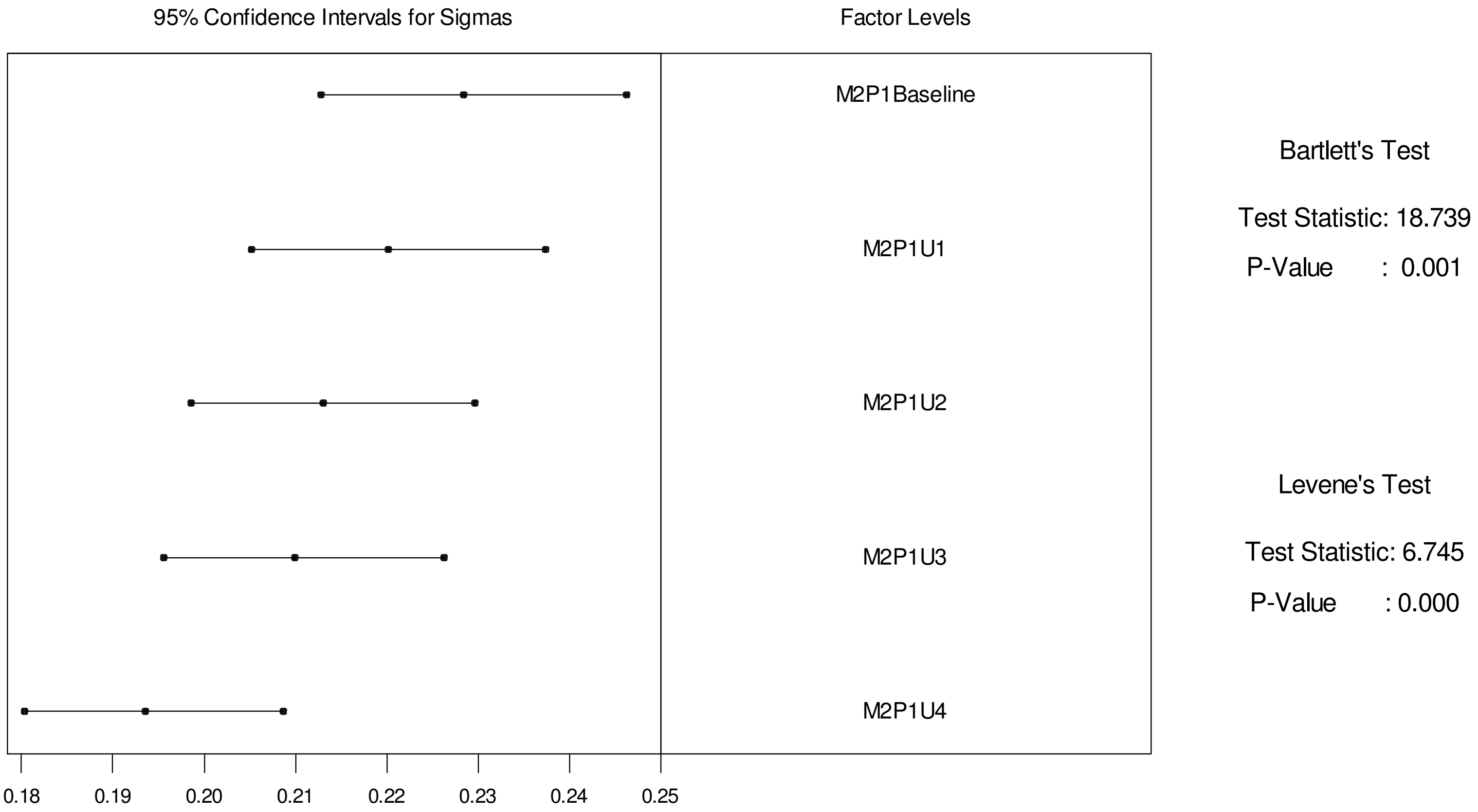


Figure 29. Test for equal variances for HVAC project 2 updates

Table 44

Tukey's Pairwise Comparisons

	M2P1Base	M2P1U1	M2P1U2	M2P1U3
M2P1U1	-0.0126			
	0.0532			
M2P1U2	0.0023	-0.0180		
	0.0681	0.0478		
M2P1U3	0.0040	-0.0163	-0.0312	
	0.0698	0.0495	0.0346	
M2P1U4	-0.0065	-0.0268	-0.0417	-0.0433
	0.0593	0.0391	0.0242	0.0225

Table 45

One-Sample T: Problem-1: HVACP M2 Baseline vs Repetitive M2 Updates

Variable	N	Mean	St.Dev.	SE Mean
M2P1Baseline	626	0.62734	0.22839	0.00913
M2P1U1	626	0.60706	0.22016	0.00880
M2P1U2	626	0.59217	0.21304	0.00851
M2P1U3	626	0.59047	0.20988	0.00839
M2P1U4	626	0.60091	0.19359	0.00774

Variable	95.0% CI	T	P
M2P1Baseline	(0.60941, 0.64526)	5.19	0.000
M2P1U1	(0.58978, 0.62434)	3.08	0.002
M2P1U2	(0.57544, 0.60889)	1.43	0.154
M2P1U3	(0.57400, 0.60695)	1.25	0.212
M2P1U4	(0.58571, 0.61610)	2.70	0.007

Method-2 is used to update M-1 expert aggregate model CPT. Project-1 evidence is applied to applicable nodes updating the model CPT. The resulting model is again updated using project-2 evidence and so on until all known evidence is propagated through the model, tweaking CPT with each learning iteration. After all learning iterations are complete, the resulting M-1 is again used to project fix effectiveness.

One can see in Project-1 ANOVA below that after the update one must fail to reject the null of no difference in means from the baseline to the update. Although the mean did shift, the shift was not statistically significant. One sample T indicates the one must reject the null of no difference in baseline and updated means versus the actual FEF. This same analysis held true for all five M-1 project updates (Tables 46-55).

Table 46

One-way ANOVA: Project-1: M1 Baseline, M1 Update

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	1	0.0714	0.0714	1.64	0.201
Error	908	39.5388	0.0435		
Total	909	39.6102			

Individual 95% Cis for Mean Based on Pooled StDev

Level	N	Mean	StDev	
				`-------+---------+---------+---------`
M1P1	455	0.5209	0.2055	` (-----------*-----------)`
M1P1Ud	455	0.5286	0.2118	`           (-----------*-----------)`
Pooled StDev =		0.2087		`-------+---------+---------+---------`
				0.512 0.528 0.544

Table 47

One-Sample T: Project-1: M1 Baseline, M1 Update

Variable	N	Mean	St.Dev.	SE Mean
M1P1	455	0.52088	0.20553	0.00964
M1P1Ud	455	0.53859	0.21157	0.00993

Variable	95.0% CI	T	P
M1P1	(0.50194, 0.53981)	-6.14	0.000
M2P1Ud	(0.51908, 0.55810)	-4.17	0.000

Note: Test of mu = 0.58 vs mu not = 0.58

Table 48

One-way ANOVA: Project-2: M1 Baseline, M1 Update

Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	1	0.1086	0.1086	1.82	0.177
Error	1362	81.1645	0.0596		
Total	1363	81.2730			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level      N     Mean     StDev    -------+---------+---------+---------
M1P2      682   0.4415   0.2416   (-----------*-----------)
M1P2Ud    682   0.4594   0.2466               (-----------*-----------)
Pooled StDev =  0.2441             -------+---------+---------+---------
                                        0.435     0.450     0.465
```

Table 49

One-Sample T: Project-2: M1 Baseline, M1 Update

Variable	N	Mean	St.Dev.	SE Mean
M1P2	682	0.44153	0.24159	0.00925
M1P2Ud	682	0.45937	0.24661	0.00944

Variable	95.0% CI	T	P
M1P2	(0.42336, 0.45969)	-27.94	0.000
M2P2Ud	(0.44083, 0.47791)	-25.48	0.000

Note: Test of mu = 0.7 vs mu not = 0.7

Table 50

One-way ANOVA: Project-3: M1 Baseline, M1 Update

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	1	0.0021	0.0021	0.07	0.785
Error	750	21.3827	0.0285		
Total	751	21.3848			

Individual 95% Cis for Mean Based on Pooled StDev

Level	N	Mean	StDev	--+---------+---------+---------+----
M1P3	376	0.2348	0.1748	(-------------*-------------)
M1P3Ud	376	0.2314	0.1625	(-------------*-------------)
Pooled StDev =		0.1688		--+---------+---------+---------+----
				0.216 0.228 0.240 0.252

Table 51

One-Sample T: Project-3: M1 Baseline, M1 Update

Variable	N	Mean	St.Dev.	SE Mean
M1P3	376	0.23475	0.17493	0.00902
M1P3Ud	376	0.23139	0.16254	0.00838

Variable	95.0% CI	T	P
M1P3	(0.21701, 0.25249)	-7.23	0.000
M2P3Ud	(0.21491, 0.24787)	-8.18	0.000

Note: Test of mu = 0.3 vs mu not = 0.3

Table 52

One-way ANOVA: Project-4: M1 Baseline, M1 Update

Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	1	0.0416	0.0416	0.98	0.323
Error	1022	43.4429	0.0425		
Total	1023	43.4846			

Individual 95% Cis for Mean Based on Pooled StDev				
Level	N	Mean	StDev	--+---------+---------+---------+----
M1P4	512	0.5090	0.2100	(--------------*-------------)
M1P4Ud	512	0.5218	0.2023	(--------------*-------------)
Pooled StDev =		0.2062		--+---------+---------+---------+----
				0.495 0.510 0.525 0.540

Table 53

One-Sample T: Project-4: M1 Baseline, M1 Update

Variable	N	Mean	St.Dev.	SE Mean
M1P4	512	0.50902	0.21000	0.00928
M1P4Ud	512	0.52177	0.20227	0.00894

Variable	95.0% CI	T	P
M1P4	(0.49078, 0.52725)	-11.96	0.000
M2P4Ud	(0.50421, 0.53933)	-10.99	0.000

Note: Test of mu = 0.3 vs mu not = 0.3

Table 54

One-way ANOVA: Project-5: M1 Baseline, M1 Update

Analysis of Variance					
Source	DF	SS	MS	F	P
Factor	1	0.1302	0.1302	2.95	0.086
Error	998	43.9792	0.0441		
Total	999	44.1094			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level         N     Mean     StDev     --+---------+---------+---------+----
M1P5        500   0.4203    0.2056                  (--------*--------)
M1P5Ud      500   0.3975    0.2142       (--------*--------)
Pooled StDev =    0.2062                --+---------+---------+---------+----
                                        0.380     0.400     0.420     0.440
```

Table 55

One-Sample T: Project-5: M1 Baseline, M1 Update

Variable	N	Mean	St.Dev.	SE Mean
M1P5	500	0.42032	0.20560	0.00919
M1P5Ud1	500	0.39749	0.21416	0.00958

Variable	95.0% CI	T	P
M1P5	(0.40225, 0.43838)	-7.58	0.000
M2P5Ud	(0.37868, 0.41631)	-9.66	0.000

Note: Test of mu = 0.49 vs mu not = 0.49

Chapter 5: Simulated SME Judgment

Research Challenges

Numerous challenges presented themselves during this research. One had to serve as a referee as SME debated model structure. It became a laborious task for SME to develop CPT for their respective model, and was indeed a challenge for the team to reach CPT consensus for the M-3 model. Due to the intensity of the sessions, many team members stated they dreaded the interface with other team members. Team member enthusiasm began to dwindle and group think became prevalent as SME spent more time together. Dominant SME pushed their thoughts on others; not budging in the negotiating process, resulting in M-3 becoming the dominant SMEs M-2 model. This was an unacceptable outcome, thus SME were required to reassemble and properly develop an M-3 CPT. Ultimately the team agreed to average their individual M-2 CPT to create an M-3 model.

One important attribute of any model is its ability to represent reality. A second challenge in this research was building models with relatively low error in FEF prediction. In general, one would expect FEF activities of most engineering communities to bound between 20% to 80% fix effectiveness. Thus BBN models FEF projection should be able to swing within this band as PSF node evidence moves from all negative to all positive. Projects under review within this research saw actual FEF range from 26.5% to 99.8%, however BBN FEF projection ranges were from 23.8% to 61.5%. This narrow projection band results in an inherent error. SME-5 M-2 model was chosen at random to perform a full surface response DOE on FEF projections as node evidence are varied. Figure 30 provides an overview of the range of FEF projection. One can see CPT developed by SME-5 allow FEF to only range from 29% to 50%, well below actual FEF for projects under review. One can also see from Appendix E that SME-5 model weight was consistently a dominant weight in SME judgment, translating into a dominant factor in FEF projection error.

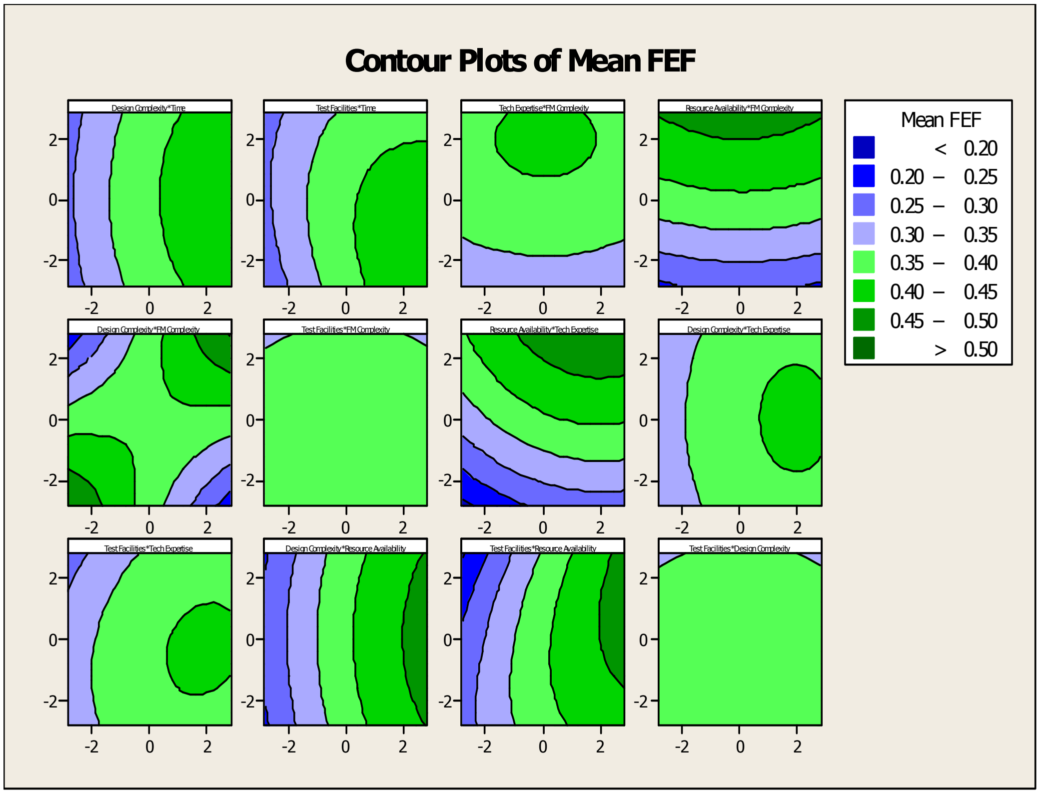


Figure 30. Contour plot of SME-5 FEF projection by PSF

Error projections from HVAC and Automotive industries provide insight as to the validity that BBN and generic PSF can provide reduced FEF error projections as compared to those provided from SME. However, the aforementioned challenges raise questions as to the method used to build a BBN and develop CPT. Is there a method that can minimize or eliminate group think, reduce the amount of time required to build CPT, and maximize ease of use of the models, all the while providing less error in FEF projection that an SME built BBN?

We propose the following methodology to simulate SME knowledge for a given model structure, whereby initial CPT are developed for a given BBN structure. FEF projections and projection error for both HVAC and Automotive industries are evaluated in the same manner as previously described.

Simulated SME Methodology

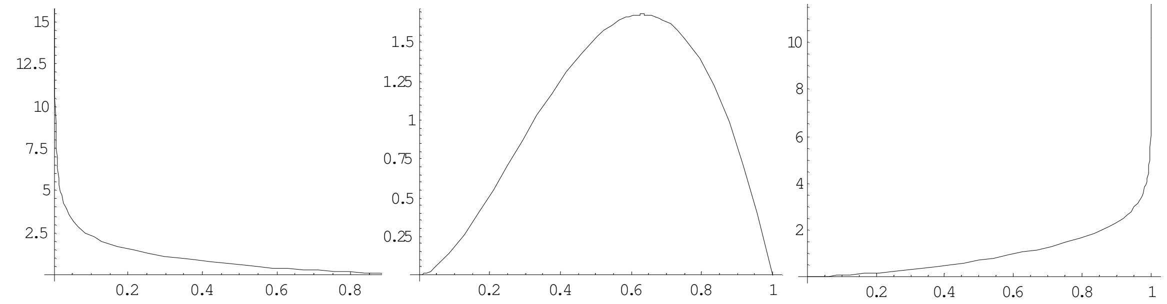
Meetings with SME spanned approximately five months. During this period SME determined PSF, built three model structures and developed associated CPT. PSF evidence from real world projects was used within each SME BBN to project that SME's FEF and FEF as a model aggregate for that model type, i.e., M-1 vs M-2, etc. Parameters for parametric distributions were defined for specific SME judgment per equations 17-18 (Martz, 1982). It was noted during this exercise that each SME judgment could be characterized in one of three categories: (a) pessimist, (b) normalist, or (c) optimist. For example, notice M-1 marginal probabilities (prior) for SME-7 in Table 28. All values are relatively low (pessimist) such as management commitment, 0.150, and quality system maturity 0.084, etc. Conversely, SME-8 M-1 marginal probabilities are relatively high (optimist), with management commitment of 0.750 and quality system maturity 0.500, whereas SME-3 and SME-5 are normalist with marginal's near the middle of the road.

Is it possible that SME judgment could have been simulated (Figure 31), cutting five months off CPT build time, and eliminate group think? Would this model provide a broader range of FEF projection say from 10% to 90%? If so, would this model show reduced FEF projection error? The strategy used to build a Simulated SME model (S-SME) was to use the SME developed M-2 structure and develop rules to populate the CPT for each PSF. Weighted posterior judgment methods will be used to assign the likelihood of each model given the observed evidence. As PSF node evidence becomes increasing inadequate, likelihood weight should favor the pessimist and as node evidence becomes increasing adequate, the model should favor the optimist. During the transition from inadequate to adequate, the normalist will be the model of choice.

A few basic S-SME rules must be established as follows:

1. Pessimist believes a worst case scenario is 90% probable.
2. Pessimist range of FEF projection is 0.1-0.5.
3. Normalist range of FEF projection is 0.25-0.75.
4. Optimist believes a best case scenario is 90% probable.
5. Optimist range of FEF projection is 0.5-0.9.
6. Projection variance is be 0.04.
7. Weight of each CPT is a function of S-SME range of belief and the number of parent nodes given by:

$$CPT = \frac{(SSME)_{BeliefRange}}{\# Parents} \tag{24}$$



Pessimist
Mean FEF=0.21

When building CPT
FEF range 0.1-0.5

Normalist
Mean FEF=0.57

When building CPT
FEF range 0.25-0.75

Optimist
Mean FEF=0.76

When building CPT
FEF range 0.5-0.9

Figure 31. Simulated SME

Examples of rule usage are shown in Figure 31. For example, within the Pessimist, one can see two nodes, management commitment and NPI maturity. Management commitment is the parent and NPI maturity the child, conditioned on management commitment. The worst case scenario within the CPT is NPI maturity is low given Management commitment being low. The pessimist notes this cell of the CPT as 90% whereas the worst case scenario for the optimist is 50% per rule 1 and 5 respectively. The best case scenario would be the cell containing the probability of NPI maturity is high given management commitment high. The pessimist best case range is 50%, whereas the optimist believes this is 90% probable.

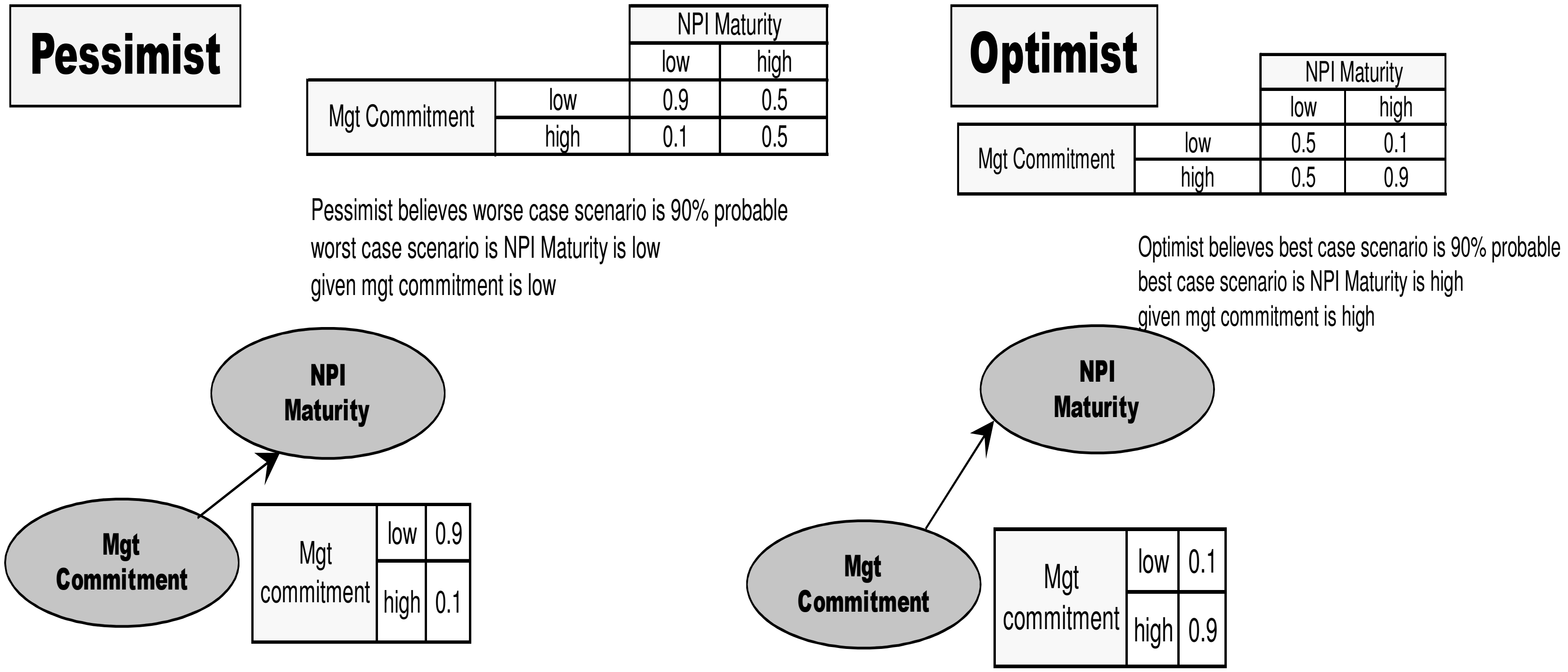
To update a pessimist node with multiple parents one can use rules 1 and 7 as stated above. For example, a pessimist FEF belief is assumed to be between 0.1-0.5. Thus a belief range of 0.4 results in 0.2 FEF weight (Rule 7) per parent state change. Rule 1 indicates the pessimist believes a worst case scenario is 90% probable, thus evidence of parent 1 low and parent 2 low yields a CPT of 90%. As noted in Figure 31, a weight of

0.2 is in effect for every parent state change. This process is repeated to update all nodes CPT with the exception of the FEF node.

Consider a pessimist with all parents adequate or in a positive state. Rule-2 suggests mean FEF projection is 0.5 and given a variance of 0.04 (Rule-6) the pdf for FEF can be determined. This allows for one to populate the FEF CPT given varying parent states (Figure 32, Figure 33).

Simulated SME Detail

One objective for developing the S-SME BBN was to provide a model that has a broader range of FEF projection than those developed by SME. The model should have the ability to trend toward low FEF projection when evidence is negative and high FEF values when evidence is positive. One can see in Table 55 when all evidence is negative, the Pessimist model weight is 97.6%, lowering FEF projection to 18.6%. Conversely Table 56 indicates when all evidence is positive; the Optimist model dominates with a weight of 98% and an FEF projection of 81.6%. In general, the S-SME model range of FEF prediction is from 20-80%, a range that should reduce model error as compared to SME BBN.



Pessimist

		NPI Maturity	
		low	high
Mgt Commitment	low	0.9	0.5
	high	0.1	0.5

Pessimist believes worse case scenario is 90% probable
worst case scenario is NPI Maturity is low
given mgt commitment is low

Mgt commitment	low	0.9
	high	0.1

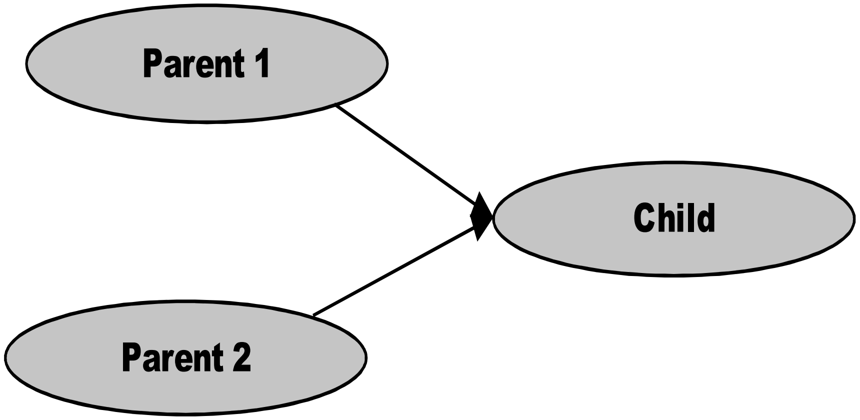
Optimist

		NPI Maturity	
		low	high
Mgt Commitment	low	0.5	0.1
	high	0.5	0.9

Optimist believes best case scenario is 90% probable
best case scenario is NPI Maturity is high
given mgt commitment is high

Mgt commitment	low	0.1
	high	0.9

Figure 32. Simulated SME logic



Variables with two parents			
		Pessimist	Optimist
Parent State		Prob of low state	Prob of high state
Parent 1	Parent 2		
low	low	0.9	0.5
low	high	0.7	0.7
high	low	0.7	0.7
high	high	0.5	0.9

Figure 33. Simulated SME 2-variable logic

One can see in the HVAC projects of Table 56, the simulated SME model, S-SME has less error on all five projects and only 14.6% HVAC overall error versus SME FEF projection error of 61.8%. Thus the S-SME model had 76.4% less error that SME. Additional HVAC project detail can be found in Appendix N. For the automotive projects, one can see in Table 57 S-SME FEF projections has less error in 4 of 5 projects that projections made by SME. S-SME overall automotive average error was 64.8% compared with 96.2 for SME, 32.2% less for the simulated model. Automotive project detail can be found in Appendix O.

Table 58-59 provide an overview of all models with Tables 60-63 providing individual project detail. The simulated SME model has less error on all five projects with only 14.6% HVAC overall error versus SME FEF projection error of 61.8%. Thus the S-SME model demonstrated 76.4% less error that SME within the HVAC industry. For the automotive projects, one can see in Tables 62 and 63, S-SME FEF projections has less error in 4 of 5 projects that projections made by SME. S-SME overall automotive average error was 64.8% compared with 96.2 for SME, 32.2% less for the simulated model. HVAC and automotive S-SME statistical analysis can be found in Appendix P.

Table 56

Simulated SME All Negative Evidence – Automotive Fix Effectiveness (di)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	0	0	0	0.000
Quality system maturity	0.140	0.500	0.860	0	0	0	0.000
Project time	0.161	0.475	0.848	0	0	0	0.000
Failure mode complexity	0.140	0.500	0.860	0	0	0	0.000
Technical expertise	0.159	0.499	0.844	0	0	0	0.000
Resource availability	0.204	0.494	0.850	0	0	0	0.000
Design complexity	0.100	0.500	0.900	0	0	0	0.000
Test facilities	0.193	0.478	0.848	0	0	0	0.000
Fix effectiveness 30%	0.603	0.064	0.008	0.746	0.305	0.018	0.735
Fix effectiveness 30%	0.228	0.198	0.042	0.137	0.316	0.200	0.414
Fix effectiveness 50%	0.123	0.256	0.107	0.078	0.204	0.500	0.081
Fix effectiveness 70%	0.045	0.255	0.330	0.039	0.103	0.190	0.040
Fix effectiveness 90%	0.002	0.228	0.513	0.001	0.037	0.093	0.002
Likelihood	-	-	-	3.36E-01	8.33E-03	7.02E-07	
Prob (Mj)	0.333	0.333	0.333	0.976	0.024	0.000	
				Mean	Variance	Alpha	Beta
				0.186	0.027	0.862	3.765

Table 57

Simulated SME All Positive Evidence – Automotive Fix Effectiveness (di) Projection Trial

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	1	1	1	1.000
Technical expertise	0.159	0.499	0.844	1	1	1	1.000
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 10%	0.603	0.064	0.008	0.095	0.003	0.012	0.012
Fix effectiveness 30%	0.228	0.198	0.042	0.200	0.048	0.040	0.040
Fix effectiveness 50%	0.123	0.256	0.107	0.500	0.137	0.075	0.0076
Fix effectiveness 70%	0.045	0.255	0.330	0.190	0.255	0.099	0.102
Fix effectiveness 90%	0.002	0.228	0.513	0.150	0.577	0.75	0.771
Likelihood	-	-	-	1.02E-06	7.31E-03	3.64E-01	
Prob (Mj)	0.333	0.333	0.333	0.000	0.020	0.980	
				Mean	Variance	Alpha	Beta
				0.816	0.031	3.122	0.703

Table 58

Model Error: BBN Model vs. Actual FEF

	Project					BBN	BBN (M1-M4)
	1	2	3	4	5	Model Avg	Industry Avg
HVAC							
M1	9	36	21	18	14	20	22
M2	6	27	19	10	47	22	
M3	24	11	23	8	54	24	
M4	1	31	20	14	43	22	
SSME	33	3	4	28	5	15	
SME	72	29	100	61	47	62	
Auto							
M1	56	2	84	67	28	47	45
M2	44	15	79	46	54	48	
M3	113	23	32	16	34	44	
M4	111	6	44	38	15	43	
SSME	159	56	21	4	84	65	
SME	240	80	10	5	146	96	

Table 59

Model Error: BBN Model FEF Projections vs. SME FEF Projections

Category	Error %	Difference
Overall BBN	34	-58%
Overal SME	79	
Overall S-SME	40	-49%
HVAC BBN	22	-65%
HVAC SME	62	
HVAC S-SME	15	-76%
Auto BBN	45	-53%
Auto SME	96	
Auto S-SME	65	-32%

Table 60

HVAC Simulated SME FEF Projections by Model Type

		S-SME				
Project	FEF Observed	Pessimist	Normalist	Optimist	S-SME Aggreg	SME
1	58	46.6	76.34	81.7	77.04	100
2	70	39.82	70.91	77.02	72.41	90
3	30	19.14	43.26	62.85	28.79	60
4	62	43.43	73.94	80.28	79.19	99.8
5	49	19.33	43.37	60.77	46.49	72

Table 61

HVAC Simulated SME FEF Projection Error by Model Type

	S-SME				
Project	Pessimist	Normalist	Optimist	S-SME Aggreg	SME
1	20%	-32%	-41%	-33%	-72%
2	43%	-1%	-10%	-3%	-29%
3	36%	-44%	-109%	4%	-100%
4	30%	-19%	-29%	-28%	-61%
5	61%	11%	-24%	5%	-47%
Average				14.6%	61.8%
Difference	76.4%				

Table 62

Automotive Simulated SME FEF Projections by Model Type

		S-SME				
Project	FEF Observed	Pessimist	Normalist	Optimist	S-SME Aggreg	SME
1	26.5	40.8	68.4	68.6	68.52	90
2	50.0	42.9	73	78.6	78.02	90
3	99.8	43.7	74.4	80.4	79.3	90
4	85.3	46.6	76.3	81.7	81.64	90
5	56.9	35.9	65.6	68.7	67.4	90

Table 63

Automotive Simulated SME FEF Projection Error

	S-SME				
Project	Pessimist	Normalist	Optimist	S-SME Aggreg	SME
1	-54%	-158%	-159%	-159%	-240%
2	14%	-46%	-57%	-56%	-80%
3	56%	25%	19%	21%	10%
4	45%	11%	4%	4%	-5%
5	2%	-79%	-88%	-84%	-146%
Average				64.8%	96.2%
Difference	32.2%				

As stated earlier one objective for developing the S-SME BBN was to provide a model with a broader range of FEF projection than those developed by SME. The model should have the ability to trend toward low FEF projection when evidence is negative and high FEF values when evidence is positive. At the time of the S-SME concept, the learning algorithm shortcomings were not known. At that time the fact the researcher assigned equal apportionment to CPT were not of concern, because the learning algorithm would adjust based real world data. Later it was found that the learning algorithm adjustments to CPT accounted for slight improvements in the model, therefore sights were set on determining other areas where model error was prominent, leading to research in dependency among fix effectiveness performance shaping factors, noted in Chapter 6.

Chapter 6: Fix Effectiveness Dependency

Overview

Assume we have two variables x and y, with z a function of x and a function of y. The variance around z is greater if x and y are correlated than if they are not correlated. Applying this analogy to BBN and FEF projections, if we can establish a correlation between FEF, one would then expect more variance in MTBF projections as a result of the correlation. Therefore it becomes a very important aspect of this research that we address whether or not FEF projections are dependent, and if so, how to model them and properly access MTBF uncertainty.

Reliability growth projection model uncertainty is a function of numerous factors. Classical factors include the choice of model that lead to variations in MTBF from choosing say Crow-AMSAA versus AMPM-Stein. Others include estimation of model parameters from rank regression versus maximum likelihood estimates or variation due to the method chosen to calculate confidence interval, i.e. likelihood ratio versus fisher inverse matrix.

Early activities within this research (Appendix Q) have shown uncertainty propagation into MTBF due to FEF projection variability. Uncertainty around each FEF was varied and its effects between Crow-AMSAA and AMPM-Stein reliability growth models analyzed. It was observed that AMPM-Stein was more robust against FEF variation than Crow-AMSAA, however this difference was observed at sample sizes well above those available for most product development programs. The main contributor to MTBF uncertainty was not the model choice, but error associated with fix effectiveness factor variation. MTBF projection error approached 20% as FEF variation rose to 0.05. Thus it became apparent a more robust method of projecting FEF was needed.

This research has provided a methodology of using BBN to project fix effectiveness that results in 38% to 67% reductions in overall FEF projection error as compared to projections made by SME. However, one assumption for the BBN methodology is independence between fix effectiveness factors. Is this assumption valid? What impact would FEF dependence have on FEF projection and ultimately MTBF projection? If FEF are proved dependent, how does uncertainty propagate from one FEF to another? This chapter will augment BBN structures previously developed to show dependency among fix effectiveness factors and provide an example of its impact on MTBF uncertainty.

Fix Effectiveness Dependency Methodology

BBN models were successful in providing a more accurate FEF projection than SME when applied to individual failure modes. Projections were made with the assumption of independence among PSF for other fix activities that might be underway within the engineering staff. One can see from Figure 34 independence implies that fix activities from team-1 have no influence on team-2 and vice versa. Therefore no matter how technically inept, no matter how complex the failure mode, no matter how much testing the lab is required to perform for team-1, independence suggests this activity does not in any way impact team-2.

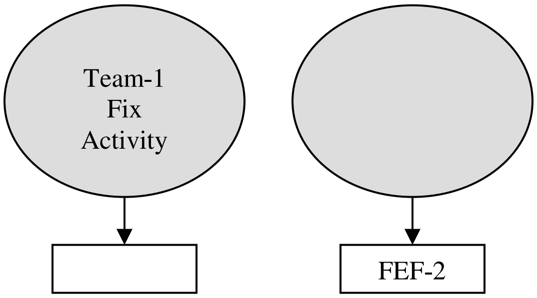


Figure 34. FEF independence

Within an engineering community, our engineering resources may be competing for lab time, or management approval for their respective project, etc. In addition engineering staffs have constraints on resources, test facilities, etc. As resources are redirected to solve one problem versus another, fix effectiveness may be impacted, thus implying dependency among FEF. Figure 35 provides a simple graphic depicting the interconnections of two fix activities. One can see two teams, team-1 and team-2 are attempting to fix their respective failure modes. Both teams report to the same management, have a specific amount of testing capacity along with a finite resource pool.

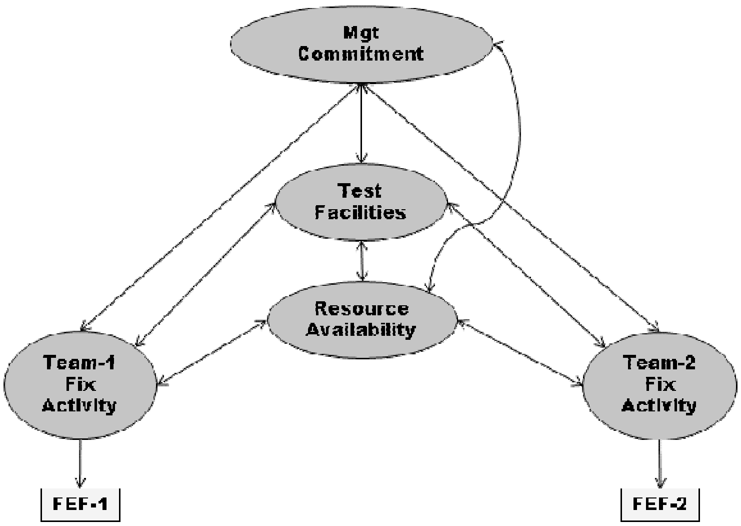


Figure 35. FEF dependency model

As each team competes for these finite resources, tradeoffs are made that impede or enhance a team's ability to fix a problem.

Fix Effectiveness Dependency Composite Model

With this in mind, SME were again reassembled to review the list of PSF and determine those they felt were common between two or more fix activities. SME were asked to imagine the scenario whereby an OEM has two products in the field experiencing one failure mode each. Defects per unit are excessive whereby your customers are adversely affected and warranty dollars charged back to your organization are beyond corporate targets. The product engineering manager attempts to divide resources into two groups, the first group (team-1) to solve failure mode 1 (FM-1) for the first product and the second group (team-2) to solve failure mode 2 (FM-2) for the second product. SME developed a new BBN, mirroring two M-2 models together, as shown in Figure 36.

One can see from Figure 36 the organization has only one group of managers, one test facility, and one resource pool of engineers to solve the two failure modes, thus these nodes are deemed common between the two teams. For example, management commitment impacts the first team's technical expertise, NPI maturity, and time of project. The same group of manager's level of commitment affects the second team's technical expertise, NPI maturity, and their time allotted for fix activities. Failure mode complexity and time of project from both team-1 and team-2 directly impact test facilities, with test facilities then impacting both FEF-1 and FEF-2. The last common node, resource availability is affected by time of project-1 and project-2, which in turn impacts technical expertise of team-1 and team-2, along with FEF-1 and FEF-2. The point to be made is that fix effectiveness is the results of a tangled web of performance shaping factors linked by common nodes. This linkage provides a path for propagation of PSF evidence from one side of the model to the other, creating dependency and impacting FEF on both sides of the model.

When the two structures were connected in Figure 36, SME reached consensus on CPT for common nodes using the same process defined earlier in this research. To understand how evidence propagates through the model, reference the prior marginal probabilities (positive state) for each PSF as shown in Table 64. One can see management commitment; resource availability and test facilities are common between FM-1 and FM-2, whereas the other PSF are not common. Prior to any node evidence, mean FEF projections for FM-1 and FM-2 are both 0.606.

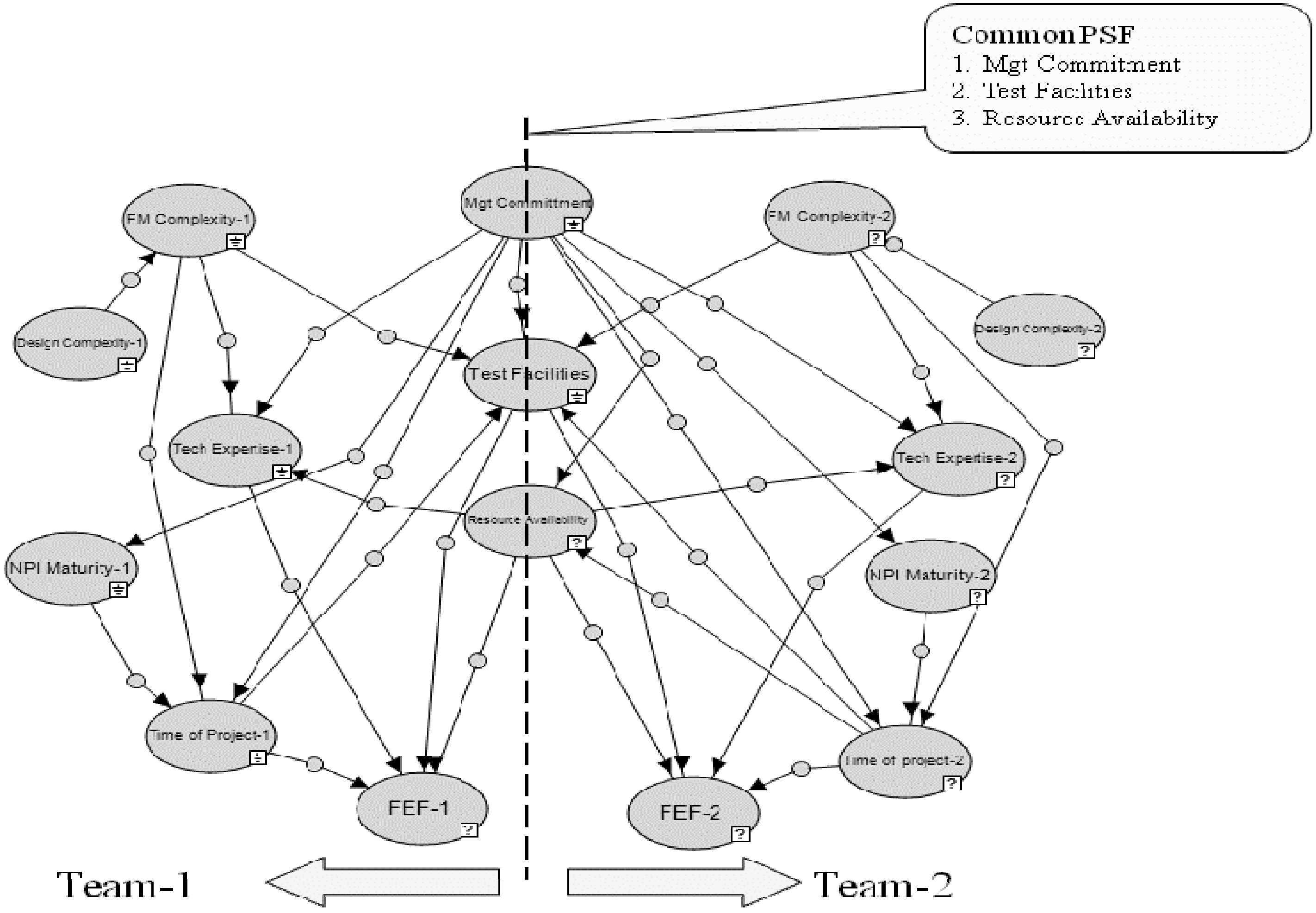


Figure 36. FEF BBN composite model for two failure modes

Table 64

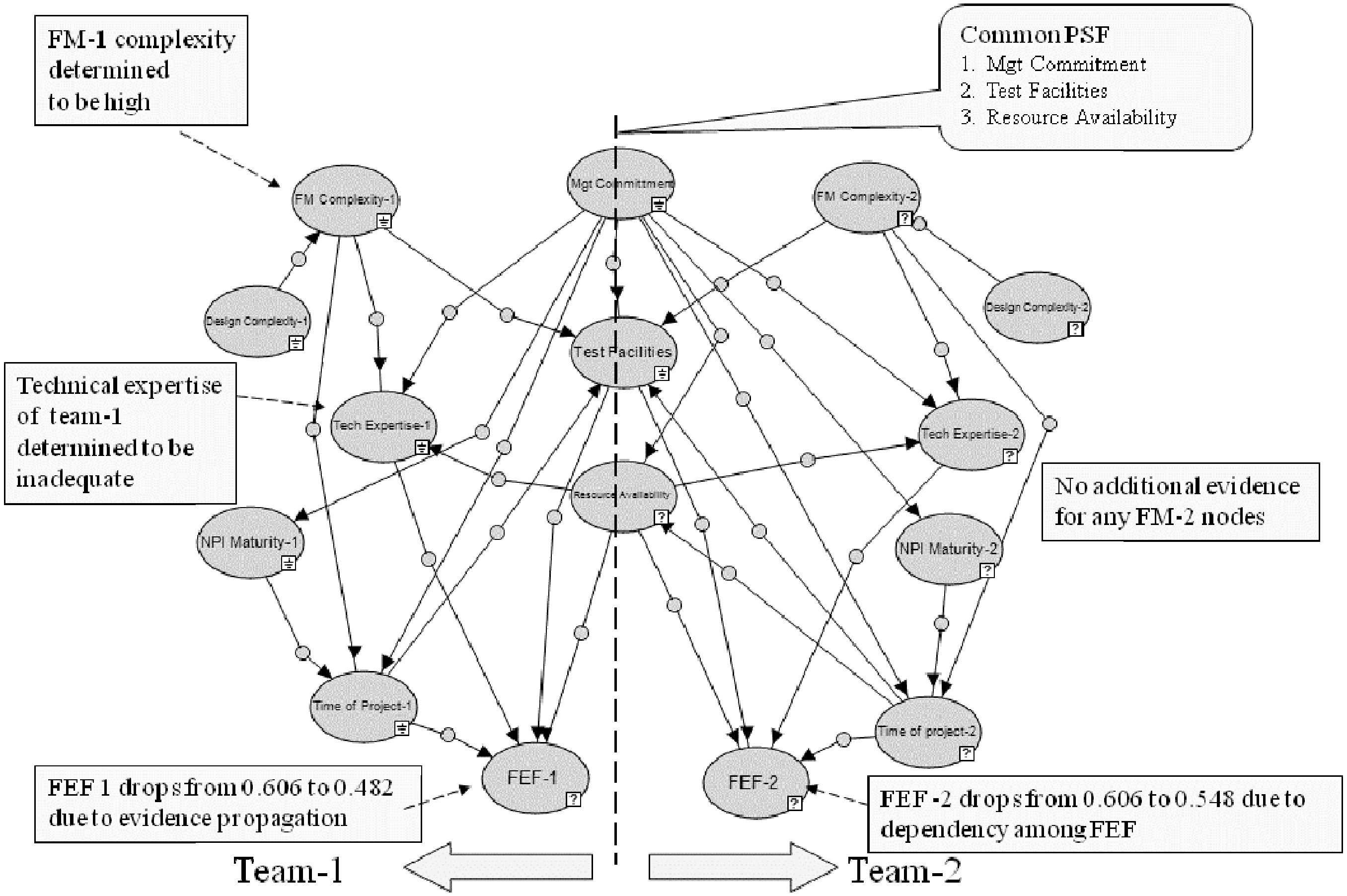
Conditional FEF on PSF – Dependency Model

	Prior		Posterior	
Node Variable	FM-1	FM-2	FM-1	FM-2
Mgt. commitment	0.6		0.442	
Quality system maturity	0.52	0.52	0.489	0.487
Project time	0.629	0.629	0.514	0.552
Failure mode complexity	0.46	0.46	0	0.459
Technical expertise	0.648	0.648	0	0.608
Resource availability	0.612		0.481	
Design complexity	0.3	0.3	0.222	0.3
Test facilities	0.589		0.552	
Fix effectiveness 10%	0.121	0.121	0.258	0.164
Fix effectiveness 30%	0.150	0.150	0.176	0.186
Fix effectiveness 50%	0.161	0.161	0.180	0.176
Fix effectiveness 70%	0.212	0.212	0.173	0.196
Fix effectiveness 90%	0.356	0.356	0.214	0.278
Mean FEF	0.606	0.606	0.482	0.548
Variance	0.079	0.079	0.089	0.084

The engineering team determines that FM-1 complexity is high and the team's technical expertise is deemed inadequate. The product engineering manager assigns these states to their respective nodes (Figure 37) and reruns the BBN simulation. Bayes theorem is used to propagate the new evidence through the model and recalculate node marginal probabilities. As one would expect, given a complex problem (FM-1) and a less than desired technical team-1, FEF-1 dropped from 0.606 to 0.482. In addition the negative evidence associated with FEF-1 negatively impacts FEF-2 via the common nodes, dropping FEF-2 from 0.606 to 0.548, a 9.6% drop, all due to the dependency between the fix effectiveness factors.

In layman terms, one can see two paths of dependency in this example, complexity of failure mode and inept technical expertise. First, given the fact that team-1 has a complex failure mode places strain on test facilities. Possibly the lab does not have the equipment to turn on and off this failure mode, or it will require more equipment etc, dropping the adequacy of the lab by 6.3%. Team-2 uses the same test facilities which are now less adequate, ultimately impacting their fix effectiveness. The same logic can be used to evaluate the impact of inept technical expertise of team-1. The pool of resources was used to supply personnel to team-1, which in turn directly impacts fix effectiveness of team-2.

Analytically one can see from equation 22 -24 the link common nodes provide between FEF-1 and FEF-2. The reduction in FEF-2 is due to the dependencies of the fix activities between the two teams. The marginal probability of FEF-2 (Equation 22) is a function of test facilities and resources (common nodes), along with time and technical expertise associated with FM-1. Looking at just one of the common nodes, P(T) probability of test facilities (Equation 23), is a function of failure mode complexity for both FM-1 and FM-2 along with time associated with the fix for both team-1 and team-2 and TI is a function of management commitment, and times associated with both team-1 and team-2 activities. The analytics of the equations provide evidence of fix effectiveness dependency among multiple fix activities.



FM-1 complexity determined to be high

Common PSF
1. Mgt Commitment
2. Test Facilities
3. Resource Availability

Technical expertise of team-1 determined to be inadequate

No additional evidence for any FM-2 nodes

FEF 1 drops from 0.606 to 0.482 due to evidence propagation

FEF -2 drops from 0.606 to 0.548 due to dependency among FEF

Figure 37. FEF BBN for two failure modes (FM-1 evidence)

$$\begin{aligned}P(FEF_2) = {} & P\left(FEF_2 \middle| T, R, Ti_2, TE_2\right) P(T) P(R) P(Ti_2) P(TE_2) + \\ & P\left(FEF_2 \middle| \overline{T}, R, Ti_2, TE_2\right) P(\overline{T}) P(R) P(Ti_2) P(TE_2) + \\ & \ldots\ldots\ldots\ldots + P\left(FEF_2 \middle| \overline{T}, \overline{R}, \overline{T}i_2, \overline{TE}_2\right) P(\overline{T}) P(\overline{R}) P(\overline{T}i_2) P(\overline{TE}_2)\end{aligned} \tag{22}$$

$$\begin{aligned}P(T) = {} & P\left(T \middle| FMC_1, Ti_i, FMC_2 Ti_2\right) P(FMC_1) P(Ti_1) P(FMC_2) P(Ti_2) \\ & + P\left(T \middle| \overline{FMC}_1, Ti_i, FMC_2 Ti_2\right) P(\overline{FMC}_1) P(Ti_1) P(FMC_2) P(Ti_2) \\ & \ldots\ldots\ldots\ldots + P\left(T \middle| \overline{FMC}_1, \overline{T}i_i, \overline{FMC}_2 \overline{T}i_2\right) P(\overline{FMC}_1) P(\overline{T}i_1) P(\overline{FMC}_2) P(\overline{T}i_2)\end{aligned} \tag{23}$$

$$\begin{aligned}P(R) = {} & P\left(R \middle| M, Ti_i, Ti_2\right) P(M) P(Ti_1) P(Ti_2) \\ & \ldots\ldots\ldots + P\left(R \middle| \overline{M}, Ti_i, Ti_2\right) P(\overline{M}) P(Ti_1) P(Ti_2) \\ & \ldots\ldots\ldots\ldots + P\left(R \middle| \overline{M}, \overline{T}i_i, \overline{T}i_2\right) P(\overline{M}) P(\overline{T}i_1) P(\overline{T}i_2)\end{aligned} \tag{24}$$

Note:

FMC=failure mode complexity
M=management commitment
R=resources
T= test facilities
TE=technical expertise
Ti=time

<u>Dependency Impact on Fix Effectiveness</u>

In order to more fully understand the impact of dependency among fix effectiveness factors a DOE was ran varying non-common nodes for team-1 side of the model and observing the impact on FEF-2. No PSF are changed for team-2 side of the model, i.e. FEF-2. The DOE chosen is a full factorial (32 treatment combinations) 2^5 Plackett-Burman. Factors and factor levels include:

1. Time of project-1 (adequate, inadequate)
2. NPI Maturity-1 (low, high)
3. Technical Expertise-1 (adequate, inadequate)

4. Design Complexity-1 (low, high)
5. Failure Mode Complexity-1(low, high)

FEF-2 error is measured as PSF factors are varied among treatment combinations. Error is defined as the percent difference in FEF-2 with no PSF evidence on either side of the model versus "with" evidence on the team-1 side of the model. If independent, FEF-2 will not change as PSF are varied on team-1.

Results of the experiment are shown in Figures 38 and Figure 39. One can see in Figure 38 that time-1, NPI Maturity-1 and technical expertise and their interactions from team-1 are significant (alpha=0.05) in affecting FEF-2. In addition the main effects plot of Figure 38 confirms that time-1 has the greatest impact on FEF-2 with technical expertise-1 a close second and NPI maturity a distant third. The regression equation indicates that (Tables 65, 66) 90.3% of FEF-2 error can be assigned to FEF-1 PSF.

As the DOE treatment combinations were administered to the composite BBN, FEF-2 error ranged from minus 11.7% to plus 21.2%. Using the mean and 95% CI from the ANOVA's ran in Chapter 4, we applied the FEF error range to the output of each BBN model, M1-M4. Table 67 indicates that inclusion of FEF dependency within the BBN models would result in 4 of 5 HVAC projects and 2, almost 3 of the automotive projects reaching statistical significance when comparing BBN FEF projections to actual fix effectiveness obtained from field data (Table 68).

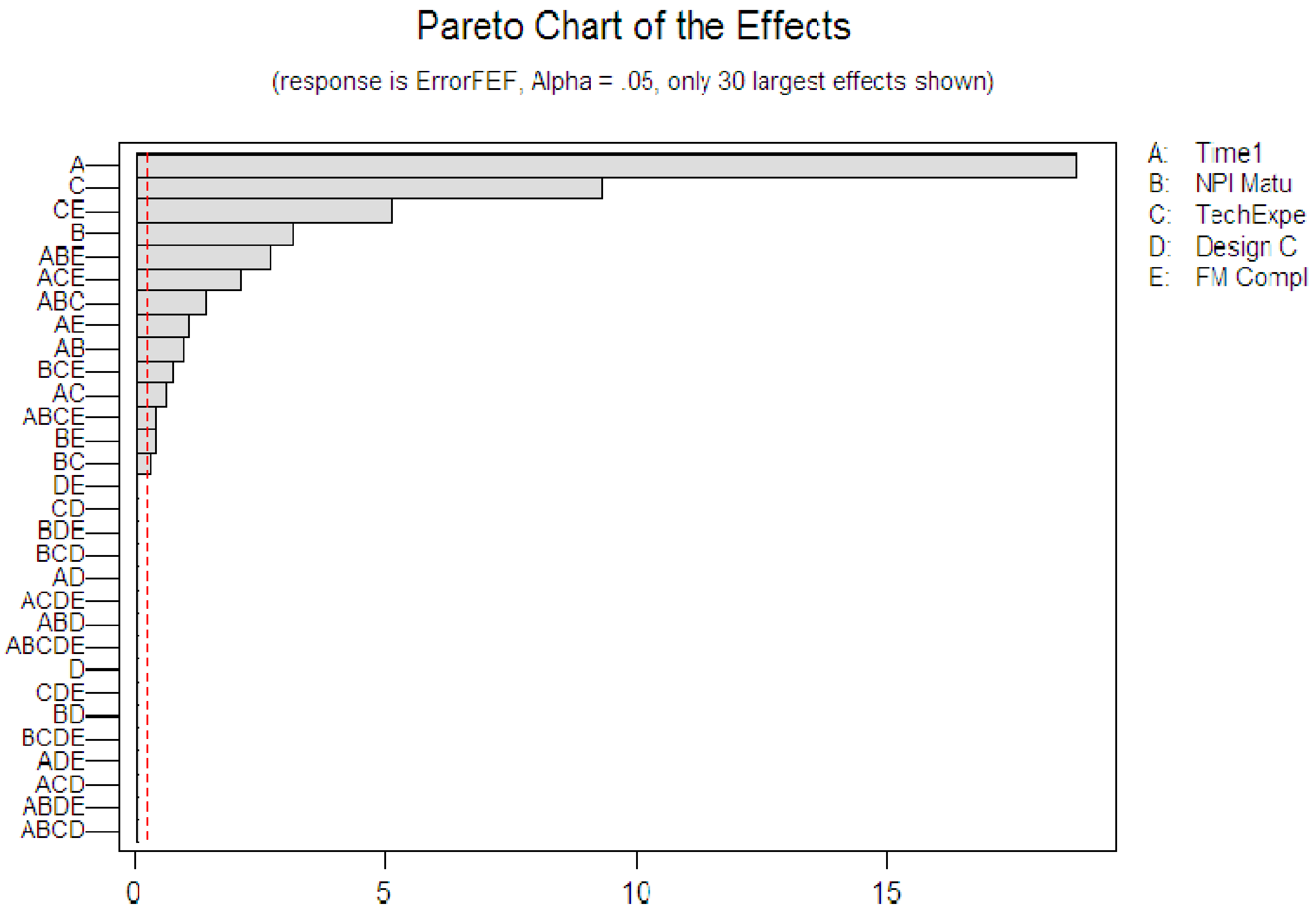


Figure 38. Pareto of DOE

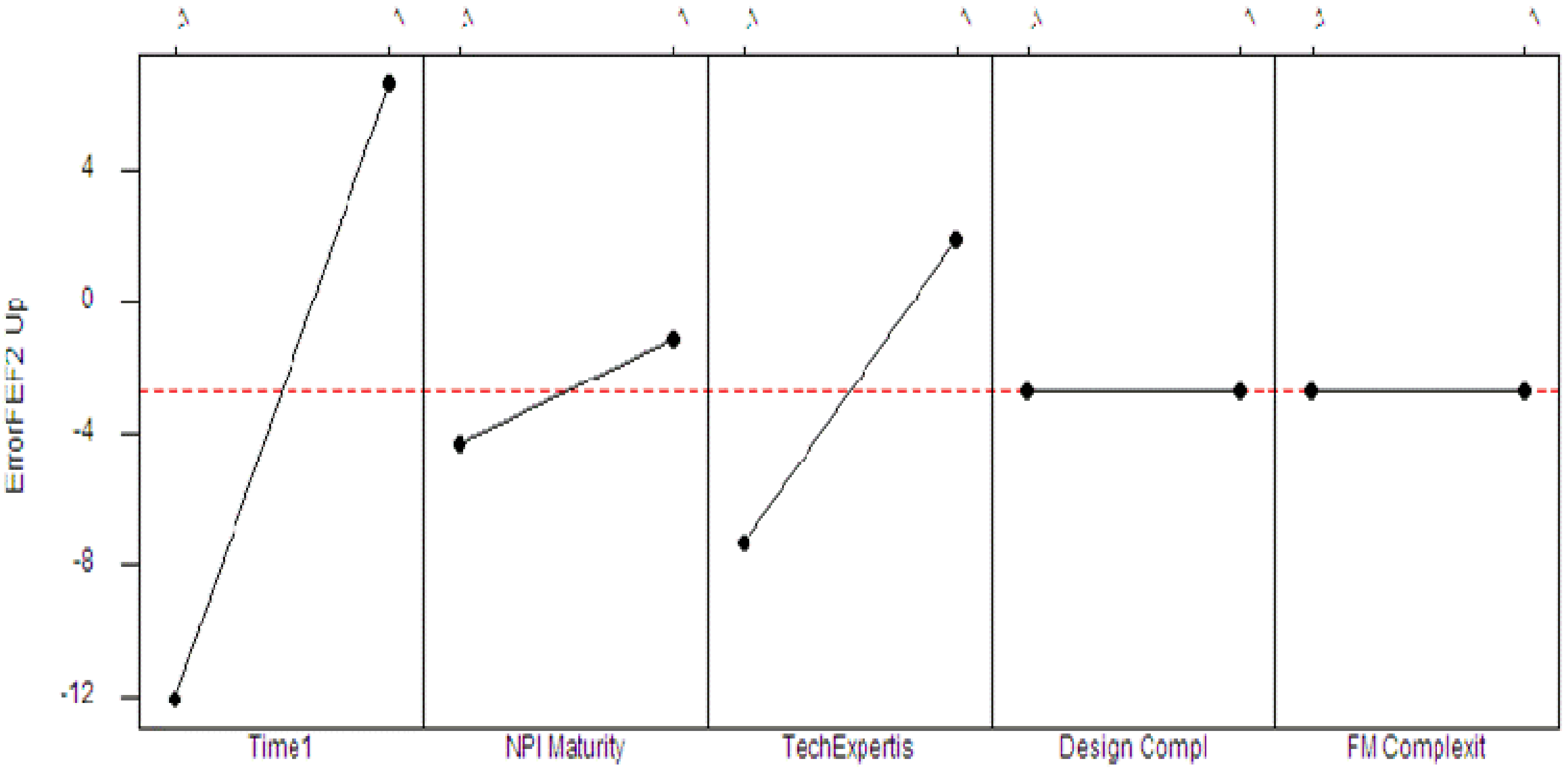


Figure 39. Main effects for FEF-2 error

Table 65

Regression Analysis: Error FEF2 Updated vs Time 1, NPI Maturity 1

Predictor	Coef	SE Coef	T	P
Constant	-2.7326	0.6194	-4.41	0.000
Time1	9.3807	0.6194	15.14	0.000
NPI Maturity	1.5702	0.6194	2.54	0.017
TechExpe	4.6560	0.6194	7.52	0.000

Note: Regression equation is ErrorFEF2 Updated = -2.736 + 9.38 Time 1 + 1.57 NPI Maturity 1 + 4.66 TechExpertise 1; S = 3.504; R-Sq = 91.3%; R-Sq(adj) = 90.3%

Table 66

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	3588.5	11963.2	97.43	0.000
Residual Error	28	343.7	12.3		
Total	31	3932.3			

Source	DF	Seq SS
Time1	1	2815.9
NPI Maturity	1	78.9
TechExpe	1	693.7

Table 67

FEF Control Limit Spread due to FEF Dependency – HVAC

			HVAC Projects Evaluated with Dependency Variation of FEF											
				Independent			Dependent Minus Error			Dependent Plus Error				
FEF Error		Auto	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL	FEF-2 Error as a Function of FEF-1 Activity		
- error	11.7%	Project 1										Dependent (-)	Independent	Dependent (+)
+ error	21.2%	M-1	0.52088	0.50194	0.53981	0.45994	0.44100	0.47887	0.63131	0.61237	0.65024	-26.1%	-141.4%	8.1%
		M-2	0.62700	0..60610	0.64790	0.55364	0.53274	0.57454	0.75992	0.73902	0.78082	-4.8%	7.5%	23.7%
		M-3	0.71450	0.63970	0.73530	0.63080	0.55610	0.65170	0.86597	0.79117	0.88677	8.1%	18.8%	33.0%
Actual FEF 0.58		M-4	0.58100	0.56070	0.60120	0.51302	0.49272	0.53322	0.70417	0.68387	0.72437	-13.1%	0.2%	17.6%
		Project 2												
		M-1	0.44153	0.42336	0.45969	0.38987	0.37170	0.40803	0.53513	0.51696	0.55329	-79.5%	-58.5%	-30.8%
		M-2	0.51770	0.49740	0.53800	0.45713	0.43683	0.47743	0.62745	0.60715	0.64775	-53.1%	-35.2%	11.6%
Actual FEF 0.7		M-3	0.60330	0.58850	0.62810	0.53713	0.51733	0.55693	0.73726	0.71746	075708	-30.3%	-15.1%	5.1%
		M-4	0.49836	0.47892	0.51780	0.44005	0.42081	0.45949	0.60401	0.58457	0.62345	-59.1%	-40.5%	-15.9%

Table 67 (*continued*).

HVAC Projects Evaluated with Dependency Variation of FEF													
		Independent			Dependent Minus Error			Dependent Plus Error					
FEF Error	Auto	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL	FEF-2 Error as a Function of FEF-1 Activity		
	Project 3												
Actual FEF 0.3	M-1	0.23475	0.21701	0.25249	0.20728	0.18954	0.22502	0.28452	0.26678	0.30226	-44.7%	-27.8%	-5.4%
	M-2	0.24800	0.22780	0.26810	0.21898	0.19878	0.23903	0.30058	0.28038	0.32068	-37.0%	-21.0%	0.2%
	M-3	023068	0.21236	0.24900	0.20369	0.1853.7	0.22201	0.27958	0.26126	0.29790	-47.3%	-30.1%	-7.3%
	M-4	0.25240	0.23230	0.27250	0.22287	0.20277	0.24297	0.30591	0.28581	0.32601	-34.6%	-18.9%	1.9%
	Project 4												
Actual FEF 0.62	M-1	0.50900	0.49078	0.52725	0.44945	0.43123	0.46770	0.61691	0.59869	0.63516	-37.9%	-21.8%	-0.5%
	M-2	0.56220	0.54020	0.58420	0.49642	0.47442	0.51842	0.68139	0.65939	0.70339	-24.9%	-10.3%	9.0%
	M-3	0.6745	0.65410	0.69670	0.59638	0.57508	0.61768	0.81858	0.79728	0.838988	-4.0%	8.2%	24.3%
	M-4	0.53860	0.51880	0.55840	0.47588	0.45578	0.49538	0.65278	0.63298	0.67258	-30.4%	-15.1%	5.0%
	Project 5												
Actual FEF 0.49	M-1	0.42032	0.40225	0.43838	0.37114	0.35307	0.38920	0.50943	0.49136	0.52749	-32.0%	-16.6%	3.8%
	M-2	0.26747	0.24978	0.28516	0.23618	0.21849	0.25387	0.32417	0.30648	0.34186	-107.5%	-83.2%	-51.2%
	M-3	0.22453	0.20957	0.23950	0.19826	0.18330	0.21213	0.27213	0.25717	028710	-147.2%	-118.2%	-80.1%
	M-4	0.29876	0.28027	0.31726	0.26381	0.24532	0.28231	0.36210	0.34381	0.38060	-85.7%	-64.0%	-35.3%

Note: Indicates statistical significance Confidence intervals are 95% Indicates very close

Table 68

FEF Control Limit Spread due to FEF Dependency – Automotive

			HVAC Projects Evaluated with Dependency Variation of FEF											
				Independent			Dependent Minus Error			Dependent Plus Error				
FEF Error		Auto	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL	FEF-2 Error as a Function of FEF-1 Activity		
- error	11.7%	Project 1										Dependent (-)	Independent	Dependent (+)
+ error	21.2%	M-1	0.60484	0.58645	0.62324	0.53407	0.51568	0.55247	0.73307	0.71468	075147	50.4%	56.2%	63.9%
		M-2	0.48290	0.46160	0.50420	.042640	0.40510	0.44770	0.58527	0.56397	0.60657	37.9%	45.1%	54.7%
Actual FEF 0.265		M-3	0.54960	0.52970	0.56950	0.48530	0.46640	0.50520	0.66612	0.64622	0.68602	45.4%	51.8%	60.2%
		M-4	0.56400	0.54340	0.58450	0.49601	0.47741	0.51851	0.68357	0.66297	0.70407	46.8%	53.0%	61.2%
		Project 2												
		M-1	0.52004	0.50164	0.53844	0.45920	0.44080	0.47760	0.63029	0.61189	0.64869	-8.9%	3.9%	20.7%
		M-2	0.57454	0.55506	0.59402	0.50732	0.48784	0.52690	0.696934	0.67686	0.71582	1.4%	13.0%	28.2%
Actual FEF 0.5		M-3	0.58950	0.56750	0.61140	0.52053	0.49853	0.54243	0.71447	0.69247	0.73637	3.9%	15.2%	30.0%
		M-4	0.52592	0.50676	0.54507	0.46439	0.44523	0.48354	0.63742	0.61826	0.65657	-7.7%	4.9%	21.6%

Table 68(*continued*).

HVAC Projects Evaluated with Dependency Variation of FEF													
			Independent			Dependent Minus Error			Dependent Plus Error				
FEF Error	Auto	Mean	LCL	UCL	Mean	LCL	UCL	Mean	LCL	UCL	FEF-2 Error as a Function of FEF-1 Activity		
	Project 3												
Actual FEF 0.998	M-1	0.53665	0.51809	0.55520	0.47386	0.45530	0.49241	0.65042	0.63186	0.66897	-110.6%	-86.0%	-53.4%
	M-2	0.57550	0.55260	0.59830	0.50817	0.48527	0.53097	0.69751	0.67481	0.720361	-96.4%	-73.4%	-43.1%
	M-3	0.67330	0.64890	0.69770	0.59452	0.57012	0.61892	0.81604	0.79164	0.84044	-67.9%	-48.2%	-22.3%
	M-4	0.56000	0.53870	0.58140	0.49448	0.47318	0.51588	0.67872	0.65742	0.70012	-101.8%	-78.2%	-47.0%
	Project 4												
Actual FEF 0.62	M-1	0.50912	0.49151	0.52673	0.44955	0.43194	0.46716	0.61705	0.59944	0.63466	-89.7%	-67.5%	-38.2%
	M-2	0.58410	0.56100	0.60730	0.51576	0.49266	0.53896	0.70793	0.69453	0.73113	-65.4%	-46.0%	-20.5%
	M-3	0.70100	0.67690	0.72500	0.61898	0.59488	0.64298	0.84981	0.82551	0.87361	-37.8%	-21.7%	-0.4%
	M-4	0.51856	0.49691	0.53821	0.45789	0.43824	0.47754	0.62849	0.60884	0.64814	-86.3%	-64.5%	-35.7%
	Project 5												
Actual FEF 0.49	M-1	0.45371	0.43580	0.47161	0.40063	0.38272	0.41853	0.54990	0.53199	0.56780	-42.0%	-25.4%	-3.5%
	M-2	0.36739	0.34791	0.38686	0.32441	0.30493	0.34388	0.44528	0.42580	0.46475	-75.4%	-54.9%	-27.8%
	M-3	0.50810	0.48450	0.53180	0.44365	0.42505	0.47235	0.61582	0.59222	0.63952	-26.8%	-12.0%	7.6%
	M-4	0.41538	0.39904	0.43173	0.36678	0.35044	0.38313	0.50344	0.48710	0.51979	-55.1%	-37.0%	-13.0%

Note: Indicates statistical significance Confidence intervals are 95% Indicates very close

Fix Effectiveness Dependency Impact on MTBF

Now let's turn our attention to understanding the impact FEF dependencies have on MTBF projection by evaluating the following four scenarios.

1. PSF evidence for FEF-1 non common nodes only.
2. Same PSF evidence applied to both sides of the composite model (non-common nodes only).
3. Different PSF evidence applied to both sides of composite model
    a. Requires common PSF to be of the same state
4. Common nodes favoring one team versus the other
    a. Requires SME augmentation of common node CPT

FEF from each of the aforementioned scenarios will be used in the Crow-AMSAA model to project MTBF within the following example. Assume a particular product has been on test for 1000 hours. SME believe at least 10 failure modes exists. Thus far we have experienced two failure modes, FM-1 and FM-2, with 4 occurrences of FM-1 occurring at 15, 102, 249 and 273 hours and 3 occurrences of FM-2 at 112, 285 and 317 hours. Two teams are established to solve FM-1 and FM-2, thus all modes will receive corrective action. Node evidence and FEF are shown in each of the scenario examples (Table 69). For example refer to Scenario-1-Example-1, for team-1 common nodes of management commitment, test facilities, and resource availably, no evidence is available. The remaining nodes, failure mode complexity, design complexity, technical expertise, NPI maturity and time of project are considered in a "negative" state.

Table 69

Scenario-1 PSF Evidence Applied to Team-1 Only

	Example 1				Example 2				Example 3			
	Team				Team				Team			
	1	2	MTBF	Error	1	2	MTBF	Error	1	2	MTBF	Error
Mgt. Commitment	-	-			-	-			-	-		
Test Facilities	-	-			-	-			-	-		
Resource Availability	-	-			-	-			-	-		
FM Complexity	High	-			High	-			Low	-		
Design Complexity	High	-		7.7% over-stated MTBF	Low	-		2.2% under stated MTBF	Low	-		5.5% under stated MTBF
NPI Maturity	Low	-			High	-			Adequate	-		
Technical Expertise	Inadequate	-			Inadequate							
Time of Project	Inadequate	-			Adequate	-			Adequate			
FEF considered independent	0.342	0.606	243.26		0.622	0.606	324.84		0.712	0.606	364.09	
FEF actual dependency	0.342	0.478	224.48		0.622	0.631	332.09		0.712	0.659	384.00	

No evidence is observed for team-2. When each FEF is considered independent, FEF-1 equates to 0.342 and FEF-2 0.606, with the resulting MTBF 243.26. However, if evidence is entered into the composite BBN dependency model, FEF-1 is 0.342 and FEF-2 falls to 0.478, pulling MTBF down to 224.48. Without consideration of the dependency between fix effectiveness factors, MTBF will be overstated by 7.7% (Reference equations below).

Per information supplied earlier, the number of surfaced failure modes m is 2, the total time on test is 1000 hours and first time occurrence of FM-1 and FM-2 is 15 and 112 hours respectively. Therefore,

MTBF using FEF considered independent

$$\rho(T)=\lambda_A+\sum_{i=1}^{m}(1-d_i)\lambda_i+\bar{d}h(t)=\lambda_A+\sum_{i=1}^{m}(1-d_i)\lambda_i+\bar{d}\hat{\lambda}_{BD}\hat{\beta}_{BD}t^{\beta-1}$$

$$\rho(T)=0+(1-0.342)\frac{4}{1000}+(1-0.606)\frac{3}{1000}+0.474(0.230)(0.313)\left(1000^{0.313-1}\right)=0.004111$$

$$MTBF=\frac{1}{\rho(T)}=\frac{1}{0.004111}=243.26$$

where $\hat{\beta}_{BD}=\dfrac{m}{\sum_{i=1}^{m}\ln\left(\dfrac{T}{X_i}\right)}$, $\hat{\lambda}_{BD}=\dfrac{m}{T^{\hat{\beta}_{BD}}}$ and $\bar{d}=\dfrac{\sum_{i=1}^{m}d_i}{m}$

MTBF using FEF with dependency

$$\rho(T)=\lambda_A+\sum_{i=1}^{m}(1-d_i)\lambda_i+\bar{d}\hat{\lambda}_{BD}\hat{\beta}_{BD}t^{\beta-1}$$

$$\rho(T)=0+(1-0.342)\frac{4}{1000}+(1-0.468)\frac{3}{1000}+0.41(0.230)(0.313)\left(1000^{0.313-1}\right)=0.004455$$

$$MTBF=\frac{1}{\rho(T)}=\frac{1}{0.004455}=224.48$$

Table 70 provides a full factorial example of non-common nodes and the effects PSF for Team-1 have on MTBF projections for team-2. Non-common PSF states are varied for team-1 with team-2 PSF unobserved. FEF for both teams are recorded. FEF-1 varies with various treatment combinations while FEF-2 varies from minus 4.4% to plus 8.7% due to the dependencies between the fix activities of the teams. Using the Crow-AMSAA model, FEF values for each team are used to project failure intensity and MTBF. For the two failure mode example, one can see MTBF error range from minus 14.5% to plus 8.5%. If the number of failure modes is increased to five, MTBF error increases as well, ranging from minus 16.5% to plus 13.7%.

Scenario-2 applies the same PSF evidence to each side of the model independently and then simultaneously to both sides of the model, while noting the impact to fix effectiveness. Example-1 indicates that FEF-1 and FEF-2 equate to 0.342 for both sides of the model when PSF are applied to one side and then the other. When simultaneously applied to both sides of the model, both fix effectiveness factors equal 0.273, indicating again a dependency between FEF. The resulting MTBF drops from 207.46 down to 190.12, thus MTBF was overstated by 8.4% when FEF are considered independent (Note examples 2-3 – Table 71).

Table 70

MTBF Error as a Function of FEF Dependency

Run	DOE Treatment Combinations					FEF-1	Independent FEF-2	Dependent FEF-2	Independent failure intensity	Dependent failure intensity	MTBF error 2 FM	MTBF error 5 FM
	Time 1	NR Maturity 1	Tech Expertise 1	Design Complexity 1	FM Complexity 1							
1	-1	-1	-1	-1	-1	0.3423	0.6060	0.4776	0.0041	0.0045	8.5%	13.7%
2	1	-1	-1	-1	-1	0.5995	0.6060	0.5900	0.0032	0.0032	0.3%	2.2%
3	-1	1	-1	-1	-1	0.3474	0.6060	0.4852	0.0041	0.0045	8.3%	13.1%
4	1	1	-1	-1	-1	0.6219	0.96060	0.6309	0.0031	0.0030	-2.9%	-3.69%
5	-1	-1	1	-1	-1	0.5670	0.6060	0.5707	0.0033	00.0034	3.8%	4.6%
6	1	-1	1	-1	-1	0.7276	0.6060	0.6771	0.0027	0.0025	-8.9%	-11.8%
7	-1	1	1	-1	-1	0.5257	0.6060	0.5848	0.0034	0.0035	2.6%	2.8%
8	1	1	1	-1	-1	0.7403	0.6060	0.7006	0.0026	0.0023	-14.0%	-16.5%
9	-1	-1	-1	1	-1	0.3423	0.6060	0.4776	0.0041	0.0045	8.5%	13.7%
10	1	-1	-1	1	-1	0.6010	0.6060	0.5900	0.0032	0.0032	0.3%	2.2%
11	-1	1	-1	1	-1	0.3474	0.6060	0.4852	0.0041	0.0045	8.3%	13.1%
12	1	1	-1	1	-1	0.6220	0.6060	0.6309	0.0031	0.0030	-2.9%	-3.6%
13	-1	-1	1	1	-1	0.5070	0.9060	0.5707	0.0035	0.00366	3.5%	4.5%
14	1	-1	1	1	-1	0.7276	0.6060	0.6771	0.0027	0.0025	-8.9%	-11.8%
15	-1	1	1	1	-1	0.5747	0.6060	0.5847	0.0033	0.0033	2.8%	2.9%
16	1	1	1	1	-1	0.7403	0.6060	0.7006	0.0026	0.0023	-13.3%	-16.5%

Table 70 (*continued*).

Run	DOE Treatment Combinations					FEF-1	Independent FEF-2	Dependent FEF-2	Independent failure intensity	Dependent failure intensity	MTBF error 2 FM	MTBF error 5 FM
	Time 1	NR Maturity 1	Tech Expertise 1	Design Complexity 1	FM Complexity 1							
17	-1	-1	-1	-1	1	0.3844	0.6060	0.5203	0.0040	0.0042	6.2%	9.8%
18	1	-1	-1	-1	1	0.6174	0.6060	0.6225	0.0039	0.0041	-2.3%	-2..4%
19	-1	1	-1	-1	1	0.3932	0.6060	0.5416	0.0039	0.0041	4.8%	7.6%
20	1	1	-1	-1	1	0.6173	0.6060	0.6182	0.0031	0.0030	-2.2%	-1.7%
21	-1	-1	1	-1	1	0.5246	0.6060	0.5124	0.0034	0.0037	8.0%	11.2%
22	1	-1	1	-1	1	0.7192	0.6060	0.6688	0.0027	0.0025	-7.9%	-10.2%
23	-1	1	1	-1	1	0.5580	0.6060	0.5682	0.0033	0.0034	3.8%	4.9%
24	1	1	1	-1	1	0.7122	0.6060	0.6587	0.0027	0.0026	-7.2%	-8.4%
25	-1	-1	-1	1	1	0.3844	0.6060	0.5203	0.0040	0.0042	6.2%	9.8%
26	1	-1	-1	1	1	0.6174	0.6060	0.6225	0.0031	0.0030	-2.3%	-2.4%
27	-1	1	-1	1	1	0.3932	0.6060	0.5416	0.0039	0.0041	4.8%	7.6%
28	1	1	-1	1	1	0.6177	0.6060	0.6252	0.0031	0.0030	-2.9%	-2.8%
29	-1	-1	1	1	1	0.5246	0.6060	0.5124	0.0034	0.0037	8.0%	11.2%
30	1	-1	1	1	1	0.7186	0.6060	0.6688	0.0027	0.0025	-7.9%	-10.2%
31	-1	1	1	1	1	0.5580	0.6060	0.5682	0.0033	0.0034	3.8%	4.9%
32	1	1	1	1	1	0.7122	0.6060	0.6587	0.0027	0.0024	-14.5%	-8.4%

Table 71

Scenario-2: Same PSF Applied to Both Sides of Composite Model

	Example 1				Example 2				Example 3			
	Team				Team				Team			
	1	2	MTBF	Error	1	2	MTBF	Error	1	2	MTBF	Error
Mgt. Commitment	-	-			-	-			-	-		
Test Facilities	-	-			-	-			-	-		
Resource Availability	-	-			-	-			-	-		
FM Complexity	High	High			High	High			Low	Low		
Design Complexity	High	High		8.4% over-stated MTBF	Low	Low		5.2% under stated MTBF	Low	Low		8.1% under stated MTBF
NPI Maturity	Low	Low-			High	High			Adequate	Adequate		
Technical Expertise	Inadequate	Inadequate			Inadequate	Inadequate			Adequate	Adequate		
Time of Project	Inadequate	Inadequate			Adequate	Adequate			Adequate	Adequate		
FEF considered independent	0.342	0.342	207.46		0.622	0.622	329.44		0.712	0.712	406.21	
FEF actual dependency	0.273	0.273	190.12		0.596	0.596	312.39		0.741	0.741	439.19	

Scenario-3 applies different PSF evidence to each side of the model independently and then simultaneously to both sides of the model, while noting the impact to fix effectiveness. Example-1 indicates that FEF-1 equates to 0.188 and FEF-2 to 0.225 when PSF are applied to one side and then the other. When simultaneously applied to both sides of the model, both fix effectiveness factors remain the same indicating no dependency. This is significant in that for the first time all common nodes PSF evidence is known resulting in a minimal impact on MTBF variation (Table 72).

Scenario-4 involves SME reconstruction of PSF CPT for management commitment and resource availability such that PSF states include no evidence, positive states for team-1 and team-2. This allows one to study “direction of influence” effects on FEF. Example 1 indicates management is committed to help team-1 by showing their commitment to the team and providing adequate resources relative to team-2. Evidence is applied independently to each side of the composite model with FEF-1 equating to 0.352 and FEF-2 0.336. As evidence is simultaneously applied to both sides one can see the favored team is virtually unaffected by team-2, however team-2 FEF indicates a slight dependency on team-1 and reduces MTBF projections by 0.3%. In the second example, team-2 is favored resulting in FEF-2 independent of FEF-1, but the un-favored team’s FEF-1 again indicates a minute level of dependency on FEF-2 (Table 73).

Table 72

Scenario-3 Different PSF Applied to Both Sides of Composite Model

	Example 1				Example 2			
	Team				Team			
	1	2	MTBF	Error	1	2	MTBF	Error
Mgt. Commitment	Low	Low-		MTBF Properly Stated 0% Error	-	-		0.8% under stated MTBF
Test Facilities	Inadequate	Inadequate			-	-		
Resource Availability	Inadequate	Inadequate			-	-		
FM Complexity	High	High			High	High		
Design Complexity	High	Low			Low	Low		
NPI Maturity	Low	High			High	High		
Technical Expertise	Inadequate	Inadequate			Inadequate	Inadequate		
Time of Project	Inadequate	Inadequate			Adequate	Adequate		
FEF considered independent	0.188	0.225	175.37		0.622	0.622	329.44	
FEF actual dependency	0.188	0.225	175.37		0.596	0.596	312.39	

Table 73

Scenario-4 Common Nodes Favoring One Team More Than Another

	Example 1				Example 2			
	Team				Team			
	1	2	MTBF	Error	1	2	MTBF	Error
Mgt. Commitment	H Team 1	H Team 1		9% over stated MTBF	H Team 2	H Team 2		0.1% under stated MTBF
Test Facilities	-	-			-	-		
Resource Availability	H Team 1	H Team 1			A Team 2	A Team 2		
FM Complexity	High	Low			High	Low		
Design Complexity	High	Low			High	Low		
NPI Maturity	Low	High			Low	Adequate		
Technical Expertise	Inadequate	Inadequate			Inadequate	Inadequate		
Time of Project	Inadequate	Inadequate			Inadequate	Inadequate		
FEF considered independent	0.3582	0.336	208.36		0.286	0.468	213.22	
FEF actual dependency	0.351	0.332	207.74		0.287	0.468	213.26	

Summary: Fix Effectiveness Dependency

In summary, we have used four scenarios to show that when common nodes are unobserved, regardless as to the state of non-common node evidence (same or different), FEF are dependent. In addition, for this model structure, to break dependency, evidence must be observed for all common nodes.

These are significant discoveries. First, we have developed a methodology of predicting fix effectiveness prior to obtaining test results and second, augmented that structure to show dependencies among fix effectiveness factors whereby we ultimately quantify the effect of FEF dependencies on MTBF uncertainty. Previous modeling does not include FEF variability in SME prediction of fix effectiveness, the variability caused by dependencies between FEF nor can any of the models predict fix effectiveness during early phases of product development prior to test results. Given these model shortcomings, we recommend altering reliability grown projection models, such as equation 1 to include FEF projection and dependence per equation 25.

$$\rho(T) = \lambda_A + \sum_{i=1}^{m} \left(1 - D_{BBN_{d_i}}\right)\lambda_i + D_{BBN_{\bar{d}}}\, h(t) \tag{25}$$

The term $D_{BBN_{di}}$ represents fix effectiveness factors obtained from the composite BBN model of Figure-24. Common nodes link activities from each team that are attempting to solve or fix their respective failure mode. The term $D_{BBN_{\bar{d}_i}}$ represents the average FEF projection obtained from the composite BBN for failure modes receiving corrective action. This methodology provides for repeatable fix effectiveness projection during early stages of product development. This research has experienced MTBF error from minus16.5% to plus 13.7% as a result of dependency between fix effectiveness factors. Inclusion of FEF variability and FEF dependence provide a more accurate projection of MTBF than projections made with the assumption of FEF independence.

Chapter 7: Conclusions

Review of existing methods to predict fix effectiveness reveal shortcomings in the method. These shortcomings may result in mis-allocation of resources, over estimation of project success etc. SME FEF projection error has a significant impact on reliability projection metrics; however examination of performance judgment error has been noted as non-conservative. This led to an exploration of a structure that would provide more accurate FEF projections and reduce the associated error. Therefore, the significance of this research is two-fold. First no structured method of projecting fix effectiveness factors during planning exists. Second, no reliability growth projection model accounts for dependencies among fix effectiveness factors. This author has defined the process to account for both of these short comings and provided test cases for confirmation the process works. Both of these concepts, structured FEF projection and FEF dependencies have been overlooked since the beginning of reliability growth projection modeling.

This research has provided an insight into the methodology of collecting fix effectiveness performance shaping factors and organizing them in such a way as to provide FEF projections with less error than those of SME. In addition this research has provided two test cases. The first occurred within a major HVAC OEM, with their SME developing the PSF, building model structures, and ultimately projecting FEF with an average of 64% less error than the very SME that developed the models and associated CPT. The second test case involved a major automotive supplier. The models and CPT developed within the HVAC firm were used to project FEF for the automotive organization and again, BBN FEF projections demonstrated a 53% reduction in FEF projection error than automotive SME, proving that generic PSF and BBN structures can provide more accurate FEF estimates across diverse industries. Research discoveries include identification of factors (PSF) affecting the values of FEF, development of paths and magnitude of PSF in the form of a BBN, testing, and empirical confirmation of the model.

BBN model error was further explored in two ways, the first involved development of a simulated SME (SSME) model. Time to develop the original BBN models and populate CPT was approximately 5 months. SME FEF output resembled a pessimist, normalist and an optimist, thus it became apparent we could have simulated SME judgment; possibly saving many months of research and eliminated group think. Rules were established for the span of FEF judgment for each category along with methods to populate CPT based on the number of PSF node parents. The S-SME model proved to be statistical significant in predicting FEF for HVAC but experienced difficulty in prediction of fix effectiveness for the automotive industry, which led to the next error evaluation.

The second method of addressing BBN model error was the advent of a composite BBN model. SME concluded that management commitment, test facilities and resource availability are common to multiple engineering teams attempting to fix a failure mode. Therefore two M-2 BBN were joined at the common nodes and CPT redeveloped. A full factorial DOE for non-common nodes, revealed that fix effectiveness factors are in fact dependent on one another. Treatment combinations for one team attempting to fix a failure mode were varied and the impact to a second team noted. If independent the second teams FEF would not vary, however the second team's FEF varied -11.7% to + 22.2%. FEF dependency is indeed a significant find within this research, with only 5 failure modes and conventional reliability growth projection methods, MTBF projections are misstated by minus 16.5% to plus 13.7%.

Contributions for this research include:

1. This research has evaluated the effects of FEF variation on the widely used reliability growth models, Crow-AMSAA and an emerging growth model, AMPM-Stein. Results show that test of equal variances indicated the null hypothesis of no difference in variances cannot be statistically rejected at the 95% significance level. Two sample T-tests however, indicated one must reject the null of no difference in mean error. In every instance, when a difference in mean error was detected, the AMPM Stein mean error was lower. Statistically, the AMPM-Stein model is the more robust model against the effects of FEF variability. FEF variance of 0.05 leads to 20% reliability growth projection error, thus model to model variability is secondary to the impact of FEF projection.
2. Experts from 15 diverse industries with 270 years of combined experience, worked to develop fix effectiveness performance shaping factors. The diversity of the industries used in this research provided a generalized list of FEF PSF applicable to most engineering communities attempting to solve a failure mode.
3. Within this research three methods have been presented to define relationships between PSF, use SME judgment to build BBN, and project fix effectiveness during project planning (Figure-21). The first method (M-1), expert aggregate,

allows SME to organize PSF, define parent-child relationship, and associated CPT for their respective model. The second method, (M-2), is a fixed structure method whereby SME reach consensus on model structure and parent-child relationships with each SME defining CPT for the agreed upon structure. The third method, (M-3) was a consensus model whereby SME reach consensus on CPT for the M-2 structure. The Stiber et al. (2004) algorithm was used to develop node likelihood functions whereby Bayesian methods update SME judgment weights leading to a weighted posterior aggregate SME judgment.

4. Lastly, we have developed a composite BBN structure showing the dependencies among FEF and their impact on MTBF projections. This led the proposal of a new failure intensity function that includes FEF dependency.

Therefore we have augmented the Crow-AMSAA model as shown below.

$$\rho(T) = \lambda_A + \sum_{i=1}^{m} \left(1 - D_{BBN_{d_i}}\right)\lambda_i + D_{BBN_{\bar{d}}}\, h(t)$$

Although an augmentation to Crow-AMSAA reliability growth projection model has been shown, the process for FEF projection defined within this research can be used for any reliability model requiring an assessment of fix effectiveness. Fix effectiveness dependency must be addressed in all reliability growth models.

Whether the FEF projection is generated from an M-1 expert aggregate, M-2 fixed structure, or M-3 consensus model, this research has defined and demonstrated the process necessary to collect SME judgment, build a BBN with associated CPT, and project FEF during product development reliability planning. Current reliability planning models are inept in this endeavor. In addition, we have developed a dependency composite BBN model whereby more accurate FEF projections can be made by the inclusion of FEF dependency among fix activities.

Future Research

Likelihood Based on both Observed and Unobserved PSF

Future research would include expanding the Stiber et al. (2004) likelihood function to provide judgment weights based on observed and unobserved or missing data. Currently the Stiber algorithm estimates the likelihood that a particular SME model is correct by measuring how close the PSF marginal probabilities are to PSF evidence; missing evidence nodes are not considered. In numerous cases within this research, it was noted the weight of a particular SME would calculate low, but their FEF might be closer to reality than other SME. Possibly a likelihood adjusting factor due to the aggressive nature of a particular SME CPT can be factored into the evidence, etc.

Breaking FEF Dependency

A second research area involves development of methodology to break FEF dependencies. We have developed a methodology proving dependence among PSF, but have not developed methodology to break dependences for all structures. Breaking dependencies will prove crucial research that will reduce the level of complexity of implementation of FEF projection via BBN methodology previously described.

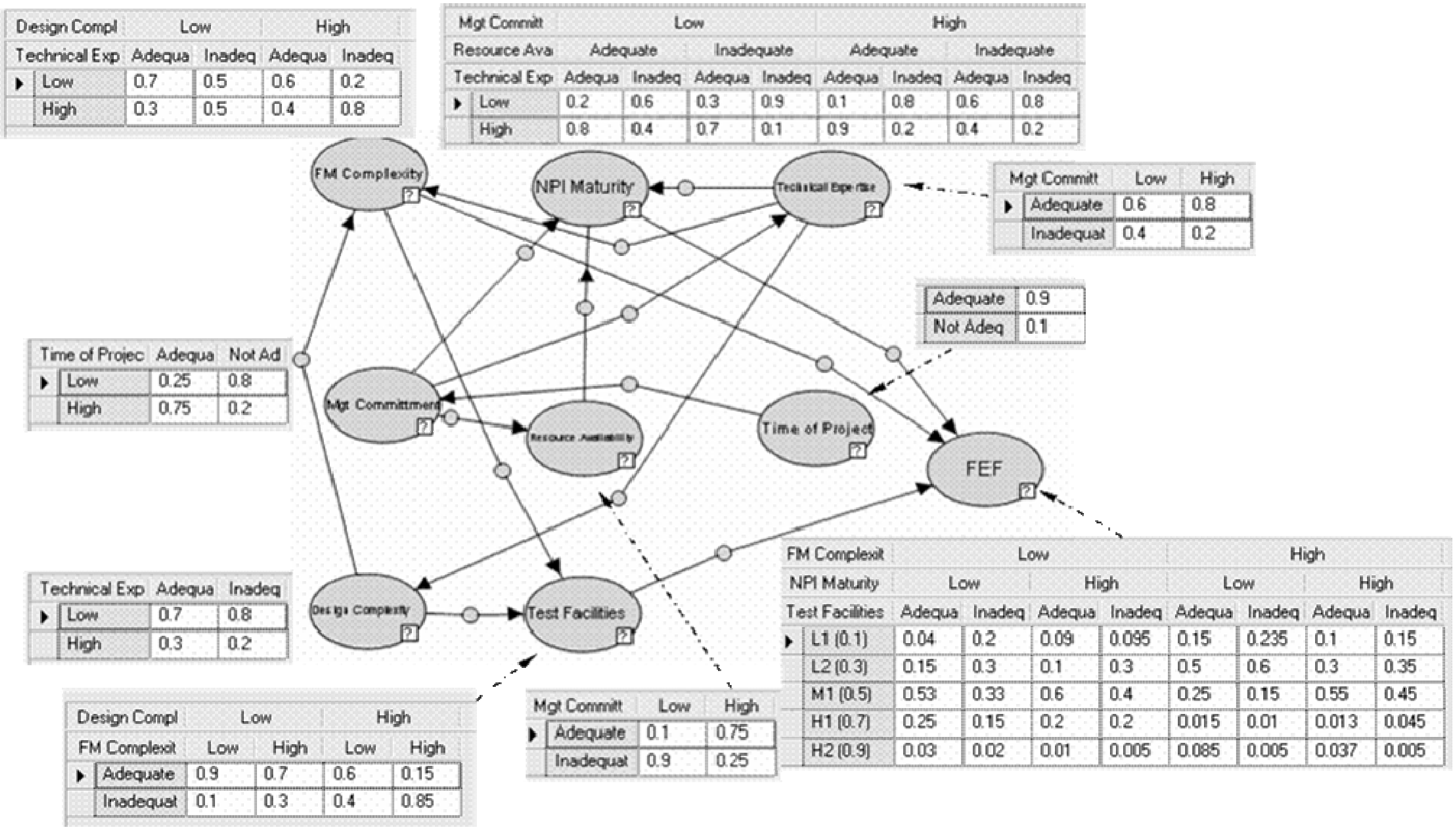
Optimizing BBN Structures

A third research area involves optimization of the BBN structures for minimization of FEF projection error. Numerous optimization algorithms exists today whereby an organization’s field data and PSF can be used to develop a BBN structure whereby FEF projection can be minimized for that organization.

Appendices

Appendix A

M-1 BBN



Design Compl	Low		High	
Technical Exp	Adequa	Inadeq	Adequa	Inadeq
Low	0.7	0.5	0.6	0.2
High	0.3	0.5	0.4	0.8

Mgt Committ	Low				High			
Resource Ava	Adequate		Inadequate		Adequate		Inadequate	
Technical Exp	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Low	0.2	0.6	0.3	0.9	0.1	0.8	0.6	0.8
High	0.8	0.4	0.7	0.1	0.9	0.2	0.4	0.2

Mgt Committ	Low	High
Adequate	0.6	0.8
Inadequat	0.4	0.2

Adequate	0.9
Not Adeq	0.1

Time of Projec	Adequa	Not Ad
Low	0.25	0.8
High	0.75	0.2

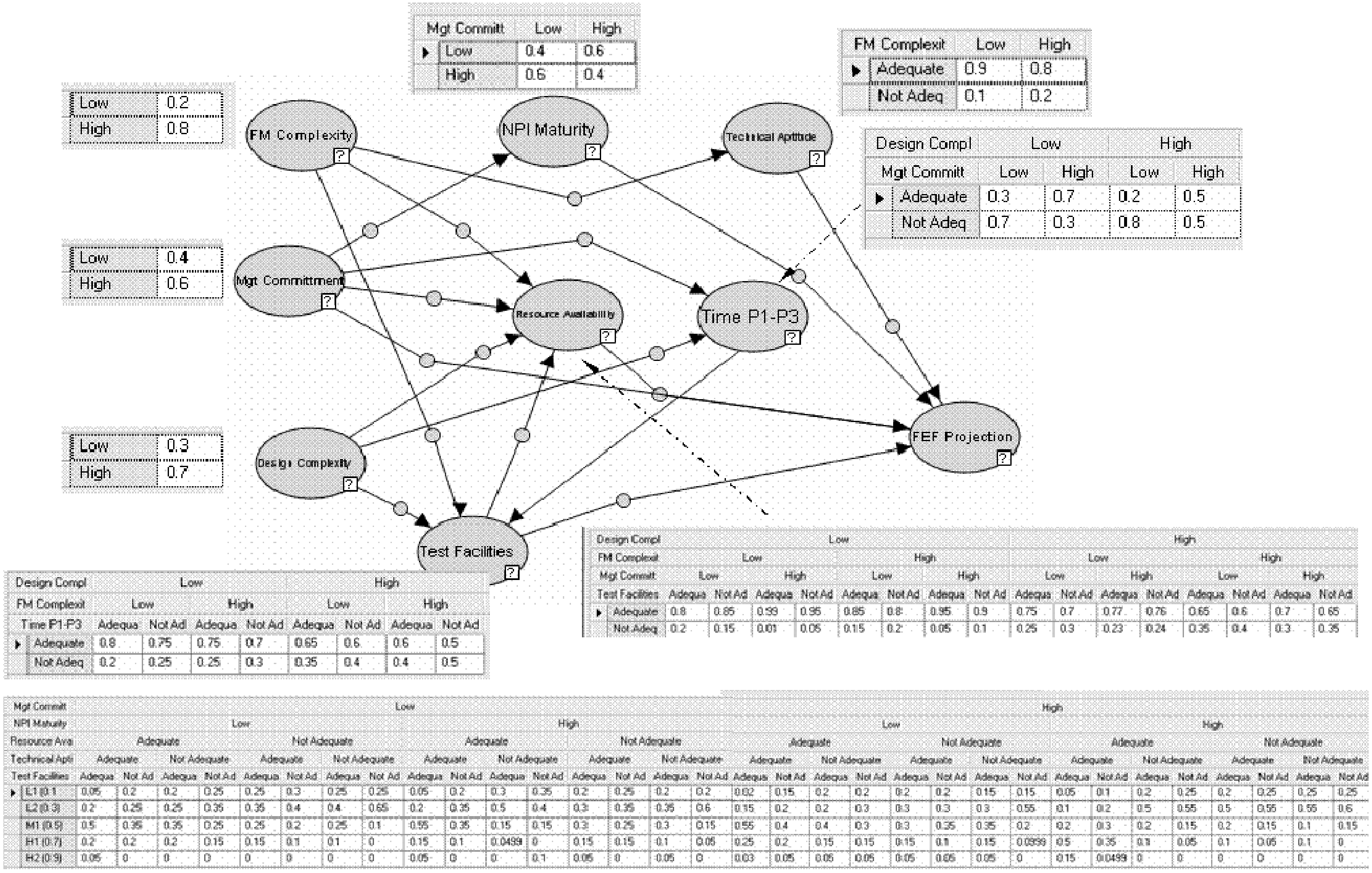
Technical Exp	Adequa	Inadeq
Low	0.7	0.8
High	0.3	0.2

Design Compl	Low		High	
FM Complexit	Low	High	Low	High
Adequate	0.9	0.7	0.6	0.15
Inadequat	0.1	0.3	0.4	0.85

Mgt Committ	Low	High
Adequate	0.1	0.75
Inadequat	0.9	0.25

FM Complexit	Low				High			
NPI Maturity	Low		High		Low		High	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.04	0.2	0.09	0.095	0.15	0.235	0.1	0.15
L2 (0.3)	0.15	0.3	0.1	0.3	0.5	0.6	0.3	0.35
M1 (0.5)	0.53	0.33	0.6	0.4	0.25	0.15	0.55	0.45
H1 (0.7)	0.25	0.15	0.2	0.2	0.015	0.01	0.013	0.045
H2 (0.9)	0.03	0.02	0.01	0.005	0.085	0.005	0.037	0.005

Figure A1. M-1 BBN – Expert-3



Mgt Committ	Low	High
Low	0.4	0.6
High	0.6	0.4

FM Complexit	Low	High
Adequate	0.9	0.8
Not Adeq	0.1	0.2

Low	0.2
High	0.8

Design Compl	Low		High	
Mgt Committ	Low	High	Low	High
Adequate	0.3	0.7	0.2	0.5
Not Adeq	0.7	0.3	0.8	0.5

Low	0.4
High	0.6

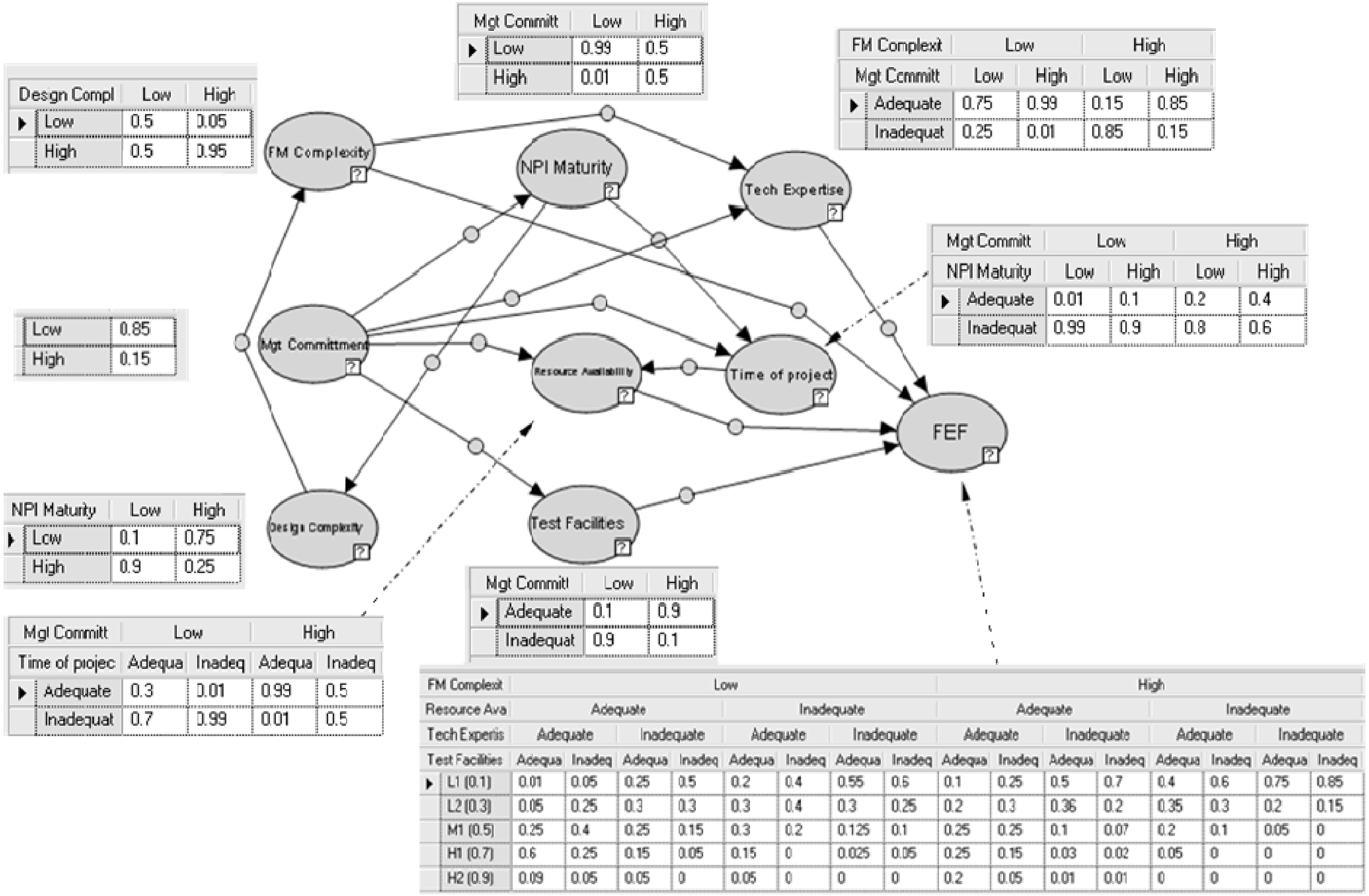
Low	0.3
High	0.7

Design Compl	Low								High							
FM Complexit	Low				High				Low				High			
Mgt Committ	Low		High		Low		High		Low		High		Low		High	
Test Facilities	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad
Adequate	0.8	0.85	0.99	0.95	0.85	0.8	0.95	0.9	0.75	0.7	0.77	0.76	0.65	0.6	0.7	0.65
Not Adeq	0.2	0.15	0.01	0.05	0.15	0.2	0.05	0.1	0.25	0.3	0.23	0.24	0.35	0.4	0.3	0.35

Design Compl	Low				High			
FM Complexit	Low		High		Low		High	
Time P1-P3	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad
Adequate	0.8	0.75	0.75	0.7	0.65	0.6	0.6	0.5
Not Adeq	0.2	0.25	0.25	0.3	0.35	0.4	0.4	0.5

Mgt Committ	Low																High															
NPI Maturity	Low								High								Low								High							
Resource Avai	Adequate				Not Adequate				Adequate				Not Adequate				Adequate				Not Adequate				Adequate				Not Adequate			
Technical Apti	Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate		Adequate		Not Adequate	
Test Facilities	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad	Adequa	Not Ad
L1 (0.1	0.05	0.2	0.2	0.25	0.25	0.3	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.2	0.2	0.02	0.15	0.2	0.2	0.2	0.2	0.15	0.15	0.05	0.1	0.2	0.25	0.2	0.25	0.25	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.4	0.65	0.2	0.35	0.5	0.4	0.3	0.35	0.35	0.6	0.15	0.2	0.2	0.3	0.3	0.3	0.3	0.55	0.1	0.2	0.5	0.55	0.5	0.55	0.55	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.2	0.25	0.1	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.55	0.4	0.4	0.3	0.3	0.35	0.35	0.2	0.2	0.3	0.2	0.15	0.2	0.15	0.1	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.1	0	0.15	0.1	0.0499	0	0.15	0.15	0.1	0.05	0.25	0.2	0.15	0.15	0.15	0.1	0.15	0.0999	0.5	0.35	0.1	0.05	0.1	0.05	0.1	0
H2 (0.9)	0.05	0	0	0	0	0	0	0	0.05	0	0	0.1	0.05	0	0.05	0	0.03	0.05	0.05	0.05	0.05	0.05	0.05	0	0.15	0.0499	0	0	0	0	0	0

Figure A2. M-1 BBN – Expert 5



Desgn Compl	Low	High
Low	0.5	0.05
High	0.5	0.95

Mgt Committ	Low	High
Low	0.99	0.5
High	0.01	0.5

FM Complexit	Low		High	
Mgt Committ	Low	High	Low	High
Adequate	0.75	0.99	0.15	0.85
Inadequat	0.25	0.01	0.85	0.15

Low	0.85
High	0.15

Mgt Committ	Low		High	
NPI Maturity	Low	High	Low	High
Adequate	0.01	0.1	0.2	0.4
Inadequat	0.99	0.9	0.8	0.6

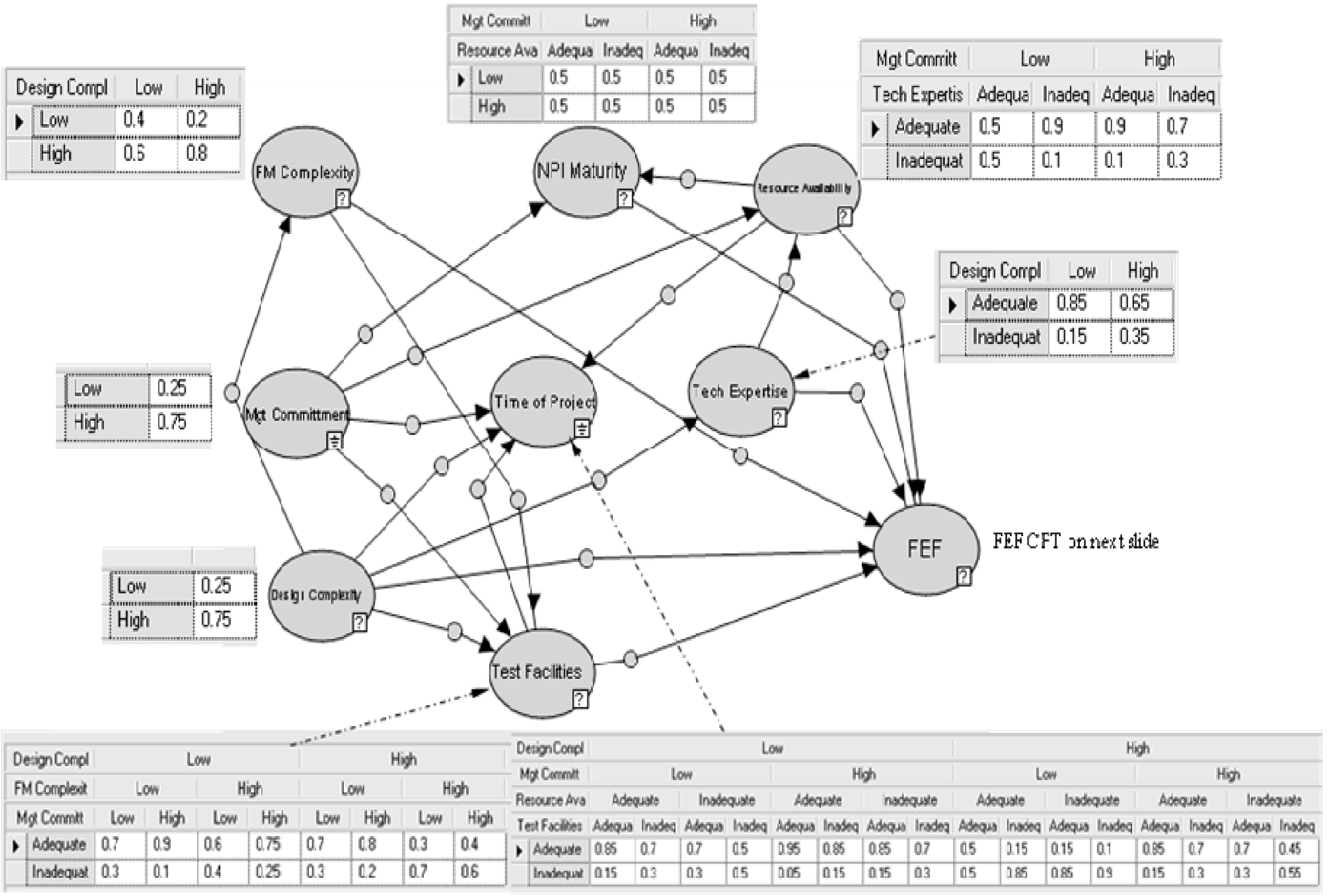
NPI Maturity	Low	High
Low	0.1	0.75
High	0.9	0.25

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.3	0.01	0.99	0.5
Inadequat	0.7	0.99	0.01	0.5

Mgt Committ	Low	High
Adequate	0.1	0.9
Inadequat	0.9	0.1

FM Complexit	Low								High							
Resource Ava	Adequate				Inadequate				Adequate				Inadequate			
Tech Expertis	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.01	0.05	0.25	0.5	0.2	0.4	0.55	0.6	0.1	0.25	0.5	0.7	0.4	0.6	0.75	0.85
L2 (0.3)	0.05	0.25	0.3	0.3	0.3	0.4	0.3	0.25	0.2	0.3	0.36	0.2	0.35	0.3	0.2	0.15
M1 (0.5)	0.25	0.4	0.25	0.15	0.3	0.2	0.125	0.1	0.25	0.25	0.1	0.07	0.2	0.1	0.05	0
H1 (0.7)	0.6	0.25	0.15	0.05	0.15	0	0.025	0.05	0.25	0.15	0.03	0.02	0.05	0	0	0
H2 (0.9)	0.09	0.05	0.05	0	0.05	0	0	0	0.2	0.05	0.01	0.01	0	0	0	0

Figure A3. M-1 BBN – Expert 7



Design Compl	Low	High
Low	0.4	0.2
High	0.6	0.8

Mgt Commit	Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq
Low	0.5	0.5	0.5	0.5
High	0.5	0.5	0.5	0.5

Mgt Committ	Low		High	
Tech Expertis	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.5	0.9	0.9	0.7
Inadequat	0.5	0.1	0.1	0.3

Design Compl	Low	High
Adequate	0.85	0.65
Inadequat	0.15	0.35

Low	0.25
High	0.75

Low	0.25
High	0.75

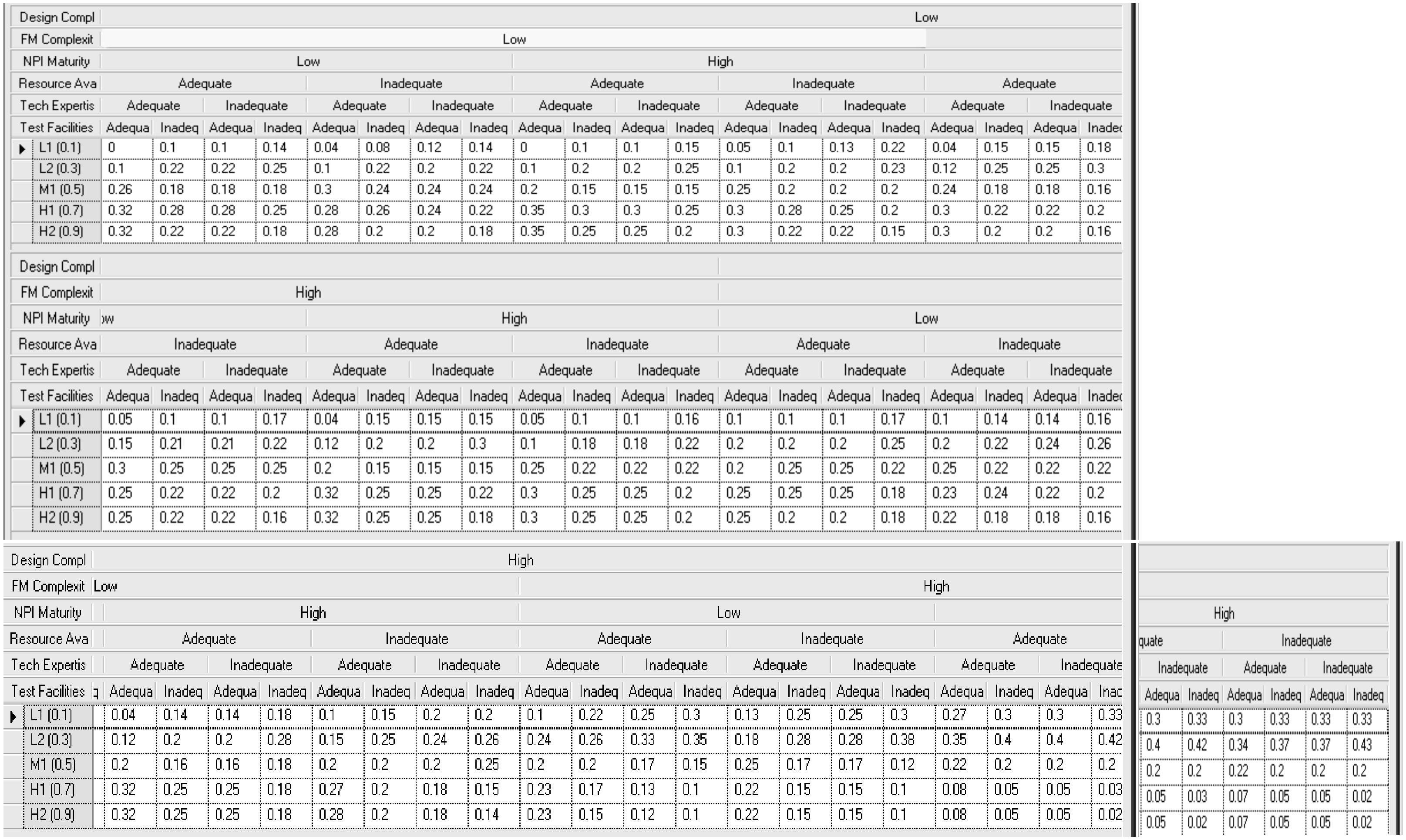
FEF CPT on next slide

Design Compl	Low				High			
FM Complexit	Low		High		Low		High	
Mgt Committ	Low	High	Low	High	Low	High	Low	High
Adequate	0.7	0.9	0.6	0.75	0.7	0.8	0.3	0.4
Inadequat	0.3	0.1	0.4	0.25	0.3	0.2	0.7	0.6

Design Compl	Low								High							
Mgt Commit	Low				High				Low				High			
Resource Ava	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.85	0.7	0.7	0.5	0.95	0.85	0.85	0.7	0.5	0.15	0.15	0.1	0.85	0.7	0.7	0.45
Inadequat	0.15	0.3	0.3	0.5	0.05	0.15	0.15	0.3	0.5	0.85	0.85	0.9	0.15	0.3	0.3	0.55

Figure A4. M-1 BBN – Expert 8

Figure A4. (*continued.*)



Design Compl	Low																			
FM Complexit	Low																			
NPI Maturity	Low								High											
Resource Ava	Adequate				Inadequate				Adequate				Inadequate				Adequate			
Tech Expertis	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadec
L1 (0.1)	0	0.1	0.1	0.14	0.04	0.08	0.12	0.14	0	0.1	0.1	0.15	0.05	0.1	0.13	0.22	0.04	0.15	0.15	0.18
L2 (0.3)	0.1	0.22	0.22	0.25	0.1	0.22	0.2	0.22	0.1	0.2	0.2	0.25	0.1	0.2	0.2	0.23	0.12	0.25	0.25	0.3
M1 (0.5)	0.26	0.18	0.18	0.18	0.3	0.24	0.24	0.24	0.2	0.15	0.15	0.15	0.25	0.2	0.2	0.2	0.24	0.18	0.18	0.16
H1 (0.7)	0.32	0.28	0.28	0.25	0.28	0.26	0.24	0.22	0.35	0.3	0.3	0.25	0.3	0.28	0.25	0.2	0.3	0.22	0.22	0.2
H2 (0.9)	0.32	0.22	0.22	0.18	0.28	0.2	0.2	0.18	0.35	0.25	0.25	0.2	0.3	0.22	0.22	0.15	0.3	0.2	0.2	0.16

Design Compl																				
FM Complexit	High																			
NPI Maturity	ow				High								Low							
Resource Ava	Inadequate				Adequate				Inadequate				Adequate				Inadequate			
Tech Expertis	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadec
L1 (0.1)	0.05	0.1	0.1	0.17	0.04	0.15	0.15	0.15	0.05	0.1	0.1	0.16	0.1	0.1	0.1	0.17	0.1	0.14	0.14	0.16
L2 (0.3)	0.15	0.21	0.21	0.22	0.12	0.2	0.2	0.3	0.1	0.18	0.18	0.22	0.2	0.2	0.2	0.25	0.2	0.22	0.24	0.26
M1 (0.5)	0.3	0.25	0.25	0.25	0.2	0.15	0.15	0.15	0.25	0.22	0.22	0.22	0.2	0.25	0.25	0.22	0.25	0.22	0.22	0.22
H1 (0.7)	0.25	0.22	0.22	0.2	0.32	0.25	0.25	0.22	0.3	0.25	0.25	0.2	0.25	0.25	0.25	0.18	0.23	0.24	0.22	0.2
H2 (0.9)	0.25	0.22	0.22	0.16	0.32	0.25	0.25	0.18	0.3	0.25	0.25	0.2	0.25	0.2	0.2	0.18	0.22	0.18	0.18	0.16

Design Compl	High																				
FM Complexit	Low									High											
NPI Maturity		High								Low											
Resource Ava		Adequate				Inadequate				Adequate				Inadequate				Adequate			
Tech Expertis		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Test Facilities	q	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inac
L1 (0.1)		0.04	0.14	0.14	0.18	0.1	0.15	0.2	0.2	0.1	0.22	0.25	0.3	0.13	0.25	0.25	0.3	0.27	0.3	0.3	0.33
L2 (0.3)		0.12	0.2	0.2	0.28	0.15	0.25	0.24	0.26	0.24	0.26	0.33	0.35	0.18	0.28	0.28	0.38	0.35	0.4	0.4	0.42
M1 (0.5)		0.2	0.16	0.16	0.18	0.2	0.2	0.2	0.25	0.2	0.2	0.17	0.15	0.25	0.17	0.17	0.12	0.22	0.2	0.2	0.2
H1 (0.7)		0.32	0.25	0.25	0.18	0.27	0.2	0.18	0.15	0.23	0.17	0.13	0.1	0.22	0.15	0.15	0.1	0.08	0.05	0.05	0.03
H2 (0.9)		0.32	0.25	0.25	0.18	0.28	0.2	0.18	0.14	0.23	0.15	0.12	0.1	0.22	0.15	0.15	0.1	0.08	0.05	0.05	0.02

	High					
	quate		Inadequate			
	Inadequate		Adequate		Inadequate	
	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
	0.3	0.33	0.3	0.33	0.33	0.33
	0.4	0.42	0.34	0.37	0.37	0.43
	0.2	0.2	0.22	0.2	0.2	0.2
	0.05	0.03	0.07	0.05	0.05	0.02
	0.05	0.02	0.07	0.05	0.05	0.02

Appendix B

HVAC Fix Effectiveness (*di*) Project 1-5: M-1 SME Models

Table B1

HVAC Fix Effectiveness (di) Projection Project 1 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	0	0	0	0	0.000
Quality system maturity	0.604	0.480	0.084	0.500	0	0	0	0	0.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	0.769	0.240	0.344	0.407	0.472
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	0.824	0.499	0.217	0.555	0.627
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.065	0.050	0.069	0.057	0.057
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.231	0.200	0.148	0.162	0.200
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.465	0.500	0.250	0.227	0.412
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.196	0.200	0.370	0.277	0.220
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.043	0.050	0.162	0.277	0.111
Likelihood	-	-	-		3.13E-02	3.30E-02	3.12E-04	2.47E-02	
Prob (M_j)	0.25	0.25	0.25	0.25	0.35	0.37	0.00	0.28	
	Mean	Var.	Alpha	Beta					
	0.525	0.043	2.517	2.275					

Table B2

HVAC Fix Effectiveness (di) Projection Project 2 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	0.971	0.545	0.569	0.850	0.737
Failure mode complexity	0.610	0.200	0.119	0.250	0	0	0	0	0.000
Technical expertise	0.739	0.820	0.310	0.700	0.931	0.800	0.850	0.705	0.766
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	0	0	0	0	0.000
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.100	0.080	0.160	0.279	0.180
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.300	0.180	0.224	0.365	0.283
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.550	0.200	0.228	0.214	0.241
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.013	0.420	0.217	0.071	0.208
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.037	0.120	0.172	0.071	0.088
Likelihood	-	-	-		1.75E-02	7.30E-02	2.22E-04	8.83E-02	
Prob (M_j)	0.25	0.25	0.25	0.25	0.10	0.41	0.00	0.49	
	Mean	Var.	Alpha	Beta					
	0.448	0.060	1.405	1.730					

Table B3

HVAC Fix Effectiveness (di) Projection Project 3 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	0	0	0	0	0.000
Quality system maturity	0.604	0.480	0.084	0.500	0	0	0	0	0.000
Project time	0.900	0.428	0.054	0.674	0	0	0	0	0.000
Failure mode complexity	0.610	0.200	0.119	0.250	0.299	0.272	0.500	0.333	0.459
Technical expertise	0.739	0.820	0.310	0.700	0.259	0.817	0.450	0.876	0.511
Resource availability	0.552	0.743	0.108	0.785	0.069	0.809	0.010	0.433	0.131
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	0	0	0	0	0.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.225	0.225	0.629	0.115	0.546
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.510	0.302	0.267	0.230	0.276
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.204	0.303	0.096	0.216	0.130
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.052	0.170	0.007	0.233	0.039
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.009	0.000	0.000	0.205	0.009
Likelihood	-	-	-		2.56E-03	1.40E-02	8.87E-02	4.76E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.02	0.13	0.81	0.04	
	Mean	Var.	Alpha	Beta					
	0.238	0.033	1.079	3.456					

Table B4

HVAC Fix Effectiveness (di) Projection Project 4 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	0	0	0	0	0.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	0.622	0.217	0.500	0.444	0.517
Technical expertise	0.739	0.820	0.310	0.700	0.315	0.822	0.920	0.879	0.525
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.082	0.052	0.086	0.035	0.066
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.282	0.159	0.138	0.126	0.227
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.424	0.523	0.239	0.241	0.397
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.161	0.232	0.406	0.301	0.205
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.051	0.034	0.131	0.297	0.105
Likelihood	-	-	-		7.01E-02	1.81E-02	2.74E-05	2.64E-02	
Prob (M_j)	0.25	0.25	0.25	0.25	0.61	0.16	0.00	0.23	
	Mean	Var.	Alpha	Beta					
	0.511	0.045	2.358	2.256					

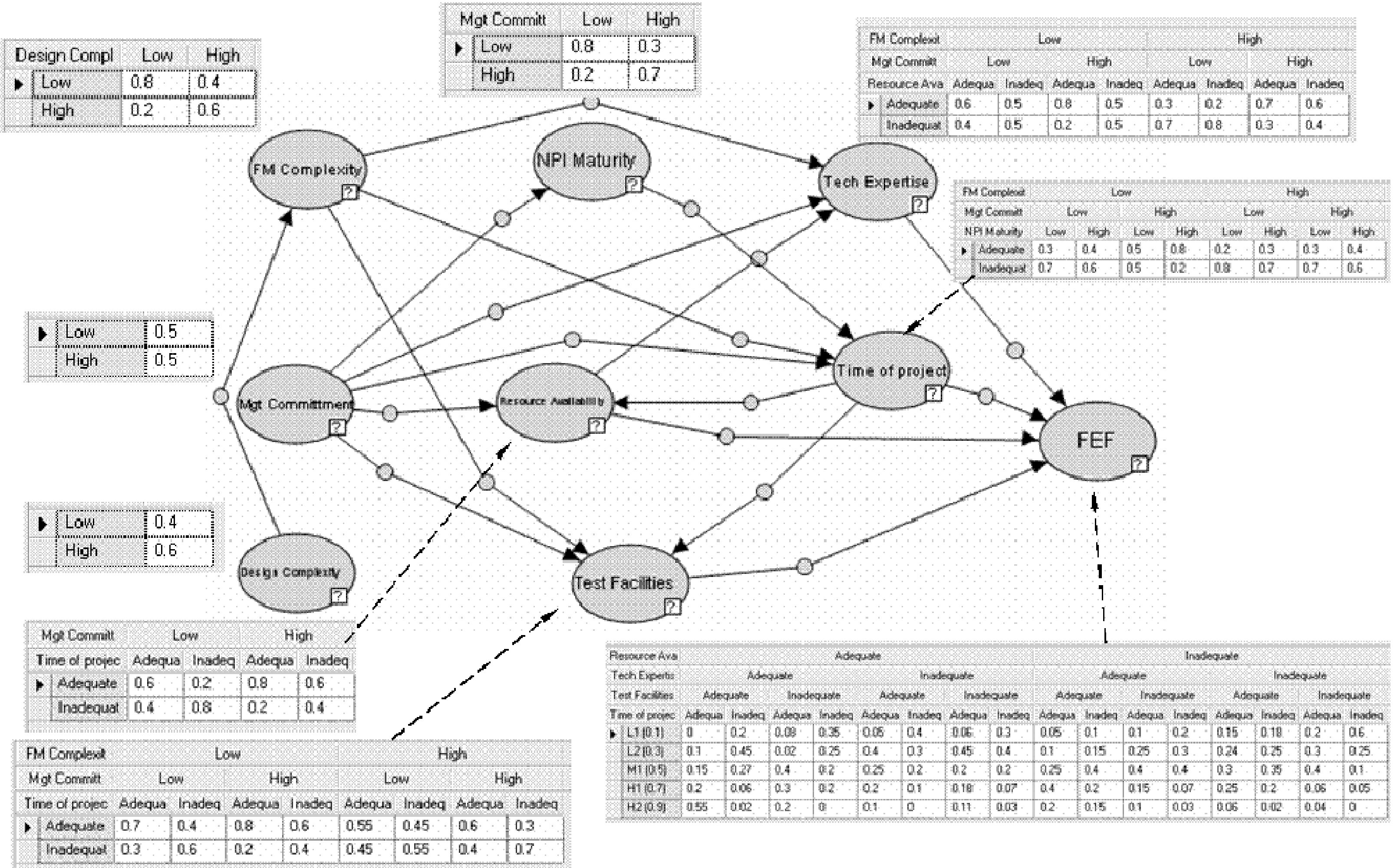
Table B5

HVAC Fix Effectiveness (di) Projection Project 5 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	0	0	0	0	0.000
Failure mode complexity	0.610	0.200	0.119	0.250	1	1	1	1	1.000
Technical expertise	0.739	0.820	0.310	0.700	0.907	0.900	0.990	0.654	0.864
Resource availability	0.552	0.743	0.108	0.785	0	0	0	0	0.000
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	0	0	0	0	0.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.095	0.250	0.402	0.142	0.149
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.300	0.555	0.399	0.210	0.346
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.400	0.150	0.199	0.200	0.301
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.200	0.045	0.001	0.252	0.167
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.005	0.000	0.000	0.196	0.037
Likelihood	-	-	-		2.44E-03	9.93E-04	1.52E-04	7.67E-04	
Prob (M_j)	0.25	0.25	0.25	0.25	0.56	0.23	0.03	0.18	
	Mean	Var.	Alpha	Beta					
	0.419	0.044	1.909	2.642					

Appendix C

M-2 Fixed Structure BBN



Design Compl	Low	High
Low	0.8	0.4
High	0.2	0.6

Mgt Commit	Low	High
Low	0.8	0.3
High	0.2	0.7

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.5	0.8	0.5	0.3	0.2	0.7	0.6
Inadequat	0.4	0.5	0.2	0.5	0.7	0.8	0.3	0.4

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.3	0.4	0.5	0.8	0.2	0.3	0.3	0.4
Inadequat	0.7	0.6	0.5	0.2	0.8	0.7	0.7	0.6

Low	0.5
High	0.5

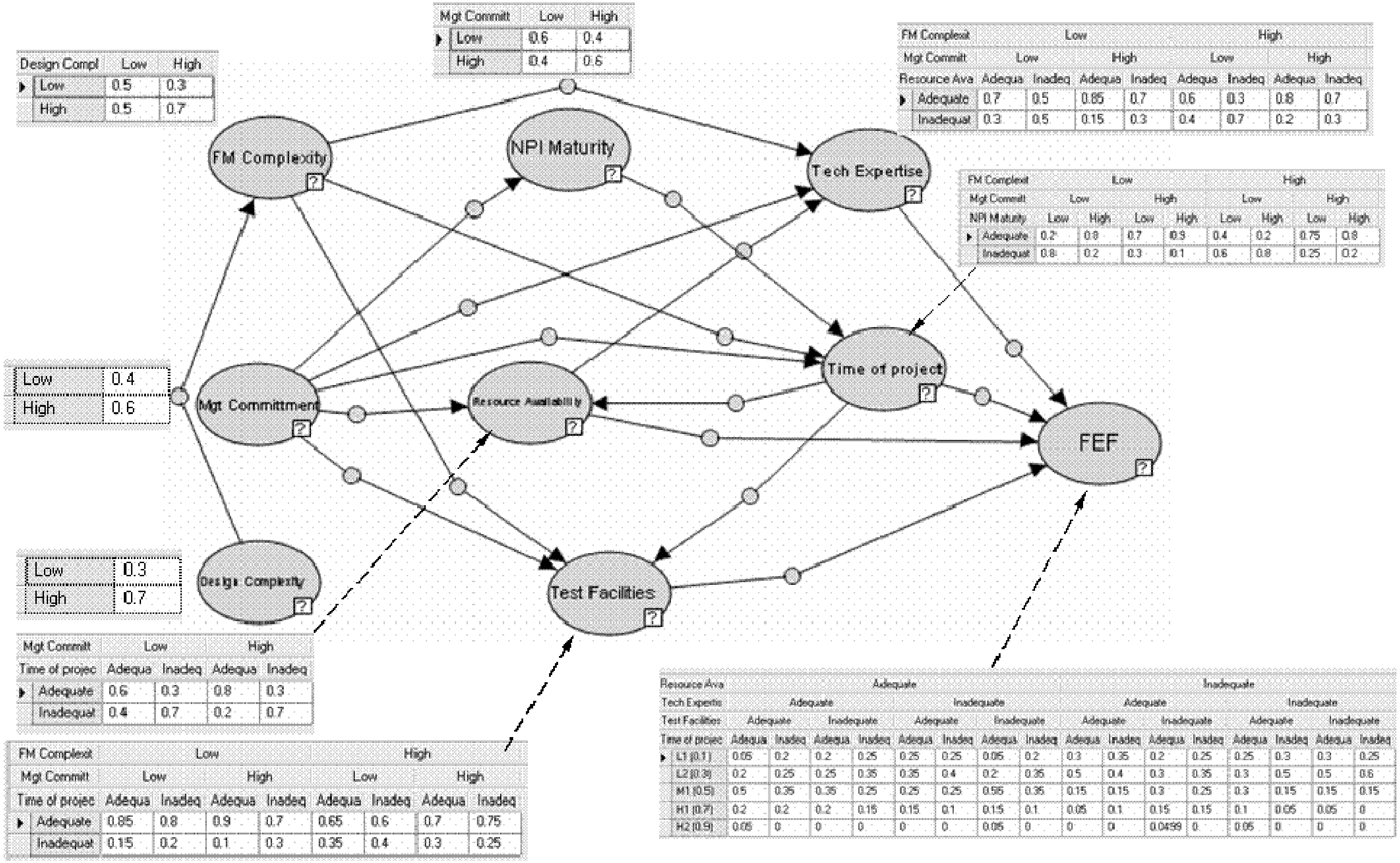
Low	0.4
High	0.6

Mgt Commit	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.2	0.8	0.6
Inadequat	0.4	0.8	0.2	0.4

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.4	0.8	0.6	0.55	0.45	0.6	0.3
Inadequat	0.3	0.6	0.2	0.4	0.45	0.55	0.4	0.7

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0	0.2	0.08	0.35	0.05	0.4	0.06	0.3	0.05	0.1	0.1	0.2	0.15	0.18	0.2	0.6
L2 (0.3)	0.1	0.45	0.02	0.25	0.4	0.3	0.45	0.4	0.1	0.15	0.25	0.3	0.24	0.25	0.3	0.25
M1 (0.5)	0.15	0.27	0.4	0.2	0.25	0.2	0.2	0.2	0.25	0.4	0.4	0.4	0.3	0.35	0.4	0.1
H1 (0.7)	0.2	0.06	0.3	0.2	0.2	0.1	0.18	0.07	0.4	0.2	0.15	0.07	0.25	0.2	0.06	0.05
H2 (0.9)	0.55	0.02	0.2	0	0.1	0	0.11	0.03	0.2	0.15	0.1	0.03	0.06	0.02	0.04	0

Figure C1. M-2 fixed structure BBN – Expert 3



Design Compl	Low	High
Low	0.5	0.3
High	0.5	0.7

Mgt Commit	Low	High
Low	0.6	0.4
High	0.4	0.6

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.5	0.85	0.7	0.6	0.3	0.8	0.7
Inadequat	0.3	0.5	0.15	0.3	0.4	0.7	0.2	0.3

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.2	0.8	0.7	0.9	0.4	0.2	0.75	0.8
Inadequat	0.8	0.2	0.3	0.1	0.6	0.8	0.25	0.2

Low	0.4
High	0.6

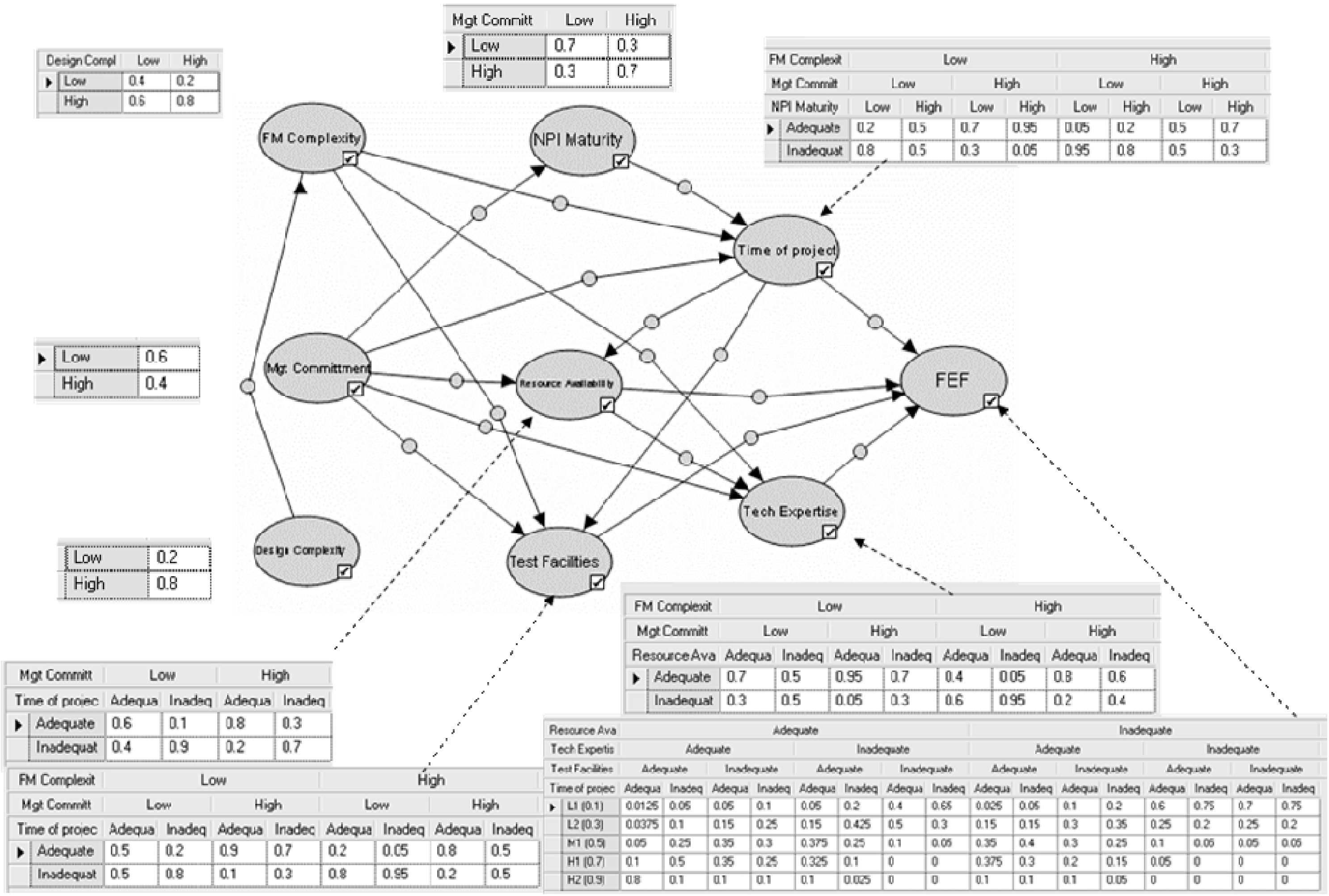
Low	0.3
High	0.7

Mgt Commit	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.3	0.8	0.3
Inadequat	0.4	0.7	0.2	0.7

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.85	0.8	0.9	0.7	0.65	0.6	0.7	0.75
Inadequat	0.15	0.2	0.1	0.3	0.35	0.4	0.3	0.25

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.05	0.2	0.2	0.25	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.25	0.3	0.3	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.2	0.35	0.5	0.4	0.3	0.35	0.3	0.5	0.5	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.25	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.15	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.15	0.1	0.05	0.1	0.15	0.15	0.1	0.05	0.05	0
H2 (0.9)	0.05	0	0	0	0	0	0.05	0	0	0	0.0499	0	0.05	0	0	0

Figure C2. M-2 fixed structure BNN – Expert 5



Design Compl	Low	High
Low	0.4	0.2
High	0.6	0.8

Mgt Committ	Low	High
Low	0.7	0.3
High	0.3	0.7

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.2	0.5	0.7	0.95	0.05	0.2	0.5	0.7
Inadequat	0.8	0.5	0.3	0.05	0.95	0.8	0.5	0.3

Low	0.6
High	0.4

Low	0.2
High	0.8

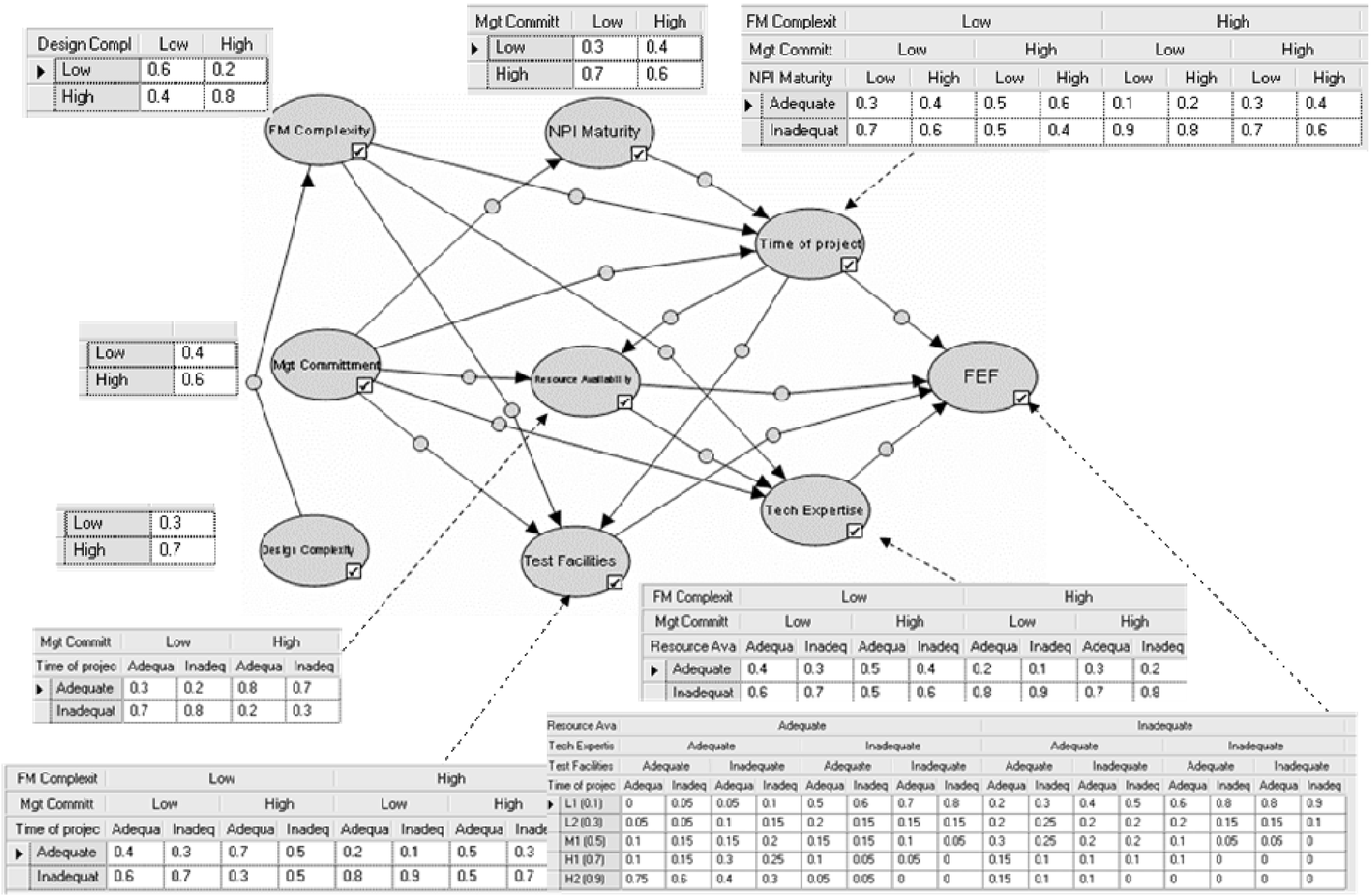
FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.5	0.95	0.7	0.4	0.05	0.8	0.6
Inadequat	0.3	0.5	0.05	0.3	0.6	0.95	0.2	0.4

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.1	0.8	0.3
Inadequat	0.4	0.9	0.2	0.7

FM Complexit	Low				High			
Mgt Committ	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.5	0.2	0.9	0.7	0.2	0.05	0.8	0.5
Inadequat	0.5	0.8	0.1	0.3	0.8	0.95	0.2	0.5

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.0125	0.05	0.05	0.1	0.05	0.2	0.4	0.65	0.025	0.05	0.1	0.2	0.6	0.75	0.7	0.75
L2 (0.3)	0.0375	0.1	0.15	0.25	0.15	0.425	0.5	0.3	0.15	0.15	0.3	0.35	0.25	0.2	0.25	0.2
M1 (0.5)	0.05	0.25	0.35	0.3	0.375	0.25	0.1	0.05	0.35	0.4	0.3	0.25	0.1	0.05	0.05	0.05
H1 (0.7)	0.1	0.5	0.35	0.25	0.325	0.1	0	0	0.375	0.3	0.2	0.15	0.05	0	0	0
H2 (0.9)	0.8	0.1	0.1	0.1	0.1	0.025	0	0	0.1	0.1	0.1	0.05	0	0	0	0

Figure C3. M-2 fixed structure BBN – Expert 7



Design Compl	Low	High
Low	0.6	0.2
High	0.4	0.8

Mgt Committ	Low	High
Low	0.3	0.4
High	0.7	0.6

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.3	0.4	0.5	0.6	0.1	0.2	0.3	0.4
Inadequat	0.7	0.6	0.5	0.4	0.9	0.8	0.7	0.6

Low	0.4
High	0.6

Low	0.3
High	0.7

Mgt Committ	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.3	0.2	0.8	0.7
Inadequat	0.7	0.8	0.2	0.3

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.4	0.3	0.5	0.4	0.2	0.1	0.3	0.2
Inadequat	0.6	0.7	0.5	0.6	0.8	0.9	0.7	0.8

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inade
Adequate	0.4	0.3	0.7	0.5	0.2	0.1	0.5	0.3
Inadequat	0.6	0.7	0.3	0.5	0.8	0.9	0.5	0.7

Resource Ava	Adequate								Inadequate							
Tech Expertis	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0	0.05	0.05	0.1	0.5	0.6	0.7	0.8	0.2	0.3	0.4	0.5	0.6	0.8	0.8	0.9
L2 (0.3)	0.05	0.05	0.1	0.15	0.2	0.15	0.15	0.15	0.2	0.25	0.2	0.2	0.2	0.15	0.15	0.1
M1 (0.5)	0.1	0.15	0.15	0.2	0.15	0.15	0.1	0.05	0.3	0.25	0.2	0.2	0.1	0.05	0.05	0
H1 (0.7)	0.1	0.15	0.3	0.25	0.1	0.05	0.05	0	0.15	0.1	0.1	0.1	0.1	0	0	0
H2 (0.9)	0.75	0.6	0.4	0.3	0.05	0.05	0	0	0.15	0.1	0.1	0	0	0	0	0

Figure C4. M-2 fixed structure BBN – Expert 8

Appendix D

HVAC Fix Effectiveness (*di*) Projection Project 1-5: M-2 SME Models

Table D1

HVAC Fix Effectiveness (di) Projection Project 1 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.5	0.6	0.4	0.6	0	0	0	0	0
Quality system maturity	0.45	0.52	0.46	0.64	0	0	0	0	0
Project time	0.4182	0.62912	0.36364	0.348	1	1	1	1	1
Failure mode complexity	0.56	0.46	0.24	0.32	0.829337	0.39387	0.846774	0.849558	0.583122615
Technical expertise	0.548676	0.670108	0.430393	0.27748	1	1	1	1	1
Resource availability	0.51124	0.584544	0.36182	0.5348	1	1	1	1	1
Design complexity	0.4	0.3	0.2	0.3	0.504937	0.288819	0.306452	0.504425	0.360058837
Test facilities	0.5409	0.753132	0.364994	0.34424	1	1	1	1	1
Fix effectiveness 10%	0.203979	0.195919	0.412897	0.187668	0	0.05	0.0125	0	0.029923339
Fix effectiveness 30%	0.241155	0.328483	0.191188	0.267649	0.1	0.2	0.0375	0.05	0.148394066
Fix effectiveness 50%	0.255573	0.317633	0.133318	0.201005	0.15	0.5	0.05	0.1	0.337268564
Fix effectiveness 70%	0.163718	0.138569	0.102362	0.175045	0.2	0.2	0.1	0.1	0.185151423
Fix effectiveness 90%	0.135575	0.0193967	0.160235	0.168633	0.55	0.05	0.8	0.75	0.299262608
Likelihood	-	-	-		0.017449159	0.035634267	0.006696689	0.002559919	
Prob (M_j)	0.25	0.25	0.25	0.25	0.27990294	0.57161129	0.107421972	0.041063798	
	Mean	Var.	Alpha	Beta					
	0.615087179	0.052766512	2.144711082	1.342129736					

Table D2

HVAC Fix Effectiveness (di) Projection Project 2 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.5	0.6	0.4	0.6	1	1	1	1	1
Quality system maturity	0.45	0.52	0.46	0.64	1	1	1	1	1
Project time	0.4182	0.62912	0.36364	0.348	0.64	0.908722	0.908722	0.559441	0.77026689
Failure mode complexity	0.56	0.46	0.24	0.32	0	0	0	0	0
Technical expertise	0.548676	0.670108	0.430393	0.27748	0.7	0.8	0.8	0.3	0.641845049
Resource availability	0.51124	0.584544	0.36182	0.5348	1	1	1	1	1
Design complexity	0.4	0.3	0.2	0.3	0	0	0	0	0
Test facilities	0.5409	0.753132	0.364994	0.34424	1	1	1	1	1
Fix effectiveness 10%	0.203979	0.195919	0.412897	0.187668	0.1032	0.100953	0.0254767	0.387448	0.174280712
Fix effectiveness 30%	0.241155	0.328483	0.191188	0.267649	0.2674	0.234564	0.0695842	0.1958	0.207138491
Fix effectiveness 50%	0.255573	0.317633	0.133318	0.201005	0.20484	0.439047	0.127323	0.141608	0.280946651
Fix effectiveness 70%	0.163718	0.138569	0.102362	0.175045	0.15392	0.189087	0.170101	0.0911888	0.153526034
Fix effectiveness 90%	0.135575	0.0193967	0.160235	0.168633	0.27064	0.0363489	0.607515	0.240175	0.200309269
Likelihood	-	-	-		0.016425865	0.051920002	0.014774055	0.033650453	
Prob (M_j)	0.25	0.25	0.25	0.25	0.140668087	0.444633345	0.126522284	0.288176284	
	Mean	Var.	Alpha	Beta					
	0.507789509	0.074427483	1.197448403	1.1607106					

Table D3

HVAC Fix Effectiveness (di) Projection Project 3 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.5	0.6	0.4	0.6	0	0	0	0	0
Quality system maturity	0.45	0.52	0.46	0.64	0	0	0	0	0
Project time	0.4182	0.62912	0.36364	0.348	0	0	0	0	0
Failure mode complexity	0.56	0.46	0.24	0.32	0.792453	0.5	0.150588	0.475728	0.487327958
Technical expertise	0.548676	0.670108	0.430393	0.27748	0.457736	0.475	0.150506	0.215146	0.300953342
Resource availability	0.51124	0.584544	0.36182	0.5348	0.2	0.3	0.1	0.2	0.173633583
Design complexity	0.4	0.3	0.2	0.3	1	1	1	1	1
Test facilities	0.5409	0.753132	0.364994	0.34424	0	0	0	0	0
Fix effectiveness 10%	0.203979	0.195919	0.412897	0.187668	0.405302	0.24475	0.657222	0.776233	0.564184491
Fix effectiveness 30%	0.241155	0.328483	0.191188	0.267649	0.281377	0.455	0.223672	0.125612	0.237594694
Fix effectiveness 50%	0.255573	0.317633	0.133318	0.201005	0.225057	0.2185	0.0823271	0.0500777	0.136747162
Fix effectiveness 70%	0.163718	0.138569	0.102362	0.175045	0.0749849	0.08175	0.0270276	0.0303689	0.049210758
Fix effectiveness 90%	0.135575	0.0193967	0.160235	0.168633	0.0132792	0	0.00975118	0.0177087	0.012262963
Likelihood	-	-	-		0.029381482	0.005273764	0.026185189	0.018470398	
Prob (M_j)	0.25	0.25	0.25	0.25	0.370459881	0.066494877	0.330159043	0.232886199	
	Mean	Var.	Alpha	Beta					
	0.241554635	0.036909791	0.957429576	3.006185428					

Table D4

HVAC Fix Effectiveness (di) Projection Project 4 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.5	0.6	0.4	0.6	1	1	1	1	1
Quality system maturity	0.45	0.52	0.46	0.64	0	0	0	0	0
Project time	0.4182	0.62912	0.36364	0.348	1	1	1	1	1
Failure mode complexity	0.56	0.46	0.24	0.32	0.898876	0.642857	0.512195	0.777778	0.725555952
Technical expertise	0.548676	0.670108	0.430393	0.27748	0.798888	0.823143	0.876829	0.455556	0.782952973
Resource availability	0.51124	0.584544	0.36182	0.5348	1	1	1	1	1
Design complexity	0.4	0.3	0.2	0.3	1	1	1	1	1
Test facilities	0.5409	0.753132	0.364994	0.34424	1	1	1	1	1
Fix effectiveness 10%	0.203979	0.195919	0.412897	0.187668	0.0105056	0.0835714	0.0171189	0.272222	0.076944485
Fix effectiveness 30%	0.241155	0.328483	0.191188	0.267649	0.163034	0.225179	0.0513567	0.131667	0.189276711
Fix effectiveness 50%	0.255573	0.317633	0.133318	0.201005	0.171011	0.458036	0.0900305	0.127222	0.323042554
Fix effectiveness 70%	0.163718	0.138569	0.102362	0.175045	0.2	0.191607	0.127713	0.1	0.182135741
Fix effectiveness 90%	0.135575	0.0193967	0.160235	0.168633	0.455449	0.0416071	0.71378	0.368889	0.228600626
Likelihood	-	-	-		0.01272092	0.023929606	0.002074597	0.004151519	
Prob (M_j)	0.25	0.25	0.25	0.25	0.296686487	0.558103554	0.048385244	0.096824715	
	Mean	Var.	Alpha	Beta					
	0.559234321	0.060235018	1.72924165	1.362917728					

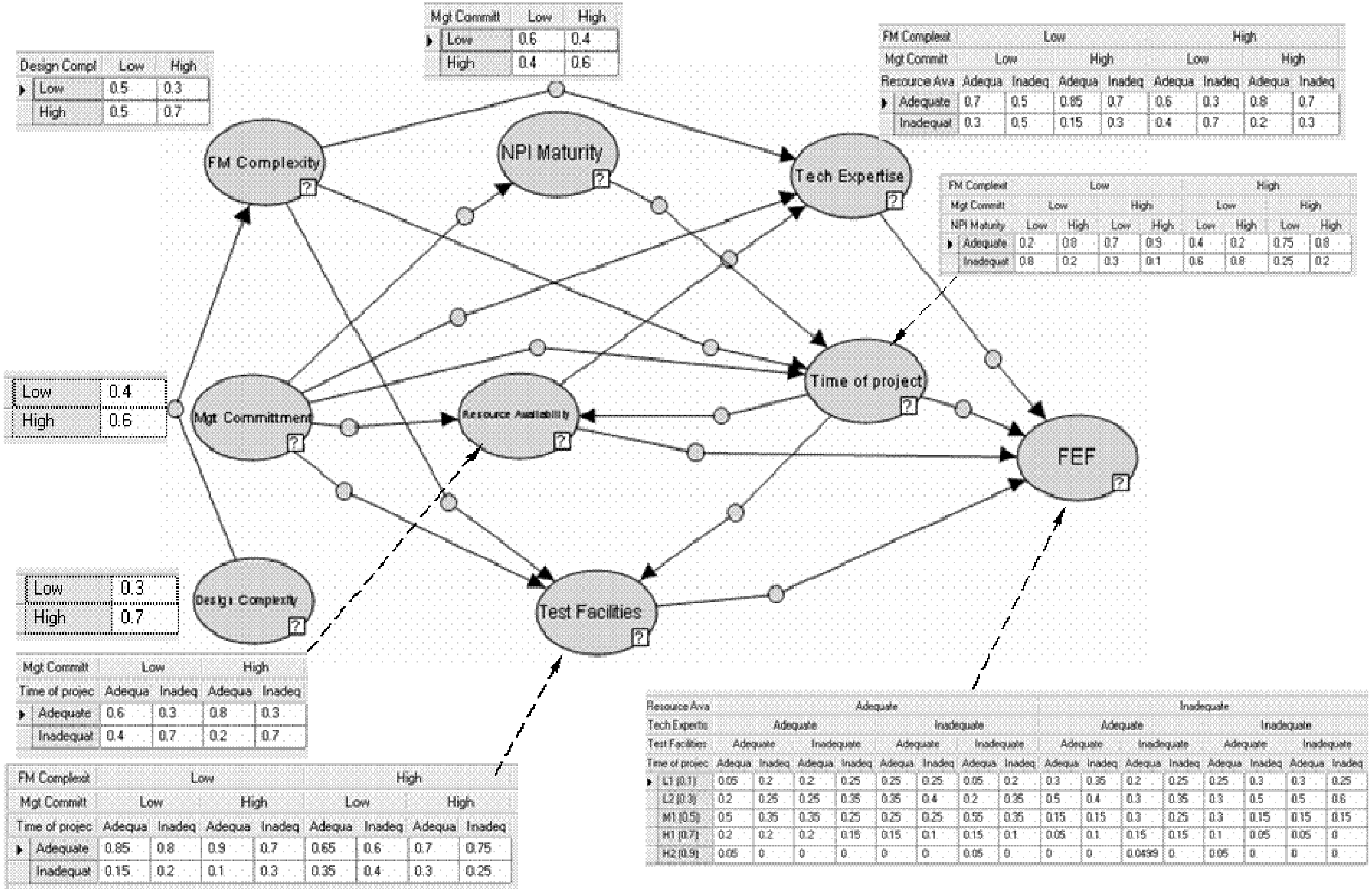
Table D5

HVAC Fix Effectiveness (di) Projection Project 5 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.5	0.6	0.4	0.6	1	1	1	1	1
Quality system maturity	0.45	0.52	0.46	0.64	1	1	1	1	1
Project time	0.4182	0.62912	0.36364	0.348	0	0	0	0	0
Failure mode complexity	0.56	0.46	0.24	0.32	1	1	1	1	1
Technical expertise	0.548676	0.670108	0.430393	0.27748	0.5	0.7	0.7	0.4	0.502796949
Resource availability	0.51124	0.584544	0.36182	0.5348	0	0	0	0	0
Design complexity	0.4	0.3	0.2	0.3	1	1	1	1	1
Test facilities	0.5409	0.753132	0.364994	0.34424	0	0	0	0	0
Fix effectiveness 10%	0.203979	0.195919	0.412897	0.187668	0.4	0.25	0.365	0.74	0.521586604
Fix effectiveness 30%	0.241155	0.328483	0.191188	0.267649	0.275	0.425	0.305	0.14	0.237089054
Fix effectiveness 50%	0.255573	0.317633	0.133318	0.201005	0.25	0.22	0.19	0.08	0.169658985
Fix effectiveness 70%	0.163718	0.138569	0.102362	0.175045	0.06	0.105	0.105	0.04	0.061657539
Fix effectiveness 90%	0.135575	0.0193967	0.160235	0.168633	0.015	0	0.035	0	0.010007819
Likelihood	-	-	-		0.006579717	0.001637785	0.002277629	0.007332206	
Prob (M_j)	0.25	0.25	0.25	0.25	0.36908018	0.091869309	0.127760462	0.411290049	
	Mean	Var.	Alpha	Beta					
	0.260282183	0.03954034	1.007120335	2.862219948					

Appendix E

M-3 Fixed Structure



Design Compl	Low	High
Low	0.5	0.3
High	0.5	0.7

Mgt Commit	Low	High
Low	0.6	0.4
High	0.4	0.6

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Resource Ava	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.7	0.5	0.85	0.7	0.6	0.3	0.8	0.7
Inadequat	0.3	0.5	0.15	0.3	0.4	0.7	0.2	0.3

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
NPI Maturity	Low	High	Low	High	Low	High	Low	High
Adequate	0.2	0.8	0.7	0.9	0.4	0.2	0.75	0.8
Inadequat	0.8	0.2	0.3	0.1	0.6	0.8	0.25	0.2

Low	0.4
High	0.6

Low	0.3
High	0.7

Mgt Commit	Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.6	0.3	0.8	0.3
Inadequat	0.4	0.7	0.2	0.7

FM Complexit	Low				High			
Mgt Commit	Low		High		Low		High	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
Adequate	0.85	0.8	0.9	0.7	0.65	0.6	0.7	0.75
Inadequat	0.15	0.2	0.1	0.3	0.35	0.4	0.3	0.25

Resource Ava	Adequate								Inadequate							
Tech Experts	Adequate				Inadequate				Adequate				Inadequate			
Test Facilities	Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate		Adequate		Inadequate	
Time of projec	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq	Adequa	Inadeq
L1 (0.1)	0.05	0.2	0.2	0.25	0.25	0.25	0.05	0.2	0.3	0.35	0.2	0.25	0.25	0.3	0.3	0.25
L2 (0.3)	0.2	0.25	0.25	0.35	0.35	0.4	0.2	0.35	0.5	0.4	0.3	0.35	0.3	0.5	0.5	0.6
M1 (0.5)	0.5	0.35	0.35	0.25	0.25	0.25	0.55	0.35	0.15	0.15	0.3	0.25	0.3	0.15	0.15	0.15
H1 (0.7)	0.2	0.2	0.2	0.15	0.15	0.1	0.15	0.1	0.05	0.1	0.15	0.15	0.1	0.05	0.05	0
H2 (0.9)	0.05	0	0	0	0	0	0.05	0	0	0	0.0499	0	0.05	0	0	0

Figure E1. M-3 fixed structure concensus CPT BBN

Appendix F

HVAC Fix Effectiveness (*di*) Projection Method: M-3 SME Models

Table F1

HVAC Fix Effectiveness (di) Projection Project 1 – Method: M-3 SME Models

Node Variable	Prior M-3	Posterior Project 1	Project 2	Project 3	Project 4	Project 5
mgt commitment	0.525	0	1	0	1	1
quality system maturity	0.531	0	1	0	0	1
project time	0.444	1	0.756	0	1	0
failure mode complexity	0.365	0.652	0	0.506	0.690	1
technical expertise	0.491	1	0.650	0.344	0.736	0.575
resource availability	0.481	1	1	0.200	1	0
design complexity	0.300	0.378	0	1	1	1
test facilities	0.518	1	1	0	1	0
fix effectiveness 10%	0.299	0.016	0.073	0.549	0.044	0.553
fix effectiveness 30%	0.245	0.097	0.146	0.296	0.127	0.295
fix effectiveness 50%	0.209	0.200	0.236	0.121	0.215	0.125
fix effectiveness 70%	0.123	0.150	0.198	0.027	0.170	0.021
fix effectiveness 90%	0.124	0.538	0.348	0.007	0.443	0.006
Likelihood	-	-	-	-	-	-
Prob (Mj)	-	-	-	-	-	-
Mean Projection		0.719	0.620	0.230	0.668	0.226
Variance		0.050	0.067	0.029	0.062	0.027
Alpha		2.171	1.574	1.182	1.734	1.233
Beta		0.847	0.963	3.961	0.861	4.214

Appendix G

HVAC Fix Effectiveness (*di*) Project 1-5: M-4 SME Models

Table G1

HVAC Fix Effectiveness (di) Project 1 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	0	0	0	0	0	0	0	0	0	0.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	0	0	0	0	0	0	0	0	0	0.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	.769	.240	.344	.407	.829	.394	.847	.850	.652	0.528
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	.824	.499	.217	.555	.505	.289	.306	.504	.378	0.507
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table G1 (*continued*).

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effective-ness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.065	.050	.069	.057	.000	.050	.013	.000	.016	.044
Fix effective-ness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.231	.200	.148	.162	.100	.200	.038	.050	.097	.173
Fix effective-ness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.465	.500	.250	.227	.150	.500	.050	.100	.200	.368
Fix effective-ness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.196	.200	.370	.277	.200	.200	.100	.100	.150	.202
Fix effective-ness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.043	.050	.162	.277	.550	.050	.800	.750	.538	.214
Likeli-hood	-		-		-		-			3.13 E-02	3.30 E-02	3.12 E-04	2.47 E-02	1.74 E-02	3.56 E-02	6.70 E-03	2.56 E-03	1.21 E-02	
Prob (M$_j$)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.19	.20	.00	.15	.11	.22	.04	.02	.07	
																Mean	Var.	Alpha	Beta
																0.574	0.051	2.190	1.627

Note: *Aggregate

Table G2

HVAC Fix Effectiveness (di) Project 2 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	.971	.545	.569	.850	.640	.909	.909	.559	.756	.751
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	0	0	0	0	0	0	0	0	0	.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	.931	.800	.850	.705	.700	.800	.800	.300	.650	.711
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	0	0	0	0	0	0	0	0	0	.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table G2 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.100	.080	.160	.279	.103	.101	.025	.387	.073	.168
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.300	.180	.224	.365	.267	.235	.070	.196	.146	.243
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.550	.200	.228	.214	.205	.439	.127	.142	.236	.255
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.013	.420	.217	.071	.154	.189	.170	.091	.198	.187
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.037	.120	.172	.071	.271	.036	.608	.240	.348	.153
Likelihood	-		-		-		-			1.75 E-02	7.30 E-02	2.22 E-04	8.83 E-02	1.64 E-02	5.19 E-02	1.48 E-02	3.37 E-02	3.08 E-02	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.05	.22	.00	.27	.05	.16	.05	.10	.09	
																Mean	Var.	Alpha	Beta
																0.486	0.068	1.293	1.369

Note: *Aggregate

Table G3

HVAC Fix Effectiveness (di) Project 3 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	0	0	0	0	0	0	0	0	0	.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	0	0	0	0	0	0	0	0	0	.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	0	0	0	0	0	0	0	0	0	.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	.299	.272	.500	.333	.792	.500	.151	.476	.506	.474
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	.259	.817	.450	.876	.458	.475	.151	.215	.344	.416
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	.069	.809	.010	.433	.200	.300	.100	.200	.200	.153
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	0	0	0	0	0	0	0	0	0	.000

Table G3 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.225	.225	.629	.115	.405	.245	.657	.776	.549	.553
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.510	.302	.267	.230	.281	.455	.224	.126	.296	.263
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.204	.303	.096	.216	.225	.219	.082	.050	.121	.132
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.052	.170	.007	.233	.075	.082	.027	.030	.027	.042
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.009	.000	.000	.205	.013	.000	.010	.018	.007	.010
Likelihood	-		-		-		-			2.56 E-03	1.40 E-02	8.87 E-02	4.76 E-03	2.94 E-02	5.27 E-03	2.62 E-02	1.85 E-02	1.79 E-02	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.01	.07	.43	.02	.14	.03	.13	.09	.09	
																Mean	Var.	Alpha	Beta
																0.239	0.034	1.036	3.305

Note: *Aggregate

Table G4

HVAC Fix Effectiveness (di) Project 4 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	0	0	0	0	0	0	0	0	0	.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	.622	.217	.500	.444	.899	.643	.512	.778	.690	.579
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	.315	.822	.920	.879	.799	.823	.877	.456	.736	.602
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table G4 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.082	.052	.086	.035	.011	.084	.017	.272	.044	.068
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.282	.159	.138	.126	.163	.225	.051	.132	.127	.212
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.424	.523	.239	.241	.171	.458	.090	.127	.215	.369
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.161	.232	.406	.301	.200	.192	.128	.100	.170	.197
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.051	.034	.131	.297	.455	.042	.714	.369	.443	.153
Likelihood	-		-		-		-			7.01 E-02	1.81 E-02	2.74 E-05	2.64 E-02	1.27 E-02	2.39 E-02	2.07 E-03	4.15 E-03	8.15 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.42	.11	.00	.16	.08	.14	.01	.03	.05	
																Mean	Var.	Alpha	Beta
																0.531	0.051	2.071	1.828

Note: *Aggregate

Table G5

HVAC Fix Effectiveness (di) Project 5 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	0	0	0	0	0	0	0	0	0	.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	1	1	1	1	1	1	1	1	1	1.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	.907	.900	.990	.654	.500	.700	.700	.400	.575	.574
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	0	0	0	0	0	0	0	0	0	.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	0	0	0	0	0	0	0	0	0	.000

Table G5 (*continued*).

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effective-ness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.095	.250	.402	.142	.400	.250	.365	.740	.553	.465
Fix effective-ness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.300	.555	.399	.210	.275	.425	.305	.140	.295	.264
Fix effective-ness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.400	.150	.199	.200	.250	.220	.190	.080	.125	.184
Fix effective-ness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.200	.045	.001	.252	.060	.105	.105	.040	.021	.072
Fix effective-ness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.005	.000	.000	.196	.015	.000	.035	.000	.006	.014
Likeli-hood	-		-		-		-			2.44 E-03	9.93 E-04	1.52 E-04	7.67 E-04	6.58 E-03	1.64 E-03	2.28 E-03	7.33 E-03	4.26 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.09	.04	.01	.03	.25	.06	.09	.28	.16	
																Mean	Var.	Alpha	Beta
																0.281	0.042	1.065	2.725

Note: *Aggregate

Appendix H

HVAC Case Study Analysis

HVAC Case Study Analysis

The process for analysis involved the following:

1. Power value for each experiment was held constant at 0.9. M-1 BBN model standard deviation was used as the baseline reference for sample size calculation. The difference in FEF projection we wished to detect was considered 0.05.
2. Perform random realizations per power and sample size calculations of previous step.
3. Perform test of equal variances.
4. Perform ANOVA to determine statistical significance among model means, with the null hypothesis equating to no difference in means. Perform Tukey pairwise comparisons as necessary.
5. Perform a one sample T-test comparing each model mean against the actual FEF obtained from field data.
6. Repeat analysis for the next project

Results for: HVAC Project-1

Power and sample size calculations indicate 455 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Analysis of variances indicated one must reject the null of no difference among variances with M-1 representing the lowest variance among the models.

One way ANOVA analysis indicated one must reject the null hypothesis (Ho) of no difference in the models. Finally, the one sample T looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject Ho of no difference for all models except M-4. M-4 proved to be statistically significance whereby one must fail to reject the null hypothesis of no difference.

Power and Sample Size

Table H1

One-way ANOVA

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	455	0.9000	0.9004	0.05

Note: Sigma = 0.2, Alpha = 0.05, Number of Levels = 4

Table H2

One-way ANOVA: HVAC Project-1: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
Factor	3	9.0942	3.0314	62.81	0.000
Error	1816	87.6515	0.0483		
Total	1819	96.7457			

				Individual 95% CIs for Mean Based on Pooled StDev
Level	N	Mean	StDev	---------+---------+---------+-------
M1	455	0.5209	0.2055	(-*--)
M2	455	0.6270	0.2267	(--*-)
M3	455	0.7145	0.2261	(--*--)
M4	455	0.5810	0.2198	(--*--) ---------+---------+---------+-------
Polled StDev =		0.2197		0.560 0.630 0.700

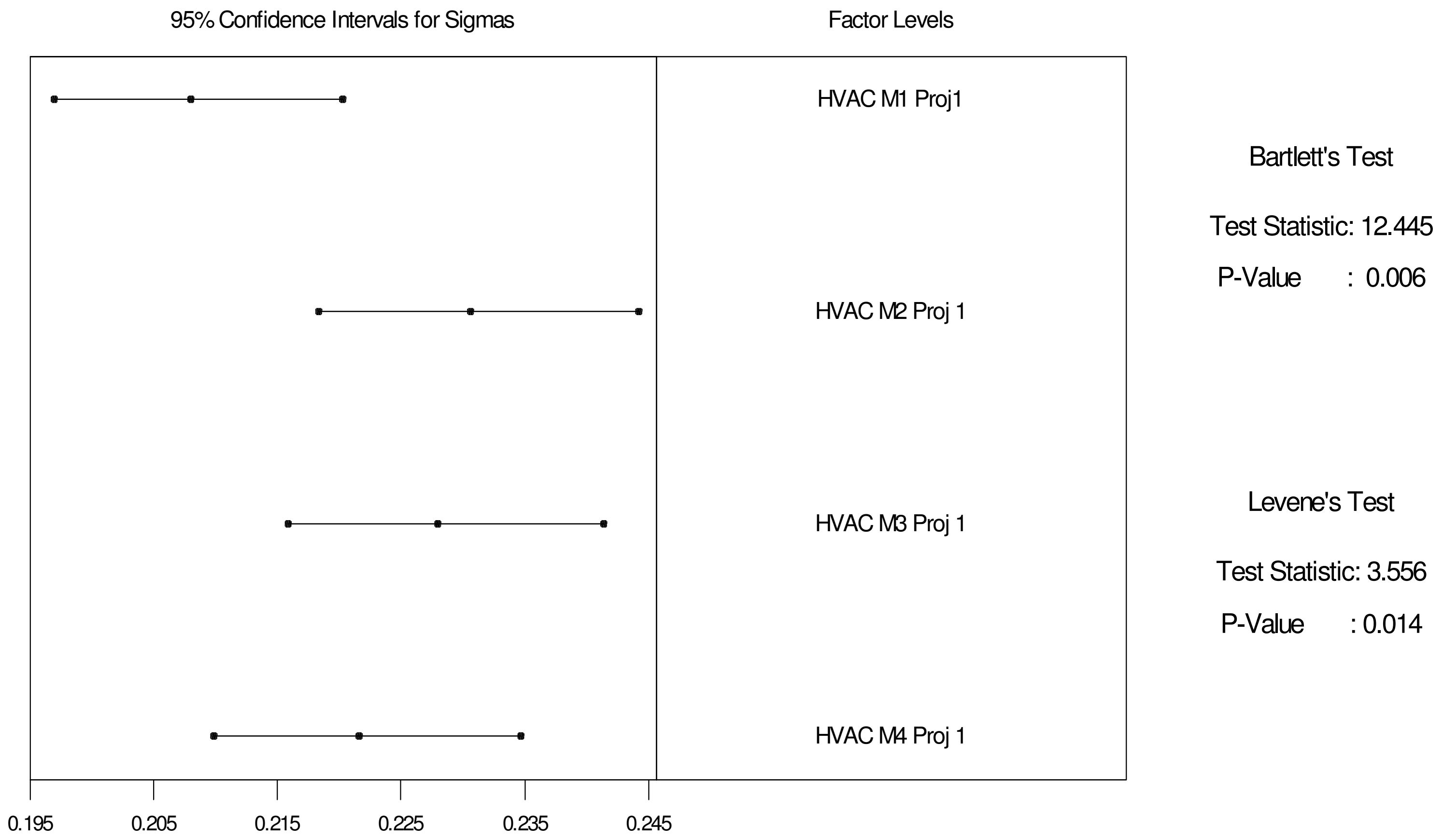


Figure H1. Test for equal variances for HVAC project 1

Table H3

Tukey's Pairwise Comparisons

	HVAC M1	HVAC M2	HVAC M3
HVAC M2	-0.1131		
	-0.0621		
HVAC M3	-0.2173	-0.1298	
	-0.1663	-0.0787	
HVAC M4	-0.0812	0.0064	0.1107
	-0.0301	0.0574	0.1617

Table H4

One-Sample T: HVAC Project-1: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1	455	0.52088	0.20553	0.00964
M2	455	0.6270	0.2267	0.0106
M3	455	0.7145	0.2261	0.0106
M4	455	0.5810	0.2198	0.0103

Variable	95.0% CI	T	P
M1	(0.50194, 0.53981)	-6.14	0.000
M2	(0.6061, 0.6479)	4.42	0.000
M3	(0.6937, 0.7353)	12.69	0.000
M4	(0.5607, 0.6012)	0.10	0.923

Note: Test of mu = 0.58 vs mu not = 0.58

Results for: HVAC Project-2

Power and sample size calculations indicated 682 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Analysis of variances indicated one must reject the null of no difference among variances with M-1 representing the lowest variance among the models.

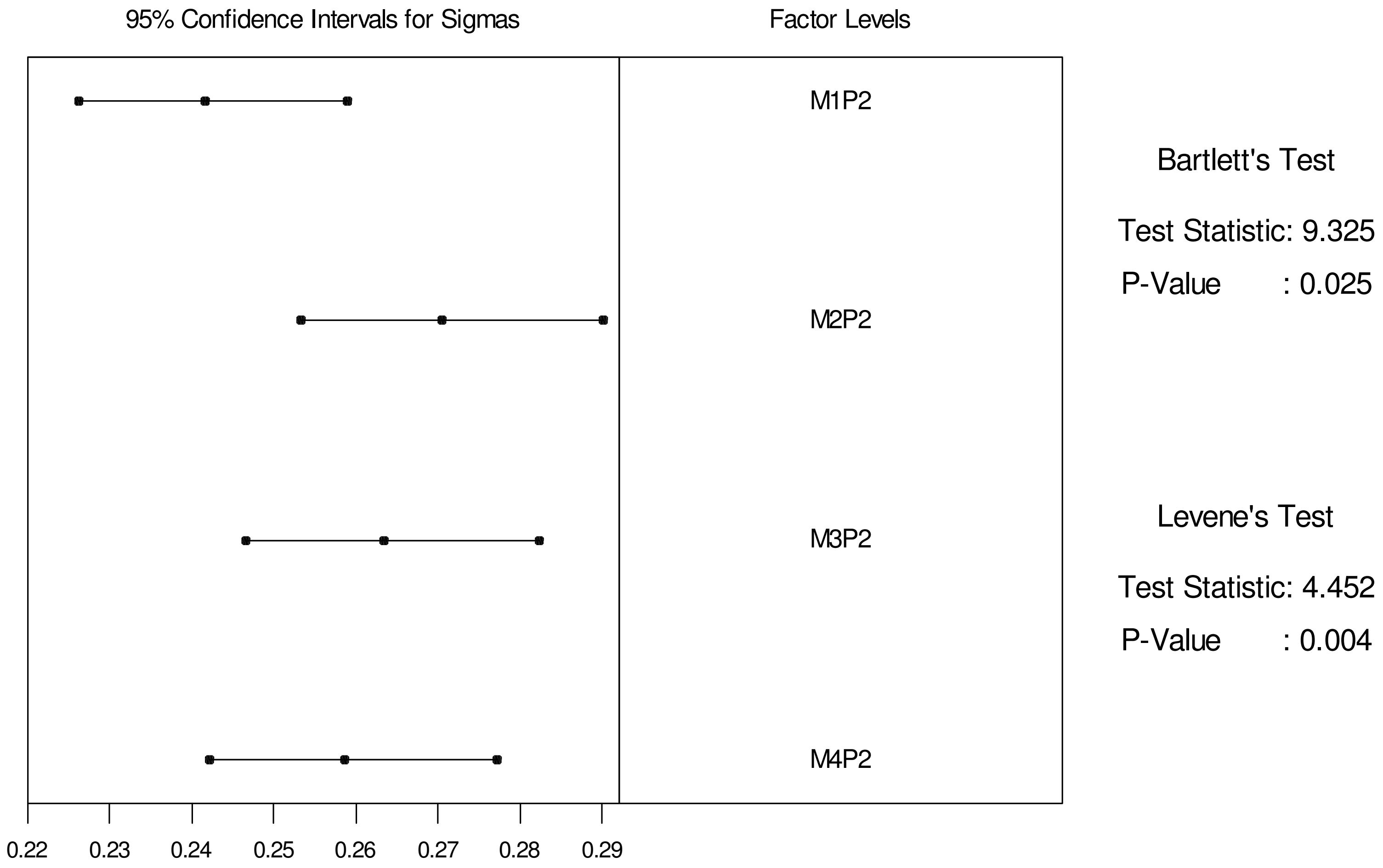
One way ANOVA analysis indicated one must reject the null hypothesis of no difference in the models. However Tukey pariwise comparisons indicated statistical significance in FEF projection between M-2 and M-4. Finally, the one sample T looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject Ho of no difference for all models.

Table H5

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	682	0.9000	0.9004	0.05

Note: Sigma = 0.24495, Alpha = 0.05, Number of Levels = 4



Bartlett's Test

Test Statistic: 9.325

P-Value : 0.025

Levene's Test

Test Statistic: 4.452

P-Value : 0.004

Figure H2. Test for equal variances for HVAC project 2

Table H6

One-way ANOVA: HVAC Project-2: Model to Model Evaluation

Analysis of Variance for HVACP2St

Source	DF	SS	MS	F	P
HVACP2St	3	9.8018	3.2673	48.81	0.000
Error	2724	182.3336	0.0669		
Total	2727	192.1354			

Level	N	Mean	StDev
M1P2	682	0.4415	0.2416
M2P2	682	0.5177	0.2705
M3P2	682	0.6083	0.2633
M4P2	682	0.4984	0.2586
Polled StDev =		0.2587	

```
Individual 95% CIs for Mean
Based on Pooled StDev
---------+---------+---------+-------
 (---*--)
                (--*----)
                            (--*---)
        (--*--)
---------+---------+---------+-------
       0.480     0.540     0.600
```

Table H7

Tukey's Pairwise Comparisons

	M1P2	M2P2	M3P2
M2P2	-0.1121		
	-0.0402		
M3P2	-0.2027	-0.1265	
	-0.1308	-0.0546	
M4P2	-0.0928	-0.0166	0.0739
	-0.0209	0.0553	0.1459

Table H8

One-Sample T: HVAC Project-2: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P2	682	0.44153	0.24159	0.00925
M2P2	682	0.5177	0.2705	0.0104
M3P2	682	0.6083	0.2633	0.0101
M4P2	682	0.49836	0.25857	0.00990

Variable	95.0% CI	T	P
M1	(0.42336, 0.45969)	-27.94	0.000
M2	(0.4974, 0.5380)	-17.60	0.000
M3	(0.5885, 0.6281)	-9.10	0.000
M4	(0.47892, 0.51780)	-20.37	0.000

Note: Test of mu = 0.7 vs mu not = 0.7

Results for: HVAC Project-3

Power and sample size calculations indicated 376 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Analysis of variances indicated one must reject the null of no difference among variances with M-1 representing the lowest variance among the models.

One way ANOVA analysis indicated one must fail to reject the null hypothesis of no difference in the models. FEF projections among all models are statistically significant. Finally, the one sample T looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject Ho of no difference for all models.

Table H9

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	376	0.9000	0.9008	0.05

Note: Sigma = 0.181659, Alpha = 0.05, Number of Levels = 4

Table H10

One-way ANOVA: HVAC Project-3: Model to Model Evaluation

Analysis of Variance for HVACP3St

Source	DF	SS	MS	F	P
HVACP3St	3	0.1213	0.0404	1.14	0.332
Error	1500	53.1991	0.0355		
Total	1503	53.3204			

```
                                Individual 95% CIs for Mean
                                Based on Pooled StDev
Level    N     Mean     StDev   -----+---------+---------+---------+-
M1P3    376   0.2348   0.1749      (--------*---------)
M2P3    376   0.2480   0.1985             (---------*---------)
M3P3    376   0.2307   0.1807   (--------*---------)
M4P3    376   0.2524   0.1981                 (--------*---------)
                                -----+---------+---------+---------+-
Pooled StDev = 0.1883             0.220     0.240     0.260      0.280
```

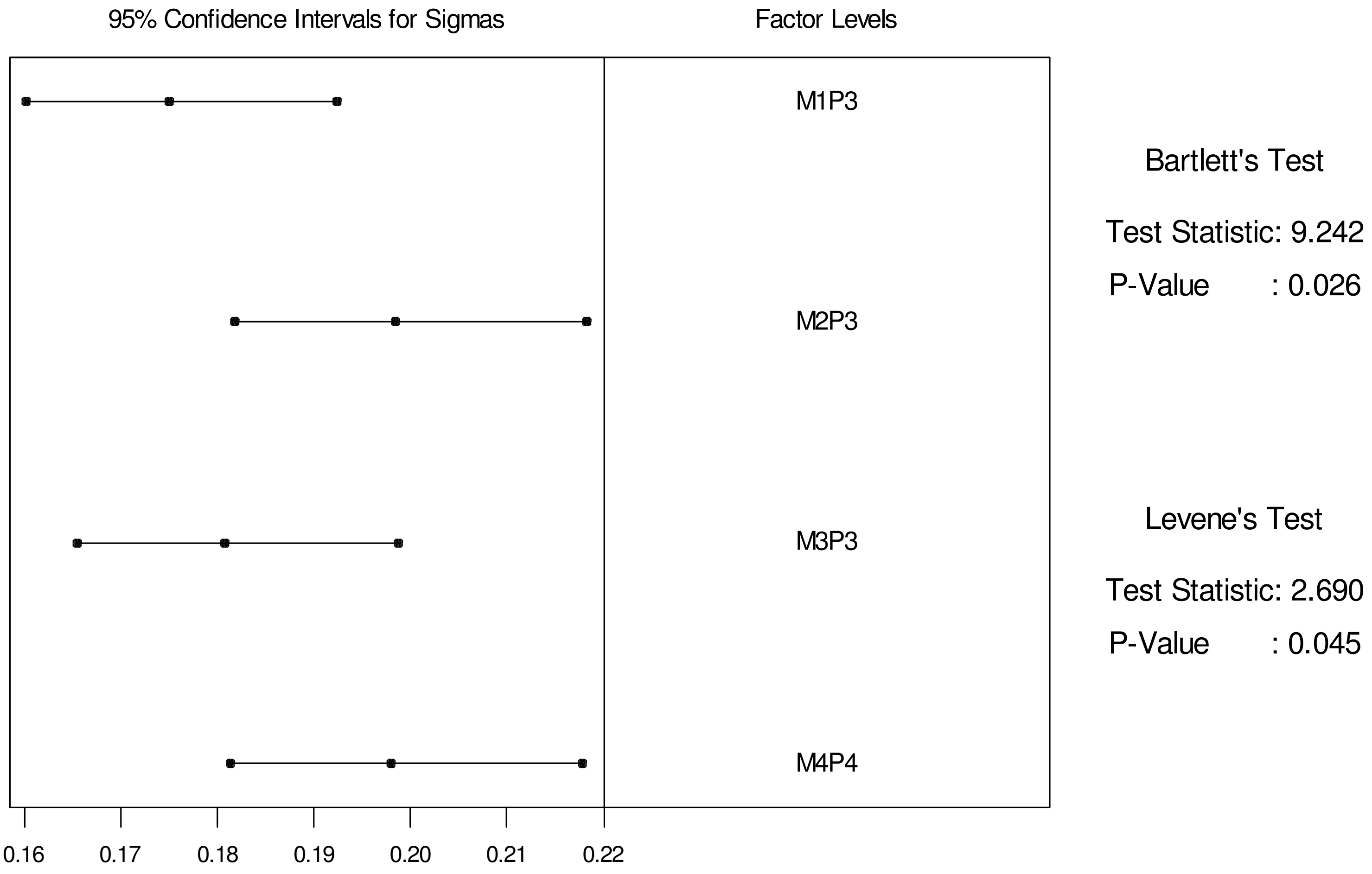


Figure H3. Test for equal variances for HVAC project-3

Table H11

Tukey's Pairwise Comparisons

	M1P3	M2P3	M3P3
M2P3	-0.0485		
	0.0220		
M3P3	-0.0312	-0.0180	
	0.0393	0.0525	
M4P4	-0.0529	-0.0397	-0.0570
	0.0176	0.0308	0.0136

Table H12

One-Sample T: HVAC Project-3: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P3	376	0.23475	0.17493	0.00902
M2P3	376	0.2480	0.1985	0.0102
M3P3	376	0.23068	0.18068	0.00932
M4P4	376	0.2524	0.1981	0.0102

Variable	95.0% CI	T	P
M1P3	(0.21701, 0.25249)	-7.23	0.000
M2P3	(0.2278, 0.2681)	-5.08	0.000
M3P3	(0.21236, 0.24900)	-7.44	0.000
M4P4	(0.2323, 0.2725)	-4.66	0.000

Note: Test of mu = 0.3 vs mu not = 0.3

Results for: HVAC Project-4

Power and sample size calculations indicated 512 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Analysis of variances indicated one must reject the null of no difference among variances with M-1 representing the lowest variance among the models.

One way ANOVA analysis indicated one must reject the null hypothesis of no difference in the models. Tukey pariwise comparisons indicated statistical significance in FEF projection between M-1 and M-4 and statistical significance between M-2 and M-4. Finally, the one sample T looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject Ho of no difference for all models.

Table H13

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	512	0.9000	0.9005	0.05

Note: Sigma = 0.212132, Alpha = 0.05, Number of Levels = 4

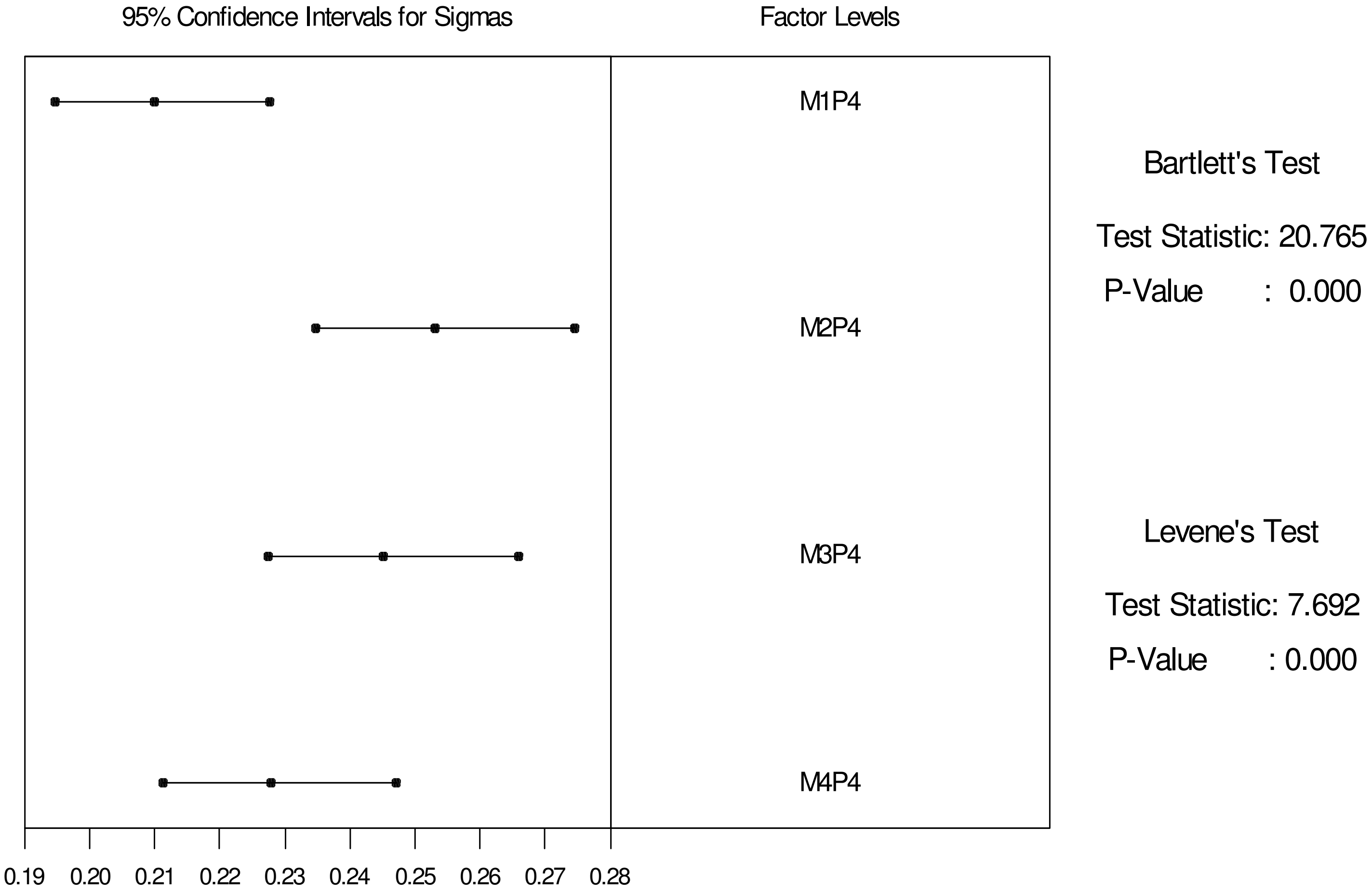


Figure H4. Test for equal variances for HVAC project 4

Table H14

One-way ANOVA: HVAC Project-4: Model to Model Evaluation

Analysis of Variance for HVACP4St

Source	DF	SS	MS	F	P
HVACP3St	3	8.1270	2.7090	49.21	0.000
Error	2044	112.5186	0.0550		
Total	2047	120.6455			

```
                                          Individual 95% CIs for Mean
                                           Based on Pooled StDev

Level     N      Mean      StDev     ---------+-----------+---------+---------
M1P4     512    0.5090    0.2100     (---*--)
                                               (---*--)
M2P4     512    0.5622    0.2531
                                                                    (---*--)
M3P4     512    0.6754    0.2452
                                        (---*--)
M4P4     512    0.5386    0.2278     ---------+-----------+---------+---------
Pooled StDev =  0.2346                    0.540      0.600      0.660
```

Table H15

Tukey's Pairwise Comparisons

	M1P4	M2P4	M3P4
M2P4	-0.0908		
	-0.0156		
M3P4	-0.2040	-0.1508	
	-0.1288	-0.0756	
M4P4	-0.0672	-0.0140	0.0992
	0.0081	0.0613	0.1745

Table H16

One-Sample T: HVAC Project-4: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P4	512	0.50902	0.21000	0.00928
M2P4	512	0.5622	0.2531	0.0112
M3P4	512	0.6754	0.2452	0.0108
M4P4	512	0.5386	0.2278	0.0101

Variable	95.0% CI	T	P
M1P4	(0.49078, 0.52725)	-11.96	0.000
M2P4	(0.5402, 0.5842)	-5.17	0.000
M3P4	(0.6541, 0.6967)	5.11	0.000
M4P4	(0.5188, 0.5584)	-8.09	0.000

Note: Test of mu = 0.62 vs mu not = 0.62

Results for: HVAC Project-5

Power and sample size calculations indicated 500 samples were required given the variance of M-1 such that the power of the experiment can be held constant at 0.9. Analysis of variances indicated one must reject the null of no difference among variances with M-3 representing the lowest variance among the models.

One way ANOVA analysis indicated one must reject the null hypothesis of no difference in the models. Tukey pariwise comparisons indicate statistical significance in FEF projection between M-2 and M-4. One sample T looked at each model FEF projection mean to actual FEF obtained from field data and indicated one must reject Ho of no difference for all models.

Table H17

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	500	0.9000	0.9004	0.05

Note: Sigma = 0.209762, Alpha = 0.05, Number of Levels = 4

Table H18

One-way ANOVA: HVAC Project-5: Model to Model Evaluation

Analysis of Variance for HVACP5St

Source	DF	SS	MS	F	P
HVACP5St	3	10.6002	3.5334	90.53	0.000
Error	1996	77.9026	0.0390		
Total	1999	88.5028			

```
                                    Individual 95% CIs for Mean
                                    Based on Pooled StDev
Level    N    Mean   StDev   -+-----------+-----------+-----------+-----
M1P5   500  0.4203  0.2056                                       (-*--)
M2P5   500  0.2675  0.2014            (-*--)
M3P5   500  0.2245  0.1703   (-*--)
M4P5   500  0.2988  0.2105                (--*-)
                             -+-----------+-----------+-----------+-----
Pooled StDev = 0.2346       0.210       0.280       0.350       0.420
```

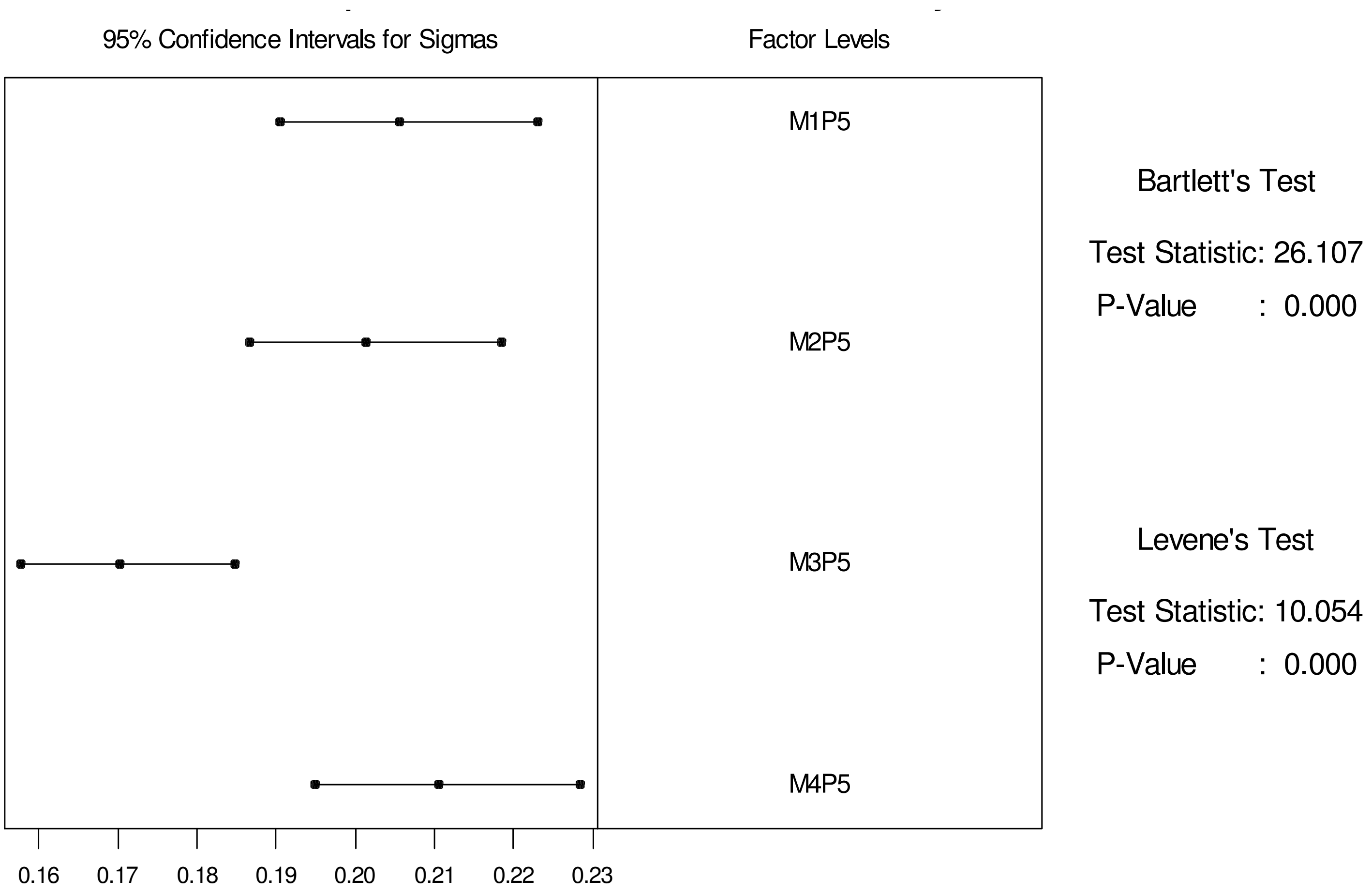


Figure H5. Test for equal variances for HVAC project 5

Table H19

Tukey's Pairwise Comparisons

	M1P5	M2P5	M3P5
M2P5	0.1208		
	0.1849		
M3P5	0.1637	0.0109	
	0.2279	0.0750	
M4P5	0.0895	-0.0634	-0.1063
	0.1536	0.0008	-0.0422

Table H20

One-Sample T: HVAC Project-5: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P5	500	0.42032	0.20560	0.00919
M2P5	500	0.26747	0.20136	0.00901
M3P5	500	0.22453	0.17031	0.00762
M4P5	500	0.29876	0.21046	0.00941

Variable	95.0% CI	T	P
M1P5	(0.40225, 0.43838)	-7.58	0.000
M2P5	(0.24978, 0.28516)	-24.71	0.000
M3P5	(0.20957, 0.23950)	-34.85	0.000
M4P5	(0.28027, 0.31726)	-20.32	0.000

Note: Test of mu = 0.49 vs mu not = 0.49

Appendix I

Automotive Fix Effectiveness (*di*) Project 1-5: M-1 SME Models

Table I1

Automotive Fix Effectiveness (di) Projection Project 1 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	0	0	0	0	0.000
Failure mode complexity	0.610	0.200	0.119	0.250	0	0	0	0	0.000
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.10	0.05	0.10	0.04	0.054
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.30	0.10	0.20	0.12	0.133
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.55	0.20	0.25	0.20	0.248
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.01	0.50	0.25	0.32	0.384
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.04	0.15	0.20	0.32	0.180
Likelihood	-	-	-		3.43E-03	1.47E-02	1.19E-05	6.72E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.14	0.59	0.00	0.27	
	Mean	Var.	Alpha	Beta					
	0.601	0.048	2.395	1.593					

Table I2

Automotive Fix Effectiveness (di) Projection Project 2 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	0	0	0	0	0.000
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	0.700	0.760	0.900	0.770	0.733
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.115	0.064	0.110	0.065	0.090
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.315	0.124	0.210	0.138	0.223
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.520	0.224	0.258	0.189	0.361
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.023	0.464	0.240	0.304	0.197
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.027	0.126	0.182	0.303	0.130
Likelihood	-	-	-		4.37E-02	1.80E-02	3.10E-06	2.60E-02	
Prob (M_j)	0.25	0.25	0.25	0.25	0.50	0.21	0.00	0.30	
	Mean	Var.	Alpha	Beta					
	0.511	0.052	1.954	1.872					

Table I3

Automotive Fix Effectiveness (di) Projection Project 3 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	0	0	0	0	0.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	1	1	1	1	1.000
Technical expertise	0.739	0.820	0.310	0.700	0.417	0.900	0.990	0.786	0.577
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	0.885	0.487	0.526	0.468	0.722
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.040	0.038	0.012	0.059	0.045
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.150	0.155	0.053	0.162	0.154
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.530	0.535	0.250	0.231	0.449
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.250	0.240	0.596	0.280	0.257
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.030	0.032	0.090	0.269	0.095
Likelihood	-	-	-		5.89E-02	1.21E-02	2.12E-05	2.64E-02	
Prob (M_j)	0.25	0.25	0.25	0.25	0.60	0.12	0.00	0.27	
	Mean	Var.	Alpha	Beta					
	0.541	0.037	3.074	2.611					

Table I4

Automotive Fix Effectiveness (di) Projection Project 4 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	1	1	1	1	1.000
Technical expertise	0.739	0.820	0.310	0.700	1	1	1	1	1.000
Resource availability	0.552	0.743	0.108	0.785	1	1	1	1	1.000
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.09	0.05	0.01	0.00	0.081
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.10	0.10	0.05	0.10	0.100
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.60	0.20	0.25	0.20	0.547
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.20	0.50	0.60	0.35	0.227
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.01	0.15	0.09	0.35	0.045
Likelihood	-	-	-		4.83E-02	2.74E-03	9.24E-08	4.63E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.87	0.05	0.00	0.08	
	Mean	Var.	Alpha	Beta					
	0.511	0.033	3.351	3.203					

Table I5

Automotive Fix Effectiveness (di) Projection Project 4 – Method: M-1 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.695	0.600	0.150	0.750	1	1	1	1	1.000
Quality system maturity	0.604	0.480	0.084	0.500	1	1	1	1	1.000
Project time	0.900	0.428	0.054	0.674	1	1	1	1	1.000
Failure mode complexity	0.610	0.200	0.119	0.250	0	0	0	0	0.000
Technical expertise	0.739	0.820	0.310	0.700	0	0	0	0	0.000
Resource availability	0.552	0.743	0.108	0.785	0.750	0.950	0.990	0.723	0.765
Design complexity	0.726	0.300	0.154	0.250	1	1	1	1	1.000
Test facilities	0.707	0.609	0.220	0.533	1	1	1	1	1.000
Fix effectiveness 10%	0.110	0.139	0.683	0.188	0.10	0.20	0.50	0.14	0.120
Fix effectiveness 30%	0.258	0.267	0.202	0.268	0.30	0.50	0.36	0.19	0.295
Fix effectiveness 50%	0.472	0.345	0.066	0.201	0.55	0.20	0.10	0.17	0.418
Fix effectiveness 70%	0.138	0.205	0.035	0.175	0.01	0.10	0.03	0.25	0.081
Fix effectiveness 90%	0.023	0.043	0.014	0.169	0.04	0.00	0.01	0.25	0.086
Likelihood	-	-	-		1.98E-02	3.24E-03	1.40E-05	7.58E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.65	0.11	0.00	0.25	
	Mean	Var.	Alpha	Beta					
	0.443	0.045	2.000	2.511					

Appendix J

Automotive Fix Effectiveness (*di*) Project 1-5: M-2 SME Models

Table J1

Automotive Fix Effectiveness (di) Projection Project 1 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.500	0.600	0.400	0.600	1	1	1	1	1.000
Quality system maturity	0.450	0.520	0.460	0.640	1	1	1	1	1.000
Project time	0.418	0.629	0.364	0.348	0	0	0	0	0.000
Failure mode complexity	0.560	0.460	0.240	0.320	0	0	0	0	0.000
Technical expertise	0.549	0.670	0.430	0.277	1	1	1	1	1.000
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000
Design complexity	0.400	0.300	0.200	0.300	1	1	1	1	1.000
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.200	0.200	0.050	0.050	0.157
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.450	0.250	0.100	0.050	0.252
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.270	0.350	0.250	0.150	0.279
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.060	0.200	0.500	0.150	0.175
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.020	0.000	0.100	0.600	0.137
Likelihood	-	-	-		3.50E-03	5.53E-03	1.01E-03	2.61E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.28	0.44	0.08	0.21	
	Mean	Var.	Alpha	Beta					
	0.477	0.064	1.392	1.528					

Table J2

Automotive Fix Effectiveness (di) Projection Project 2 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.500	0.600	0.400	0.600	1	1	1	1	1.000
Quality system maturity	0.450	0.520	0.460	0.640	1	1	1	1	1.000
Project time	0.418	0.629	0.364	0.348	1	1	1	1	1.000
Failure mode complexity	0.560	0.460	0.240	0.320	0	0	0	0	0.000
Technical expertise	0.549	0.670	0.430	0.277	1	1	1	1	1.000
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000
Design complexity	0.400	0.300	0.200	0.300	1	1	1	1	1.000
Test facilities	0.541	0.753	0.365	0.344	0.600	0.700	0.800	0.500	0.651
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.032	0.095	0.020	0.025	0.064
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.068	0.215	0.060	0.075	0.149
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.250	0.455	0.110	0.125	0.330
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.240	0.200	0.150	0.200	0.205
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.410	0.035	0.660	0.575	0.251
Likelihood	-	-	-		4.65E-03	1.25E-02	1.58E-03	4.05E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.20	0.55	0.07	0.18	
	Mean	Var.	Alpha	Beta					
	0.586	0.057	1.894	1.339					

Table J3

Automotive Fix Effectiveness (di) Projection Project 3 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.500	0.600	0.400	0.600	1	1	1	1	1.000
Quality system maturity	0.450	0.520	0.460	0.640	0	0	0	0	0.000
Project time	0.418	0.629	0.364	0.348	1	1	1	1	1.000
Failure mode complexity	0.560	0.460	0.240	0.320	1	1	1	1	1.000
Technical expertise	0.549	0.670	0.430	0.277	0.800	0.850	0.950	0.500	0.814
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000
Design complexity	0.400	0.300	0.200	0.300	0.571	0.391	0.333	0.563	0.454
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.010	0.080	0.014	0.250	0.069
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.160	0.223	0.043	0.125	0.190
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.170	0.463	0.066	0.125	0.337
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.200	0.193	0.111	0.100	0.185
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.460	0.043	0.765	0.400	0.219
Likelihood	-	-	-		1.78E-02	3.67E-02	2.49E-03	4.43E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.29	0.60	0.04	0.07	
	Mean	Var.	Alpha	Beta					
	0.559	0.058	1.832	1.447					

Table J4

Automotive Fix Effectiveness (di) Projection Project 4 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.500	0.600	0.400	0.600	1	1	1	1	1.000
Quality system maturity	0.450	0.520	0.460	0.640	1	1	1	1	1.000
Project time	0.418	0.629	0.364	0.348	1	1	1	1	1.000
Failure mode complexity	0.560	0.460	0.240	0.320	1	1	1	1	1.000
Technical expertise	0.549	0.670	0.430	0.277	1	1	1	1	1.000
Resource availability	0.511	0.585	0.362	0.535	1	1	1	1	1.000
Design complexity	0.400	0.300	0.200	0.300	1	1	1	1	1.000
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.000	0.050	0.013	0.000	0.033
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.100	0.200	0.038	0.050	0.163
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.150	0.500	0.050	0.100	0.378
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.200	0.200	0.100	0.100	0.193
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.550	0.050	0.800	0.750	0.232
Likelihood	-	-	-		3.20E-03	7.99E-03	1.83E-04	6.55E-04	
Prob (M_j)	0.25	0.25	0.25	0.25	0.27	0.66	0.02	0.05	
	Mean	Var.	Alpha	Beta					
	0.586	0.049	2.289	1.620					

Table J5

Automotive Fix Effectiveness (di) Projection Project 5 – Method: M-2 SME Models

	Prior				Posterior				Aggregate
Node Variable	SME-3	SME-5	SME-7	SME-8	SME-3	SME-5	SME-7	SME-8	
Mgt. commitment	0.500	0.600	0.400	0.600	1	1	1	1	1.000
Quality system maturity	0.450	0.520	0.460	0.640	1	1	1	1	1.000
Project time	0.418	0.629	0.364	0.348	1	1	1	1	1.000
Failure mode complexity	0.560	0.460	0.240	0.320	0	0	0	0	0.000
Technical expertise	0.549	0.670	0.430	0.277	0	0	0	0	0.000
Resource availability	0.511	0.585	0.362	0.535	0.750	0.727	0.667	0.778	0.742
Design complexity	0.400	0.300	0.200	0.300	1	1	1	1	1.000
Test facilities	0.541	0.753	0.365	0.344	1	1	1	1	1.000
Fix effectiveness 10%	0.204	0.196	0.413	0.188	0.075	0.250	0.233	0.522	0.303
Fix effectiveness 30%	0.241	0.328	0.191	0.268	0.360	0.336	0.183	0.200	0.281
Fix effectiveness 50%	0.256	0.318	0.133	0.201	0.263	0.264	0.283	0.139	0.225
Fix effectiveness 70%	0.164	0.139	0.102	0.175	0.213	0.136	0.233	0.100	0.149
Fix effectiveness 90%	0.136	0.019	0.160	0.169	0.090	0.014	0.067	0.039	0.042
Likelihood	-	-	-		4.04E-03	7.90E-03	2.11E-03	6.78E-03	
Prob (M_j)	0.25	0.25	0.25	0.25	0.19	0.38	0.10	0.33	
	Mean	Var.	Alpha	Beta					
	0.369	0.055	1.186	2.025					

Appendix K

Automotive Fix Effectiveness (*di*) Project 1-5: M-3 SME Models

Table K1

Automotive Fix Effectiveness (di) Projection – Method: M-3 SME Models

	Prior	Posterior				
Node Variable	M-3	Project 1	Project 2	Project 3	Project 4	Project 5
mgt commitment	0.525	1	1	1	1	1
quality system maturity	0.531	1	1	0	1	1
project time	0.444	0	1	1	1	1
failure mode complexity	0.365	0	0	1	1	0
technical expertise	0.491	1	1	0.775	1	0
resource availability	0.481	1	1	1	1	1
design complexity	0.300	1	1	0.473	1	1
test facilities	0.518	1	0.650	1	1	1
fix effectiveness 10%	0.299	0.095	0.08	0.04	0.02	0.15
fix effectiveness 30%	0.245	0.130	0.16	0.12	0.10	0.22
fix effectiveness 50%	0.209	0.313	0.22	0.21	0.20	0.27
fix effectiveness 70%	0.123	0.288	0.17	0.17	0.15	0.21
fix effectiveness 90%	0.124	0.175	0.37	0.46	0.54	0.15
Likelihood	-	-	-	-	-	-
Prob (Mj)	-	-	-	-	-	-
Mean Projection		0.564	0.616	0.676	0.719	0.502
Variance		0.056	0.072	0.060	0.050	0.065
Alpha		1.917	1.395	1.779	2.171	1.419
Beta		1.485	0.870	0.854	0.847	1.410

Appendix L

Automotive Fix Effectiveness (*di*) Project 1-5: M-4 SME Models

Table L1

Automotive Fix Effectiveness (di) Project 1 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	0	0	0	0	0	0	0	0	0	.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	0	0	0	0	0	0	0	0	0	.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table L1 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.100	.05	.100	.040	.200	.200	.050	.050	.095	.089
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.300	.10	.200	.120	.450	.250	.100	.050	.130	.169
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.550	.20	.250	.200	.270	.350	.250	.150	.313	.263
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.013	.50	.250	.320	.060	.200	.500	.150	.288	.311
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.037	.15	.200	.320	.020	.000	.100	.600	.175	.167
Likelihood	-		-		-		-			3.43 E-03	1.47 E-02	1.19 E-05	6.72 E-03	3.50 E-03	5.53 E-03	1.01 E-03	2.61 E-03	3.61 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.08	.36	.00	.16	.09	.13	.02	.06	.09	
																Mean	Var.	Alpha	Beta
																0.559	0.057	1.872	1.476

Note: *Aggregate

Table L2

Automotive Fix Effectiveness (di) Project 2 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	0	0	0	0	0	0	0	0	0	.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	.700	.760	.900	.770	.600	.700	.800	.500	.650	.713

Table L2 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.115	.064	.110	.065	.032	.095	.020	.025	.085	.085
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.315	.124	.210	.138	.068	.215	.060	.075	.159	.206
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.520	.224	.258	.189	.250	.455	.110	.125	.220	.348
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.023	.464	.240	.304	.240	.200	.150	.200	.165	.197
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.027	.126	.182	.303	.410	.035	.660	.575	.371	.165
Likelihood	-		-		-		-			4.37 E-02	1.80 E-02	3.10 E-06	2.60 E-02	4.65 E-03	1.25 E-02	1.58 E-03	4.05 E-03	5.56 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.38	.16	.00	.22	.04	.11	.01	.03	.05	
																Mean	Var.	Alpha	Beta
																0.530	0.055	1.867	1.653

Note: *Aggregate

Table L3

Automotive Fix Effectiveness (di) Project 3 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	0	0	0	0	0	0	0	0	0	.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	1	1	1	1	1	1	1	1	1	1.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	.417	.900	.990	.786	.800	.850	.950	.500	.775	.675
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	.885	.487	.526	.468	.571	.391	.333	.563	.473	.610
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table L3 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.040	.038	.012	.059	.010	.080	.014	.250	.040	.053
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.150	.155	.053	.162	.160	.223	.043	.125	.123	.165
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.530	.535	.250	.231	.170	.463	.066	.125	.212	.395
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.250	.240	.596	.280	.200	.193	.111	.100	.167	.225
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.030	.032	.090	.269	.460	.043	.765	.400	.457	.161
Likelihood	-		-		-		-			5.89 E-02	1.21 E-02	2.12 E-05	2.64 E-02	1.78 E-02	3.67 E-02	2.49 E-03	4.43 E-03	9.91 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.35	.07	.00	.16	.11	.22	.01	.03	.06	
																Mean	Var.	Alpha	Beta
																0.555	0.047	2.366	1.896

Note: *Aggregate

Table L4

Automotive Fix Effectiveness (di) Project 4 – Method: M-4 BBN Aggregate Structure

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	1	1	1	1	1	1	1	1	1	1.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	1	1	1	1	1	1	1	1	1	1.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	1	1	1	1	1	1	1	1	1	1.000
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table L4 (*continued*).

Node Variable	Prior									Posterior									Ag.
	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effective-ness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.090	.05	.010	.000	.000	.050	.013	.000	.016	.071
Fix effective-ness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.100	.10	.050	.100	.100	.200	.038	.050	.097	.111
Fix effective-ness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.600	.20	.250	.200	.150	.500	.050	.100	.200	.509
Fix effective-ness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.200	.50	.600	.350	.200	.200	.100	.100	.150	.219
Fix effective-ness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.010	.15	.090	.350	.550	.050	.800	.750	.538	.089
Likeli-hood	-		-		-		-			4.83 E-02	2.74 E-03	9.24 E-08	4.63 E-03	3.20 E-03	7.99 E-03	1.83 E-04	6.55 E-04	1.65 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.70	.04	.00	.07	.05	.12	.00	.01	.02	
																Mean	Var.	Alpha	Beta
																0.529	0.038	2.941	2.617

Note: *Aggregate

Table L5

Automotive Fix Effectiveness (di) Project 5 – Method: M-4 BBN Aggregate Structure

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Mgt. commitment	.695	.600	.150	.750	.500	.600	.400	.600	.525	1	1	1	1	1	1	1	1	1	1.000
Quality system maturity	.604	.480	.084	.500	.450	.520	.460	.640	.531	1	1	1	1	1	1	1	1	1	1.000
Project time	.900	.428	.054	.674	.418	.629	.364	.348	.444	1	1	1	1	1	1	1	1	1	1.000
Failure mode complexity	.610	.200	.119	.250	.560	.460	.240	.320	.365	0	0	0	0	0	0	0	0	0	.000
Technical expertise	.739	.820	.310	.700	.549	.670	.430	.277	.491	0	0	0	0	0	0	0	0	0	.000
Resource availability	.552	.743	.108	.785	.511	.585	.362	.535	.481	.750	.950	.990	.723	.750	.727	.667	.778	1	.753
Design complexity	.726	.300	.154	.250	.400	.300	.200	.300	.300	1	1	1	1	1	1	1	1	1	1.000
Test facilities	.707	.609	.220	.533	.541	.753	.365	.344	.518	1	1	1	1	1	1	1	1	1	1.000

Table L5 (*continued*).

	Prior									Posterior									Ag.
Node Variable	M-1				M-2				M-3	M-1				M-2				M-3	
	SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		SME 3	SME 5	SME 7	SME 8	SME 3	SME 5	SME 7	SME 8		
Fix effectiveness 10%	.110	.139	.683	.188	.204	.196	.413	.188	.299	.100	.20	.503	.136	.075	.250	.233	.522	.146	.128
Fix effectiveness 30%	.258	.267	.202	.268	.241	.328	.191	.268	.245	.300	.50	.358	.194	.360	.336	.183	.200	.220	.305
Fix effectiveness 50%	.472	.345	.066	.201	.256	.318	.133	.201	.209	.550	.20	.100	.169	.263	.264	.283	.139	.268	.453
Fix effectiveness 70%	.138	.205	.035	.175	.164	.139	.102	.175	.123	.013	.10	.030	.250	.213	.136	.233	.100	.213	.062
Fix effectiveness 90%	.023	.043	.014	.169	.136	.019	.160	.169	.124	.037	.00	.010	.250	.090	.014	.067	.039	.154	.052
Likelihood	-		-		-		-			4.83 E-02	2.74 E-03	9.24 E-08	4.63 E-03	3.20 E-03	7.99 E-03	1.83 E-04	6.55 E-04	1.65 E-03	
Prob (M_j)	.111	.111	.111	.111	.111	.111	.111	.111	.111	.70	.04	.00	.07	.05	.12	.00	.01	.02	
																Mean	Var.	Alpha	Beta
																0.421	0.037	2.330	3.203

Note: *Aggregate

Appendix M

Auto BBN Analysis

Automotive BBN Analysis

The same analysis methodology used for the HVAC case study will be repeated for the automotive case study. Power and sample size calculations indicate 541 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table M1). Analysis of variances indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure M1).

Table M1

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	541	0.9000	0.9003	0.05

Note: Sigma = 0.218161, Alpha = 0.05, Number of Levels = 4

One way ANOVA (Table M2) indicate one must reject Ho of no difference in the models with Tukey pairwise comparisons (Table M3) indicating M3 FEF projection mean is statistically equal to M4. One sample T (Table M4) looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference. Overall summary is the models did not perform well in predicting actual FEF on this project. One would expect these results though given most node evidence were very positive however actual FEF was low at 26.5%.

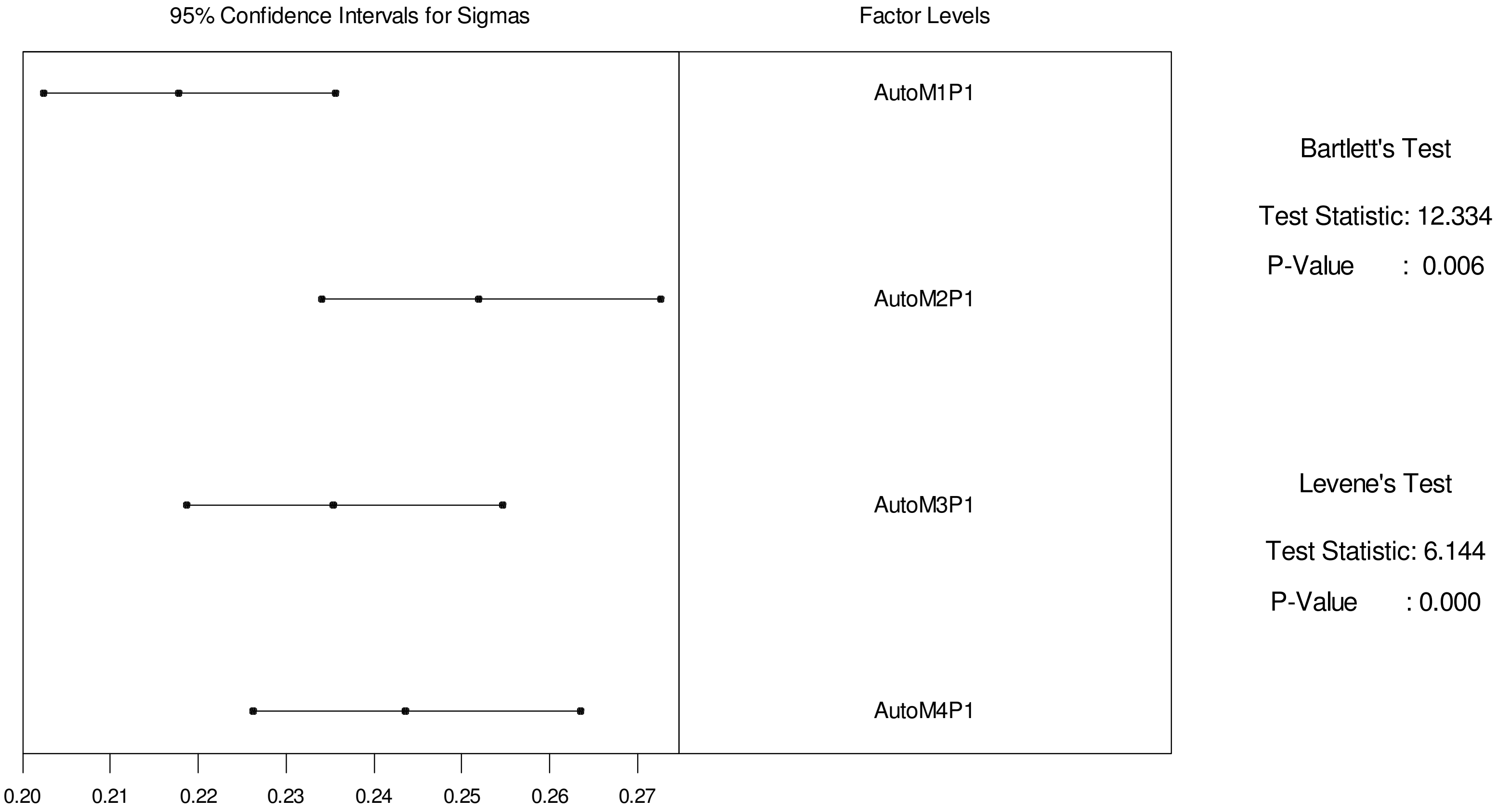


Figure M1. Test for equal variances for automotive project 1

Table M2

One-way ANOVA: Auto Project-1: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP1St	3	4.1682	1.3894	24.63	0.000
Error	2160	121.8382	0.0564		
Total	2163	126.0064			

Individual 95% CIs for Mean Based on Pooled StDev

```
Level        N     Mean     StDev   --------+---------+---------+--------
AutoM1P1    541   0.6048   0.2178                            (---*---)
AutoM2P1    541   0.4829   0.2520     (---*---)
AutoM3P1    541   0.5496   0.2354                (---*---)
AutoM4P1    541   0.5640   0.2435                  (---*---)
                                    --------+---------+---------+--------
Pooled StDev =    0.2375                  0.500     0.500     0.600
```

Table M3

Tukey's Pairwise Comparisons

	AutoM1P1	AutoM2P1	AutoM3P1
AutoM2P1	0.0849		
	0.1590		
AutoM3P1	0.0182	-0.1038	
	0.0923	-0.0296	
AutoM4P1	0.0038	-0.1181	-0.0514
	0.0779	-0.0440	0.0227

Table M4

One-Sample T: Auto Project-1: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
AutoM1P1	541	0.60484	0.21780	0.00936
AutoM2P1	541	0.4829	0.2520	0.0108
AutoM3P1	541	0.5496	0.2354	0.0101
AutoM4P1	541	0.5640	0.2435	0.0105

Variable	95.0% CI	T	P
AutoM1P1	(0.58645, 0.62324)	36.29	0.000
AutoM2P1	(0.4616, 0.5042)	20.12	0.000
AutoM3P1	(0.5297, 0.5695)	28.12	0.000
AutoM4P1	(0.5434, 0.5845)	28.55	0.000

Note: Test of mu = 0.265 vs mu not = 0.265

Results for: Auto Project-2

Power and sample size calculations indicate 591 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table M5). Analysis of variances indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure M2).

One way ANOVA (Table M6) analysis indicates one must reject Ho of no difference in the models with Tukey pairwise comparisons (Table M7) indicating M-1 FEF projection mean is statistically equal to M-4 and M-2 is statistically equal to M-3. One sample T (Table M8) looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference for all models.

Table M5

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	591	0.9000	0.9003	0.05

Note: Sigma = 0.228035, Alpha = 0.05, Number of Levels = 4

Table M6

One-way ANOVA: Auto Project-2: Model to Model Evaluation

Analysis of Variance					
Source	DF	SS	MS	F	P
AutoP2St	3	2.1345	0.7115	11.86	0.000
Error	2360	141.5508	0.0600		
Total	2363	143.6852			

Individual 95% CIs for Mean Based on Pooled StDev

Level	N	Mean	StDev	----+---------+---------+---------+--
M1P2	591	0.5200	0.2277	(----*----)
M2P2	591	0.5745	0.2411	(----*----)
M3P2	591	0.5895	0.2715	(-----*-----)
M4P2	591	0.5259	0.2371	(-----*-----)
Pooled StDev =		0.2449		----+---------+---------+---------+--
				0.510 0.540 0.570 0.600

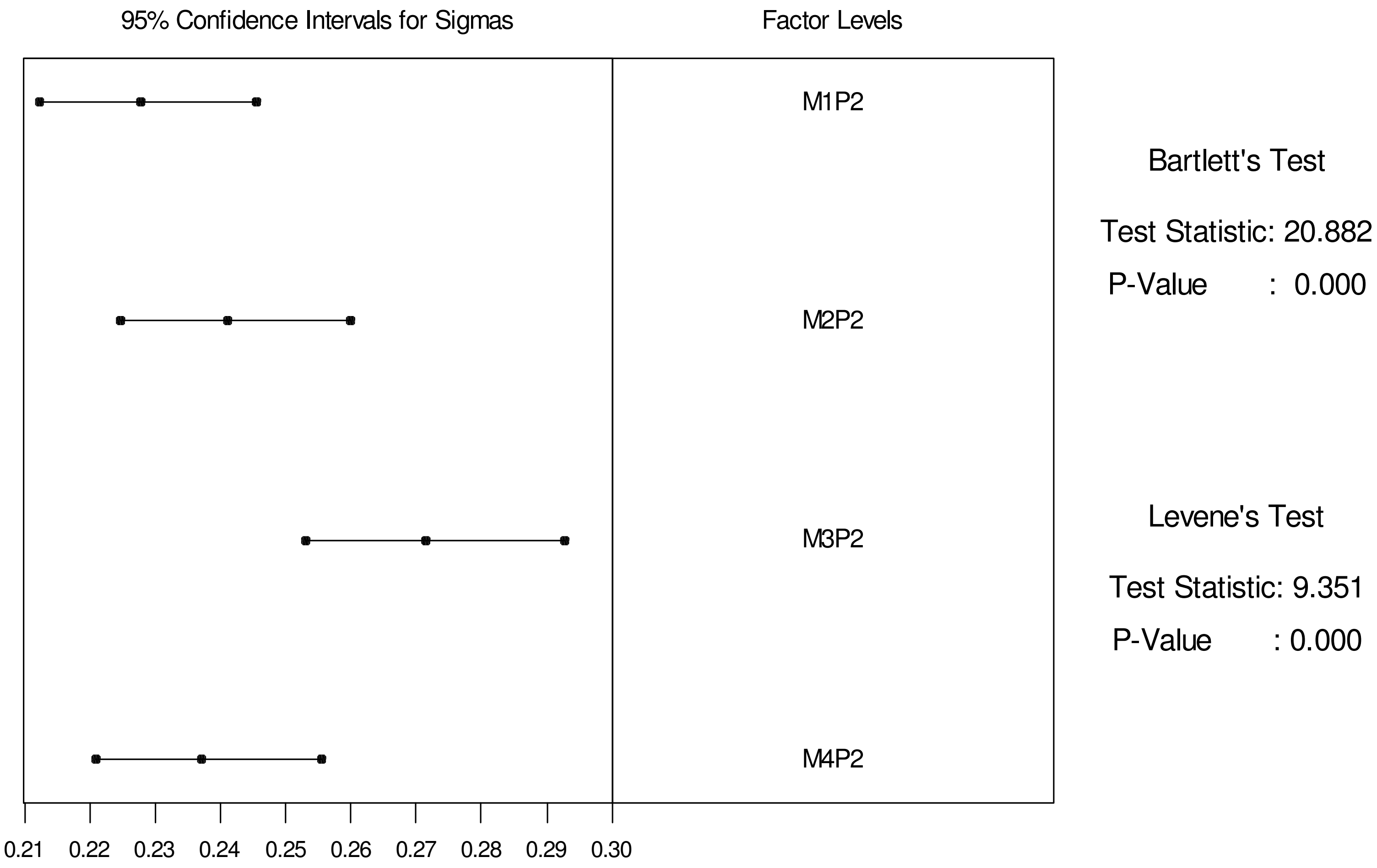


Figure M2. Test for equal variances for auto project 2

Table M7

Tukey's Pairwise Comparisons

	M1P2	M2P2	M3P2
M2P2	-0.0911		
	-0.0179		
M3P2	-0.1060	-0.0515	
	-0.0328	0.0217	
M4P2	-0.0424	0.0121	0.0270
	0.0307	0.0852	0.1001

Table M8

One-Sample T: Auto Project-2: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P2	591	0.52004	0.22774	0.00937
M2P2	591	0.57454	0.24112	0.00992
M3P2	591	0.5895	0.2715	0.0112
M4P2	591	0.52592	0.23706	0.00975

Variable	95.0% CI	T	P
M1P2	(0.50164, 0.53844)	2.14	0.033
M2P2	(0.55506, 0.59402)	7.52	0.000
M3P2	(0.5675, 0.6114)	8.01	0.000
M4P2	(0.50676, 0.54507)	2.66	0.008

Note: Test of mu = 0.5 vs mu not = 0.5

Results for: Auto Project-3

Power and sample size calculations indicate 421 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table M9). Analysis of variances indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure M3).

One way ANOVA (Table M10) analysis indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table M11) indicating M-1 FEF projection mean is statistically equal to M-2 and M-4. One sample T (Table M12) looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference for all models.

Table M9

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	421	0.9000	0.9004	0.05

Note: Sigma = 0.192354, Alpha = 0.05, Number of Levels = 4

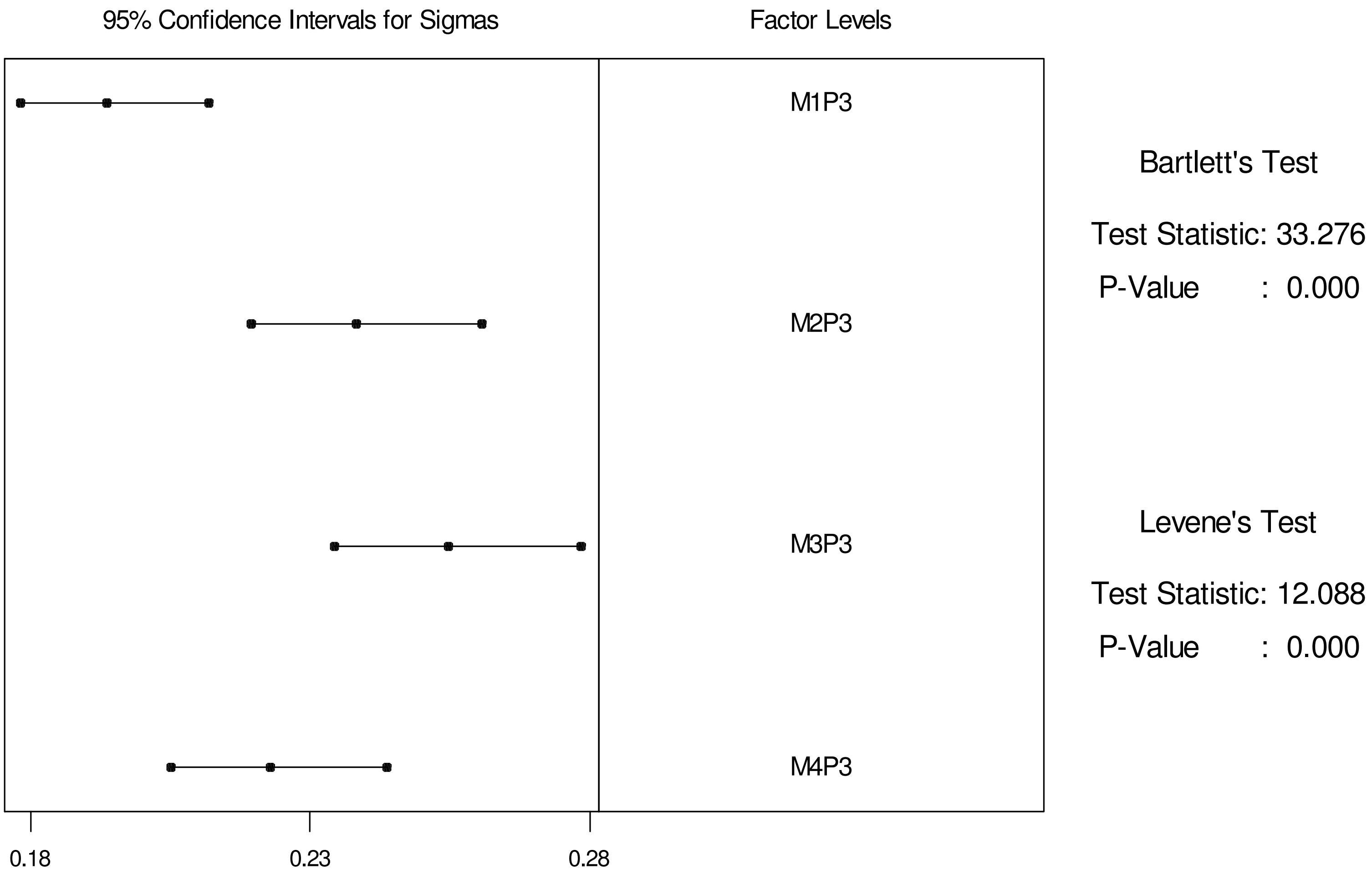


Figure M3. Test for equal variances for auto project 3

Table M10

One-way ANOVA: Auto Project-3: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP3St	3	4.5634	1.5211	29.14	0.000
Error	1680	87.7094	0.0522		
Total	1683	92.2728			

Individual 95% CIs for Mean Based on Pooled StDev

```
Level      N    Mean   StDev   --------+---------+---------+--------
M1P3     421  0.5366  0.1937     (---*----)
M2P3     421  0.5755  0.2384             (---*----)
M3P3     421  0.6733  0.2547                               (----*---)
M4P3     421  0.5600  0.2228         (---*---)
Pooled StDev = 0.2285          --------+---------+---------+--------
                                     0.550     0.600     0.650
```

Table M11

Tukey's Pairwise Comparisons

	M1P3	M2P3	M3P3
M2P3	-0.0792		
	0.0016		
M3P3	-0.1771	-0.1382	
	-0.0962	0.0574	
M4P3	-0.0638	-0.0250	0.0728
	0.0170	0.0558	0.1537

Table M12

One-Sample T: Auto Project-3: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P3	421	0.53665	0.19371	0.00944
M2P3	421	0.5755	0.2384	0.0116
M3P3	421	0.6733	0.2547	0.0124
M4P3	421	0.5600	0.2228	0.0109

Variable	95.0% CI	T	P
M1P3	(0.51809, 0.55520)	-48.87	0.000
M2P3	(0.5526, 0.5983)	-36.37	0.000
M3P3	(0.6489, 0.6977)	-26.16	0.000
M4P3	(0.5387, 0.5814)	-40.34	0.000

Note: Test of mu = 0.998 vs mu not = 0.998

Results for: Auto Project-4

Power and sample size calculations indicate 376 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table M13). Analysis of variances indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure M4).

One way ANOVA analysis (Table M14) indicates one must reject Ho of no difference in the models with Tukey pairwise comparisons (Table M15) indicating M-1 FEF projection mean is statistically equal to M-4. One sample T (Table M16) looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference for all models.

Table M13

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	376	0.9000	0.9008	0.05

Note: Sigma = 0.181659, Alpha = 0.05, Number of Levels =

Table M14

One-way ANOVA: Auto Project-4: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP4St	3	8.8130	2.9377	66.77	0.000
Error	1500	65.9938	0.0440		
Total	1503	74.8068			

Individual 95% CIs for Mean Based on Pooled StDev

```
Level       N    Mean    StDev   -+---------+---------+---------+-----
M1P4      376  0.5091   0.1737   (--*--)
M2P4      376  0.5841   0.2282             (--*--)
M3P4      376  0.7010   0.2370                                 (--*--)
M4P4      376  0.5186   0.1938    (--*--)
Pooled StDev = 0.2098            -+---------+---------+---------+-----
                                 0.490    0.560     0.630     0.700
```

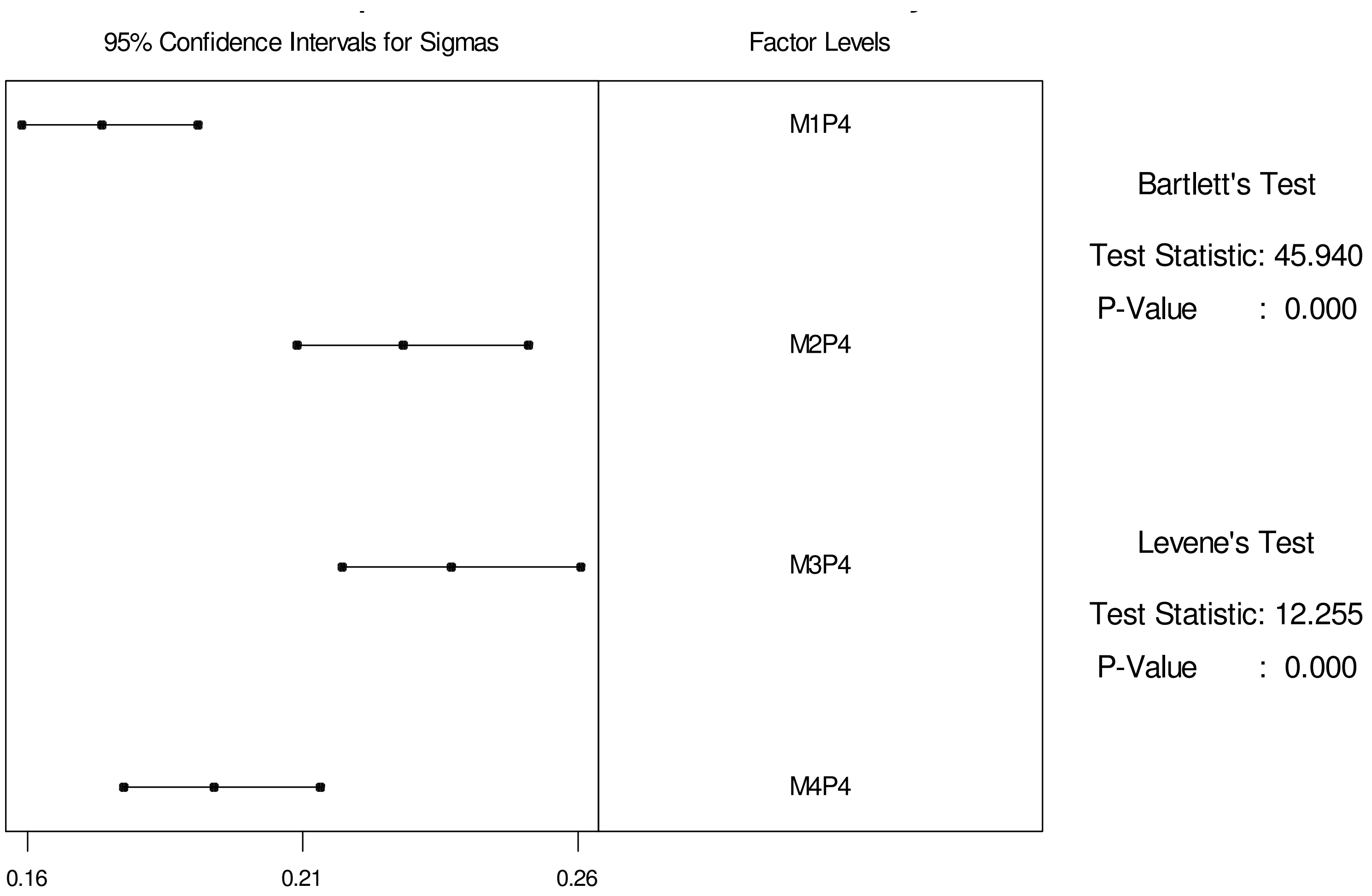


Figure M4. Test for equal variances for auto project 4

Table M15

Tukey's Pairwise Comparisons

	M1P4	M2P4	M3P4
M2P4	-0.1143		
	0.0358		
M3P4	-0.2311	-0.1561	
	-0.1526	-0.0776	
M4P4	-0.0487	0.0263	0.1432
	0.0298	0.1049	0.2217

Table M16

One-Sample T: Auto Project-4: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P4	376	0.50912	0.17365	0.00896
M2P4	376	0.5841	0.2282	0.0118
M3P4	376	0.7010	0.2370	0.0122
M4P4	376	0.51856	0.19379	0.00999

Variable	95.0% CI	T	P
M1P4	(0.49151, 0.52673)	-38.40	0.000
M2P4	(0.5610, 0.6073)	-22.84	0.000
M3P4	(0.6769, 0.7250)	-12.44	0.000
M4P4	(0.49891, 0.53821)	-33.46	0.000

Note: Test of mu = 0.853 vs mu not = 0.853

Results for: Auto Project-5

Power and sample size calculations indicate 512 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table M17). Analysis of variances indicates one must reject the null of no difference among variances with M-4 representing the lowest variance among the models (Figure M5).

One way ANOVA analysis (Table M18) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table M19) indicating no statistical significance among the models. One sample T (Table M20)looks at each model FEF projection mean to actual FEF obtained from field data and indicates one must reject Ho of no difference for all models.

Table M17

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	512	0.9000	0.9005	0.05

Note: Sigma = 0.212132, Alpha = 0.05, Number of Levels = 4

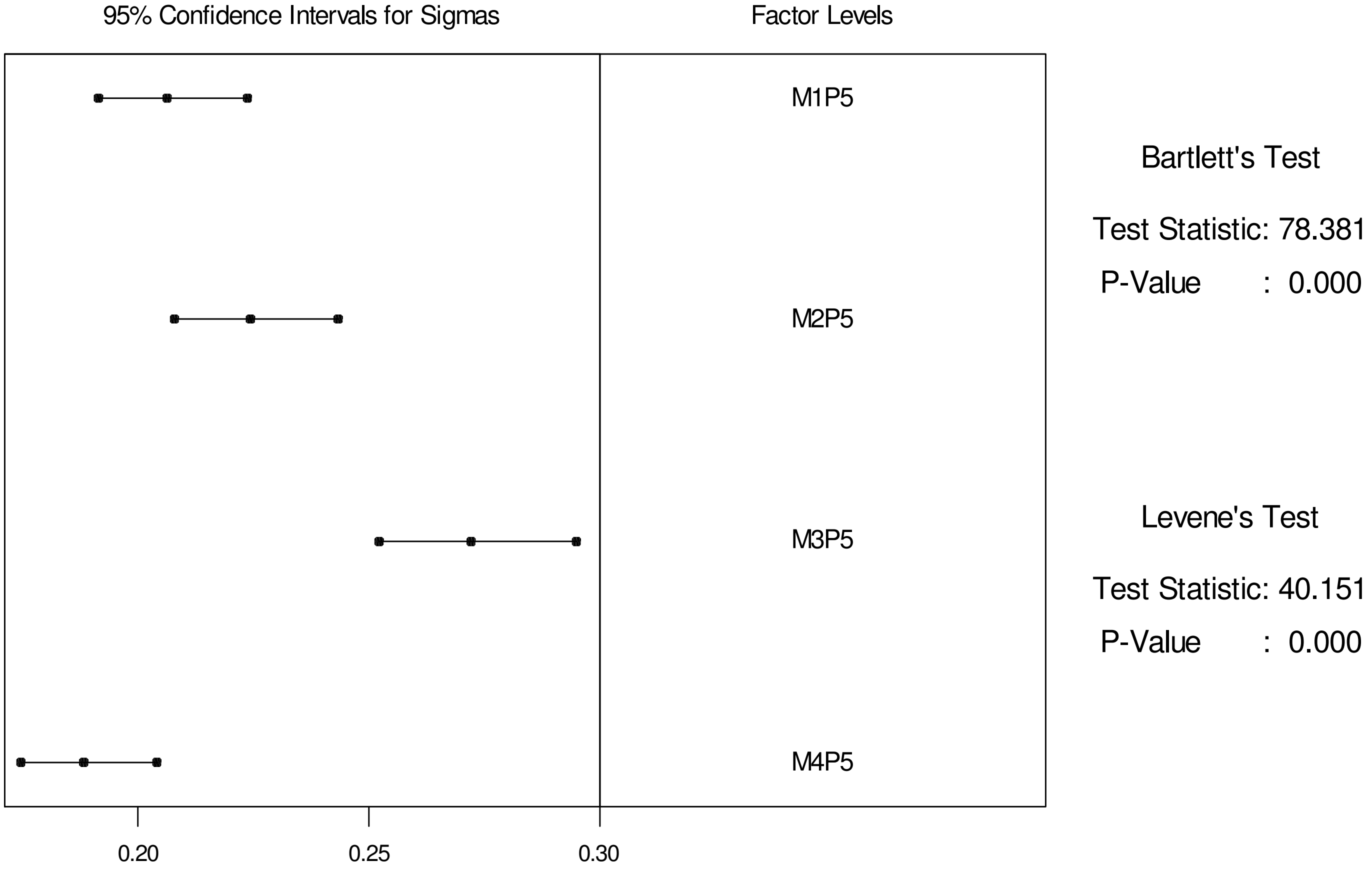


Figure M5. Test for equal variances for auto project 5

Table M18

One-way ANOVA: Auto Project-5: Model to Model Evaluation

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP5St	3	5.4536	1.8179	35.94	0.000
Error	2044	103.3967	0.0506		
Total	2047	108.8503			

Individual 95% CIs for Mean Based on Pooled StDev

Level	N	Mean	StDev	
				-+---------+---------+---------+-----
M1P5	512	0.4537	0.2062	(--*---)
M2P5	512	0.3674	0.2243	(--*---)
M3P5	512	0.5081	0.2722	(--*---)
M4P5	512	0.4154	0.1883	(---*---)
Pooled StDev =		0.2249		-+---------+---------+---------+-----
				0.350 0.400 0.450 0.500

Table M19

Tukey's Pairwise Comparisons

	M1P5	M2P5	M3P5
M2P5	0.0502		
	0.1224		
M3P5	-0.0905	-0.1768	
	-0.0184	-0.1047	
M4P5	0.0022	-0.0841	0.0567
	0.0744	-0.0119	0.1288

Table M20

One-Sample T: Auto Project-5: Model to Actual FEF Evaluation

Variable	N	Mean	St.Dev.	SE Mean
M1P5	512	0.45371	0.20619	0.00911
M2P5	512	0.36739	0.22428	0.00991
M3P5	512	0.5081	0.2722	0.0120
M4P5	512	0.41538	0.18829	0.00832

Variable	95.0% CI	T	P
M1P5	(0.43580, 0.47161)	-12.65	0.000
M2P5	(0.34791, 0.38686)	-20.34	0.000
M3P5	(0.4845, 0.5318)	-5.06	0.000
M4P5	(0.39904, 0.43173)	-18.46	0.000

Note: Test of mu = 0.569 vs mu not = 0.569

Appendix N

HVAC Simulated SME Project 1-5

Table N1

HVAC Fix Effectiveness (di) Projection – Project 1 – Method: M-3 Simulated SME (baseline)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	0	0	0	0.000
Quality system maturity	0.140	0.500	0.860	0	0	0	0.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	0.257	0.647	0.919	0.722
Technical expertise	0.159	0.499	0.844	1	1	1	1.000
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	0.135	0.515	0.918	0.633
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.095	0.003	0.012	0.009
Fix effectiveness 30%	0.228	0.198	0.042	0.200	0.048	0.040	0.051
Fix effectiveness 50%	0.123	0.256	0.107	0.500	0.137	0.075	0.129
Fix effectiveness 70%	0.045	0.255	0.330	0.190	0.255	0.099	0.202
Fix effectiveness 90%	0.002	0.228	0.513	0.015	0.557	0.775	0.610
Likelihood	-	-	-	7.79E-04	1.40E-02	7.21E-03	
Prob (Mj)	0.333	0.333	0.333	0.035	0.636	0.328	
				Mean	Variance	Alpha	Beta
				0.770	0.036	3.024	0.901

Table N2

HVAC Fix Effectiveness (di) Projection – Project 2 – Method: M-3 Simulated SME (baseline)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	0.603	0.746	0.836	0.774
Failure mode complexity	0.140	0.500	0.860	0	0	0	0.000
Technical expertise	0.159	0.499	0.844	0.367	0.583	0.767	0.642
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	0	0	0	0.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.102	0.006	0.009	0.009
Fix effectiveness 30%	0.228	0.198	0.042	0.422	0.083	0.038	0.074
Fix effectiveness 50%	0.123	0.256	0.107	0.367	0.197	0.090	0.164
Fix effectiveness 70%	0.045	0.255	0.330	0.103	0.288	0.320	0.295
Fix effectiveness 90%	0.002	0.228	0.513	0.007	0.426	0.543	0.458
Likelihood	-	-	-	4.26E-04	1.48E-02	7.80E-03	
Prob (Mj)	0.333	0.333	0.333	0.019	0.642	0.339	
				Mean	Variance	Alpha	Beta
				0.724	0.039	2.963	1.129

Table N3

HVAC Fix Effectiveness (di) Projection – Project 3 – Method: M-3 Simulated SME (baseline)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	0	0	0	0.000
Quality system maturity	0.140	0.500	0.860	0	0	0	0.000
Project time	0.161	0.475	0.848	0	0	0	0.000
Failure mode complexity	0.140	0.500	0.860	0.580	0.493	0.829	0.546
Technical expertise	0.159	0.499	0.844	0.169	0.374	0.622	0.251
Resource availability	0.204	0.494	0.850	0.100	0.250	0.500	0.160
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	0	0	0	0.000
Fix effectiveness 30%	0.603	0.064	0.008	0.709	0.178	0.008	0.497
Fix effectiveness 30%	0.228	0.198	0.042	0.166	0.309	0.106	0.223
Fix effectiveness 50%	0.123	0.256	0.107	0.085	0.261	0.319	0.155
Fix effectiveness 70%	0.045	0.255	0.330	0.038	0.175	0.370	0.093
Fix effectiveness 90%	0.002	0.228	0.513	0.001	0.077	0.197	0.032
Likelihood	-	-	-	5.24E-02	3.43E-02	2.92E-04	
Prob (Mj)	0.333	0.333	0.333	0.603	0.394	0.003	
				Mean	Variance	Alpha	Beta
				0.288	0.052	0.840	2.078

Table N4

HVAC Fix Effectiveness (di) Projection – Project 4 – Method: M-3 Simulated SME (baseline)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	0	0	0	0.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	0.682	0.687	0.942	0.899
Technical expertise	0.159	0.499	0.844	0.458	0.698	0.892	0.859
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.077	0.004	0.011	0.010
Fix effectiveness 30%	0.228	0.198	0.042	0.331	0.062	0.039	0.043
Fix effectiveness 50%	0.123	0.256	0.107	0.445	0.165	0.079	0.093
Fix effectiveness 70%	0.045	0.255	0.330	0.137	0.273	0.168	0.185
Fix effectiveness 90%	0.002	0.228	0.513	0.010	0.497	0.704	0.668
Likelihood	-	-	-	5.44E-05	1.40E-02	6.92E-02	
Prob (Mj)	0.333	0.333	0.333	0.001	0.168	0.831	
				Mean	Variance	Alpha	Beta
				0.792	0.032	3.234	0.850

Table N5

HVAC Fix Effectiveness (di) Projection – Project 5 – Method: M-3 Simulated SME (baseline)

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	0	0	0	0.000
Failure mode complexity	0.140	0.500	0.860	1	1	1	1.000
Technical expertise	0.159	0.499	0.844	0.367	0.583	0.767	0.617
Resource availability	0.204	0.494	0.850	0	0	0	0.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	0	0	0	0.000
Fix effectiveness 30%	0.603	0.064	0.008	0.700	0.163	0.009	0.139
Fix effectiveness 30%	0.228	0.198	0.042	0.174	0.319	0.123	0.280
Fix effectiveness 50%	0.123	0.256	0.107	0.087	0.274	0.370	0.290
Fix effectiveness 70%	0.045	0.255	0.330	0.037	0.174	0.317	0.200
Fix effectiveness 90%	0.002	0.228	0.513	0.002	0.070	0.181	0.091
Likelihood	-	-	-	1.06E-04	8.67E-03	2.09E-03	
Prob (Mj)	0.333	0.333	0.333	0.010	0.798	0.192	
				Mean	Variance	Alpha	Beta
				0.465	0.055	1.651	1.900

Appendix O

Automotive Simulated SME Project 1-5

Table O1

Automotive Fix Effectiveness (di) Projection – Project 1 – Method: Simulated SME

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	0	0	0	0.000
Failure mode complexity	0.140	0.500	0.860	0	0	0	0.000
Technical expertise	0.159	0.499	0.844	1	1	1	1.000
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.062	0.006	0.003	0.004
Fix effectiveness 30%	0.228	0.198	0.042	0.441	0.094	0.033	0.051
Fix effectiveness 50%	0.123	0.256	0.107	0.398	0.230	0.110	0.146
Fix effectiveness 70%	0.045	0.255	0.330	0.093	0.314	0.740	0.612
Fix effectiveness 90%	0.002	0.228	0.513	0.005	0.356	0.114	0.187
Likelihood	-	-	-	6.32E-06	3.87E-03	9.03E-03	
Prob (Mj)	0.333	0.333	0.333	0.000	0.300	0.700	
				Mean	Variance	Alpha	Beta
				0.685	0.023	5.828	2.677

Table O2

Automotive Fix Effectiveness (di) Projection – Project 2 – Method: Simulated SME

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	0	0	0	0.000
Technical expertise	0.159	0.499	0.844	1	1	1	1.000
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	0.367	0.583	0.767	0.747
Fix effectiveness 30%	0.603	0.064	0.008	0.074	0.004	0.010	0.009
Fix effectiveness 30%	0.228	0.198	0.042	0.353	0.067	0.038	0.041
Fix effectiveness 50%	0.123	0.256	0.107	0.436	0.176	0.083	0.093
Fix effectiveness 70%	0.045	0.255	0.330	0.129	0.280	0.248	0.252
Fix effectiveness 90%	0.002	0.228	0.513	0.009	0.474	0.621	0.605
Likelihood	-	-	-	6.29E-06	7.31E-03	5.93E-02	
Prob (Mj)	0.333	0.333	0.333	0.000	0.110	0.890	
				Mean	Variance	Alpha	Beta
				0.780	0.031	3.478	0.980

Table O3

Automotive Fix Effectiveness (di) Projection – Project 3 – Method: Simulated SME

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	0	0	0	0.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	1	1	1	1.000
Technical expertise	0.159	0.499	0.844	0.500	0.750	0.900	0.873
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	0.643	0.550	0.942	0.873
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.07851	0.003	0.011	0.010
Fix effectiveness 30%	0.228	0.198	0.042	0.32069	0.060	0.039	0.043
Fix effectiveness 50%	0.123	0.256	0.107	0.44919	0.160	0.078	0.093
Fix effectiveness 70%	0.045	0.255	0.330	0.14151	0.270	0.163	0.181
Fix effectiveness 90%	0.002	0.228	0.513	0.01010	0.507	0.709	0.673
Likelihood	-	-	-	7.61E-05	1.40E-02	6.61E-02	
Prob (Mj)	0.333	0.333	0.333	0.001	0.175	0.824	
				Mean	Variance	Alpha	Beta
				0.793	0.032	3.226	0.842

Table O4

Automotive Fix Effectiveness (di) Projection – Project 4 – Method: Simulated SME

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	1	1	1	1.000
Technical expertise	0.159	0.499	0.844	1	1	1	1.000
Resource availability	0.204	0.494	0.850	1	1	1	1.000
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.095	0.003	0.012	0.012
Fix effectiveness 30%	0.228	0.198	0.042	0.200	0.048	0.040	0.040
Fix effectiveness 50%	0.123	0.256	0.107	0.500	0.137	0.075	0.075
Fix effectiveness 70%	0.045	0.255	0.330	0.190	0.255	0.099	0.100
Fix effectiveness 90%	0.002	0.228	0.513	0.015	0.557	0.775	0.773
Likelihood	-	-	-	1.97E-07	3.50E-03	3.09E-01	
Prob (Mj)	0.333	0.333	0.333	0.000	0.011	0.989	
				Mean	Variance	Alpha	Beta
				0.816	0.031	3.126	0.703

Table O5

Automotive Fix Effectiveness (di) Projection – Project 5 – Method: Simulated SME

	Prior			Posterior			Aggregate
Node Variable	Pessimist	Normalist	Optimist	Pessimist	Normalist	Optimist	
Mgt. commitment	0.100	0.500	0.900	1	1	1	1.000
Quality system maturity	0.140	0.500	0.860	1	1	1	1.000
Project time	0.161	0.475	0.848	1	1	1	1.000
Failure mode complexity	0.140	0.500	0.860	0	0	0	0.000
Technical expertise	0.159	0.499	0.844	0	0	0	0.000
Resource availability	0.204	0.494	0.850	0.452	0.682	0.851	0.784
Design complexity	0.100	0.500	0.900	1	1	1	1.000
Test facilities	0.193	0.478	0.848	1	1	1	1.000
Fix effectiveness 30%	0.603	0.064	0.008	0.133	0.010	0.003	0.006
Fix effectiveness 30%	0.228	0.198	0.042	0.514	0.122	0.036	0.070
Fix effectiveness 50%	0.123	0.256	0.107	0.281	0.252	0.121	0.173
Fix effectiveness 70%	0.045	0.255	0.330	0.068	0.313	0.705	0.550
Fix effectiveness 90%	0.002	0.228	0.513	0.004	0.303	0.135	0.201
Likelihood	-	-	-	3.14E-05	7.11E-03	1.09E-02	
Prob (Mj)	0.333	0.333	0.333	0.002	0.393	0.605	
				Mean	Variance	Alpha	Beta
				0.674	0.028	4.688	2.268

Appendix P

HVAC/Automotive S-SME Statistical Analysis

Results for: HVAC Project-1: M1-M4 & S-SME

Power and sample size calculations indicate 531 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P1). Analysis of variances indicates one must reject the null of no difference among variances with M-1 representing the lowest variance among the models (Figure P1).

Table P1

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	531	0.9000	0.9001	0.05

Note: Sigma = 0.207364, Alpha = 0.05, Number of Levels = 5

One way ANOVA analysis (Table P2) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P3) indicating statistical significance between M-1 and M-4, M-2 and S-SME, as well as M-4 and S-SME. One sample T (Table P4) looks at each model FEF projection mean to actual FEF obtained from field data and indicates statistical significance of M-4 and S-SME to actual FEF.

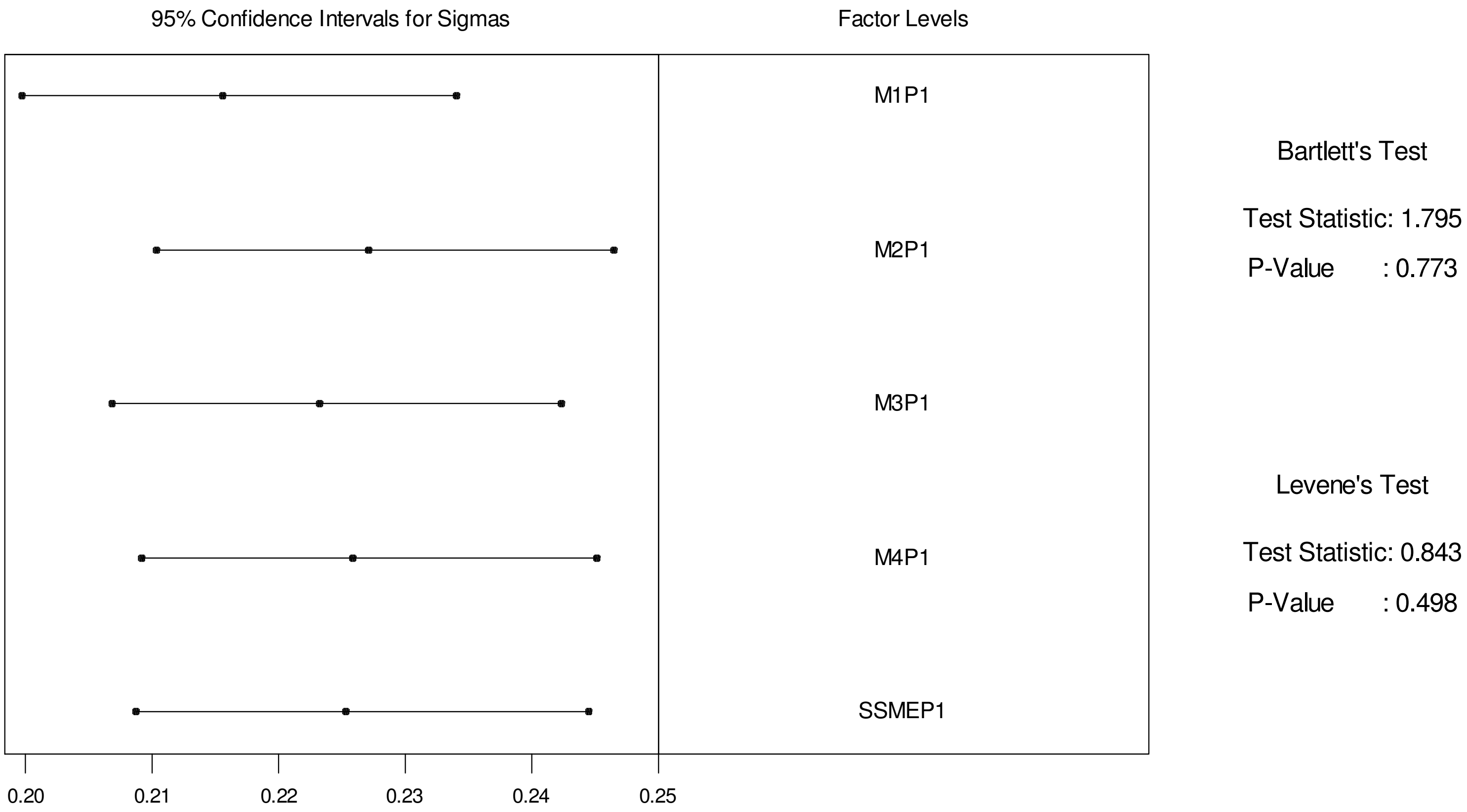


Figure P1. Test for equal variances for HVAC project 1: M1-M4 & S-SME

Table P2

One-way ANOVA: HVAC Project-1: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
HVACP1SM	4	9.7246	2.4311	48.68	0.000
Error	2650	132.3372	0.0499		
Total	2654	142.0618			

Individual 95% CIs for Mean Based on Pooled StDev

Level	N	Mean	StDev	---+---------+---------+---------+---
M1P1	531	0.5291	0.2156	(--*--)
M2P1	531	0.6096	0.2271	(---*--)
M3P1	531	0.7082	0.2233	(--*--)
M4P1	531	0.5630	0.2259	(---*---)
SSMEP1	531	0.5892	0.2253	(---*---)
Pooled StDev =		0.2235		---+---------+---------+---------+---
				0.540 0.600 0.660 0.720

Table P3

Tukey's Pairwise Comparisons

	M1P1	M2P1	M3P1	M4P1
M2P1	-0.1179			
	-0.0430			
M3P1	-0.2165	-0.1361		
	-0.1417	-0.0612		
M4P1	-0.0713	0.0091	0.1078	
	0.0035	0.0840	0.1827	
SSMEP1	-0.0975	-0.0171	0.0816	-0.0636
	-0.0227	0.0578	0.1564	0.0112

Table P4

One-Sample T: HVAC Project-1: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P1	531	0.52913	0.21562	0.00936
M2P1	531	0.60960	0.22712	0.00986
M3P1	531	0.70824	0.22326	0.00969
M4P1	531	0.56302	0.22586	0.00980
SSMEP1	531	0.58923	0.22530	0.00978

Table P4 (*continued*).

Variable	95.0% CI	T	P
M1P1	(0.51075, 0.54751)	-5.44	0.000
M2P1	(0.59024, 0.62896)	3.00	0.003
M3P1	(0.68921, 0.72727)	13.24	0.000
M4P1	(0.54377, 0.58228)	-1.73	0.084
SSMEP1	(0.57003, 0.60844)	0.94	0.345

Note: Test of mu = 0.58 vs mu not = 0.58

Results for: HVAC Project-2: M1-M4 & S-SME

Power and sample size calculations indicate 741 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P5). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P2).

One way ANOVA analysis (Table P6) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P7) indicating statistical significance between M-1 and M-4, as well as between M-2 and M-4. One sample T (Table P8) measures statistical significance at each model FEF projection mean to actual FEF obtained from field data and indicates no statistical significance. However, the p-value of S-SME is 0.011, but mean projection of the simulated SME is 0719 with 95% lower confidence interval at 0.7044, just off the mark of 0.7. Experts predicted 90% FEF, so while the S-SME didn't exhibit statistical success; it was very successful in less error than the experts.

Table P5

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	741	0.9000	0.9003	0.05

Note: Sigma = 0.244949, Alpha = 0.05, Number of Levels = 5

Table P6

One-way ANOVA: HVAC Project-2: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
HVACP2SM	4	33.5757	8.3939	133.19	0.000
Error	3700	233.1795	0.0630		
Total	3704	266.7552			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level             N    Mean   StDev  ------+--------+--------+--------+---
M1P2            741  0.4593  0.2458      (-*-)
M2P2            741  0.5159  0.2721         (-*)
M3P2            741  0.6063  0.2627                   (-*)
M4P2            741  0.4840  0.2637       (*-)
SSMEP2          741  0.7192  0.2052                        (-*-)
Pooled StDev =       0.2510          ---+---------+---------+---------+---
                                      0.50      0.60      0.70      0.80
```

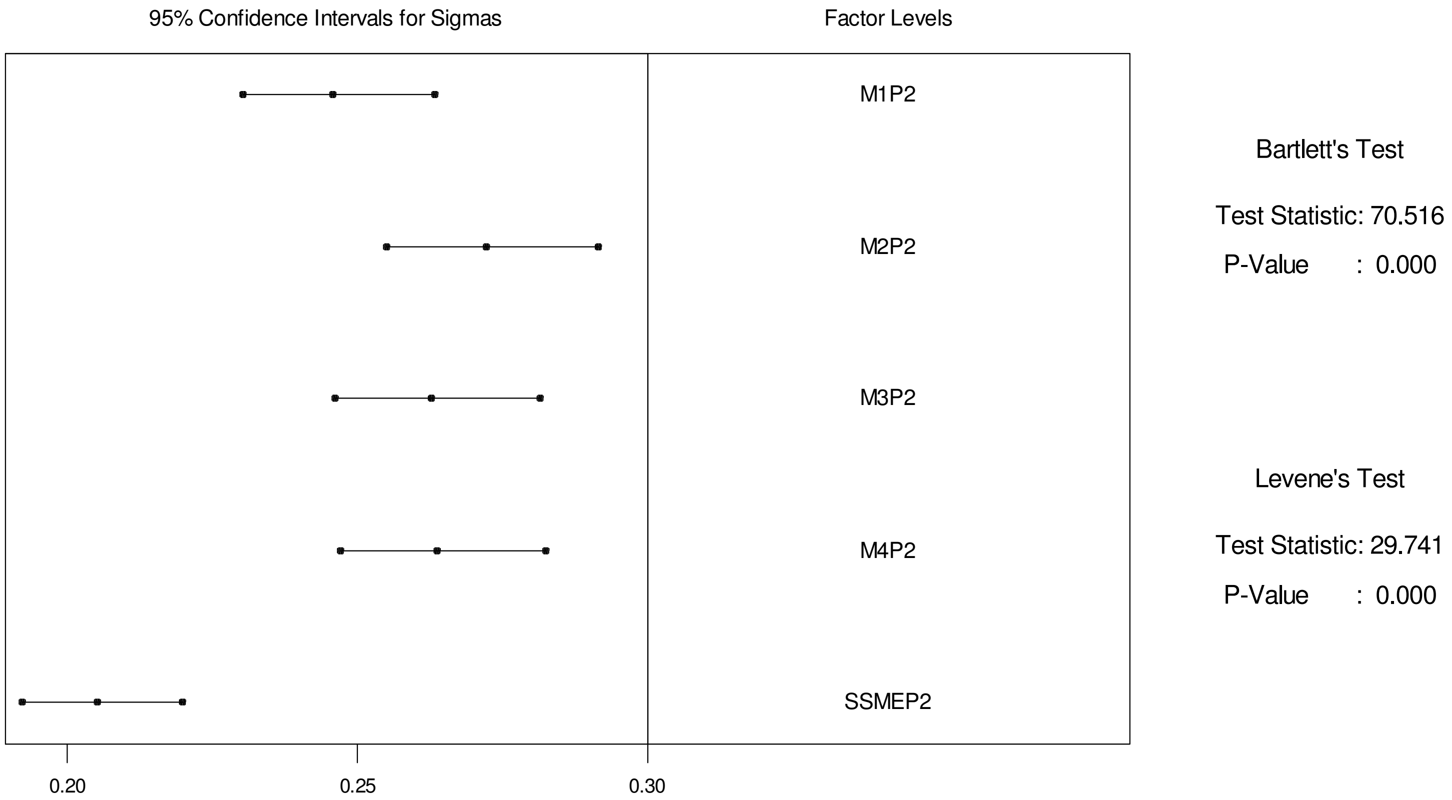


Figure P2. Test for equal variances for HVAC project 2: M1-M4 & S-SME

Table P7

Tukey's Pairwise Comparisons

	M1P2	M2P2	M3P2	M4P2
M2P2	-0.0922			
	-0.0210			
M3P2	-0.1826	-0.1260		
	-0.1114	-0.0548		
M4P2	-0.0603	0.0037	0.0868	
	0.0109	0.0675	0.1579	
SSMEP2	-0.2955	-0.2389	-0.1485	-0.2708
	-0.2243	-0.1677	-0.0773	-0.1996

Table P8

One-Sample T: HVAC Project-2: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P2	741	0.45930	0.24576	0.00903
M2P2	741	0.5159	0.2721	0.0100
M3P2	741	0.60631	0.26271	0.00965
M4P2	741	0.48396	0.26369	0.00969
SSMEP2	741	0.71920	0.20519	0.00754

Table P8 (*continued*).

Variable	95.0% CI	T	P
M1P2	(0.44158, 0.47703)	-26.66	0.000
M2P2	(0.4962, 0.5355)	3.00	0.000
M3P2	(0.58736, 0.62526)	-9.71	0.000
M4P2	(0.46494, 0.50298)	-22.30	0.000
SSMEP2	(0.70440, 0.73400)	2.55	0.011

Note: Test of mu = 0.7 vs mu not = 0.7

Results for: HVAC Project-3: M1-M4 & S-SME

Power and sample size calculations indicate 408 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P9). Analysis of variances indicates one must reject the null of no difference among variances with M-3 representing the lowest variance among the models (Figure P3).

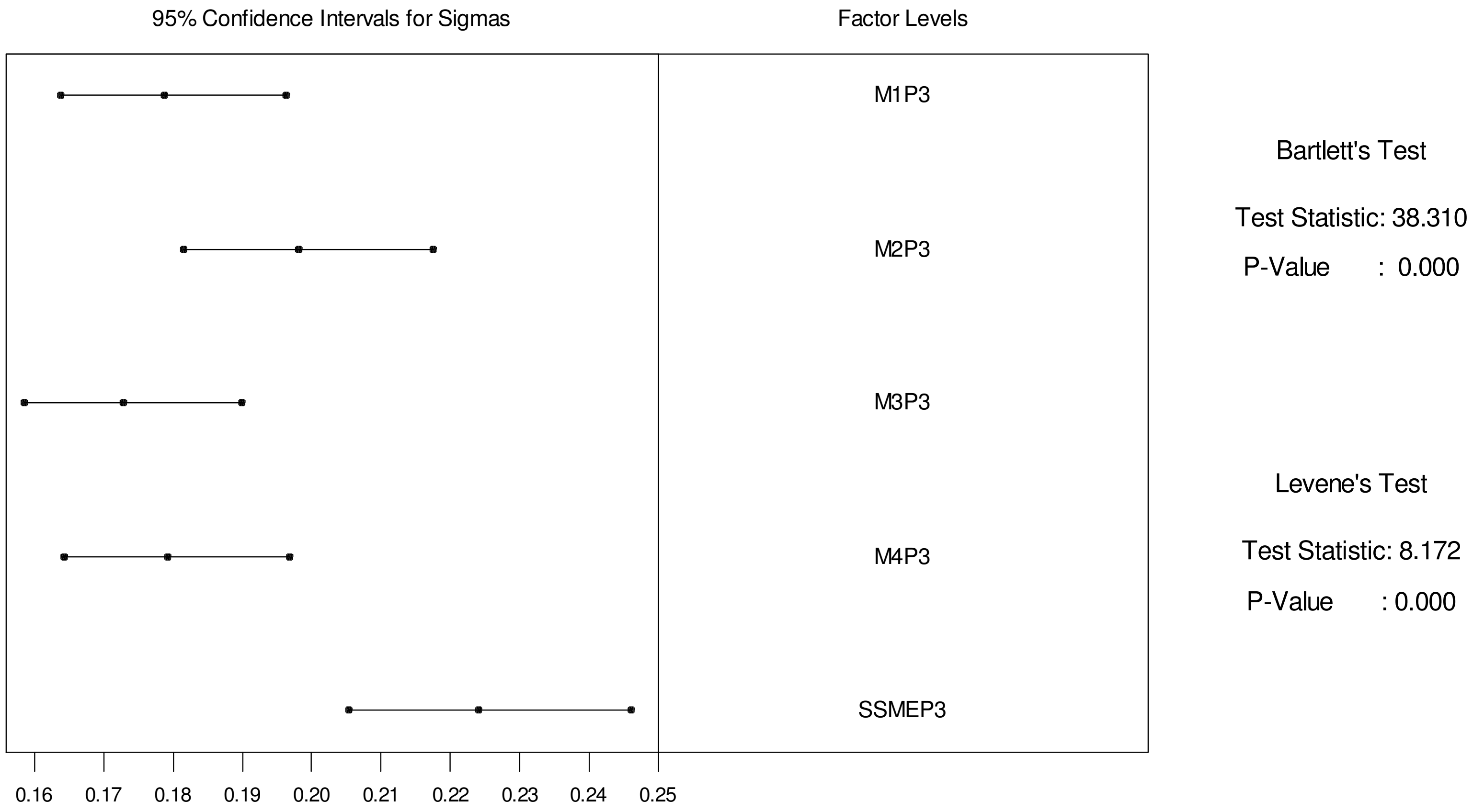
One way ANOVA analysis (Table P10) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P11) indicating statistical significance between all models except S-SME. However one sample T (Table P12) indicates statistical significance between S-SME and actual FEF obtained from field data.

Table P9

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	408	0.9000	0.9003	0.05

Note: Sigma = 0.18165, Alpha = 0.05, Number of Levels = 5



Bartlett's Test

Test Statistic: 38.310
P-Value : 0.000

Levene's Test

Test Statistic: 8.172
P-Value : 0.000

Figure P3. Test for equal variances for HVAC project 3: M1-M4 & S-SME

Table P10

One-way ANOVA: HVAC Project-3: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
HVACP3SM	4	0.9647	0.2412	6.58	0.000
Error	2035	74.6328	0.0367		
Total	2039	75.5975			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level          N   Mean     StDev   ------+---------+---------+---------+---
M1P3         408   0.2272   0.1787         (-----*-----)
M2P3         408   0.2369   0.1981            (-----*-----)
M3P3         408   0.2294   0.1729         (-----*-- ---)
M4P3         408   0.2312   0.1792          (-----*-----)
SSMEP3       408   0.2850   0.2240                           (-----*-----)
Pooled StDev =     0.1915               ---+---------+---------+---------+---
                                       0.210     0.240     0.270     0.300
```

Table P11

Tukey's Pairwise Comparisons

	M1P3	M2P3	M3P3	M4P3
M2P3	-0.0463			
	0.0269			
M3P3	-0.0388	-0.0291		
	0.0344	0.0441		
M4P3	-0.0406	-0.0309	-0.0384	
	0.0326	0.0423	0.0348	
SSMEP3	-0.0943	-0.0847	-0.0921	-0.0904
	-0.0211	-0.0115	-0.0189	-0.0172

Table P12

One-Sample T: HVAC Project-3: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P3	408	0.22722	0.17872	0.00885
M2P3	408	0.23689	0.19807	0.00981
M3P3	408	0.22942	0.17289	0.00856
M4P3	408	0.23119	0.17922	0.00887
SSMEP3	408	0.2850	0.2240	0.0111

Table P12 (*continued*).

Variable	95.0% CI	T	P
M1P3	(0.20982, 0.24461)	-8.23	0.000
M2P3	(0.21762, 0.25617)	-6.44	0.000
M3P3	(0.21259, 0.24625)	-8.25	0.000
M4P3	(0.21375, 0.24864)	-7.75	0.000
SSMEP3	(0.2631, 0.3068)	-1.36	0.176

Note: Test of mu = 0.3 vs mu not = 0.3

Results for: HVAC Project-4: M1-M4 & S-SME

Power and sample size calculations indicate 556 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P13). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P4).

One way ANOVA analysis (Table P14) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P15) indicating statistical significance between M-1 and M-4. One sample T (Table P16) indicates no statistical significance between the models and actual FEF obtained from field data.

Table P13

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	556	0.9000	0.9003	0.05

Note: Sigma = 0.212132, Alpha = 0.05, Number of Levels = 5

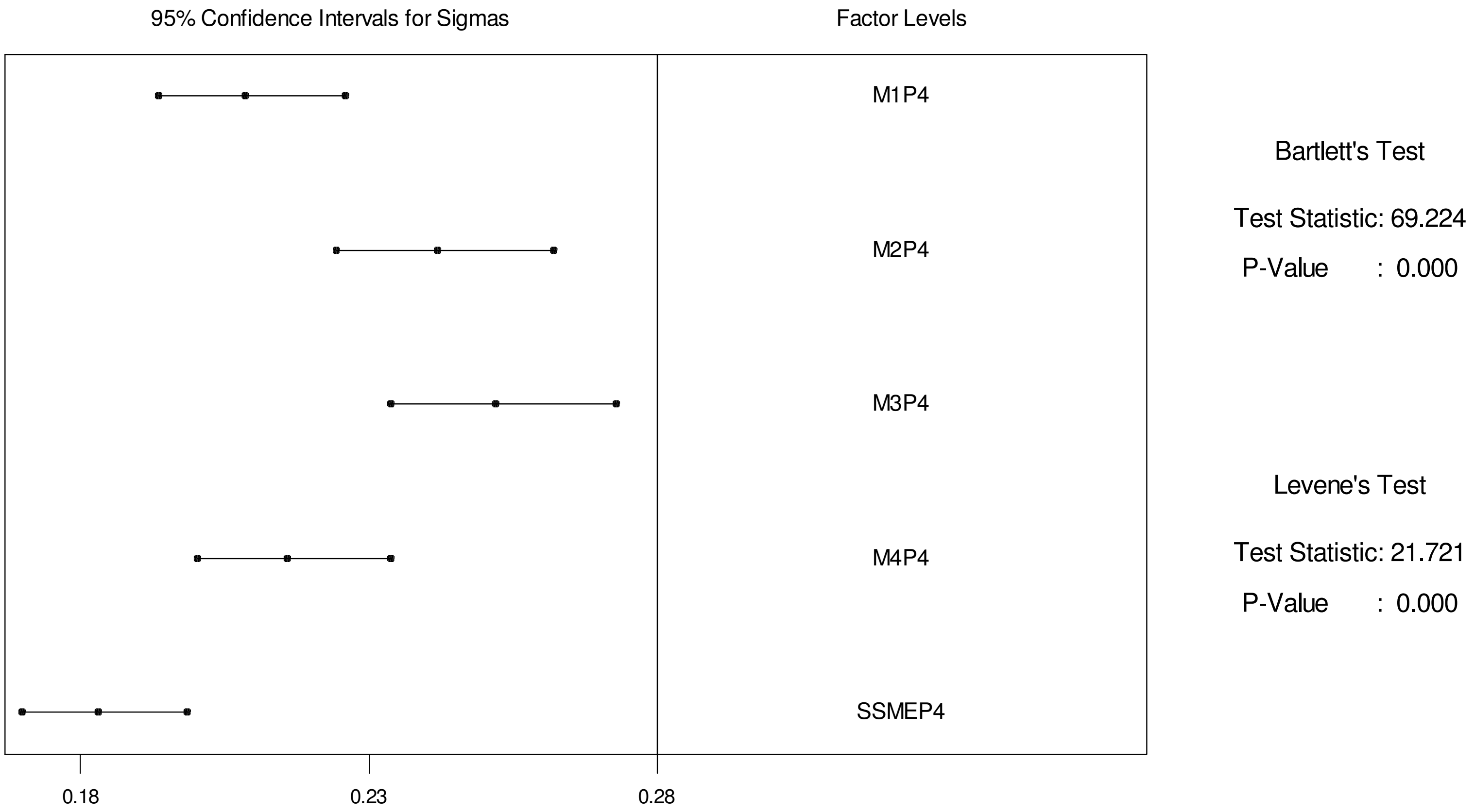


Figure P4. Test for equal variances for HVAC project 4: M1-M4 & S-SME

Table P14

One-way ANOVA: HVAC Project-4: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
HVACP4SM	4	31.4762	7.8691	160.17	0.000
Error	2775	136.3321	0.0491		
Total	2779	167.8083			

Individual 95% Cis for Mean Based on Pooled StDev

Level	N	Mean	StDev	-------+-------+-------+-------
M1P4	556	0.5209	0.2086	(-*-)
M2P4	556	0.5634	0.2419	(-*-)
M3P4	556	0.6816	0.2520	(-*-)
M4P4	556	0.5247	0.2158	(*-)
SSMEP4	556	0.7954	0.1832	(-*)
				-------+-------+-------+-------
Pooled StDev =		0.2216		0.60 0.70 0.80

Table P15

Tukey's Pairwise Comparisons

	M1P4	M2P4	M3P4	M4P4
M2P4	-0.0787			
	-0.0062			
M3P4	-0.1969	-0.1545		
	-0.1244	-0.0819		
M4P4	-0.0401	0.0024	0.1206	
	0.0325	0.0750	0.1932	
SSMEP4	-0.3107	-0.2683	-0.1501	-0.3070
	-0.2382	-0.1957	-0.0775	-0.2344

Table P16

One-Sample T: HVAC Project-4: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P4	556	0.52092	0.20862	0.00885
M2P4	556	0.5634	0.2419	0.0103
M3P4	556	0.6816	0.2520	0.0107
M4P4	556	0.52470	0.21580	0.00915
SSMEP3	556	0.79538	0.18317	0.00777

Table P16 (*continued*).

Variable	95.0% CI	T	P
M1P4	(0.50355, 0.53830)	-11.20	0.000
M2P4	(0.5432, 0.5835)	-5.52	0.000
M3P4	(0.6606, 0.7026)	5.76	0.000
M4P4	(0.50673, 0.54268)	-10.41	0.000
SSMEP4	(0.78012, 0.81063)	22.58	0.000

Note: Test of mu = 0.62 vs mu not = 0.62

Results for: HVAC Project-5: M1-M4 & S-SME

Power and sample size calculations indicate 544 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P17). Analysis of variances indicates one must reject the null of no difference among variances with M-3 representing the lowest variance among the models (Figure P5).

One way ANOVA analysis (Table P18) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P19) indicating statistical significance between M-2 and M-4. One sample T (Table P20) indicates statistical significance between S-SME and actual FEF obtained from field data.

Table P17

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	544	0.9000	0.9005	0.05

Note: Sigma = 0.20976, Alpha = 0.05, Number of Levels = 5

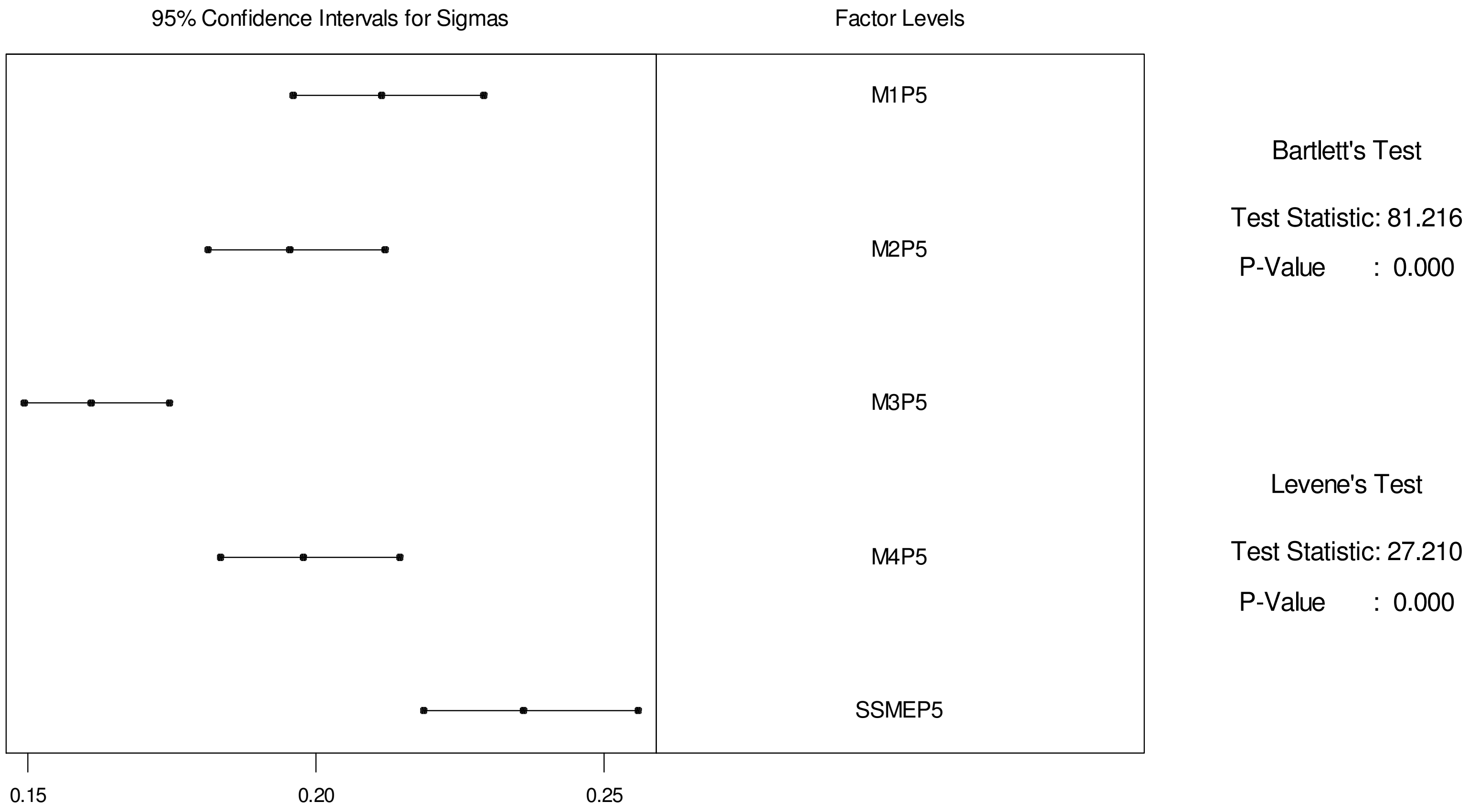


Figure P5. Test for equal variances for HVAC project 5: M1-M4 & S-SME

Table P18

One-way ANOVA: HVAC Project-5: M1-M4 & S-SME

Analysis of Variance					
Source	DF	SS	MS	F	P
HVACP5SM	4	28.9568	7.2392	177.74	0.000
Error	2715	110.5826	0.0407		
Total	2719	139.5394			

Individual 95% Cis for Mean Based on Pooled StDev

Level	N	Mean	StDev	------+---------+---------+---------+
M1P5	544	0.4271	0.2114	(-*--)
M2PS	544	0.2571	0.1955	(-*-)
M3P5	544	0.2151	0.1610	(-*-)
M4P5	544	0.2622	0.1978	(*-)
SSMEP5	544	0.4740	0.2359	(-*-)
Pooled StDev =		0.2018		------+---------+---------+---------+
				0.240 0.320 0.400 0.480

Table P19

Tukey's Pairwise Comparisons

	M1P5	M2P5	M3P5	M4P5
M2P5	0.1366			
	0.2034			
M3P5	0.1785	0.0086		
	0.2453	0.0754		
M4P5	0.1314	-0.0385	-0.0805	
	0.1982	0.0283	-0.0137	
SSMEP5	-0.0804	-0.2503	-0.2923	-0.2452
	-0.0136	-0.1835	-0.2255	-0.1784

Table P20

One-Sample T: HVAC Project-5: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P5	544	0.42706	0.21139	0.00906
M2P5	544	0.25710	0.19549	0.00838
M3P5	544	0.21513	0.16104	0.00690
M4P5	544	0.26224	0.19785	0.00848
SSMEP5	544	0.4740	0.2359	0.0101

Table P20 (*continued*).

Variable	95.0% CI	T	P
M1P5	(0.40926, 0.44487)	-6.94	0.000
M2P5	(0.24064, 0.27357)	-27.79	0.000
M3P5	(0.20157, 0.22869)	-39.81	0.000
M4P5	(0.24557, 0.27890)	-26.85	0.000
SSMEP5	(0.4542, 0.4939)	-1.58	0.115

Note: Test of mu = 0.49 vs mu not = 0.49

Thus far three of five S-SME models have proven statistically significant, with the fourth just missing the mark. Let's now turn our efforts to data collected within the automotive industry.

Power and sample size calculations indicate 593 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P21). Analysis of variances indicates one must reject the null of no difference among variances with M-3 representing the lowest variance among the models (Figure P6).

One way ANOVA analysis (Table P22) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P23) indicating statistical significance between M-1 and M-3. One sample T (Table P24) indicates no statistical significance between the models and actual FEF obtained from field data.

Table P21

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	593	0.9000	0.9003	0.05

Note: Sigma = 0.219089, Alpha = 0.05, Number of Levels = 5

Table P22

One-way ANOVA: Automotive Project-1: M1-M4 & S-SME

Analysis of Variance					
Source	DF	SS	MS	F	P
AutoP1SM	4	14.9265	3.7316	77.68	0.000
Error	2960	142.1908	0.0480		
Total	2964	157.1173			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level     N    Mean   StDev  ------+---------+---------+---------+
M1P1    593  0.6068  0.2148               (--*-)
M2P1    593  0.4746  0.2444 (--*-)
M3P1    593  0.5805  0.2387             (--*-)
M4P1    593  0.5433  0.2352         (--*-)
SSMEP1  593  0.6896  0.1483                       (--*-)
Pooled StDev = 0.2192        ------+---------+---------+---------+
                                 0.490     0.560     0.630    0.700
```

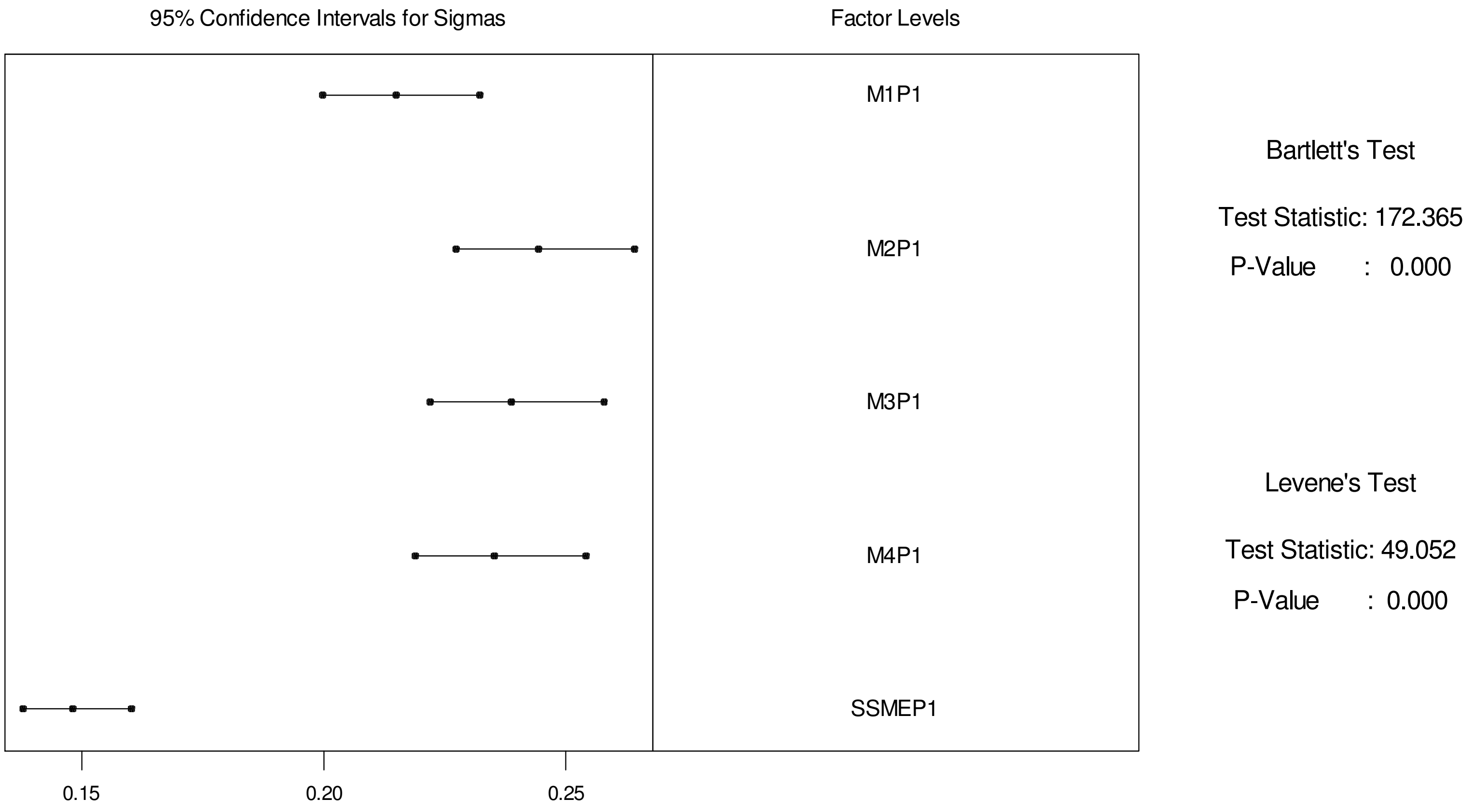


Figure P6. Test for equal variances for auto project 1: M1-M4 & S-SME

Table P23

Tukey's Pairwise Comparisons

	M1P1	M2P1	M3P1	M4P1
M2P1	0.0975			
	0.1669			
M3P1	-0.0084	-0.1406		
	0.0611	-0.0711		
M4P1	0.0288	-0.1034	0.0024	
	0.0983	-0.0339	0.0719	
SSMEP1	-0.1175	-0.2497	-0.1438	-0.1810
	-0.0480	-0.1802	-0.0744	-0.1116

Table P24

One-Sample T: Automotive Project-1: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P1	593	0.60682	0.21485	0.00882
M2P1	593	0.4746	0.2444	0.0100
M3P1	593	0.58045	0.23867	0.00980
M4P1	593	0.54326	0.23524	0.00966
SSMEP1	593	0.68955	0.14830	0.00609

Table P24 (*continued*).

Variable	95.0% CI	T	P
M1P1	(0.58949, 0.62415)	38.74	0.000
M2P1	(0.4549, 0.4943)	20.89	0.000
M3P1	(0.56120, 0.59970)	32.19	0.000
M4P1	(0.52429, 0.56224)	28.80	0.000
SSMEP1	(0.67759, 0.70152)	69.72	0.000

Note: Test of mu = 0.265 vs mu not = 0.265

Results for: Automotive Project-2: M1-M4 & S-SME

Power and sample size calculations indicate 642 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P25). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P7).

One way ANOVA analysis (Table P26) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P27) indicating statistical significance between M-2 and M-4. One sample T (Table P28) indicates no statistical significance between the models and actual FEF obtained from field data.

Table P25

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	642	0.9000	0.9001	0.05

Note: Sigma = 0.228035, Alpha = 0.05, Number of Levels = 5

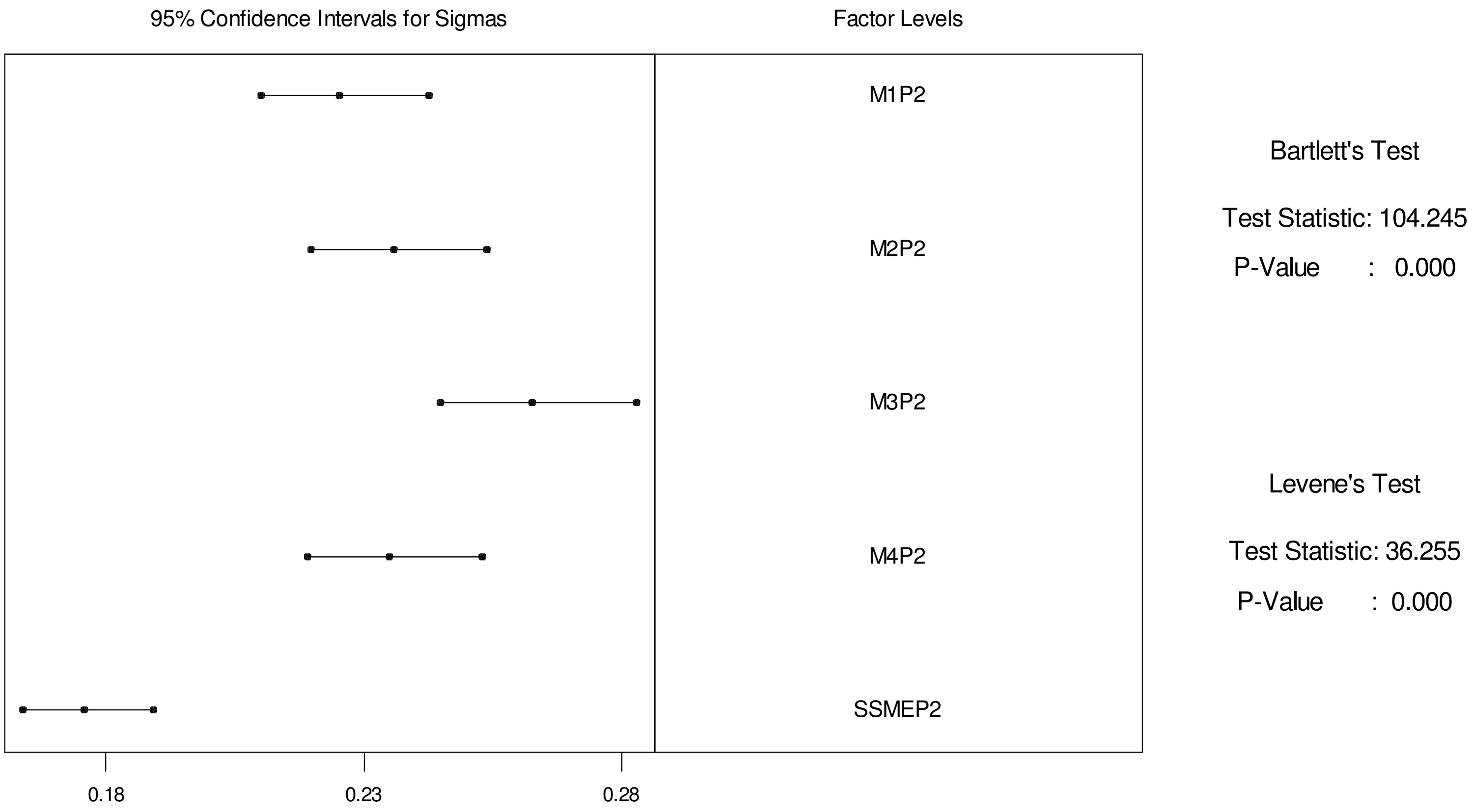


Figure P7. Test for equal variances for auto project 2 M1-M4 & S-SME

Table P26

One-way ANOVA: Automotive Project-2: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP2SM	4	30.4565	7.6141	145.68	0.000
Error	3205	167.5134	0.0523		
Total	3209	197.9699			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level          N    Mean    StDev   ------+---------+---------+---------+
M1P2         643  0.5052  0.2253         (-*)
M2P2         642  0.5695  0.2357            (-*-)
M3P2         642  0.6334  0.2625                  (*-)
M4P2         642  0.5448  0.2349          (*-)
SSMEP2       642  0.7834  0.1758                              (*-)
Pooled StDev =    0.2286            ------+---------+---------+---------+
                                        0.50      0.60     0.70      0.80
```

Table P27

Tukey's Pairwise Comparisons

	M1P2	M2P2	M3P2	M4P2
M2P2	-0.0991			
	-0.0294			
M3P2	-0.1630	-0.0987		
	-0.0933	-0.0291		
M4P2	-0.0744	-0.0102	0.0537	
	-0.0048	0.0595	0.1234	
SSMEP2	-0.3130	-0.2487	-0.1848	-0.2734
	-0.2433	-0.1791	-0.1152	-0.2037

Table P28

One-Sample T: Automotive Project-2: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P2	642	0.50521	0.22534	0.00889
M2P2	642	0.56946	0.23573	0.00930
M3P2	642	0.6334	0.2625	0.0104
M4P2	642	0.54481	0.23487	0.00927
SSMEP2	642	0.78337	0.17577	0.00694

Table P28 (*continued*).

Variable	95.0% CI	T	P
M1P2	(0.48774, 0.52267)	0.59	0.558
M2P2	(0.55119, 0.58773)	7.47	0.000
M3P2	(0.6130, 0.6537)	12.87	0.000
M4P2	(0.52661, 0.56301)	4.83	0.000
SSMEP2	(0.76975, 0.79699)	40.85	0.000

Note: Test of mu = 0.5 vs mu not = 0.5

Results for: Automotive Project-3: M1-M4 & S-SME

Power and sample size calculations indicate 457 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P29). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P8).

One way ANOVA analysis (Table P30) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P31) indicating statistical significance between M-2 and M-4. One sample T (Table P32) indicates no statistical significance between S-SME and actual FEF obtained from field data.

Table P29

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	457	0.9000	0.9000	0.05

Note: Sigma = 0.192354, Alpha = 0.05, Number of Levels = 5

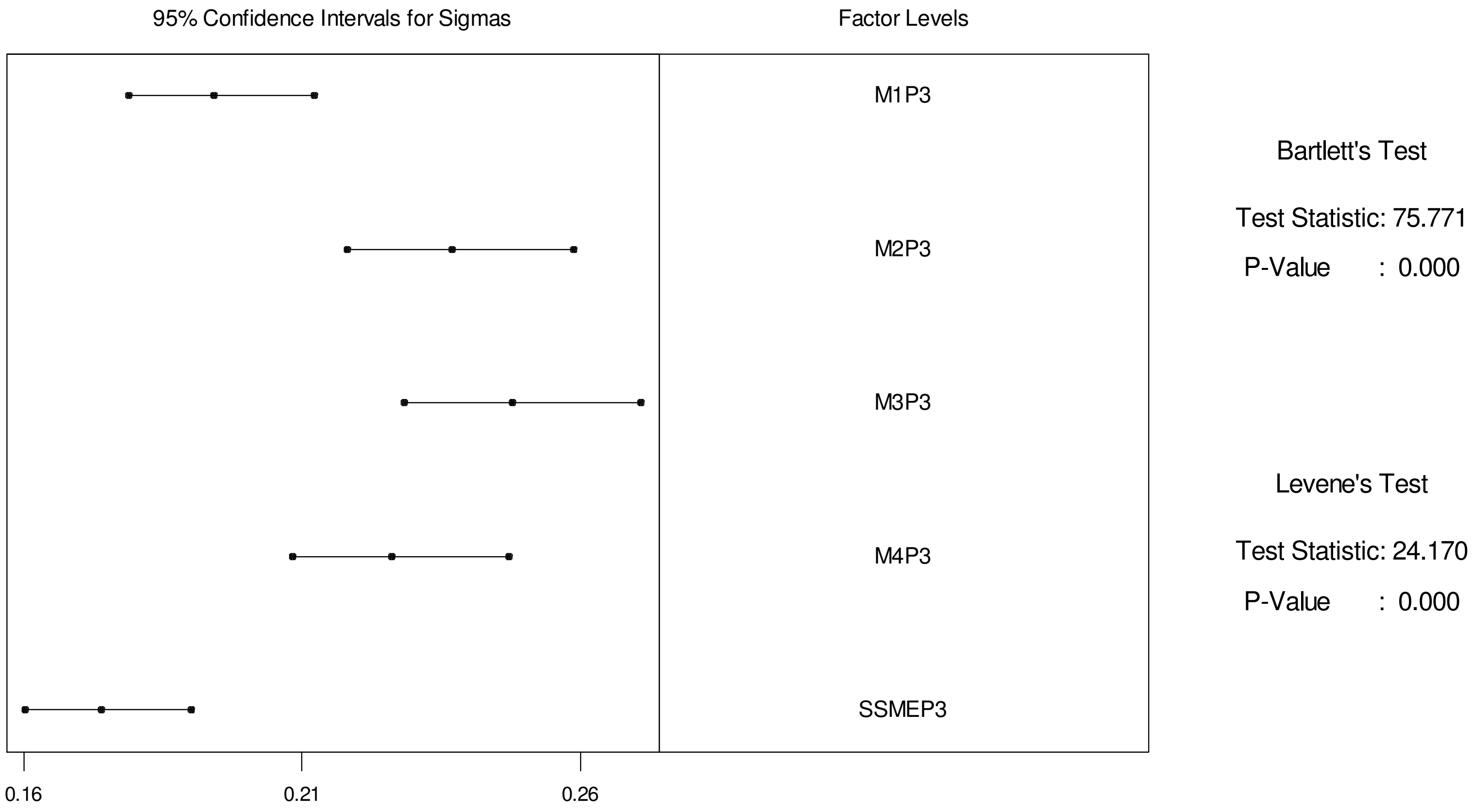


Figure 8. Test for equal variances for auto project 3: M1-M4 & SSME

Table P30

One-way ANOVA: Automotive Project-3: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP3SM	4	25.5159	6.3790	134.78	0.000
Error	2280	107.9058	0.0473		
Total	2284	133.4216			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level           N    Mean    StDev   -----------+------------+-----------+------------
M1P3          457  0.5439   0.1941        (-*-)
M2P3          457  0.5571   0.2369          (-*-)
M3P3          457  0.6676   0.2478                   (-*-)
M4P3          457  0.5529   0.2261        (-*-)
SSMEP2        457  0.8191   0.1740                                  (-*-)
Pooled StDev =     0.2178            -----------+------------+-----------+------------
                                              0.60          0.70        0.80
```

Table P31

Tukey's Pairwise Comparisons

	M1P3	M2P3	M3P3	M4P3
M2P3	-0.0525			
	0.0260			
M3P3	-0.1630	-0.1497		
	-0.0844	-0.0712		
M4P3	-0.0484	-0.0351	0.0754	
	0.0302	0.0435	0.1539	
SSMEP3	-0.3145	-0.3013	-0.1908	-0.3055
	-0.2360	-0.2227	-0.1123	-0.2269

Table P32

One-Sample T: Automotive Project-3: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P3	457	0.54387	0.19413	0.00908
M2P3	457	0.5571	0.2369	0.0111
M3P3	457	0.6676	0.2478	0.0116
M4P3	457	0.5529	0.2261	0.0106
SSMEP3	457	0.81913	0.17396	0.00814

Table P32 (*continued*).

Variable	95.0% CI	T	P
M1P3	(0.52602, 0.56171)	-50.01	0.000
M2P3	(0.05354, 0.5789)	-39.78	0.000
M3P3	(0.5322, 0.5737)	42.07	0.000
M4P3	(0.52661, 0.56301)	4.83	0.000
SSMEP3	(0.80314, 0.83513)	-21.98	0.000

Note: Test of mu = 0.998 vs mu not = 0.998

Results for: Automotive Project-4: M1-M4 & S-SME

Power and sample size calculations indicate 408 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P33). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P9).

One way ANOVA analysis (Table P34) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P35) indicating no statistical significance between the models. One sample T (Table P36) indicates no statistical significance between S-SME and actual FEF obtained from field data.

Table P33

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	408	0.9000	0.9003	0.05

Note: Sigma = 0.181659, Alpha = 0.05, Number of Levels = 5

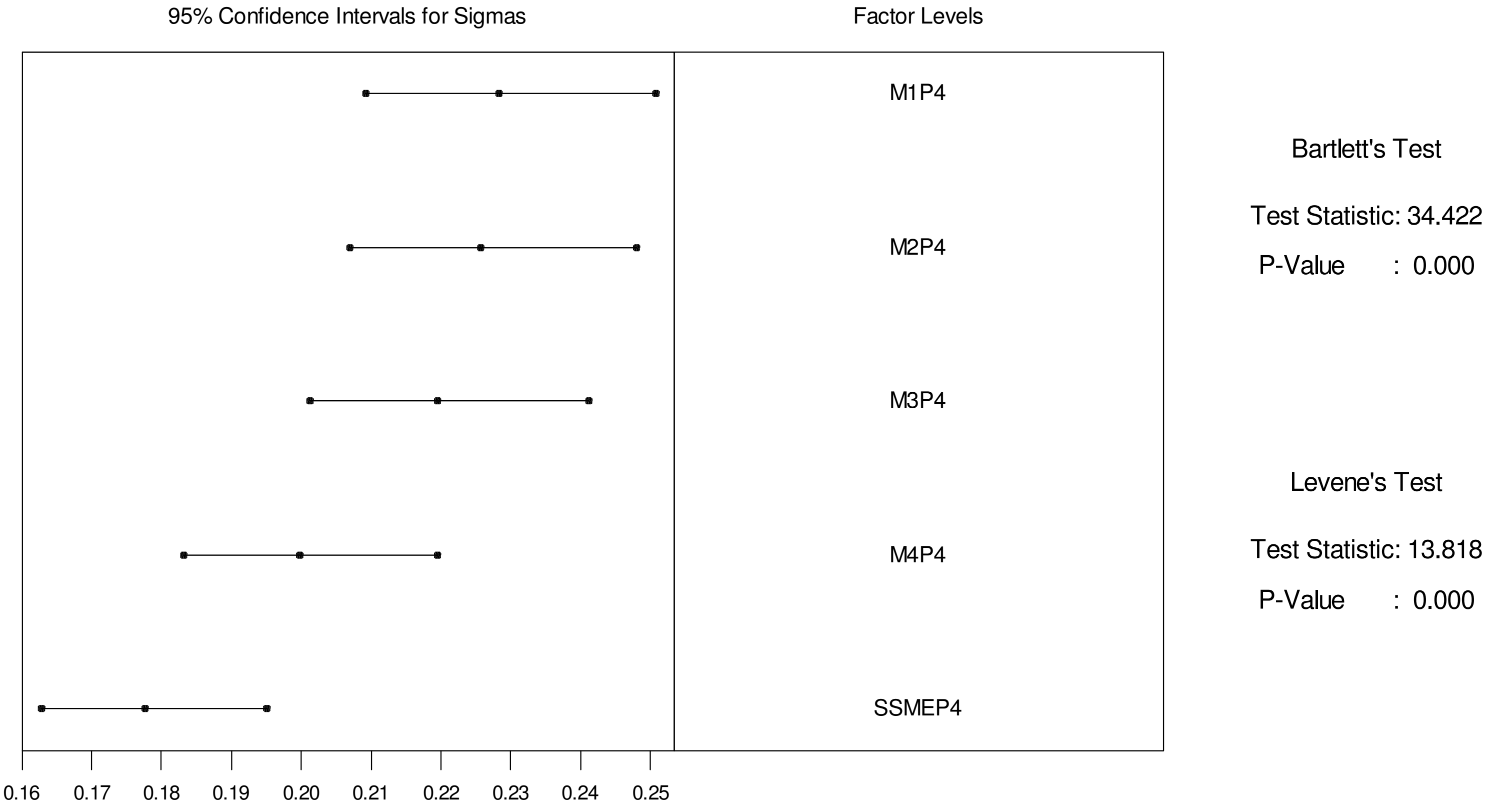


Figure P9. Test for equal variances for auto project 4: M1-M4 & S- SME

Table P34

One-way ANOVA: Automotive Project-4: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP4SM	4	37.7329	9.4332	211.81	0.000
Error	2035	90.6313	0.0445		
Total	2039	128.3642			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level       N    Mean   StDev   ------+---------+---------+---------+
M1P4      408  0.4361  0.2282      (*-)
M2P4      408  0.6007  0.2257              (-*-)
M3P4      408  0.7147  0.2195                    (-*)
M4P4      408  0.5098  0.1998         (*-)
SSMEP4    408  0.8132  0.1776                              (-*)
                                ------+---------+---------+---------+
Pooled StDev = 0.2110
                                    0.48      0.60      0.72      0.84
```

Table P35

Tukey's Pairwise Comparisons

	M1P4	M2P4	M3P4	M4P4
M2P4	-0.2048			
	-0.1242			
M3P4	-0.3189	-0.1543		
	-0.2382	-0.0737		
M4P4	-0.1140	0.0505	0.1646	
	-0.0333	0.1312	0.2452	
SSMEP4	-0.4174	-0.2529	-0.1388	-0.3437
	-0.3367	-0.1722	-0.0582	-0.2631

Table P36

One-Sample T: Automotive Project-4: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P4	408	0.4361	0.2282	0.0113
M2P4	408	0.6007	0.2257	0.0122
M3P4	408	0.7147	0.2195	0.0109
M4P4	408	0.50980	0.19982	0.00989
SSMEP4	408	0.81320	0.17757	0.00879

Table P36 (*continued*).

Variable	95.0% CI	T	P
M1P4	(0.4139, 0.4583)	-36.89	0.000
M2P4	(0.5787, 0.6226)	-22.58	0.000
M3P4	(0.6933, 0.7360)	-12.73	0.000
M4P4	(0.49035, 0.52924)	-34.69	0.000
SSMEP4	(0.79592, 0.83048)	-4.53	0.000

Note: Test of mu = 0.853 vs mu not = 0.853

Results for: Automotive Project-5: M1-M4 & S-SME

Power and sample size calculations indicate 556 samples will be required given the variance of M-1 such that the power of the experiment can be held constant at 0.9 (Table P37). Analysis of variances indicates one must reject the null of no difference among variances with S-SME representing the lowest variance among the models (Figure P10).

One way ANOVA analysis (Table P38) indicates one must reject the null of no difference in the models with Tukey pairwise comparisons (Table P39) indicating no statistical significance between the models and one sample T (Table P40) indicates no statistical significance between the models and actual FEF obtained from field data.

Table P37

Power and Sample Size

SS Means	Sample Size	Target Power	Actual Power	Maximum Difference
0.00125	556	0.9000	0.9003	0.05

Note: Sigma = 0.21213, Alpha = 0.05, Number of Levels = 5

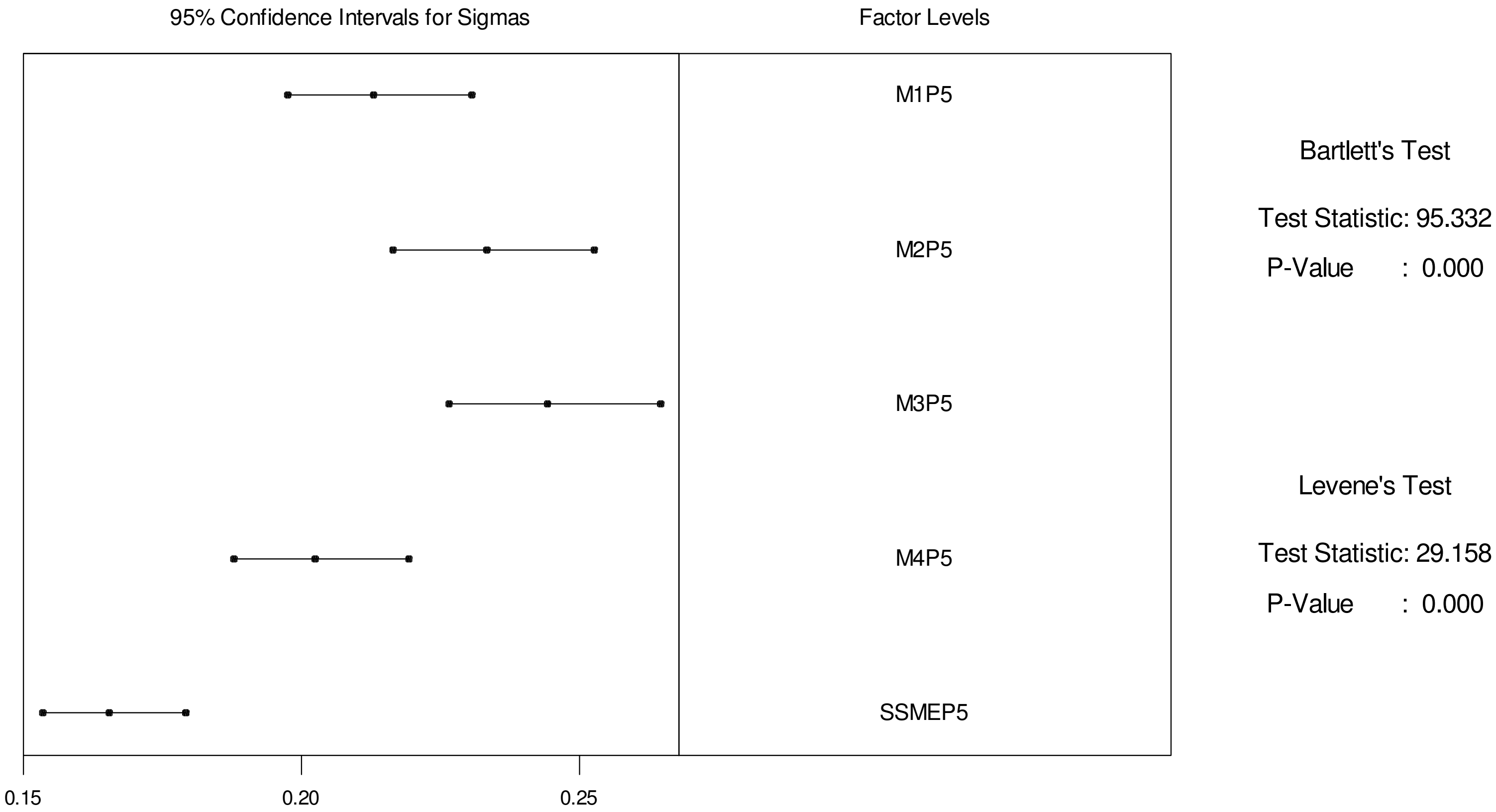


Figure 10. Test for equal variances for auto project 5: M1-M4 & SSME

Table P38

One-way ANOVA: Automotive Project-5: M1-M4 & S-SME

Analysis of Variance

Source	DF	SS	MS	F	P
AutoP5SM	4	32.4915	8.1229	178.25	0.000
Error	2775	126.4552	0.0456		
Total	2779	158.9467			

Individual 95% Cis for Mean Based on Pooled StDev

```
Level        N    Mean    StDev     ------+---------+---------+---------+
M1P5       556  0.4500   0.2129             (-*-)
M2P5       556  0.3701   0.2333        (-*-)
M3P5       556  0.5159   0.2442                  (-*)
M4P5       556  0.4091   0.2025           (-*-)
SSMEP5     556  0.6781   0.1655                                (-*-)
                                    ------+---------+---------+---------+
Pooled StDev =  0.2135
                                        0.40      0.50       0.60      0.70
```

Table P39

Tukey's Pairwise Comparisons

	M1P5	M2P5	M3P5	M4P5
M2P5	0.0449			
	0.1148			
M3P5	-0.1009	-0.1807		
	-0.0310	-0.1108		
M4P5	0.0060	-0.0739	0.0719	
	0.0759	-0.0040	0.1418	
SSMEP5	-0.2631	-0.3429	-0.1972	-0.3040
	-0.1932	-0.2730	-0.1273	-0.2341

Table P40

One-Sample T: Automotive Project-5: M1-M4 & S-SME

Variable	N	Mean	St.Dev.	SE Mean
M1P5	556	0.44999	0.21291	0.00903
M2P5	556	0.37013	0.23334	0.00990
M3P5	556	0.5159	0.2442	0.0104
M4P5	556	0.40908	0.20255	0.00859
SSMEP5	556	0.67813	0.16552	0.00702

Table P40 (*continued*).

Variable	95.0% CI	T	P
M1P5	(0.43225, 0.46772)	-13.18	0.000
M2P5	(0.35070, 0.38957)	-20.10	0.000
M3P5	(0.4956, 0.5363)	-5.13	0.000
M4P5	(0.39221, 0.42595)	-18.62	0.000
SSMEP5	(0.66434, 0.69192)	15.55	0.000

Note: Test of mu = 0.569 vs mu not = 0.569

Appendix Q

Reliability Growth Projection Error

RELIABILTY GROWTH PROJECTION ERROR

Q.1. Reliability Growth Projection Error as a Function of Fix Effectiveness Variation

A simulation process was developed to evaluate the effects fix effectiveness variation has on reliability growth projection for both Crow-AMSAA and AMPM-Stein models. The simulation involved use of time to fail data for m of k failure modes. Failure rates for the i^{th} failure modes are calculated afterwards of which FEF variability is introduced into projections of $\rho(T)$. Model error for mean reliability projection $r(T)$ is calculated for both Crow-AMSAA Projection and AMPM-Stein reliability models where $r(T) = \frac{1}{\rho(T)}$.

The following questions will be answered from the simulation of both AMPM-Stein and Crow-AMSAA Projection models: *(i)* How does variability around FEF influence each model and *(ii)* which is the more robust model relative to FEF variability?

The simulation steps are noted below:

1. <u>Specify simulation inputs</u>. The simulation requires five inputs *(i)* *r*, the total number of simulation tests; *(ii)* T, the total of continuous run hours; *(iii)* *m*, the total number

   of surfaced system failure modes; *(iv)* time to fail for surfaced failure modes, and *(v)* the mean and variance of a beta distribution used to generate *FEF* variability.
2. <u>Identify A and B failure modes</u>. For each of the m modes, designate them as A-modes or B-modes. A-modes are failure modes that will not receive corrective action, whereas B-modes will receive corrective actions and will be

susceptible to FEF and the associated variability. B-mode first occurrence times will be noted in this step.

3. Introduce Fix Effectiveness Variability. Utilize given *FEF* and variances to calculate Beta shape parameters [Martz and Waller (1982)] of which size *r* realizations of *FEF* will be randomly drawn.

$$\alpha = \left(\frac{1-d_i}{FEF_{VAR}}\right)\left(d_i^{\,2}\right) - d_i \qquad (17) \qquad \beta = \left(\frac{d_i(1-d_i)^2}{FEF_{Var}}\right) - (1-d_i) \quad (18)$$

FEF Var	0.001	0.0025	0.005	0.0075	0.01	0.02	0.03	0.04	0.05

4. Calculate projected reliability $r(T)$. Utilizing random realizations of FEF, calculate $r(T)$ for Crow-AMSAA and AMPM-Stein projection models where $r(t) = \frac{1}{\rho(t)}$.
5. Error estimation. Estimate error, $\breve{E}$, between the true MTBF $r(T)$ and $\breve{r}(t)$, the realized MTBF based on variation in fix effectiveness where $\breve{E} = \frac{r(T) - \breve{r}(T)}{r(T)}$.
6. Replication. Repeat steps 3-5 *r* times.
7. Evaluate. Analysis of Variance Method is used to evaluate *FEF* variance impact on reliability projection. The process begins with determination of reliability projection error standard deviation of the *FEF* in question, determining sample sizes for a statistical difference of one sigma, one-half sigma…one-tenth sigma such that the power assumption of 0.9 is maintained.

Instruments

The instruments used in this experiment were Mathematica version 5.2 and Mini-Tab version 14. Mathematica was used to perform *r* simulations of failure intensities, while introducing random *FEF* and associated variability, calculate *MTBF* and model error. Mini-Tab was used to perform ANOVA's of model error comparing *FEF* impact on Crow-AMSAA Projection and AMPM-Stein reliability growth models.

Study Findings

Data used in the simulation consists of both A and B failure modes, the number of occurrences, and time to failure for each surfaced failure mode (ReliaSoft, 1999). Table Q1 shows 42 surfaced failures, 10 A-modes, 32 B-modes with 16 distinct BD modes, with total time on test of 400 hours. Table Q2 shows first occurrence BD modes and their respective *FEF*.

The simulation process involved calculation of the true failure intensity function $\rho(t)$ for both Crow-AMSAA Projection and AMPM-Stein using TTF data, the number of A and B failure modes, the rate of occurrence of B-modes, and the mean FEF values of Table Q2 with $r(t) = \frac{1}{\rho(t)}$.

Beta distribution shape parameters were developed using mean FEF and variation values whereby random FEF values are generated ultimately leading to $\breve{\rho}(t)$, the realized failure intensity based. Model error was stored for each iteration of random FEF and evaluated using analysis of variance.

Table Q1

Failure Modes

i	Xi	Mode	i	Xi	Mode
1	15	BD1	22	260.1	BD1
2	25.3	BD2	23	263.5	BD8
3	47.5	BD3	24	273.1	A
4	54	BD4	25	274.7	BD6
5	56.4	BD5	26	285	BD13
6	63.6	A	27	304	BD9
7	72.2	BD5	28	315.4	BD4
8	99.6	BD6	29	317.1	A
9	100.3	BD7	30	320.6	A
10	102.5	A	31	324.5	BD12
11	112	BD8	32	324.9	BD10
12	120.9	BD2	33	342	BD5
13	125.5	BD9	34	350.2	BD3
14	133.4	BD10	35	364.6	BD10
15	164.7	BD9	36	364.9	A
16	177.4	BD10	37	366.3	BD2
17	192.7	BD11	38	373	BD8
18	213	A	39	379.4	BD14
19	244.8	A	40	389	BD15
20	249	BD12	41	394.9	A
21	250.8	A	42	395.2	BD16

Note: Published with permission of ReliaSoft Corporation

Table Q2

First Time Occurrence B-Modes

BD Mode	Number Ni	First Occurrence	FEF di
1	2	15	0.67
2	3	25.3	0.72
3	2	47.5	0.77
4	2	54	0.77
5	3	54	0.87
6	2	99.6	0.92
7	1	100.3	0.50
8	3	112	0.85
9	3	125.5	0.89
10	4	133.4	0.74
11	1	192.7	0.70
12	2	249	0.63
13	1	285	0.64
14	1	379.4	0.72
15	1	389	0.69
16	1	395.2	0.46

Note: Published with permission of ReliaSoft Corporation

D.1.1 Crow-AMSAA

Utilizing *FEF* noted in the first occurrence B-modes and equation (6), one can show the true value for $\rho(T) = 0.067493$ where $\lambda_A = \frac{N_A}{T} = \frac{10}{400} = 0.025$,

$\sum_{i=1}^{m}(1-d_i)\lambda_i = 0.01955$, $\hat{\beta}_{BD} = \frac{m}{\sum_{i=1}^{m}\ln\left(\frac{T}{X_i}\right)} = 0.79524$, $\hat{\lambda}_{BD} = \frac{m}{T^{\hat{\beta}_{BD}}} = 0.1364$

and $\bar{d} = \frac{\sum_{i=1}^{m} d_i}{m} = 0.72125$, $h(t) = 3.18095x10^{-2}$ and $\bar{d}\,h(t) = 2.294x10^{-2}$.

Random FEF were generated from a Beta distribution with variances of simulation step 3 and a mean equal to the FEF of the ith BD mode (Table Q3). This process is repeated for all BD modes with $\breve{r}(T)$ recalculated for each group of randomly distributed FEF. The value of $r(T)$, calculated from the given FEF, is used as the true value. Relative error is calculated for each group of FEF and replicated 50000 times.

Q.1.2 AMPM-Stein

Repeating the process used on the Crow-AMSAA model, failure intensity projections are made using the data from Tables Q3 and Q4.

Equation (11) is expanded to show $\rho(T)$ as a function of both A and B failure modes where;

$$\rho_s(T) = \frac{N_A}{T} + \sum_{i\in obsB}(1-d_i)\tilde{\lambda}_{i,B} + \sum_{i\in \overline{obsB}}\tilde{\lambda}_{i,B} = 0.025 + 0.019287 + 0.02389 = 0.068177$$

$\tilde{\lambda}_{i,B} = \theta_s\,\hat{\lambda}_{i,B} + (1-\theta_{s,B})\frac{\sum_{i\in B}\hat{\lambda}_i}{k_B}$, $\theta_s = \frac{\sum_{i\in B}(\lambda_i - \bar{\lambda}_B)^2}{\left(\frac{\lambda_B}{kT}\right)\left(1-\frac{1}{k_B}\right) + \sum_{i\in B}(\lambda_i - \bar{\lambda}_B)^2} = 0.96872$, and

$$\sum_{i \in \overline{obs}B} \tilde{\lambda}_i \approx \frac{\hat{\lambda}_B}{1+(\beta_{s,B})T} = 2.389x10^{-2} \text{ where } \hat{\lambda}_B = \frac{N_B}{T} = 0.08 \text{, and}$$

$$\beta_{s,B} = \frac{\sum_{i \in B} \lambda_i^2}{\lambda_B} = 5.87x10^{-3}.$$

Analysis Method

To answer the question "how does variability around *FEF* influence each model," three evaluations were performed, *(i)* test of equal variances between models for a given *FEF-var, (ii)* ANOVA of mean error between models for a given *FEF-var* and *(iii)* ANOVA of mean error within each model for varying levels of *FEF-var.*

Q.1.3 Test of Equal Variance Between Models For a Given FEF-Var

To statistically compare models at a given *FEF-var* a power value of 0.9 is assumed and a sample size is calculated for a difference in variation of one standard deviation, one-half standard deviation etc . . . up to one-tenth standard deviation difference. Test of equal variance is used to evaluate p-values to determine if one must reject the null of no difference in variance (p-value <0.05) or fail to reject the null for each difference of interest (p-value>0.05). Table Q3 shows p-values in excess of 0.05, indicating one must fail to reject the null of no difference in variation between the Crow-AMSAA and AMPM-Stein reliability growth projection models error standard deviations as a function of *FEF-var*.

Table Q3

Error Evaluation at Each Variance – Comparison of Model Error Variance at Each FEF-Variance Level

			Diff.	σ	σ/2	σ/3	σ/4	σ/5	σ/6	σ/7	σ/8	σ/9	σ/10
Crow-AMSAA Std. Dev.	AMPM-Stein Std. Dev.	FEF-Var	Sample Size	23	86	191	338	527	758	1031	1346	1704	2103
0.031355	0.03054	0.05	P-Value	0.622	0.653	0.460	0.099	0.143	0.244	0.466	0.103	0.247	0.479
0.028057	0.027332	0.04		0.772	0.114	0.312	0.704	0.380	0.634	0.342	0.680	0.155	0.400
0.024154	0.023671	0.03		0.121	0.313	0.996	0.175	0.484	0.905	0.666	0.300	0.150	0.054
0.019864	0.019257	0.02		0.965	0.743	0.342	0.698	0.013	0.091	0.174	0.197	0.524	0.438
0.013955	0.013597	0.01		0.308	0.149	0.530	0.138	0.709	0.108	0.386	0.237	0.845	0.494
0.012127	0.011816	0.0075		0.142	0.543	0.606	0.660	0.175	0.450	0.080	00323	0.133	0.285
0.009877	0.009657	0.005		0.304	0.316	0.686	0.846	0.174	0.070	0.093	0.785	0.585	0.778
0.006992	0.006865	0.0025		0.969	0.282	0.273	0.315	0.566	0.261	0.350	0.088	0.037	0.088
0.004421	0.004301	0.001		0.751	0.107	0.233	0.115	0.865	0.988	0.593	0.081	0.123	0.226

Q.1.4 ANOVA of Mean Reliability Projection Error Between Models For a Given FEF-Var

Once again a power value of 0.9 was assumed, and a sample size was calculated for a difference in mean reliability projection error of one standard deviation, one-half standard deviation etc…up to one-tenth standard deviation difference. A two sample-T test was used to evaluate if one must reject the null of no difference in means at a given *FEF-var*, (p-value <0.05) or fail to reject the null (p-value>0.05).

Figure Q3. AMPM Stein Error

Table Q4 suggests however, that for mean error differences less than 1/7 standard deviation, one can repeatedly detect a difference between Crow-AMSAA Projection and AMPM-Stein. In every case, when the p-value was < 0.05, the reliability projection mean error associated with the AMPM-Stein was less than that of the Crow-AMSAA Projection model.

Table Q4

Mean Error Comparison of Model Error Means at Each FEF-Variance Level

			Diff.	σ	σ/2	σ/3	σ/4	σ/5	σ/6	σ/7	σ/8	σ/9	σ/10
Crow-AMSAA Std. Dev.	AMPM-Stein Std. Dev.	FEF-Var	Sample Size	23	86	191	338	527	758	1031	1346	1704	2103
0.031355	0.03054	0.05	P-Value	0.572	0.956	0.642	0.914	0.448	0.449	0.795	0.229	0.001	0.088
0.028057	0.027332	0.04		0.168	0.805	0.454	0.675	0.832	0.535	0.419	0.774	0.144	0.348
0.024154	0.023671	0.03		0.188	0.557	0.674	0.986	0.722	0.722	0.008	0.392	0.489	0.437
0.019864	0.019257	0.02		0.974	0.616	0.237	0.941	0.502	0.091	0.417	0.661	0.244	0.113
0.013955	0.013597	0.01		0.991	0.474	0.896	0.631	0.453	0.514	0.805	0.014	0.027	0.012
0.012127	0.011816	0.0075		0.132	0.429	0.447	0.450	0.050	0.733	0.470	0.037	0.073	0.362
0.009877	0.009657	0.005		0.109	0.934	0.776	0.120	0.766	0.809	0.001	0.256	0.030	0.003
0.006992	0.006865	0.0025		0.373	0.945	0.012	0.000	0.929	0.261	0.102	0.000	0.008	0.001
0.004421	0.004301	0.001		0.417	0.736	0.153	0.005	0.182	0.001	0.015	0.000	0.000	0.000

Q.1.5 ANOVA of Mean Reliability Projection Error Within Each Model

A one-way ANOVA was used to analyze the effects that 9 levels of *FEF-var* have on reliability projection means within Crow-AMSAA and AMPM-Stein reliability projection models respectively. Again, a power value of 0.9 was assumed and sample sizes were calculated for a difference in mean reliability projection error of one standard deviation, one-half standard deviation etc…up to one-tenth standard deviation difference. Note, the standard deviation for a *FEF-var* of 0.001 was used as the base standard deviation for sample size calculations. Table Q5 shows a significant effect takes place at a difference of 1/9 standard deviation for the Crow-AMSAA model and 1/10 standard deviation for the AMPM-Stein.

In summary, test of equal variances indicate that one must fail to reject the null hypothesis of no difference in variances between each model for a given FEF-var. Two sample T tests however, indicate one must reject the null of no difference in mean error. In every instance, when a difference in mean error was detected, the AMPM Stein mean error was lower. Statistically, the AMPM-Stein model is the more robust model against the effects of FEF variability.

Table Q5

Error Within Models – Mean Error Evaluation within Models

		Diff.	σ	σ/2	σ/3	σ/4	σ/5	σ/6	σ/7	σ/8	σ/9	σ/10
FEF-Var used as Base	Base Std. Dev.	Sample Size	40	154	345	612	955	1375	1871	2444	3092	3818
C0.001	0.004421	P-Value	0.529	0.214	0.535	0.294	0.704	0.092	0.085	0.547	0.000	0.01
S0.001	0.004301		0.127	0.55	0.83	0.122	0.902	0.822	0.652	0.326	0.205	0.004

Bibliography

Aroef, M. (1957). *Study of learning curves of industrial manual operations.* Unpublished Master's Thesis, Cornell University.

Ashton, A. H, & Ashton, R. H. (1985). Aggregating subjective forecasts: Some empirical results. *Management Science, 31,* 1499-1508.

Ashton, R. H. (1986). Combining the judgments of experts: How many and which ones? *Organizational Behavior and Human Decision Processes, 38,* 405-414.

Ayyub, B. M. (2001). *Elicitation of expert opinions for uncertainty and risk.* Boca Ratan, FL: CRC Press.

Barlow, R. E., & Scheuer, E. M. (1966). Reliability growth during a development testing program. *Technometrics, 8*(1), 53-60.

Bedford, T., & Cook, R. (2001). *Probabilistic risk analysis.* Cambridge, MA: Cambridge University Press.

Budescu, D. V., & Rantilla, A. K. (2000). Confidence in aggregation of expert opinions. *Acta Psychologica, 104,* 371-398.

Chernoff, H., & Woods, W. M. (1962, February). *Reliability growth models-Analysis and applications* (Field Memo). *CEIR*, Los Altos, CA.

Corcoran, W. J., Weingarten, H., & Zehna, P. W. (1964). Estimating reliability after corrective action. *Management Science, 10*(4), 786-798.

Crow, L. (2006). Useful metrics for managing failure mode corrective action. Fort Worth, TX: RAMS.

Crow, L. H. (1982). *An improved methodology for reliability growth projection,* Tech. Rep. TR-357. Aberdeen Proving Ground, MD: AMSAA.

Crow, L. H. (1983). *Reliability growth projection from delayed fixes*. IEEE Proceedings of the Annual Reliability and Maintainability Symposium, New York.

Crow, L. H. (2004). *An extended reliability growth model for managing and assessing corrective actions*. IEEE Proceedings of the Annual Reliability and Maintainability Symposium, January 26-29, Los Angeles, CA.

Davis, D., Hwang, H., & Tsang, L. (1994). *Solving remote sensing inverse problems using Bayesian modeling to incorporate information sources.* Presentation at the International Geo-Science and Remote Sensing Symposium, pp. 1395-1397, Pasadena, CA, August 1994.

Droguett, E., & Mosleh, A. (2008). *Use of Heritage data and other partially observable information in reliability prediction of systems under development.* Fort Worth, TX: RAMS.

Duane, J. (1964). Learning curve approach to reliability monitoring. *IEEE Transactions on Aerospace, 2,* 563-566.

Ellner, P., & Hall, J. (2005). *AMSAA maturity projection model based on Stein estimation.* IEEE Proceedings of the Annual Reliability and Maintainability Symposium, Alexandria, VA.

Ellner, P., & Wald, L. (1995). *AMSAA maturity projection model.* Paper presented at the 1995 Reliability and Maintainability Symposium, Washington, DC.

Ellner, P., McCarthy, M., Mortin, D., & Querido, D. (1995). Developing a subsystem reliability growth program using SPLAN. *Proceedings of the 1995 Annual Reliability and Maintainability Symposium*, Washington, DC.

Ellner, P. M., Broemm, W. J., & Woodworth, W. J. (2000). *Reliability growth guide, Tech. Rep. TR-652.* Aberdeen Proving Ground, MD: AMSAA.

Ellner, P. M., McCarthy, M. J., Mortin, D. E., & Querido, D. M. (1994). *Developing a subsystem reliability growth program using SPLAN, Tech. Rep. 555*. Aberdeen Proving Ground, MD: AMSAA.

Farquhar, K., & Mosleh, A. (1995). An approach to quantifying reliability growth effectiveness. *Proceedings of the 1995 Annual Reliability and Maintainability Symposium*, Washington, DC.

Friedman, N., & Goldszmidt, M. (1998). *Learning Bayesian networks from data.* SRI International.

Gibson, G., & Crow, L. (1989). Reliability fix effectiveness factor estimation. *Reliability and Maintainability Symposium* (pp. 171-177). Fort Worth, TX: RAMS.

Gueston, C. (2007). *Bayesian learning*. Retrieved March 13, 2008, from Carnegie Melton University Web site: http://www.cs.cmu.edu~guestrin/class/10701

Hall, B. (2007). *Expert engineering judgment in probabilistic risk assessment.* Unpublished Manuscript, College Park, MD: University of Maryland

Hall, J. B. (2008). *Methodology for Evaluating Reliability Growth Programs of Discrete Systems.* Dissertation, University of Maryland.

Hodge, R., Evans, M., Marshall, J., Quigley, J., & Walls, L. (2001). Eliciting engineering knowledge about reliability during design-lessons learnt from implementation. *Quality and Reliability Engineering International*, 169-179.

Hogarth, R. M. (1978). A note on aggregating opinions. *Organizational Behavior and Human Performance, 21,* 40-46.

Howard, R. (1989). Knowledge maps. *Management Science*, 903-922.

Hsu, W. (1999, October 5). *Introduction to Bayesian networks*. Retrieved June 29, 2008, from Kansas State University Web site: http://www.cis.ksu.edu/~bhsu

Lloyd, D. K., & Lipow, M. (1962). *Reliability growth models.* Reliability: Management, Methods, and Mathematics, Prentice-Hall Space Technology Series. Englewood Cliffs, New Jersey: Prentice-Hall Inc.

Martz, H. F., & Waller, R. A. (1991). *Selecting beta prior distributions based on observed binomial sampling data in Bayesian Reliability Analysis* (2nd ed.). Malabar, FL: Krieger.

Martz, H. W. (1982). *Selecting beta prior distributions based on observed binomial sampling data.* San Francisco, CA: John Wiley & Sons.

Mosleh, A. (2002). "Lecture notes: Advanced methods in reliability engineering." University of Maryland, College Park.

Mosleh, A., & Apostolakis, G. (1986). The assessment of probability distributions from expert opinions with an application to seismic fragility curves. *Risk Analysis, 6,* 447-461.

Mosleh, A., Bier, V. M., & Apostolakis, G. (1988). A critique of current practice for the use of expert opinions in probabilistic risk assessment. *Reliability Engineering and System Safety, 20,* 63-85.

Nyberg, J., Marcot, B., & Sulyma, R. (2006). Using Bayesian belief networks in adaptive management. *NRC Research Press*, 3104-3116.

Pollock, S. M. (1968, December). A Bayesian reliability growth model. *IEEE Trans. Reliability, R-17*(4), 187-198.

ReliaSoft. (1999). *Reliability growth and repariable system data analysis.* San Diego, CA: ReliaSoft .

Rosner, N. (1961). *System analysis - Nonlinear estimation techniques.* Proceeding of the National Symposium on Reliability and Quality Control, Philadelphia, PA.

Selby, J., & Miller, S. (1970). Reliability planning and management (RPM). *ASQC/SRE Seminar,* Milwaukee, WI.

Shirazi, C. & Mosleh, A. (2009). *Correlation between number of experts and accuracy of aggregated estimate.* Paper approved to be presented at the Esrel Conference, Prague, September 7-10, 2009.

Sigurdsson, J., Walls, L., & Quigley, J. (2001). Bayesian belief nets for managing expert judgement and modelling reliability. *Quality and Reliability Engineering International*, 181-190.

Smith, J. (1989). Influence diagrams for statistical modelling. *Annals of Statistics*, 654-672.

Stiber, N., Small, M., & Pantazidou, M. (2004). Site specific updating and aggregation of Bayesian belief network models for multiple experts. *Risk Analysis*, 1529-1538.

Tang, Z., & McCabe, B. (2007). Developing complete conditional probability tables from fractional data for Bayesian belief networks. *Journal of Computing in Civil Engineering*, 265-276.

Trucco, P. (2007). A Bayesian belief network modeling of organizational factors in risk analysis: A case study in maritime transportation. *Reliability Engineering and System Safety*, 1-12.

U.S. Department of Defense. (1981). *Military handbook 189.* Philadelphia, PA: Naval Publications and Forms Center.

Virene, E. P. (1968). *Reliability growth and the upper reliability limit, How to use available data to calculate*, Report No. D2-24142-1, AD-804 150. Seattle, WA: Boeing Company, Seattle, WA.

Walls, L., & Quigley, J. (2001). Building prior distributions to support Bayesian reliability growth modelling using expert judgment. *Reliability Engineering and System Safety*, 117-128.

Weiss, H. K. (1956). Estimation of reliability growth in a complex system with poisson-type failure. *Operations Research, 4*(5), 532-545.

Wilson, A., & Huzurbazar, A. (2007). Bayesian networks for multilevel system reliability. *Reliability Engineering and System Safety*, 1413-1420.

Wolman, W. (1963). *Problems in system reliability analysis*, *Statistical theory of reliability*, Marvin Zelen., ed. Madison, WI: University of Wisconsin Press.

Author Bibliography

Stephen Brown is the Director of Reliability Engineering for Lennox International. In this position Mr. Brown is responsible for global reliability of Lennox products that include residential and commercial HVAC, refrigeration, and furnace products. Mr. Brown has 28 years of progressive engineering experience in facility operations, manufacturing engineering, quality management, as well as product and reliability engineering. Mr. Brown received his BS in Engineering Technology from the University of Tennessee; an MBA from Union University, and is currently pursuing a PhD in Reliability Engineering from the University of Maryland; College Park.

Office Phone 972-497-7396

e-mail steve.m.brown@lennoxind.com