2017

WEIGHTED QUANTILE SUM REGRESSION FOR ANALYZING CORRELATED PREDICTORS ACTING THROUGH A MEDIATION PATHWAY ON A BIOLOGICAL OUTCOME

Bhanu M. Evani
Virginia Commonwealth University

WEIGHTED QUANTILE SUM REGRESSION FOR ANALYZING CORRELATED PREDICTORS ACTING THROUGH A MEDIATION PATHWAY ON A BIOLOGICAL OUTCOME

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

BHANU MURTHY EVANI
B. Tech., Indian Institute of Technology-Madras, Chennai, INDIA
MBA, University of Wisconsin-Milwaukee, WI, USA
M.S., Virginia Commonwealth University, Richmond, VA, USA

Directors: ROBERT A. PERERA, Ph.D.
ASST. PROFESSOR, BIOSTATISTICS, VCU, VA
CHRIS GENNINGS, Ph.D.
PROFESSOR, ICAHN SCHOOL OF MEDICINE AT MOUNT SINAI
DEPT. OF ENVIRONMENTAL MEDICINE AND PUBLIC HEALTH, NY

Virginia Commonwealth University
Richmond, Virginia
April 2017

This research was supported in part by a NIEHS T32 training grant 5 T32 ES 7334-13

Acknowledgements

I thank my adviser Dr. Robert A. Perera and my thesis committee for their guidance. Without their support this body of work would not have been possible. The faculty and staff of the department of Biostatistics offered me educational opportunities which enabled me to conduct my studies, especially the Biostatistics Linux Cluster which was an asset in completing tedious computations that drove the simulation studies. I am grateful to the university library resources, and the writing center staff who reviewed my technical writing for this thesis.

Standing steadfast behind my efforts were family and friends led by my wife, who partnered with me since 1984. She has been an integral part of the VCU Health System Physician Network for the past two decades, constantly motivating me to accomplish my goals and being a shining example of a compassionate human being. This endeavor is directed to our children, who graduated from VCU and UVA, and are poised to excel beyond our own accomplishments, to represent the future. I dedicate this work to Dr. & Mrs. Evani (my parents), for having been an inspiration and for paving the path for my education.

I thank Dr. Roy Sabo, who saw me in an intellectual struggle with Advanced Statistical Inference and found it within himself to share his lecture notes from when he was a graduate student. Such gestures of compassion from several self-appointed mentors helped me succeed. A sincere thank you to all of you.

Table of Contents **Page**

List of Tables

Chapter 2

Chapter 4

Chapter 5

List of Figures

WQS 3-variable statistical analysis

OLS Method 2-Variable Mediation Analysis

OLS Method 3-Variable Mediation Analysis

LASSO Method 2-Variable Mediation Analysis

Pairwise comparison of 2-variable performance for WQS, LASSO & OLS

Chapter 5

Appendix-1

Appendix-2

Appendix-3

Abstract

WEIGHTED QUANTILE SUM REGRESSION FOR ANALYZING CORRELATED PREDICTORS ACTING THROUGH A MEDIATION PATHWAY ON A BIOLOGICAL OUTCOME

By

Bhanu M. Evani, Ph.D.

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2017.

Major Director: Robert A. Perera, Asst. Professor, Department of Biostatistics

This work examines mediated effects of a set of correlated predictors using the recently developed Weighted Quantile Sum (WQS) regression method. Traditionally, mediation analysis has been conducted using the multiple regression method, first proposed by Baron and Kenny (1986), which has since been advanced by several authors like MacKinnon (2008).

Mediation analysis of a highly correlated predictor set is challenging due to the condition of multicollinearity. Weighted Quantile Sum (WQS) regression can be used as an alternative method to analyze the mediated effects, when predictor correlations are

high. As part of the WQS method, a weighted quartile sum index (WQS$_{index}$) is computed to represent the predictor set as an entity. The predictor variables in classic mediation are then replaced with the WQS$_{index}$, allowing for the estimation of the total indirect effect between all the predictors and the outcome. Predictors having a high relative importance in their association with the outcome can be identified by examining the empirical weights for the individual predictors estimated by the WQS regression method. Other constrained optimization methods (e.g. LASSO) focus on reducing dimensionality of the correlated predictors to reduce multicollinearity.

WQS regression in the context of mediation is studied using Monte Carlo simulation for mediation models with two and three correlated predictors. WQS regression's performance is compared to the classic OLS multiple regression and the regularized LASSO regression methods. An application of these three methods to the National Health and Nutrition Examination Survey (NHANES) dataset examines the effect of serum concentrations of Polychlorinated Biphenyls (independent variables) on the liver enzyme, alanine aminotransferase ALT (outcome), with chromosomal telomere length as a potential mediator.

Keywords: Multicollinearity, Weighted Quantile Sum Regression, Mediation Analysis

Vita

Name: Bhanu M. Evani B.Tech., MBA, M.S., Ph.D.

Date of Birth: October 06, 1954

Place of Birth: Masulipatam, Andhra Pradesh, India

Citizenship: India, USA

High School: Campion School, Mumbai, India (1972)

Undergraduate Degree: Bachelor's in Mechanical Engineering, graduated in First Class, from the Indian Institute of Technology - Madras, Chennai, India (1978).

Postgraduate Degree: Masters in Business Administration (MBA, 1980) from the University of Wisconsin – Milwaukee, Wisconsin, USA.

Postgraduate Degree: Master of Science in Mathematical Sciences (M.S., 1989), from Virginia Commonwealth University, Richmond, Virginia, USA.

Postgraduate Degree: Doctor of Philosophy in Biostatistics (Ph.D., 2017), from the School of Medicine, Virginia Commonwealth University, Richmond, VA, USA.

Work Experiences: Teaching Assistant (1/2 time) Undergraduate course in Business Computing 210, (1978 – 1980).

Philip Morris USA – Altria Client Services (Mechanical Engineer 1980 – Manager, Industrial and Quality Engineering in 2009) completed 30 years in the industry. Awarded US patents in Engineering and Product Innovation and four promotions.

VCU School of Dentistry, Medical College of Virginia – VCU, Research Assistant (2012- 2013). Published as co-author in Pediatrics (2013). Tegwyn H. Brickhouse, DDS, PhD[1], Robin R. Haldiman, CEO[2], Bhanu Evani, MS[3]; "The impact of a home visiting program on children's utilization of dental services".

1 Introduction

A mediated effect is traced from a predictor through an intermediate variable along a causal path to the outcome. Mediation analysis sheds light on *how* a predictor produces its effect and enables researchers and clinicians to affect outcomes. As an example, the Multiple Risk Factor Intervention Trial (MRFIT) conducted by a research group in 1990, attempted to design effective prevention programs (independent variables) that would have an effect on smoking, cholesterol and blood pressure levels (mediating variables), which clinicians have hypothesized to have a causative effect on the outcome of heart disease (NHLBI,NIH). A simplified mediation model would be to solve the problem with just one intervening variable, considered as the mediator, with other variables relevant to the model included as covariates, to predict a single outcome. Covariates are included in the mediation model to increase the accuracy of the predictor's effect on the mediator and the outcome by reducing variable omission bias.

This study of mediation analysis starts with Baron and Kenny's (1986) seminal paper, which uses multiple regression as the statistical analysis tool to estimate the regression coefficients used to compute the mediated effect. However, traditional regression methods produce unstable coefficient estimates in conditions of high multicollinearity. These estimates lack precision (i.e. large estimate *s.e.*), and the unstable regression coefficient estimates change in magnitude and sometimes their sign, with small changes in the covariates or minor data perturbations.

A variable selection method that reduces the effects of multicollinearity in the design matrix (***X***) is the Least Absolute Shrinkage and Selection Operator method called LASSO regression (Tibshirani, 1993)**.** Zhou and Hastie (2005) have written about the strengths and weaknesses of LASSO method's variable selection from a correlated cluster of predictors as determined by the constrained optimization function. The LASSO regression method provides additional control for the analyst to set weights favoring a certain independent variable for inclusion into the final solution set based on some design criteria. The emphasis in LASSO is to reduce the effects of multicollinearity by reducing the number of correlated predictors in the dataset. A tuning parameter controls the amount of shrinkage for all the regression coefficients, reducing them all towards zero. Absent any *a priori* knowledge about the variable's association with the other variables in the model, LASSO admits variables based on a MSE criteria for the final model selection, which is determined by n-fold cross-validation. What needs to be determined for this method in the context of mediation is whether individual predictors are admitted based on the strength of their association with the outcome or primarily by their correlations with other variables in the regression model.

When conducting mediation analysis researchers are interested in identifying all the predictors from an initial variable set that have a significant indirect effect on the outcome. For regression in general, the method which could provide more precise regression coefficient estimates in the presence of multicollinearity, and deliver the predictor sub-set which is purported to have a higher relative importance in its association with the outcome is the Weighted Quantile Sum (WQS) Regression, discussed in Carrico's dissertation (2013). Extensive simulation studies done in

characterizing the performance of WQS as a regression method have shown it to provide greater stability for the regression estimates within the context of multicollinearity. Carrico's thesis showed the WQS method as having high sensitivity and improved selectivity for independent variables associated with the outcome in regression, limited only by small effect sizes and high pairwise predictor correlations.

Chapter 2 reviews the current literature on recognizing and dealing with the condition of multicollinearity in regression models, and the methods one could use to analyze mediated effects in a multicollinear dataset. Current literature on mediation analysis methods is reviewed to answer the question of *the mechanism by which* a predictor affects the outcome. This research investigates alternative methods to the traditional multiple regression approach and contrasts it against LASSO which addresses the issues associated with multicollinearity through dimensionality reduction. A more recently developed Weighted Quantile Sum Regression (WQS), deals with clusters of correlated predictors and identifies those predictors that have an association strength above a chosen cut-off value with the outcome. A related area of research is to explore statistical methods for testing the significance of the mediated effects. Comparative strengths and weaknesses of the OLS, LASSO, and WQS regression methods as published in the literature are also discussed. The goal of this review is to outline an unsolved statistical problem of general interest in mediated effects from correlated predictors, using currently developed regression methods and define the subsequent scope of work.

Chapter 3 is devoted to the design of the simulation study, starting with the two predictors and advancing to three predictor population models, by defining their model

parameters. Using simulation studies the strengths and weaknesses are assessed for the WQS, LASSO, and OLS regression methods, and possibly select a preferred method for use in analyzing a given dataset for mediated effects, given the degree of multicollinearity that exists amongst the predictors, the sample size and the magnitude of the component effects of mediation. The simulation design chapter ends with a proposed method for comparing the performances of WQS, LASSO and OLS regressions in mediation.

Chapter 4 documents the total and individual predictor indirect and direct effect inferences drawn from analyzing the simulation study. Chapter 5 provides an application of the three regression methods in mediation, using a National Health and Nutrition Examination Survey data set (NHANES, 2001-2002). Chapter 6 summarizes the thesis findings, outlines the limitations of the study and suggests a direction for future work in this area of study.

The novelty of this research is to extend the WQS method by its ability to analyze mediation models with correlated clusters of predictors and to compare these results to other contemporary methods of analysis. The goal is to identify the subset of individual predictors that are relevant to the proposed mediation model, and to define each predictor's association with other correlated predictors, the intermediate variable and the outcome.

2 Literature Review

2.1 Multicollinearity

Multicollinearity exists in regression when a predictor's variance is explained, in a large part, by the other variables in the model. Such a dataset having correlated independent variables is termed to be an "*ill-conditioned*" dataset (Belsley, 2006). Published literature suggests that multicollinearity has little impact on the overall regression model and test statistics, e.g. the model's R^2 or *MSE*, but a large impact on the individual predictor's regression estimate and test statistics, e.g. imprecise estimates due to high standard errors caused by the condition of multicollinearity. A statistical method which tests the strength of an individual predictor's effect, in order to include or exclude that variable from the model, can be unreliable if the data used for the test is ill-conditioned, due to reduced statistical power (Gunst & Mason, 1980; Marquardt & Snee, 1975). Problems arising from the presence of multicollinearity in regression models include: a) decreased statistical power due to an inflated standard error of the individual effect's estimate and b) variable exclusion bias in the regression estimates, if the objective is to reduce multicollinearity by deleting correlated variables from the regression model (Graham, 2003). Multicollinearity makes the estimates of the regression coefficients β unreliable i.e. the slope parameter of the regression line is not estimable with precision. Farrar and Glauber (1967) place multicollinearity in the context of a spectrum, with multicollinearity being at the opposite extreme from orthogonality, by introducing a definition of the "degree of multicollinearity" as being

a departure from orthogonality for columns of predictor values in the design matrix. Two of the common methods used for measuring the effect of other predictors on a regression coefficient are: 1) calculating the degree of multicollinearity as measured by the variance inflation factor *VIF* for each predictor variable's regression coefficient estimate and 2) finding the eigenvalues or characteristic roots of the correlation matrix ***X`X*** and taking the square root of the ratio of the largest eigenvalue to the smallest eigenvalue, to get the condition number *k*, which is compared to a threshold. The VIF coefficient for the j^{th} predictor is $VIF_j = \frac{1}{1-R_j^2}$, where R_j^2 is the coefficient of determination for predictor X_j, obtained by regressing X_j on all other predictor variables in the model. When R_j^2 increases, the denominator decreases and subsequently the variance inflation factor VIF_j for the j^{th} independent variable becomes increasingly large and the standard error of regression coefficient $\hat{\beta}_j$ is inflated by $\sqrt{VIF_j}$. VIF_j is the multiple for $\hat{\beta}_j$'s variance had predictor X_j been uncorrelated with all other predictors in the model (Neter, Wasserman, & Kunter, 1990).

An example of the VIF for three independent variables is shown in Table 2.1. Let $\mathbf{r_{jj}}$ be the 3x3 bivariate correlation matrix for standardized independent variables X_1, X_2, and X_3. The variance inflation factors VIF_j of these standardized independent variables fall along the diagonal of $\mathbf{r_{jj}^{-1}}$, (the inverse correlation matrix), given by the relationship $VIF_j = Diag\left(r_{jj}^{-1}\right)$ (Mansfield & Helms, 1982).

Table 2.1

Correlation matrix r_{jj} of standardized independent variables for X_1, X_2 and X_3

r_{jj}	X_1	X_2	X_3
X_1	1	0.90	0.40
X_2		1	0.25
X_3			1

r_{jj}^{-1}	X_1	X_2	X_3
X_1	**6.40**	-1.27	-0.35
X_2		**5.70**	0.03
X_3			**1.30**

The decision to declare a VIF_k as high is subjective, but typically VIF_k >10 is troublesome and is considered to represent a condition of multicollinearity (Kutner, Nachtsheim, & Neter, 2004). A $VIF_{X_1} = 6.40$ means that the standard error of the regression coefficient for X_1 is inflated $\sqrt{6.4} = 2.5$ times as compared to what it would have been had X_1 been uncorrelated with the other predictors, X_2 and X_3. Hair et al. (2009) suggest threshold values for the condition number *k* (derived from the eigenvalue solutions of the bivariate correlation matrix for the independent variables), which falls within the range of 15 to 30, with 30 being the most commonly used threshold for high multicollinearity. A comprehensive assessment of the condition of multicollinearity is not possible by examining only the correlation matrix (showing predictor bivariate correlations). To diagnose the amount of multicollinearity present and to identify the regressors exhibiting high multicollinearity, a regression coefficient variance-decomposition matrix is constructed showing the proportion of variance for each regression coefficient (its associated predictor) attributable to each condition index (relative magnitude of the eigenvalues of the correlation matrix) above a commonly used threshold of 30. For all condition indices exceeding the threshold of 30 identify the

predictors with variance proportions above 90%. A collinearity problem is indicated when a condition index identified in the first step (>30), accounts for a substantial proportion of the overall variance of Y (>0.90) for ***two or more*** coefficients.

The consequences of a high VIF or condition number k, associated with a predictor's coefficient estimate are: 1) reduced power for the predictor i.e. its effect might fail to achieve statistical significance in a high multicollinear setting 2) the regression coefficient is unstable having poorly estimated effects because of its changing magnitude and sign resulting from minor changes to the dataset The condition of multicollinearity makes each predictor's role in the model difficult to interpret.

2.2 Mediation Analysis

Some research studies stop at establishing a significant association between study variables. Taking the hypothesis a step further would be to understand the mechanism through which the independent variable (*X*) affects the response (*Y*). Finding a significant indirect effect adds further insight into the mechanism of the action of *X*'s influence on *Y*. The indirect effect of *X* acting through *M* on *Y*, is the study of mediation.

The work of Sewall Wright (1918, 1934) on estimating causal path coefficients by decomposing the correlations among the variables in a path diagram, could be the origin of causal and mediation analysis. A causal relationship is represented by a unidirectional arrow between the predictor and response, while a non-causal relationship, e.g. the correlation between predictors (X_1, X_2) is represented by a two-headed curvilinear arrow. Residual variables are shown with their respective response

variables having model error variances. Given the correct model specification, the covariance between two variables in a path diagram can be expressed using covariance algebra. Alternatively, the covariance between variables in a path diagram can be expressed as the product of the path coefficients for each segment, and summed over the alternate paths traced between them, while conforming to the tracing rules. This established the foundation for generating simulated data for correlated predictors using either the estimated regression coefficients, or equivalently by using the empirical correlation matrix.

In the mediation model X→M→Y, the total effect of predictor *X* on the outcome *Y*, can be shown using tracing rules to be the sum of the direct effect of *X* on *Y*, and *X*'s indirect effect on the outcome *Y* through the mediating variable *M*. The indirect effect being the product of the path coefficients for the two segments X→M and M→Y.

2.2.1 Classic Mediation Analysis- OLS Multiple Regression

Baron and Kenny (1986) presented the regression-based approach for testing mediation patterns given by the regression of mediator *M* on just the predictor *X* and the regression of output *Y* on predictor *X* adjusting for mediator *M*,

$$M = \alpha_1 + \theta X + \varepsilon_1 \qquad 2.1$$
$$Y = \alpha_2 + \beta X + \gamma M + \varepsilon_2 \qquad 2.2$$

shown in equation 2.1 and 2.2. When *Y* is regressed on *X* alone (absent the mediator variable), the total effect of *X* on *Y* is defined by

$$Y = \alpha_3 + \beta_0 X + \varepsilon_3 \,. \qquad 2.3$$

They proposed that if $\beta_0, \theta \text{ and } \gamma$ were each significant regression coefficients, then a mediated effect is present provided β_0 becomes significantly diminished (as compared to β) in the presence of the mediating variable *M*, i.e. $(\beta < \beta_0)$, due to a significant indirect effect being present from predictor *X* acting through the mediator *M* on output *Y*. The influence of *X* on *Y* is completely mediated if the null hypothesis $H_0 : \beta = 0$ fails to be rejected, indicating that no residual direct effect is present with mediator *M* in the model. Baron and Kenny (1986) used multiple regression to provide evidence of mediation.

Wright (1920, 1921) displayed the relationships between variables in a causal model using path diagrams which are now referred to as Directed Acyclical Graphs. He used tracing rules and path analysis to generate quantitative estimates for path coefficients based on the observed correlations between variables. This provided a quantitative estimate for the indirect effect as being the product of the path segment's regression coefficients $\hat{\theta} \text{ and } \hat{\gamma}$, as an alternative method for providing evidence of mediation. Regressing *M* on *X* yields the regression coefficient $\hat{\theta}$, represented by the slope $\frac{\Delta M}{\Delta X}$ of the regression line, and regressing *Y* on *M* adjusting for *X*, yields $\hat{\gamma}$, represented by the slope $\frac{\Delta(Y \mid X, M_1 - Y \mid X, M_2)}{\Delta M}$ of the regression line. Therefore, the product of the path coefficients $\hat{\theta}\hat{\gamma}$ is equivalent to taking the product of the slopes $\frac{\Delta M}{\Delta X} . \frac{\Delta(Y \mid X, M_1 - Y \mid X, M_2)}{\Delta M} = \frac{\Delta(Y \mid X_1, M_1 - Y \mid X_1, M_2) - \Delta(Y \mid X_2, M_1 - Y \mid X_2, M_2)}{\Delta X}$, which represents $\frac{\Delta(Y \mid X_1, M_1 - Y \mid X_2, M_1) - \Delta(Y \mid X_1, M_2 - Y \mid X_2, M_2)}{\Delta X} = \frac{\Delta(Y \mid X_1, M - Y \mid X_2, M)}{\Delta X}$ the

regression of *Y* on *X* adjusting for *M* $\frac{\Delta(Y \mid M)}{\Delta X}$. MacKinnon et al. (1995) showed that a beta coefficient difference $\left(\hat{\beta}_0 - \hat{\beta}\right)$ is algebraically equivalent to the product $\hat{\theta}\hat{\gamma}$ obtained from ordinary least-squares regression. Therefore, rejecting either the null hypothesis $H_b:(\beta_0 - \beta) = 0$ or $H_c:\ \theta\gamma = 0$ is identical.

2.2.2 Testing for Indirect and Direct Effects

The literature categorizes methods for testing the statistical significance of mediated effects into four categories: 1) Normal distributed assumption for $\hat{\theta}\hat{\gamma}$ while using the *z*-test for H_0: $\theta\gamma = 0$ (Sobel, 1982); 2) Joint significance tests for $\hat{\theta}$ and $\hat{\gamma}$ where H_0: $\left(\theta = 0 \text{ and } \gamma = 0\right)$(Cohen & Cohen, 1983; Kenny et al., 1998); 3) bootstrap data resampling methods to address the asymmetric distribution of the product of two regression coefficients $\hat{\theta}\hat{\gamma}$ (Shrout & Bolger, 2002) and 4) Indirect effect's confidence limits from the distribution of the product of regression coefficients (MacKinnon, Fritz, Williams, & Lockwood, 2007).

MacKinnon, Lockwood, Hoffman West and Sheets (2002) present an extensive comparison of several historical and existing methods for testing mediation and other intervening variable effects. The method uses two regression equations to test for the significant indirect effect.

$$M = \alpha_5 + \sum_{k=1}^{p} \theta_k X_k + \varepsilon_5 \qquad 2.4$$

$$Y = \alpha_6 + \sum_{k=1}^{p} \beta_k X_k + \gamma M + \varepsilon_6 \qquad 2.5$$

Baron and Kenny (1986) suggested using Sobel's *z*-test statistic, $z = \frac{\hat{\theta}\hat{\gamma}}{\sqrt{\hat{\theta}^2 s_{\hat{\gamma}}^2 + \hat{\gamma}^2 s_{\hat{\theta}}^2}}$,

$\hat{\theta}$ and $\hat{\gamma}$ obtained from equations 2.4 and 2.5, to test the indirect effect estimate for a mediated effect. Sobel (1982) used the multivariate delta method to calculate the asymptotic standard error $s_{\hat{\theta}\hat{\gamma}} = \sqrt{\hat{\theta}^2 s_{\hat{\gamma}}^2 + \hat{\gamma}^2 s_{\hat{\theta}}^2}$ after dropping the usually small covariance term $s_{\hat{\theta}}^2 s_{\hat{\gamma}}^2$, which Aroian (1944) had suggested as an exact standard error $s_{\hat{\theta}\hat{\gamma}} = \sqrt{\hat{\theta}^2 s_{\hat{\gamma}}^2 + \hat{\gamma}^2 s_{\hat{\theta}}^2 + s_{\hat{\theta}}^2 s_{\hat{\gamma}}^2}$ based on the Taylor series expansion for the product $\hat{\theta}\hat{\gamma}$. Goodman (1960), Sampson and Breunig (1971) suggested an unbiased standard error of the product of two normally distributed variables, which subtracted the product of the sample variances $s_{\hat{\theta}\hat{\gamma}} = \sqrt{\hat{\theta}^2 s_{\hat{\gamma}}^2 + \hat{\gamma}^2 s_{\hat{\theta}}^2 - s_{\hat{\theta}}^2 s_{\hat{\gamma}}^2}$. These constitute alternative *z*-statistics $\left(\frac{\hat{\theta}\hat{\gamma}}{s_{\hat{\theta}\hat{\gamma}}} \right)$ to test $H_0 : \theta\gamma = 0$, given $s_{\hat{\theta}}^2$ and $s_{\hat{\gamma}}^2$ are estimate variances for the non-standardized regression estimates obtained from equations 2.1 and 2.2.

When the distribution for the indirect effect is heavy-tailed (small sample sizes) and skewed because the sampling distribution of the indirect effect is non-normal, MacKinnon and Dwyer (1993) showed that Sobel's z-statistic for the indirect effect has reduced statistical power to detect the indirect effect. MacKinnon et al. (1998) showed that any z_statistic used to test the significance of an indirect effect has low statistical power because the distribution of the product of regression coefficients is not normally distributed, but is skewed with high kurtosis. Shrout and Bolger (2002) found that the product of two normally distributed variables with positive means forms a sampling

distribution which tends to be positively skewed, while the product of two normally distributed variables having means of opposite signs, are typically negatively skewed. These findings favor a bootstrap data resampling method for calculating the confidence interval for $\hat{\theta}\hat{\gamma}$, to address the non-Gaussian distribution of the product of regression coefficients.

Bollen & Stine (1990) and Lockwood & MacKinnon (1998) found by using simulated data and bootstrap analysis that the product method's indirect effect confidence intervals tend to lie to the left of the true value for positive effects and to the right for negative effects, yielding a low coverage probability. Non-parametric bootstrapped confidence limits of the product $\hat{\theta}\hat{\gamma}$, when used to test the null hypothesis $H_0 : \theta\gamma = 0$, have better power than tests which assume a normal distribution for the indirect effect (MacKinnon et al. 2004). Based on these findings, the preference is for bootstrapped data generated percentile confidence limits to test for the significance of the indirect effect estimate.

2.2.3 The Study of Mediation Analysis Utilizing Competing Regression Methods

A preliminary step to analyzing a mediation model is to visually encode the correlations between observed variables using a Directed Acyclic Graph (DAG) incorporating *a priori* knowledge of the relationship between the model variables and the utilizing the rules for drawing Directed Acyclic Graphs (DAGs). The diagram separates 1) the causal relationships X→Y, between independent variable's (X) direct influence on the outcome variable (Y), and 2) the confounding variables (C) which affects both *X* and *Y* in a non-causal relationship X←C→Y, giving the appearance of an association

between X and Y, which is visible in a scatter plot of their observed values. Researchers are interested in quantifying the strengths of causal relationships identified in a DAG structure after adjusting for the minimally sufficient set of covariates related to the confounding variables.

OLS regression can be applied to analyze the mediation model using multiple regression equations. The multi-predictor mediation regression equations are:

$$M = \alpha_1 + \theta_j \sum_{j=1}^{p} X_j + \varepsilon_1 \qquad j \in \{1..p\} \qquad 2.6$$

$$\mathrm{Y} = \alpha_2 + \beta_j \sum_{j=1}^{p} X_j + \gamma M + \varepsilon_2 \qquad j \in \{1..p\}\ . \qquad 2.7$$

Indirect effects defined as the product $\theta_j \gamma$ and the direct effects defined by β_j from equations 2.6 and 2.7, can be estimated and the hypothesis $H_{01} : \theta_j \gamma = 0$ used to test for any significant indirect effects $j \in \{1..p\}$.

When conducting mediation analysis with multiple correlated predictors, each predictor's (*X's)* regression coefficient estimate has a VIF associated with it. Uncertainty in the estimate brought on by small perturbations of the model, like a covariate change or small changes in the dataset causes large changes in either the coefficient estimate's magnitude, direction, or both. The purpose for using a particular method in mediation analysis is to identify the subset of the predictors that are significantly associated with the outcome through the mediating variable. This requirement calls for methods that can differentiate between highly correlated independent variables that have an association with variables along the causal chain, which will define the subset of variables that have a significant effect on the outcome through the mediating variable. The adverse effects

of multicollinearity are a drawback for OLS regression in estimating the indirect and direct effects.

Alternative regression methods to OLS produce biased estimates but have reduced standard errors for their regression estimates. Although the Bias-Variance trade-off usually refers to the consequence of adding predictors to a regression model, the same underlying concept can be used to consider alternative methods which trade some estimate bias for reduced estimate variance. OLS regression is at one end of this bias-variance trade-off spectrum, since it produces unbiased estimates with inflated variance in the presence of multicollinearity. The mean-squared error statistic for a parameter estimate $\hat{\theta}$ is defined by $MSE_{\theta} =$ (Bias^2 + Variance) $= (E_{\theta}(\hat{\theta}) - \theta)^2 + E_{\theta}(\hat{\theta} - \theta)^2$ and in the case of an OLS regression producing an unbiased estimate, the MSE would consist entirely of the estimate's variability which is influenced by the predictor's VIF.

A methodological alternative to solving the multiple regression equations using OLS in the mediation model is to construct the structural equations model (SEM). In SEM the implied covariance matrix (model) is compared to the observed covariance matrix (data) using a goodness-of-fit statistic, to test the validity of the causal models entered in the structural equations. However, one cannot test this model if full mediation exists i.e. if the direct influence of *X*→ *Y* goes to zero in the presence of the mediating variable, since the model gets saturated. The SEM approach allows for an easier analysis of more complex mediation models and is deferred to future research.

Carrico et al. (2013, 2015) have shown that WQS handles collinearity issues well except in conditions with a combination of small sample sizes, small effect sizes and high multicollinearity. The WQS algorithm incorporates a bootstrapped data step when

determining the empirical predictor weights, which improves the method's sensitivity with regards variable selection without compromising specificity. Unlike the WQS method, (Carrico et al. 2013) found that LASSO regression has higher sensitivity but is often accompanied by lower specificity. This provides motivation to study mediation using OLS, WQS and LASSO regression methods to understand how each method deals with multicollinearity. The measure of quality for the three regression methods could be the regression estimate's Mean Squared Error since it combines estimate bias and the estimate variance into a single statistic.

2.3 Least Absolute Shrinkage Selection Operator (LASSO) Regression

A regularization method can reduce the effects of multicollinearity in regression by retaining a few variables from a correlated group of variables, while shrinking the rest of the variables' regression coefficients within the group to be zero. Removing some of the highly correlated variables from the regression model, introduces model specification bias in the regression estimates and the excluded variable's effect on the remaining variables is not accounted for when estimating the remaining regression coefficients. However, a benefit of reducing dimensionality in the number of correlated predictors is less multicollinearity in the design matrix ($\boldsymbol{X}$). Collection of more data is another solution to dealing with multicollinearity in multiple regression but in many cases it is either impracticable or prohibitively expensive. Common regularization methods to consider are Ridge regression, LASSO regression and the Elastic Net. In Ridge regression method (Hoerl & Kennard 1970) the coefficient estimates are shrunk towards zero and the regularized estimates have reduced variability but there is no

dimensionality reduction. This is accomplished by augmenting the diagonal of the $\boldsymbol{X^T X}$ matrix with small positive quantities which allows a nearly singular $\boldsymbol{X^T X}$ matrix to stabilize, while remaining positive definite, so that the regression estimate vector $\hat{\beta}^{ridge}$ can be solved for. Ridge regression imposes an upper, positive-valued constraint on the coefficients using the tuning parameter t, on the sum of the squared regression coefficients $\sum_{k=1}^{p} \beta_k^2 < t$ for t>0. The Ridge regression method exhibits a grouping effect for highly correlated variables, where the regression coefficients for variables within the correlated sub-group tend to be similar (Zhou & Hastie, 2005). In such cases, Ridge regression does not allow variable selection since the relative importance between these predictors based on their regression coefficients, does not allow for much differentiation.

The LASSO (Tibshirani 1996) penalized regression method could achieve a lower regression coefficient's MSE than its unbiased OLS regression estimate, when the condition of multicollinearity is high, by increasing the estimate's bias through penalization, for a larger reduction in the estimate's variability. The Least Absolute Shrinkage Selection Operator (LASSO) method constrains the sum of the absolute regression coefficients to be below a positive-valued tuning parameter value t, $\sum_{k=1}^{p} |\beta_k| < t$ while simultaneously minimizing the sum of the squared residuals (Hastie, Tibshirani, & Friedman, 2009). This method shrinks some of the regression estimates to zero thus excluding them from the final model.

$$\hat{\beta}^{Lasso} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij} \beta_j \right)^2 \quad \text{subject to} \sum_{j=1}^{p} \left| \beta_j \right| \le t \qquad 2.8$$

with $\hat{\beta}^{Lasso}$ being the solution vector of LASSO regression coefficient estimates. This method shrinks the weak regression estimates to zero, thus eliminating some of the correlated variables and consequently reducing the effects of multicollinearity. The optimal tuning parameter *t* is obtained from an *n*-fold cross validation procedure.

$$\hat{\beta}^{Lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right| \right\} \qquad 2.9$$

Equation 2.9 is similar to equation 2.8 but the constraint has been written in the Lagrange multiplier form to enable determination of the optimum value for the Lagrange multiplier λ . The LASSO regression coefficients associated with the minimum average n-fold cross-validation prediction error and the more conservative solution associated the minimum + 1 *s.e.*, are usually reported. Within a cluster of highly correlated predictors, the LASSO algorithm admits the predictor having the strongest association with the outcome, while shrinking the remaining regression coefficients in that correlated cluster to zero (Zou & Hastie, 2005). The number of non-zero regression coefficients in the final LASSO model can be controlled by selecting the tuning parameter *t* in equation 2.8. Multicollinearity is reduced because of the excluded correlated predictors at the expense of an increased regression coefficient bias for those included variables in the final LASSO solution set. The problems with LASSO's method of variable selection are: the excluded variables' covariances affect the admitted predictor's regression estimate bias and the excluded variables could also have an important association with the response, which is not reported. The Elastic Net's penalization method attempts to capture the benefits of Ridge and LASSO regressions, but suffers from the "grouping" effect, where the variables in a correlated cluster could all have their regression

coefficients shrunk to zero simultaneously, or be admitted as a group into the final solution set (Zou & Hastie, 2005). The grouping effects are characteristic of the Ridge regression and the Elastic Net methods. When conducting mediation analysis accurate variable selection is more important than accurate model predictions for future datasets. This makes the LASSO regression the preferred method for a comparative performance when analyzing mediation models, over the other two methods.

2.4 Weighted Quantile Sum (WQS) Regression

The Weighted Quantile Sum regression method was proposed and developed by Gennings, Sabo and Carney (2010), Christensen, Carrico, Sanyal and Gennings (2013), and Carrico, Gennings, Wheeler and Factor-Litvak (2015). WQS is a constrained optimization multiple regression method, which uses the data to estimate the individual predictor weights to determine their relative importance in selecting the most important subset of predictors that influence the response. Feldman (2005) introduced “the proportional marginal variance decomposition” using data driven weights, to decompose the full model R^2 into contributions from the different regressors as a way of assigning shares of relative importance to each regressor in the multiple regression model.

In the WQS method, the independent variable values are scored into quartiles to make predictor values independent of the scales used, to reduce some multicollinearity and to address any outliers which could influence the process of selecting the variables associated with the outcome, from a set of correlated predictors. Scoring in quartiles allows predictors on different scales to be merged into a WQS index. Quartiles are

denoted by q_j taking on values 0,1,2,3 representing the 1st to 4th quartiles for $j \in \{1..p\}$ predictors in each observation. The WQS model is defined by

$$g(\mu) = \beta_0 + \beta_1\left(\sum_{j=1}^{p} w_j q_{ij}\right) + \gamma' M \qquad 2.10$$

where w_j is the weight parameter for the j^{th} predictor, β_1 is the regression slope parameter for the weighted sum of the predictors, for each observation represented by $\sum_{j=1}^{p} w_j q_{ij}$; *M* is a vector of covariates and γ is the corresponding vector of regression coefficients, and g is any monotonic differentiable link function as in a generalized linear model which links the outcome to the predictor variables. The WQS weights for all the predictors in a dataset are constrained to sum to 1, $\sum_{j=1}^{p} w_j = 1$, the regression coefficient β_1 is constrained to be either positive or negative i.e. $\beta_1 > 0$ or $\beta_1 < 0$ based on the direction of association between the set of predictors and the outcome that the researcher is investigating, and each weight is bound by the limits $0 \le w_j \le 1$.

The estimator in equation 2.10 can be written in the least squares form or in a maximum likelihood form. The least squares parameter estimates and the WQS predictor weights are obtained from

$$\hat{\theta}_{WQS} = \arg\min_{\theta} \sum_{i=1}^{N}\left(y_i - \left(\beta_0 + \beta_1\left(\sum_{j=1}^{p} w_j q_j\right) + \gamma' M\right)\right)^2 ; \sum_{j=1}^{p} w_j = 1 \text{ and } \beta_1 > 0 \text{ or } \beta_1 < 0 \qquad 2.11$$

where θ is the vector of the unknown parameters and $\hat{\beta}_1$ is the regression slope parameter for the weighted quartile sum index in equation 2.11, under the specified

constraints. Alternatively, the maximum likelihood form can be used to estimate the WQS parameter estimates and weights

$$\hat{\theta}_{WQS} = \arg\max_{\theta}\left[\ln\left(L\left(\theta; y\right)\right)\right] \text{ subject to } \sum_{j=1}^{p} w_j = 1 \text{ and } \beta_1 > 0 \text{ or } \beta_1 < 0 \qquad 2.12$$

where θ represents the unknown parameters in equation 2.12, under the specified constraints. A trust region method of optimization is used since it allows for non-linear inequality constraints $\beta_1 > 0$ or $\beta_1 < 0;$, to be applied to the objective function (Byrd, R., Gilbert, J. & Nocedal, (2000). To increase sensitivity in detecting important predictors through data perturbation as suggested in Meinshausen and Buhlmann (2010), the original dataset is bootstrapped (typically 100 times) and the least squares regression parameters and the predictor WQS weights are estimated for each of the bootstrapped dataset. A pre-specified "signal function" $f\left(\beta_{1(b)}\right)$ may be used with the estimated slope parameters $\hat{\beta}_{1(b)}$ which is a measure of the signal strength associated with estimating the b^{th} set of individual predictor weights $w_{j(b)}$, from the b^{th} bootstrap sample. The weights are used to estimate a weighted quantile sum index $WQS_{index} = \left(\sum_{j=1}^{p} \tilde{w}_j q_j\right)$ where $\tilde{w}_j = f\left(\beta_{1(b)}\right).w_{j(b)}; b \in (1..., B)$. Examples of some possible signal functions have been suggested by Carrico, Gennings, Wheeler et al. (2015). The signal function $f\left(\beta_{1(b)}\right)$ can be defined such that bootstrap samples having a higher signal strength are given a higher relative weight in estimating the individual predictor WQS weights $\tilde{w}_j$. Another definition for the signal function $f\left(\beta_{1(b)}\right)$ could be a test statistic associated with $\beta_{1(b)}$: i.e.,

if $S_{(b)}$ is defined as the test statistic from the *b*th bootstrap sample, then the signal function is $\frac{S_{(b)}}{\sum_{b=1}^{B} S_{(b)}}$. Alternatively, the signal function may be based on an indicator of whether β_1 is significant for the *b*th bootstrap sample. A uniform function may also be used so that all the bootstrapped samples equally influence the estimation of the WQS predictor weights by defining $\tilde{w}_j = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)}$ as the arithmetic average of the individual predictor weights from the *B* bootstrapped datasets.

Carrico et al. (2014) found that variable selection was not affected by the use of different signal functions for $w_{j(b)}$ in $\tilde{w}_j = f\left(\beta_{1(b)}\right).w_{j(b)};\ b \in (1..., B)$ that affect the forming of the WQS_{index}. Carrico et.al (2013, 2015) evaluated and characterized the accuracy of WQS regression in variable selection through extensive simulation studies. WQS had the ability to correctly identify the subset of original correlated predictors that were strongly associated with the outcome, while shrinking the individual weights of those predictors having weak associations with the outcome. WQS regression method in the context of multicollinearity, estimates the predictors' overall effect and in addition identifies those individual predictors with high weights in the WQS_{index} as having a strong association with the outcome.

In summary, WQS regression (2010, 2013 and 2015) is a more recent weighted regression method that has been used for selecting relevant predictor variables from a cluster of highly correlated predictors with small effect sizes of the predictor on the

outcome. The WQS method was characterized by Carrico et al. (2013 and 2015), and more recently was applied to environmental chemical mixtures (Czarnota et al., 2015).

Most importantly the WQS method has been shown to identify additional predictors from the original set of correlated predictors that were associated with the outcome (verified by current scientific literature), as compared with other variable selection methods. WQS method's limitations are in its performance as measured by sensitivity and selectivity, which deteriorate as the predictor correlations increase while sample sizes and effect sizes decrease. Also, the WQS_{index} does not identify associations in different directions because of the constraint on the regression slope coefficient $\beta_1 > 0$ or $\beta_1 < 0$ for the WQS_{index}. However, the paper finds those single predictors that exhibit inverse associations with the outcome as having negligible weights in a positively associated WQS_{index}, with the outcome.

2.5 Literature on the Strengths and Weakness of OLS, WQS and LASSO

The OLS method is widely used to estimate parameter values that fit a function to the observed data and to characterize the statistical properties of the estimates. OLS regression has the drawback of being highly sensitive to influential outliers in the data because the squaring of model errors magnifies the differences between the observed and the predicted values for extreme observations, giving them more leverage in determining the parameter estimates. Techniques that attempt to minimize the effect of such outliers are called robust regression techniques. One important class of such robust techniques is the iterated reweighted least squares (IRLS) method (Huber 1972,1981), which iteratively performs a weighted least squares fit using a chosen

weighting function, whose observation weights are derived from current residuals starting with the OLS residuals until the residuals remain unchanged between consecutive iterations. This procedure brings stability to the unstable slope parameter in OLS regression, and when it is applied to ill-conditioned datasets. Correlated predictors with inflated regression estimate variances are imprecise and unreliable, since small changes in the model covariates or the observed data could drastically change the estimate's magnitude and/or direction. Moreover, classic OLS regression estimators with large MSE's could perhaps be improved on by using other regression techniques, especially when multicollinearity exists.

In WQS regression the weights for predictors that are weakly associated with the outcome are smaller, highlighting their relative importance in their association with the outcome. Carrico et al. (2015) have shown that WQS method has weaknesses in estimating individual predictor effects and in selecting the important predictor variables that are associated with the outcome, especially when the effect sizes are small in the presence of high multicollinearity.

LASSO is a popular variable selection penalized regression method developed by Tibshirani (1996). In this regression method, clusters of predictors exhibiting multicollinearity and having an association with the outcome, are reduced to a parsimonious set of predictors with non-zero regression coefficients. This addresses the multicollinearity issue but the LASSO solution is not unique because of the constraint on the summed absolute values of the regression coefficients which is applied while the sum of squared residuals is being minimized. Alternatively, for large samples under i.i.d. conditions and an outcome variable distributional assumption, the MLE estimates could

be obtained using the same constraint on the summed absolute values of the regression coefficients, imposed on the minimization of the LASSO objective consisting of the negative log likelihood function.

A mediation analysis research study on the comparative performance of OLS, WQS and LASSO methods has not been done to date.

2.6 Research Problem and Scope of Work

The research focus is to evaluate the performance of the Weighted Quantile Sum regression method against the LASSO and OLS regression methods in conducting mediation analysis on data exhibiting a wide rage on multicollinearity. An approach for conducting this research is to first tackle two independent variables, a mediator and an outcome in the mediation model. This could be followed by the study of three correlated predictors, acting through a single mediator M on an outcome Y. The research attempts to determine the advantages and limitations of the WQS method (Carrico et al., 2013; 2015), by studying its performance in mediation and contrasting it against OLS regression and LASSO penalized regression.

Another first step towards understanding the performance of the Weighted Quantile Sum regression method as applied to mediation with multiple correlated predictors, would have the mediator and outcome variables Gaussian distributed. Given that LASSO is a regularization method offering a parsimonious solution set with the benefit of reduced multicollinearity, it is compared to the WQS method from the standpoint of variable selection. Conducting OLS multiple regression on just the LASSO regression admitted variables (having non-zero coefficients), the individual predictor

mediated effects can be estimated for comparison with the regular OLS procedure results. Using the population model, the classic OLS multiple regression method is used to obtain the overall mediated effect and the individual predictor mediated effects. This provides the necessary information to compare WQS and LASSO for accurate variable selection and LASSO results post processed using OLS to compare the individual predictor mediated effects to those from the classic multiple regression method. Since OLS is not a variable selection method, its individual predictor mediated results cannot be compared truly with the WQS method’s list of individual predictors that have an influence on its overall mediated effect.

Based on the patterns of the results seen in the two and three variable cases, recommendations can be made on how to conduct mediation analysis using the WQS method for multiple correlated predictors acting through a mediator on an outcome.

3 Simulation Design

3.1 Population model and Parameters

A Monte Carlo simulation was conducted to determine recommendations for applying WQS to mediation analysis and assess its performance relative to OLS regression and LASSO penalized regression methods. Two separate population models shown in Figures 3.1 and 3.2, each with varying parameters were examined. The first population model with two correlated predictors acting through a mediator affecting an outcome was chosen to understand the effect of collinearity amongst the simplest case of two predictors on a mediated outcome. The second population model with three correlated predictors acting through a mediator affecting an outcome was chosen to understand the effect of multicollinearity amongst the simplest case of a mixture of three predictors on a mediated outcome.

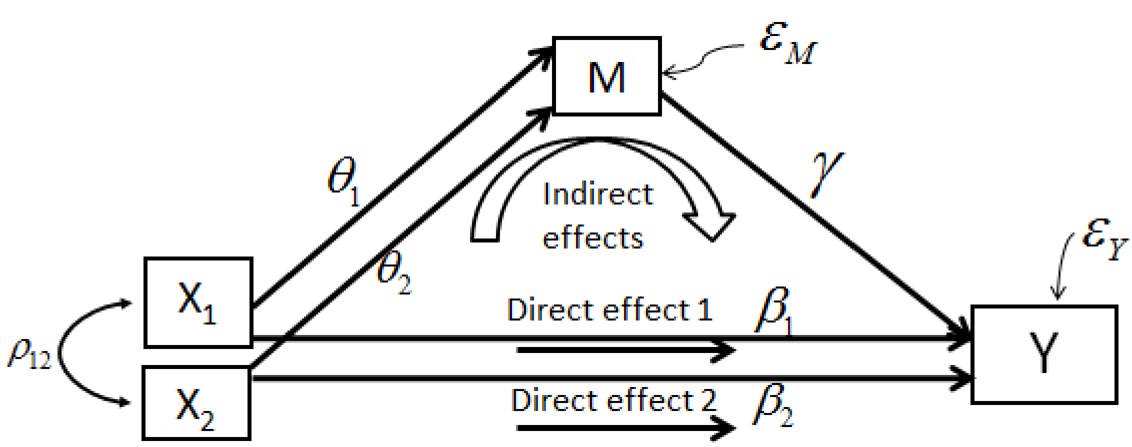


Figure 3.1 Two Correlated Predictors Acting Through Mediator *M* on Outcome *Y*

Each parameter in the path model equations was used to determine the population correlation matrix to generate the p independent variables $X_1 \ldots X_p$, mediator variable *M,* and dependent variable *Y,* following a multivariate normal distribution. The NHANES 2003-04 datasets were used to determine appropriate pairwise correlations between the independent variables; while the regression parameters were chosen such that the pairwise correlations between any *X*, *M* and *Y* were not much greater than 0.5, a typical upper bound for variable associations in a mediation model.

The mediation model with two independent variables had these parameter combinations:

1. Sample size N = 300 and 1000 (2 conditions)
2. Correlation between X_1 and X_2 , ρ_{12} = 0, 0.5, 0.95 (3 conditions)
3. Population parameter set - $(\theta_{1,}\theta_2)$ (0,0), (0,0.15), (0.15,0.30) (3 conditions)
4. Population parameter set - (β_1, β_2) (0.15, 0.30), (0.30, 0.0) (2 conditions)
5. Population parameter γ - 0, 0.10, 0.30 (3 conditions)

These factors 2* 3* 3* 2* 3 equal 108 conditions for the two predictor model. The parameter sets e.g.(β_1, β_2) = (0.15, 0.30) were used as a single combination in order to limit the total number of conditions used in simulation.

The second population model with three correlated predictors acting through a mediator affecting an outcome was chosen to understand the effect of multicollinearity amongst three predictors on a mediated outcome.

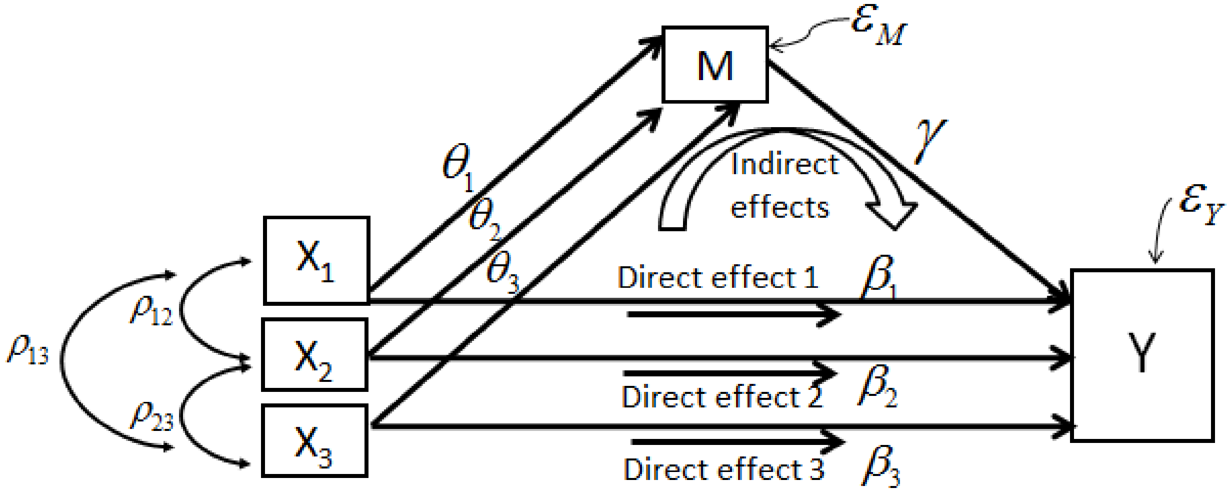


Figure 3.2 Three Correlated Predictor Mixture Acting via Mediator *M* on Outcome *Y*

The mediation model with three independent variables had these parameter combinations:

1. Sample size N = 300 and 1000 (2 conditions)
2. Correlation between X_1X_2, X_2X_3 and X_1X_3 $(\rho_{12}, \rho_{13}, \rho_{23})$ were (0.9, 0.40, 0.25), (0.68, 0.30, 0.19) and (0.45, 0.20, 0.13) (3 conditions)
3. Population parameter set $(\theta_1, \theta_2, \theta_3)$ were set - (0.15, 0.45, 0), (0, 0.25, 0.30), and (0.35, 0, 0.20) (3 conditions)
4. Population parameter set $(\beta_1, \beta_2, \beta_3)$ were set – (0.15, 0.20, 0), (0.30, 0, 0.20), and (0, 0.15, 0.30) (3 conditions)
5. Population parameter γ - 0, 0.25, 0.35 (3 conditions)

These factors 2* 3* 3* 3* 3 equal 162 conditions for the three predictor model. The parameter sets e.g. $(\rho_{12}, \rho_{23}, \rho_{13})$ = (0.9, 0.40, 0.25) were used as a single combination in order to limit the total number of conditions used in simulation.

Conditions were chosen to represent the null hypothesis to enable the evaluation of the Type I error rate, by choosing the theta parameter to be zero, the gamma parameter to be zero or both the theta and gamma parameters to be zero simultaneously. The remaining conditions had non-zero regression parameters to estimate the statistical power for finding a significant indirect or direct effect under varying conditions.

The population models shown Figures 3.1 and 3.2 can each be represented by equations 3.1 and 3.2. The first is the dependent variable mediator *M* regressed on the independent variables $X_1 \ldots X_p$.

$$M = \sum_{j=1}^{p} \theta_j X_j + \varepsilon_M \qquad (3.1)$$

with zero intercept, each regression coefficient, $\theta_j \in \left(\theta_1 \ldots \theta_p\right)$ representing the unique individual effects of an X_j adjusting for the remaining X's on M in the model and ε_M represents the model errors. The second is the dependent variable response *Y* regressed on the independent variables $X_1 \ldots X_p$ and mediator M,

$$Y = \sum_{j=1}^{p} \beta_j X_j + \gamma M + \varepsilon_Y \qquad (3.2)$$

with zero intercept, each regression coefficient, $\beta_j \in \left\{\beta_1 \ldots \beta_p\right\}$ representing the unique individual effects of X_j adjusting for the remaining *X*'s and *M*, on *Y* in the model with model errors ε_Y . Correlated variables for the simulation of two and three predictor data sets were generated using each parameter in the path model equations to determine the population correlation matrix by tracing rules (Wright, 1934) following a multivariate

normal distribution using Cholesky's decomposition with 500 replications of each condition.

3.2 Mediation Analysis Using WQS, LASSO, OLS Methods

Three statistical methods for mediation analysis were considered namely, weighted quantile sum (WQS) regression, least absolute shrinkage selection operator (LASSO) regression, and ordinary least squares (OLS) regression. Within WQS regression three ways of estimating the WQS weights were considered leading to three different estimates of the direct and indirect effects. LASSO regression with two different shrinkage parameter estimates were considered leading to two different analyses; and the standard OLS regression analysis was considered. This leads to six analyses that were examined, three WQS, two LASSO, and the classic OLS multiple regression method.

3.2.1 Weighted Quantile Sum Regression Estimation

The essence of the WQS mediation model is described in Figure 3.3.

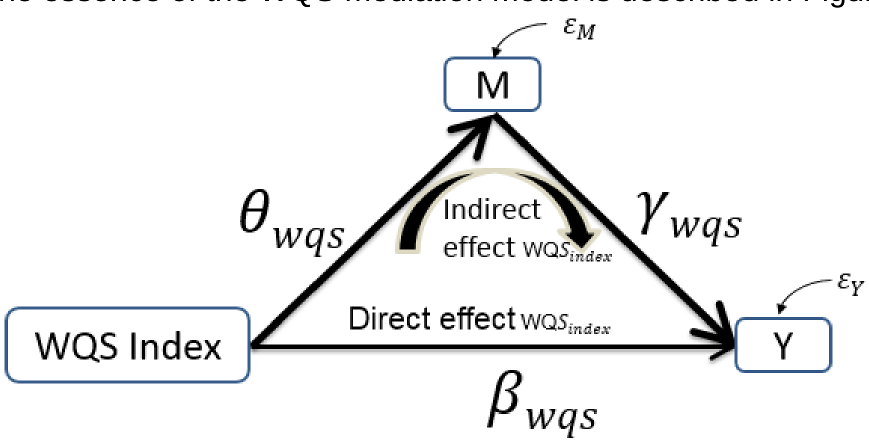


Figure 3.3 WQS Mediation Model with Composite Weighted Sum of the Predictors

The predictors $X_1 \ldots X_p$ are represented as a weighted quartile sum index WQS_{Index} have an effect on outcome Y, which acts partially through mediator M while the remainder acts directly on Y. The weighted sum parameters shown in Figure 3.3 can be estimated as $\hat{\theta}_{wqs}, \hat{\beta}_{wqs}$ and $\hat{\gamma}_{wqs}$ by equations 3.3 and 3.4

$$M = \beta_{0,M} + \theta_{wqs} WQS_{index} + \varepsilon_M \qquad (3.3)$$

$$Y = \beta_{0,Y} + \beta_{wqs} WQS_{index} + \gamma_{wqs} M + \varepsilon_Y \qquad (3.4)$$

The WQS analysis models for mediation that are comparable to the population mediation models in Figures 3.1 and 3.2, are shown in Figures 3.4 and 3.5. As part of the WQS method, a weighted quartile sum index (WQS_{index}) is computed to represent the predictor set as an entity. The predictor variables in classic mediation shown in Figures 3.1 and 3.2 are then replaced with the WQS_{index}, allowing for the estimation of the total indirect effect between all the predictors and the outcome. The independent variable values in a dataset are scored in quartiles by each predictor *j* and these ranks are used in place of the X_{ij} values in WQS regression. The quartile rank q_{ij} for each independent variable is multiplied by the bootstrapped predictor weight $\tilde{w}_j$ and summed over all the p predictors for each observation i to obtain a WQS_i index, in a replicated dataset.

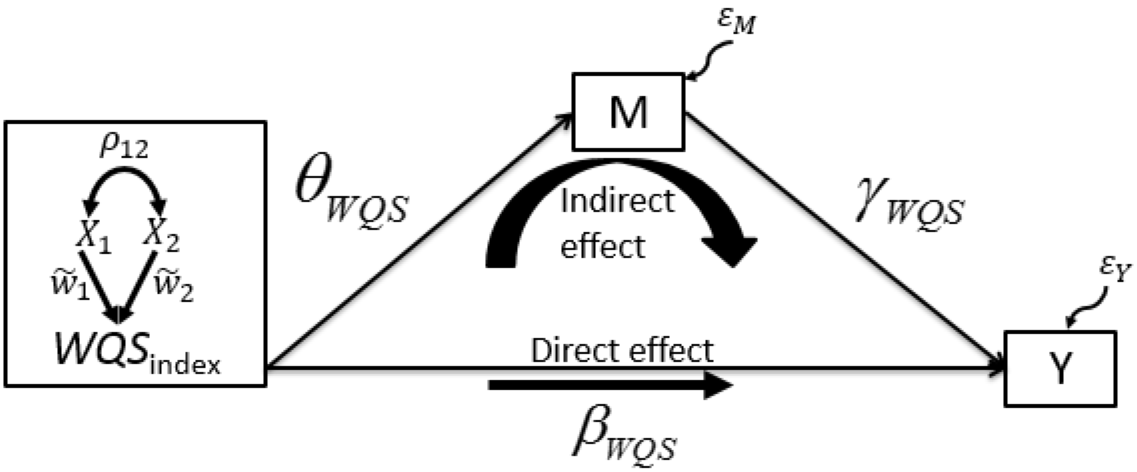


Figure 3.4 WQS Index as a Composite Weighted Sum of the Two Correlated Predictors

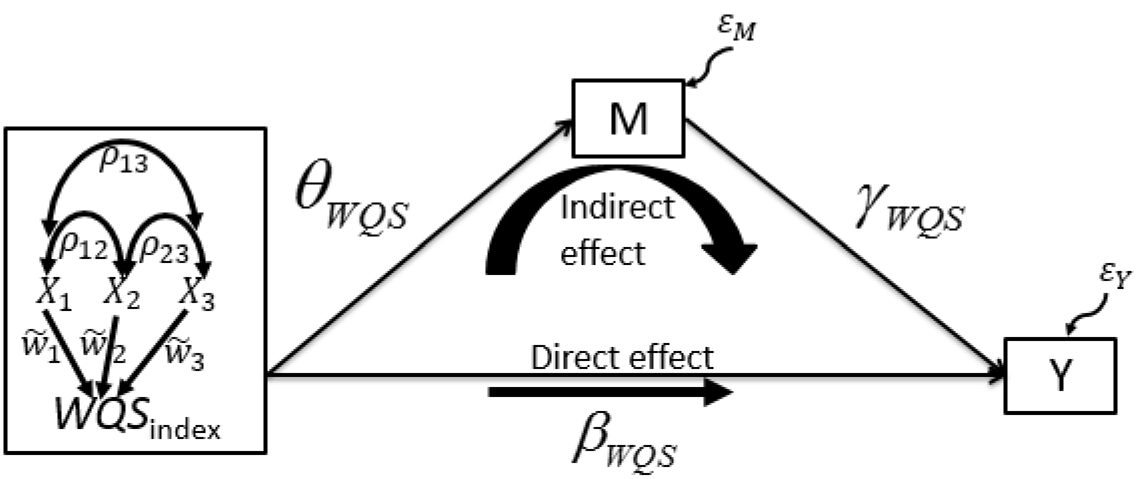


Figure 3.5 WQS Index as Composite Weighted Sum of the Three Correlated Predictors

To obtain the WQS weights, each dataset is bootstrapped *B* =100 times. A WQS weight for the X_j^{th} predictor is $\bar{w}_j = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)}$ which is defined as the mean of all bootstrapped weights $W_{j_{(b)}}$ of the independent variable $X_{j_{(b)}}$. Three methods for calculating the w_j weights were considered: a) *M* is regressed on the independent predictors denoted as *M*|*X*

$$M = \beta_{0,M} + \beta_{wsm}^{M|X} \sum_{j=1}^{p} w_j^{M|X} q_{ij} + \varepsilon_M \qquad (3.5)$$

giving $\bar{w}_j^{M|X} = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)}^{M|X}$ b) *Y* regressed on the independent predictors and *M*

denoted as *Y*|*X, M* in

$$Y = \beta_{0,Y} + \beta_{wsm}^{Y|X,M} \sum_{j=1}^{p} w_j^{Y|X,M} q_{ij} + \gamma M + \varepsilon_Y \quad (3.6)$$

giving $\bar{w}_j^{Y|X,M} = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)}^{Y|X,M}$; and c) *Y* regressed on the independent predictors

without *M*, denoted as *Y*|*X* in

$$Y = \beta_{0,Y^*} + \beta_{wsm}^{Y|X} \sum_{j=1}^{p} w_j^{Y|X} q_{ij} + \varepsilon_{Y^*} \qquad (3.7)$$

giving $\bar{w}_j^{Y|X} = \frac{1}{B}\sum_{b=1}^{B} w_{j(b)}^{Y|X}$

The $w_{j_{(b)}}$ weights are estimated using least squares in SAS 9.4 PROC OPTMODEL under the constraints that the regression slope parameter is positive $\beta \geq 0$, the p predictor weights for each observation add to unity, $\sum_{j=1}^{p} w_j = 1$,and each predictor weight $(w_{ij},\ i^{th}\ \text{obervation and}\ j^{th}\ \text{predictor})$ is constrained between zero and one, $0 <= w_{ij} <= 1$. The three different WQS$_i$ indices are defined as $\mathrm{WQS}_i^{M|X} = \sum_{j=1}^{p} \bar{w}_j^{M|X} q_{ij}$, $\mathrm{WQS}_i^{Y|X,M} = \sum_{j=1}^{p} \bar{w}_j^{Y|X,M} q_{ij}$ and $\mathrm{WQS}_i^{Y|X} = \sum_{j=1}^{p} \bar{w}_j^{Y|X} q_{ij}$; where each is used in a separate WQS regression mediation analysis. The appropriate WQS_i index is

used in the two regression equations 3.3 and 3.4 to obtain three sets of regression coefficients $\hat{\theta}_{wqs}^{M|X}, \hat{\beta}_{wqs}^{M|X}$ and $\hat{\gamma}_{wqs}^{M|X}$; $\hat{\theta}_{wqs}^{Y|X,M}, \hat{\beta}_{wqs}^{Y|X,M}$ and $\hat{\gamma}_{wqs}^{Y|X,M}$; and $\hat{\theta}_{wqs}^{Y|X}, \hat{\beta}_{wqs}^{Y|X}$ and $\hat{\gamma}_{wqs}^{Y|X}$.

Each mediation model now has a single WQS_{index} as the independent variable, a mediator *M* and an outcome *Y* as shown in Figure 3.3 to obtain the WQS mediated effects. A WQS weighted sum mediated effect for p predictors is defined to exist when the product of theta and gamma $\hat{\theta}_{wsm}\hat{\gamma}_{wsm}$ is significant. The three methods for estimating the WQS weights yield three different estimates of the indirect and direct effects. 1) $\hat{\theta}_j^{M|X}\hat{\gamma}^{M|X}$ and $\hat{\beta}_j^{M|X}$ 2) $\hat{\theta}_j^{Y|X,M}\hat{\gamma}^{Y|X,M}$ and $\hat{\beta}_j^{Y|X,M}$ 3) $\hat{\theta}_j^{Y|X}\hat{\gamma}^{Y|X}$ and $\hat{\beta}_j^{Y|X}$ respectively.

3.2.2 Weighted Quantile Sum Regression Testing for Indirect Effects

Two nonparametric bootstrap methods for testing the indirect effects were considered, namely a nested bootstrap method and a sequential bootstrap method. The goal of both methods is to form a percentile confidence interval for the indirect effect. In both methods we use 500 bootstrap replications of the original datasets to determine the indirect effect's percentile confidence interval and use *k* to denote the k^{th} bootstrap replication where $k \in \{1...500\}$. The nested method for computing the indirect effect's confidence interval is termed nested bootstrap because the bootstrap for the mean weights is nested within the bootstrap for the indirect effects. The nested bootstrap method treats the observed predictors as the data that is resampled in the WQS bootstrap procedure and uses *b* to denote the b^{th} WQS bootstrap to determine

mean weights, where $b \in \{1...100\}$ for each of the 500 bootstrapped replications. This incorporates variability of $\bar{w}_{j(k)}^{M|X}$, $\bar{w}_{j(k)}^{Y|X,M}$, $\bar{w}_{j(k)}^{Y|X}$ over the 500 bootstrapped datasets for determining the indirect effect's percentile confidence interval. Therefore, once a bootstrap replication is obtained, each bootstrapped replication is further bootstrapped 100 times to calculate the WQS mean predictor weights $\bar{w}_{j(k)}^{M|X}$, $\bar{w}_{j(k)}^{Y|X,M}$, $\bar{w}_{j(k)}^{Y|X}$, which are different for each of the 500 bootstrap replications. The WQS_i indices for each observation *i,* are calculated $WQS_{i(k)}{}^{M|X}$, $WQS_{i(k)}{}^{Y|X,M}$ *and* $WQS_{i(k)}{}^{Y|X}$ $k \in \{1^{st}...500^{th}\}$ bootstrap replication using $\bar{w}_{j(k)}^{M|X}$, $\bar{w}_{j(k)}^{Y|X,M}$, $\bar{w}_{j(k)}^{Y|X}$ respectively. This bootstrap method is applied to each of the three methods of estimating the WQS predictor weights and the estimated indirect effects $\hat{\theta}_{wsm\,(k)}^{M|X}\ \hat{\gamma}_{wsm\,(k)}^{M|X}$, $\hat{\theta}_{wsm\,(k)}^{Y|X,M}\ \hat{\gamma}_{wsm\,(k)}^{Y|X,M}$, $\hat{\theta}_{wsm\,(k)}^{Y|X}\ \hat{\gamma}_{wsm\,(k)}^{Y|X}$; $k \in \{1...500\}$ from 500 bootstrap replications, yielding the 95% percentile confidence interval for the indirect effects.

The sequential bootstrap method treats the WQS_i index as the unit for resampling compared to the observed data in the nested bootstrap method. The method is termed sequential bootstrap because the bootstrap for the indirect effects follows the bootstrap for the WQS predictor weights. Therefore, the mean weights $\bar{w}_j^{Y|X}, \bar{w}_j^{M|X}, \bar{w}_j^{Y|X,M}$ are estimated to calculate the WQS_i indices; $WQS_i^{M|X}$, $WQS_i^{Y|X,M}$ and $WQS_i^{Y|X}$ and these WQS_i indices are used as the data that is sampled with replacement. Five hundred bootstrap replications are then used to determine the indirect effect's percentile confidence intervals. This method for testing the indirect effect was applied to each of the three methods for estimating the WQS weights.

The essential difference between the sequential and nested bootstrap methods for testing the indirect effects is that, in the sequential bootstrap method the WQS indices are sampled with replacement 500 times, where as in the nested bootstrapped method the original data of the predictors are resampled yielding different mean predictor weights $\bar{w}_j^{Y|X}, \bar{w}_j^{M|X}, \bar{w}_j^{Y|X,M}$ for each of the 500 bootstrapped original datasets. The nested method is computationally more intensive than the sequential method since each nonlinear optimization algorithm is executed100 times for each of the 500 resampled original datasets. However, since the mean predictor weights are allowed to vary in each bootstrapped dataset, they account for the sampling error in the weights in addition to the sampling error in regression coefficients. The WQS direct effects obtained from PROC GLM use the *t-test* of the regression coefficient estimates for the predictor $\text{WQS}_{\text{index}}$ in equation 3.4 i.e. $\hat{\beta}_{wsm}^{M|X}, \hat{\beta}_{wsm}^{Y|X,M}$ and $\hat{\beta}_{wsm}^{Y|X}$ to detect statistical significance for the direct effects DE; $DE^{M|X} = \hat{\beta}_{wsm}^{M|X}$, $DE^{Y|X,M} = \hat{\beta}_{wsm}^{Y|X,M}$ or $DE^{Y|X} = \hat{\beta}_{wsm}^{Y|X}$ respectively.

3.2.3 Ordinary Least Squares Regression (OLS)

In the mediation models shown Figures 3.1 and 3.2, the effects of independent variables $X_1 \ldots X_p$ on a dependent variable Y could be partially transmitted through a mediator variable M, while the remainder acts directly on Y. Analyses of the data in the population models with p predictors uses two regression equations Equation 3.8 and Equation 3.9 to estimate coefficients that are used to calculate the indirect and direct

effects of mediation (Barron & Kenny 1986). The first regression is of the dependent variable mediator *M* on the independent variables $X_1 \ldots X_p$,

$$M = \beta_{0,M} + \sum_{j=1}^{p} \theta_j X_j + \varepsilon_M \qquad (3.8)$$

with $\beta_{0,M}$ being the intercept, each regression coefficient, $\theta_j \in (\theta_1 \ldots \theta_p)$,representing the unique individual effects of an X_j adjusting for the remaining X's on M in the model and ε_M representing the model errors. The second regression,

$$Y = \beta_{0,Y} + \sum_{j=1}^{p} \beta_j X_j + \gamma M + \varepsilon_Y \qquad (3.9)$$

is of the dependent variable response *Y* on the independent variables $X_1 \ldots X_p$ and the mediator *M*, with $\beta_{0,Y}$ being the intercept, and each regression coefficient, $\beta_j \in \{\beta_1 \ldots \beta_p\}$ representing the unique individual effects of X_j adjusting for the remaining *X*'s and *M*, on *Y* in the model with errors ε_Y . The regression estimates needed are $\hat{\theta}_j$ from equation 3.8 and $\hat{\beta}_j$, $\hat{\gamma}$ from equation 3.9. The indirect effects are calculated from the estimates as $\hat{\theta}_j \hat{\gamma}$ while direct effects are estimated directly from the regression coefficient for predictors X_j; $\hat{\beta}_j$ in equation 3.9. (Baron and Kenny, 1986). Mediation analysis for OLS regression was conducted in SAS 9.4 PROC REG yielding regression coefficients for the *p* individual indirect effects $\hat{\theta}_j \hat{\gamma}$ and *p* individual direct effects $\hat{\beta}_j$. The OLS direct effects use the *t-test* to test the significance of the regression coefficient for X_j

from equation 3.9 to detect statistical significance for $\hat{\beta}_j;\ j \in \{1...p\}$. The nonparametric percentile bootstrap method with 500 replications was used to form a confidence interval for the indirect effects $\hat{\theta}_j \hat{\gamma};\ j \in \{1...p\}$.

3.2.4 Least Absolute Shrinkage Operator Regression (LASSO)

In LASSO, two methods for obtaining the parameter shrinkage estimates are considered: 1) at the minimum of the n-fold cross-validation average prediction error plotted as a function of a range of Lagrange multiplier (λ) values to determine the optimum value for λ , which is used to estimate the LASSO regression coefficients and 2) at the minimum plus one standard error of the average n-fold cross-validation prediction errors, for the more conservative regression estimates. A post processing step includes fitting only admitted predictors from LASSO regression into the regression equations 3.1 and 3.2 to solve for the two sets of parameter estimates to calculate the two LASSO indirect effects $\hat{\theta}_j^{L_{\min}} \hat{\gamma}^{L_{\min}}$, $\hat{\theta}_j^{L_{\min+1se}} \hat{\gamma}^{L_{\min+1se}}$ and the two direct effects $\hat{\beta}_j^{L_{\min}}$, $\hat{\beta}_j^{L_{\min+1se}}$.

LASSO analyses were conducted in in R Version 3.0.1 using the *glmnet* package (Friedman, Hastie and Tibshirani (2010)). A lambda multiplier of zero for differential shrinkage with regards the variable *M* was used, to ensure that *M* was always admitted as a variable in the LASSO regression.

Post-processing only variables that were admitted into the final solution from LASSO regression using OLS multiple regression equations for mediation analysis with PROC REG, yields regression coefficients that estimate the two sets of indirect effects $\hat{\theta}_j^{L_{\min}} \hat{\gamma}^{L_{\min}}$, $\hat{\theta}_j^{L_{\min+1se}} \hat{\gamma}^{L_{\min+1se}}$ and direct effects $\hat{\beta}_j^{L_{\min}}$, $\hat{\beta}_j^{L_{\min+1se}}$. The non-parametric

percentile bootstrap method was used on "*up to* p" individual indirect effects because of possible dimension reduction occurring due to shrinkage of the estimates during LASSO optimization. Statistically significant LASSO direct effects were identified using the *t-test* results for "*up to* p" LASSO admitted predictors in equation 3.2.

3.3 Comparison of Methods

To provide a fair comparison of OLS and LASSO methods with the WQS method, each observation's estimated WQS_{index} must be standardized, since X_{ij}s and *M* in OLS and LASSO regressions are obtained from the multinormal distribution, $N_{p+1}(0,1)$. The next difference for the WQS method is the WQS_{index}, which takes the place of the individual predictors in the OLS and LASSO regressions. A weighted sum of the predictors defined by WQS_{index} has regression coefficient estimates from equations 3.3 and 3.4 that will be the joint indirect effect $\hat{\theta}_{wqs}\hat{\gamma}_{wqs}$ and the joint direct effect $\hat{\beta}_{wqs}$. In LASSO regression and the OLS method, individual mediated effects are estimated from the regression equations 3.8 and 3.9. A *joint* mediated effect must be defined for the OLS and LASSO methods, to make the estimated indirect and direct effects from the WQS_{index} comparable. A mediated effect present for one or more of the predictors will aggregate to a *joint* mediated effect estimate. An individual predictor's mediated effect may not be significant but the joint mediated effect could be significant through an aggregation of the other predictor's significant mediated effects. Hence, the joint power is higher than that for a single predictor.

A joint effect for the LASSO method is defined to exist, if any of the "up to *p*" predictors, admitted into the final LASSO shrunk estimates has a statistically significant

indirect effect. The same method is used to determine if there is any direct effect present amongst the up to *p* predictors that were admitted into the final LASSO shrunk estimates, which were used in the post processing step to calculate the LASSO individual direct effect estimates.

In WQS the regression coefficients associated with WQS_{index} are used to estimate the joint mediated effects. Comparing the joint mediated effect for the WQS_{index} to an OLS or LASSO predictor's individual mediated effect is not a fair comparison. While the joint mediated effects across the three methods are comparable for type1 error and power, individual mediated effects can be compared only for OLS and LASSO. In WQS, given a significant WQS_{index} mediated effect, an Individual predictor's influence on its mediated effect can be estimated using various cut-off values for the WQS bootstrapped weights $\bar{w}_j^{Y|X}, \bar{w}_j^{M|X}, \bar{w}_j^{Y|X,M}$ that decide whether the j^{th} predictor is an important contributor to the joint mediated effect. The range and increments for cut-off values for *p* predictors is motivated by the heuristic $\frac{1}{1+p}$ as being the mid-value for the range of WQS cut-off values to be explored. Predetermined fractional values of 0.2 to 0.5 in steps of 0.1 were used for the two predictor mediation models and values of 0.1 and 0.4 in steps of 0.1 were used for the three predictor mediation models. These cut-off values represent the floor below which a WQS predictor is considered have a zero weight and minimal influence on the outcome through the $\mathrm{WQS}_{\mathrm{index}}$. Provided the joint mediated effect exists, the cut-off value will determine whether an individual predictor is important in the mediation model. The influence of a predictor in the WQS method is quantified by its type1 error rate and power, which is compared to the corresponding

OLS and LASSO results. The measures used to compare the WQS, LASSO and OLS regression methods for the point estimates are the standard error of the estimate, estimate's bias, and an estimate's mean squared error, and for the interval estimates are its coverage probability, type1 error and power.

4 Simulation Results

4.1 Weighted Quantile Sum Regression 2 & 3 Variable Mediation Analysis

The results from the simulation studies for 2-predictor conditions and 3-predictor conditions are discussed and displayed in this section. The 2-predictor population model shown in Chapter 3 – Figure 3.1 had seven parameters to be chosen that provided the possibility of estimating the method's type1 error and power. Similarly, the 3-predictor population model shown in Chapter 3 – Figure 3.2 had eleven parameters to be chosen that provided the possibility of estimating the method's type1 error and power. The parameter combinations resulted in 108 unique combinations of parameter settings for the 2-predictor mediation model and 162 different parameter combinations for the 3-predictor mediation model.

Each parameter setting was used to estimate three point estimate related statistics and three interval estimate related statistics: Average indirect and direct estimate for the combined parameters and the individual predictor indirect and direct effects from the 500 replicated datasets for each condition, providing the sampling variability of the estimate. Since the true parameter values were known a priori (chosen), the bias for the average indirect and direct effects could be estimated. The Root Mean Squared Error (RMSE) for the average indirect and direct effects was also estimated. For the interval estimates the reported statistics for coverage probability, type1 error and power were estimated for the 2-predictor WQS method using Sequential and Nested 95th percentile confidence intervals. However, the results show that the

added computational burden associated with the Nested bootstrap method gave lower type1 error and lower power than the Sequential bootstrap 95th percentile confidence intervals for the indirect estimates. The preferred method was chosen to be the Sequential bootstrap method for estimating all 3-predictor indirect effect interval estimates for coverage probability, type1 error rate and power.

The sequence of the discussions in this section are as follows:

WQS 2-predictor mediation models

1. Compare average mediated estimates, bias and RMSE values for all 108 conditions, and contrast coverage probability, type1 error rate and power from Nested vs. Sequential methods used for estimating the indirect effect's 95th percentile confidence interval.
2. The indirect effect estimates in sub-section 1 (above) is followed by the corresponding combined and individual direct effect estimates for average estimate, the bias, RMSE, coverage probability based on the 95th percent confidence interval (average estimate ± 1.96 * s.e. of the estimate), the type1 error rates and the power to detect the combined or individual direct effect estimate.

WQS 3-predictor mediation models

3. The 2-predictor mediation models are followed by the 162 conditions, 3-predictor mediation models, for combined and individual indirect and direct effect estimates. The indirect effect 95th percentile confidence intervals were estimated using the Sequential bootstrap method.

LASSO 2-predictor mediation models

4. Two different ways of estimating the LASSO regression coefficients are discussed in this sub-section: a) the regression coefficients associated with the minimum point on the n-fold cross-validation prediction error curve and b) the regression coefficients associated with the minimum + 1 s.e. for the n-fold cross-validation prediction errors (as suggested in the literature).These were designated as L_{min} and $L_{min+1se}$ (or L_{1se}).
5. A decision between the two ways of estimating the LASSO regression coefficients (L_{min} vs. $L_{min+1s.e.}$) is made based on the results for the bias, RMSE, coverage probability, type1 error and power for the indirect and direct effects.

OLS 2-predictor mediation models

6. The 2-predictor mediation models and 3-predictor mediation models are used to estimate the average indirect and direct mediated effects using the OLS method for each population model, the estimate bias, estimate RMSE, coverage probability, type1 error and power.

Pairwise comparisons between OLS, LASSO and WQS methods

7. First the 2-predictor mediation models are compared on the statistics of average indirect and direct estimates for combined and individual predictor effects, estimate bias, estimate RMSE, coverage probability, type1 error and power, followed by the 3-predictor mediation models.
8. These comparative plots are summarized in a table to make recommendations based on the strengths and weaknesses of OLS, LASSO and WQS methods for 2-predictor and 3-predictor population models. LASSO and WQS are compared based on variable selection while LASSO and OLS are compared based on

individual predictor estimates. WQS is compared to LASSO and OLS based on a joint mediated effect defined as LASSO or OLS having one or more significant mediated effects when compared to the WQS index mediated effect from the WQS regression method. Also, the set of influencing individual predictors from the WQS method is determined by a series of selected cut-off values for the individual WQS weights in a given WQSindex: $WQS_{indsex}^{M|X}, WQS_{indsex}^{Y|X}, WQS_{indsex}^{Y|X,M}$.

4.1.1 WQS 2-Variable Mediation Analysis

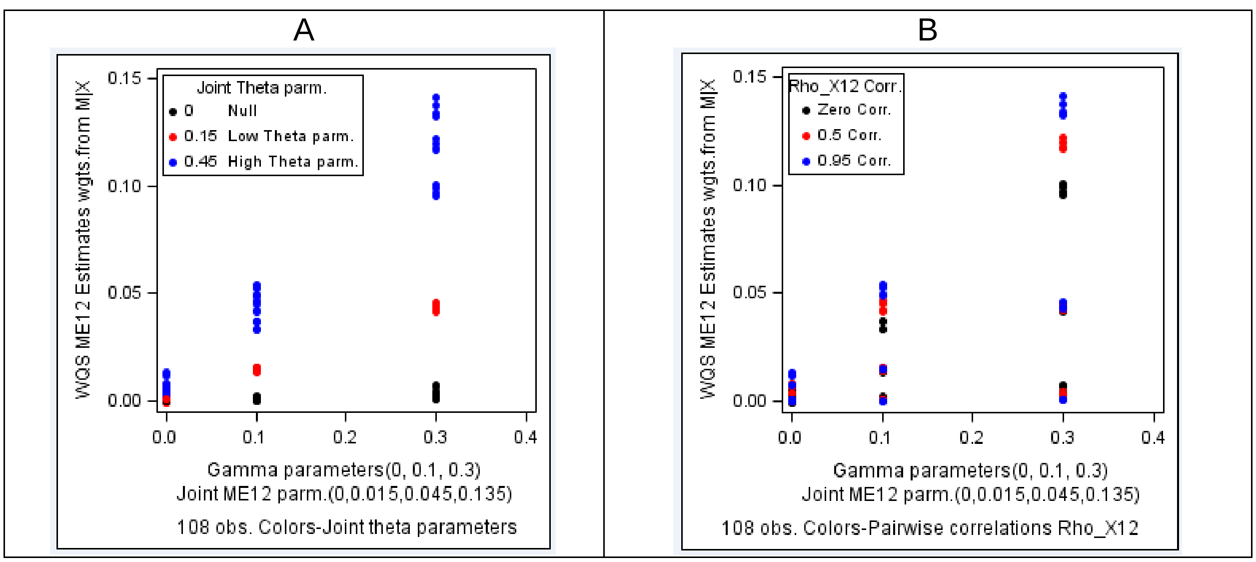
4.1.1.1 Joint Indirect Effects' WQS *Estimate*, *Bias* and *RMSE* – *WQS*$^{M|X}$ Weights

Joint indirect effects are calculated by predictor weights using three different methods. 1) $ME_{12}^{M|X}$ using $WQS_{index}^{M|X}$ when *M* is regressed on *X*, 2) $ME_{12}^{Y|X}$ using $WQS_{index}^{Y|X}$ when *Y* is regressed on *X*, and 3) $ME_{12}^{Y|X,M}$ using $WQS_{index}^{Y|X,M}$ when *Y* is regressed on *X* with M as a covariate. The estimated joint indirect effects in Figures 4.1A and B show the relationships between the joint indirect effect $ME_{12}^{M|X}$ estimates and the gamma parameter γ, the joint theta parameter $\theta_{12} = (\theta_1 + \theta_2)$ and pairwise correlations ρ_{12} . The joint indirect effect $ME_{12}^{M|X}$ estimates are larger for increasing values of gamma, joint theta and pairwise correlation. The figures also supports the fact that the WQS method estimates the joint indirect effect with almost zero bias, when $\gamma = 0$ and $\theta_1 = 0, \theta_2 = 0$, which is triple null case. When $\gamma = 0$ and $\theta_1 = 0, \theta_2 = 0.15$, the small positive bias on $ME_{12}^{M|X}$ shifts towards a zero bias as $\rho_{12} = \{0, 0.5, 0.95\}$ increases. When $\gamma = 0$ and $\theta_1 = (0.15, 0.30)$, the positive bias on $ME_{12}^{M|X}$ increases (0.004 to 0.013) as $\rho_{12} = \{0, 0.5, 0.95\}$ increases, and is

shown in Figure 4.1 A. The $ME_{12}^{M|X}$ estimate bias is positive (0 to 0.01) for the true value $ME_{12} = 0$, and negative (0 to -0.03) for positive non-zero ME_{12} parameter values.

The $ME_{12}^{M|X}$ estimate bias becomes more negative with larger $ME_{12}^{M|X}$ estimates shown in Figure 4.1 C.

The $ME_{12}^{M|X}$ estimate's *RMSE* has an inverse relationship with square-root of sample size and by definition, a direct relationship with the joint mediated effect's parameter value ME_{12}. Larger correlations (black ($\rho_{12} = 0$), red ($\rho_{12} = 0.5$), blue ($\rho_{12} = 0.95$) increase the estimate's *RMSE* as shown in Figure 4.1 D. However, the pattern for $ME_{12}^{M|X}$'s *RMSE* inverts for large values of $(\gamma = 0.3, \theta_{12} = 0.45)$, showing smaller *RMSE* values for larger pairwise correlations ρ_{12}.



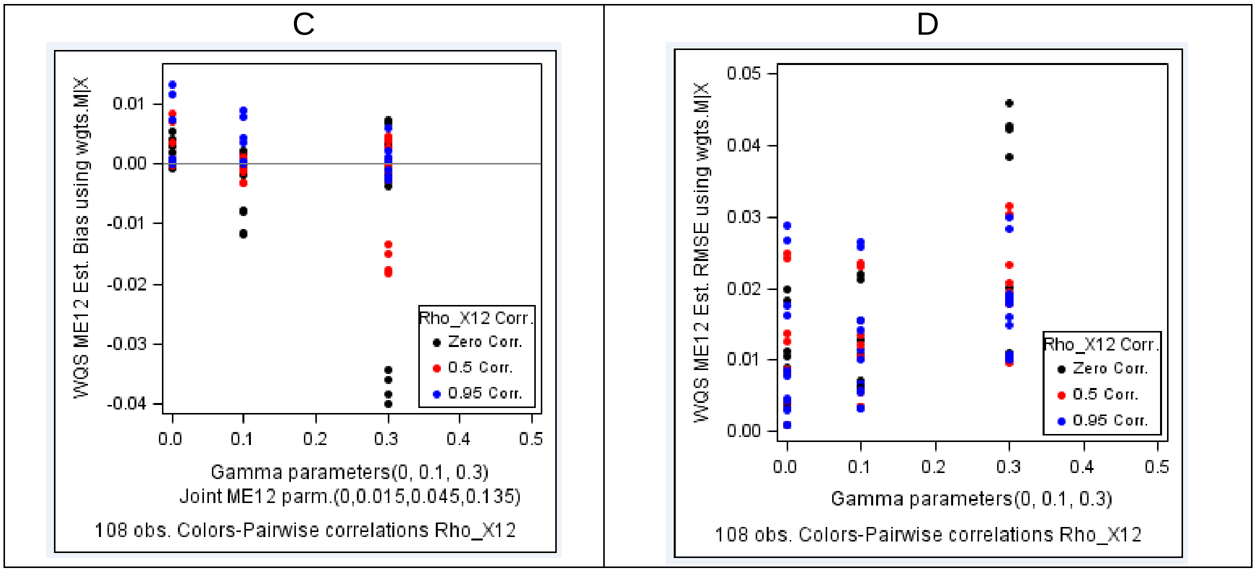


Figure 4.1 Joint Indirect Effects *Estimate* (A-B), *Bias*, *RMSE* (C-D) Using $WQS_{index}^{M|X}$

4.1.1.2 Coverage Probability, Type1 Error and Power for Joint Mediated Effects

Coverage probability in Figure 4.2 A vs. B shows that the sequential bootstrap method performs markedly better than the nested bootstrap method for the joint mediated effect's coverage probability. The nested bootstrapping method uses a distribution of WQS weights, while the sequential bootstrapping method uses fixed WQS weights for determining the point estimate's confidence interval. The same conclusion is true if the sample size is large. The choice of index $WQS_{index}^{M|X}$ over $WQS_{index}^{Y|X}$ or $WQS_{index}^{Y|X,M}$ is evident by comparing Figures 4.2 A to C and noting that several conditions in plot C have coverage probabilities less than 0.2 and some conditions with near zero coverage when using $WQS_{index}^{Y|X}$ or $WQS_{index}^{Y|X,M}$ indices. Coverage probabilities for the joint indirect effects using the index $WQS_{Index}^{Y|X}$ (Seq. vs. Nested) shown in Figures 4.2 C and D do not have the same marked difference between the two methods when using the index $WQS_{Index}^{M|X}$, as shown in Figures 4.2 A and B. There is no preference between the sequential and the nested

bootstrap methods when determining the predictor WQS weights using the index $WQS_{Index}^{Y|X}$. However, the sequential bootstrap method is preferred over the nested bootstrap method if the chosen index is $WQS_{Index}^{M|X}$, since it provides better coverage probabilities for the 108 conditions as shown in Figure 4.2 A.

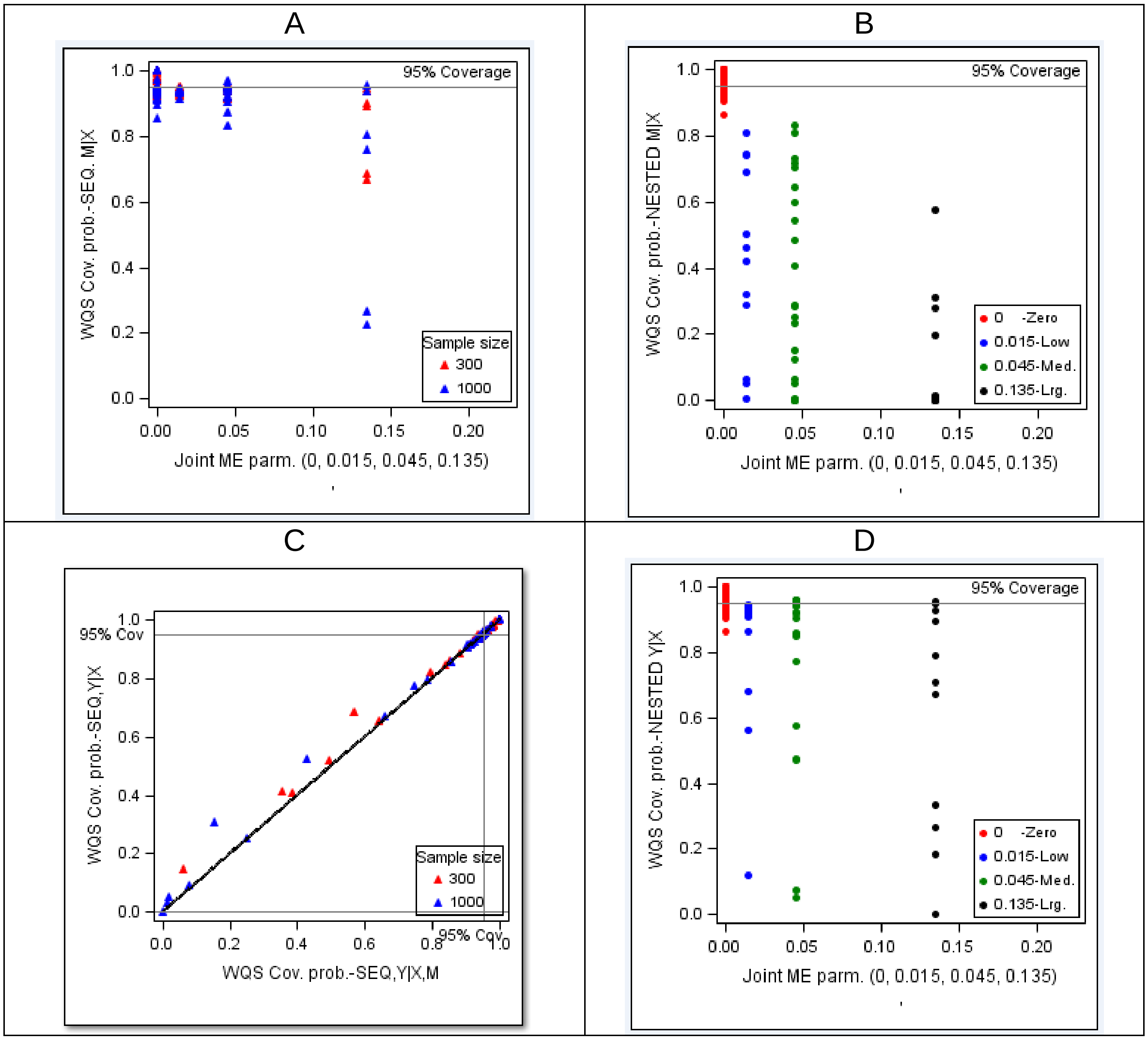


Figure 4.2 Coverage Probability for ME_{12} A) Sequential $WQS_{index}^{M|X}$, B) Nested $WQS_{index}^{M|X}$ C) Sequential $WQS_{index}^{Y|X}$ vs. Sequential $WQS_{index}^{Y|X,M}$ D) Nested $WQS_{index}^{Y|X}$

Next, the type 1 error for the joint mediated effect $ME_{12}^{M|X}$ using $WQS_{index}^{M|X}$ is compared between the sequential and nested bootstrap methods. The influencing

parameters that determine the joint mediated effect $ME_{12}^{M|X}$ are $\theta_{12} \in \{0, 0.15, 0.30\}$ and $\gamma \in \{0, 0.1, 0.3\}$. Exceptions to the type1 error confidence limit of 0.075, set *a priori*, for $ME_{12}^{M|X}$, $ME_{12}^{Y|X,M}$ and $ME_{12}^{Y|X}$ using the sequential and nested methods are presented next.

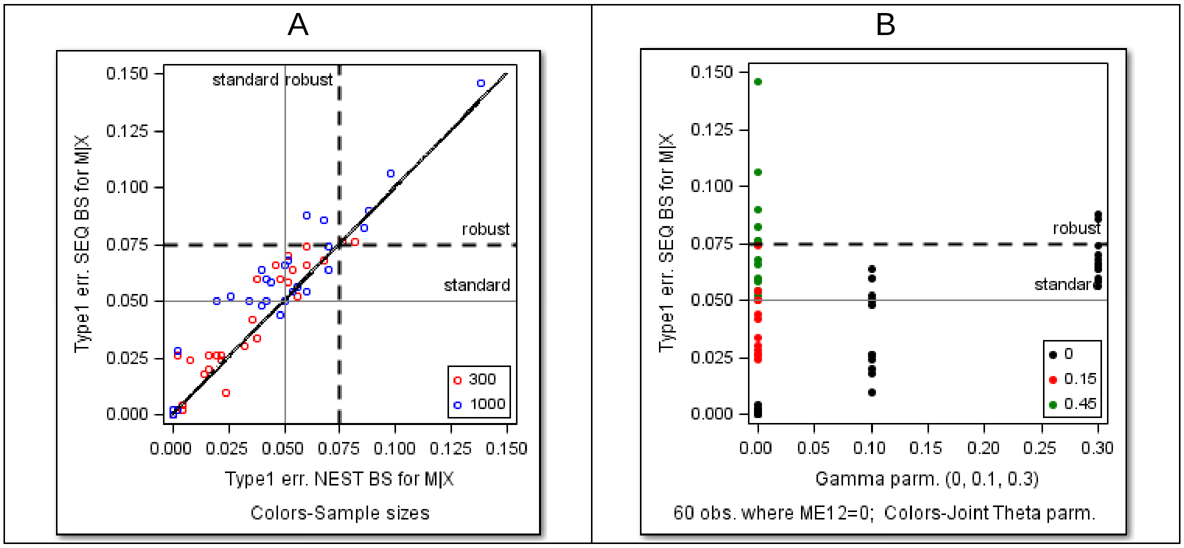


Figure 4.3 A-B Seq. vs. Nested Methods Type1 Error Analysis Using $WQS_{index}^{M|X}$ for $ME_{12}^{M|X}$

The table of exceptions shown below details the conditions where type1 errors>0.075

Table 4.1

Type1 err. >0.075 for ME_{12}; $WQS_{index}^{M|X}$ $WQS_{index}^{Y|X}$ $WQS_{index}^{Y|X,M}$ for Nested & Sequential

Sample size	ρ_{12}	θ_1	θ_2	γ	β_1	β_2	Seq. *M|X*	Nest. *M|X*	Seq. *Y|X*	Nest. *Y|X*	Seq. *Y|X,M*	Nest. *Y|X,M*
1000	0.5	0	0	0.3	0.15	0.3	0.088					
1000	0.95	0	0.15	0	0.15	0.3			0.076		0.076	
1000	0	0.15	0.3	0	0.15	0.3	0.082	0.086	0.082	0.086	0.082	0.086
300	0.5	0.15	0.3	0	0.15	0.3	0.076	0.082	0.076	0.082	0.080	0.082
1000	0.5	0.15	0.3	0	0.15	0.3	0.090	0.088	0.088	0.088	0.088	0.088
300	0.95	0.15	0.3	0	0.15	0.3	0.076	0.076	0.078	0.076	0.078	0.076
1000	0.95	0.15	0.3	0	0.15	0.3	0.146	0.138	0.144	0.138	0.146	0.138
1000	0.5	0	0	0.3	0.3	0	0.086					
1000	0.95	0.15	0.3	0	0.3	0	0.106	0.098	0.098	0.098	0.098	0.098

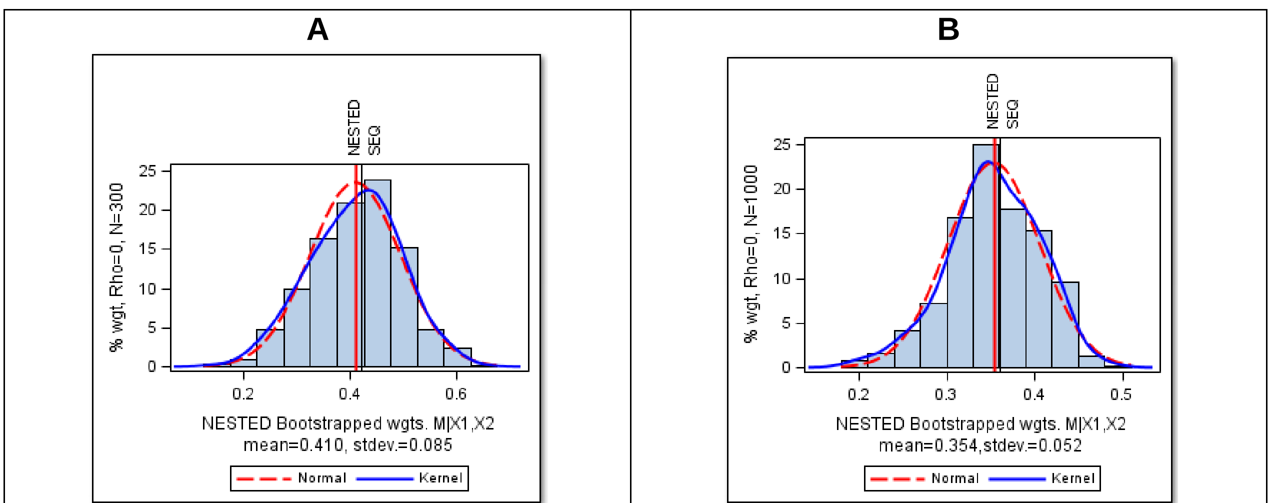
Joint mediated effects' type1 error rates exceptions based on the *a priori* chosen parameters are shown in Table 4.1. There were 8 exceptions out of a possible 60 type1 errors for $ME_{12}^{M|X}$, 7 for $ME_{12}^{Y|X}$, and 7 for $ME_{12}^{Y|X,M}$ using the sequential method, and 6 exceptions using the nested method for each. There is no mediated effect either because $\gamma = 0$ while one or more theta parameters are non-zero (conditions 2 – 7 and 9), or all the theta parameters are zero when the gamma parameter is non-zero (conditions 1 & 8). WQS regression being a constrained optimization method introduces a positive bias on zero parameters and a negative bias on positive, non-zero parameters. Pairwise correlation between the predictors, and sample size, play an important part in determining the regression coefficient estimates. Since a weighted quartile sum regression method is used, the zero mediated effects are biased positively, and the larger sample sizes have narrower confidence intervals resulting in high type1 errors as seen in conditions 2 through 7 and 9. The mediated estimate bias changes from negative to positive as the pairwise predictor correlation increases from 0 to 0.95. The large sample size have smaller confidence interval widths around the positively biased zero estimate, raising its type1 error rate.

The empirically determined WQS weights, when multiplied by the quartile ranks of predictor values, determine the index WQS_{index} for X_1, X_2. In Table 4.1 most of the conditions have a joint theta regression coefficient $\hat{\theta}_{12} > 0$ that is negatively biased, while the regression coefficient $\hat{\gamma}$ is a positively biased zero estimate. The type1 error is determined by the magnitude of the positive bias and the sample size of the data set. A larger biased estimate might have a reduced type1 error (0.076) if the sample size is small (N=300) as in conditions 4 and 6. On the other hand, a smaller positively biased

estimate from a large dataset (N=1000) results in a high type1 error (0.082 to 0.146), which increases with increasing pairwise correlation $\rho_{12} \in \{0, 0.5, 0.95\}$.

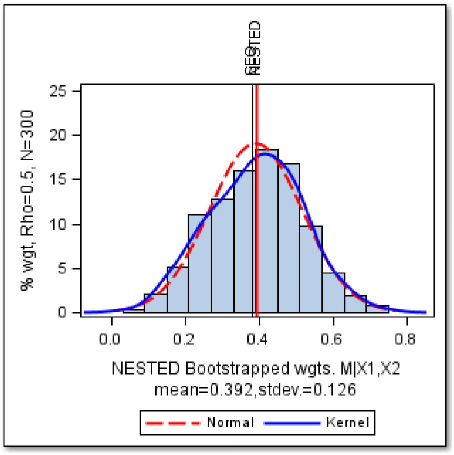
No determination could be made about the preferential choice of the WQS index based on Type 1 error results shown in Table 4.1, but Figure 4.3 A suggests that the nested $WQS_{index}^{M|X}$ method (Nested bootstrap method has 6 exceptions with Type1 errors >0.075 vs. Sequential method of 8 exceptions for Type1 errors >0.075) is slightly preferred over the sequential $WQS_{index}^{M|X}$ when calculating the mediated effects.

The Nested method for determining the percentile confidence interval has a type1 error rate which is influenced by the pairwise correlation between the predictors $\rho_{12} \in \{0, 0.5, 0.95\}$, because the WQS weights are allowed to vary between bootstrapped datasets. WQS weight distributions for predictor variables when using the sequential and nested methods for the joint and individual mediated effects are discussed below.

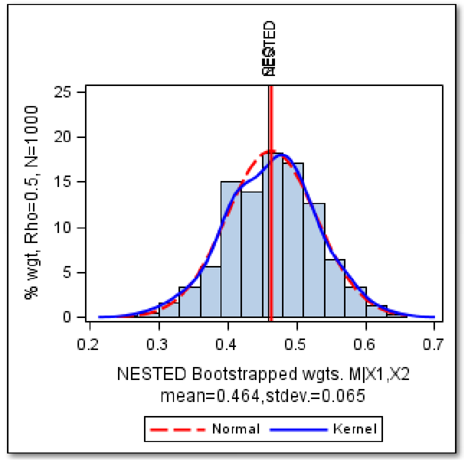


A-B Distribution of weights for wX_1 when $M \mid X$, ρ_{12}=0 ;sample size =300 (left),1000 (right);

C

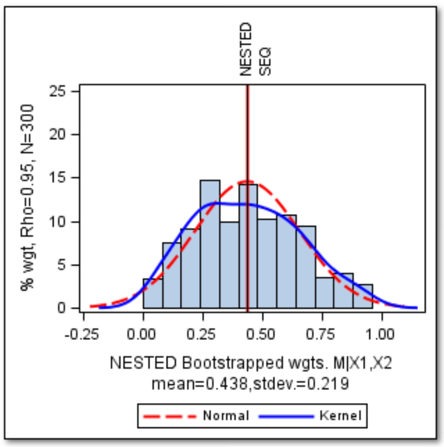


D

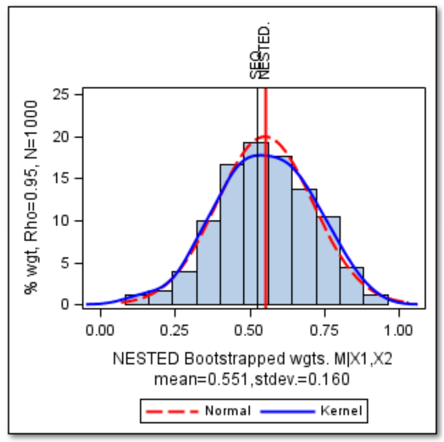


C-D Distribution of weights for $wX_1$when $M \mid X$, ρ_{12}=0.5 ;sample size =300 (left),1000 (right);

E



F



E-F Distribution of weights for $wX_1$when $M \mid X$, ρ_{12}=0.95 ;sample size =300 (left),1000 (right);

Figure 4.4 A-F Nested Method Distribution of WQS Wgts of 500 Bootstrapped Datasets

The distribution of weights for wX_1 from the nested method for conditions $(\theta_1 = 0.15, \theta_2 = 0.30)$, $\gamma = 0$, and $\rho_{12} \in \{0, 0.5, 0.95\}$ are shown in Figures 4.4. Ideally, the weights (wX_1, wX_2) should be proportioned as $1:2$ when $\rho_{12} = 0$, since $\theta_1 : \theta_2 = 1:2$ in the above condition. The distribution for wX_1 is centered closest to 0.33 when $\rho_{12} = 0,\ N = 1000$, and centered closest to 0.5 when $\rho_{12} = 0.95,\ N = 1000$ showing the effect of pairwise correlation on the WQS weights, which determine the WQS indices

$WQS_{index}^{M|X}, WQS_{index}^{Y|X}$ and $WQS_{index}^{Y|X,M}$ and subsequently the mediated estimates, e.g. $ME_{12}^{M|X}$ and $DE_{12}^{M|X}$. The WQS weight wX_1 from the sequential bootstrap method, is shown as a black vertical line, superimposed on the kernel and normal distributions for the histogram of 500 wX_1 weights, for the nested bootstrap method. Type1 errors and coverage probability for the two methods for a given condition, changing only the sample size and pairwise correlations, are shown in the table below.

Table 4.2

Type1 Error and Coverage Probability for $ME_{12} = 0$ Nested vs. Sequential Bootstrapping for the Condition with $(\theta_1 = 0.15, \theta_2 = 0.30)$, $\gamma = 0$ and $\rho_{12} \in \{0, 0.5, 0.95\}$

	N=300				N=1000			
Rho_X12	Type1 error		Cov. Prob.		Type1 error		Cov. Prob.	
	Seq.	Nested	Seq.	Nested	Seq.	Nested	Seq.	Nested
$\rho_{12} = 0$	0.052	0.056	0.948	0.944	0.082	0.086	0.918	0.914
$\rho_{12} = 0.5$	0.076	0.082	0.924	0.918	0.090	0.088	0.910	0.912
$\rho_{12} = 0.95$	0.076	0.076	0.924	0.924	0.146	0.138	0.854	0.862

When $\gamma = 0$, the true value for ME_{12} is zero. This table shows that the sequential method performs slightly better than the nested method with regards type1 error and coverage probability, for small datasets (N=300). For large sample sizes (N=1000), the type1 error and coverage probability is better for the sequential method only when $\rho_{12} = 0$, but type1 error increases while coverage decreases when the pairwise correlation is at 0.95. The nested method includes both the sampling variability in WQS weights and sampling variability in the data that determine the mediated estimates, resulting in a wider percentile confidence interval, a lower Type 1 error, e.g. (0.138 vs. 0.146), than that of the sequential bootstrap method.

The difference between the two methods is very small, lending preference to the less computationally intensive method of sequential bootstrapping when calculating the percentile confidence intervals for the mediated effects. The positive mediated estimate bias and the *RMSE* for the $ME_{12}^{M|X}$ increase as ρ_{12} increases, since wX_1 and wX_2 regress towards their mean ($\frac{wX_1 + wX_2}{2}$) for increasing correlations, but the variability for wX_1, wX_2 increase as both converge to 0.5. The variability in wX_1, wX_2 reflects in the variability of $\hat{\theta}_{12}$, $\hat{\gamma}$ giving $ME_{12}^{M|X}$ a high *RMSE* for higher pairwise correlations. Although the bias does not change much with sample size, the *RMSE* decreases as sample size increases narrowing the confidence interval around the biased estimate. This confidence interval increases the type 1 error rate for large sample sizes, while the coverage probability for the mediated effect decreases.

Indices $WQS_{index}^{M|X}, WQS_{index}^{Y|X}$ and $WQS_{index}^{Y|X,M}$ are used in place of the individual predictor values in the system of regression equations for mediation. Statistics such as the estimated value, the estimate's bias, the estimate's *RMSE*, type1 error, power, and coverage probability are evaluated for the joint and individual direct and indirect effects. The performance of the nested vs. the sequential bootstrap methods are summarized for the sixty parameter settings given that $ME_{12} = 0$. The nested method's lowest power for mediated effects was when theta or beta had a null parameter value for X_1 or X_2 in *Y*|*X*, *M* or *M*|*X* regression equations given high gamma parameter values, since the mediated effect's variability increases when a parameter takes a zero value in the WQS method. When $\beta_1 > \beta_2 = 0$ but $\theta_1 = 0 < \theta_2$, X_1 participates in the regression equation *Y*|*X*, *M*, but does not in *M*|*X*. On the other hand, X_2 participates in the regression equation *M*|*X,*

but not in *Y*|*X*, *M*. This reduces the magnitude of $ME_{12}^{M|X}$ since $\hat{\gamma}$ from *Y*|*X*, *M* is multiplied by $\hat{\theta}_{12}^{M|WQS_{index}^{M|X}}$ from the *M*|*X* to get the $ME_{12}^{M|X}$ estimate. The nested method has improved type1 error when $\rho_{12} = \{0.5, 0.95\}$, because the correlation between X_1, X_2 makes the null values $\beta_2 = 0$ and $\theta_1 = 0$ less critical.

Table 4.3 shows that the nested bootstrap method has a lower type1 error in more conditions (36) than the sequential bootstrap method (24=60-36), due to its larger confidence interval width. Consequently, the nested method has a higher power in more conditions (31) than the sequential method (17=48-31). This is shown in the 60% of the 60 conditions where the nested method has a lower type1 error rate than the sequential method.

Table 4.3

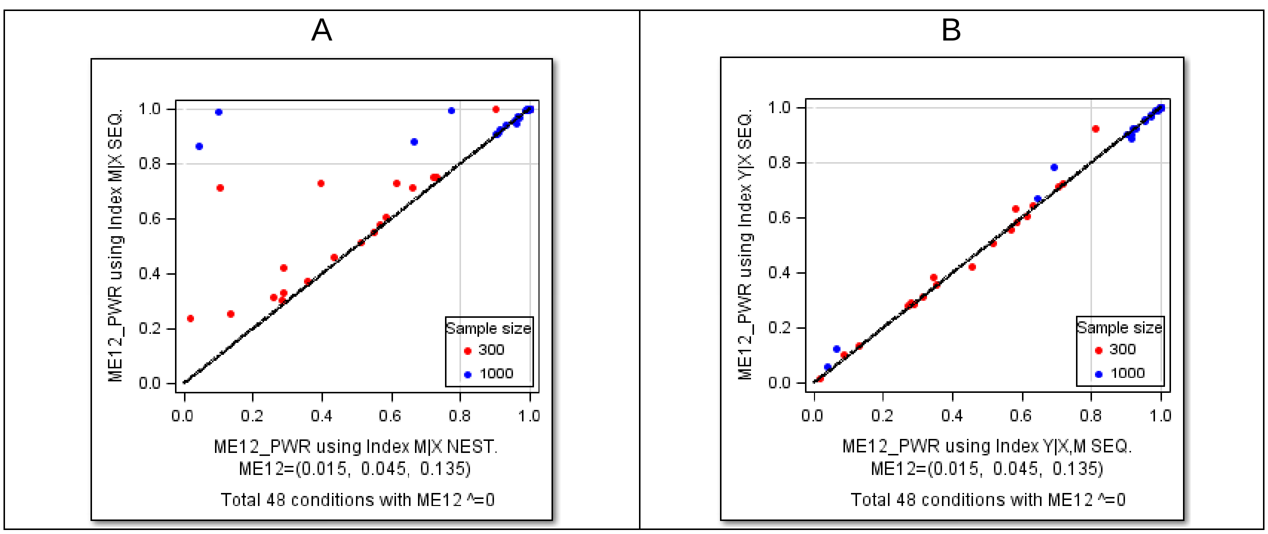
Nested vs. Sequential Bootstraps for $ME_{12}^{M|X}$ Type1 Errors (60) vs. Power (48)

Sample size	Smaller type1 error nested (36) vs. sequential	Larger power sequential (31) vs. nested
300	17	18
1000	19	13
	60 conditions	48 conditions

The difference in the type1 error rate between the nested vs. sequential methods is (0.03), but the difference in the sequential method having a higher power over the nested method is a wide range (0, 0.89). The nested method for determining the percentile confidence interval for conditions with opposing associations between the theta and beta parameters $(\theta_2 > \theta_1$ and $\beta_2 < \beta_1)$ and non-zero $\gamma \in \{0.3, 0.1\}$ with large

sample sizes (N=1000), performs at almost zero coverage probability and power. The wider confidence interval for the nested method combined with the small effect sizes (0.045, 0.015) affects the power to detect the joint mediated effect $ME_{12}^{M|X}$. However, the sequential method performs consistently better for the same conditions. Based on this information and supported by Figures 4.2, the sequential bootstrap method is preferred over the nested bootstrap method for calculating the percentile confidence intervals for all indirect effects ME_{12}^{WQS} .

WQS sequential bootstrapped method performed well over the 48 conditions when $ME_{12} \neq 0$ for detecting $ME_{12}^{M|X}$ as compared to the nested bootstrap method, used to determine the joint mediated effect's percentile confidence interval Figure 4.5A. Comparing Figures4.5A - $WQS_{index}^{M|X}$, B - $WQS_{index}^{Y|X,M}$ vs. $WQS_{index}^{Y|X}$ shows that $WQS_{index}^{M|X}$ is preferred over the latter two in Figure4.5B. $WQS_{index}^{Y|X,M}$ & $WQS_{index}^{Y|X}$ have some conditions that result in almost zero power to detect ME_{12} .

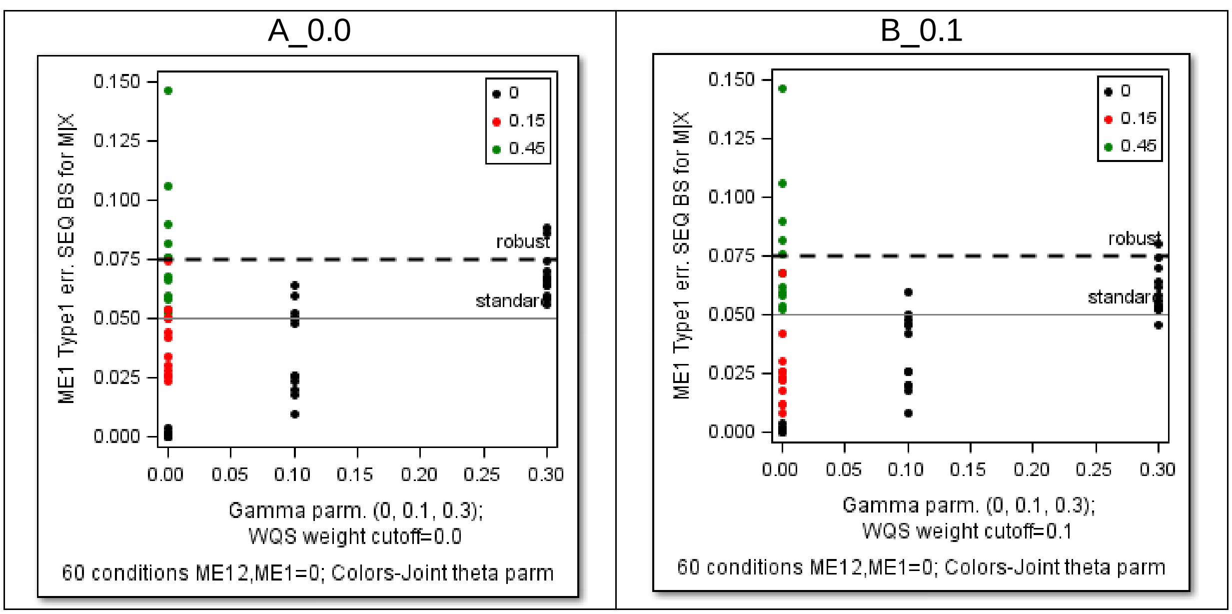


Figures 4.5 Power Detecting $ME_{12}^{M|X}$ A) Seq. vs. Nested, B) Seq. y= $WQS_{index}^{Y|X}$, x= $WQS_{index}^{Y|X,M}$

4.1.1.3 WQS Type 1 Analysis of 2-Variable Individual Mediated Effects

The WQS method used to calculate the type 1 error for an individual mediated effects is to first verify that the joint effect has a type1 error (a significant effect when the true value is zero), and subsequently examine if the j^{th} predictor's weight $\bar{w}_j^{M|X}, \bar{w}_j^{Y|X}, \bar{w}_j^{Y|X,M}$, depending on $WQS_{index}^{M|X}, WQS_{index}^{Y|X}, WQS_{index}^{Y|X,M}$ used in the mediation regression equations is above a threshold (cut-off values 0.1 to 0.5 in steps of 0.1). If the j^{th} predictor's weight is above the threshold, then it is considered an important predictor to be included in the set which influences the joint mediated effect's type1 error. This is a method for WQS variable selection based on the importance of the independent variable in determining the joint mediated effect on the outcome.

Type 1 error for indirect effects $WQS^{M|X}$ for ME_1 with cut-off values 0.0 to 0.5 by 0.1 is shown in Figure 4.6.



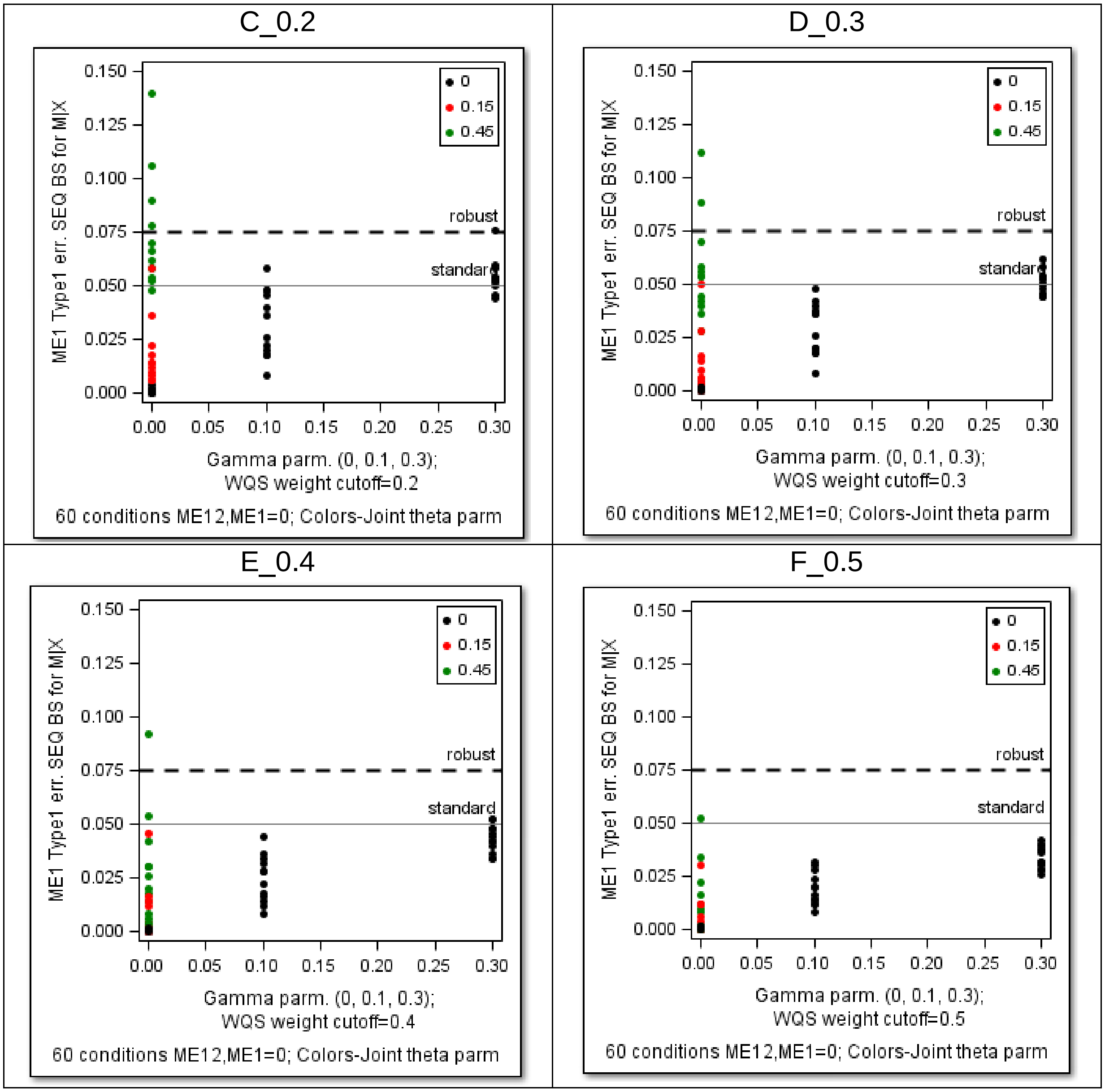


Figure 4.6 A0.1-F0.5 X_1 Type1 $ME_1^{M|X}$ Using $WQS_{index}^{M|X}$ 0 to 0.5 by 0.1 Sequential Method

Exceptions from the plots in Figures 4.6 where type1 error >0.075, which decrease as the cut-off value is increased, are detailed in Table 4.4. The index $WQS_{index}^{M|X}$ is used in the mediation model regression equations for calculating the X_1 mediated effect's type1 error $ME_1^{M|X}$.

Table 4.4

Exceptions in Type1 Errors (>0.075), Where $ME_1 = 0$ But $ME_{12}^{M|X} \,\&\, ME_1^{M|X} > 0$

N	ρ_{12}	Theta1	Theta2	Gamma	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.5	0	0	0.3	0.088	0.080	0.076			
1000	0	0.15	0.30	0	0.082	0.082	0.078	0.056		
300	0.5	0.15	0.30	0	0.076	0.076	0.070			
1000	0.5	0.15	0.30	0	0.090	0.090	0.090	0.070		
300	0.95	0.15	0.30	0	0.076	0.068				
1000	0.95	0.15	0.30	0	0.146	0.146	0.140	0.112	0.092	0.052
1000	0.5	0	0	0.3	0.086	0.074				
1000	0.95	0.15	0.30	0	0.106	0.106	0.106	0.088	0.054	

Sixty conditions potentially had type1 errors when $ME_1 = 0$ but with $ME_{12}^{M|X} \,\&\, ME_1^{M|X} > 0$, of which 8 conditions exceeded the a priori set type1 error limit of 0.075. The type1 error rate for predictor X_1 was investigated only if the joint mediated effect had a type1 error and amongst those cases, where the mediated effect for X_1 also had a type1 error. The eight exceptions are shown in Table 4.5. Six of eight exceptions for $ME_1^{M|X}$ had a zero gamma parameter and both non-zero theta parameters $\gamma = 0,\ \theta_1 = 0.15,\ \text{and}\ \theta_2 = 0.30$. In WQS regression, $\hat{\theta}_{12}^{M|WQS_{index}^{M|X}}$ is obtained by regressing *M* on $WQS_{index}^{M|X}$ giving it a negative bias, and $\hat{\gamma}_{12}^{M|WQS_{index}^{M|X}}$ is obtained by regressing *Y* on $WQS_{index}^{M|X}$ adjusting for *M*, giving $\hat{\gamma}_{12}^{M|WQS_{index}^{M|X}}$ a positive bias, resulting in a joint mediated estimate $ME_{12}^{M|X}$ with a positive bias. $ME_1^{M|X}$ is determined by the WQS weight $wX_1^{M|X}$. The relative WQS weights $wX_1^{M|X}, wX_2^{M|X}$ depend on the pairwise correlation between the two predictors as $\rho_{12} \in \{0, 0.5, 0.95\}$. $ME_1^{M|X}$ has a positive bias which increases as ρ_{12} increases from 0 to 0.95. This pattern of type1 errors (0.080, 0.090, and 0.146) can be

inferred from Table 4.4 in the column with 0cutoff, N=1000, with $\theta_1 = 0.15,\ \theta_2 = 0.30$ and γ=0, as the pairwise correlation increases $\rho \in \{0, 0.5, 0.95\}$. However, if the sample size drops to N=300, the standard error of the individual mediated estimate increases, resulting in a wider confidence interval around $ME_1^{M|X}$ and a lower type1 error of 0.076. Cases 1 and 7 have $(\theta_1, \theta_2) = 0,\ \gamma$=0.3, $\rho_{12} = 0.5;\ N = 1000$. The WQS weights are positively biased for $\hat{\theta}_{12}^{M|WQS_{index}^{M|X}}$ which when multiplied by $\hat{\gamma}_{12}^{M|WQS_{index}^{M|X}}$ gives a positive biased $ME_1^{M|X}$ with a high *RMSE*, increasing the type1 error to 0.088 and 0.086 respectively. Case 1 has a slightly higher type1 error since its parameter set has a high joint beta parameter value of $\beta_{12} = 0.45$(not shown in Table 4.5), while case 7 has $\beta_{12} = 0.30$, and $\gamma = 0.3$, which affects the positive bias on $\hat{\theta}_{12}^{M|WQS_{index}^{M|X}}\gamma$ more for case 1 than case 7.

In the last case listed in Table 4.4 with N=1000, the beta parameters for X_1, X_2 ($\beta_1 = 0.30 > \beta_2 = 0$) (not shown in Table 4.4), are opposing in magnitude to the theta parameters for X_1, X_2 ($\theta_1 = 0.15 < \theta_2 = 0.30$). These conditions create a mediated effect bias which changes sign as the pairwise correlation increases from 0 to 0.95, resulting in a lower type1 error (0.106), as compared to case 6 with similar conditions, but $\beta_1 = 0.15 < \beta_2 = 0.30$, $\theta_1 = 0.15 < \theta_2 = 0.30$ having a type1 error (0.146). The individual mediated effect for X_1 using $WQS_{index}^{Y|X}$ is shown in Table 4.5.

Table 4.5

Exceptions in Type1 Errors (>0.075), when $ME_1 = 0$ But $ME_{12}^{Y|X} \& ME_1^{Y|X} > 0$

N	ρ_{12}	Theta1	Theta2	Gamma	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.95	0	0.15	0	0.076	0.076	0.070			
1000	0	0.15	0.30	0	0.082	0.080	0.078	0.052		
300	0.5	0.15	0.30	0	0.076	0.076	0.064			
1000	0.5	0.15	0.30	0	0.088	0.088	0.088	0.074		
300	0.95	0.15	0.30	0	0.078	0.070				
1000	0.95	0.15	0.30	0	0.144	0.144	0.138	0.114	0.072	
1000	0.95	0.15	0.30	0	0.098	0.098	0.098	0.098	0.088	0.080

Sixty conditions potentially had type1 errors when $ME_1 = 0$ but $ME_{12}^{Y|X} \& ME_1^{Y|X} > 0$ of which 7 conditions exceeded the a priori type1 error limit of 0.075. A similar analysis of the type 1 errors for different parameter settings than those shown Table 4.5 can be done, resulting in similar inferences for $ME_1^{Y|X}$. The individual mediated effect for *X₁* when using $WQS_{index}^{Y|X,M}$ is shown in Table 4.6.

Table 4.6

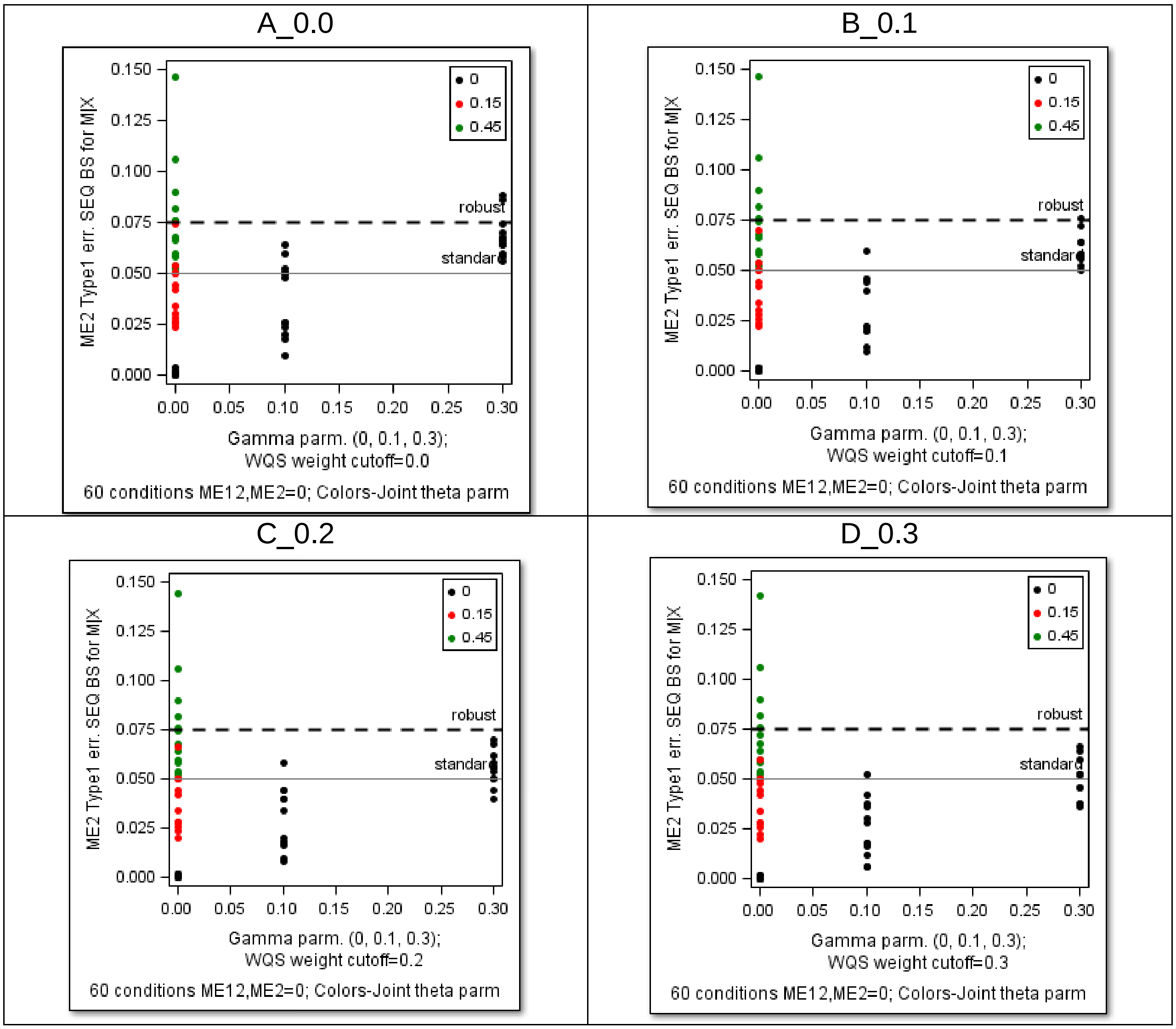
Exceptions in Type1 Errors (>0.075), Where $ME_1 = 0$ But $ME_{12}^{Y|X,M} \& ME_1^{Y|X,M} > 0$

N	ρ_{12}	θ_1	θ_2	γ	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.95	0	0.15	0	0.076	0.076	0.070			
1000	0	0.15	0.30	0	0.082	0.080	0.076	0.052		
300	0.5	0.15	0.30	0	0.080	0.080	0.070			
1000	0.5	0.15	0.30	0	0.088	0.088	0.088	0.074		
300	0.95	0.15	0.30	0	0.078	0.068				
1000	0.95	0.15	0.30	0	0.146	0.146	0.136	0.114	0.072	
1000	0.95	0.15	0.30	0	0.098	0.098	0.098	0.098	0.086	0.082

Sixty conditions potentially had type1 errors where $ME_1 = 0$, $ME_{12}^{Y|X,M} \& ME_1^{Y|X,M} > 0$ of which 7 conditions exceeded the *a priori* set type1 error limit of 0.075. A similar

analysis of the type 1 errors for different parameter settings than those shown in Table 4.5 can be done, resulting in very similar inferences for $ME_1^{Y|X,M}$.

Individual predictor X_2 type1 errors for indirect effect *using* $WQS_{index}^{M|X}$ for ME_2 cutoffs 0 to 0.5 by 0.1 are show in Figure 4.7.



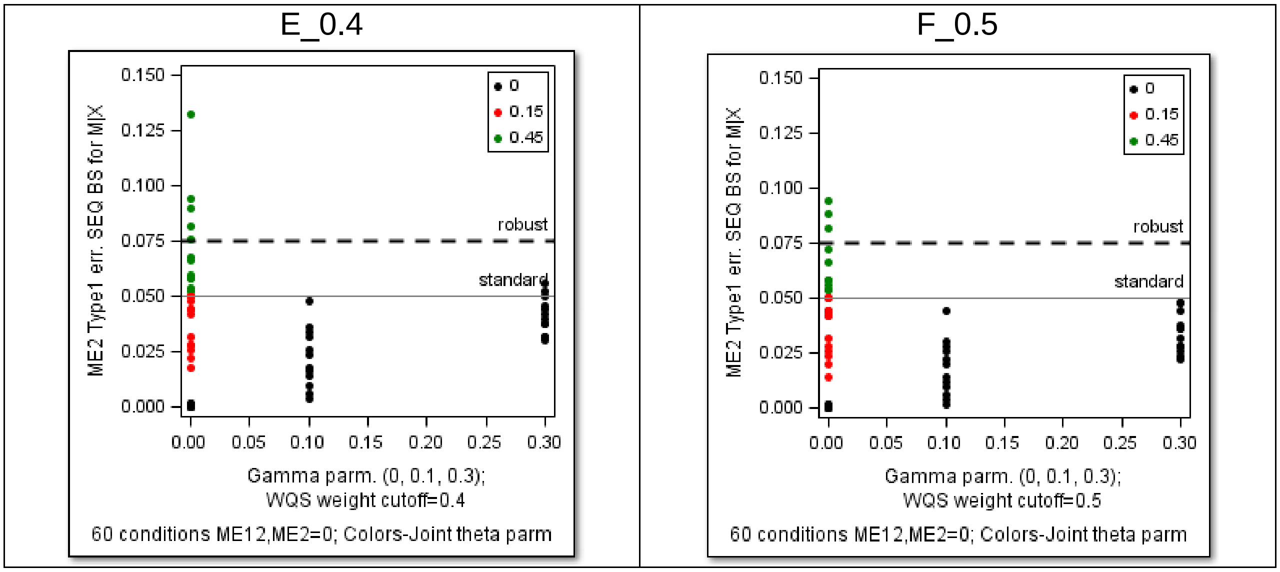


Figure 4.7 A0.1-F0.5 X_2 Type1 Errors $ME_2^{M|X}$ w/ $WQS_{index}^{M|X}$ 0-0.5 by 0.1 Sequential Method

Exceptions from the plots in Figures 4.7 where increased cut-off values decrease the type1 error >0.075 exceptions, are tabulated in Table 4.7.

Table 4.7

Exceptions in Type1 Errors (>0.075), Where $ME_2 = 0$ But $ME_{12}^{M|X}$ & $ME_2^{M|X} > 0$

N	ρ_{12}	Theta1	Theta2	Gamma	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.5	0	0	0.3	0.088	0.076	0.070			
1000	0	0.15	0.30	0	0.082	0.082	0.082	0.082	0.082	0.082
300	0.5	0.15	0.30	0	0.076	0.076	0.076	0.076	0.076	0.066
1000	0.5	0.15	0.30	0	0.090	0.090	0.090	0.090	0.090	0.088
300	0.95	0.15	0.30	0	0.076	0.074				
1000	0.95	0.15	0.30	0	0.146	0.146	0.144	0.142	0.132	0.094
1000	0.5	0	0	0.3	0.086	0.072				
1000	0.95	0.15	0.30	0	0.106	0.106	0.106	0.106	0.094	0.072

Sixty conditions potentially had type1 errors where $ME_2 = 0$, $ME_{12}^{M|X}$ & $ME_2^{M|X} > 0$ of which 8 conditions exceeded the *a priori* type1 error limit of 0.075. The type1 error rate for predictor X_2 was investigated only if the joint mediated effect had a type1 error and within those cases if the mediated effect for X_2 also had a type1 error. The eight exceptions are shown in Table 4.8. Six of the eight conditions for $ME_2^{M|X}$ had a zero

gamma parameter and both non-zero theta parameters, $\gamma = 0,\ \theta_1 = 0.15$ and $\theta_2 = 0.30$, while the last two conditions had beta parameters for X_1, X_2 ($\beta_1 = 0.30, \beta_2 = 0$), which were opposing in magnitudes to the theta parameters for X_1 & X_2 ($\theta_1 = 0.15,\ \theta_2 = 0.30$). The 1st and the 7th cases have a non-zero gamma parameter with both zero theta parameters $\gamma = 0.3,\ \theta_1 = 0 \text{ and } \theta_2 = 0$. The relative WQS weights depend on the pairwise correlation between the two predictors as $\rho_{12} \in \{0,\ 0.5,\ 0.95\}$ increases. For this reason, $ME_2^{M|X} = 0$ starts with a positive bias and increases as ρ_{12} increases from 0 to 0.95. The pattern of type1 errors (0.082, 0.090, and 0.146) can be inferred from Table 4.8 in the column with $\hat{\theta}_2^{weight}$ for cut-off =0 as pairwise correlations increase $\rho \in \{0,\ 0.5,\ 0.95\}$. If the sample size drops to N=300, $ME_2^{M|X}$'s standard error increases resulting in a lower type1 error of 0.076. In the remaining two cases with N=1000, with $\theta_1 = 0,\ \theta_2 = 0,\ \gamma = 0.3$ and $\theta_1 = 0.15 < \theta_2 = 0.30,\ \gamma = 0$, the beta parameters for X_1, X_2 ($\beta_1 = 0.30 > \beta_2 = 0$), are opposing in magnitude to the theta parameters for X_1, X_2. These parameter settings cause $ME_2^{M|X}$ bias to change sign as the pairwise correlation increases from 0 to 0.95 resulting in a lower type1 error (Case 7 vs. 9 with $\rho_{12} = 0.95$, decreasing the Type 1 error from 0.146 to 0.106; Case 1 vs. 8 both with $\rho_{12} = 0.5$, decreasing the Type 1 error from 0.088 to 0.086). The reduction in type1 error is smaller for the latter comparative pair since there is a point of inflection around $\rho_{12} = 0.5$.

Exceptions in type1 errors (>0.075) where $ME_2 = 0$, but $ME_{12}^{Y|X}$ and $ME_2^{Y|X} > 0$ are tabulated in Table 4.8.

Table 4.8

Exceptions in Type1 Errors (>0.075), Where $ME_2 = 0$ But $ME_{12}^{Y|X}$ & $ME_2^{Y|X} > 0$

N	ρ_{12}	Theta1	Theta2	Gamma	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.95	0	0.15	0	0.076	0.076	0.076	0.074		
1000	0	0.15	0.30	0	0.082	0.082	0.082	0.082	0.082	0.082
300	0.5	0.15	0.30	0	0.076	0.076	0.076	0.076	0.076	0.074
1000	0.5	0.15	0.30	0	0.088	0.088	0.088	0.088	0.088	0.088
300	0.95	0.15	0.30	0	0.078	0.078	0.076	0.074		
1000	0.95	0.15	0.30	0	0.144	0.144	0.144	0.138	0.128	0.108
1000	0.95	0.15	0.30	0	0.980	0.098	0.084	0.066		

Sixty conditions potentially had type1 errors where $ME_2 = 0$, $ME_{12}^{Y|X}$ & $ME_2^{Y|X} > 0$ of which 7 conditions exceeded the a priori type1 error limit of 0.075. A similar analysis of the type 1 errors for the different parameter settings than those shown in the Table 4.8 can be done resulting in similar inferences for $ME_2^{Y|X}$. The mediated effect's type1 error exceptions using index $WQS_{index}^{Y|X,M}$ for *X₂* $ME_2^{Y|X,M}$ are tabulated in Table 4.9.

Table 4.9

Exceptions in Type1 Errors (>0.075), Where $ME_2 = 0$ But $ME_{12}^{Y|X,M}$ & $ME_2^{Y|X,M} > 0$

N	ρ_{12}	Theta1	Theta2	Gamma	0cutoff	0.1cutoff	0.2cutoff	0.3cutoff	0.4cutoff	0.5cutoff
1000	0.95	0	0.15	0	0.076	0.076	0.076	0.074		
1000	0	0.15	0.30	0	0.082	0.082	0.082	0.082	0.082	0.082
300	0.5	0.15	0.30	0	0.080	0.080	0.080	0.080	0.080	0.076
1000	0.5	0.15	0.30	0	0.088	0.088	0.088	0.088	0.088	0.088
300	0.95	0.15	0.30	0	0.078	0.078	0.074	0.072		
1000	0.95	0.15	0.30	0	0.146	0.146	0.144	0.136	0.130	0.110
1000	0.95	0.15	0.30	0	0.098	0.078	0.060			

Sixty conditions potentially had type1 errors where $ME_2 = 0$, $ME_{12}^{Y|X,M}$ & $ME_2^{Y|X,M} > 0$ of which 7 conditions exceeded the a priori type1 error limit of 0.075. A similar analysis

of the type 1 errors for the different parameter settings than those shown in the Table 4.8 can be done resulting in very similar inferences for $ME_2^{Y|X,M}$.

The inference from the results in Tables 4.4 through 4.10 is that $WQS_{index}^{M|X}$ is the preferred WQS index to use when predicting the individual mediated effects, but with a 0.3 cut-off for $wX_1^{M|X}, wX_2^{M|X}$ below which these WQS weights should be considered as zero. $WQS_{index}^{Y|X}$, $WQS_{index}^{Y|X,M}$ are poor WQS indices to use when estimating individual mediated effects, especially in the conditions when the regression coefficient $\hat{\gamma}_{12}$ which is obtained by regressing *Y|M*, WQS_{index} is close to zero or when $(\theta_1, \theta_2) = 0 \text{ and } \gamma \neq 0$. These conditions may be avoided by starting with the correct choice of mediator *M*, which has a correlation with the outcome *Y*, so that $\hat{\gamma}_{12}^{M|WQS_{index}^{M|X}} \neq 0$ and *M* has a correlation with at least one predictor so that $\hat{\theta}_{12}^{M|WQS_{index}^{M|X}} \neq 0$. There is a very small mediated effect to detect when either $\hat{\gamma}_{12}^{M|WQS_{index}}, \hat{\theta}_{12}^{M|WQS_{index}^{M|X}}$ is close to zero, resulting in increased type 1 errors.

4.1.1.4 Power for Indirect Effect Using $WQS_{index}^{M|X}$ for *ME₁*, Cutoff Values 0-0.5 by 0.1

Using any value for the cut-offs other than 0.0 with WQS weights reduces the power to detect $ME_1^{M|X}$ as can be seen in the figures below. The most conservative cut-off value for type1 errors seen in Figure 4.6 and Figure 4.7 for the individual 2-variable mediated effects $ME_1^{M|X}$ and $ME_2^{M|X}$, was 0.30 below which the individual WQS weight could be considered to be zero i.e. that variable has no effect on the outcome. The

power to detect the individual effect $ME_1^{M|X}$ rapidly deteriorates for cut-off values greater than 0.20.

The power to detect individual mediated effects with parameter values $ME_1 = (0.015,\ 0.045)$, and $ME_2 = (0.015,\ 0.030,\ 0.045, 0.090)$ are shown in Figure 4.8.

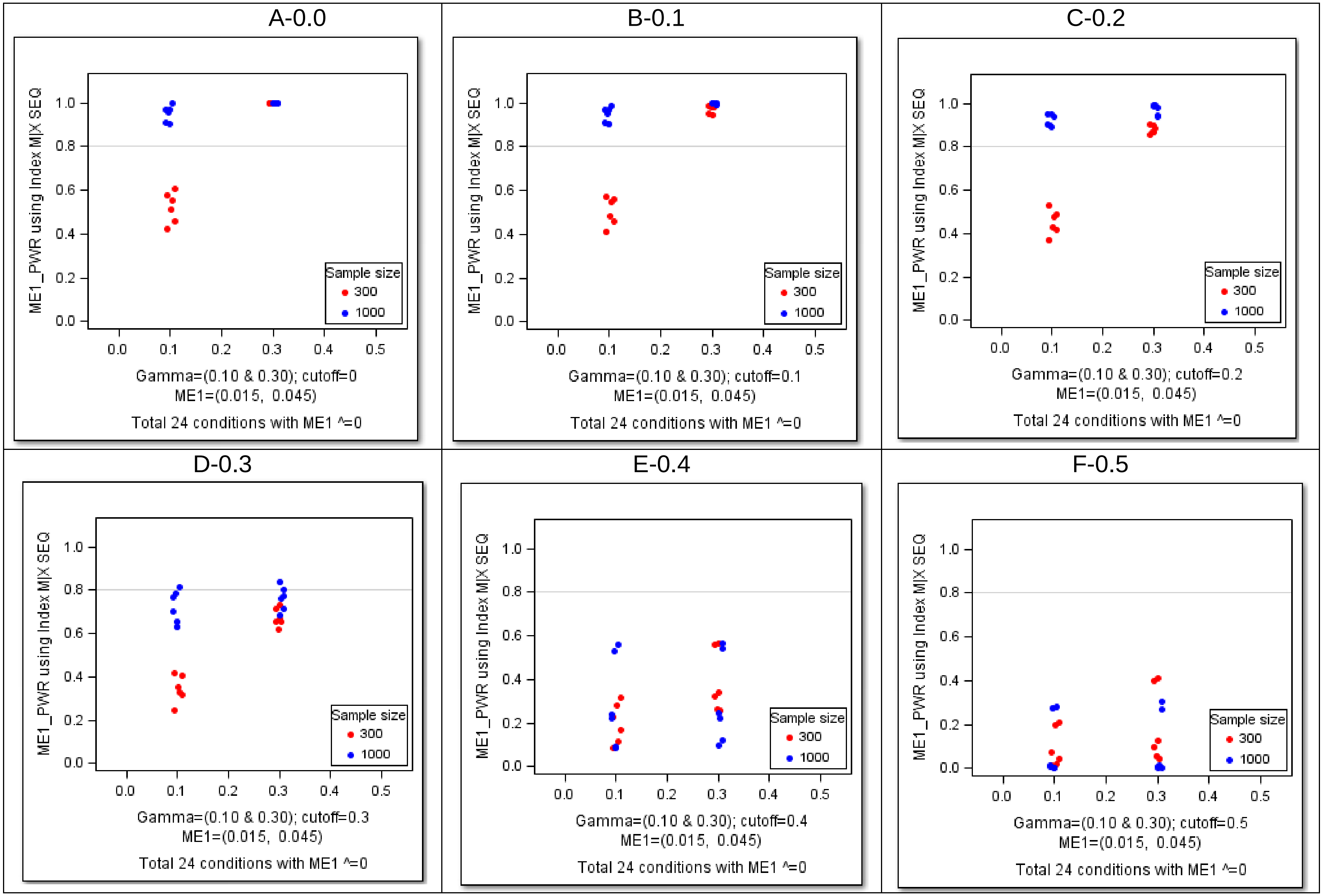
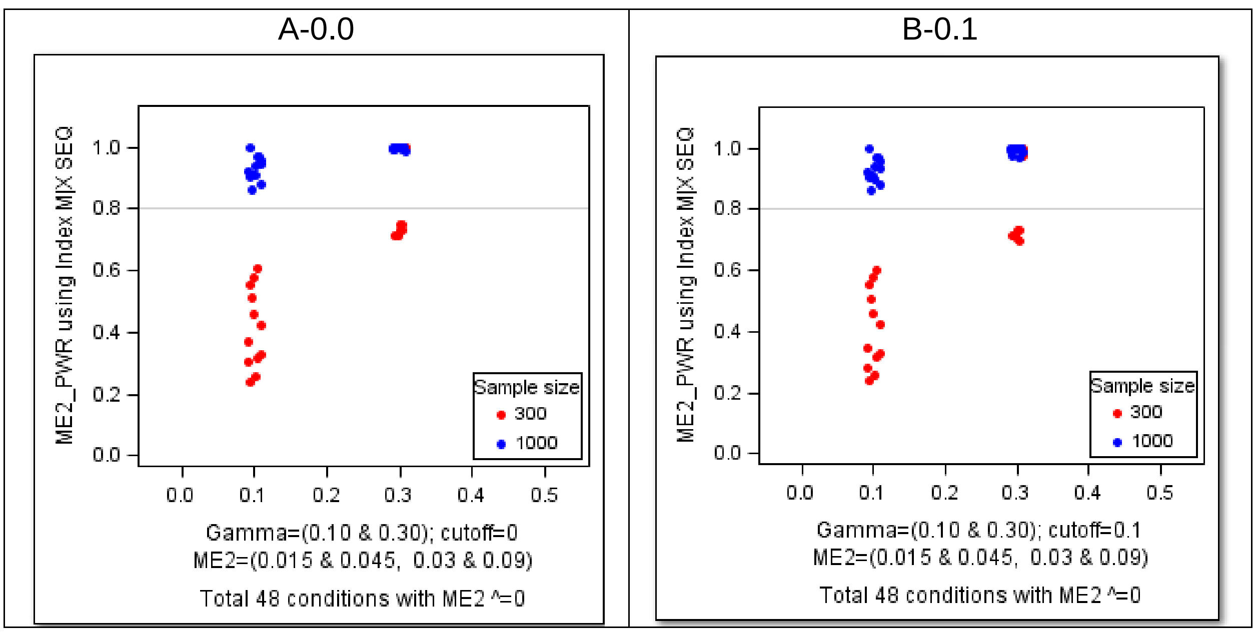


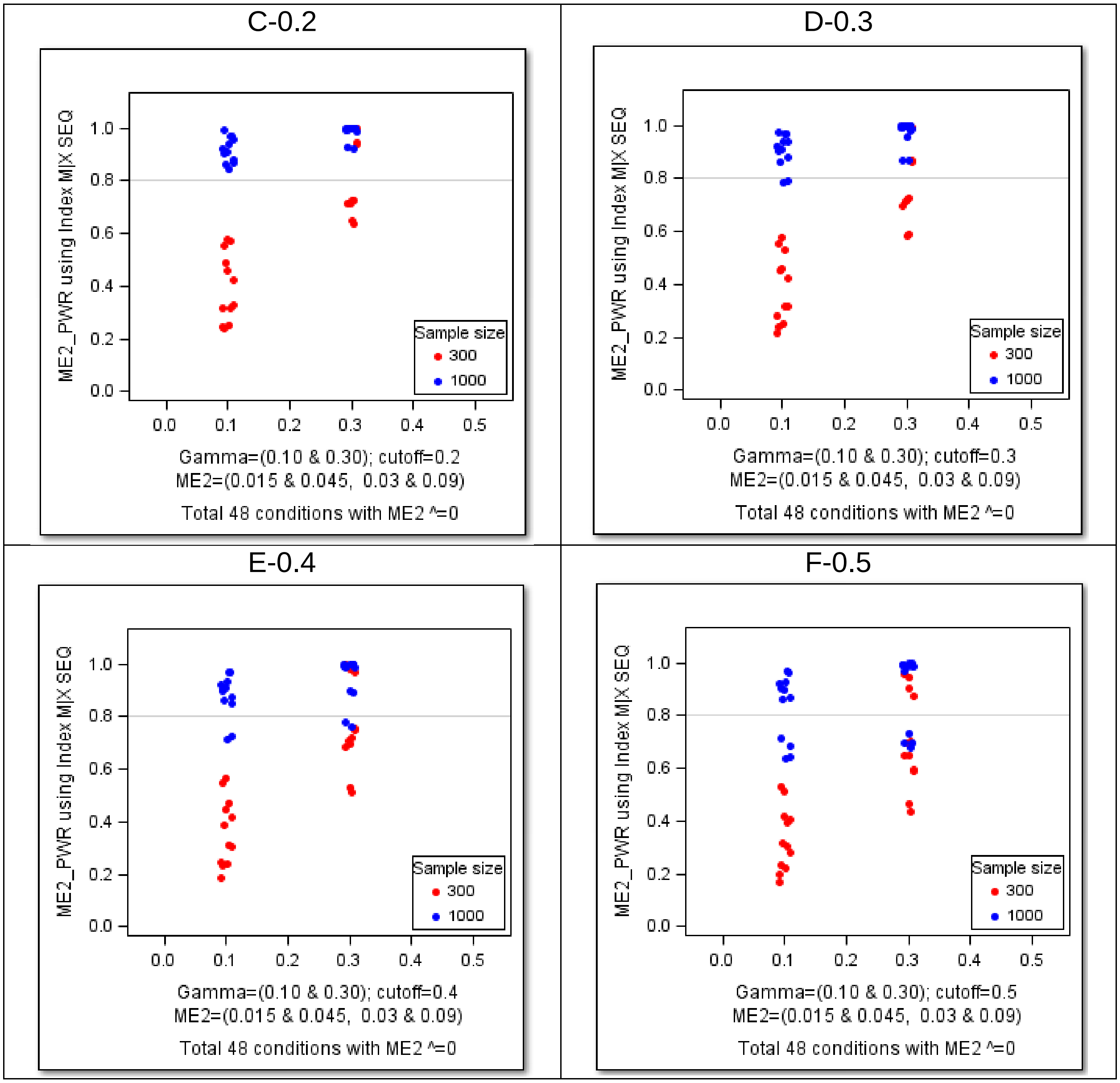
Figure 4.8 A-0.0 to F-0.5 by 0.1 Power to Detect $ME_1^{M|X}$ Using Index $WQS_{index}^{M|X}$ for Individual Predictor Weight Cut_Offs

4.1.1.5 Power for Indirect Effect using $WQS_{index}^{M|X}$ for ME_2, Cut-off Values 0-0.5 by 0.1

Using any value for the cut-offs other than 0.0 to 0.20 for the individual WQS weights reduces the power to detect $ME_1^{M|X}$ significantly as can be seen in the figures below. The most conservative cut-off value for type1 errors seen in Figures 4.6 and Figures 4.7 for the individual mediated effects $ME_1^{M|X}$ & $ME_2^{M|X}$ was 0.30 below which the WQS weight could be considered to be zero i.e. the variable has no effect on the outcome. In the case the power for these individual effects $ME_1^{M|X}$ & $ME_2^{M|X}$, the power to detect the individual effect $ME_2^{M|X}$ rapidly deteriorates for cut-off greater than 0.30.

$WQS_{index}^{M|X}$ Used in Mediation for X_2 Mediated Effect's Power to Detect $ME_2^{M|X}$

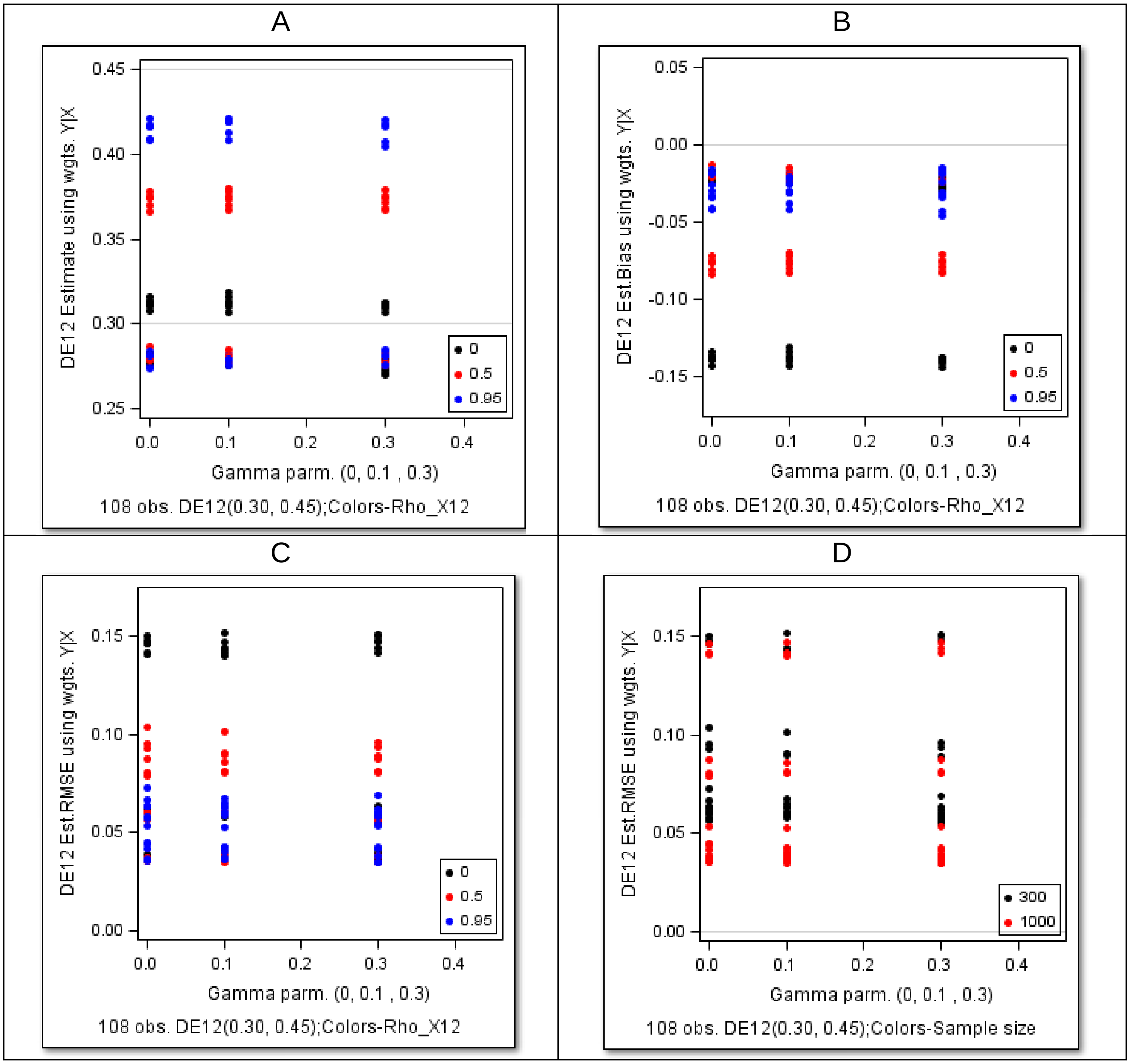




Figures 4.9 Power to Detect $ME_2^{M|X}$; $ME_2 = (0.015 \ \& \ 0.045, \ 0.03 \ \& \ 0.09)$ Using $WQS_{index}^{M|X}$

4.1.1.6 WQS Method Joint Direct effects for Two-Variable Mediation DE_{12}

Joint Direct Effects' WQS *Estimate*, *Bias* and *RMSE* w/Weights from $WQS_{index}^{Y|X}$



Figures 4.10 Joint Direct Effects *Estimate*, *Bias* (A-B), *RMSE* (C-D) Using $WQS_{index}^{Y|X}$

The estimated joint direct effects in Figures 4.10 show that the pairwise predictor correlation influences the $DE_{12}^{Y|X}$ estimate value and its negative bias, when theta and beta parameters related to X_1, X_2 are congruent as in $\theta_1 < \theta_2, \beta_1 < \beta_2$, while the gamma parameter and sample size have little influence. However, when theta and beta parameters related to (X_1, X_2) are incongruent as in $\theta_1 < \theta_2, \beta_1 > \beta_2 = 0$, neither the correlation parameter nor the gamma parameter have much influence on the estimate

$DE_{12}^{Y|X}$ or its negative bias. The parameter $\beta_2 = 0$ (*X_2* having no association with *Y* adjusted for M in *Y|X, M*), but *X_2* having a strong association with *M* in *M|X* since $\theta_2 > \theta_1 > 0$, results in a negative bias for the joint indirect effects but not the joint direct effects. In both conditions discussed above, gamma has little influence on the joint direct effect or its negative bias since it is a covariate in the regression *Y|X*, *M* but it is *X_1*, *X_2* pairwise correlation which influences $WQS_{index}^{Y|X}$ used in calculating the mediated effects. Increased correlation between *X_1*, *X_2* increases the estimate's value for $DE_{12}^{Y|X}$ and reduces the bias on the joint estimate. When $\rho_{12} = 0, \gamma = 0$, the $DE_{12}^{Y|X}$ estimate bias is most negative for a large sample size (-0.143), but an increased correlation between *X_1*, *X_2*, decreases the negative bias for the $DE_{12}^{Y|X}$ estimate (-0.041) and remains unchanged with increasing values for the gamma parameter. The $DE_{12}^{Y|X}$ estimate's *RMSE* increases with reduced sample size but decreases with increased pairwise predictor correlations. It has the highest values (0.03, 0.05) for a small sample size and a zero pairwise correlation as shown in Figures 4.10 C & D. By definition, an estimate's *RMSE* directly increases with its parameter value. The coverage probability for DE_{12} for108 conditions are discussed below.

4.1.1.7 Coverage Probability for DE_{12} Using $WQS^{Y|X}_{index}$ vs. $WQS^{Y|X,M}_{index}$, $WQS^{M|X}_{index}$

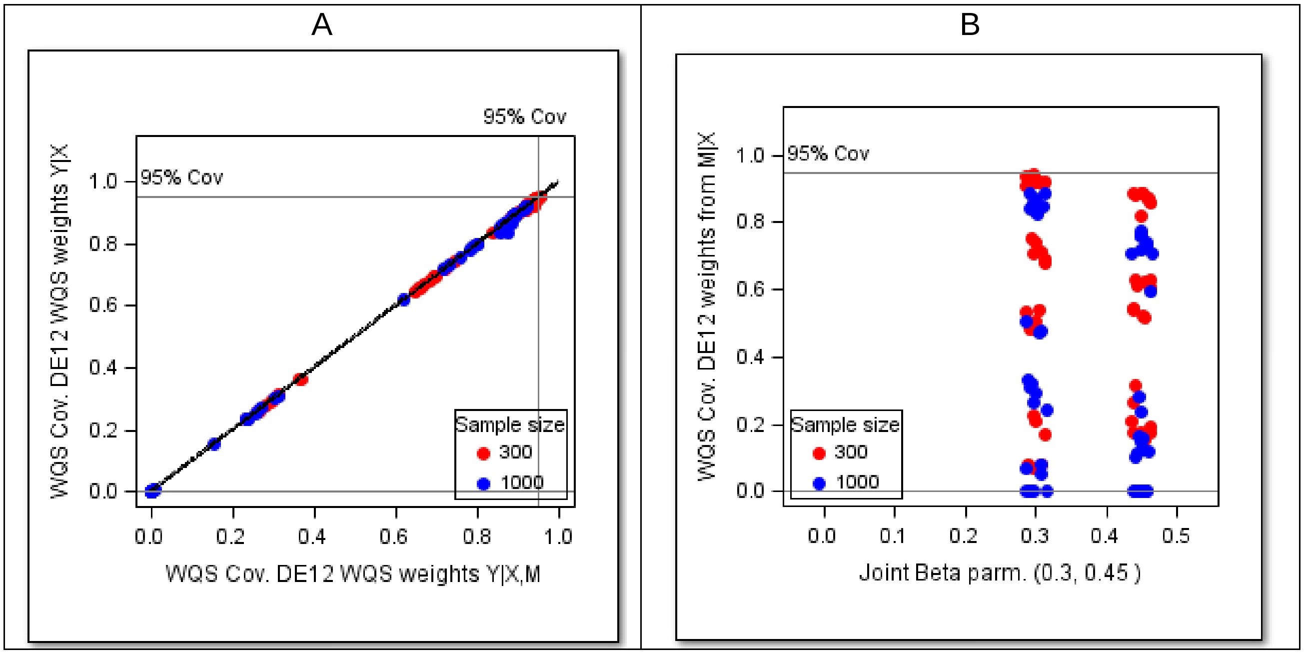


Figure 4.11 Coverage Probability for DE_{12} Using A) $WQS^{Y|X}_{index}$ vs. $WQS^{Y|X,M}_{index}$ B) $WQS^{M|X}_{index}$

All three WQS indices: A) $WQS^{Y|X}_{index}$ vs. $WQS^{Y|X,M}_{index}$ and B) $WQS^{M|X}_{index}$ have conditions with parameter values that have zero coverage. A detailed examination of the near zero coverage cases of the joint direct effect DE_{12} for $WQS^{Y|X}_{index}$ shows that the effect estimates result from the predictors having zero pairwise correlations, large sample sizes, and large negative biases (proportional to the parameter value). Conditions with $\theta_1 \le \theta_2$, $\beta_1 > \beta_2 = 0$ for the joint direct effect DE_{12}=0.3, have a high coverage regardless of gamma, predictor correlations and theta parameter values, since the effect size is a non-zero, positive value. The three conditions resulting in a high coverage for the joint direct effect are for a small sample size, high pairwise correlation, congruent theta and beta parameters for *X1*,*X2* ($\theta_1 > \theta_2$ and $\beta_1 > \beta_2 = 0$), and a large effect size $DE_{12} = 0.45$.

The inference that can be drawn from the coverage of the joint direct effect is that there were extremes in the coverage probabilities over the 108 conditions for two variable mediation, regardless of the weighted index used for X_1, X_2 and the narrow confidence interval due to a large sample size. The determining factor was the magnitude of the joint direct effect parameter, needing to be covered, by the WQS estimate's 95% confidence interval around.

4.1.1.8 Power to Detect Joint Direct Effects using Index $WQS_{index}^{Y|X}$

The power to detect $DE_{12}^{Y|X}$ using index $WQS_{index}^{Y|X}$ and $DE_{12}^{M|X}$ using $WQS_{index}^{M|X}$ are shown in Figure 4.12 for 108 conditions having the ability to detect a joint direct effect.

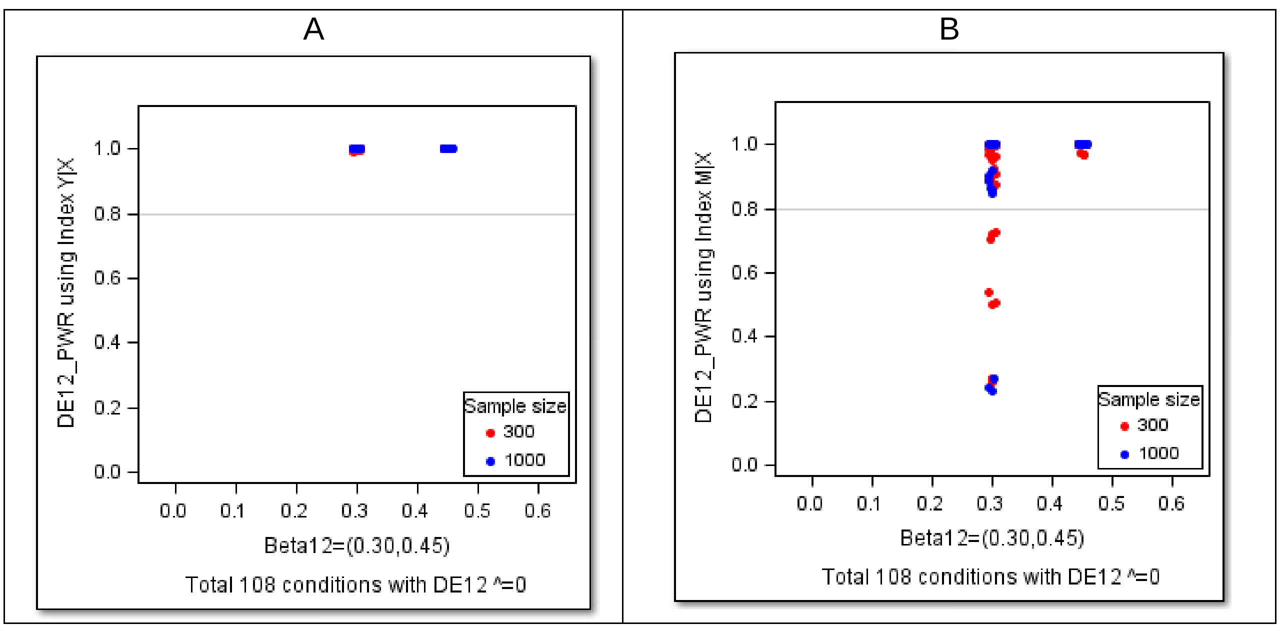


Figure 4.12 Power to detect DE_{12} When Using A) $WQS_{index}^{Y|X}$ and B) $WQS_{index}^{M|X}$

Inference is that $WQS_{index}^{Y|X}$ set of weights are preferred to $WQS_{index}^{M|X}$ when estimating DE_{12}.

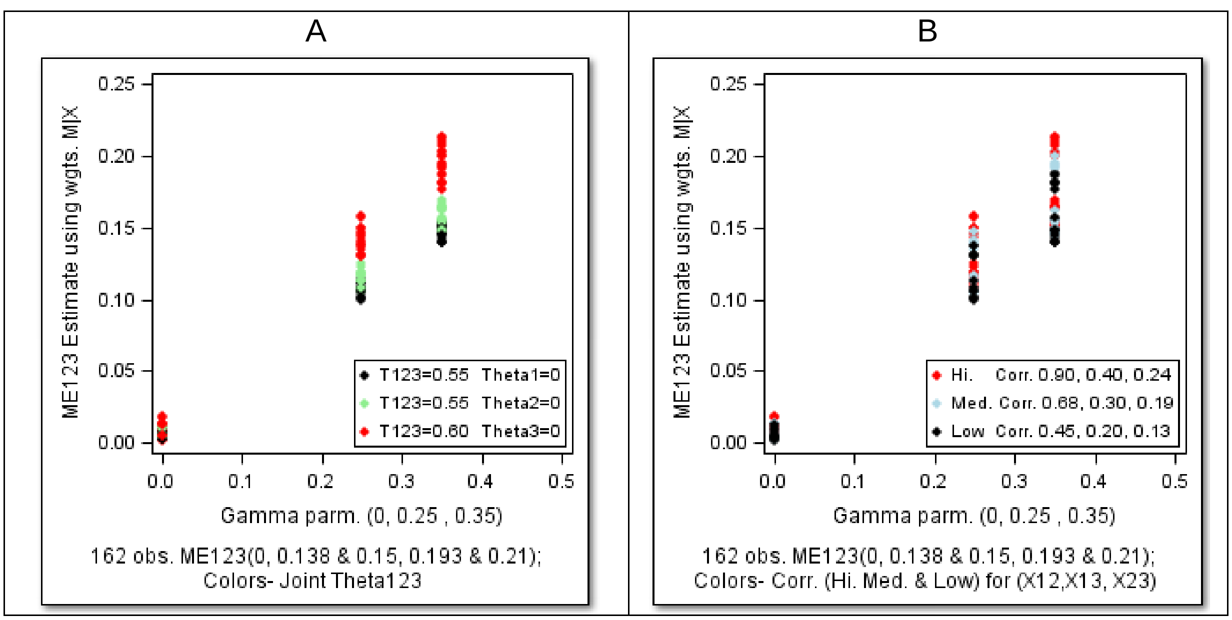
The cut-off for maximizing the power is a zero-cut-off value since the power declines as the cut-off value is increased to contain the exceptions for the type1 errors>0.075.

4.1.1.9 Individual Predictor's Influence on Joint Direct Effect using $WQS_{index}^{Y|X}$

There were 108 conditions with a non-zero beta1 parameter value having the power to detect a joint direct effect $DE_{12}^{Y|X} = \{0.45, 0.30\}$ using the index $WQS_{index}^{Y|X}$. The first 54 conditions had $\beta_1 = 0.15$, and the next 54 conditions had $\beta_1 = 0.30$. In all 108 conditions the power related to variable X_1's influence on the joint direct effect $\beta_{12} = \{0.45, 0.30\}$ was 1, since the beta parameter values were all large, non-zero values even with the cut-off for WQS predictor weights at 0.10 and 0.20. However, when the cut-off value was set to 0.3, there were 40 conditions for $\beta_1 = 0.15$ to detect the joint direct effect $DE_{12}^{Y|X} = 0.45$ with a much lower power.

WQS 3-Variable Mediation Analysis

4.1.2.1 Joint Mediated Effect Using $WQS_{index}^{M|X}, WQS_{index}^{Y|X} or WQS_{index}^{Y|X,M}$ *Est., Bias, RMSE*



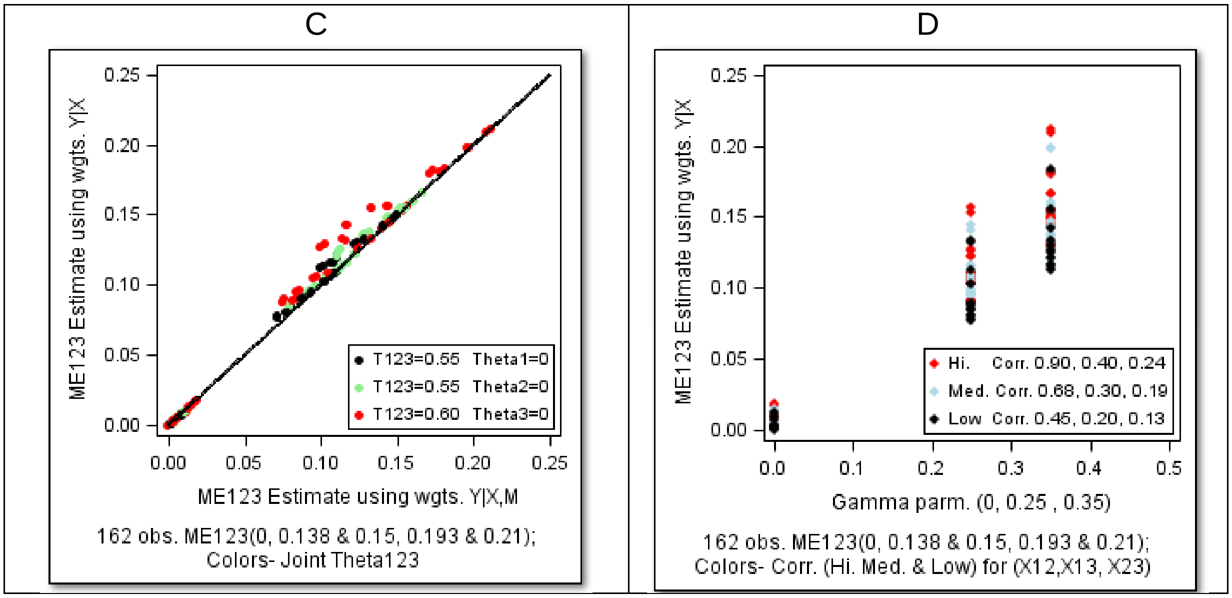


Figure 4.13 Joint Indirect Effect Est. A&B for $WQS^{M|X}_{index}$;C- $WQS^{Y|X}_{index}$ vs. $WQS^{Y|X,M}_{index}$ & D) $WQS^{Y|X}_{index}$

The estimated joint direct effects in Figures 4.13 A and B show the relationship between the joint mediated effect $ME^{M|X}_{123}$ estimates and the gamma parameter, grouped by the joint theta parameter $\theta_{123} = (\theta_1 + \theta_2 + \theta_3)$ and by the pairwise correlations. There were 54 $ME^{M|X}_{123}$ estimates associated with each gamma parameter value $\gamma = 0,$ and γ={0.25, 0.35}, with each gamma parameter value had three clusters of 18 estimates with joint theta parameters $\theta_{123} = \{0.60(\theta_3 = 0),\ 0.55(\theta_2 = 0)\ \&\ 0.55(\theta_1 = 0)\}$, and each joint theta parameter cluster had three sub-clusters of 6 estimates each with the joint beta parameter $\beta_{123} = \{0.35(\beta_3 = 0),\ 0.50(\beta_2 = 0)\ \&\ 0.45(\beta_1 = 0)\}$. The inference from Figure 4.13 A-D is that $WQS^{M|X}_{index}$ (panel A and B) or $WQS^{Y|X}_{index}$ (panel D) is better at estimating the joint mediated effect ME_{WQS} than $WQS^{Y|X,M}_{index}$ (panel C). All three indices work equally well.

Table 4.10

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0$, $\theta_{123} = 0.60(\theta_1 + \theta_2 + 0)$, $\rho_{12} = \{0.90, 0.68, 0.45\}$

Sl. #	$\theta_{123}^{M\|X}$ parm.=0.60	$\gamma_{WQS}^{M\|X}$ parm.= 0	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ parm.=0	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.554	0.033	$\rho_{12}=$ 0.90	0.019	0.019	0.041	0.920
2	0.555	0.031	$\rho_{12}=$ 0.90	0.017	0.017	0.026	0.864
3	0.532	0.025	$\rho_{12}=$ 0.68	0.014	0.014	0.039	0.934
4	0.532	0.027	$\rho_{12}=$ 0.68	0.014	0.014	0.024	0.884
5	0.503	0.024	$\rho_{12}=$ 0.45	0.012	0.012	0.034	0.924
6	0.503	0.025	$\rho_{12}=$ 0.45	0.012	0.012	0.022	0.878
7	0.550	0.016	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.009	0.009	0.036	0.950
8	0.554	0.024	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.013	0.013	0.024	0.900
9	0.532	0.010	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.006	0.006	0.033	0.950
10	0.531	0.016	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.009	0.009	0.020	0.932
11	0.508	0.003	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.001	0.001	0.033	0.954
12	0.506	0.011	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.006	0.006	0.018	0.958
13	0.552	0.005	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.003	0.003	0.037	0.940
14	0.552	0.015	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.008	0.008	0.021	0.944
15	0.537	0.008	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.004	0.004	0.036	0.936
16	0.531	0.006	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.003	0.003	0.019	0.944
17	0.504	0.007	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.004	0.004	0.033	0.944
18	0.502	0.009	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.005	0.005	0.018	0.952

When $\gamma = 0$ and $\theta_{123} = 0.60(\theta_3 = 0)$, there were 18 estimates for $ME_{123}^{M|X}$ grouped by $\beta_{123} = \{0.35(\beta_3 = 0),\ 0.50(\beta_2 = 0),\ \text{and}\ 0.45(\beta_1 = 0)\}$ in Table 4.10. The interest is in looking at the beta parameter set and the associated correlation between the influencing predictors for each of these 18 conditions. The focus is on

$\beta_{123} = 0.35(\beta_1 + \beta_2 + \beta_3\theta_3 = 0)$ and ρ_{12} for the first six cases,

$\beta_{123} = 0.50(\beta_1 + 0 + \beta_3\theta_3 = 0)$ and ρ_{13} for cases 7-12 and $\beta_{123} = 0.45(0 + \beta_2 + \beta_3\theta_3 = 0)$ and

ρ_{23} for the final six cases.

Predictors X_1 and X_2 are the important predictors for the set $\theta_{123} = 0.60(\theta_3 = 0),$ $\gamma = 0,$ and $\beta_{123} = 0.35(\beta_3 = 0)$ in determining the $ME_{123}^{M|X}$ estimate. The value for the regression coefficient $\hat{\theta}_{123}^{M|X}$ is stable, being influenced only by the pairwise correlation coefficient $\rho_{12} \in \{0.90,\ 0.68,\ 0.45\}$, since $\theta_{123} = (\theta_1 + \theta_2 + 0)$ =0.60 is fixed for all 18 conditions in the set.

The next set of six cases (7-12) are influenced by X_1 and $\beta_{13} = 0.50(\beta_2 = 0)$. *X1* and *X2* determine $WQS_{index}^{Y|X}$, since $\theta_3 = 0$, and furthermore $\beta_2 = 0$ for regression $Y|WQS_{index}^{Y|X}$, *M* leaving only X_1 as the important factor which determines the regression estimate $\hat{\gamma}_{WQS}^{M|X}$ and subsequently the joint mediated effect $ME_{WQS}^{M|X} = \hat{\theta}_{WQS}^{M|X}\hat{\gamma}_{WQS}^{M|X}$. The pairwise correlation $\rho_{12} \in \{0.9,\ 0.68,\ 0.45\}$ is important in determining $WQS_{index}^{Y|X}$ and $\hat{\theta}_{12}$ given $\theta_3 = 0$ and $\rho_{13} \in \{0.40, 0.30,\ 0.20\}$ is important in determining $\hat{\gamma}_{WQS}^{M|X}$ and $\hat{\beta}_{13}$ given $\beta_2 = 0$, and both estimates are used in $ME_{WQS}^{M|X} = \hat{\theta}_{WQS}^{M|X}\hat{\gamma}_{WQS}^{M|X}$ to calculate $ME_{WQS}^{M|X}$.

The next set of six cases (12-18) are influenced by the predictor X_2 , $\beta_{23} = 0.45(0 + \beta_2 + \beta_3 = 0)$ and ρ_{23} . *X1* and *X2* determine $WQS_{index}^{Y|X}$, given $\theta_3 = 0$, and since $\beta_1 = 0$ for the regression $Y|WQS_{index}^{Y|X}$, *M* that leaves only X_2 as the important factor which determines the regression estimate $\hat{\gamma}_{WQS}^{M|X}$ and subsequently the joint mediated effect $ME_{WQS}^{M|X} = \hat{\theta}_{WQS}^{M|X}\hat{\gamma}_{WQS}^{M|X}$. The pairwise correlation $\rho_{12} \in \{0.9,\ 0.68,\ 0.45\}$ is important for determining $WQS_{index}^{Y|X}$ and $\hat{\theta}_{12}$, since $\theta_3 = 0$ and $\rho_{23} \in \{0.25, 0.19, 0.13\}$ is important for determining $\hat{\gamma}_{WQS}^{M|X}$ and $\hat{\beta}_{23}$, since $\beta_1 = 0$. Consider case 2 with $\hat{\gamma}_{WQS}^{M|X}$ (bias=0.033, true

value=0) and a negatively biased $\hat{\theta}_{WQS}^{M|X}$ (bias= -0.045 with true value=0.60), used to produce the $ME_{WQS}^{M|X} = \hat{\theta}_{WQS}^{M|X}\hat{\gamma}_{WQS}^{M|X}$ estimate $ME_{123}^{M|X}$ =0.019 (given true value=0) with *RMSE* of 0.041, and a coverage probability of 0.92. The $ME_{WQS}^{M|X}$ estimate's value and *RMSE* determine the estimate's coverage probability. In case 11, $\beta_2 = 0$ and X_1 becomes the influencing factor in the regression $Y|WQS_{index}^{Y|X}$, *M* given $\beta_1 = 0.3$ and $\rho_{12} = 0.45$, determining $\hat{\gamma}_{WQS}^{M|X}$ =0.003 with a true value=0, and resulting in an almost unbiased estimate for $ME_{123}^{M|X}$ =0.001 with RMSE (0.033) and a coverage probability of 0.95. Case 13 has a parameter set $(\theta_3 = 0,\ \beta_3 = 0.30,\ \beta_1 = 0)$, $\theta_2 = 0.45, \text{ and } \beta_2 = \left(\frac{\theta_2}{3}\right)$ making X_2 the main variable in regression *Y|X, M*, with a diminished effect on *Y*, because $\beta_2 = \left(\frac{\theta_2}{3}\right)$. The underlying factors acting through the index $WQS_{index}^{M|X}$ that produce the joint mediated effect $ME_{123}^{M|X}$ are the predictors X_1 and X_2, with associated parameters $\theta_{12} = 0.60(\theta_3 = 0)$ in the regression $M\,|WQS_{index}^{M|X}$, and $\beta_{23} = 0.45(\beta_1 = 0)$ in the regression $Y|WQS_{index}^{M|X}$, *M*. Predictor X_3 has a weak association with outcome *Y* in $Y|WQS_{index}^{M|X}$, *M* (even though $\beta_3 = 0.30$), since the index $WQS_{index}^{M|X}$ being used to represent the predictors is influenced only by X_1 and X_2, because $\theta_3 = 0$. Therefore, X_2 is the predictor which determines the joint mediated effect in this case. The joint mediated effect is diminished by the high correlation $\rho_{12} = 0.9$ between X_1 and X_2, $(\theta_1 = 0.15,\ \theta_2 = 0.45)$, decreasing X_2's influence while raising X_1's influence in $WQS_{index}^{M|X}$, which is used in the regression $Y|WQS_{index}^{M|X}$, *M* . Given that $\beta_2 = 0.15 \text{ and } \beta_3 = 0.30$ in the regression *Y|X, M*, X_2*'s* influence is increased to

a lesser extent, aided by *X₃*'s beta parameter value and the lower correlation $\rho_{23} = 0.25$. The gamma estimate $\hat{\gamma}_{WQS}^{M|X}$ =0.005 (true value=0) when multiplied by $\hat{\theta}_{WQS}^{M|X}$ (bias= -0.048 with true value=0.60) produces an almost unbiased joint mediated effect $ME_{123}^{M|X}$ =0.003 (true value=0), with *RMSE* 0.037 and a high coverage probability of 0.94.

When $\gamma = 0$ and $\theta_{13} = 0.55(\theta_2 = 0),$ there were 18 estimates for $ME_{123}^{M|X}$ grouped by $\beta_{123} = \{0.35(\beta_3 = 0),\ 0.50(\beta_2 = 0),\ \text{and}\ 0.45(\beta_1 = 0)\}$ in Table 4.11. The interest is in looking at the beta parameters and the associated correlations between influencing predictors.

Table 4.11

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0$, $\theta_{123} = 0.55(\theta_1 + 0 + \theta_3)$, $\rho_{13} = \{0.40, 0.30, 0.20\}$

Sl.#	$\theta_{123}^{M\|X}$ parm.=0.55	$\gamma_{WQS}^{M\|X}$ parm.= 0	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ parm.=0	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.447	0.015	ρ_{13}= 0.40, ρ_{12}= 0.90	0.007	0.007	0.028	0.932
2	0.441	0.016	ρ_{13}= 0.40, ρ_{12}= 0.90	0.007	0.007	0.017	0.918
3	0.428	0.007	ρ_{13}= 0.30, ρ_{12}= 0.68	0.003	0.003	0.027	0.94
4	0.430	0.010	ρ_{13}= 0.30, ρ_{12}= 0.68	0.004	0.004	0.014	0.942
5	0.410	0.010	ρ_{13}= 0.20, ρ_{12}= 0.45	0.004	0.004	0.027	0.934
6	0.410	0.010	ρ_{13}= 0.20, ρ_{12}= 0.45	0.004	0.004	0.015	0.942
7	0.448	0.027	ρ_{13}= 0.40	0.012	0.012	0.029	0.94
8	0.441	0.024	ρ_{13}= 0.40	0.011	0.011	0.018	0.878
9	0.432	0.029	ρ_{13}= 0.30	0.013	0.013	0.028	0.94
10	0.428	0.027	ρ_{13}= 0.30	0.012	0.012	0.018	0.866
11	0.410	0.023	ρ_{13}= 0.20	0.009	0.009	0.027	0.91
12	0.412	0.024	ρ_{13}= 0.20	0.01	0.01	0.017	0.88
13	0.445	0.018	ρ_{13}= 0.40, ρ_{23}= 0.25	0.008	0.008	0.029	0.936
14	0.443	0.018	ρ_{13}= 0.40, ρ_{23}= 0.25	0.008	0.008	0.016	0.924
15	0.428	0.015	ρ_{13}= 0.30, ρ_{23}= 0.19	0.004	0.004	0.036	0.936
16	0.430	0.015	ρ_{13}= 0.30, ρ_{23}= 0.19	0.003	0.003	0.019	0.944
17	0.416	0.009	ρ_{13}= 0.20, ρ_{23}= 0.13	0.004	0.004	0.033	0.944
18	0.411	0.010	ρ_{13}= 0.20, ρ_{23}= 0.13	0.005	0.005	0.018	0.952

The focus is on X_1, $\beta_{12} = 0.35, (\beta_1 + (\beta_2 = 0) + 0),$ ρ_{12} for cases1-6, and cases 7-12 X_1, X_3, $\beta_{13} = 0.50,$ $(\beta_1 + (\beta_2\theta_2 = 0) + \beta_3),$ ρ_{13} and X_3, $\beta_{23} = 0.45, (0 + (\beta_2\theta_2 = 0) + \beta_3),$ ρ_{23} for

cases 13-18. The value for the regression coefficient $\hat{\theta}_{123}^{M|X}$ is stable, being influenced only by the pairwise correlation coefficient $\rho_{13} \in \{0.40, 0.30, 0.20\}$, since $\theta_{123} = 0.55$ is fixed for all 18 cases under the $\gamma = 0$ and $\theta_{123} = 0.55, (\theta_2 = 0)$ condition. The joint mediated effects $ME_{123}^{M|X}$ is investigated for the 18 cases contained in this set. The difference between the previous parameter values $\rho_{12} \in \{0.90, 0.68, 0.45\}$ and $\theta_{12} = 0.60, (\theta_3 = 0)$ discussed in Tables 4.10, and 4.11 is that $\rho_{13} \in \{0.40, 0.30, 0.20\}$ and $\theta_{13} = 0.55, (\theta_2 = 0)$ will now determine the WQS index $WQS_{index}^{M|X}$. The joint mediated effect $ME_{123}^{M|X}$ estimates are now grouped by the predictors (X_1), (X_1, X_3), and (X_3) that influence the determination of $\hat{\gamma}$ in *Y*|*X*, *M*. The positively biased estimates for $ME_{123}^{M|X}$ with a true value=0, occur for the parameter set (Condition 9) where $(\beta_2, \theta_2) = 0$. The reason for the positively biased $ME_{123}^{M|X}$ estimates is that $\beta_1, \beta_3 \neq 0$, and adding in the influence of the pairwise correlations results in an increased positive bias for the gamma regression coefficient. The higher pairwise correlations in cases 7-8, diminish the influence of X_1 and X_3, while increasing the influence of X_2 (but $\beta_2 = 0$), in the regression *Y*|$WQS_{index}^{M|X}$, *M*, therefore X_1 and X_3 are influential in determining $\hat{\gamma}_{WQS}^{M|X}$ (bias=0.029) ,which in turn gives the positively biased estimate for $ME_{WQS}^{M|X}$ (bias=0.013). In contrast, $(\theta_2 = 0, \beta_3 = 0)$ in case 3 leaves only predictor X_1 ($\beta_1 = 0.15$), to determine the regression coefficient's value $\hat{\gamma}$ (bias=0.007) from *Y*|$WQS_{index}^{M|X}$, *M*. As before, the estimate's *RMSE* (0.027) is increased because of the small sample size of N=300, and when combined with the value of the $ME_{123}^{M|X}$ estimate, determines $ME_{123}^{M|X}$'s high coverage probability of 0.94. In the final set of

null cases, the joint mediated effects $ME_{123}^{M|X}$ when $\gamma = 0$, $\theta_{123} = 0.55, (0 + \theta_2 + \theta_3)$ is investigated for the 18 conditions contained in this set.

Table 4.12

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0$, $\theta_{23} = 0.55(0 + \theta_2 + \theta_3)$ $\rho_{23} = \{0.25, 0.19, 0.13\}$

Sl.#	$\theta_{123}^{M\|X}$ parm.=0.55	$\gamma_{WQS}^{M\|X}$ parm.= 0	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ parm.=0	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.414	0.011	$\rho_{23}=$ 0.25, $\rho_{12}=$ 0.90	0.004	0.004	0.026	0.956
2	0.412	0.012	$\rho_{23}=$ 0.25, $\rho_{12}=$ 0.90	0.005	0.005	0.015	0.932
3	0.404	0.009	$\rho_{23}=$ 0.19, $\rho_{12}=$ 0.68	0.003	0.003	0.025	0.956
4	0.401	0.010	$\rho_{23}=$ 0.19, $\rho_{12}=$ 0.68	0.004	0.004	0.014	0.960
5	0.393	0.009	$\rho_{23}=$ 0.13, $\rho_{12}=$ 0.45	0.003	0.003	0.025	0.928
6	0.386	0.007	$\rho_{23}=$ 0.13, $\rho_{12}=$ 0.45	0.003	0.003	0.013	0.950
7	0.414	0.020	$\rho_{23}=$ 0.25, $\rho_{13}=$ 0.40	0.008	0.008	0.025	0.936
8	0.413	0.018	$\rho_{23}=$ 0.25, $\rho_{13}=$ 0.40	0.008	0.008	0.015	0.904
9	0.408	0.011	$\rho_{23}=$ 0.19, $\rho_{13}=$ 0.30	0.005	0.005	0.024	0.944
10	0.401	0.013	$\rho_{23}=$ 0.19, $\rho_{13}=$ 0.30	0.005	0.005	0.014	0.940
11	0.391	0.005	$\rho_{23}=$ 0.13, $\rho_{13}=$ 0.20	0.002	0.002	0.023	0.954
12	0.387	0.009	$\rho_{23}=$ 0.13, $\rho_{13}=$ 0.20	0.003	0.003	0.013	0.946
13	0.415	0.025	$\rho_{23}=$ 0.25	0.010	0.010	0.026	0.932
14	0.410	0.022	$\rho_{23}=$ 0.25	0.009	0.009	0.016	0.898
15	0.404	0.021	$\rho_{23}=$ 0.19	0.009	0.009	0.027	0.926
16	0.401	0.019	$\rho_{23}=$ 0.19	0.008	0.008	0.015	0.886
17	0.390	0.023	$\rho_{23}=$ 0.13	0.009	0.009	0.026	0.926
18	0.391	0.020	$\rho_{23}=$ 0.13	0.008	0.008	0.015	0.916

The correlations between X_2 and X_3 are all smaller values $\rho_{23} = \{0.25, 0.19, 0.13\}$, than the first set in Figure 4.13, where correlations between X_1, X_2 come into play when determining the WQS index $WQS_{index}^{Y|X}$. In this set the estimate for $ME_{123}^{M|X}$ occurs for reasons very similar to the conditions discussed above in Table 4.10. The parameter set for cases 13-18 in Figure 4.13 have $(\beta_1, \theta_1) = 0$. Predictors X_2 and X_3 with parameter values $\beta_2 = 0.15$, $\beta_3 = 0.30$, and $\rho_{23} = \{0.25, 0.19, 0.13\}$ determine the value of the

regression coefficient $\hat{\gamma}_{WQS}^{M|X}$, which when combined with $\hat{\theta}_{WQS}^{M|X}$ from *M*|$WQS_{index}^{M|X}$ to determine the value of $ME_{123}^{M|X}$ (bias=0.010 for a true value=0), with an estimate's *RMSE* of 0.026, and a coverage probability of 0.93. In contrast, the low estimate for $ME_{123}^{M|X}$ occurs for case 11 in the second set of cases 7-12, with the parameter set $(\beta_2, \theta_1) = 0$, making predictor *X₃* $(\beta_3 = 0.2)$ determine the value of the regression coefficient $\hat{\gamma}_{WQS}^{M|X}$. Since correlation $\rho_{23} = 0.13$ is small, the influence of predictor X_3 $(\beta_3 = 0.2)$ remains small in determining $\hat{\gamma}$ (bias=0.005, with true value=0), and consequently the $ME_{123}^{M|X}$ estimate (bias=0.002), is almost unbiased estimate. From the null conditions discussed in Tables 4.16 to 4.18, the influence of predictor correlations on $ME_{123}^{M|X}$ as plotted in Figure.4.13, B is evident. Higher pairwise correlations (red diamond) increase the joint mediated estimate's value.

Simulated conditions to calculate $ME_{123}^{M|X}$ estimates based on gamma parameters γ=$\{0, 0.25, 0.35\}$ are organized by joint beta parameters $\beta_{123} = 0.35, (\beta_3 = 0)$, $\beta_{123} = 0.50, (\beta_2 = 0)$ and $\beta_{123} = 0.45, (\beta_1 = 0)$ each with 54 different conditions. The three joint beta parameter sets are each organized into three sub-groups of 18 $ME_{123}^{M|X}$ estimates based on the joint theta values $\theta_{123} = \{0.60(\theta_3 = 0),\ 0.55(\theta_2 = 0) \text{ and } 0.55(\theta_1 = 0)\}$, and each sub-group of 18 has three clusters of six conditions based on the gamma parameter γ=$\{0, 0.25, 0.35\}$ as shown in Figure 4.13A & B. The 54 null conditions based on $\gamma = 0$ were discussed using tables 4.16 to 4.18 and the remaining 108 $ME_{123}^{M|X}$ estimates for non-zero gamma parameters, 54 conditions for $\gamma = 0.25$ and 54 conditions

for $\gamma = 0.35$ are discussed below. Figure 4.13B shows that a higher predictor correlation will produce a higher joint mediated effect estimate $ME_{123}^{M|X}$ with a positive (if >0) or smaller negative bias (if <0). This is evident from comparing cases 1 and 2 to cases 5 and 6 in the following table, since both pairs of cases have the same beta and theta parameter sets but different pairwise predictor correlation sets.

Table 4.13

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0.25$, $\theta_{123} = 0.6, (\theta_1 + \theta_2 + 0)$

Sl.#	$\theta_{123}^{M\|X}$ =0.60	$\gamma_{WQS}^{M\|X}$ =0.25	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ =0.15	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.552	0.284	$\rho_{12}=$ 0.90	0.157	0.007	0.035	0.954
2	0.555	0.284	$\rho_{12}=$ 0.90	0.158	0.008	0.021	0.930
3	0.531	0.272	$\rho_{12}=$ 0.68	0.145	-0.005	0.037	0.912
4	0.531	0.276	$\rho_{12}=$ 0.68	0.147	-0.003	0.020	0.938
5	0.504	0.274	$\rho_{12}=$ 0.45	0.138	-0.012	0.035	0.928
6	0.502	0.273	$\rho_{12}=$ 0.45	0.137	-0.013	0.023	0.886
7	0.554	0.271	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.150	0	0.036	0.958
8	0.553	0.272	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.150	0	0.020	0.934
9	0.534	0.260	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.139	-0.011	0.037	0.922
10	0.530	0.267	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.141	-0.009	0.021	0.918
11	0.501	0.259	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.130	-0.020	0.037	0.912
12	0.503	0.259	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.130	-0.020	0.026	0.798
13	0.557	0.258	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.144	-0.006	0.039	0.928
14	0.554	0.263	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.146	-0.004	0.021	0.940
15	0.533	0.254	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.135	-0.015	0.038	0.912
16	0.531	0.263	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.140	-0.010	0.022	0.906
17	0.506	0.260	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.131	-0.019	0.038	0.902
18	0.505	0.260	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.131	-0.019	0.026	0.830

In cases 1 and 2, the theta parameters are $(\theta_1 = 0.15, \theta_2 = 0.45, \theta_3 = 0)$ with pairwise predictor correlations $\rho_{12} = 0.90, \rho_{13} = 0.40, \rho_{23} = 0.25$. Pairing the theta parameters with their corresponding correlations shows $\theta_1 = 0.15 \text{ and } \theta_2 = 0.45$ for

X_1 and X_2 as $\rho_{12} = 0.90$, $\theta_1 = 0.15$ and $\theta_3 = 0$ for X_1, X_3 as $\rho_{13} = 0.40$, and $\theta_2 = 0.45$ and $\theta_3 = 0$ for X_2 and X_3 as $\rho_{23} = 0.25$. In cases 5 and 6, the pairwise correlations decrease to mush smaller values $\rho_{12} = 0.45, \rho_{13} = 0.20, \rho_{23} = 0.13$. Pairing the theta parameters with their corresponding correlations shows $(\theta_1 = 0.15, \theta_2 = 0.45)$ for X_1 and X_2 as $\rho_{12} = 0.45$, $\theta_1 = 0.15$ and $\theta_3 = 0$ for X_1 and X_3 as $\rho_{13} = 0.19$, and $\theta_2 = 0.45$ and $\theta_3 = 0$ for X_2 and X_3 as $\rho_{23} = 0.13$. Correlations $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.90, 0.68, 0.45)$ produce a WQS joint theta estimate $\hat{\theta}_{WQS}^{M|X}$ with a smaller negative bias as compared to $\left(\rho_{12}, \rho_{13}, \rho_{23}\right) = (0.45, 0.20, 0.13)$, resulting in $\hat{\theta}_{WQS}^{M|X}$ =0.555 vs 0.502 (true value for $\hat{\theta}_{WQS}^{M|X}$ =0.60), because of the higher correlations between the predictors in case 2 vs. 6. The corresponding joint mediated effect estimates shown in Table 4.19 associated with $\hat{\theta}_{WQS}^{M|X}$ using $WQS_{index}^{M|X}$ for this example are 0.158 vs. 0.137, with the true value for $ME_{123}^{M|X}$ =0.15. The 18 conditions listed in Figure 4.13 produce three sub-clusters of six $ME_{123}^{M|X}$ estimates seen in Figure 4.13 A and B for $\gamma = 0.25$: $(\theta_3 = 0, \beta_3 = 0)$, $(\theta_3 = 0, \beta_2 = 0)$ and $(\theta_3 = 0, \beta_1 = 0)$, $ME_{123}^{M|X}$ having a true value of 0.15, arranged in descending order of pairwise correlations and subsequently the estimate values ranging from 0.158 to 0.130. Other tables (not shown) associated with $\gamma = 0.25$, $\theta_{123} = 0.55, (\theta_1 + 0 + \theta_3)$ produce three additional sub-clusters of six $ME_{123}^{M|X}$ estimates each, as seen in Figure 4.13 A and B: $(\theta_2 = 0, \beta_2 = 0)$, $(\theta_2 = 0, \beta_1 = 0)$ and $(\theta_2 = 0, \beta_3 = 0)$ with $ME_{123}^{M|X}$ having a true value of 0.138, arranged in descending order pairwise correlations and their associated estimate values ranging from 0.125 to 0.105; while other tables

(not shown) associated with $\gamma = 0.25$, $\theta_{123} = 0.55, (0 + \theta_2 + \theta_3)$ produce three additional sub-clusters of six $ME_{123}^{M|X}$ estimates each, as seen in Figure 4.13 A and B: $(\theta_1 = 0, \beta_1 = 0)$, $(\theta_1 = 0, \beta_2 = 0)$, and $(\theta_1 = 0, \beta_3 = 0)$ with $ME_{123}^{M|X}$ having a true value of 0.138, arranged in descending order pairwise correlations and their associated estimate values ranging from 0.114 to 0.100.

The value for the regression coefficient $\theta_{123}^{M|X}$ is not associated with the gamma parameter coefficient, since $\hat{\theta}_{WQS}^{M|X}$'s values only determine $WQS_{index}^{Y|X}$ and $\{\rho_{12}, \rho_{13}, \rho_{23}\}$, given $\theta_{123} = (\theta_1 + \theta_2 + 0)$ =0.60 is fixed for all 18 conditions in Table 4.19. The interest is in looking at the remaining parameters $\beta_{123} = 0.35(\beta_1 + \beta_2 + 0), \rho_{12}$ for the first six cases, $\beta_{123} = 0.50(\beta_1 + 0 + \beta_3\theta_3 = 0), \rho_{13}$ for the second six cases and for the final six cases $\beta_{123} = 0.45(0 + \beta_2 + \beta_3\theta_3 = 0), \rho_{23}$ based on γ=0.25. From the previous analysis we know that the highest positively biased $ME_{123}^{M|X}$ estimate will be in the first two cases of the first set, because of β_1 & β_2 are both influential in determining the value of the gamma regression coefficient $\hat{\gamma}$, and $\rho_{12} = 0.90$ (highest correlation in the set) gives the gamma estimate a bias=0.034 (with true value=0.25). The regression coefficient $\hat{\theta}_{WQS}^{M|X}$ has a low negative bias= -0.045, resulting in an $ME_{123}^{M|X}$ estimate of 0.158 (with true value= 0.15), and its estimate's *RMSE* of 0.021, with a coverage of 0.93. The most negative biased $ME_{123}^{M|X}$ estimate comes from $\beta_{13} = 0.50(\beta_1 + 0 + \beta_3\theta_3 = 0), \rho_{13} = 0.20$ for the middle six cases, with only X_1 left to determine the gamma coefficient estimate $\hat{\gamma}$, given $\rho_{12} = 0.45$ and $\rho_{13} = 0.20$. The correlation $\rho_{12} = 0.45$ acts on X_1, X_2, $(\beta_1, \beta_2) = (0.3, 0)$ giving

$\hat{\beta}_1$ a negative bias while increasing $\hat{\beta}_2$, and $\rho_{13} = 0.20$ acts on *X1, X3,* $(\beta_{1,}\beta_3) = (0.3, 0.2)$ giving $\hat{\beta}_1$ an additional negative bias, while increasing $\hat{\beta}_3$. Since $\hat{\beta}_1$ is the most important coefficient determining $\hat{\gamma}$, the positive bias on $\hat{\gamma}$ is diminished, resulting on reduced $ME_{123}^{M|X}$ estimate of 0.13 (with true value=0.15), a *RMSE* of 0.026, and coverage of 0.80.

Table 4.14

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0.35$, $\theta_{123} = 0.6, (\theta_1 + \theta_2 + 0)$

Sl. #	$\theta_{123}^{M\|X}$ =0.60	$\gamma_{WQS}^{M\|X}$ =0.35	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ =0.21	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.557	0.381	$\rho_{12}=$ 0.90	0.212	0.002	0.034	0.964
2	0.554	0.385	$\rho_{12}=$ 0.90	0.213	0.003	0.020	0.948
3	0.539	0.377	$\rho_{12}=$ 0.68	0.203	-0.007	0.038	0.932
4	0.531	0.377	$\rho_{12}=$ 0.68	0.200	-0.010	0.022	0.906
5	0.504	0.371	$\rho_{12}=$ 0.45	0.187	-0.023	0.042	0.894
6	0.501	0.372	$\rho_{12}=$ 0.45	0.187	-0.023	0.030	0.746
7	0.557	0.376	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.209	-0.001	0.035	0.948
8	0.553	0.375	$\rho_{12}=$ 0.90, $\rho_{13}=$ 0.40	0.208	-0.002	0.020	0.948
9	0.534	0.365	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.195	-0.015	0.040	0.906
10	0.532	0.364	$\rho_{12}=$ 0.68, $\rho_{13}=$ 0.30	0.194	-0.016	0.024	0.874
11	0.509	0.355	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.181	-0.029	0.044	0.866
12	0.503	0.361	$\rho_{12}=$ 0.45, $\rho_{13}=$ 0.20	0.182	-0.028	0.034	0.704
13	0.557	0.359	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.200	-0.010	0.040	0.930
14	0.553	0.365	$\rho_{12}=$ 0.90, $\rho_{23}=$ 0.25	0.202	-0.008	0.022	0.924
15	0.532	0.359	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.191	-0.019	0.041	0.898
16	0.531	0.363	$\rho_{12}=$ 0.68, $\rho_{23}=$ 0.19	0.193	-0.017	0.026	0.854
17	0.502	0.353	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.177	-0.033	0.047	0.836
18	0.504	0.359	$\rho_{12}=$ 0.45, $\rho_{23}=$ 0.13	0.181	-0.029	0.035	0.658

The interest is in looking at the parameters $\beta_{123} = 0.35, (\beta_1 + \beta_2 + 0), \rho_{12}$ for the first six cases, $\beta_{123} = 0.50, (\beta_1 + 0 + \beta_3\theta_3 = 0), \rho_{13}$ for the second six cases and for the final six with the parameters $\beta_{123} = 0.45, (0 + \beta_2 + \beta_3\theta_3 = 0), \rho_{23}$ based on γ=0.35. From the previous

analysis we know that the highest positively biased $ME_{123}^{M|X}$ estimate will be in the first two cases of the first set, because of β_1 and β_2 are both influential in determining the value of the gamma regression coefficient $\hat{\gamma}$, and $\rho_{12} = 0.90$ (highest correlation in the set), giving a gamma estimate a bias=0.035 (with true value=0.35). The joint theta coefficient $\hat{\theta}_{WQS}^{M|X}$ has a bias = -0.046, resulting in an $ME_{123}^{M|X}$ estimate of 0.213 (with true value=0.21), and its estimate's *RMSE* of 0.020 with a coverage of 0.95. The most negative biased $ME_{123}^{M|X}$ estimate comes from $\beta_{23} = 0.45, (0 + \beta_2 + \beta_3\theta_3 = 0), \rho_{23} = 0.13$ for the last six cases, with only X_2 left to determine the gamma coefficient estimate $\hat{\gamma}$ since $\theta_3 = 0$ when $WQS_{index}^{M|X}$ was determined, given $\rho_{12} = 0.45$ and $\rho_{23} = 0.13$. The correlation $\rho_{12} = 0.45$ acts on X_1, X_2, $(\beta_1, \beta_2) = (0, 0.15)$ giving $\hat{\beta}_1$ a positive bias while decreasing $\hat{\beta}_2$, and $\rho_{23} = 0.13$ acts on X_2, X_3, $(\beta_2, \beta_3) = (0.15, 0.3)$ without much increase to $\hat{\beta}_2$ as the correlation is small $\rho_{23} = 0.13$. Since $\hat{\beta}_2$ is the important coefficient determining $\hat{\gamma}$, the positive bias on $\hat{\gamma}$ is diminished at 0.009, resulting on reduced $ME_{123}^{M|X}$ estimate of 0.18 (with true value=0.21), a *RMSE* of 0.026, and reduced coverage of 0.80.

The interest is in looking at the remaining parameters for the first six cases, $\beta_{123} = 0.50(\beta_1 + 0 + \beta_3\theta_3 = 0), \rho_{13}$ for the second six cases and the final six, the parameters are $\beta_{123} = 0.45(0 + \beta_2 + \beta_3\theta_3 = 0), \rho_{23}$. From the previous analysis we know that the highest positively biased $ME_{123}^{M|X}$ estimate will be in the first two cases of the first set, because of β_1 & β_2 are both influential in determining the value of the gamma regression coefficient $\hat{\gamma}$, and $\rho_{12} = 0.90$ giving the gamma estimate a bias=0.035 (with true value=0.35) and

$\theta_{123}^{M|X}$ a bias= -0.061, resulting in an $ME_{123}^{M|X}$ estimate of 0.213 (with true value=0.21) and a large estimate's *RMSE* of 0.035 with low coverage probability of 0.66.

Table 4.15

Joint Mediated Effect $ME_{123}^{M|X}$ When $\gamma = 0.35$, $\theta_{123} = 0.55, (0 + \theta_2 + \theta_3)$

Sl. #	$\theta_{123}^{M\|X}$ =0.55	$\gamma_{WQS}^{M\|X}$ =0.35	$\rho_{12}/\rho_{13}/\rho_{23}$	$ME_{WQS}^{M\|X}$ =0.19	$ME_{WQS}^{M\|X,Bias}$	$ME_{WQS}^{M\|X,RMSE}$	$ME_{WQS}^{M\|X,Cov.}$
1	0.419	0.362	$\rho_{23}=$ 0.25, $\rho_{12}=$ 0.90	0.152	-0.041	0.051	0.720
2	0.414	0.363	$\rho_{23}=$ 0.25, $\rho_{12}=$ 0.90	0.150	-0.042	0.045	0.316
3	0.405	0.359	$\rho_{23}=$ 0.19, $\rho_{12}=$ 0.68	0.146	-0.047	0.055	0.672
4	0.401	0.359	$\rho_{23}=$ 0.19, $\rho_{12}=$ 0.68	0.144	-0.048	0.051	0.190
5	0.394	0.361	$\rho_{23}=$ 0.13, $\rho_{12}=$ 0.45	0.143	-0.050	0.058	0.594
6	0.390	0.357	$\rho_{23}=$ 0.13, $\rho_{12}=$ 0.45	0.139	-0.053	0.056	0.122
7	0.414	0.364	$\rho_{23}=$ 0.25, $\rho_{13}=$ 0.40	0.151	-0.042	0.051	0.686
8	0.413	0.370	$\rho_{23}=$ 0.25, $\rho_{13}=$ 0.40	0.153	-0.040	0.043	0.312
9	0.406	0.361	$\rho_{23}=$ 0.19, $\rho_{13}=$ 0.30	0.147	-0.046	0.054	0.642
10	0.401	0.365	$\rho_{23}=$ 0.19, $\rho_{13}=$ 0.30	0.146	-0.046	0.049	0.218
11	0.395	0.355	$\rho_{23}=$ 0.13, $\rho_{13}=$ 0.20	0.140	-0.052	0.059	0.546
12	0.388	0.361	$\rho_{23}=$ 0.13, $\rho_{13}=$ 0.20	0.140	-0.052	0.055	0.118
13	0.418	0.369	$\rho_{23}=$ 0.25	0.154	-0.038	0.048	0.752
14	0.412	0.371	$\rho_{23}=$ 0.25	0.153	-0.040	0.043	0.342
15	0.403	0.368	$\rho_{23}=$ 0.19	0.148	-0.044	0.053	0.670
16	0.402	0.373	$\rho_{23}=$ 0.19	0.150	-0.043	0.045	0.270
17	0.393	0.375	$\rho_{23}=$ 0.13	0.148	-0.045	0.054	0.680
18	0.390	0.371	$\rho_{23}=$ 0.13	0.144	-0.048	0.050	0.164

The influence of low pairwise correlations $(\rho_{12} = 0.45, \rho_{13} = 0.20, \rho_{23} = 0.13)$ associated with low individual theta parameters values $(\theta_1 = 0,\ \theta_2 = 0.25,\ \theta_3 = 0.3)$ and low beta parameter values $(\beta_1 = 0.15, \beta_2 = 0.2, \beta_3 = 0)$, results in negatively biased $\hat{\theta}_{WQS}^{M|X}$ estimate=0.39 (bias= -0.16, with true value $\hat{\theta}_{WQS}^{M|X}$ =0.55) and a positively biased $\hat{\gamma}_{WQS}^{M|X}$ (bias=0.007), and a diminished $ME_{123}^{M|X}$ estimate=0.139 (bias= -0.053, with true value

$ME_{123}^{M|X}$ =0.193) having a high *RMSE* of 0.056 and low coverage probability of 0.122. The same pattern is seen in cases 11 and 12, 17 and 18, for $WQS_{index}^{M|X}$ with low theta parameter values $\theta_{123} = 0.55, (0 + \theta_2 = 0.25 + \theta_3 = 0.30)$, the low beta parameter values $\beta_{123} = 0.2,\ (\theta_1\beta_1 = 0 + \beta_2 = 0.2 + \beta_3 = 0),\ \ \rho_{12} = 0.45$ for cases 5 and 6, $\beta_{123} = 0.2,\ (\theta_1\beta_1 = 0 + \beta_2 = 0 + \beta_3 = 0.2), \rho_{13} = 0.20$ for cases 11 and12, and $\beta_{123} = 0.45,\ (\theta_1\beta_1 = 0 + \beta_2 = 0.15 + \beta_3 = 0.3),\ \rho_{23} = 0.13$, for cases 17 and18, the pairwise correlations being $(\rho_{12} = 0.45, \rho_{13} = 0.20, \rho_{23} = 0.13)$. In the last example, $\beta_{123} = 0.45$ is diminished because $\hat{\beta}_2$ is negatively biased in $(\theta_1\beta_1 = 0 + \beta_2 = 0.15)$ and $\rho_{12} = 0.45$, and $\hat{\beta}_3$ is negatively biased in $(\theta_1\beta_1 = 0 + \beta_3 = 0.3)$ and $\rho_{13} = 0.20$. The combined negative bias for $\hat{\beta}_2, \hat{\beta}_3$ reduces the joint beta estimate and $\hat{\gamma}_{WQS}^{M|X}$ has a positive bias of 0.021 but $\hat{\theta}_{WQS}^{M|X}$ =0.39, with a negative bias of -0.16 (true value=0.55), resulting in $ME_{WQS}^{M|X}$ =0.144 with true value=0.193, a large RMSE of 0.50, and a low coverage of 0.164.

The next plots show the joint mediated estimate $ME_{WQS}^{M|X}$'s *Bias* and *RMSE*.

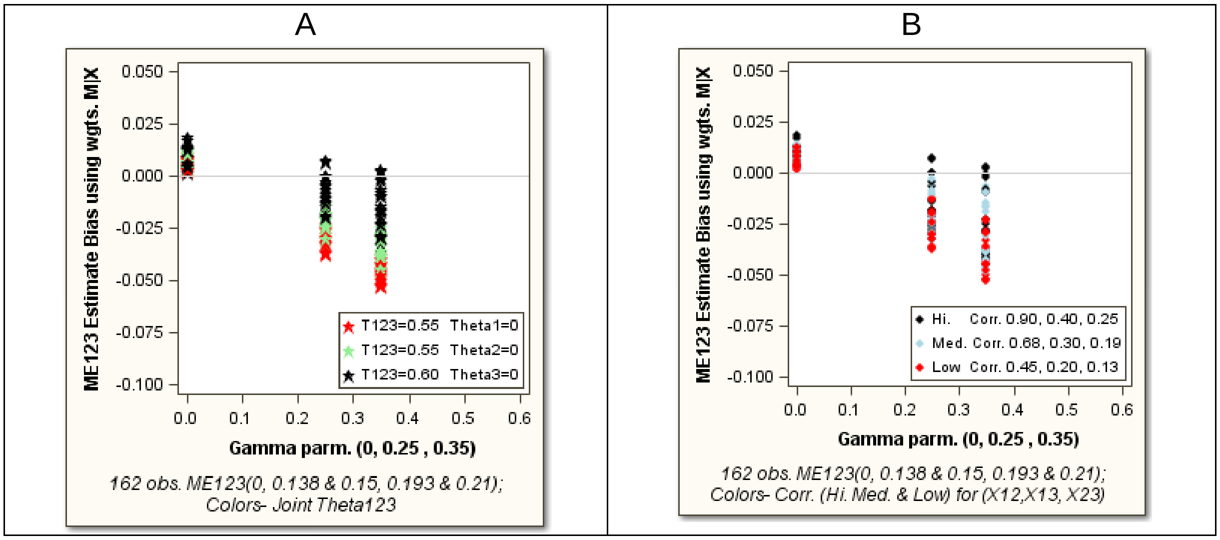


Figure 4.14 A-B Joint Mediated Estimate $ME_{WQS}^{M|X}$'s *Bias* w/ WQS Weights $\mathrm{WQS}_{index}^{M|X}$

Figures 4.14 A shows that when $\theta_3 = 0$ (black star) the positive biased or less negatively biased $ME_{WQS}^{M|X}$ estimates are associated with the set of high pairwise correlation values between X_1, X_2, X_3 of $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.90, 0.40, 0.25)$ shown in Figure 4.14 B. The more negatively biased $ME_{WQS}^{M|X}$ estimates are associated with $\theta_1 = 0$ and the set of low pairwise correlation values $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.45, 0.20, 0.13)$. This is because when $\theta_3 = 0$ the influencing predictors determining the index $WQS_{index}^{Y|X}$ are X_1 and X_2, and all the correlation values between X_1, X_2, X_3 have the highest values for X_1 and X_2 namely, $\rho_{12} = (0.90,\ 0.68,\ 0.45)$ as compared to $\rho_{23} = (0.25, 0.19, 0.13)$, while the correlation between X_1, X_3 can take on values $\rho_{13} = (0.40, 0.30, 0.20)$. The predictor *X₁* is associated with ρ_{12}, ρ_{13}, whereas predictor X_3 is associated with ρ_{13}, ρ_{23} which are lower valued correlations. Simulation design parameter values associated with X_1, i.e. $(\theta_1, \beta_1) \in (0,\ \text{low, high})$ values and their associated correlations $(\rho_{12},\ \rho_{13})$, influence the mediated effects and their statistics, since they are directly dependent on the pairwise correlations for their results. In contrast, when $\theta_1 = 0$ (red star) and $\beta_2 = 0$, predictor X_3 is the only influencing variable in the regression $Y \mid WQS_{index}, M$, which estimates $\hat{\gamma}_{WQS}$. The parameter values associated with X_3, i.e. $(\theta_3, \beta_3) \in (0,\ \text{low, high})$ values and their associated correlations (ρ_{13}, ρ_{23}) make X_3 the least important amongst $X_1, X_2,$ and X_3 in determining the mediated effects and their statistics. This explains the patterns of high and low negative bias shown in Figure 4.14 A – B. The estimate's RMSE and coverage probability are discussed in Figure 4.15

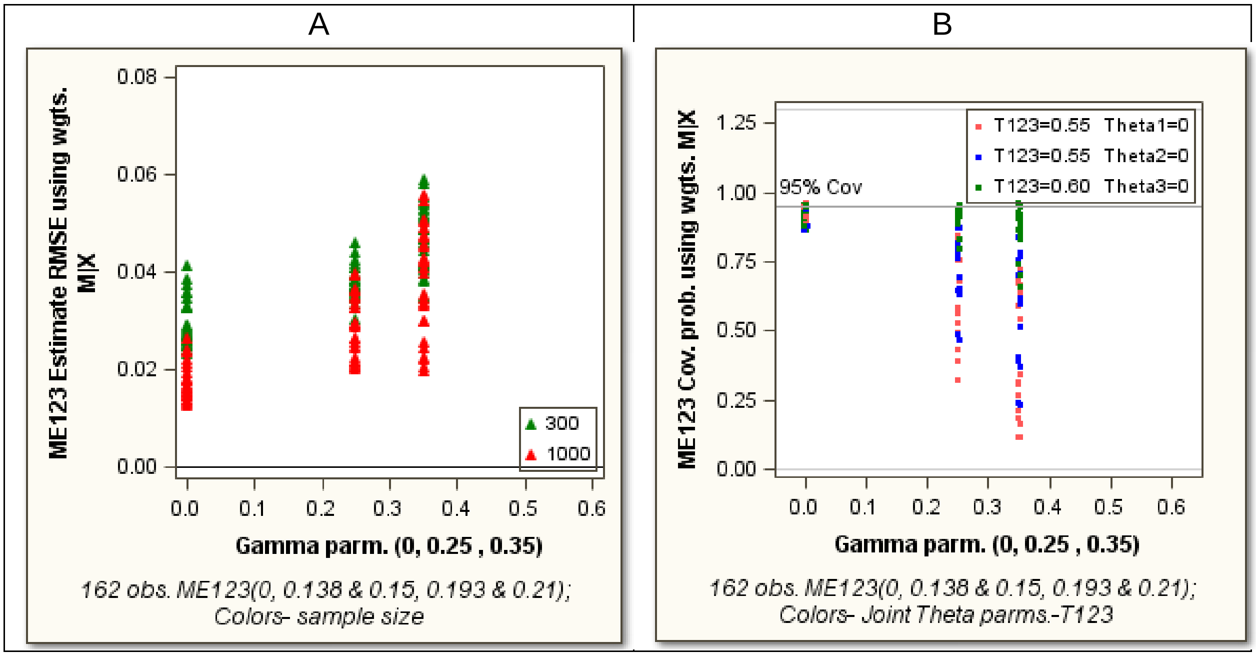


Figure 4.15 A-B Joint Mediated Estimate $ME_{WQS}^{M|X}$'s *RMSE* and Coverage probability

Figures 4.15 A shows that $ME_{WQS}^{M|X}$'s *RMSE* is inversely proportional to the sample size and its magnitude and variability depend on the value of the estimate. Figure 4.15 B shows the coverage probability for the joint mediated effect displayed by the estimate's two influencing parameters the γ (x-axis) and the color coded grouping by $\theta_{123}\left(\theta_3 = 0(green), \theta_2 = 0(blue), \theta_1 = 0(red)\right)$. Earlier analysis has shown that the order of magnitude of influence for the predictors based on their parameter values and the relevant correlations was in the order of X_1, X_2 and X_3. It is then conceivable that the highest coverage probabilities $\theta_3 = 0(green)$ will align with the lesser biased $ME_{WQS}^{M|X}$ estimates where $(X_1, X_2) \neq 0$, with larger *RMSE* values due to the small sample size. The null conditions for the joint mediated effects have adequate (close to 95%) coverage probability. When $\theta_3 = 0$ (green) the influential predictors X_1 and X_2 are associated with high pairwise correlations, whereas when $\theta_1 = 0$ (red) the influential

predictors X_2 and X_3 are associated with weaker correlations, $\rho_{23}=(0.25, 0.19, 0.13)$ and $\rho_{31}=(0.40, 0.30, 0.20)$, resulting in diminished, more negatively biased joint mediated effects with low coverage.

The joint mediated effect's type1 errors and the individual influential predictors are discussed next.

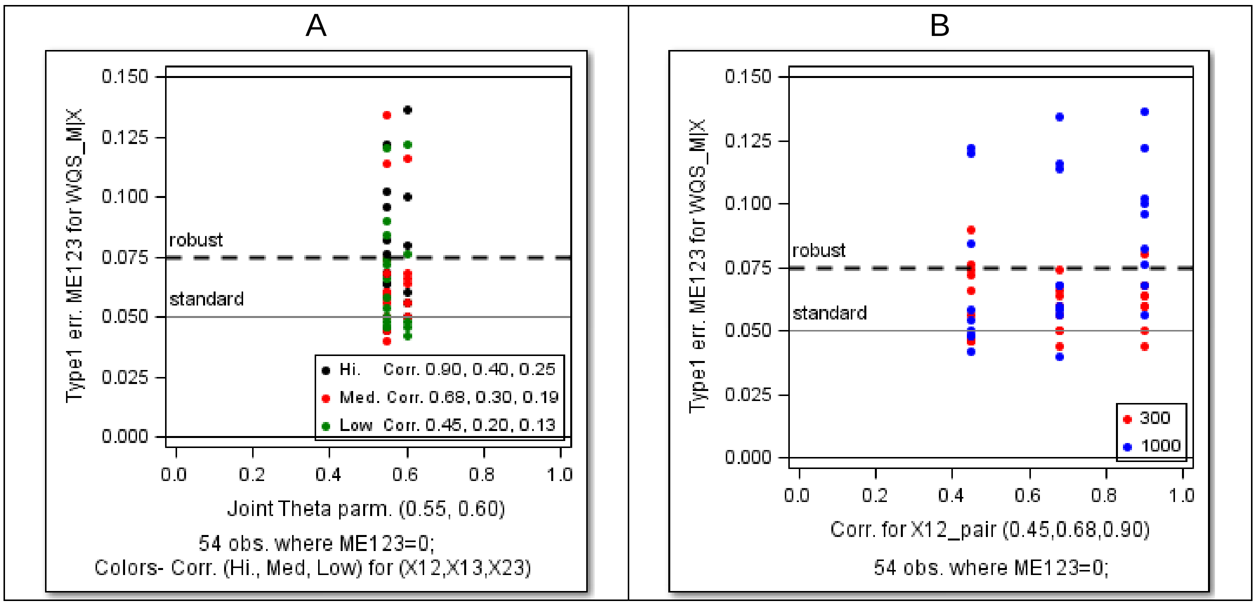


Figure 4.16 A-B Joint Mediated Estimate's Type1 Error $ME_{WQS}^{\substack{Type1\\M|X}}$ by θ_{123}, Correlations and *N*

The plots in Figure 4.16 A-B show how the joint mediated effect $ME_{WQS}^{M|X}$ performed under Type1 error rate. Panel A shows the sixteen of the possible 54 conditions with an exception for exceeding the *a priori* set robust limit for Type1 error of 0.075, having a false significant mediated effect when $ME_{123}=0$.

There were 16 of the possible 54 conditions of the joint mediated effect $ME_{123}^{M|X}$ (ME_{123}=0) that had a type1 error rate >0.075. These exceptions are listed in Table 4.22 and are discussed here. Cases 1 and 2 may be compared between themselves and with Cases 5 and 6 to see the effect of an increased sample size (*N*=300 vs. 1000) and

a decreased pairwise correlation $\left(\rho_{12}=0.90 \text{ vs. } 0.45), \rho_{13}\right)$, all other parameters remaining the same. The larger sample size (*N*=1000) for case 2, results in a smaller *RMSE* (0.026 vs. 0.041), which puts a narrower confidence interval around the average estimate $ME_{123}^{M|X}$ =0.017, causing higher type1 errors (0.136 vs. 0.080). The effect of a larger sample size is always an increased power and an increased type1 error rate.

Cases 3 and 5 may be compared since all their parameters are similar except for the correlation $\rho_{12}=0.68$ for case 3 vs. $\rho_{12}=0.45$ for case 5. From the previous analysis it can be learned that a higher correlation between X_1 and X_2 will result in a higher joint mediated effect, all other parameters remaining the same. This is true here as case 3 with $\rho_{12}=0.68$ has a higher $ME_{123}^{M|X}$ (0.014) as compared to case 5 with $\rho_{12}=0.45$, having a lower estimate value $ME_{123}^{M|X}$ (0.012). The *RMSE* for the higher estimate is also larger (0.024 vs. 0.022), because of the larger estimate for $ME_{123}^{M|X}$, both cases having the same sample size (*N*=1000). The larger *RMSE* (0.024) provides a wider confidence interval around $ME_{123}^{M|X}$ (0.014), lowering the type1 error to 0.116 vs. 0.112, for case 5.

The higher type1 error for case 6 (0.100) vs.7 (0.082) is based on X_1 being influential for both in determining the joint mediated estimate $ME_{123}^{M|X}$, but the β_1 parameter for case 6 is twice as large (0.30 vs. 0.15), producing a larger estimate for $ME_{123}^{M|X}$ (0.024 vs. 0.016) and lower coverage probability (0.90 vs. 0.92), which is reflected in the higher type1 error for case 6 (0.100) vs. 7(0.082). The higher correlation acting on smaller association strengths, i.e. $\rho_{12}=0.90$ for case 7, does not increase its type1 error over the lower pairwise correlation acting on the higher association strengths in the

regression $Y \mid WQS_{index}, M$, i.e. $\rho_{13} = 0.40$ for case 6. Cases 8-11 behave predictably similar to the cases discussed above. The higher $ME_{123}^{M|X}$ estimates for null conditions 8, 9, and 11 ($ME_{123} = 0$) have progressively higher type1 error rates (0.120, 0.122, and 0.134), with all other parameters remaining the same. Case 10 has a smaller sample size resulting in a higher *RMSE* and lower type 1 error (0.090). Cases 12 and 13 are similar to cases 6 and 7 discussed earlier. The higher estimate goes with the higher regression coefficients, but the magnitude of the correlations having much lesser influence on the joint mediated effect estimate. Cases 6 and 13 have the larger beta parameters $(\beta_1, \beta_3) = (0.30, 0.20)$, as compared to case 7 $(\beta_1, \beta_2) = (0.15, 0.20)$ and case 12 $(\beta_2, \beta_3) = (0.15, 0.30)$. Since the sample size is the same, the *RMSE* of the estimate is not changing and therefore cases 6 and 13 have type1 errors (0.100, 0.096) compared to cases 7 and 12 with type1 errors (0.082, 0.076). Cases 14-16 have X_2 and X_3 as the influencing variables; however, the earlier discussion had shown that X_1 had the most influence in the regression equations, because X_1's correlations with X_2 and X_3 are $\rho_{12} \in \{0.90, 0.68, 0.45\}$ and $\rho_{13} \in \{0.40, 0.30, 0.20\}$. In these three cases having $(\beta_1, \theta_1) = 0$, one would expect a lower positively biased estimate for $ME_{123}^{M|X}$. However, X_2 and X_3 have similar regression coefficients in both $M \mid WQS_{index}$ and $Y \mid WQS_{index}, M$. This synergy increases the estimate's value. Hence, we have the three exceptions for these null cases $(14,\ 15,\ 16)$ with type1 errors $(0.102,\ 0.114,\ 0.84)$ with estimate bias of $(\ 0.016, 0.016, \text{and } 0.014)$.

Table 4.16

Joint Mediated Effect $ME_{123}^{M|X}$ Type1 Error >0.075 Exceptions, 16 of 54 conditions

Sl #	N	β_1	β_2	β_3	Correlations	θ_1	θ_2	θ_3	Type1 err.	T_{WQS} Bias	G_{WQS} Bias	ME_{123} Bias	ME_{123} RMSE
1	300	0.15	0.20	0	ρ_{12} =0.90	0.15	0.45	0	0.080	-0.046	0.033	0.019	0.041
2	1000	0.15	0.20	0	ρ_{12} =0.90	0.15	0.45	0	0.136	-0.045	0.031	0.017	0.026
3	1000	0.15	0.20	0	ρ_{12} =0.68	0.15	0.45	0	0.116	-0.068	0.027	0.014	0.024
4	300	0.15	0.20	0	ρ_{12} =0.45	0.15	0.45	0	0.076	-0.097	0.024	0.012	0.034
5	1000	0.15	0.20	0	ρ_{12} =0.45	0.15	0.45	0	0.122	-0.097	0.025	0.012	0.022
6	1000	0.30	0	0.20	ρ_{12} =0.90, ρ_{13} =0.40	0.15	0.45	0	0.100	-0.046	0.024	0.013	0.024
7	1000	0.15	0.20	0	ρ_{13} =0.40, ρ_{12} =0.90	0.35	0	0.20	0.082	-0.109	0.016	0.007	0.017
8	1000	0.30	0	0.20	ρ_{13} =0.40	0.35	0	0.20	0.122	-0.109	0.024	0.011	0.018
9	1000	0.30	0	0.20	ρ_{13} =0.30	0.35	0	0.20	0.134	-0.122	0.027	0.012	0.018
10	300	0.30	0	0.20	ρ_{13} =0.20	0.35	0	0.20	0.090	-0.140	0.023	0.009	0.027
11	1000	0.30	0	0.20	ρ_{13} =0.20	0.35	0	0.20	0.120	-0.138	0.024	0.010	0.017
12	1000	0	0.15	0.30	ρ_{13} =0.40, ρ_{23} =0.25	0.35	0	0.20	0.076	-0.107	0.018	0.008	0.016
13	1000	0.30	0	0.20	ρ_{23} =0.25, ρ_{13} =0.40	0	0.25	0.30	0.096	-0.137	0.018	0.018	0.015
14	1000	0	0.15	0.30	ρ_{23} =0.25	0	0.25	0.30	0.102	-0.140	0.022	0.016	0.016
15	1000	0	0.15	0.30	ρ_{23} =0.19	0	0.25	0.30	0.114	-0.149	0.019	0.016	0.015
16	1000	0	0.15	0.30	ρ_{23} =0.13	0	0.25	0.30	0.084	-0.159	0.020	0.014	0.015

4.1.2.2 Individual Predictors that Influence the Joint Mediated Effect

Table 4.17

Exceptions to Type1 Error >0.075 While Examining X_1's Influence on $ME_{123}^{M|X}$

SL #	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.0	Type1 error when $(\gamma \,\&\, \theta_1) = 0$	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.10	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.20	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.30	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.40
13	0.096	0.096	0.064	-	-	-
14	0.102	0.102	0.076	0.028	-	-
15	0.114	0.114	0.038	-	-	-
16	0.084	0.084	0.014	-	-	-

The influence of the X_1 predictor in the 16 exceptions of the 54 conditions, to the joint mediated effect's type1 error exceeding the *a priori* set robust limit of 0.075 when $(\gamma \text{ and } \theta_1) = 0$ is shown in Table 4.17. There were zero conditions left with type1 errors that still exceeded the 0.075 limit after the cut-off value of 0.20 for the WQS weights related to the X_1 predictor. The single condition 14 from the Table 4.17 discussed earlier, remained above 0.075 at the cut-off value of 0.10 and is highlighted in bold. Condition 14 had a large sample size of *N*=1000 which decreased the estimate's *RMSE* and increased its type1 error to 0.102. Of all the 54 null conditions where $\gamma = 0$, X_1 was influential in the joint mediated effect's type1 error exceeding 0.075 for this case, since it had a weight exceeding 0.10 when the parameters $(\gamma \text{ and } \theta_1) = 0$ were both zero.

The influence of the X_2 predictor in the 16 exceptions to the joint mediated effect's type1 error exceeding the *a priori* set robust limit of 0.075 when $(\gamma \text{ and } \theta_2) = 0$ is shown in Table 4.18. There were zero conditions left with type1 errors that still exceeded the 0.075 limit after the cut-off value of 0.20 for the WQS weights related to

the X_2 predictor. The single condition 8 from the Table 4.18 remained above 0.075 at the cut-off value of 0.10 and is highlighted in bold.

Table 4.18

Exceptions to Type1 Error >0.075 While Examining X_2's Influence on $ME_{123}^{M|X}$

SL #	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.0	Type1 error when $(\gamma\ \&\ \theta_2)=0$	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.10	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.20	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.30	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.40
7	0.082	0.082	0.050	-	-	-
8	0.122	0.122	0.082	0.022	-	-
9	0.134	0.134	0.052	-	-	-
10	0.090	0.090	0.026	-	-	-
11	0.120	0.120	0.020	-	-	-
12	0.076	0.076	0.042	-	-	-

Case 8 had a large sample size of *N*=1000 with $\rho_{13}=0.40$ which was the higher correlation value between X_1 and X_3. The corresponding theta and beta parameters were $(\theta_1,\theta_3)=(0.35,0.20)$ and $(\beta_1,\beta_3)=(0.30,0.20)$ respectively. Since their magnitudes were very similar, the synergy increased the joint mediated effect estimate to 0.011 when the true value=0, resulting in an increased type 1 error of 0.122. Of all the 54 null conditions where $\gamma=0$, X_2 was influential in the joint mediated effect's type1 error exceeding 0.075 for this case, since it had a weight exceeding 0.10 when the parameters $(\gamma \text{ and } \theta_2)=0$ were both zero.

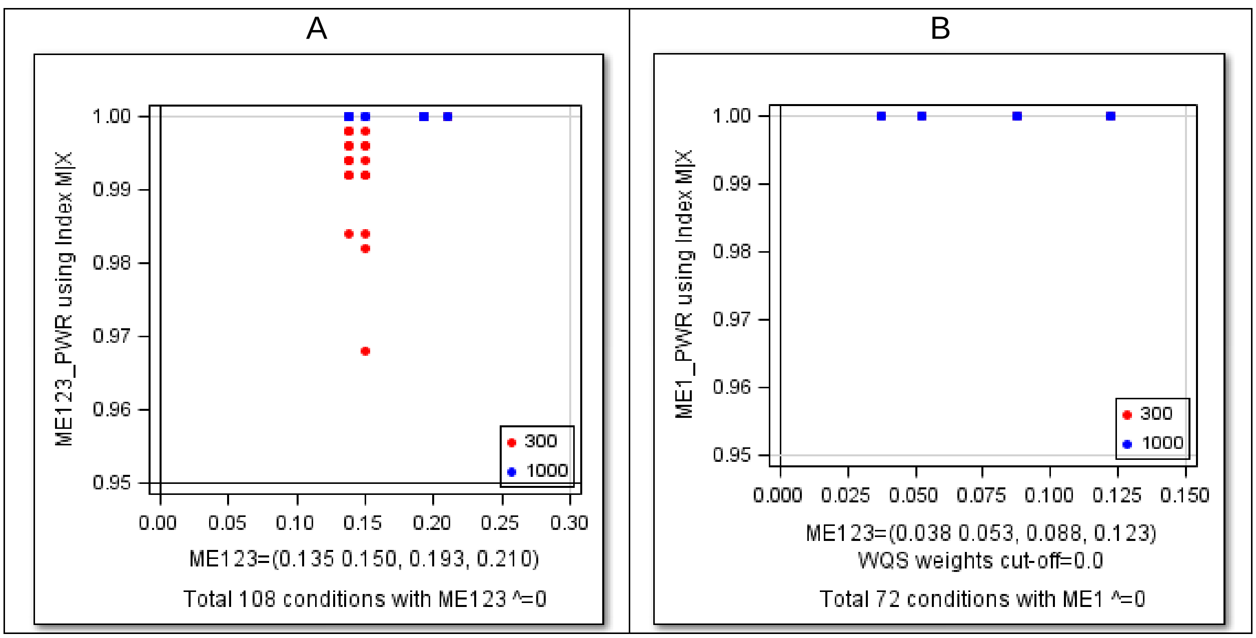
The influence of the X_3 predictor in the 16 exceptions to the joint mediated effect's type1 error exceeding the *a priori* set robust limit of 0.075 when $(\gamma\ \&\ \theta_2)=0$ is shown in Table 4.19. There were no conditions that exceeded the 0.075 limit after the cut-off value of 0.10 for the WQS weights related to the X_3 predictor.

Table 4.19

Exceptions to Type1 Error >0.075 While Examining X_3's Influence on $ME_{123}^{M|X}$

SL #	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.0	Type1 error when $(\gamma \;\&\; \theta_3)=0$	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.10	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.20	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.30	$ME_{WQS}^{M\|X,Typ1}$ cut-off=0.40
1	0.080	0.080	0.014	-	-	-
2	0.136	0.136	0.002	-	-	-
3	0.116	0.116	0	-	-	-
4	0.076	0.076	0.016	-	-	-
5	0.122	0.122	0.002	-	-	-
6	0.100	0.100	0.008	-	-	-

The joint mediated effect's power and the influential predictors that determine its value are discussed next.



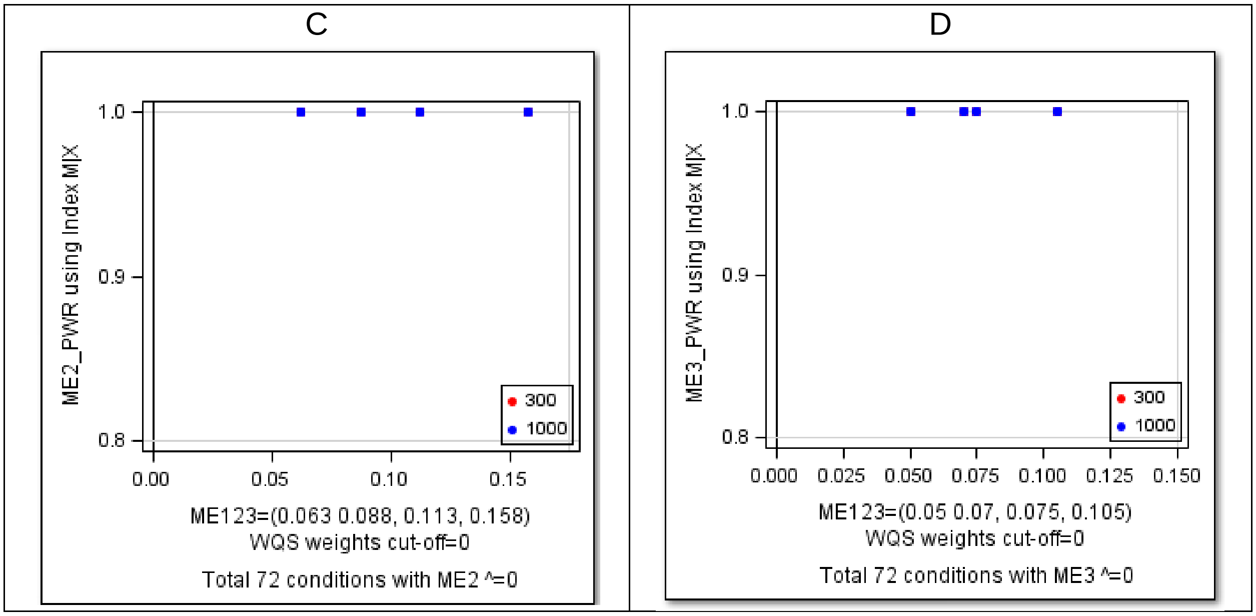


Figure 4.17 Joint Mediated Est. Power A) $ME_{WQS}^{M|X}$ B) $ME_{X_1}^{M|X}$ C) $ME_{X_2}^{M|X}$ D) $ME_{X_3}^{M|X}$ Cut-off=0.0

There were 0 of 108 possible joint mediated power estimates that had less than 80% power to detect $ME_{WQS}^{M|X}$. However, due to the smaller sample size N=300 (red dots) there were 23 cases between (0.95, 0.99) in Figure 4.17 A. Figure 4.17 B shows the power for the 72 cases where X_1 was influencing the power to detect the joint mediated effect $ME_{WQS}^{M|X}$ cut-off for WQS weights was zero. Figure 4.17 C shows the power for the 72 cases where X_2 was influencing the power to detect the joint mediated effect $ME_{WQS}^{M|X}$ cut-off for WQS weights was zero. Figure 4.17 D shows the power for the 72 cases where X_3 was influencing the power to detect the joint mediated effect $ME_{WQS}^{M|X}$ cut-off for WQS weights was zero.

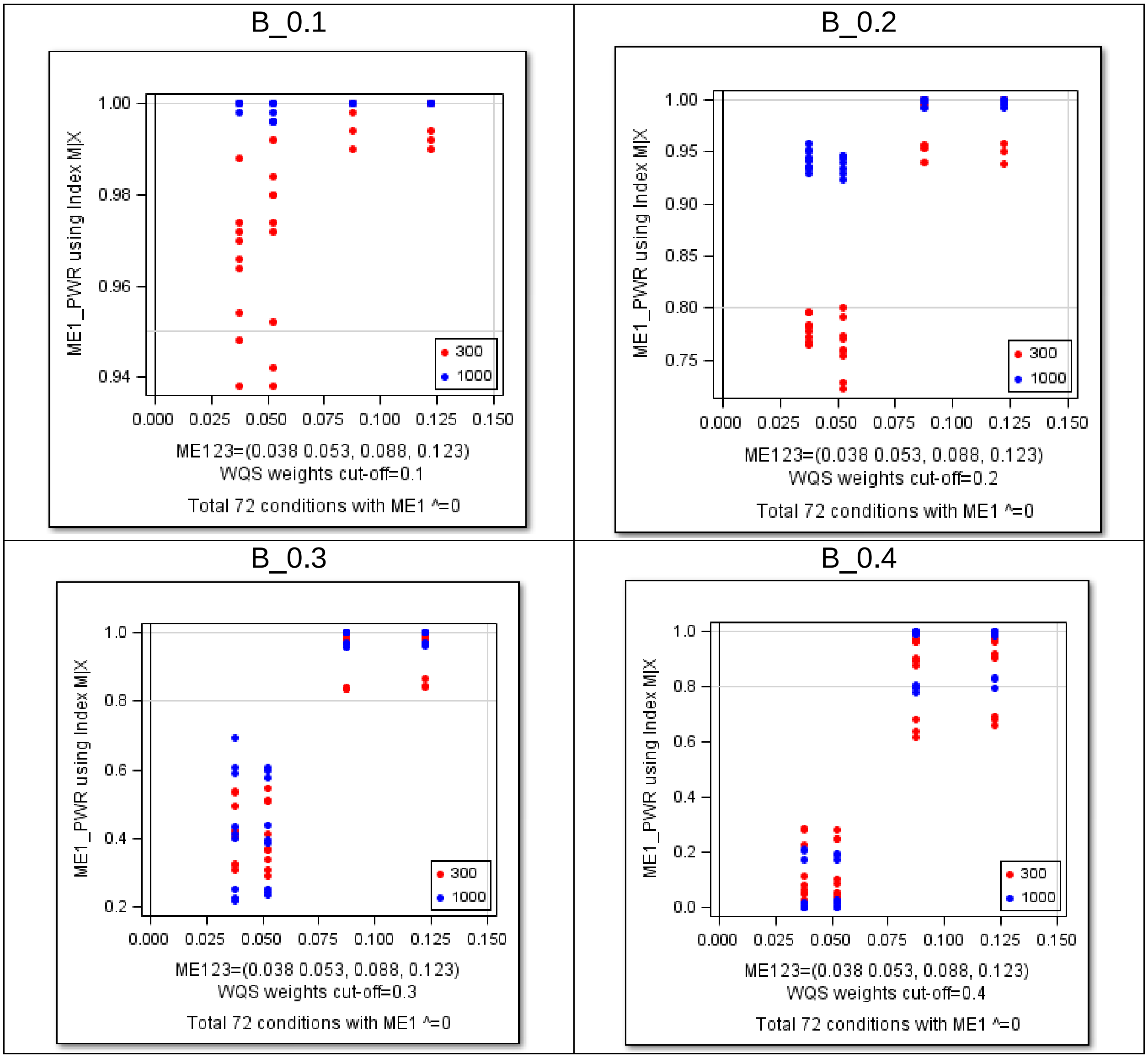


Figure 4.18 $ME_{WQS}^{M|X}$'s Power B_0.1 to B_0.4 X_1's Influence at 0.1 to 0.4 by 0.1 Cut-offs

Of the 108 cases where $ME_{123} \neq 0$, predictor X_1 contributed to the joint mediated effect in 72 conditions given that $ME_1 \neq 0$ and the power to detect the mediated effect was 1. At a cut-off = 0.2 for the WQS weight related to predictor X_1, had a range=0.72 to 0.80; all having the smaller sample size of *N*=300. At a cut-off = 0.3 for the WQS weight related to predictor X_1, had a range=0.22 to 0.69 to detect $ME_{WQS}^{M|X}$, and half having the smaller sample size. At a cut-off = 0.4 for the WQS weight related to predictor X_1, had a

range=0 to 0.79, to detect $ME_{WQS}^{M|X}$. A maximum reasonable cut-off to include X_1's influence in the joint mediated effect would be 0.2, provided the sample size could be increased, but the type1 error determines the choice for a uniformly applied cut-off value for all individual effects.

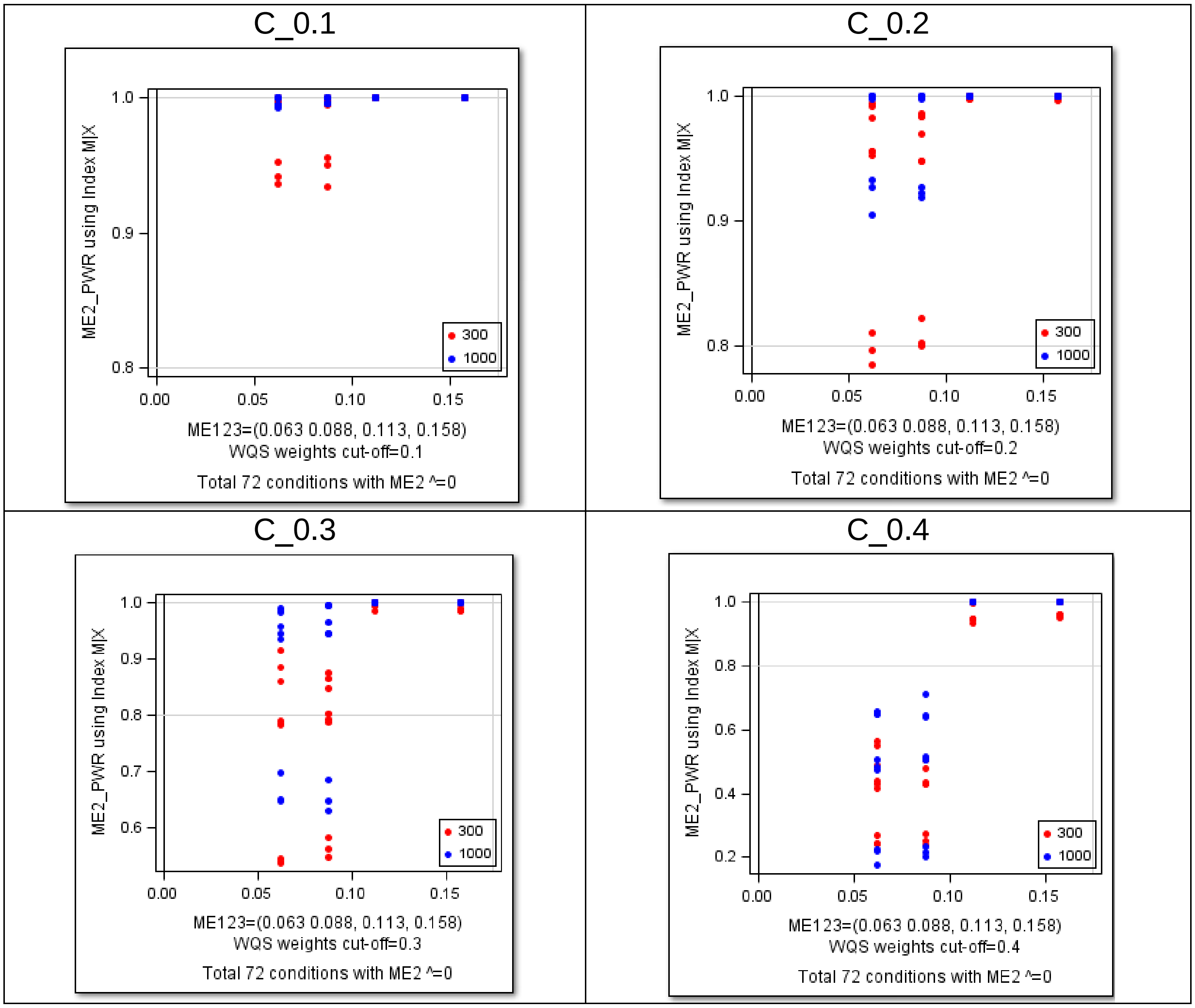


Figure 4.19 $ME_{WQS}^{M|X}$'s Power C_0.1 to C_0.4 X_2's Influence at 0.1 to 0.4 by 0.1 Cut-offs

Of the 108 conditions where $ME_{123} \neq 0$, predictor X_2 contributed to the joint mediated effect in 72 conditions given that $ME_2 \neq 0$ and the power to detect the mediated effect was 1. At a cut-off = 0.2 for the WQS weight related to predictor X_2, had a range=0.54 to 0.55; all having the smaller sample size of N=300. At a cut-off = 0.3 for

the WQS weight related to predictor X_2, had a range=0.54 to 0.79, to detect $ME_{WQS}^{M|X}$, and11 of 17 had the smaller sample size. At a cut-off = 0.4 for the WQS weight related to predictor X_2, had a range=0.17 to 0.71 to detect $ME_{WQS}^{M|X}$, half having the smaller sample size. A maximum reasonable cut-off to include X_2's influence in the joint mediated effect would be 0.2, provided the sample size could be increased to reduce the number of exceptions down from 2, but the type1 error determines the choice for a uniformly applied cut-off value for all individual effects.

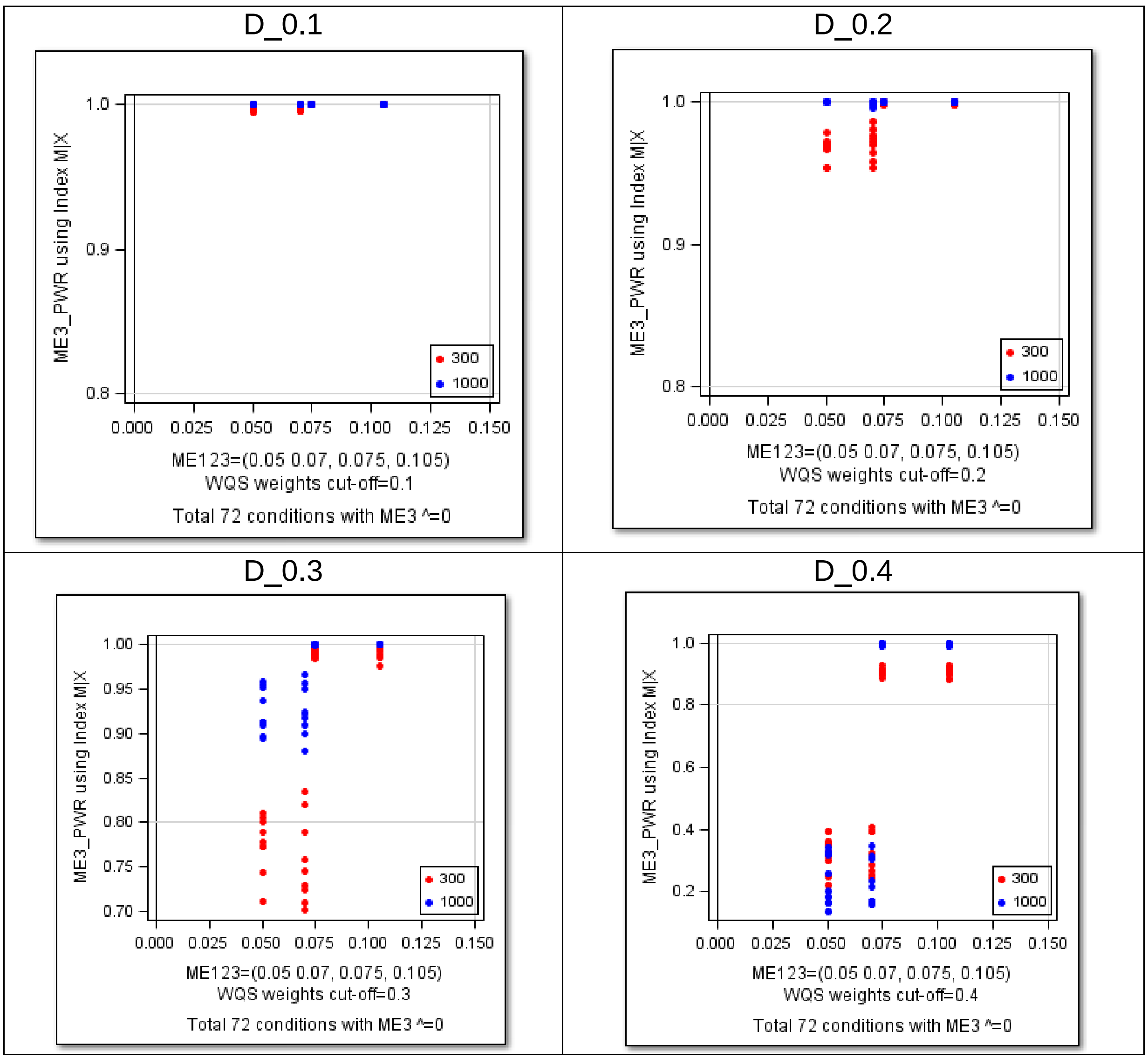
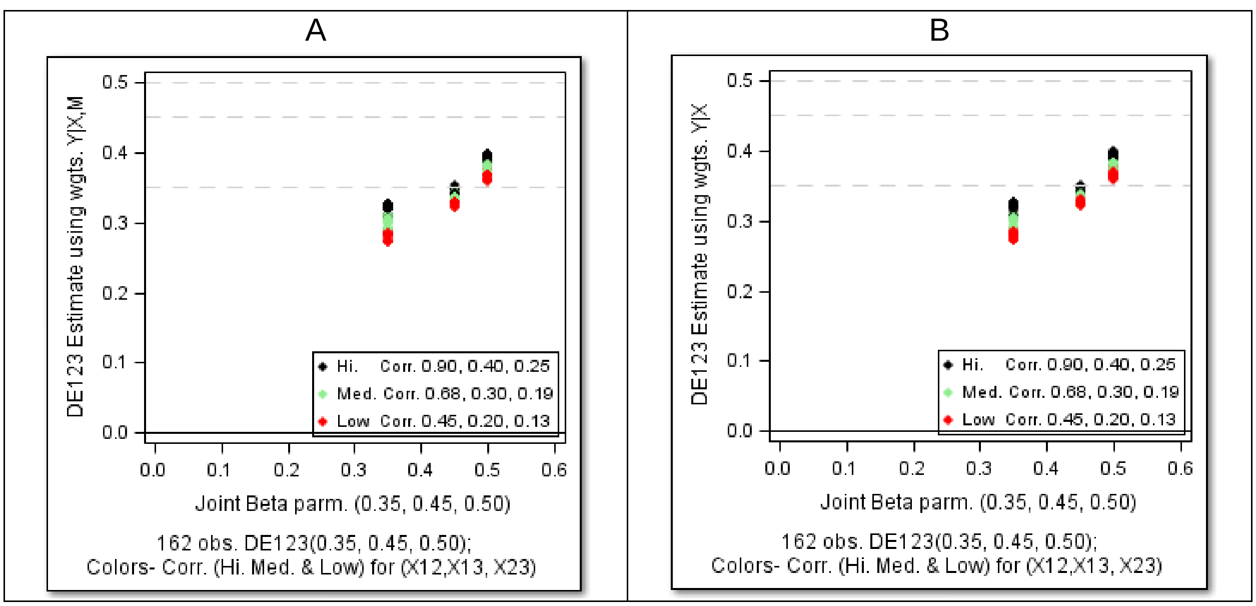


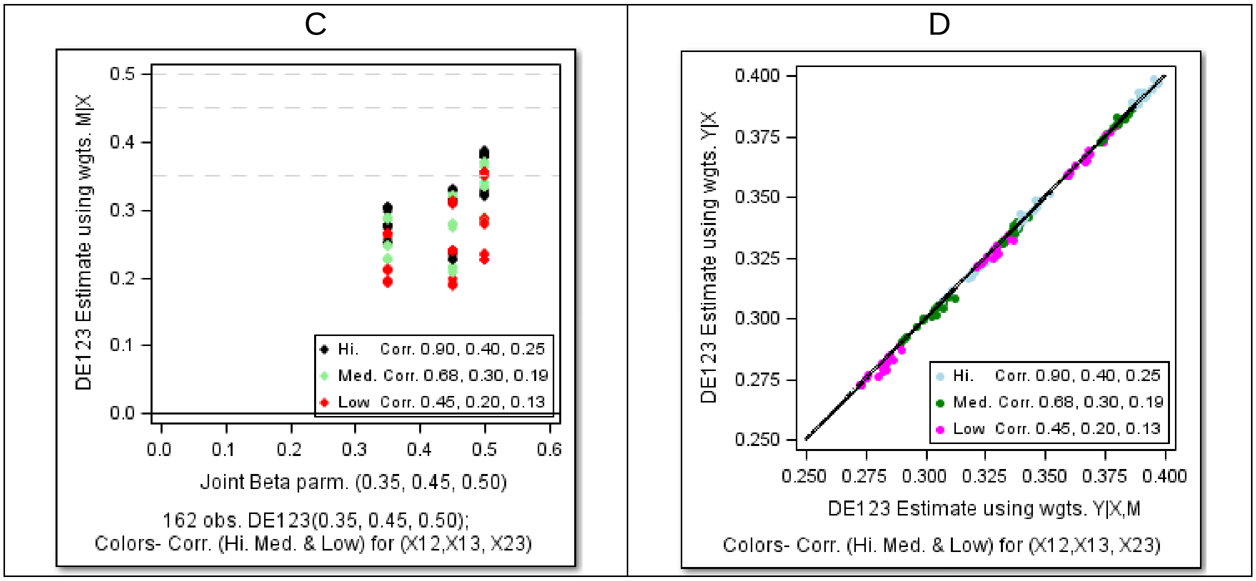
Figure 4.20 $ME_{WQS}^{M|X}$'s Power D_0.1 to D_0.4 X_3's influence at 0.1 to 0.4 by 0.1 cut-offs

Of the 108 conditions where $ME_{123} \neq 0$, predictor X_3 contributed to the joint mediated effect in 72 conditions given that $ME_3 \neq 0$ and the power to detect the mediated effect was 1. At a cut-off = 0.3 for the WQS weight related to predictor X_3, had a range=0.70 to 0.79, all having the smaller sample size of N=300. At a cut-off = 0.4 for the WQS weight related to predictor X_3, had a range=0.13 to 0.41, to detect $ME_{WQS}^{M|X}$, and 11 of 36 had the smaller sample size. A maximum reasonable cut-off to include X_3's influence in the joint mediated effect would be 0.3, provided the sample size could be increased, but the type1 error determines the choice for a uniformly applied cut-off value for all individual effects.

The joint direct effects for 3-variable mediation will be studied in the next section.

4.1.2.3 Joint Direct Effects using $WQS_{index}^{Y|X}$, $WQS_{index}^{Y|X,M}$ or $WQS_{index}^{M|X}$ *Est., Bias & RMSE*



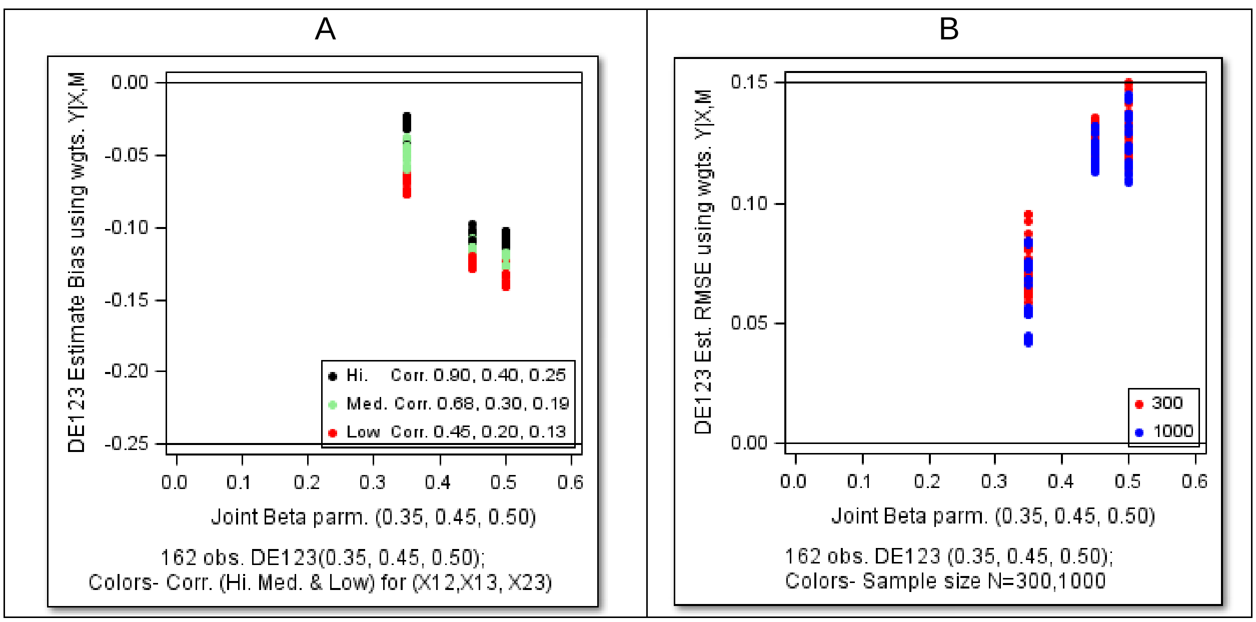


Figures 4.21 Joint Direct Effect for $WQS_{index}^{Y|X}$ A-D) $WQS_{index}^{Y|X,M}$ $WQS_{index}^{Y|X}$ $WQS_{index}^{M|X}$ & $Y|X$ vs. $Y|X,M$

Figure 4.21 A and B are very similar since either $WQS_{index}^{Y|X,M}$ or $WQS_{index}^{Y|X}$ WQS index serves well in estimating the joint direct effect, A slight preference for $WQS_{index}^{Y|X,M}$ over $WQS_{index}^{Y|X}$ is evident from the comparative plot D, especially for low correlations (magenta dots) to medium correlations (green dots) as related ($\rho_{12}, \rho_{13}, \rho_{23}$) . Panel C shows $DE_{WQS}^{M|X}$ estimates having estimates with a larger negative bias as compared to plots A and B, for all 162 conditions. This suggests that $WQS_{index}^{M|X}$ is an inappropriate index to use when estimating the joint direct effect.

There are nine sub-clusters of estimates for each joint beta parameter set, shown on the x-axis, in plot C at $\beta_{123} = (0.35, \beta_3 = 0), (0.45, \beta_1 = 0),$ and $(0.50, \beta_2 = 0)$. They are grouped by $\theta_{123} = (0.55, \theta_1 = 0), (0.55, \theta_2 = 0),$ and $(0.60, \theta_3 = 0)$, having high, medium, and low $DE_{WQS}^{M|X}$ estimate values, e.g. the 54 estimates at $\beta_{123} = (0.45, \beta_1 = 0)$ have three groups of 18 estimates with average values of (0.32, 0.28, and 0.21), and each group of 18 estimates have three sub-groups of six estimates based on the pairwise correlations

$\rho_{12}, \rho_{13}, \rho_{23}$ being high, medium or low (color coded as black, light green, and red). However, when calculating $DE_{WQS}^{Y|X,M}$ using $WQS_{index}^{Y|X,M}$ as shown in panel A, and $DE_{WQS}^{Y|X}$ using $WQS_{index}^{Y|X}$ as shown in panel B, these clusters are overlapping since different indices are being used to derive estimates that are not sensitive to the theta parameters alone, as is $DE_{WQS}^{M|X}$ which is derived using $WQS_{index}^{M|X}$ (affected by theta parameters). Figure 4.21 A-C identify a consistent color pattern, showing the joint direct effect estimate's direct dependence on the pairwise predictor correlations. The figures also show that the WQS regression method applies a negative bias of varying magnitudes to the joint direct estimate, regardless of the index used to derive it. The next figures show that negative bias for the joint direct estimates $DE_{WQS}^{Y|X,M}$ and it's *RMSE*.



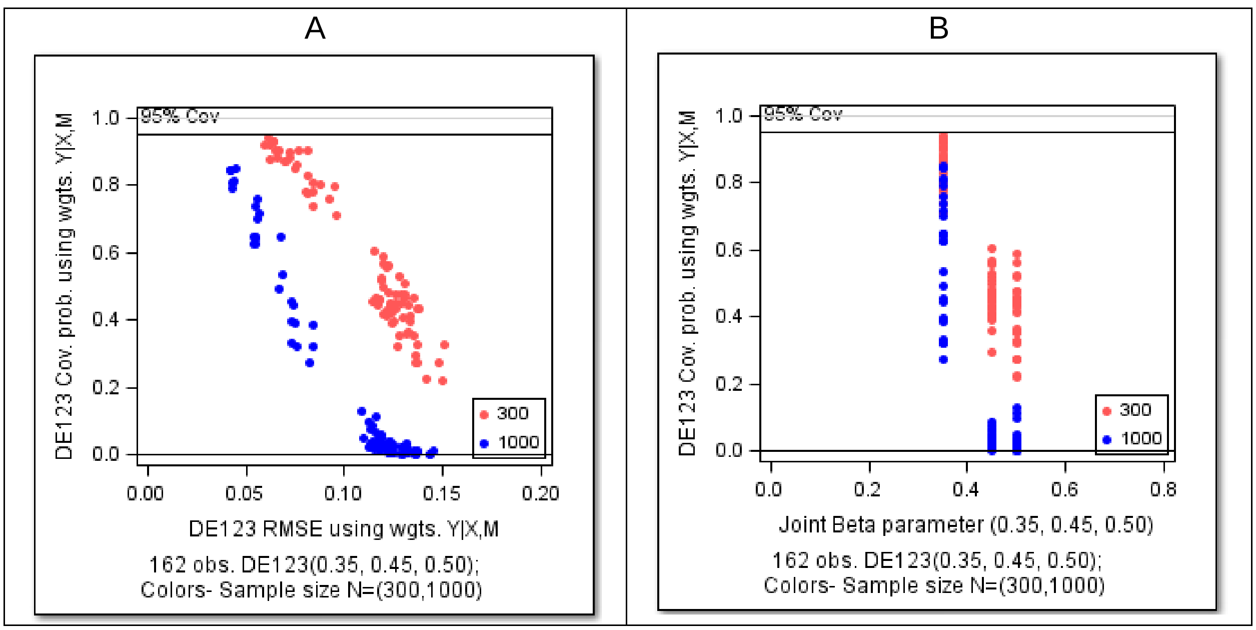
Figures 4.22 WQS Joint Direct Effect $DE_{WQS}^{Y|X,M}$ From $WQS_{index}^{Y|X,M}$ A) Bias and B) *RMSE*

The plots shown in Figure 4.22 are the $DE_{WQS}^{Y|X,M}$'s bias in panel A, and its *RMSE* in panel B, for all 162 conditions grouped by joint beta values $\beta_{123} = (0.35, \beta_3 = 0)$,

$(0.45, \beta_1 = 0)$, and $(0.50, \beta_2 = 0)$, on the x-axis. The pairwise correlations chosen for (ρ_{12}, ρ_{13}) make X_1 the most influential amongst (X_1, X_2, X_3) , and X_3 the least influential predictor, because of its correlations (ρ_{31}, ρ_{32}) , $\rho_{31} \in \{0.40, 0.30, 0.20\}$, and $\rho_{32} \in \{0.25, 0.19, 0.13\}$. The beta parameter values for predictors (X_1, X_2, X_3) , essentially determine the joint direct effects. For example, in Figure 4.22 A given $\beta_{123} = (0.35, \beta_3 = 0)$, should produce large $DE_{WQS}^{Y|X,M}$ values (least negative bias), since $(\beta_1, \beta_2) \neq 0$ which are associated with the two influencing predictors (X_1, X_2) based on the chosen pairwise correlations. In the 54 conditions with $\beta_{123} = (0.45, \beta_1 = 0)$, the influencing X_1 predictor has $\beta_1 = 0$, which should produce diminished values (larger negative bias) of $DE_{WQS}^{Y|X,M}$. In comparison $\beta_{123} = 0.50$ with $\beta_2 = 0$, and the joint beta parameter is larger (0.50 vs. 0.45) giving the resulting $DE_{WQS}^{Y|X,M}$ an additional negative bias through regression, which explains the arrangement of the estimate's bias in plot A. The joint beta estimates are directly proportional to the magnitude of the pairwise correlations, which explains why low correlated predictors (red dots) have a larger negative bias as compared to the higher correlated predictors (black dots) shown closer to the zero line. The estimate's *RMSE* is dependent on the value of the joint direct effect and the sample size. Larger sample sizes (blue dots), have a smaller standard error for the estimate mean. Referring back to Figure 4.22 A, the dispersion of the $DE_{WQS}^{Y|X,M}$ estimates about their means is largest for $\beta_{123} = (0.35, \beta_3 = 0)$ and smallest for $\beta_{123} = (0.45, \beta_1 = 0)$ which is reflected in the spread of *RMSE* values for the three clusters.

4.1.2.4 Coverage probability and Power of the joint direct effect $DE_{WQS}^{Y|X,M}$ Estimate.

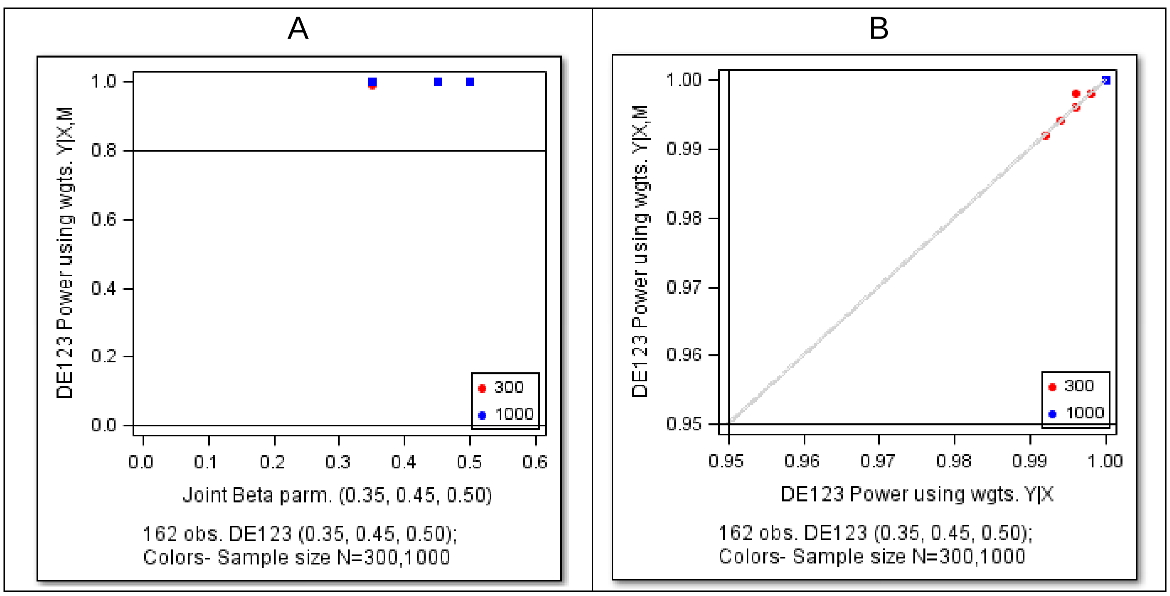
Figure 4.23A shows 162 $DE_{WQS}^{Y|X,M}$'s Coverage probabilities plotted on the y-axis with their corresponding $DE_{WQS}^{Y|X,M}$'s *RMSE* along the x-axis. The three influencing variables that determine an estimate's coverage probability are: 1) the estimate value, 2) the estimate's *RMSE* and 3) the sample size that influences the estimate's *RMSE.* The estimate's *RMSE* values are the same as those shown in Figure 4.22B and should look familiar. Figure 4.23A shows two clusters, one with 54 (2 x27) estimates and the other cluster with 27 estimates having higher coverage probabilities), by sample size (N=300(red) and N=1000(blue)).



Figures 4.23 WQS Joint Direct Effect $DE_{WQS}^{Y|X,M}$ Coverage Prob. by A) RMSE B) Joint Beta

The first cluster is composed of two nearly overlapping sub-groups, $\beta_{123} = (0.45, \beta_1 = 0)$ and $\beta_{123} = (0.50, \beta_2 = 0)$. This is because the joint theta values are close in value (0.45, 0.50) and the fact that one of the two influencing parameters (β_1, β_2) has a zero value which results in a smaller $DE_{WQS}^{Y|X,M}$ estimate with a larger

negative bias, for both the sub-groups (as shown in Figure 4.22A). For example, the 54 blue dot cluster closest to the y=0 line having coverage probabilities (0 to 0.13), belongs to the overlapping sub-groups, $\beta_{123} = (0.45, \beta_1 = 0)$ and $\beta_{123} = (0.50, \beta_2 = 0)$. The same is true for the 54 blue dot cluster just above it. The *RMSE* for these two groups is higher (range=0.1 to 0.15) as compared to the smaller 27 estimate clusters $\beta_{123} = (0.35, \beta_3 = 0)$. The power for the joint direct effect $DE_{WQS}^{Y|X,M}$ estimates and the individual predictors influencing their detection are discussed next. There were no conditions where $DE_{123} = 0$, allowing the WQS model to be assessed for its performance with regards the joint and individual type1 error rate.



Figures 4.24 WQS Joint direct effect $DE_{WQS}^{Y|X,M}$ power A) $DE_{WQS}^{Y|X,M}$ B) $DE_{WQS}^{Y|X,M}$ vs. $DE_{WQS}^{Y|X}$

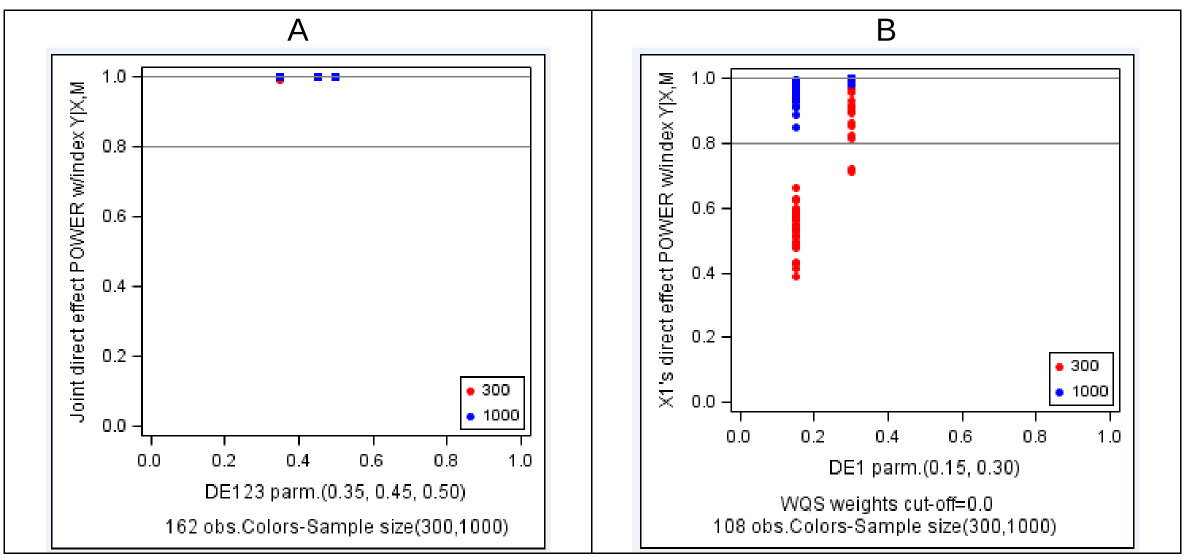
Figure 4.24A and B show that $DE_{WQS}^{Y|X,M}$ and $DE_{WQS}^{Y|X}$ are very comparable except for a single condition having a sample size of *N*=300, which is in favor of using the index $WQS_{index}^{Y|X,M}$ over $WQS_{index}^{Y|X}$ with regard the joint direct effect's power. Of the total of 162 conditions having the condition $DE_{132} \neq 0$, there were seven conditions, five related to

$\beta_{123} = (0.35, \beta_3 = 0)$ and one for each of $\beta_{123} = (0.45, \beta_1 = 0), \& \ (0.50, \beta_2 = 0)$, with the sample size of *N*=300, (shown in Figure 4.24 A and B), having a power which deviated from 1.

4.1.2.5 Individual Predictors Which Influence the Joint Direct Effects

The influence of individual predictors X_1, X_2 and X_3, in determining the joint direct effect's Type1 error, cannot be determined since there are no conditions where DE_{123}=0. If there were conditions with DE_{123}=0 then $DE_{WQS}^{M|X,Typ1_{0.1cut-off}}$ $DE_{WQS}^{M|X,Typ1_{0.2cut-off}}$ $DE_{WQS}^{M|X,Typ1_{0.3cut-off}}$ and $DE_{WQS}^{M|X,Typ1_{0.4cut-off}}$ Type1 errors >0.075 exceptions would guide in determining the optimum cut-off value for the WQS weights used to determine X_1, X_2 and X_3's Type1 error rate.

The joint direct effect's Power and the influential predictors that determine its value when the cut-off value $DE_{WQS}^{Y|X,M}$ is set to 0 are discussed next.



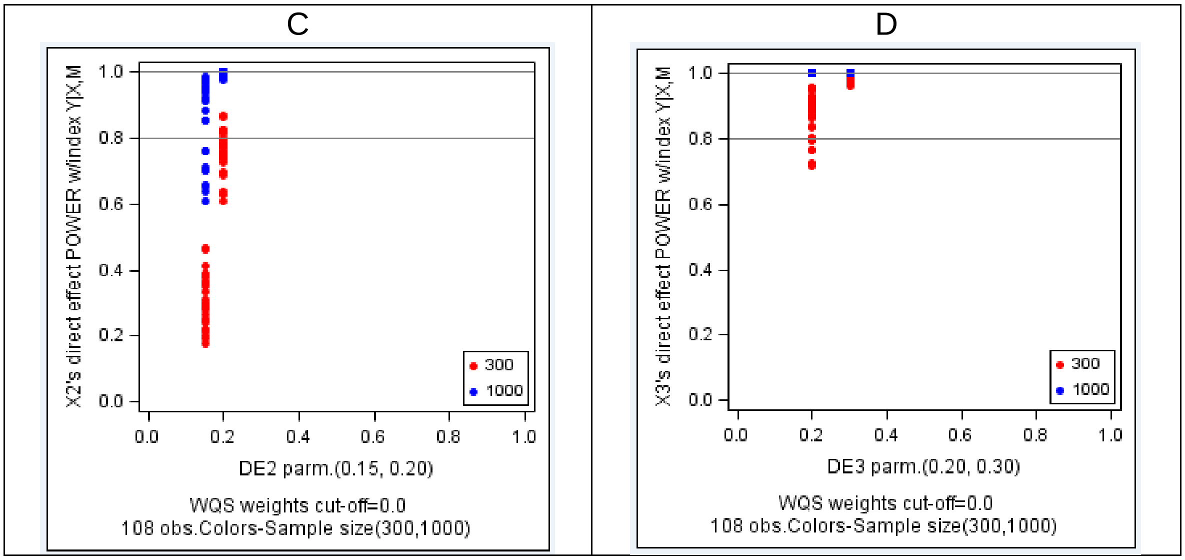


Figure 4.25 Direct effect est.'s Power A) $DE_{WQS}^{Y|X,M}$ B) $DE_{X_1}^{Y|X,M}$ C) $DE_{X_2}^{Y|X,M}$ D) $DE_{X_3}^{Y|X,M}$ cutoff=0.0

The influence of predictor X_1 in detecting a joint direct effect for the possible 108 conditions having $DE_1 \neq 0$, were all at 0.79 for variable X_1. This suggests that a reasonable cut-off value for X_1 could be 0.1, which would continue to make it influential in detecting the joint direct effect with a power greater than 0.8.

Moving towards analyzing the influence of predictor X_2 in detecting a joint direct effect having a range 0.77 to 0.79, when the cut-off is set to 0.1. The range for variable X_2 was 0.48 to 0.80 when the cut-off is set to 0.2, a range from 0.14 to 0.58 when the cut-off is set to 0.3, and a range from 0 to 0.80, when the cut-off is set to 0.4. This suggests that a reasonable cut-off value for X_2 could be 0.1, which would continue to make it influential in detecting the joint direct effect.

Moving towards analyzing the influence of predictor X_3 in detecting a joint direct effect having $DE_3 \neq 0$, with power at 0.78, for variable X_3 when the cut-off is set to 0.3. X_3 had a range from 0.38 to 0.69 when the cut-off was set to 0.4, and with $\beta_3 = 0.20$ or with $\beta_3 = 0.30$. This illustrates the importance of the strength of association between the

predictor and the outcome as denoted by the value of its beta parameter. Earlier when determining the influence of individual predictors in detecting the joint mediated effect we had noted the importance of having adequate sample sizes. Although the analysis of the individual predictor's influence in detecting a joint direct effect was optimum at 0.1 cut-off value, the choice for the cut-off value, which is applied uniformly to all individual predictor effects, is determined by the optimal individual type1 error rates. The joint direct effect's power and the predictors that determine its value when the cut-off value $DE_{WQS}^{Y|X}$ is set to 0 are discussed next.

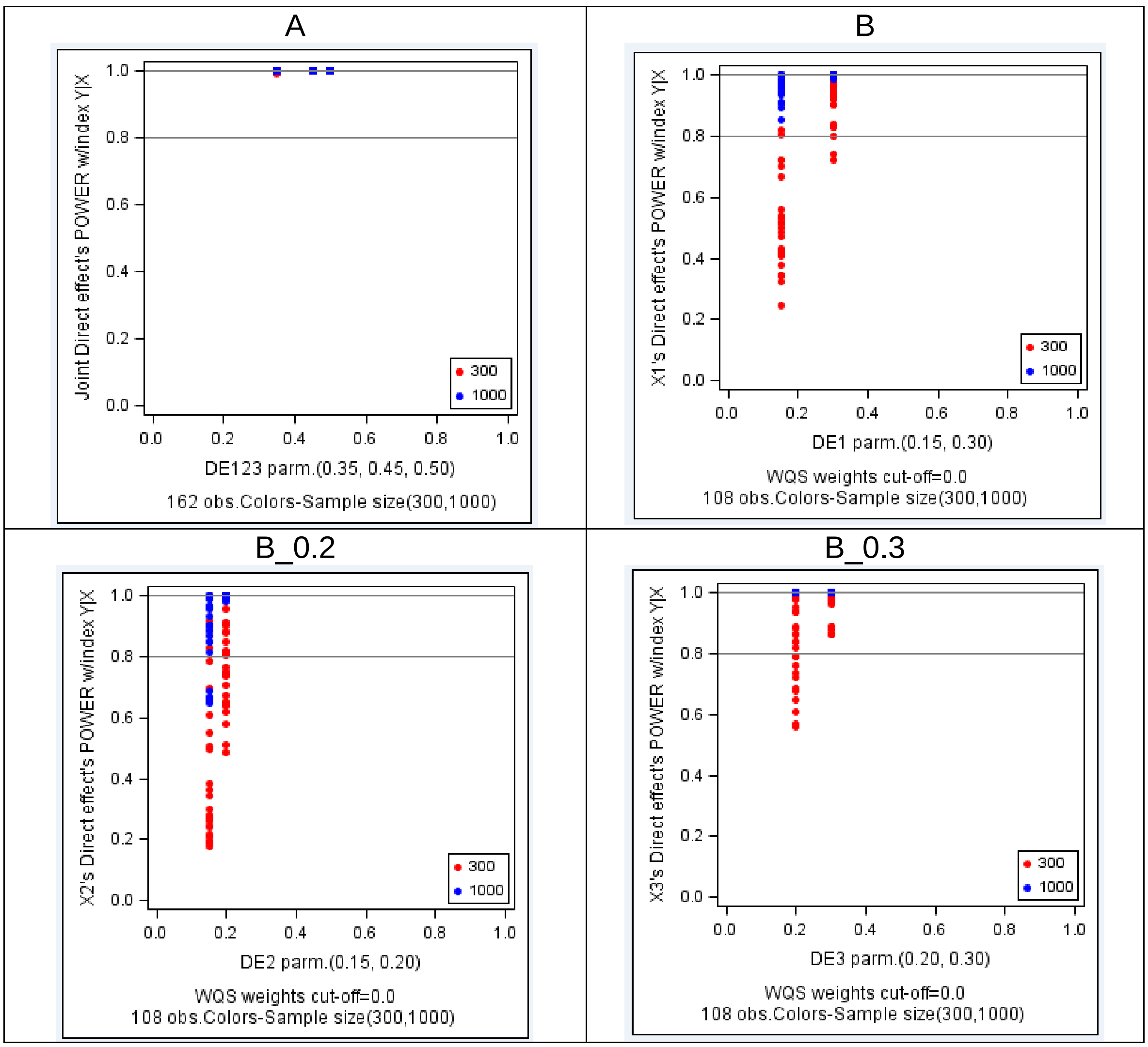


Figure 4.26 Direct effect est.'s Power A) $DE_{WQS}^{Y|X}$ B) $DE_{X_1}^{Y|X}$ C) $DE_{X_2}^{Y|X}$ D) $DE_{X_3}^{Y|X}$ cutoff=0.0

The joint direct effect's Power and the influential predictors that determine its value when the cut-off value $DE_{WQS}^{M|X}$ is set to 0 are discussed next.

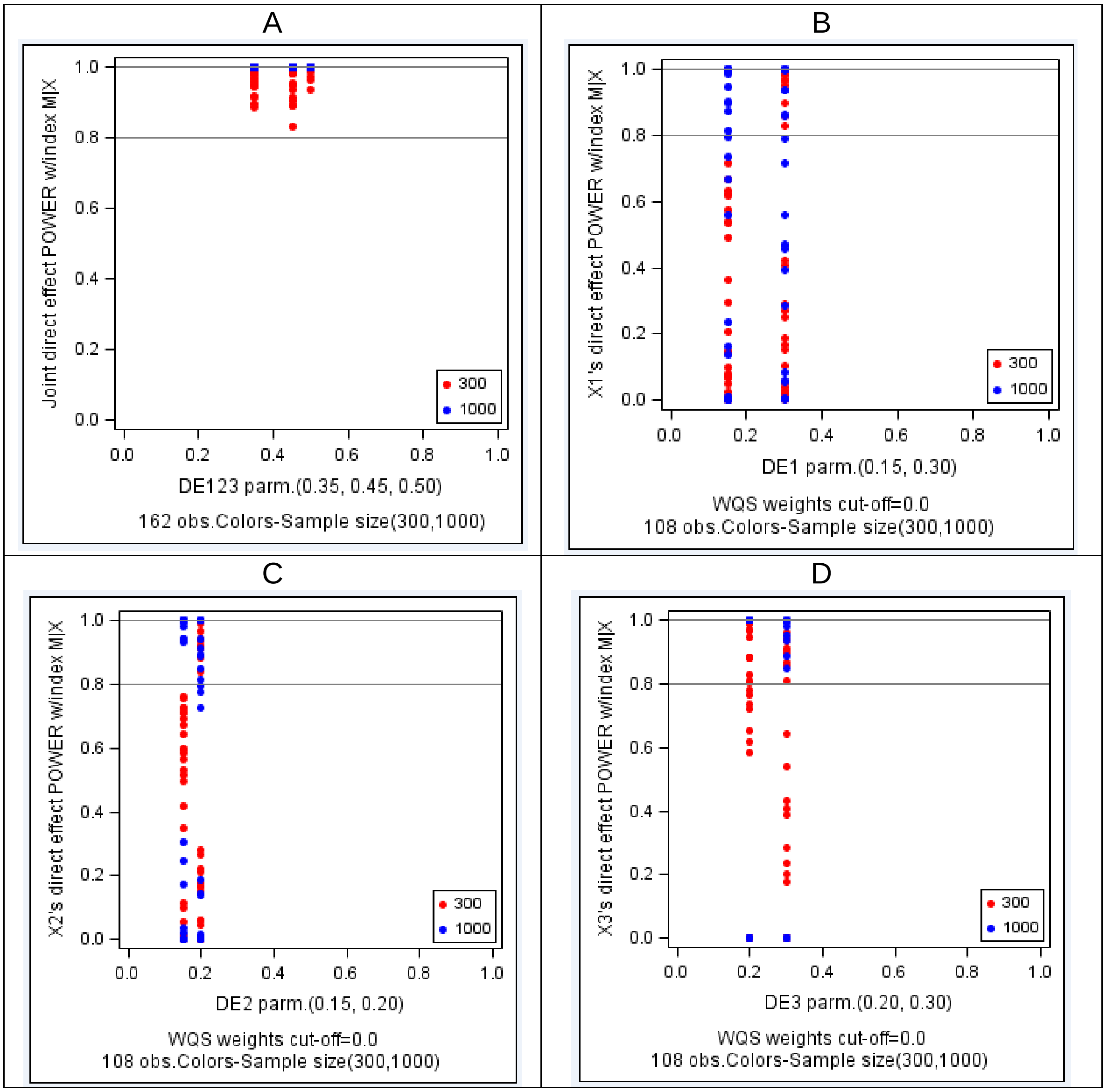


Figure 4.27 Direct Effect Est.'s Power A) $DE_{WQS}^{M|X}$ B) $DE_{X_1}^{M|X}$ C) $DE_{X_2}^{M|X}$ D) $DE_{X_3}^{M|X}$ Cutoff=0.0

From Figures 4.17, 4.18 and 4.19 have evident that the best power to detect the joint mediated effect and identify the individual predictors that contribute to its value of power, is by using the WQS index $WQS_{index}^{Y|X,M}$ in preference to $WQS_{index}^{Y|X}$ or $WQS_{index}^{M|X}$ to

detect $DE_{WQS}^{Estimate}$. The influential predictors that determine the joint direct effect's power and its value using varying cut-off values 0.1 to 0.4 by 0.10 are discussed next.

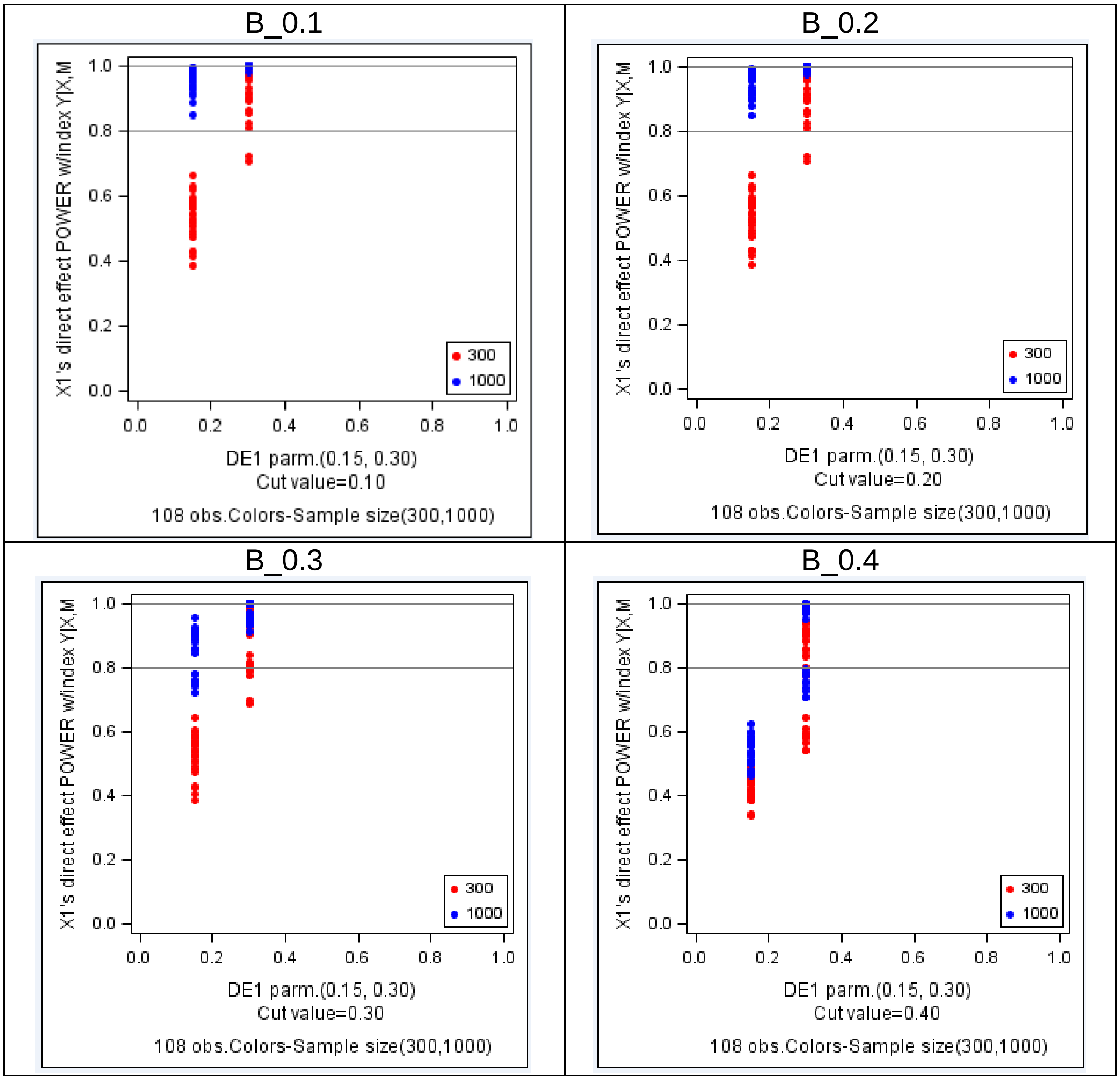


Figure 4.28 WQS Joint direct effect $DE_{WQS}^{Y|X,M}$ Power for X_1 at 0.1 to 0.4 by 0.1cutoffs

Of the 162 conditions where $DE_{123} \neq 0$, predictor X_1 contributed to the joint mediated effect in 108 conditions given that $DE_1 \neq 0$ and the power to detect the direct effect was between 0.99 and 1. A maximum reasonable cut-off to include X_1's influence

in the joint mediated effect would be 0.2, provided the sample size could be increased to reduce the 29 exceptions. The power for X_1 is especially low when the effect size is small.

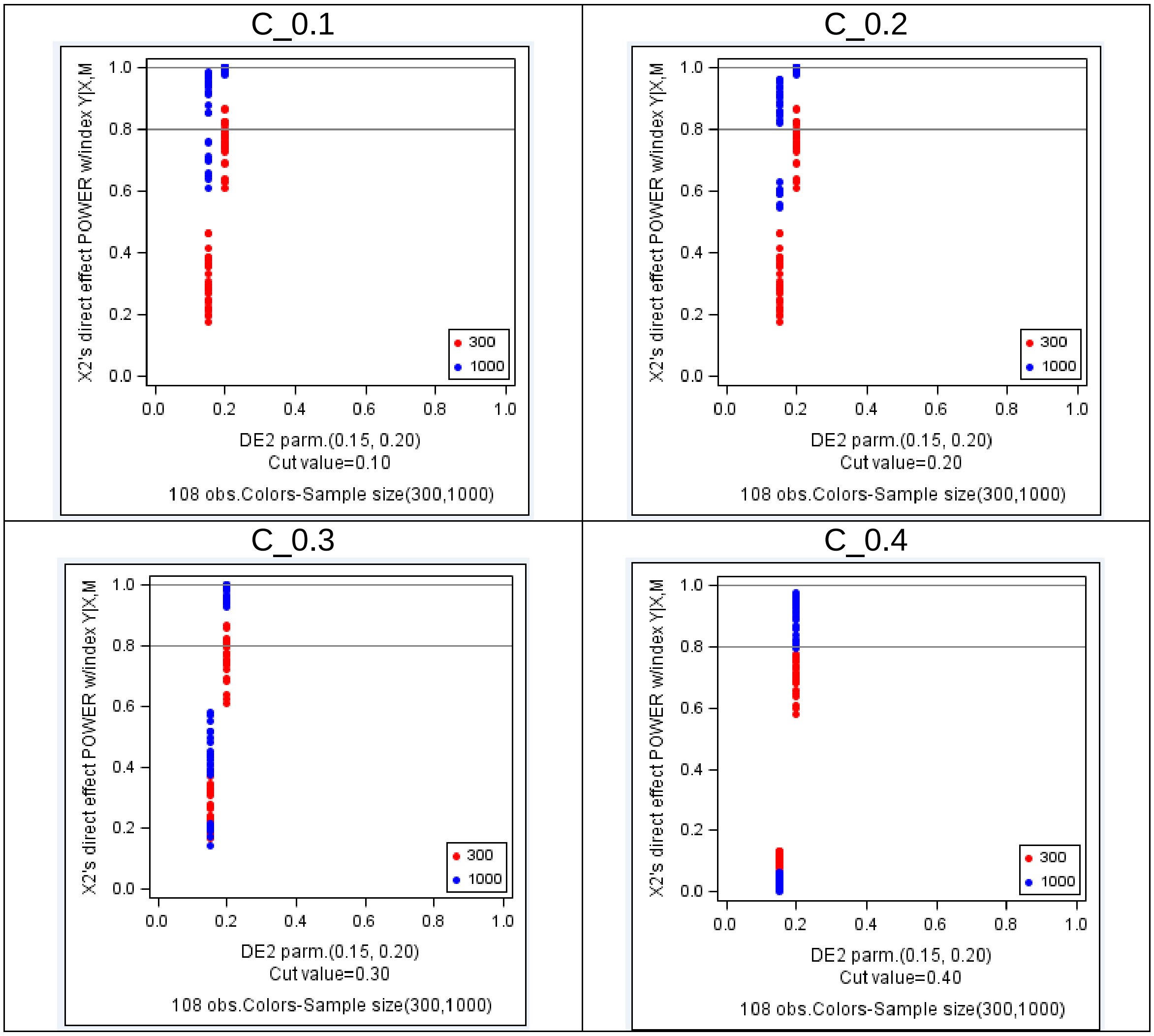


Figure 4.29 WQS Joint direct effect $DE_{WQS}^{Y|X,M}$ Power for X_2 at 0.1 to 0.4 by 0.1cutoffs

Of the 162 conditions where $DE_{123} \neq 0$, predictor X_2 contributed to the joint mediated effect in 108 conditions given that $DE_2 \neq 0$ and the power to detect the direct effect was between 0.18 and 1. A maximum reasonable cut-off to include X_2's influence in the joint mediated effect would be 0.2, provided the sample size could be increased

to reduce the 54 exceptions. X_2*'s* power is especially low when the effect size is small (0.15 vs. 0.20).

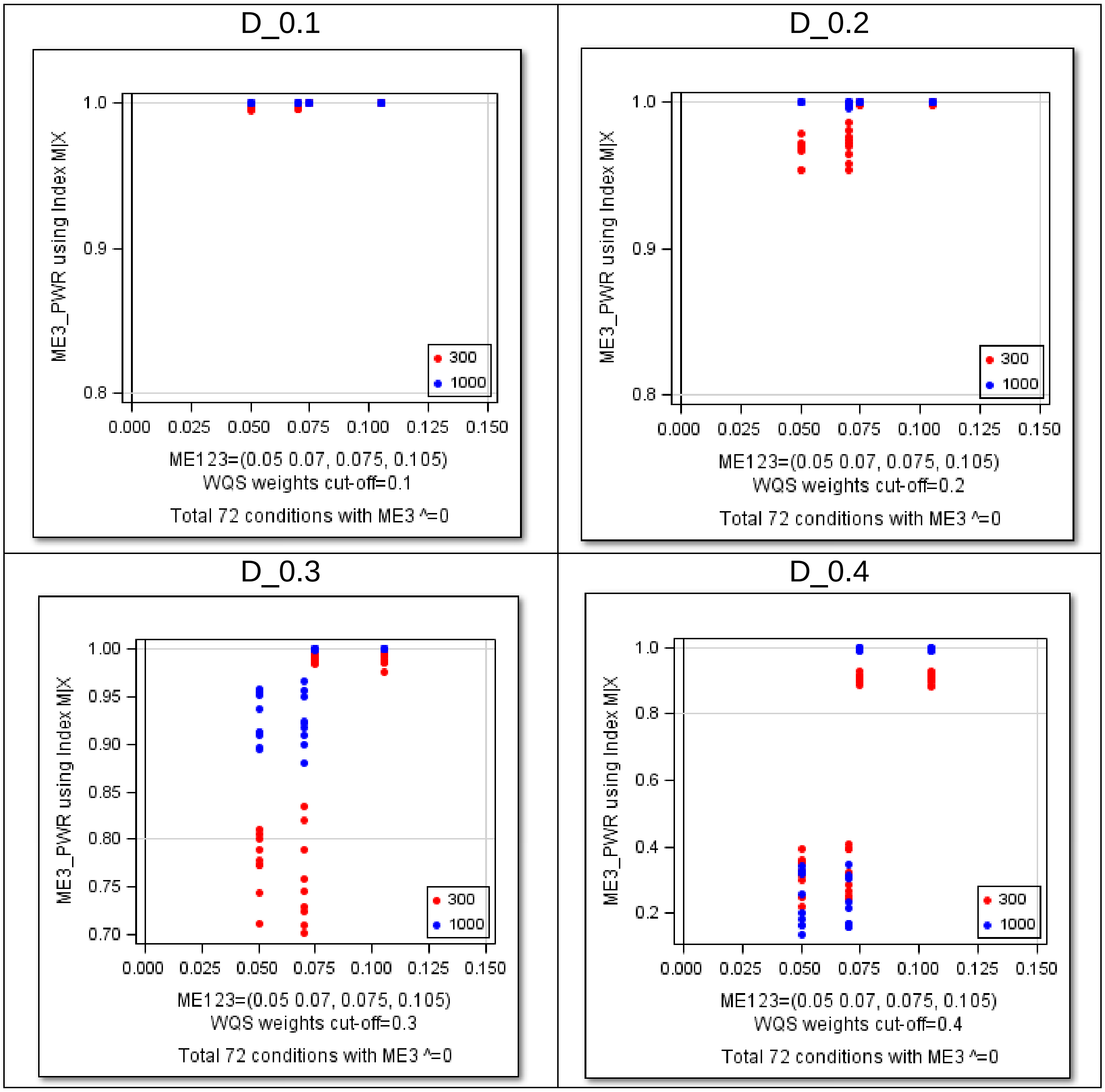


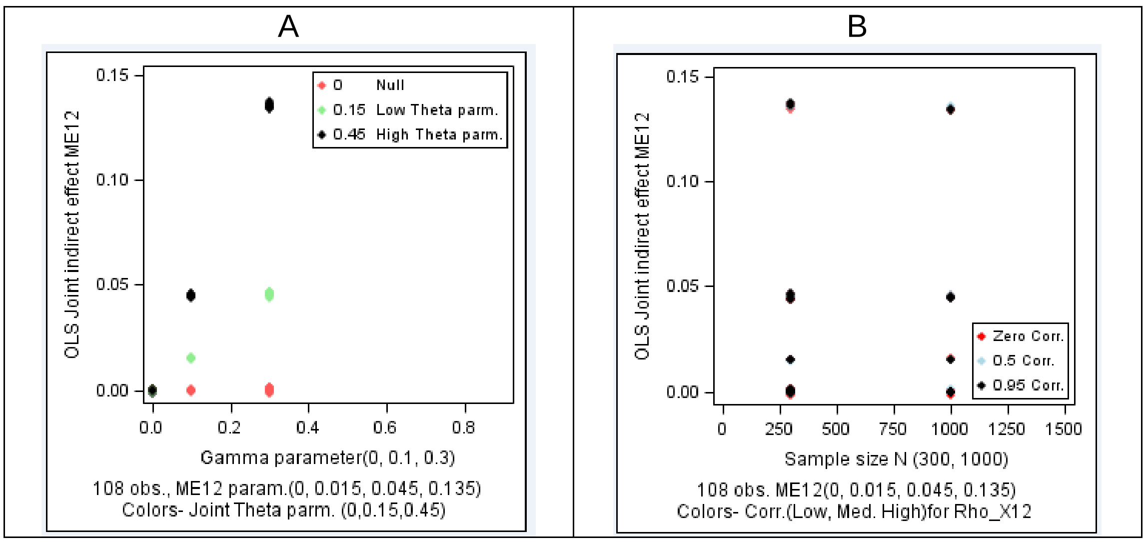
Figure 4.30 WQS Joint direct effect $DE_{WQS}^{Y|X,M}$ Power for X_3 at 0.1 to 0.4 by 0.1cutoffs

Of the 162 conditions where $DE_{123} \neq 0$, predictor X_3 contributed to the joint mediated effect in 108 conditions given that $DE_3 \neq 0$ and the power to detect the direct effect was between 0.72 and 1. A maximum reasonable cut-off to include X_3's influence in detecting the joint mediated effect would be 0.3. X_3*'s* power is especially low when

the effect size is small (0.20 vs. 0.30). The decision for an applicable cut-off value for all individual direct effects for three variable mediation models is guided primarily by the acceptable type1 error rates for the individual predictors.

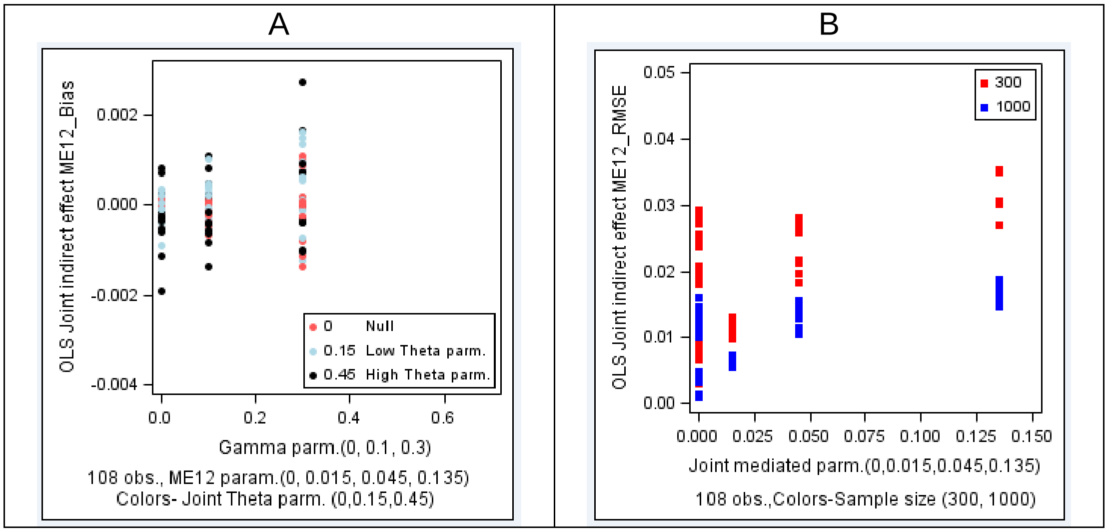
4.2 Ordinary Least Squares 2 & 3 Variable Mediation Analysis

4.2.1.1 OLS Joint 2-Variable Mediated Effect –*Estimate, Bias* & *RMSE*



Figures 4.31 A-B OLS Joint Indirect Effect Estimates for ME_{12}^{OLS}

OLS joint mediated effects increase with the joint theta and gamma parameters as shown in 4.31 A. Pairwise correlations between predictors do not influence the joint mediated effect as shown in 4.31 B. There is a low bias for each ME_{12}^{OLS} estimate value leading to the conclusion that the OLS method produces unbiased coefficient estimates.



Figures 4.32 OLS Joint Indirect Effect Estimate for ME_{12}^{OLS} A) *Bias* and B) *RMSE*

The ME_{12}^{OLS} estimate bias is small (±0.002) for estimates 0 to 0.135 attributable to the OLS regression method. The estimate's standard error and the average estimate's *RMSE* is large for the parameter being estimated. The *RMSE* is inversely proportional on the sample size but directly proportional to the mediated effect (Figure 4.32 B).

4.2.1.2 OLS Joint Mediated Effect Coverage Probability, Type1 Error & Power

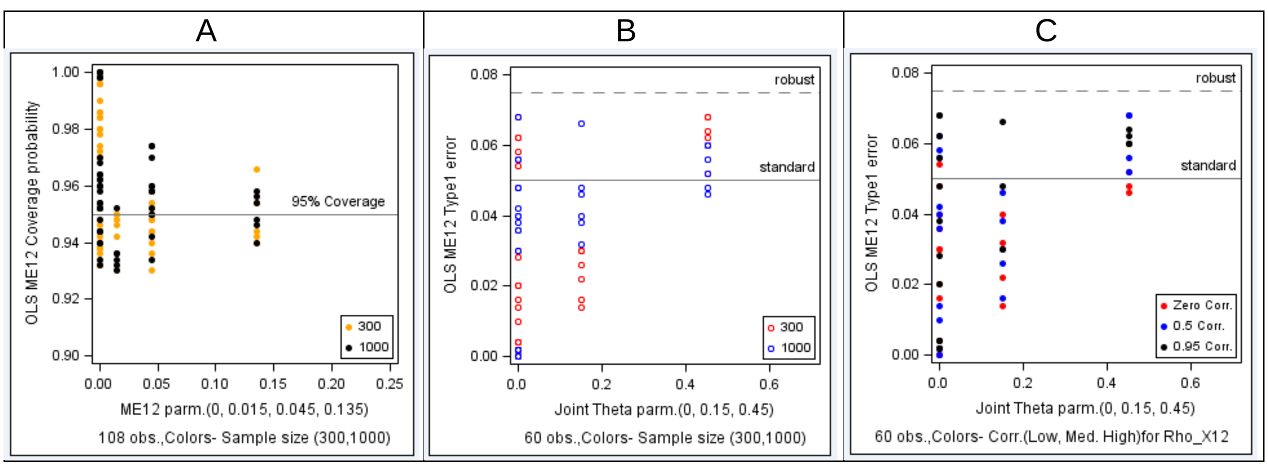


Figure 4.33 Joint Indirect Effect Estimate for ME_{12}^{OLS} A) Coverage B) & C) Type1 Error

The joint mediated effect's coverage probability for $ME_{12} = 0.045$ is the result of two different parameter arrangements. ME_{12}=0.045 when $\gamma = 0.1, \theta_{12} = 0.45$ and $\gamma = 0.3, \theta_{12} = 0.15$, and ME_{12}=0 can be achieved in three ways. Figure 4.33 B shows ME_{12} for $\gamma = (0,\ 0.1,\ 0.3), \theta_{12} = (0,\ 0.15,\ 0.45)$ and their corresponding joint mediated effect's type1 errors. Type1 error increases with increasing values of θ_{12}, since a positive bias on the gamma coefficient estimate, applied to the joint theta parameter $\hat{\gamma}^{OLS}\hat{\theta}_{12}^{OLS}$, produces a non-zero joint mediated effect with a type1 error (plots B and C). No exceptions for ME_{12}^{OLS}, all 60 are below the chosen robust type1 error limit of 0.075.

4.2.1.3 Joint Mediated Effect – OLS Power to Detect ME_{12}^{OLS} When $ME_{12}≠0$

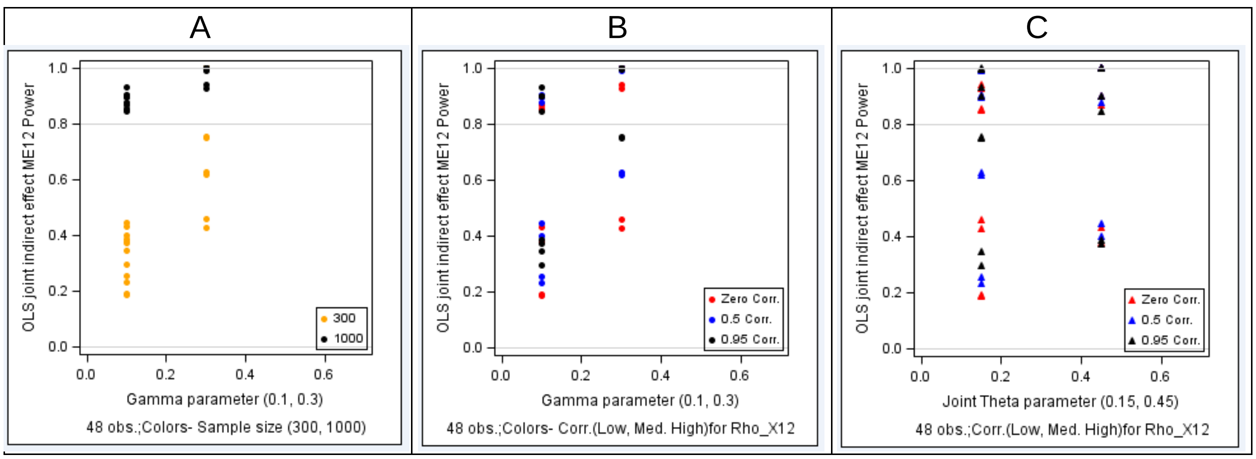


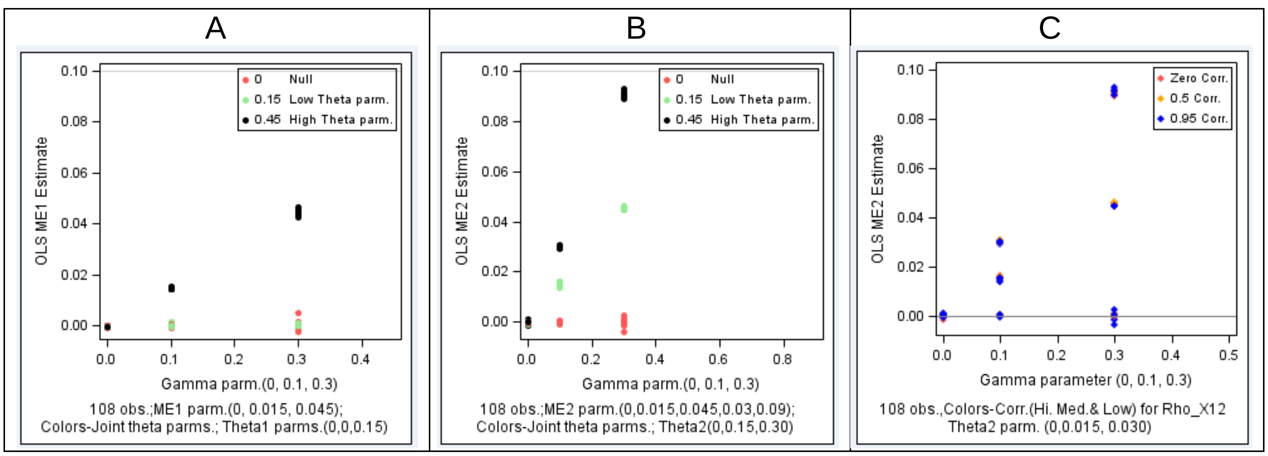
Figure 4.34 OLS Method Joint Indirect Effect Estimate A-C) Power to Detect ME_{12}^{OLS}

A larger joint mediated effect results from larger gamma and joint theta parameter estimates which also increases the power since the effect size is larger, Figure 4.34 A-C. There were 18 exceptions of the 48 conditions that had a power less than 0.8 to detect ME_{12}^{OLS} when $ME_{12}≠0$. All 18 exceptions were because of the smaller sample size (N=300). Pairwise predictor correlations directly influence the power to

detect ME_{12}^{OLS} , since higher correlations increase the ME_{12}^{OLS} joint estimate values and consequently the power to detect them. Figure 4.34 C shows that the joint theta parameter also directly influences the power of the OLS method in detecting ME_{12}^{OLS} .

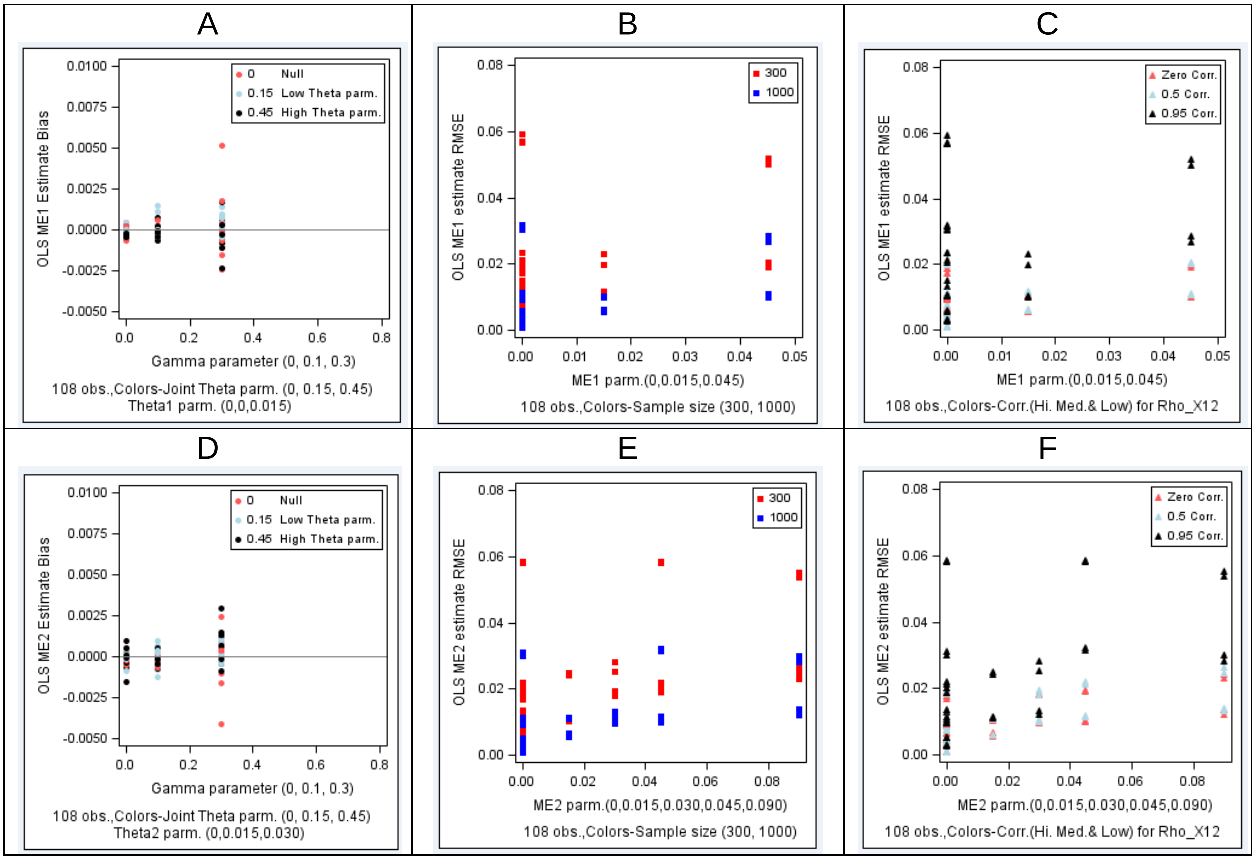
Six exceptions shown in plot C at $\theta_{12} = 0.45$ had a lower power of 0.40. Combining the information from Figure 4.34 A and C, the conclusion is that the OLS method has a lower power for predictors with a small effect size in a small sized, uncorrelated datasets.

4.2.1.3 OLS Individual Mediated Effects – *Estimate, Bias* & *RMSE*



Figures 4.35 OLS Individual Indirect Effect Estimates for $ME_{1}^{OLS}, ME_{2}^{OLS}$

OLS individual mediated effects are an increasing function of the individual theta and gamma parameters as shown in Figure 4.35 A for ME_{1}^{OLS} , Figure 4.35 B and C for ME_{2}^{OLS} . Each individual $ME_{1,2}^{OLS}$ estimate has a low bias, since all the estimates are closely centered on the true parameter value.



Figures 4.36 OLS Individual Indirect Effect Estimate's *Bias, RMSE* for ME_1^{OLS}, ME_2^{OLS}

The ME_{12}^{OLS} estimate bias is small (±0.005) for estimates (0, 0.09) attributable to the OLS regression method. Figure 4.36 A and D show that the dispersion of the estimates centered on the parameter value increases for increasing gamma parameter values $\gamma = (0,\ 0.1,\ 0.3)$. For a mediated effect to be zero given gamma is non-zero, requires that the joint theta parameter to be zero. With $\rho_{12} = 0.95$, the joint theta estimate gets multiplied by the larger gamma parameter $\hat{\theta}_{12}^{OLS}\hat{\gamma}^{OLS}$, which increases the mediated estimate bias. The mediated estimate bias is positive for *ME₁* plot A (0.005) and negative for *ME₂* plot D (-0.004), since predictors *X₁, X₂* are highly correlated $\rho_{12} = 0.95$ in the multiple regression model. The two extreme biases are for the same condition $N = 300, \rho_{12} = 0.95$. The estimate's standard error is inversely dependent on the sample

size and is the highest for small datasets (red squares) with high pairwise correlations (black triangles), as shown in Figures 4.36 B and C for *ME₁*, and E and F for *ME₂*.

4.2.1.4 OLS Individual Mediated Effect –Coverage Prob., Type1 Error & Power

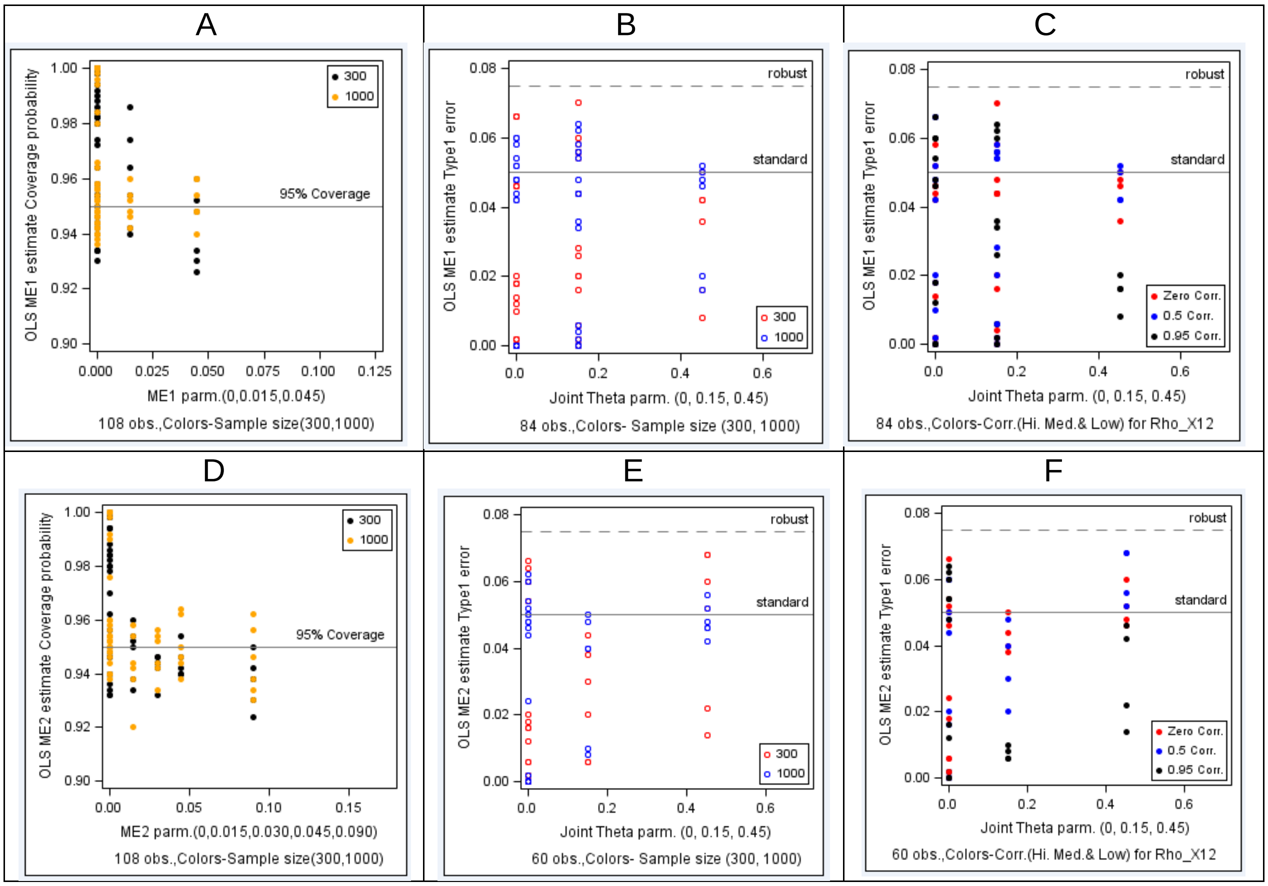


Figure 4.37 Individual indirect effect $ME_{1,2}^{OLS}$ A, D) Coverage B&C, E&F) Type1 error

The OLS method has coverage probability for individual mediated effects that exceeds 0.92 for both $ME_{1,2}^{OLS}$ for all the 108 conditions. Figures 4.37 B&C, E&F show that the individual mediated effect’s type1 error is below the chosen robust limit of 0.075. The type1 error is shown to increase (with increasing values of the non-zero theta parameter, given $\gamma = 0$) for $ME_1^{Typ1}, \theta_1 = 0.15$ in plot B and for $ME_2^{Typ1}, \theta_2 = (0.15,\ 0.30)$ in plot E. Small sample sizes decrease the type1 error rate since they provide wider

confidence intervals which includes zero. Figures 4.37 C and F show that the type1 error is lower for individual mediated effects having high pairwise correlations.

4.2.1.5 Individual Mediated Effect – OLS's Power to Detect $ME_{1,2}^{OLS}$, $ME_{1,\,2} \neq 0$

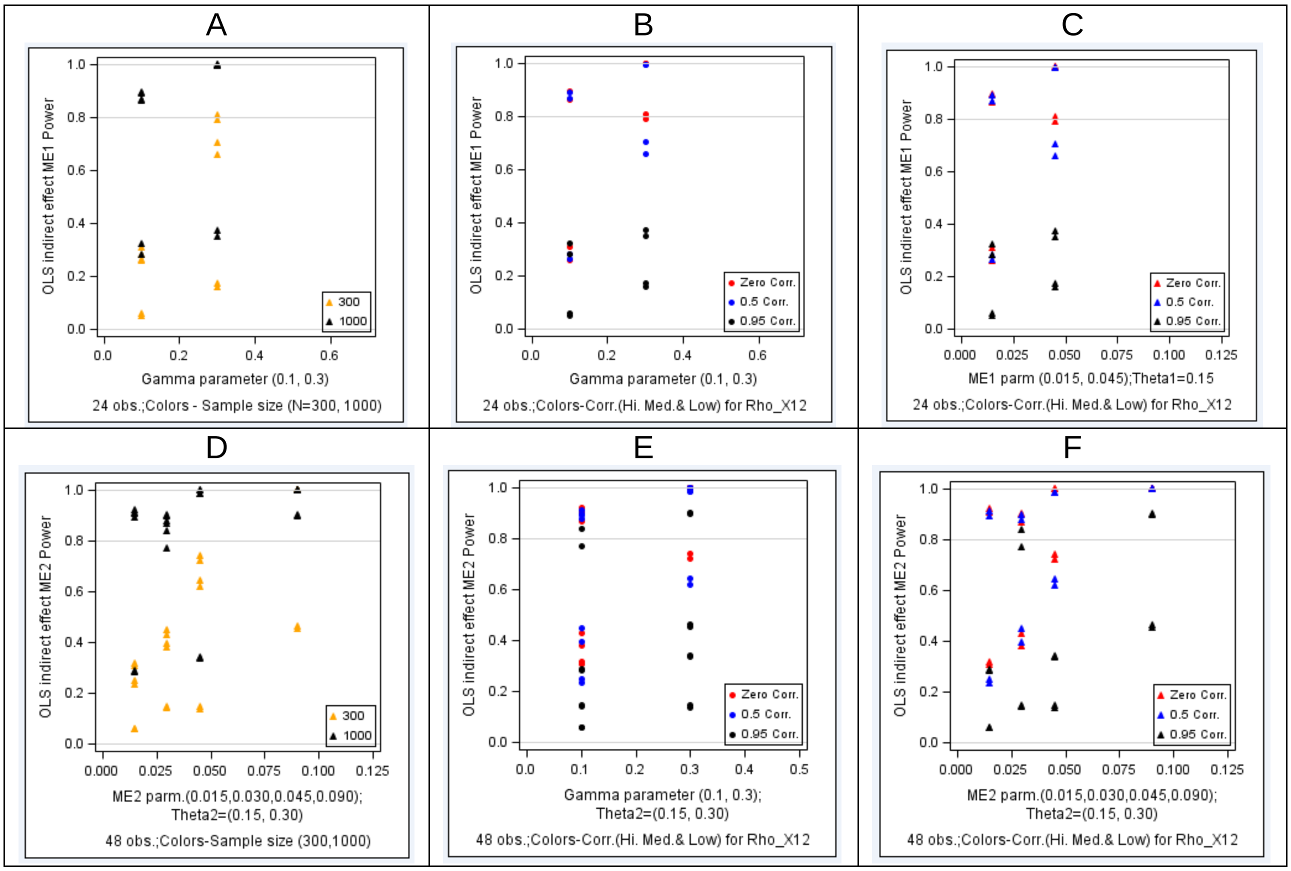


Figure 4.38 OLS Individual Indirect Effect Est.'s Power to Detect A-C) ME_1^{OLS} D-F) ME_2^{OLS}

There were 15 exceptions for 24 conditions that had power to detect ME_1^{OLS} with $ME_1 \neq 0$. Eleven of the 15 exceptions were due to a smaller sample size *N*=300, and 4 were with the larger sample size but with $\rho_{12} = 0.95$. A correlation of $\rho_{12} = 0.95$ reduces the power for ME_1 in these four conditions since the standard error for the OLS mediated estimate is large ME_1^{RMSE} =0.010 for a small effect size of 0.015. The large

standard error of the estimate results in a wide confidence interval which includes zero in 68% of the data replications with the low power of 0.32.

There were 25 exceptions for 48 conditions that had power to detect ME_2^{OLS} with *ME₂*≠0. Twenty of the 25 exceptions were because of the smaller sample size *N*=300, and five were with the larger sample size but with $\rho_{12} = 0.95$. A correlation of $\rho_{12} = 0.95$ reduces the power for ME_2 in these five conditions since the standard error for the OLS mediated estimate is large $ME_2^{RMSE} = 0.012$ for a small effect size of 0.015. The large standard error of the estimate results in a wide confidence interval which includes zero in 72% of the data replications, resulting in a low power of 0.28. Increasing the gamma parameter value $\gamma = 0.3, 0.1$, increases the power since the effect size for ME_2^{OLS} is larger (0.090 vs. 0.030). Combining the information from A-C & D-F the conclusion is that the OLS method has the worst power when the predictors have small effect sizes in small datasets with highly correlated independent predictors.

4.2.1.6 OLS Joint Direct effect – *Estimate, Bias & RMSE*

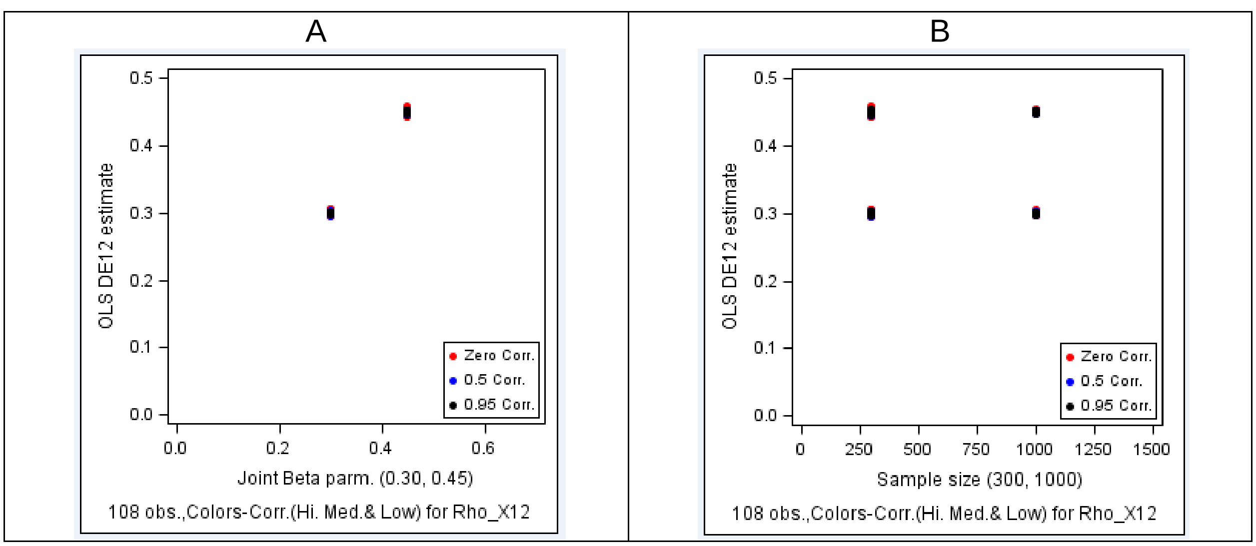
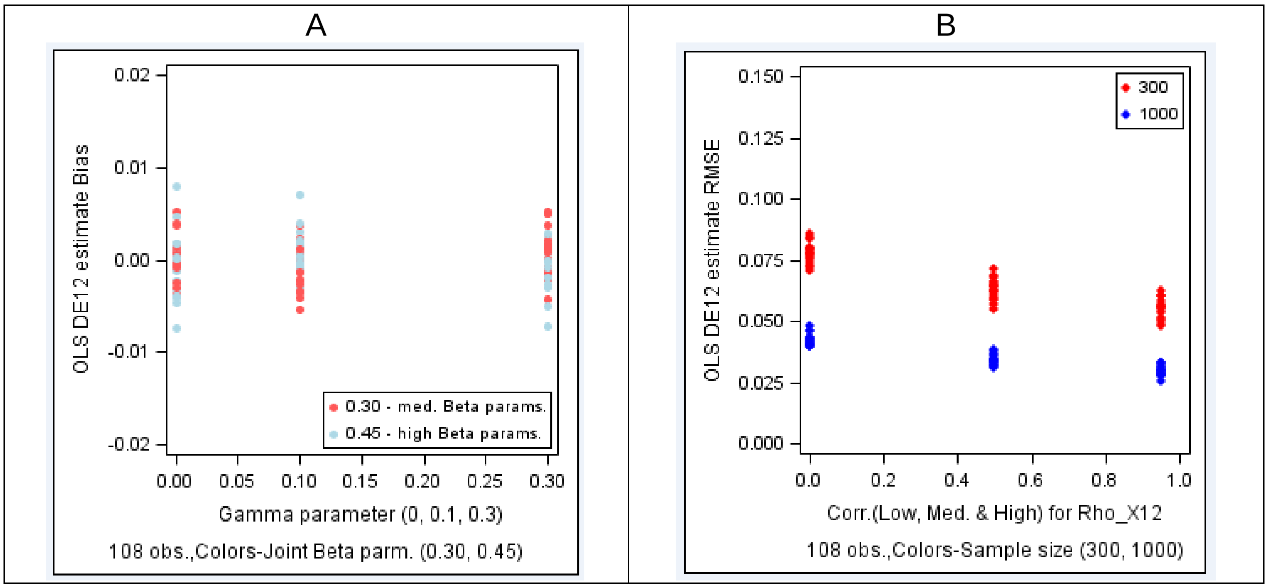


Figure 4.39 A-B Joint Direct Effect Estimate for DE_{12}^{OLS} Using OLS Regression Method

OLS joint direct effects are an increasing function of the joint beta parameter as shown in Figure 4.49 A. Gamma parameter values, pairwise correlations and sample size do not influence the joint direct effect as shown in Figure 4.49 B. Each DE_{12}^{OLS} estimate value has a low bias since the OLS method produces unbiased regression coefficient estimates.



Figures 4.40 OLS Joint Direct Effect Estimate for DE_{12}^{OLS} A) *Bias* and B) *RMSE*

The DE_{12}^{OLS} estimate bias is small (±0.01) for estimates (0.30, 0.45) attributable to the OLS regression method. The estimate's standard error and the average estimate's *RMSE* are dependent on the sample size (small *N* having a large *standard error*) as shown in Figure 4.40 B which decreases with increasing pairwise correlations.

There were no conditions where DE_{12}=0 to calculate the type1 error rate for the OLS method, in having a null joint direct effect DE_{12}^{OLS}. The same is true for the individual direct effect for X_1 since the beta parameter values are non-zero $\beta_1 = (0.15, 0.30)$.

However, the type1 error rate was determined for the individual direct effect for X_2 since the beta parameter $\beta_2 = (0,\ 0.30)$ is zero and there were 54 conditions where DE_2=0.

4.2.1.7 OLS Joint Direct Effect –Coverage Probability, Power to Detect DE_{12}^{OLS}

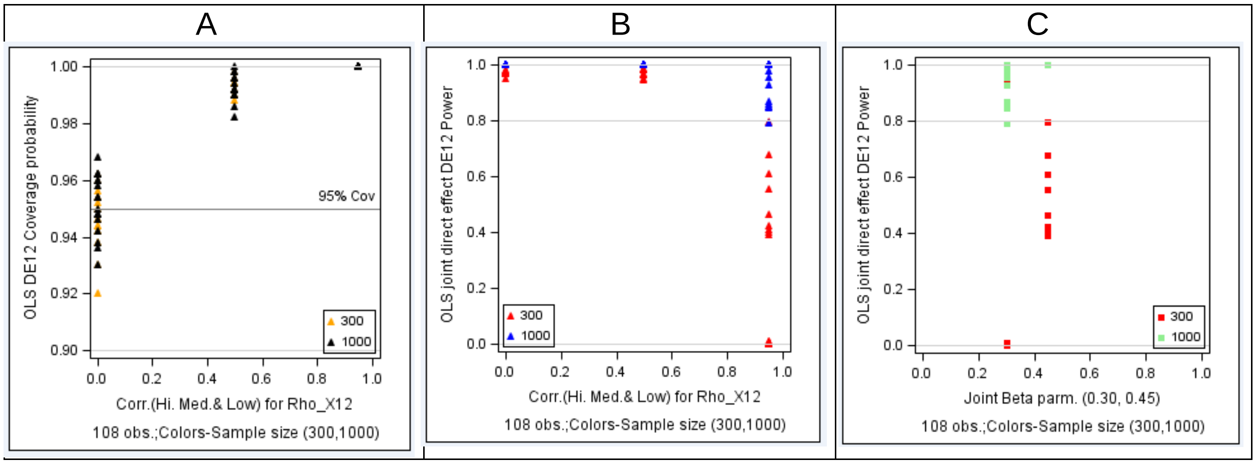


Figure 4.41 Joint Direct Effect Estimate's A) Coverage; B&C) Power to Detect DE_{12}^{OLS}

Figure 4.41 A shows that the OLS joint direct effect's coverage probability for all 108 conditions have a coverage higher than 0.92, which increases with higher pairwise correlations. Figure 4.41 B shows that the OLS method's power to detect a joint direct effect DE_{12}^{OLS} (108 conditions) had reduced power due to a low sample size and a high pairwise correlation $\rho_{12} = 0.95$. Combining the information from Figure 4.41 B and C, the conclusion is that the OLS method has the worst power when the predictors have small effect sizes in small, highly correlated predictor datasets.

4.2.1.8 OLS Individual Direct Effects – *Estimate, Bias & RMSE*

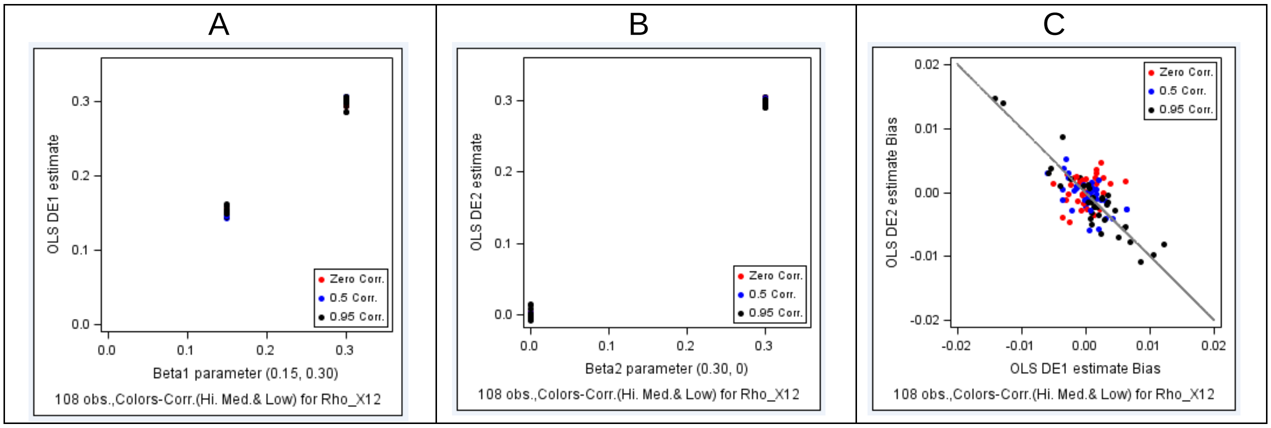


Figure 4.42 OLS Individual Direct Effect Estimates, A&B) DE_1^{OLS}, DE_2^{OLS} C) $DE_1^{Bias}{}_{OLS}$ *vs.* $DE_2^{Bias}{}_{OLS}$

OLS individual direct effects are a direct function of the individual beta parameters and are shown in 4.42 A and B. The joint beta parameter DE_{12} (0.30, 0.45) is composed of DE_1 (0.30, 0.15) and DE_2 (0, 0.30). Figure 4.42 C shows a strong negative correlation between the individual direct effect biases, and that higher pairwise correlations between predictor influences the individual direct effect bias by increasing it to an equal but opposite magnitude for DE_1^{OLS} and DE_2^{OLS} . There is a low bias (±0.015) for each $DE_{1,2}^{OLS}$ estimate value (0, 0.30), leading to the conclusion that the OLS method produces unbiased regression coefficient estimates.

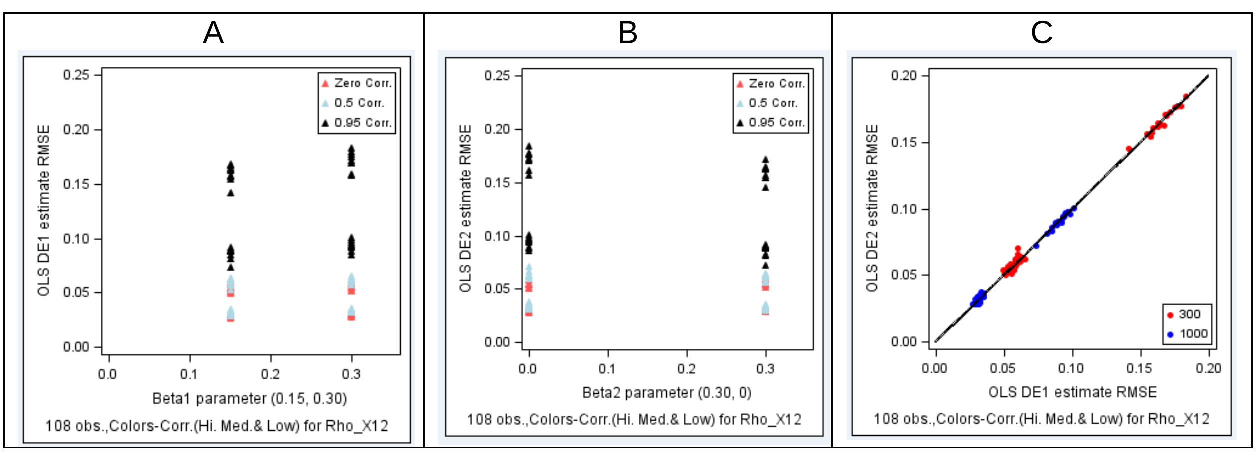


Figure 4.43 OLS Individual Direct Effect Est.'s *RMSE* A&B) $DE_{1,2}^{OLS}$ C) $DE_1^{RMSE}{}_{OLS}$ *vs.* $DE_2^{RMSE}{}_{OLS}$

The estimate's standard error and the average estimate's *RMSE* are dependent on sample size, with smaller sample sizes having larger standard errors as shown in Figure 4.43 C. The effect on an estimate's standard error and RMSE for high pairwise correlations (black triangles) and small datasets (red dots) are large *standard errors* of the estimate averages, as shown in Figure 4.43 A-C. As is the theme for the OLS regression method, the individual direct estimate's *RMSE* is higher for small datasets with high pairwise correlations between the independent variables.

4.2.1.9 OLS Individual Direct Effect Type1 Error, Coverage Probability and Power

The X_2 direct effect's type1 error rate for DE_2^{Typ1} is shown in Figure 4.44. There were zero exceptions for the 54 conditions having a true value for the individual direct effect DE_2=0. The range of the type1 errors was (0.026 to 0.070). Neither sample size nor predictor correlations had any influence on the type1 error rate DE_2^{Typ1}.

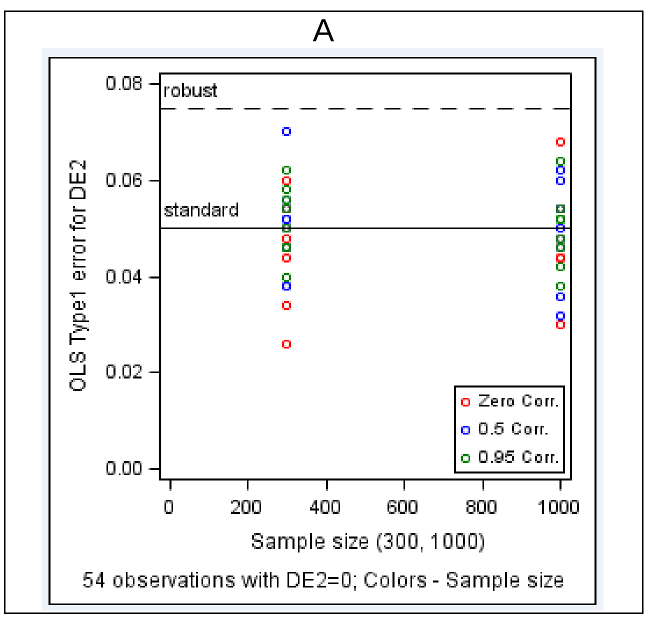


Figure 4.44 OLS method Individual direct effect, X_2 estimate's type1 error DE_2^{Typ1}

The OLS method has coverage probability for individual direct effects $DE_{1,2}^{OLS}$ ranging between 0.92 and 0.97 shown in Figure 4.45 A and D, are centered on the 95% coverage probability for all 108 conditions.

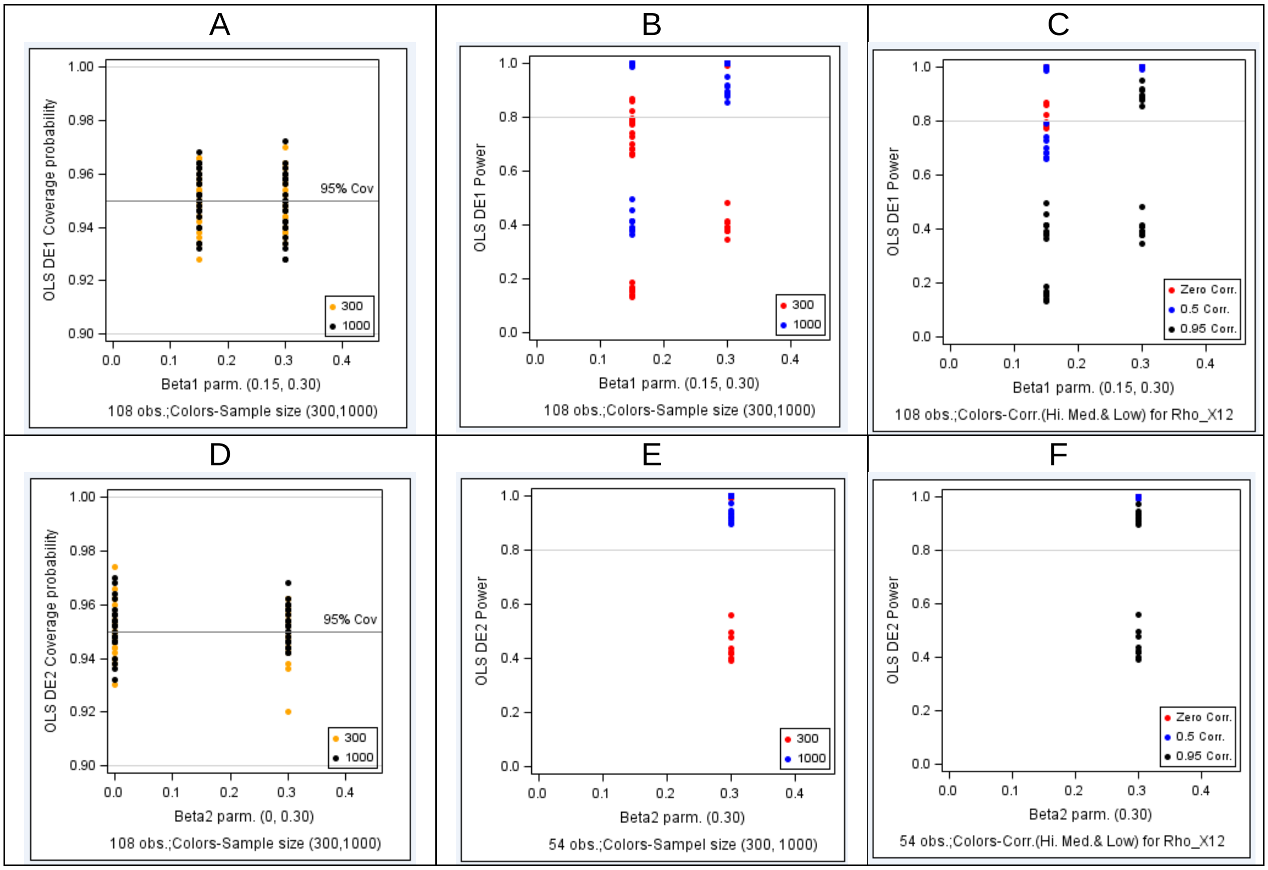


Figure 4.45 Individual Direct Effect's $DE_{1,2}^{\overset{Cov.prob.}{OLS}}$ (A&D), Power (B&C) by *N*, (E&F) by ρ_{12}

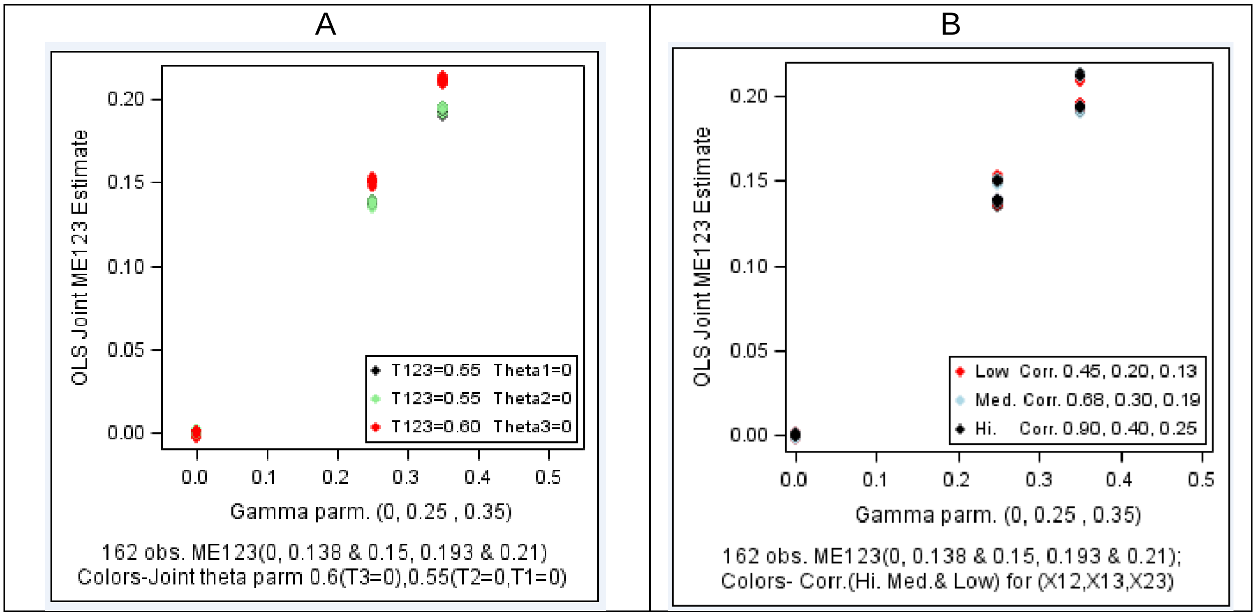
There were 42 exceptions out of 108 conditions for the power falling below 0.8 to detect DE_1^{OLS} (parm. values=0.015, 0.30) shown in Figure 4.45 B and C, and 9 exceptions out of 54 conditions for the power falling below 0.8 to detect DE_2^{OLS} (parm. value=0.30) shown in Figure 4.45 E and F. There were fewer exceptions for DE_2^{OLS} than for DE_1^{OLS} because of the differences in effect size being detected. Given the same direct effect size both $DE_{1,2}^{OLS}$ had 9 exceptions out of 54 conditions, with the power

ranging from 0.35 to 0.56. The power to detect an individual direct effect $DE_{1,2}^{OLS}$ with small sample sizes and high correlations ($N = 300, \rho_{12} = 0.95$) combined with small effect sizes ($DE_1 = 0.15$) reduces the power drastically to a range of 0.13 to 0.16, but with an increased direct effect size ($DE_1 = 0.30$) the power increases to a range of 0.35 to 0.38. However, under similar small effect size conditions but with a larger sample size, the increased power ranges from 0.37 to 0.41, and with a zero correlation the power increases further to a range from 0.77 to 0.79. The large standard error of the estimate for DE_1^{OLS} (*RMSE*=0.16) for detecting a direct effect ($DE_1 = 0.15$), results in a wide confidence interval which includes 0 in 84%of the data replications and a power of 0.16.

These results lead to the conclusion that the OLS method does not handle highly correlated independent variables in multiple regression very well, but produces on average over a large number of replicated datasets, unbiased estimates for its regression coefficients. Similar conclusions were reached earlier with regards the individual indirect effects shown in Figure 4.38 A-F, in their power to detect $ME_{1,2}^{OLS}$.

4.2.2 OLS Method 3-Variable Mediation Analysis

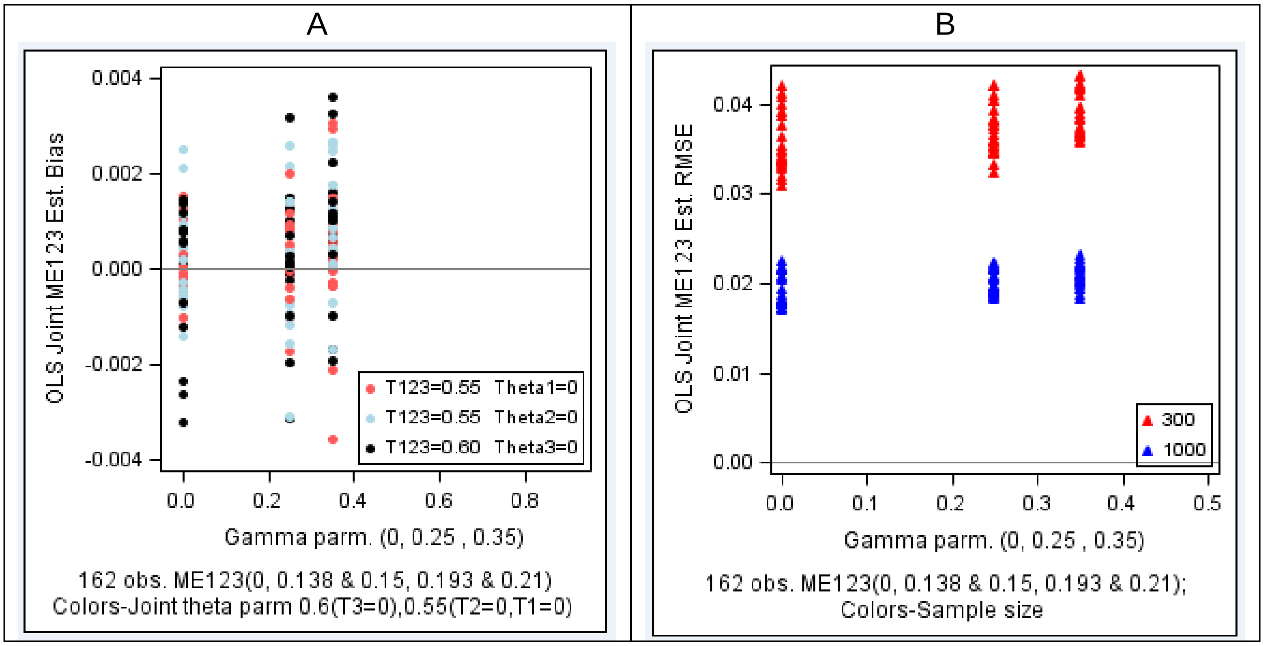
4.2.2.1 OLS Joint Mediated Effect - *Estimate, Bias* & *RMSE*



Figures 4.46 OLS Joint Indirect Effect Estimate for ME_{123}^{OLS} A&B Joint Mediated Effect

Figures 4.46 A and B show the influencing variables that determine the joint mediated effect for 3-variable mediation using the Ordinary Least Squares method. The influencing variables are: the gamma parameter $\gamma(0, 0.25, 0.35)$, and the joint theta parameter $\theta_{123}(0.60, 0.55, 0.55)$, because values for γ, θ_{123} determine the joint indirect effect ME_{123}^{OLS}. The pairwise correlation values do not seem to influence the estimate's value as shown in Figures 4.46 B.

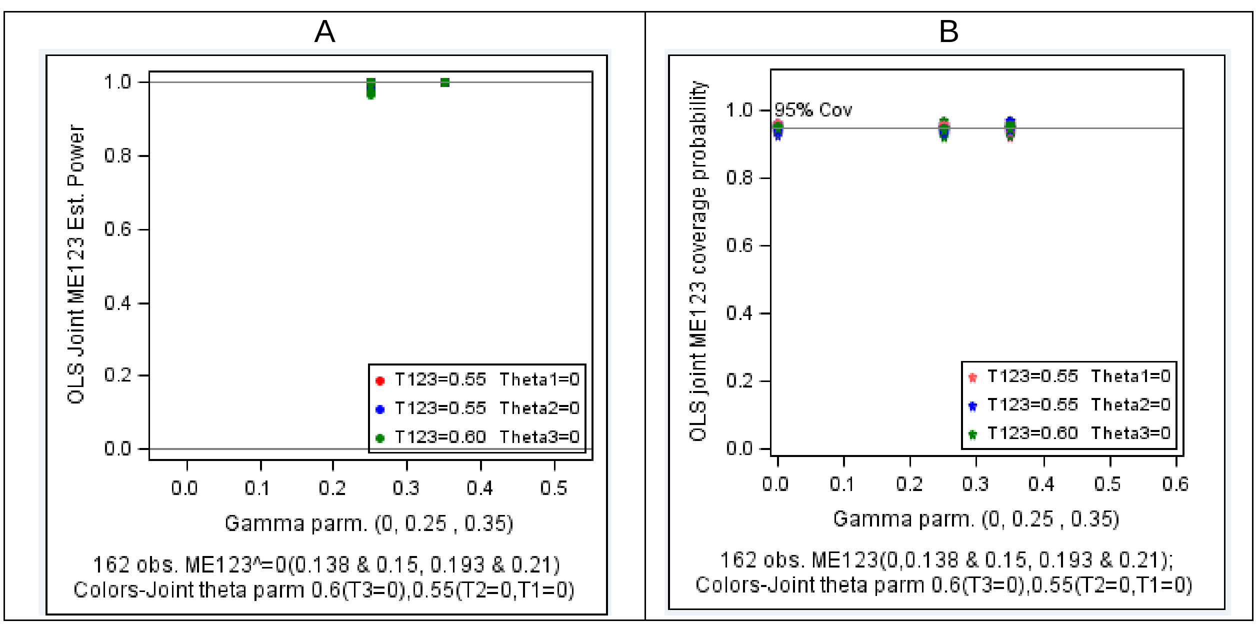
The joint mediated effect's estimate bias and *RMSE* are shown in Figures 4.47 A and B. Panel A shows that the OLS method provides an unbiased estimate (±0.004) for the estimates true value with the range 0 to 0.21. Neither the joint theta parameter nor the gamma parameter values affect the estimate bias, as they did with the OLS estimate value shown in Figure 4.46 A. Panel B shows the estimate's high *RMSE* (range 0.02 to 0.04), which is influenced by the low sample size for a given gamma.



Figures 4.47 OLS Joint Indirect Effect Estimate for ME_{123}^{OLS} A) Bias and B) RMSE

The high standard error of the joint mediated effects (ranging from 0 to 0.21) makes the unbiased estimate unreliable, because of such high variability.

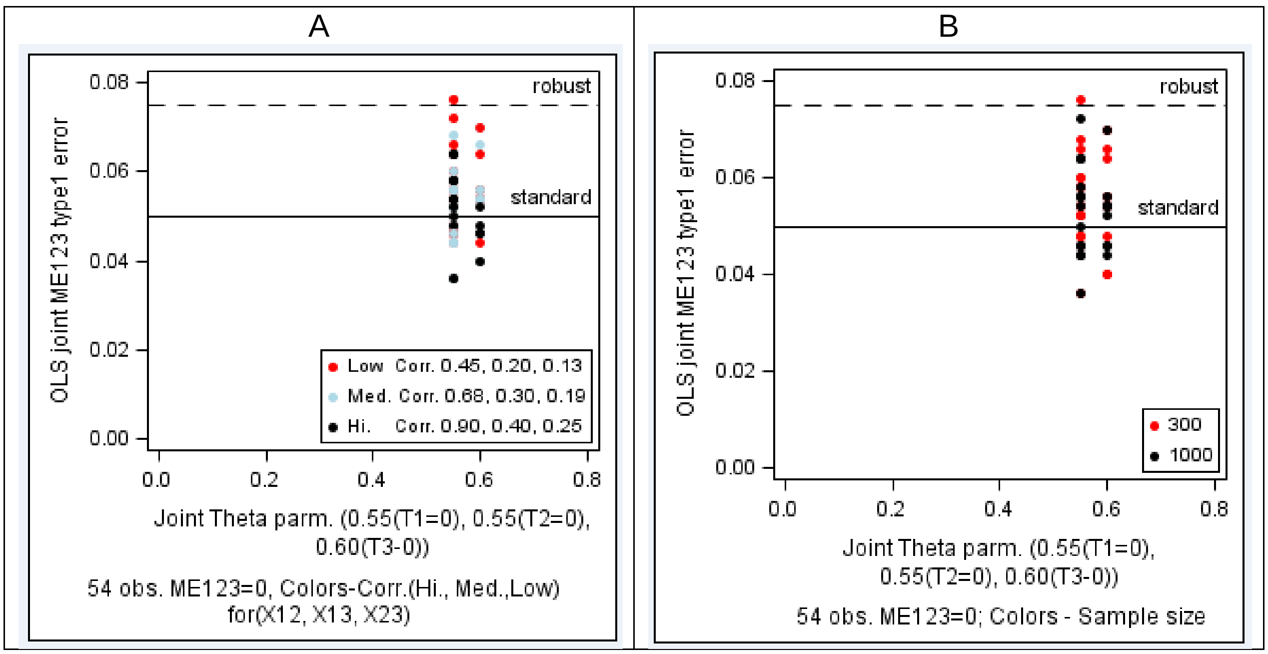
4.2.2.2 OLS Joint indirect effect for ME_{123}^{OLS} Power, Coverage prob. & Type 1 error



Figures 4.48 OLS Joint Indirect Effect Est.'s for ME_{123}^{OLS} A) Power; B) Cov. Probability

The OLS method's power to detect the joint mediated effect ME_{123}^{OLS} (range 0.13 to 0.21), and its coverage probability are shown in Figures 8.3A&B. The power to detect ME_{123}^{OLS} is almost 1 for higher values of the joint mediated effect and above 0.97 for the rest of the 108 conditions where ME123 ≠0. The coverage probability for all 162 conditions was at 95% as shown in Figure 4.48B.

The OLS method's type1 error for estimating a null joint mediated effect is shown below.



Figures 4.49 A-B OLS method's joint indirect effect type1 error for ME_{123}^{OLS}

The OLS joint mediated effect's type1 error rate with the joint theta parameter $\theta_{123}(0.60, 0.55, 0.55)$ on the x-axis with color coded marks showing increasing pairwise correlations (low(red), medium(blue), high(black)) are presented in Figures 8.4 A&B. There was one exception amongst the possible 54 conditions where $ME_{123}^{OLS} = 0$, having a type1 error (0.076) greater than the *a priori* set limit of 0.075. As shown in in the plots, the sample size was small and the pairwise correlations between $X_{12} X_{13}$ X_{23} were low.

Overall, the OLS method utilizes the type1 error limit well, and has a high power for detecting the joint indirect effect, with high coverage probability. However, the standard error for the estimates are high since it does not handle small sample sizes with large pairwise correlations effectively.

The individual mediated effects estimates, bias, standard error and coverage probability, type1 error and power for predictors (X_1, X_2, X_3) using the OLS method for mediation analysis are discussed next.

4.2.2.3 Individual Mediated effects - OLS *Estimates, Bias* & *RMSE*

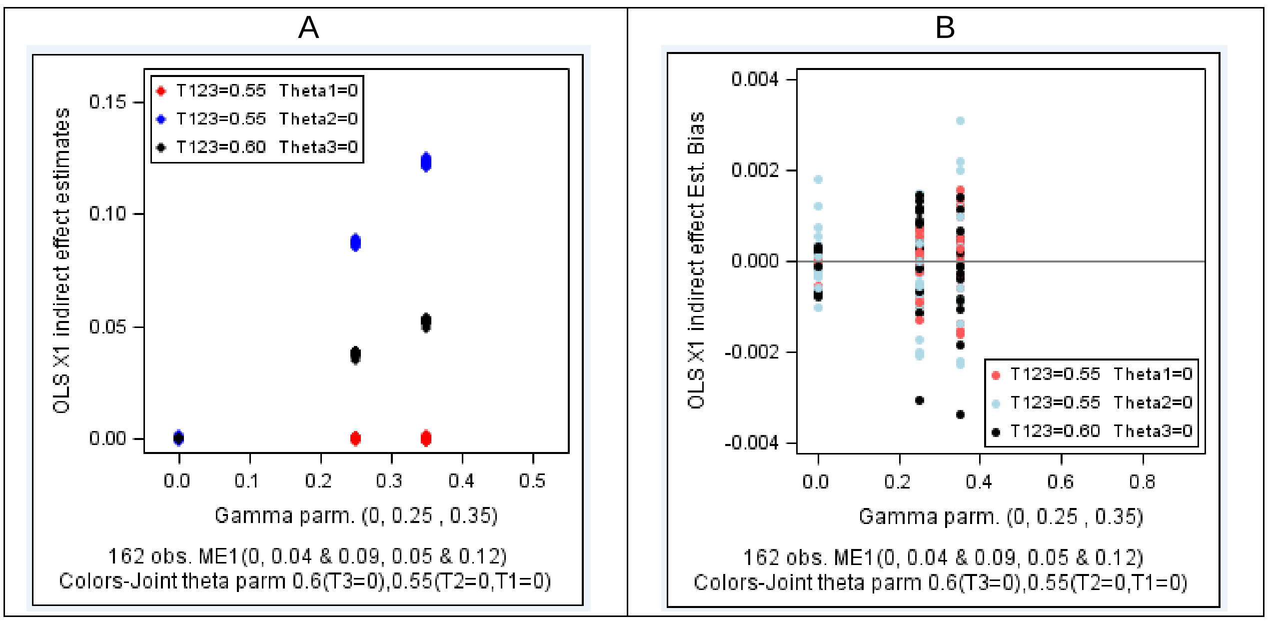
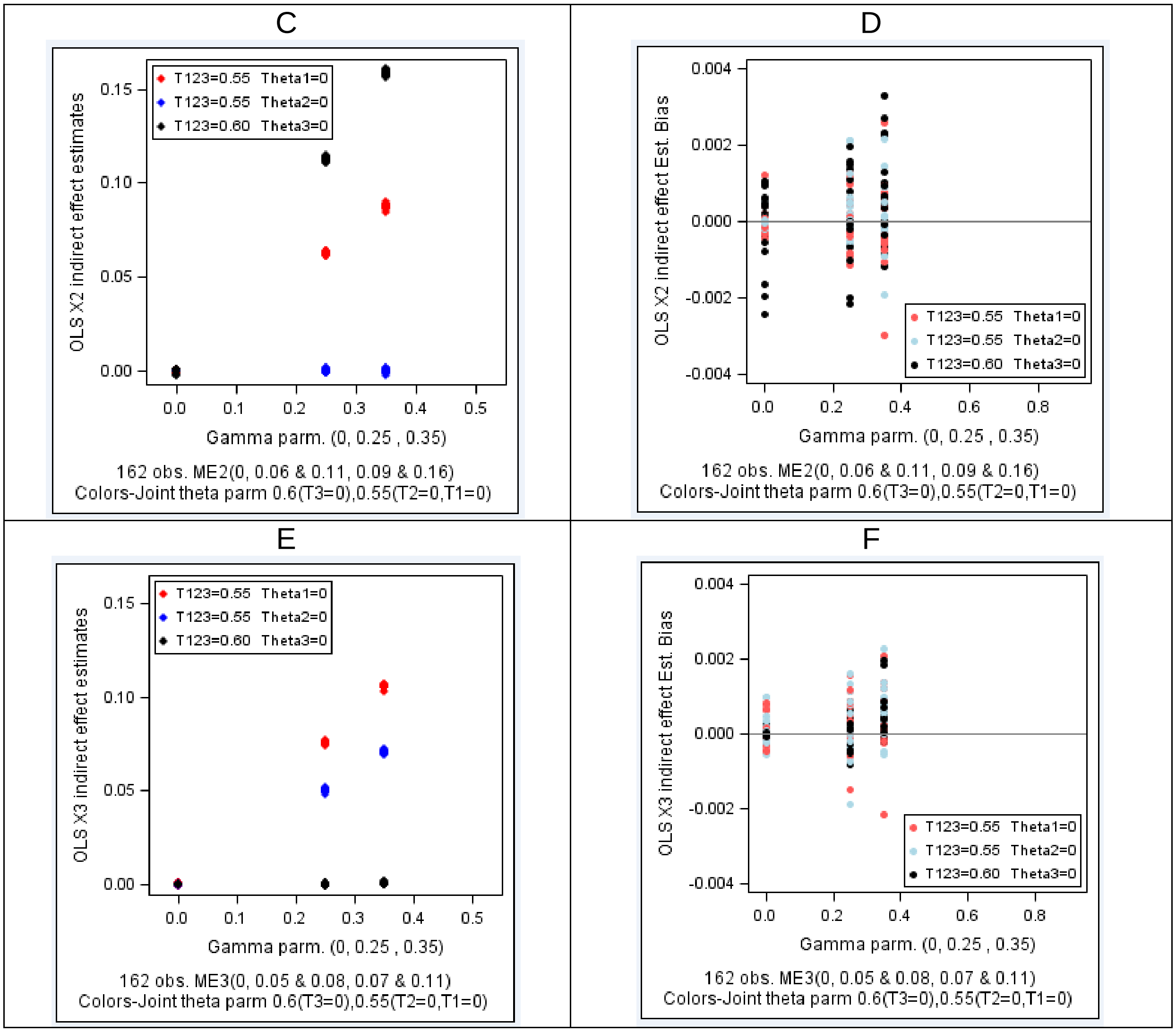


Figure 4.50 OLS Individual Indirect Effect Est's A) Bias and B) RMSE for ME_1^{OLS}

The influencing variables that determine *X₁*'s mediated effect using the OLS method are the gamma parameter $\gamma(0, 0.25, 0.35)$, and the joint theta parameter $\theta_{123}(0.55, 0.55, 0.60)$. Higher values for gamma increase the indirect effect ME_1^{OLS} for all nonzero theta parameter values. The pairwise correlations of (high, medium, low) for

$X_{12}X_{13}$ X_{23} do not markedly influence the estimate, since no other groupings are readily visible in Figure 4.50 A.

OLS Individual Predictor Estimates and Biases for ME_2^{OLS} and ME_3^{OLS} are discussed next in Figure 4.51 C-F.

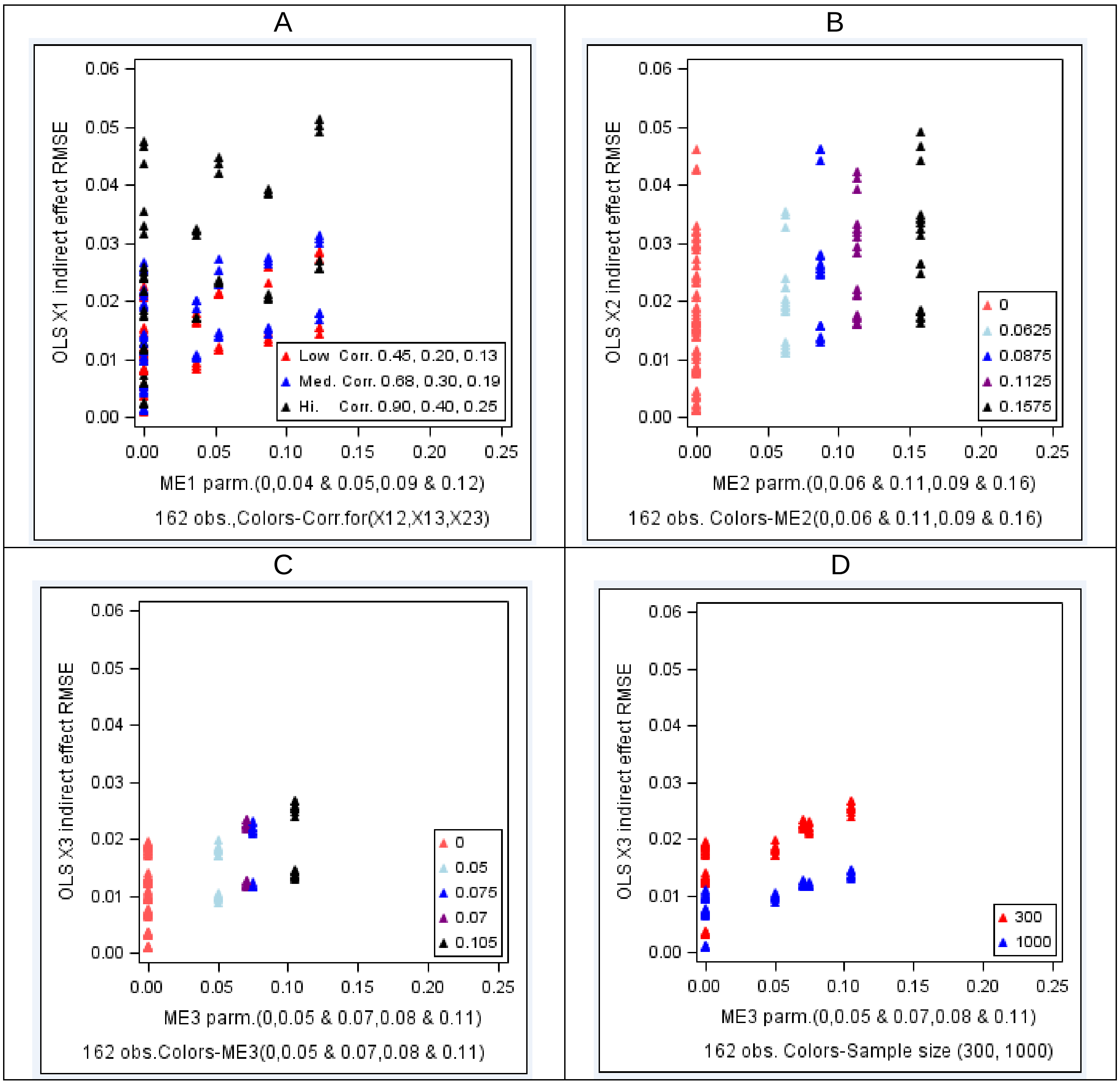


Figures 4.51 OLS Individual X_1, X_2, X_3 $ME_{1,2,3}^{OLS}$ A, C, E- Estimates & B, D, F- Est. Bias

The individual mediated effect's Estimate and *Bias* for X_2 and X_3 are shown in Figure 4.51 C - F. Panel C & E are comparable plots for ME_2 and ME_3 estimates. The corresponding plots D & F show that the OLS method provides an unbiased estimate

(bias range: ±0.004) for the estimates true value (range: 0 to 0.16). Neither the joint theta nor the gamma parameter values affect the estimate's bias since these is little dispersion amongst the estimate values in Figure 4.50 A and Figure 4.51 C and E.

The standard errors for the estimates in Figures 4.50 and 4.51are discussed next.

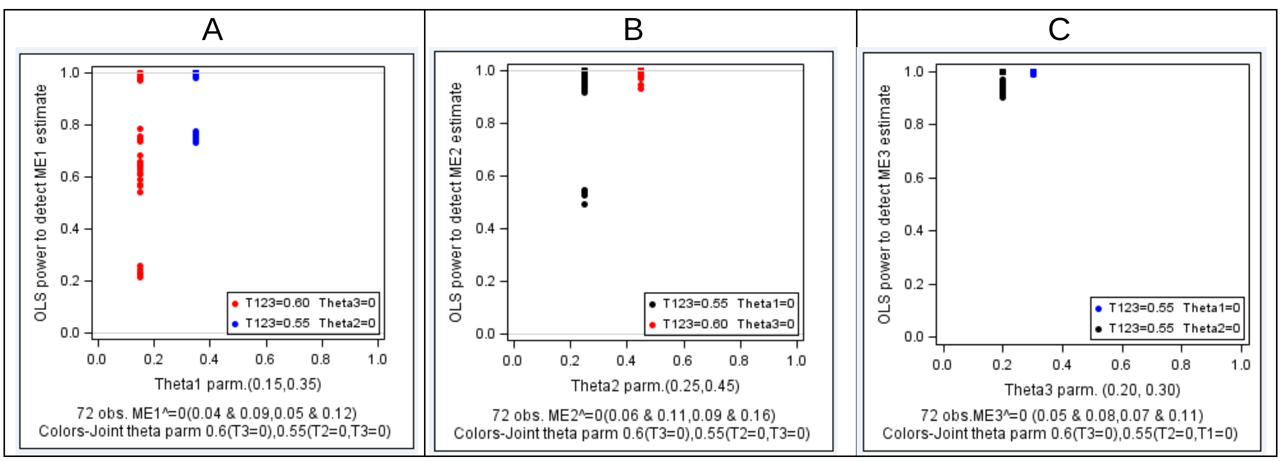


Figures 4.52 OLS Indiv. Indirect Effect's $ME^{\substack{RMSE\\OLS}}_{1,\ 2,3}$ A-D) grouped by *Corr.*, $ME_{1,\ 2,\ 3}$ and *N*

The *RMSE* for the average indirect effect estimates shown in Figure 4.40 A and Figure 4.51 C and E have standard errors that are shown in Figure 4.52 A-D. The magnitude of the *RMSE* is large as compared to the values of the individual mediated estimate, which suggests wide variability in OLS individual mediated effect estimates. A mediated estimate from a single dataset is of no practical value because of the high variability of the OLS estimate. The factors influencing the estimate's *RMSE* are the correlation between the individual predictors (panel A), proportionality to the estimate's value (panels A-C), and sample size (panel D).

4.2.2.4 Individual Mediated effects – Power, Coverage prob. and Type 1 error

The OLS method's power to detect an individual predictor's mediated effect is discussed below.



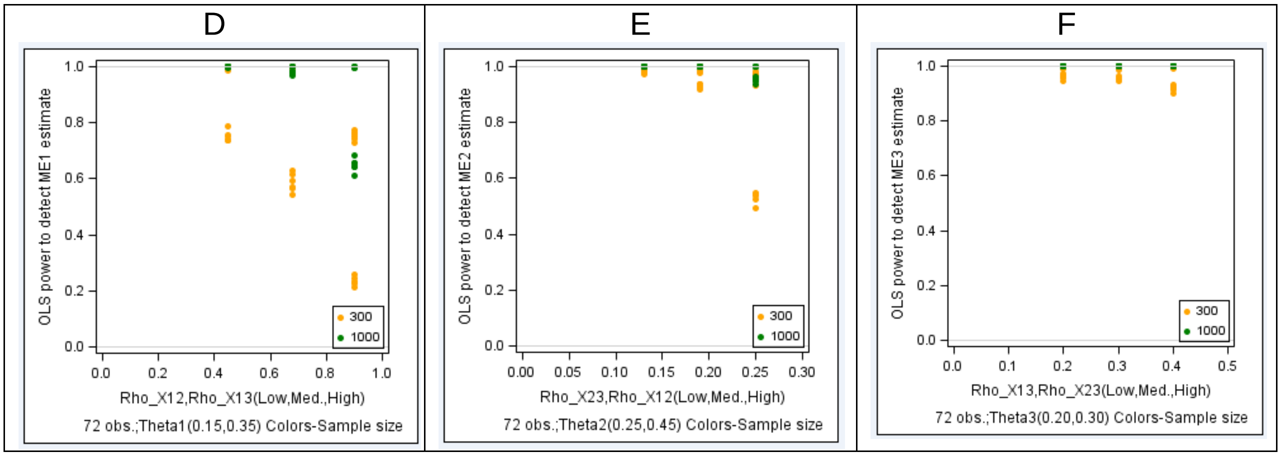
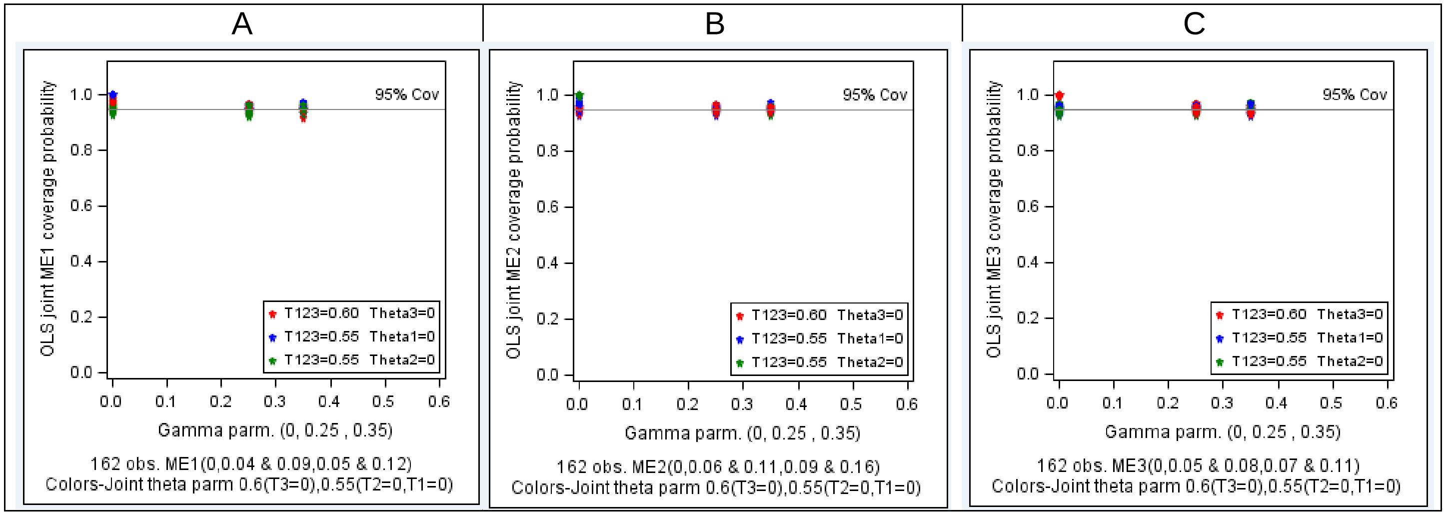


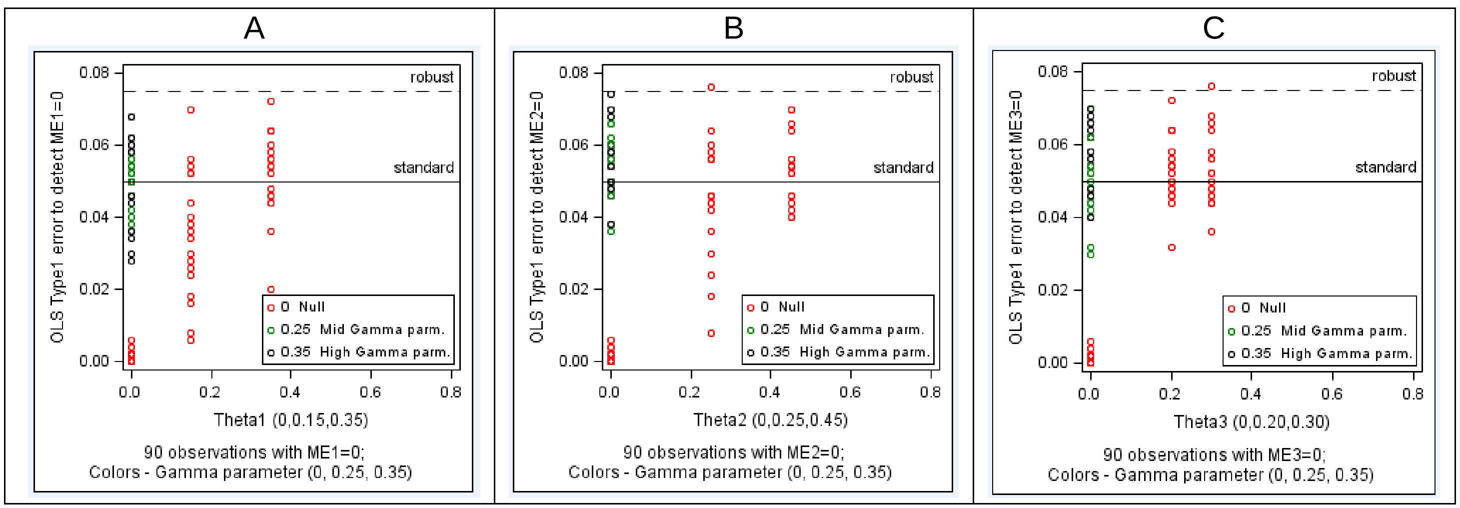
Figure 4.53 OLS Indiv. Indirect Effect $ME^{\substack{Power\\OLS}}_{1,\ 2,3}$ A, B, C are grouped by *Corr.* & D, E, F by *N*

The power for OLS to detect an individual mediated effect is influenced by three parameter values listed in decreasing order of influence, pairwise correlations (high correlations result in lower power, Figure 4.53 D, E and F, next by the sample size (small sample sizes result in lower power, shown in Figure 4.53 D, E and F, and by the individual mediated effect sizes, i.e. smaller individual theta values with lower individual mediated effects having a lower power, shown in A, B & C). A combination of high correlation, low sample size, and a small effect size, results in the lowest power to detect an individual effect using the OLS method, e.g. ME_1^{OLS} with power=0.21, shown in panel A (lowest red dot with $ME_1^{OLS} = 0.0375$) , high correlation $\rho_{12}(0.90, 0.68, 0.45)$ and low sample size (*N*=300, gold dot) shown in panel D. The power increases with reduced pairwise correlations and increased sample sizes. Figure 4.53 A-F support the inference that the OLS method performs poorly when detecting small individual mediated effects in small sample datasets, having high predictor pairwise correlations.

The individual predictor mediated effect's coverage probability using the OLS method was approximately 95% for all conditions and are shown in Figure 4.54 A-C.



Figures 4.54 OLS Individual Mediated Effect Estimate $ME_{1,2,3}^{OLS}$ A - C Coverage Probability



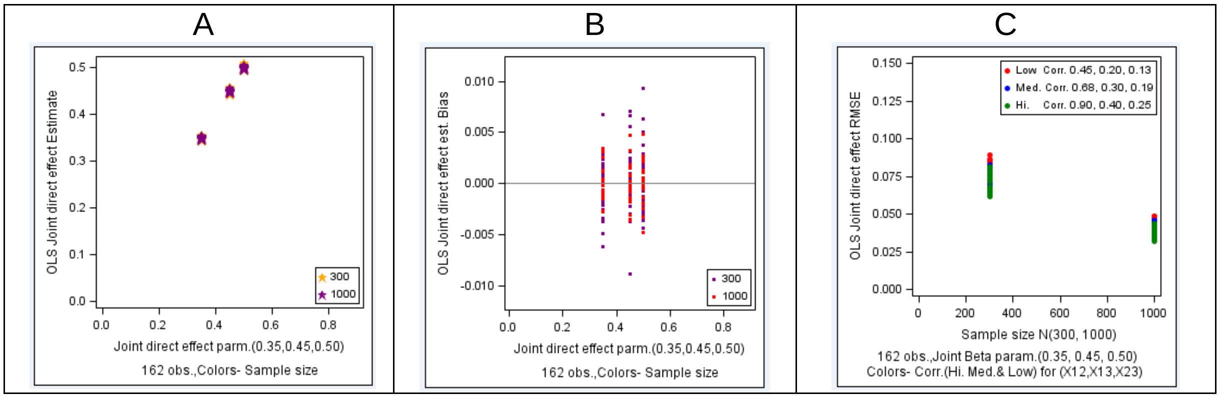
Figures 4.55 A-C OLS Individual Mediated Effect $ME_{1,2,3}^{OLS}$ Type1 Error grouped by θ_1, θ_2, and θ_3

The OLS individual mediated effect's type1 error in 3-variable mediation is shown in Figure 4.55 A, B and C. There were 90 conditions where $ME_1 = 0$ of which 36 conditions had non-zero gamma parameter values $\gamma = (0.25, 0.35), \theta_1 = 0$, and of these 36 conditions 24 conditions had $\gamma = 0.25$ (smaller gamma parameter value). The remaining 54 conditions had $\gamma = 0,\ \theta_1 = (0,\ 0.25, 0.35)$ with 18 conditions for each of the three theta values. The Figure 4.55 A shows that 0 of the 90 conditions for $ME_1 = 0$ had a type1 error that exceeded the limit of 0.075. There were 90 conditions where $ME_2 = 0$ for type1 error consideration. One exception was noted and shown in Figure 4.55 B. The highest type1 error for ME_2^{OLS} = 0.076 was for *N*=300, $\gamma = 0, \theta_2 = 0.25, \gamma_{Bias}^{OLS} = 0.002, ME_2^{RMSE} = 0.016$ (small *RMSE*) which resulted in the single exception. There were 90 conditions where $ME_3 = 0$ for type1 error consideration. One exception was noted and shown in Figure 4.55 C. The highest type1 error for ME_3^{OLS} = 0.076 was for *N*=300, $\gamma = 0, \theta_3 = 0.30, \gamma_{Bias}^{OLS} = 0.002, ME_3^{RMSE} = 0.019$ (larger RMSE since $\theta_3 > \theta_2$) which results in the single exception. All other parameters remaining the same, a larger sample size reduces the standard error of the estimate which results in a narrower confidence interval around the estimate and an increased type1 error.

4.2.2.5 OLS 3-Variable Joint Direct Effects Estimate, Bias and *RMSE*

The OLS method estimates the joint direct effect having a small bias (-0.01, 0.01) with a standard error for the estimate of 0.04 (large *N*) to 0.08 (small *N*), given the true

value of the joint direct effect (0.35, 0.45, 0.50). Figure 4.56 C shows that larger sample sizes and higher correlations reduce the joint direct estimate's *RMSE*.



Figures 4.56 A-C OLS Method's Joint Direct Effect Estimate, Bias and *RMSE*

4.2.2.6 OLS 3-Variable Joint Direct Effect's Power and Coverage Probability

There were 12 exceptions to the OLS method being able to detect a joint direct effect with a power greater than 0.8 and the common influencing variables were: 1) small sample size (shown in panel B), high predictor correlations and low joint beta parameter values (9 with $\beta_{123} = 0.35(\text{black})$ and 3 with $\beta_{123} = 0.45(\text{red})$ shown in panel A. The joint direct effect coverage probability for all 162 conditions ranged 0.97 to 1.

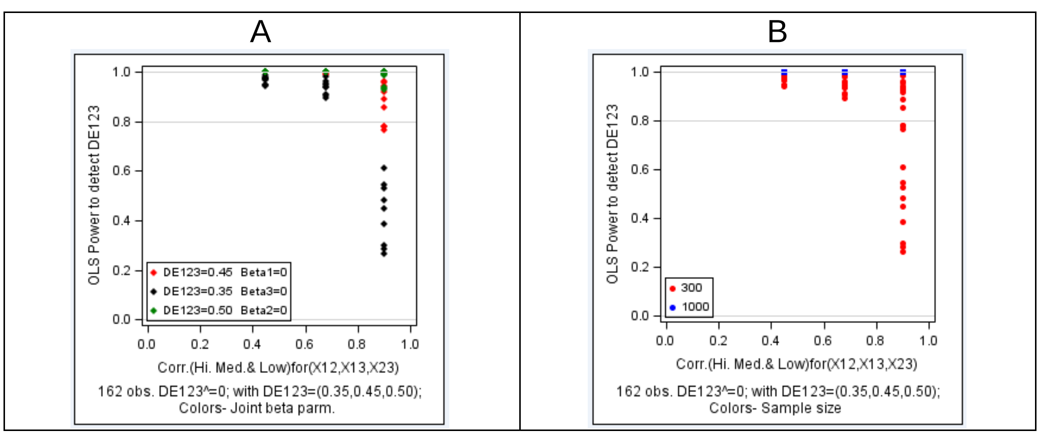
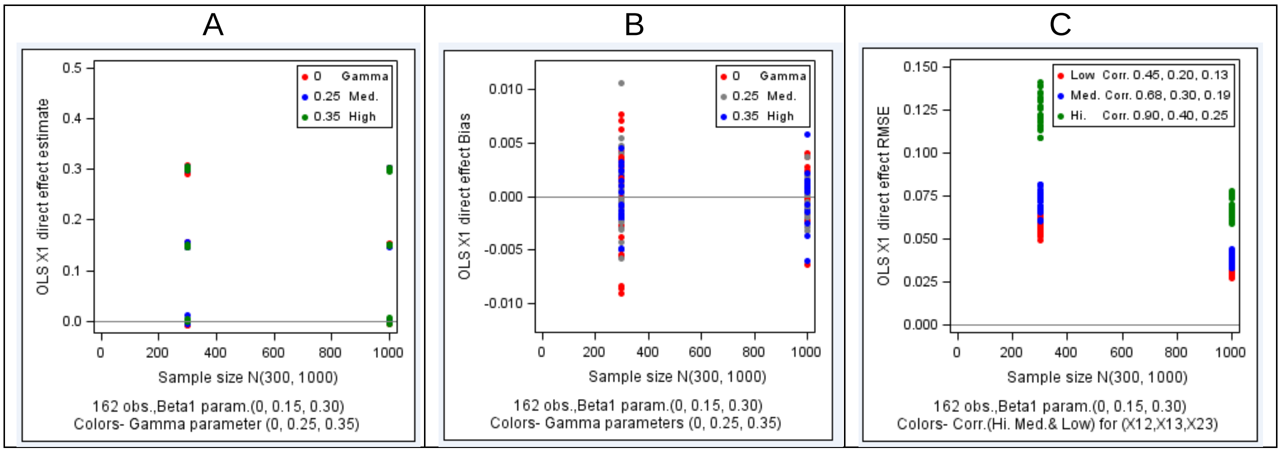


Figure 4.57 OLS Method's Joint Direct Effect Power for DE_{123}^{OLS} A & B) by *Corr.*

There were no conditions where DE_{123}=0 to assess the joint type 1 error.

4.2.2.7 OLS 3-variable Individual direct effects *Estimate*, *Bias* and *RMSE*

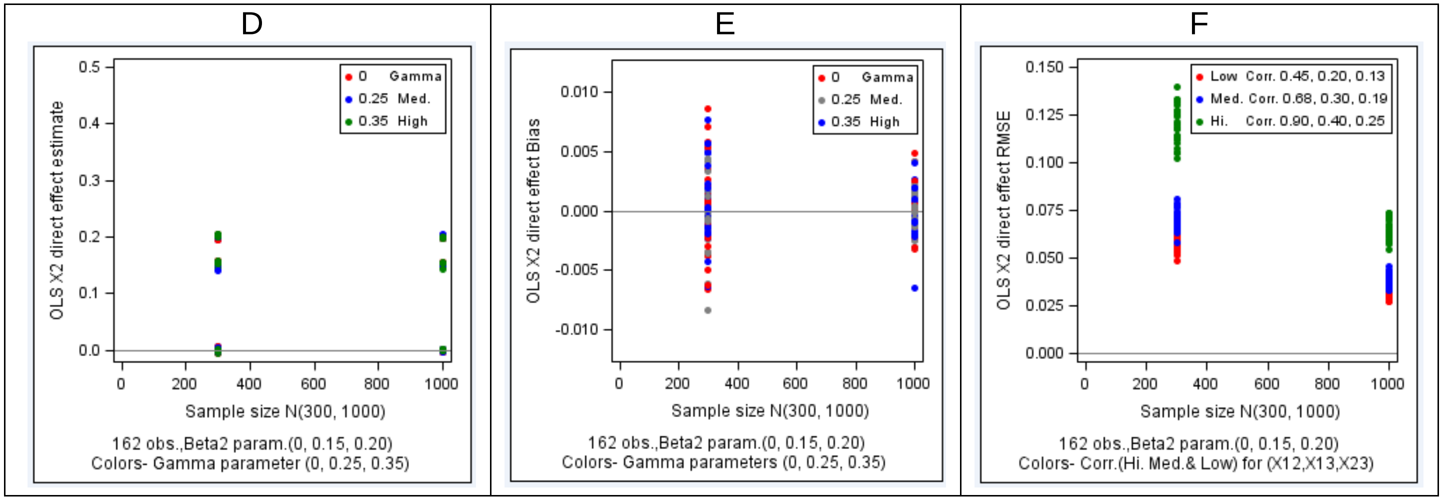
The influencing variable that determine X_1's direct effect using the OLS method is the beta parameter associated with the independent variable. Higher values for gamma and sample size reduce the variability for the estimated direct effect (Figure 4.58 B).



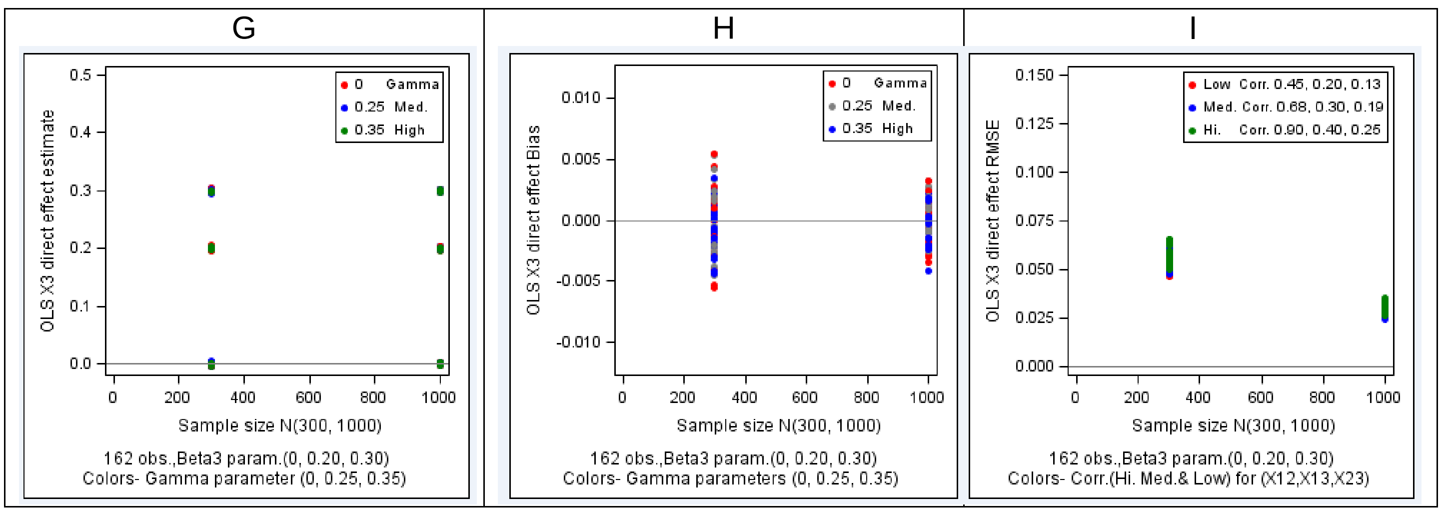
Figures 4.58 OLS Method's Individual Direct Effect for DE_1^{OLS} A) *Est.* B) *Bias* & C) *RMSE*

Neither sample size nor the pairwise correlations of (high, medium, low) for (*X_{12}, X_{13}, and X_{23}*), markedly influence the estimate's value. OLS estimates have high standard errors for the estimate when correlations are high and the sample sizes are small.

The individual direct effect's *Estimate*, *Bias* and *RMSE* are shown in Figure 4.58 A - I. Panels A, D, and G are comparable plots that display the individual direct effect estimates. The corresponding plots B, E & H show that the OLS method provides an unbiased estimate (bias range: -0.009 to 0.009) for the estimates true value (range: 0 to 0.30). The individual beta parameter values do not seem to affect the estimate's bias. Contrary to the joint direct effect's *RMSE* (Figure 4.56 C), the individual *RMSE* increases with higher pairwise correlation (Figure 4.58 C, F, & I). Larger sample sizes consistently reduce the standard error of the estimates.



Figures 4.58 OLS Method's Individual Direct Effect for DE_2^{OLS} D) *Est.* E) *Bias* & F) *RMSE*



Figures 4.58 Individual $DE_{1,2,3}^{OLS}$ A, D, & G-*Estimates*; B, E, & H-*Bias*; and C, F, & I-*RMSE*

4.2.2.8 OLS Individual 3-Variable Direct Effect's Type1 Error, Power & Coverage

The OLS method's type1 error for estimating a null joint direct effect could not be evaluated since $\beta_{123} = (0.35,\ 0.45,\ 0.50)$ and $DE_{123} \neq 0$ for the 162 conditions. The individual predictor direct effect's type1 errors, followed by the power to detect the joint and individual direct effects are discussed below.

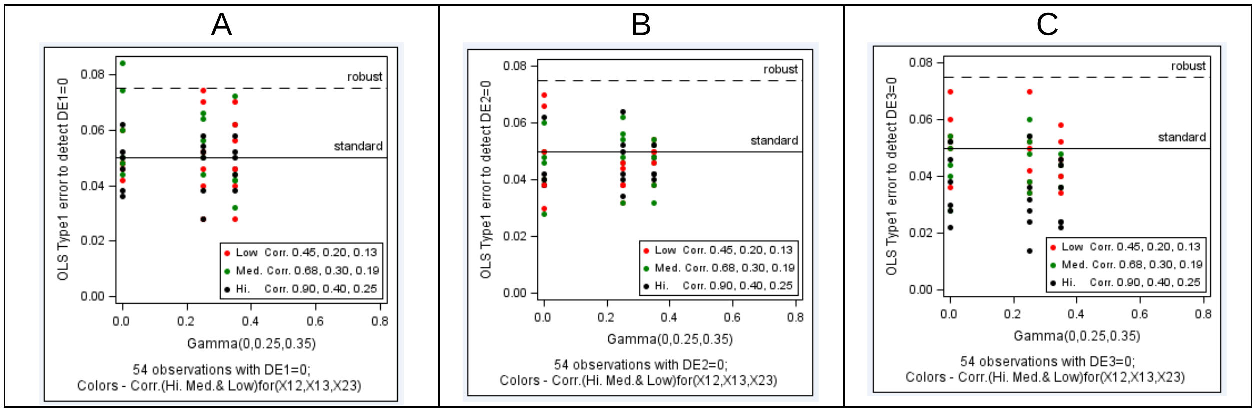


Figure 4.59 OLS Method's Individual Direct Effect Type1 Error A), B) & C) for $DE^{OLS}_{1,2,3}$

The OLS individual direct effect's type1 error rate with the gamma parameter values (0, 0.25, 0.35), color coded by predictor pairwise correlations, 54 conditions for each individual direct effect, totaling 162 conditions and shown in Figure 4.59 A, B and C. There was one exception amongst the 54 conditions where $DE^{OLS}_{1} = 0$, with type1 error= 0.084 which was greater than the *a priori* set type 1 error limit of 0.075. The joint beta parameters for the three plots A, B & C were $\beta_{123}(0.45(\beta_1 = 0), 0.50(\beta_2 = 0),$ $0.35(\beta_3 = 0))$. There were zero exceptions for the other two sets of 54 conditions shown in plots B & C where DE_2=0 and DE_3=0.

The OLS method's power to detect each individual predictor's direct effects is shown in Figure 4.60.

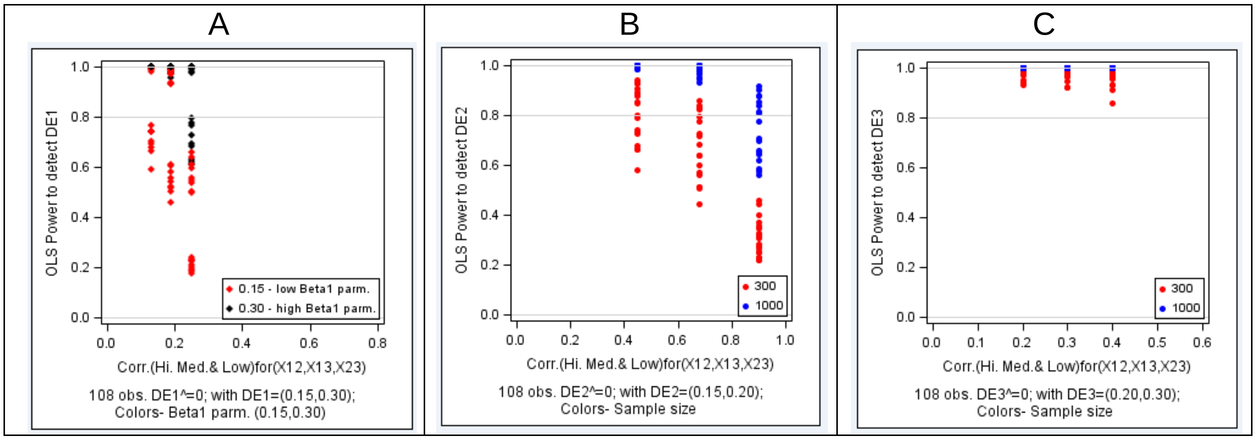
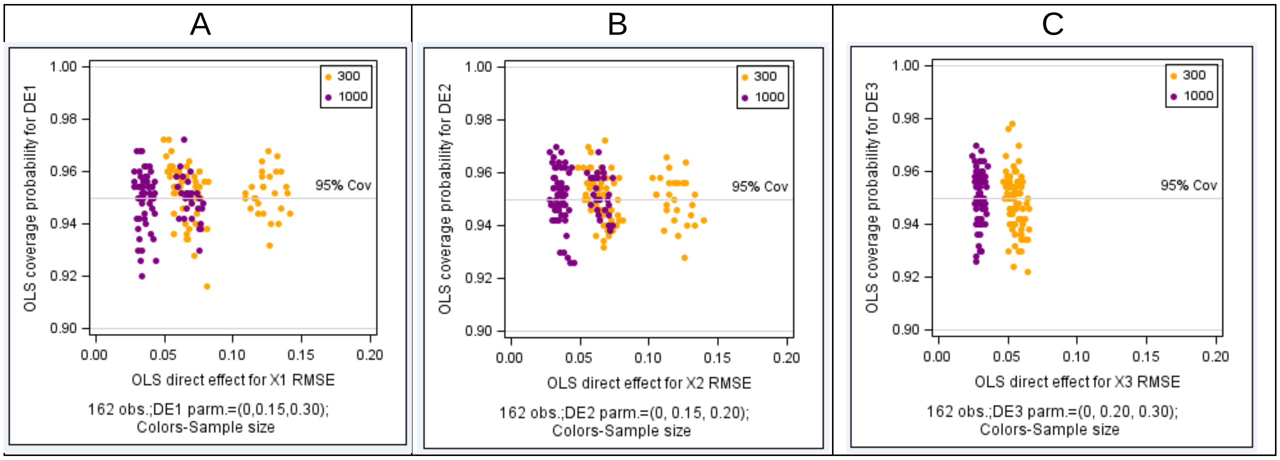


Figure 4.60 OLS Method's Individual Direct Effect Power A), B) & C) for $DE_{1,2,3}^{OLS}$

The factors influencing the power to detect an individual direct effect as inferred from plots A-C, are that high predictor correlations, small sample sizes, and small individual beta parameter values reduce power. The two clusters for a given sample size and pairwise correlation are associated with the two beta parameter values $\beta_1 = (0.15,\ 0.30)$, and shown in Figure 4.60 A as two clusters with an average power of 0.21 vs. 0.70 respectively. An overall assessment of the OLS method's power to detect an individual mediated direct effect is that it is poor for small sample sizes, high predictor correlations, and small effect sizes. The joint direct effect's coverage probability for all 162 conditions exceeded 0.974 resulting in no exceptions.

4.2.2.9 OLS 3-Variable Individual Direct Effect's *Coverage probability* vs. *RMSE*

The lowest mediated direct effect's coverage probability $DE_1^{OLS} = 0.92$ is shown in 4.61 A with a small sample size, correlations between (X_1, X_2, X_3) $\rho_{12} = 0.68, \rho_{13} = 0.20$. The OLS method provides unbiased joint and individual direct estimates having large *standard errors*, for highly correlated predictors in small datasets, but with low power.



Figures 4.61 OLS Individual Direct Effect $DE_{1,2,3}^{OLS}$ A-C) $DE_{1,2,3}^{COV}$ vs. $DE_{1,2,3}^{RMSE}$ by sample size *N*

4.3 LASSO 2 & 3 Variable Summary Statistical Analysis

4.3.1.1 Joint Mediated Effect – LASSO Regression *Estimate, Bias* & *RMSE*

LASSO joint mediated effects increase with increasing joint theta and gamma parameter values as shown in Figure 4.62 A and B.

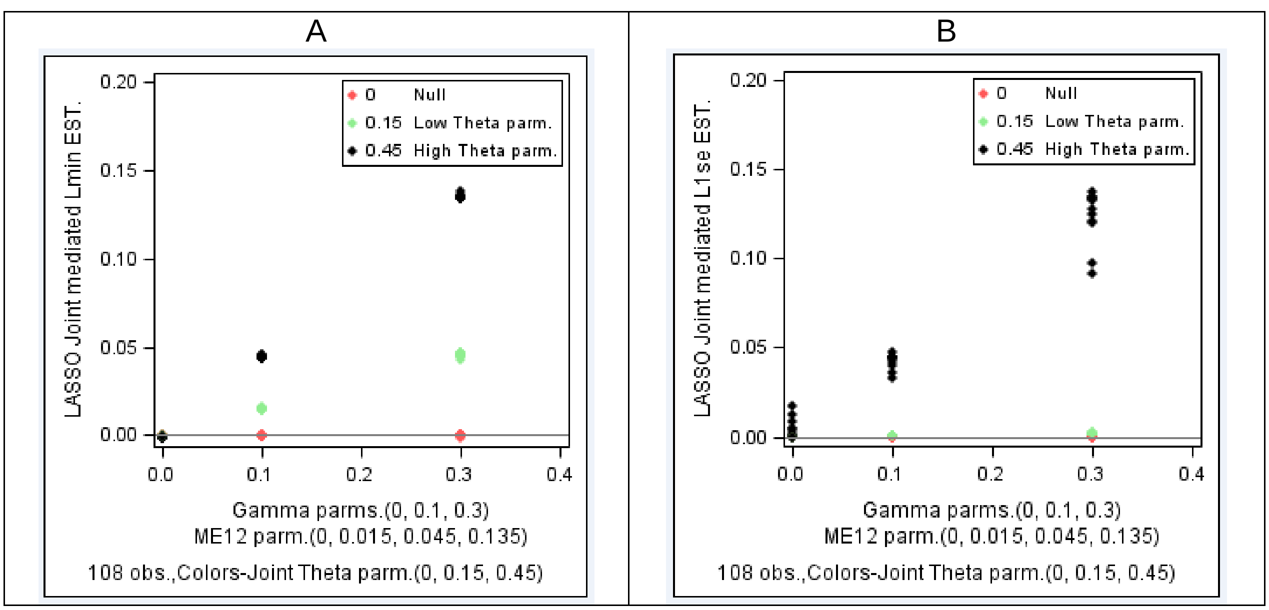


Figure 4.62 LASSO Joint Indirect Effect *Estimate* for Methods $ME_{12}^{L_{\min}}$ and $ME_{12}^{L_{\min+1s.e.}}$

LASSO joint mediated effects were obtained in two different ways; first by selecting regression coefficients with the minimum MSE (Figure 4.62 A) and next by selecting the minimum+1*s.e.* (Figure 4.62 B) regression coefficients in the n-fold cross-validation. When the joint theta value is zero $\theta_{12} = 0$, both methods $ME_{12}^{L_{\min}}, ME_{12}^{L_{\min+1se}}$, do not admit either X_1 or X_2, but when $\theta_1 = 0.15, \theta_2 = 0$, $L_{min+1se}$ does not admit X_1 until the sample size is large and until the correlation is large $\rho_{12} = 0.95$. The bias for the $ME_{12}^{L_{\min+1se}}$ estimate is larger than that for $ME_{12}^{L_{\min}}$ as shown in Figures 4.62 B & 4.62 A.

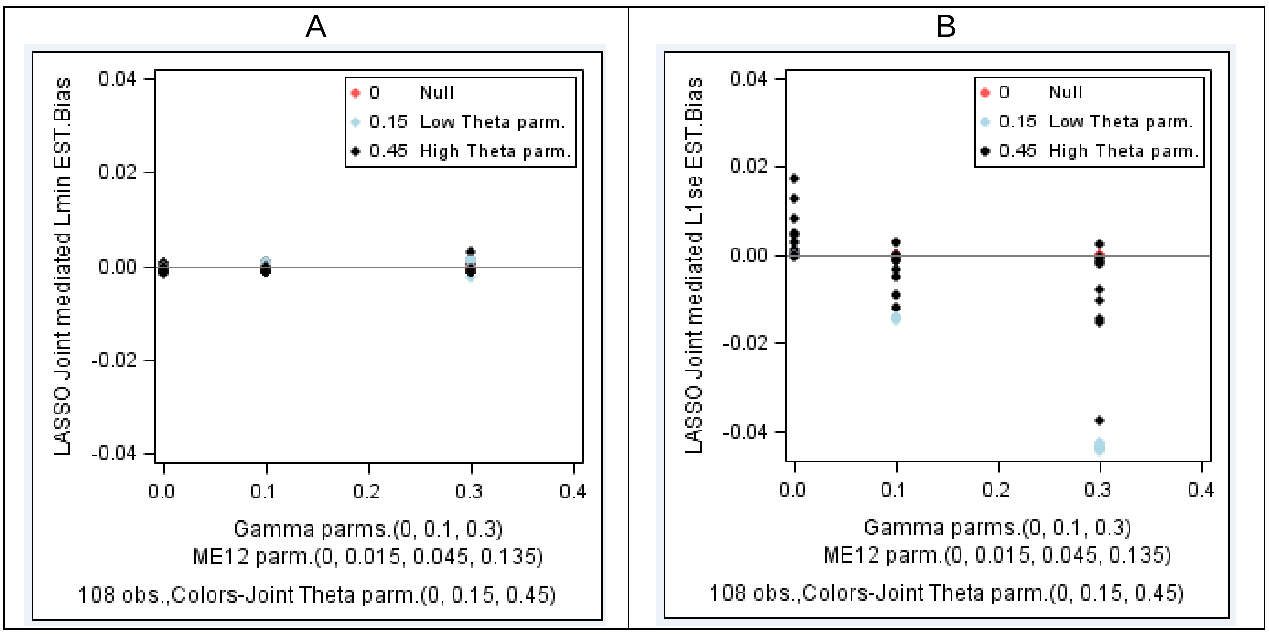


Figure 4.63 LASSO Joint Indirect Effect's *Bias* for Methods A) $ME_{12}^{L_{\min} Bias}$ B) $ME_{12}^{L_{\min+1se} Bias}$

The $ME_{12\,L_{\min}}^{Bias}$ estimate bias is small (±0.003) for the joint mediated estimate parameters ranging between 0 and 0.135. Figure 4.63 B shows that the $ME_{12\,L_{\min+1s.e.}}^{Bias}$ estimate's absolute bias is larger (-0.04) for $\theta_{12} = 0,\ \gamma=0.3$ as compared to (0.02) for $\theta_{12} = 0.45,\ \gamma=0$, given the joint mediated estimate ranging from 0 to 0.135.

The $ME_{12}^{L_{\min}RMSE}$, $ME_{12}^{L_{\min+1se}RMSE}$ average estimate's *RMSE* ranges between 0 and 0.05 for the parameter they are estimating, and are larger for small sample sizes while being directly proportional to the mediated effect estimate's values (Figure 4.64). Low pairwise correlation in small sample sized datasets increase $ME_{12}^{L_{\min+1se}RMSE}$ to 0.07 (Figure 4.64 B).

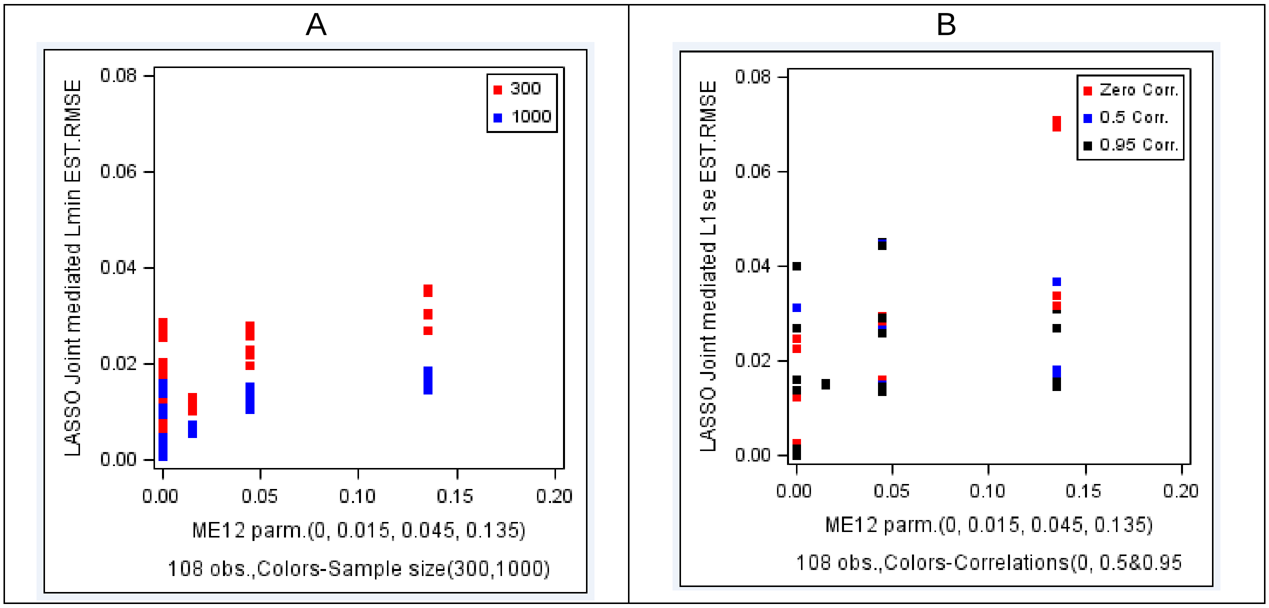
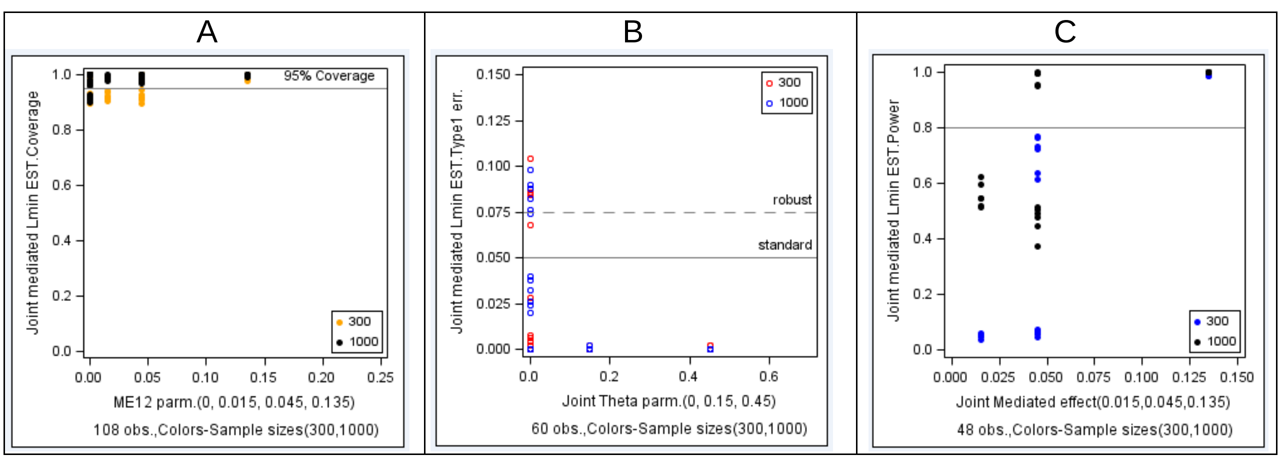


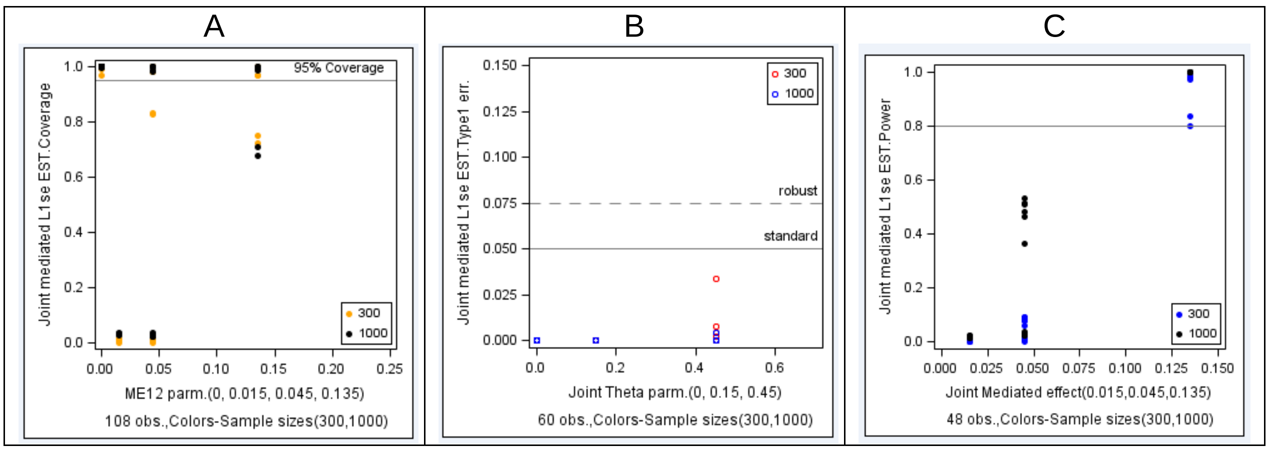
Figure 4.64 LASSO Joint Indirect Effect's *RMSE* for Methods A) $ME_{12}^{RMSE}_{L_{\min}}$ B) $ME_{12}^{RMSE}_{L_{\min+1.s.e.}}$

4.3.2.1 LASSO Joint Mediated Effect –Coverage Probability, Type1 Error & Power



Figures 4.65 Joint Indirect Effect Estimates for $A)ME_{12}^{Cov.}_{L_{\min.}}$, $B)ME_{12}^{Type1\ error}_{L_{\min.}}$, and $C)ME_{12}^{Power}_{L_{\min.}}$

The joint mediated effect's coverage probability for $ME_{12} = 0.045$ is the result of two different parameter arrangements, when 1) $\gamma = 0.1, \theta_{12} = 0.45$ and 2) $\gamma = 0.3, \theta_{12} = 0.15$. The joint mediated effect ME_{12} =0 is the result of three different parameter arrangements of θ_{12} and γ, and these clusters are reflected in the Figure 4.65 B. The coverage probability using $ME_{12}^{L_{\min}}{}^{Cov.}$ is high for the null joint mediated effect. For the alternate LASSO method $ME_{12}^{L_{\min+1s.e.}}{}^{Cov.}$ is poor, for the small effect size (0.015, 0.045), $\theta_1 = 0.15, \gamma = (0.1, 0.3)$ shown in Figures 4.66 A. Comparing the two LASSO methods for coverage probability we can state that $ME_{12}^{L_{\min+1s.e.}}{}^{Cov.}$ is consistently worse than $ME_{12}^{L_{\min}}{}^{Cov.}$ (Figure 4.65 vs. Figure 4.66).



Figures 4.66 Joint Indirect Effect Estimates for $A)ME_{12}^{L_{\min+1s.e.}}{}^{Cov.}$, $B)ME_{12}^{L_{\min+1s.e.}}{}^{Type1\ error}$, and $C)ME_{12}^{L_{\min+1s.e.}}{}^{Power}$

The type1 error for the $ME_{12}^{L_{\min}}{}^{TYP1}$ method has 9 exceptions ($\theta_{12} = 0, \gamma = 0.3$) with type1 error (0.076 to 0.104) which exceed the robust limit of 0.075, of the possible 60 estimates that have *ME₁₂*=0, and there were 0 exceptions for $ME_{12}^{L_{\min+1s.e.}}{}^{TYP1}$. This is because $ME_{12}^{L_{\min+1s.e.}}{}^{TYP1}$ is a more conservative selection method and does not admit predictors that have a weak effect on the outcome into the final solution set; hence its type 1 error is

close to zero. The nine exceptions for $ME_{12}^{TYP1}{}_{L\min}$ type1 error were occurring when $\theta_{1,2} = 0, \gamma = 0.3$ having the combination of a small uncorrelated sample size or weak correlations within a large sample sized dataset, while estimating a weak effect on the outcome. When presented with two correlated predictors the LASSO method admits one of the two correlated predictors by shrinking the other regression coefficient estimate to zero, but if both predictors have a zero effect on the outcome then the admitted variable has an almost zero regression coefficient. This results in a type1 error rate of 0.104 for small datasets, since the standard error for the null mediated effect estimate is high, ranging between 0.1 and 0.2.

The cause for the low power is the small sample size combined with a small effect size ME_{12} = (0.015, 0.045) being detected. The power to detect an effect size of 0.135 is above 0.80 for the LASSO methods. The individual mediated effects $ME_1^{L_{\min}}, ME_2^{L_{\min}}$ & $ME_1^{L_{\min+1s.e.}}, ME_2^{L_{\min+1s.e.}}$ are discussed next.

4.3.2.2 LASSO Individual Mediated Effect –*Estimate, Bias* & *RMSE*

LASSO individual mediated effects are an increasing function of the individual theta and gamma parameters as shown in Figure 4.67 A for $ME_1^{L_{\min}}$ and Figure 4.67 B for $ME_2^{L_{\min}}$. The individual mediated effect estimate bias is affected by high pairwise correlations or small sample sizes with correlated independent variables in the dataset.

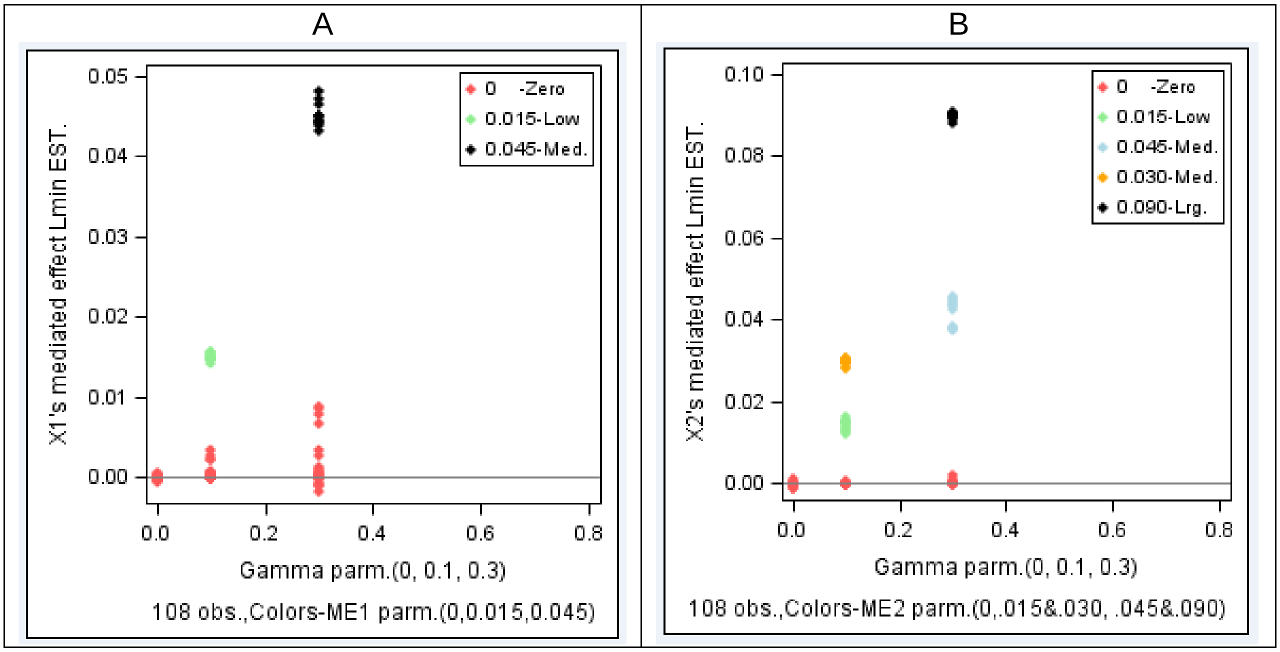
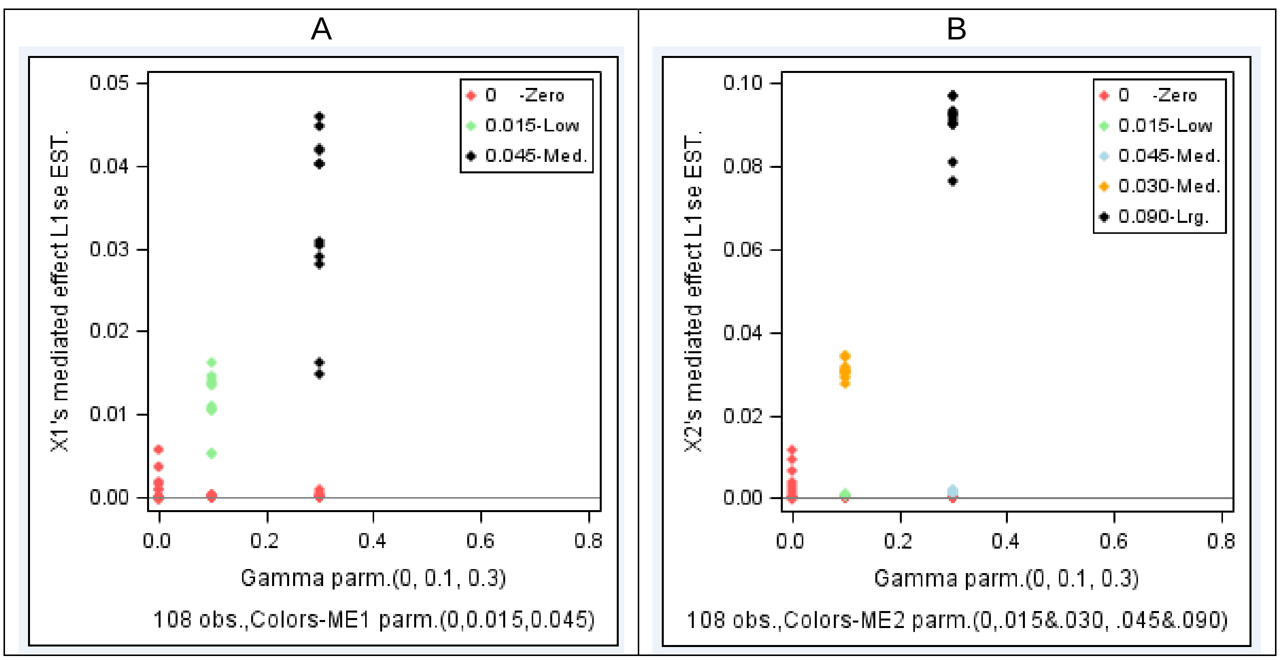


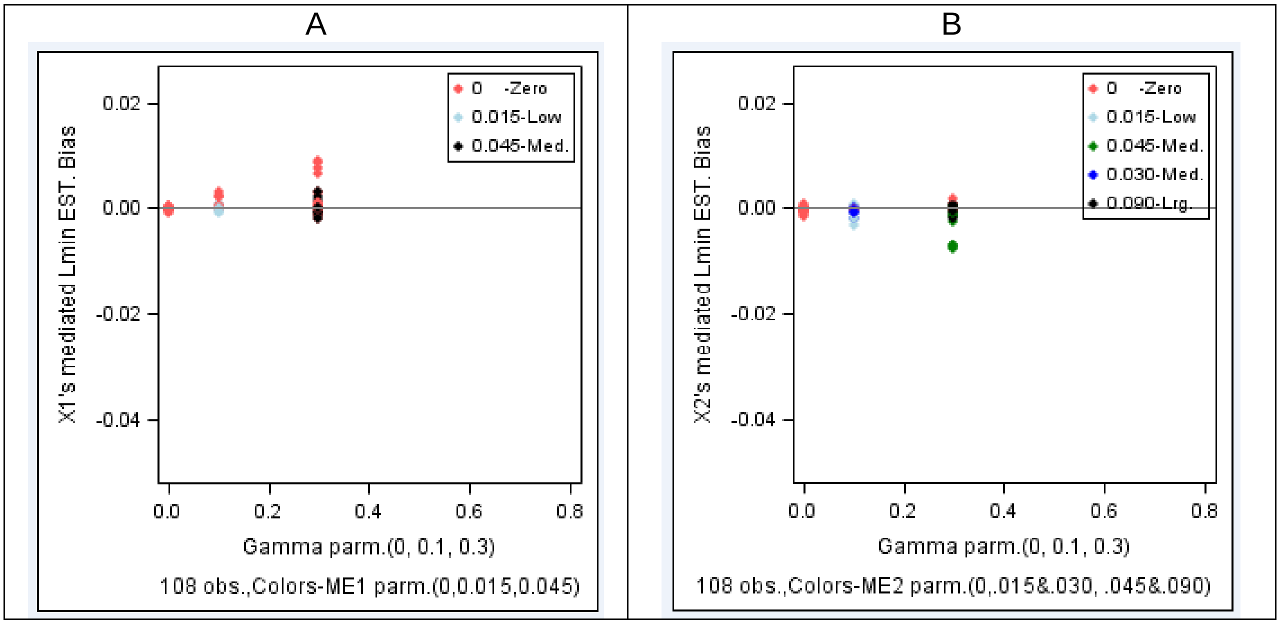
Figure 4.67 LASSO Individual Indirect Effect *Estimates* for A) $ME_1^{L_{\min}}$ and B) $ME_2^{L_{\min}}$



Figures 4.68 LASSO Individual Indirect Effect *Estimates* for A) $ME_1^{L_{\min+1s.e.}}$ and B) $ME_2^{L_{\min+1s.e.}}$

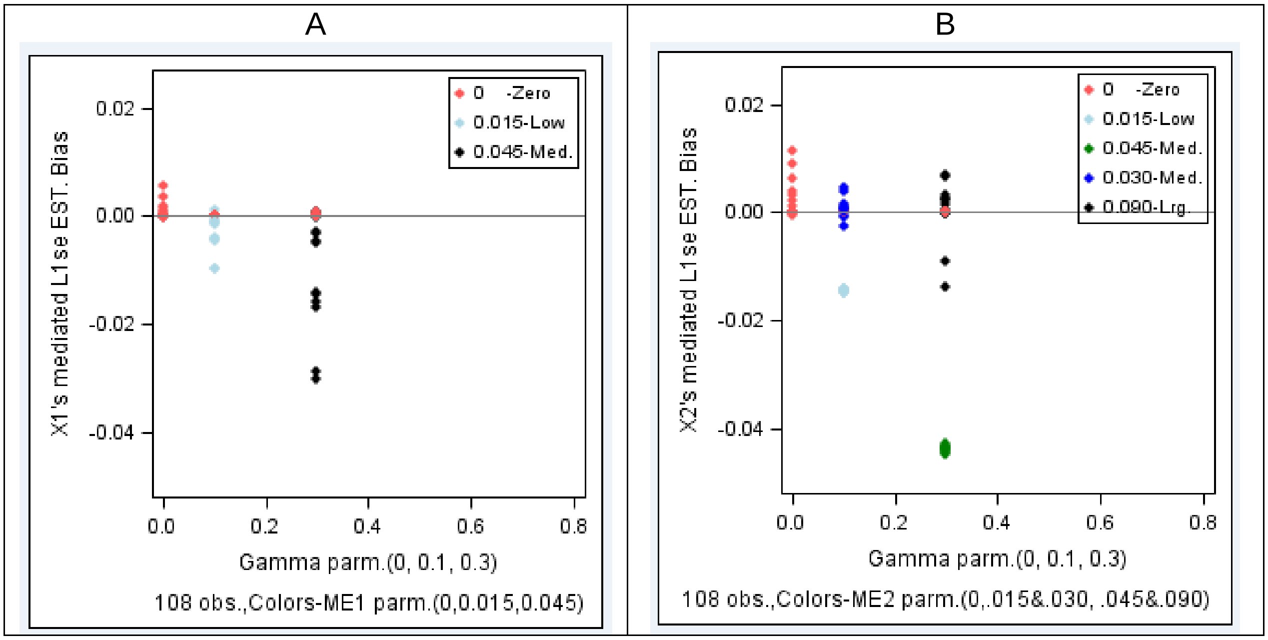
This LASSO individual estimate bias reduces when the sample size is increased. All the estimates are centered on the parameter values listed in the Figures 4.67 A and B. The $ME_1^{L\min+1s.e.}$ estimates are diminished for small mediated effects since that

predictor is not admitted into the final solution set if its effect on the outcome is small, which makes $L_{min+1s.e.}$ more conservative than the L_{min} method, but $L_{min+1s.e}$ has increased estimate variability for small correlated sample sized datasets (Figure 4.68 A and B).



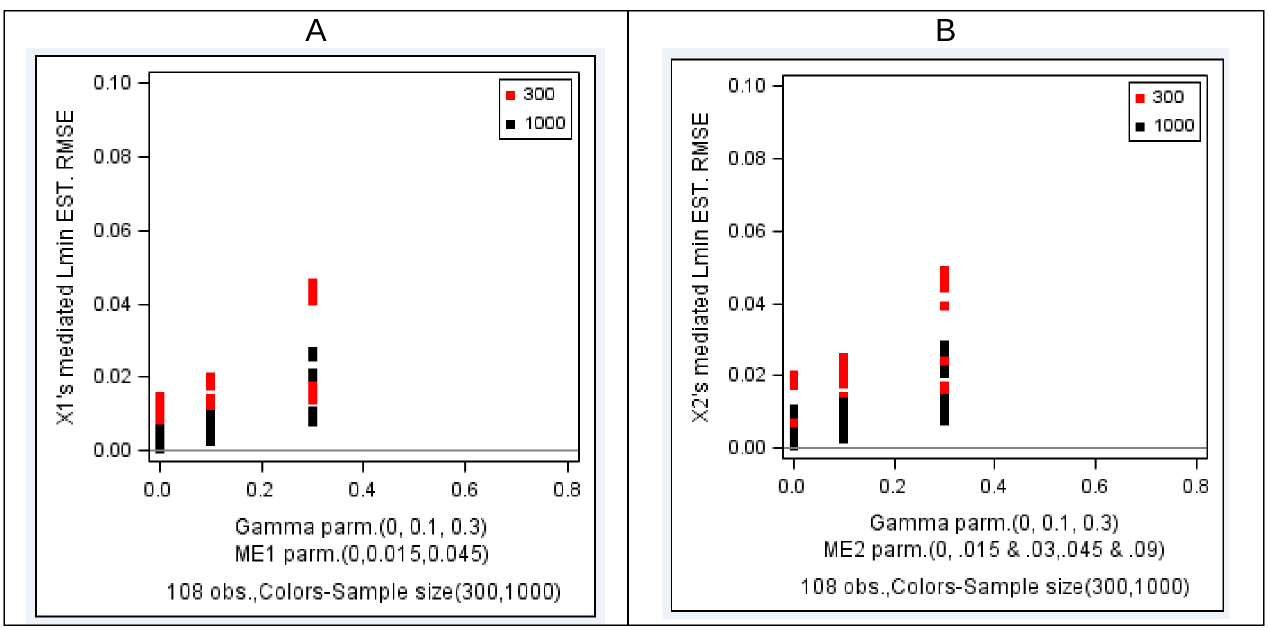
Figures 4.69 LASSO Individual Indirect Effect Est.'s *Bias* for A) $ME_1^{L_{min}}$ and B) $ME_2^{L_{min}}$.

The largest estimate bias of -0.045 for X_2's true mediated effect of 0.045 using the $L_{1se+1s.e}$. method for estimating $ME_2^{L_{min+1s.e.}}$ with parameter values of $(\theta_1, \theta_2) = 0.015$ $\& \gamma$=(0.1,0.3) shown in Figure 4.70 B. The LASSO method diminishes X_1's & X_2's effect to zero by not admitting X_1, X_2 into the final solution set, which results in a large bias of -0.045. If the effect size is larger e.g. 0.0875 then the variable X_2 is admitted into the solution set and the bias is reduced to ±0.01, but not zero because the dataset is small and with correlated independent variables. The $ME_{1,2}^{L_{min}}$ method performs better in estimating the individual mediated effect with a low bias than $ME_{1,2}^{L_{min+1s.e.}}$.



Figures 4.70 LASSO Individual Indirect Effect Est.'s *Bias* for A) $ME_1^{L_{\min+1s.e.}}$ and B) $ME_2^{L_{\min+1s.e.}}$

The mediated effect for X_1's *RMSE* when $\gamma = 0.3$ shows the three (black) clusters in Figure 4.71, which represent the three different mediated effect sizes of 0, 0.015 and 0.045 with increasing *RMSE* for increasing effect size. The *RMSE* values are consistently higher for smaller sample sizes (red).



Figures 4.71 LASSO Individual Indirect Effect Est.'s *RMSE* for A) $ME_1^{L_{\min}}$ and B) $ME_2^{L_{\min}}$

Since the $ME_{1,2}^{L_{\min+1s.e.}}$ is a more conservative in its estimates, it reduces the regression coefficients associated with small effect sizes to zero, and the overall *RMSE* statistic is smaller for the shrunken regression coefficient estimates than for $ME_{1,2}^{L_{\min}}$ as shown in Figure 4.72.

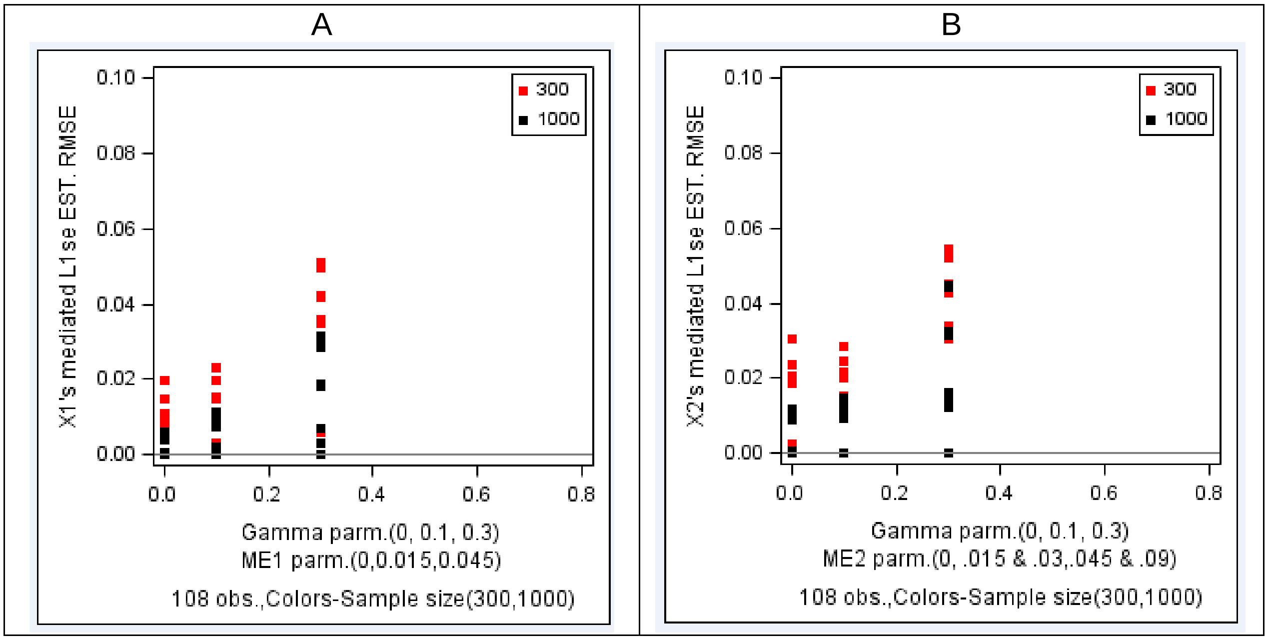


Figure 4.72 LASSO Individual Indirect Effect Est.'s *RMSE* A) $ME_{1}^{L_{\min+1s.e.}}$ and B) $ME_{2}^{L_{\min+1s.e.}}$

4.3.2.3 LASSO Individual Mediated Effect –Coverage Prob., Type1 Error & Power

The LASSO method $L_{\min+1s.e.}$ for individual mediated effects $ME_{1,2}^{L_{\min+1s.e.}}$ covers the null mediated effect with coverage probability of 1, by being conservative in admitting none of the variables into the final solution set. In Figure 4.73 E the coverage for $ME_{2}^{L^{COV}_{\min+1s.e.}}$ detecting a small mediated effect (0.015, 0.045) using a small sample size having strong correlated predictors, results in the low coverage probability of approximately 0.

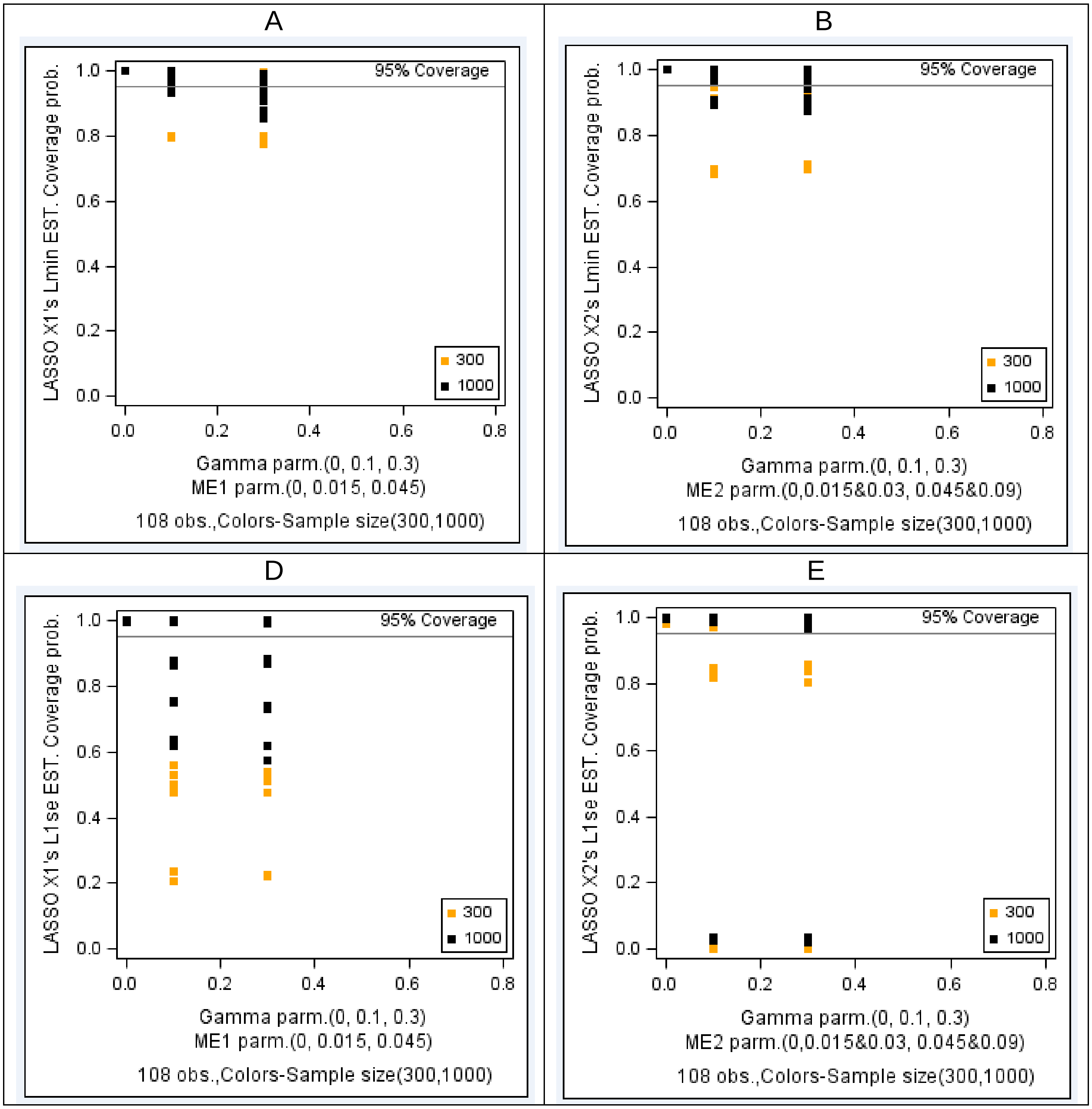


Figure 4.73 LASSO Indiv. Indirect *Est.'s* Coverage Prob. A&B) $ME_{1,2}^{Cov\ L_{min}}$ and C&D) $ME_{1,2}^{Cov.\ L_{min+1s.e.}}$

The individual mediated effect type 1error and power using the L_{min} method for estimating the mediated effects are discussed next and shown in Figure 4.74.

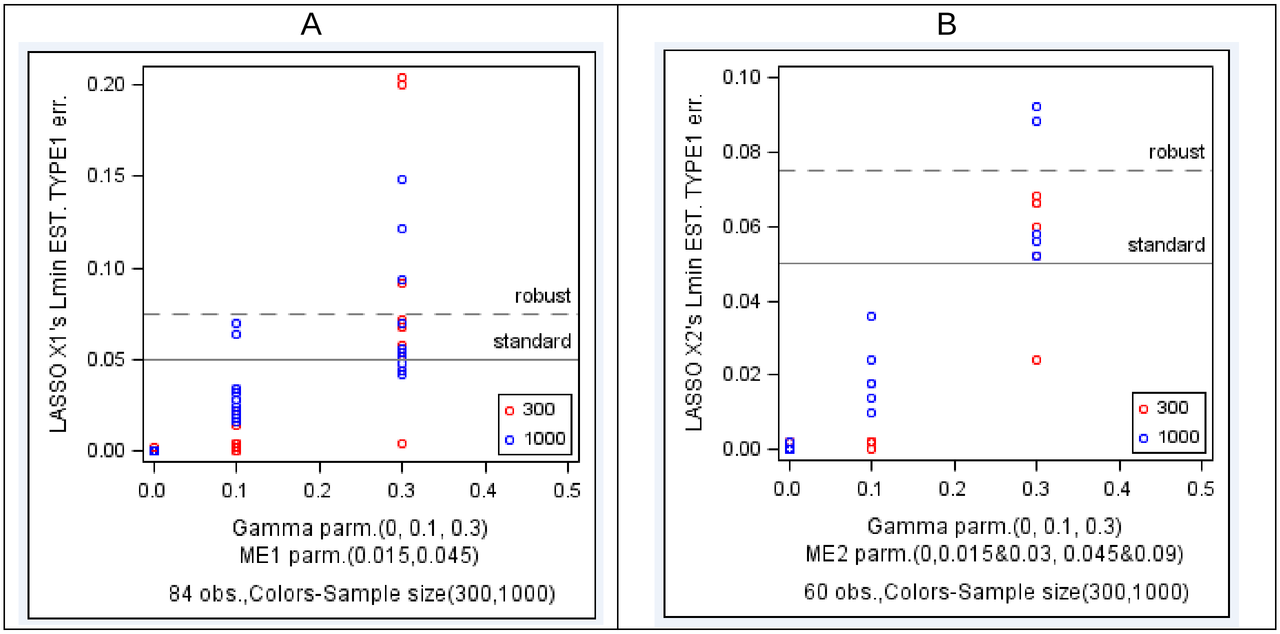


Figure 4.74 LASSO Individual Indirect Effect Est.'s *Type1 err.* A) $ME_1^{L_{\min}}$ and B) $ME_2^{L_{\min}}$

Of the 84 conditions with the individual mediated effect for X_1 being zero (ME_1=0), there were 6 exceptions with the Type 1 error exceeding the robust limit of 0.075. The six $ME_1^{L_{\min}}$ Type 1 error exceptions have a range from 0.09 to 0.20, with conditions set for $\theta_1 = 0, \gamma = 0.3, \rho_{12} = 0.95$, regardless of the sample size. There were 2 exceptions (0.09) with conditions set for $\theta_2 = 0, \gamma = 0.3, \rho_{12} = 0.95$ out of the 60 possible conditions where ME_2=0 shown in Figure 4.74 B. The type1 error rates for $ME_{1,2}^{L_{\min}}$ show that this LASSO method works well except for conditions where the pairwise correlations are high $\rho_{12} = 0.95$, with high gamma parameter $\gamma = 0.3$, since the bias on the theta estimate multiplied by a large gamma parameter estimate results in a non-zero mediated effect, which raises the type1 error rate.

The individual $ME_1^{L_{\min+1s.e.}}, ME_2^{L_{\min+1s.e.}}$ estimates from the $L_{\min+1s.e.}$ method are very conservative and Figure 4.75 A and B show that the type 1 error rates of 0.05(standard

limit) and 0.075(robust limit) are not used appropriately and is therefore not preferred for use in a 2-variable LASSO regression mediation analysis.

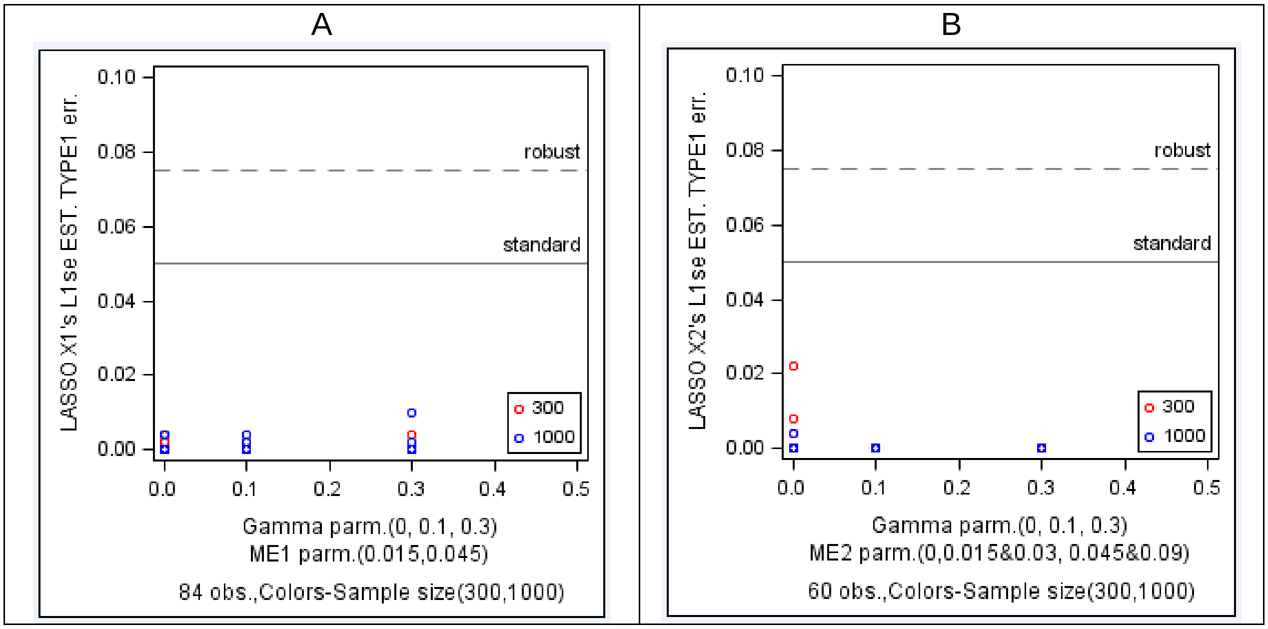


Figure 4.75 LASSO Individual Indirect Effect Est.'s*Type1 err.* A) $ME_1^{L_{\min+1s.e.}}$ and B) $ME_2^{L_{\min+1s.e.}}$

Figure 4.76 A and D show the power to detect the individual mediated effect $ME_1^{\overset{Power}{L_{\min}}}$ and $ME_1^{\overset{Power}{L_{\min+1s.e.}}}$, Figure 4.76 B&C show $ME_2^{\overset{Power}{L_{\min}}}$ & Figure 4.76 E&F show $ME_2^{\overset{Power}{L_{\min+1s.e.}}}$.

4.3.2.4 Individual Mediated Effect LASSO's Power to Detect $ME_{1,2}^{LASSO}$ When $ME_{1,2} \neq 0$

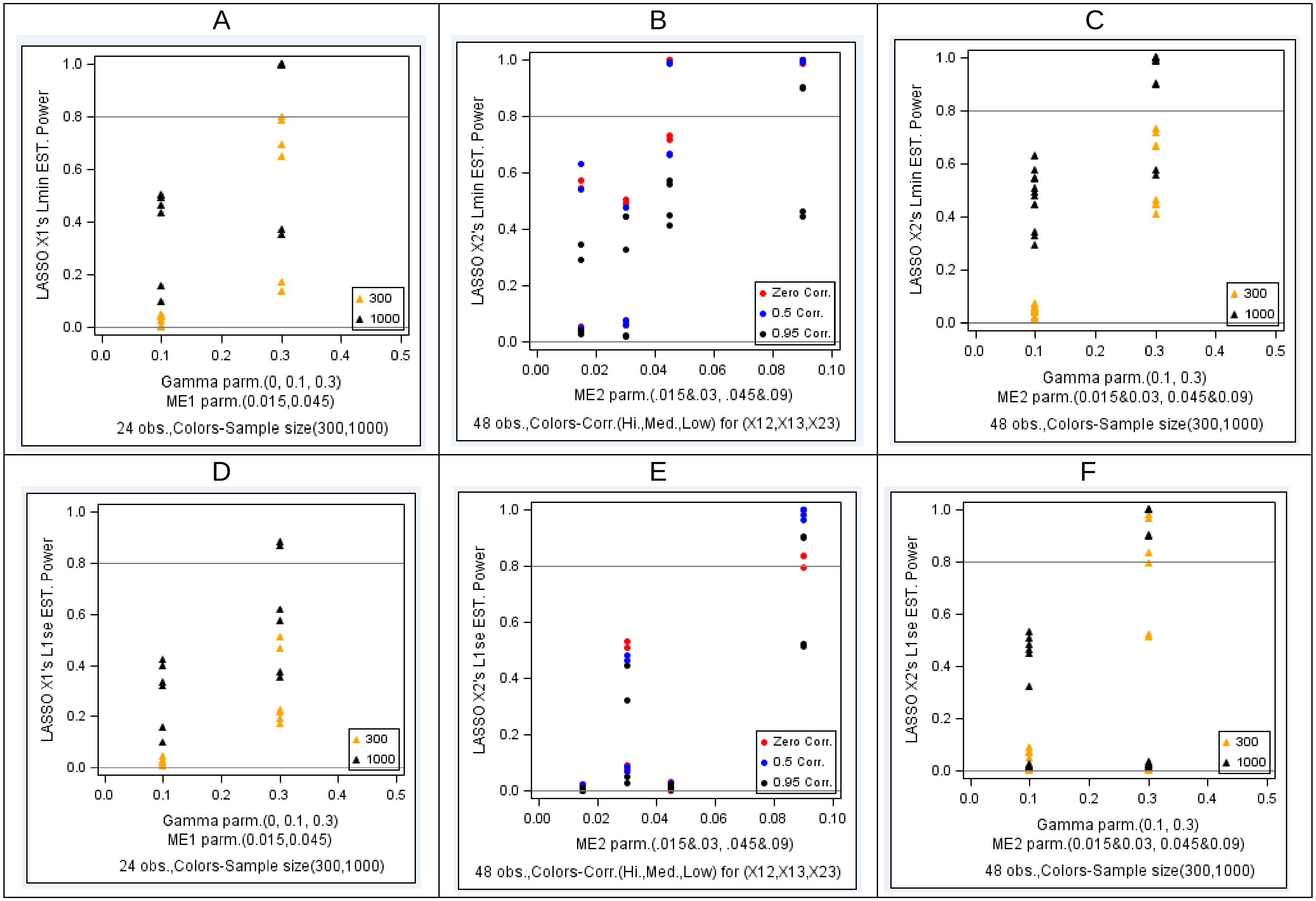


Figure 4.76 LASSO Individual Indirect Effect *Est.'s* Power A-C) $ME_{1,2}^{L_{\min}\,Power}$ and D-F) $ME_{1,2}^{L_{\min+1s.e.}\,Power}$

The conclusion is that the $ME_1^{L_{\min}}$ method has a low power to detect ME_1^{LASSO} for small highly correlated predictor datasets tying to detect a small effect size. The power increases with decreasing multicollinearity and increasing effect size or the sample size. Of the 24 conditions for the $L_{min+1se}$ method, had a power ranging from 0 to 0.62. The $L_{min+1se}$ method had 12 conditions where ME_1=0.015 (power range: 0 to 0.42) and 10 conditions when ME_1=0.045 (power range: 0.17 to 0.62). All these conditions relate to high $ME_{1\,L\min+1s.e.}^{RMSE}$ for the smaller estimates of $ME_1^{L\min+1s.e.}$ which result in the reduced power. The power to detect ME_2 is discussed next.

Figures 4.76 B&C, E&F show the power to detect the individual mediated effect $ME_2^{L\min}$ & $ME_2^{L_{\min+1s.e.}}$ Of the 48 conditions where $ME_2 \neq 0$, 34 had a power range 0.02 to 0.73 for $ME_{2\,L\min}^{PWR}$ when using the L_{min} method. Figure 4.76 B and C show that small datasets (gold triangles), with high pairwise correlations (black dots) have reduced power to detect the $ME_2^{L_{\min}}$ individual mediated effect. The power is worse (0.018) when the standard error of the estimate is larger (0.023) and better (power=0.73) for uncorrelated predictors in small datasets, if the effect being detected is larger (0.045). The power with the $L_{min+1s.e.}$ method has a power ranging from 0.01 to 0.8. The 0.01 power is because the variable X_2's mediated effect has a large standard error (0.045) for the $ME_2^{L_{\min+1s.e.}}$ estimate (true value 0.045), being estimated using a small sample size. In summary, the $ME_{1,2}^{L_{\min}}$ method has low power (0.03) when detecting a small effect (0.015) using a small sample size dataset, with highly correlated predictors. With larger uncorrelated datasets, LASSO can detect the small mediated effect (0.015) with a power of 0.574, which is still low for power.

4.3.2.4 LASSO Joint Direct Effect –*Estimate, Bias* & *RMSE*

LASSO joint direct effects are an increasing function of the joint beta parameter as shown in Figure 4.77 A and B. Gamma parameter values, pairwise correlations and sample size do not influence the joint direct effect $DE_{12}^{L_{\min}}$ as shown in Figure 4.77 A, but they do affect the estimates for $DE_{12}^{L_{\min+1s.e.}}$ as shown in Figure 4.77 B.

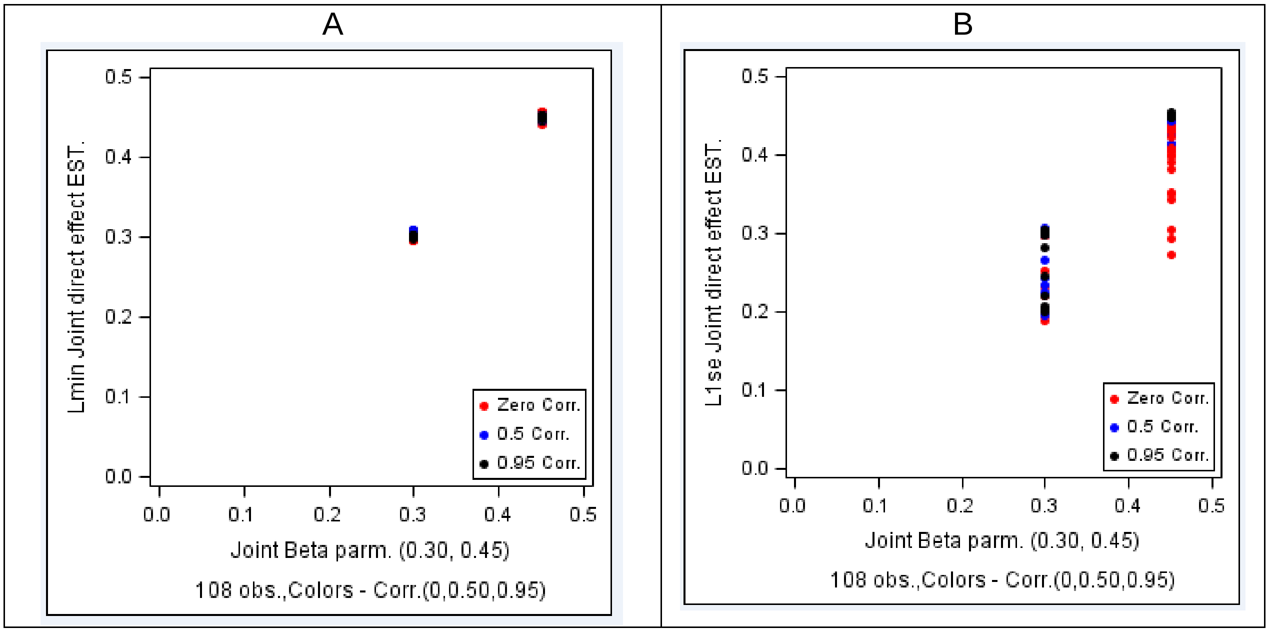


Figure 4.77 LASSO Joint Direct Effect *Estimate* for $A)\ DE_{12}^{L_{\min}}$ and $B)\ DE_{12}^{L_{\min+1s.e.}}$

Each $DE_{12}^{L_{\min}}$ estimate value has a low bias and the $L_{\min}$ method produces joint direct effect regression coefficient estimates with a low bias. There is more dispersion of the estimates under the $L_{min+1s.e.}$ method as shown in Figure 4.77 B and the uncorrelated independent variables (red dots) show increased estimate bias.

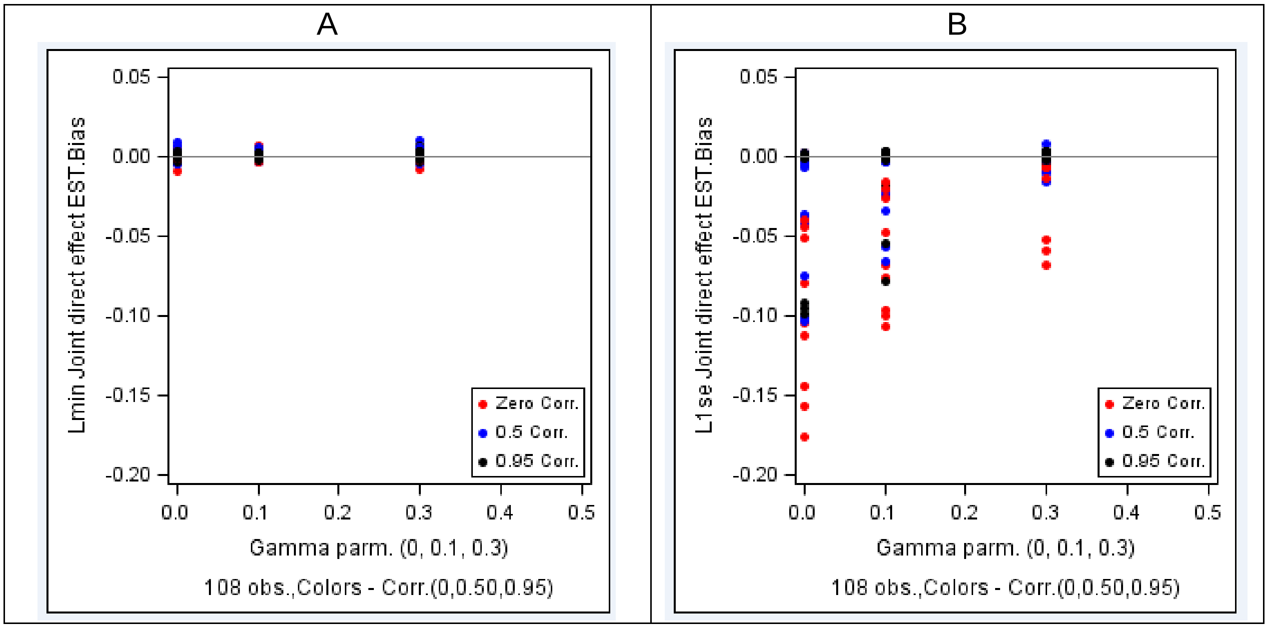


Figure 4.78 LASSO Joint Direct Effect *Est.'s Bias* for A) $DE_{12}^{L_{\min}}$ and B) $DE_{12}^{L_{\min+1s.e.}}$

The $DE_{12}^{L_{\min}}$ estimate bias in Figure 4.78 A is small (±0.01) for joint direct effect estimates (0.30, 0.45) attributable to the L_{min} LASSO regression method. However, the $DE_{12}^{L_{\min+1s.e.}}$ estimate bias (-0.18 to 0.01) is dependent on the gamma parameter and the pairwise correlations as shown in Figure 4.78 B.

The estimate's standard error and the average estimate's *RMSE* are dependent on the sample size (small *N* has a large *standard error*) shown in Figure 4.79 A and B.

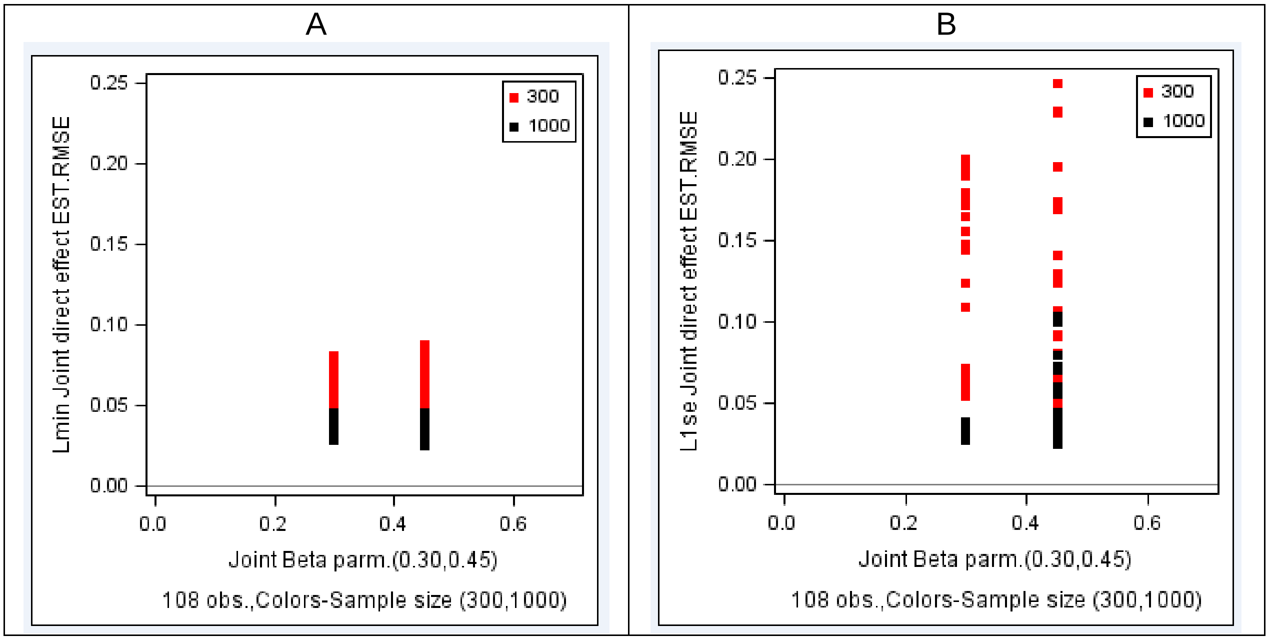
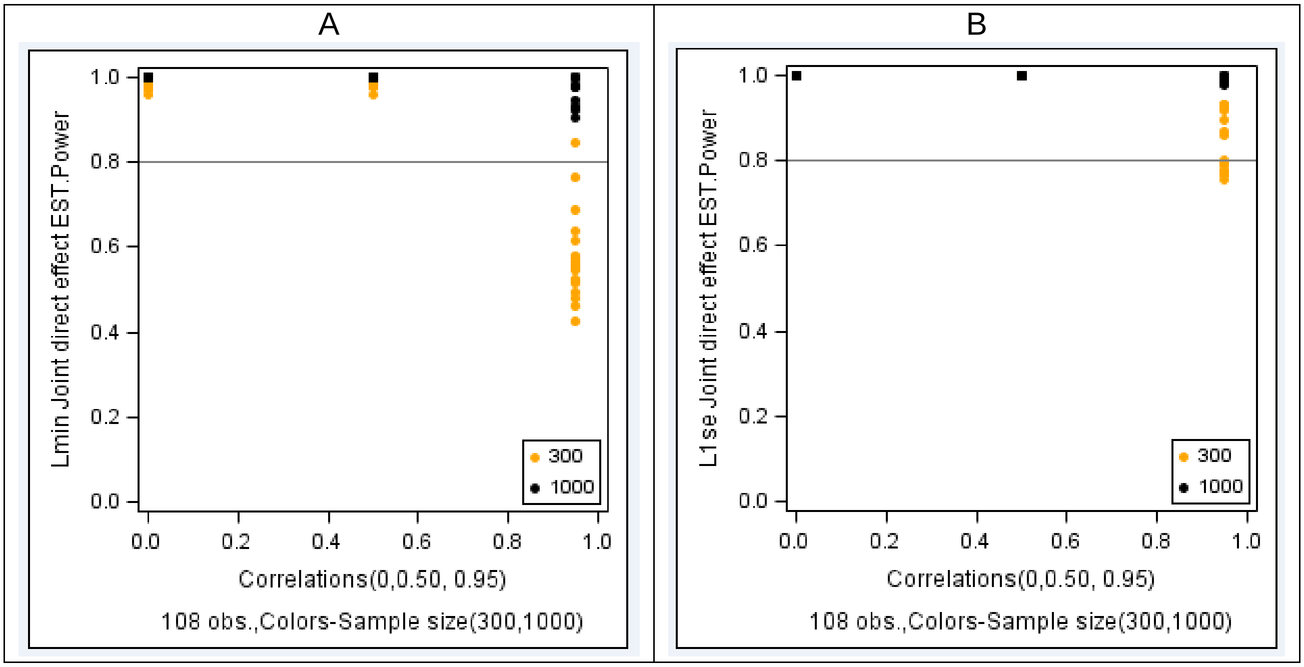


Figure 4.79 LASSO Joint Direct Effect *Est.'s RMSE* for $A)\ DE_{12}^{L_{\min}}$ and $B)\ DE_{12}^{L_{\min+1s.e.}}$

The LASSO joint direct effect $DE_{12}^{RMSE}{}_{L_{\min}}$ has estimates in the range 0.026 to 0.087 and $DE_{12}^{RMSE}{}_{L_{\min+1s.e.}}$ has estimates in the range 0.026 to 0.247. The high *RMSE* estimate $DE_{12}^{RMSE}{}_{L_{\min+1s.e.}}$ =0.257 corresponds to the uncorrelated predictors in the small sample sized dataset.

4.3.2.5 LASSO Joint Direct effect – Coverage probability, Power to detect DE_{12}^{LASSO}

There were no conditions where DE_{12}=0 to calculate the type1 error rate for the LASSO method. The same is true for the individual direct effect for X_1, since the beta parameters are non-zero $\beta_1 = (0.15, 0.30)$. However, the type1 error rate is determined for the individual X_2's direct effect, since the beta parameter takes on values $\beta_2 = (0,\ 0.30)$ and there are 54 conditions where DE_2=0.



Figures 4.80 Joint Direct Effect Estimate DE_{12}^{LASSO} *Power* A) L_{min} B) $L_{min+1s.e.}$ Methods

The L_{min} method estimating the LASSO joint direct effect for $DE_{12} \neq 0$ has a power ranging 0.43 to 0.77. Figures 4.80 A shows $DE_{12}^{PWR}{}_{L_{min}}$ and $DE_{12}^{PWR}{}_{L_{min+1s.e.}}$ having a small sample size (gold) and high pairwise correlations $\rho_{12} = 0.95$. The power looks better for $L_{min+1s.e.}$ over L_{min} since the former method shrinks the weaker associations to zero.

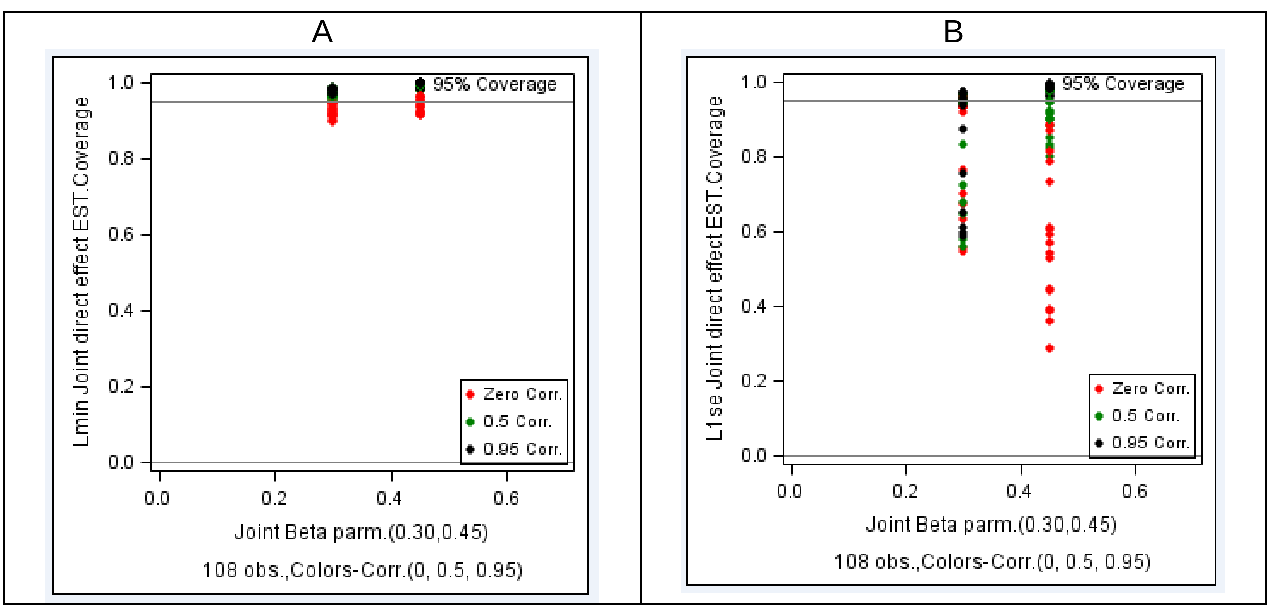


Figure 4.81 Joint Direct Effect Estimate DE_{12}^{LASSO} Coverage A) L_{min} and B) $L_{min+1s.e.}$

Figure 4.81A and B shows the LASSO joint direct effect's coverage probability under the L_{min} and $L_{min+1s.e.}$ methods for all 108 conditions. Figure 4.81A shows $DE_{12}^{COV_{L_{min}}}$ has coverage of 0.90 to 1.0, for all the 108 conditions. $L_{min+1s.e.}$ method being a more conservative method performs worse on the joint direct effect coverage (probability range 0.29 to 1.0) since small effects are shrunk to zero. This creates a negative bias on the joint direct effect estimate DE_{12}, with a large *RMSE*, resulting in poor coverage.

The individual direct effect *Estimates*, *Bias* and *RMSE* for $DE_{1,2}^{EST_{L_{min}}}, DE_{1,2}^{Bias_{L_{min}}}, DE_{1,2}^{RMSE_{L_{min}}}$ are shown in Figure 4.82 A & D for the estimate, B & E for the estimate's bias, and C & F for the estimate's *RMSE*. The sample size and pairwise correlations influence all three individual predictor estimates shown in Figure 4.82.

LASSO Individual Direct Effects –*Estimate, Bias & RMSE*

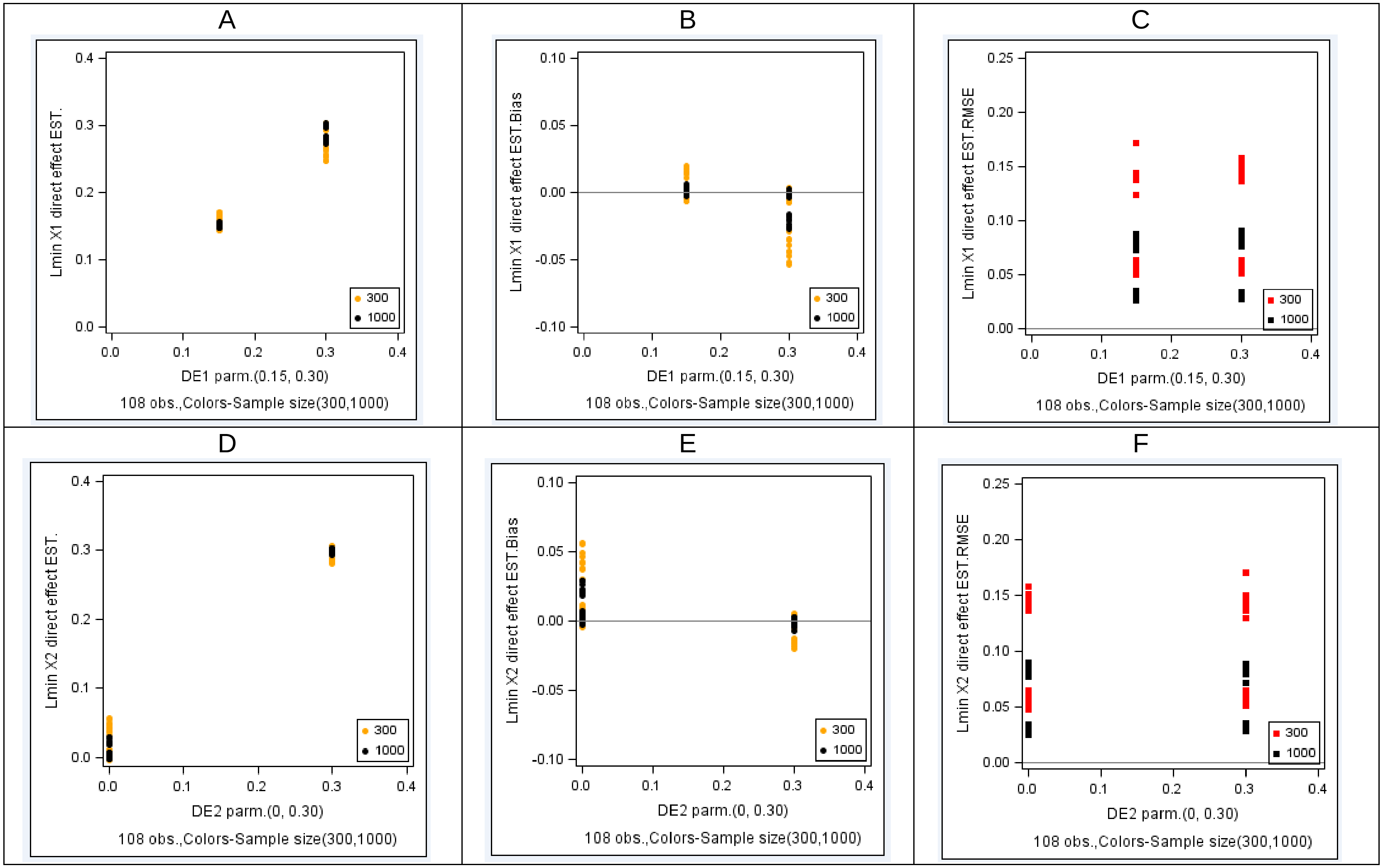


Figure 4.82 LASSO Individual Direct Effect Estimates A-F) $DE_{1,2}^{\overset{EST}{L_{\min}}}, DE_{1,2}^{\overset{Bias}{L_{\min}}}, DE_{1,2}^{\overset{RMSE}{L_{\min}}}$

4.3.2.6 LASSO Individual Direct Effect's Type1 for DE_2^{Typ1},Power & Coverage

There were three exceptions for $DE_2^{TYP1}_{L_{\min}}$ out of the possible 54 having the type1 error that exceeded the 0.075 limit, because of the small sample size with highly correlated predictors, which results in a positively biased estimate ($DE_{2\,L_{\min}}^{Bias}$ =0.05) and large RMSE ($DE_{2\,L_{\min}}^{RMSE}$ 0.16) for $DE_2^{L_{\min}}$ shown in Figure 4.83.

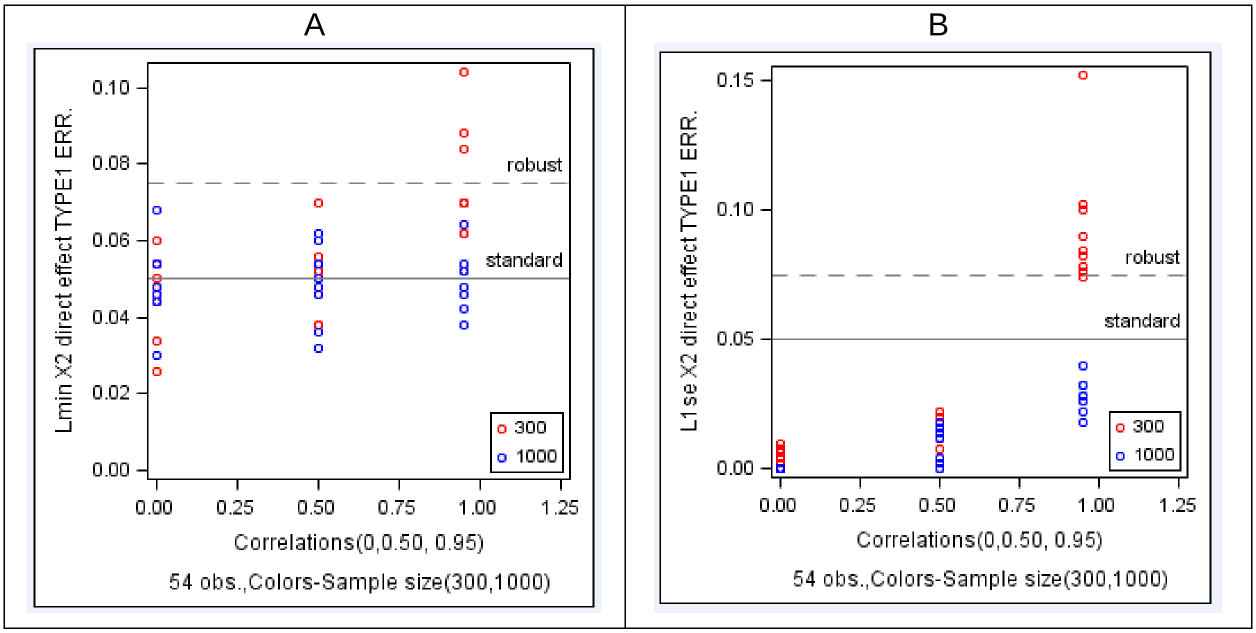


Figure 4.83 LASSO Individual Direct Effect, X_2 Est.'s type1 err. DE_2^{Typ1} A) L_{min} B) $L_{min+1se}$

The range for these type1 error exceptions for $DE_{2\,L_{\min}}^{TYP1}$ was 0.088 to 0.104. The range of the eight type1 error exceptions for $DE_{2\,L_{\min+1s.e.}}^{TYP1}$ was 0.076 to 0.152. This is evidence that $DE_{1,2\,L_{\min}}^{TYP1}$ performs better than $DE_{1,2\,L_{\min+1s.e.}}^{TYP1}$ as a LASSO regression method in 2-variable mediation analysis.

4.3.2.7 LASSO Method Individual Direct Effect's Power for $DE_1^{PWR}{}_{L\min}$ and $DE_2^{PWR}{}_{L\min}$.

The plots in Figure 4.84 show that the LASSO method's individual direct effect's power $DE_{1,2}^{PWR}{}_{L\min}$ performs poorly (Power = 0.13 to 0.19) on highly correlated predictors ($\rho_{12} = 0.95$) with small effect sizes (0.15), estimated using small sample sized datasets (N=300).

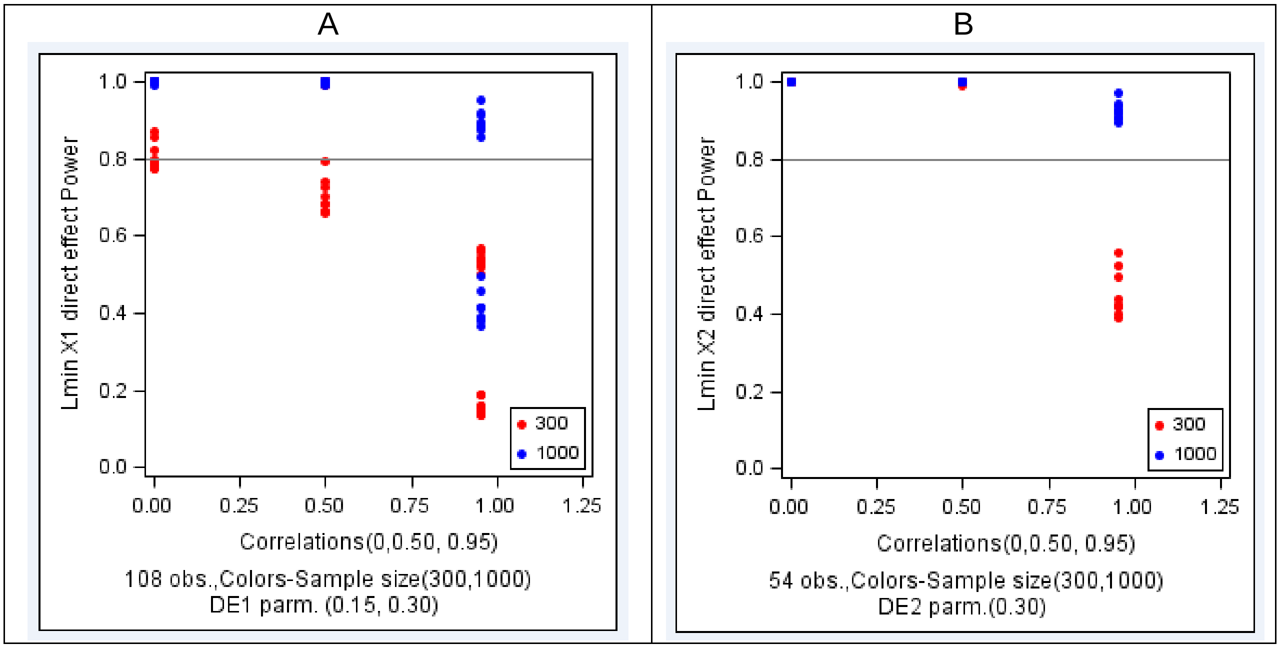


Figure 4.84 A) X1 & B) X2 LASSO Individual Direct Effect Estimate's Power $DE_{1,2}^{PWR}$

There were 42 conditions of the 108 conditions for $DE_1^{PWR}{}_{L\min}$ having a low power of which 33 exceptions were with the smaller dataset (N=300) and 9 were with the larger dataset (N=1000, $\rho_{12} = 0.95$). Figure 4.84 A shows that the smaller sized dataset (red), with the higher the pairwise correlation $\rho_{12} = 0.95$, has the smaller power. There were 9 conditions to the 54 conditions for $DE_2^{PWR}{}_{L\min}$ having a low power and all nine exceptions were because of the smaller sized dataset (N=300. This leads to the conclusion that the LASSO for two variable mediation analysis does not perform well for small datasets with

highly correlated predictor variables especially when the effect size is small. The power is still low if the dataset size increases to 1000, but with the pairwise correlation remaining high ($\rho_{12} = 0.95$).

The individual predictor coverage for direct effect is discussed using Figure 4.85.

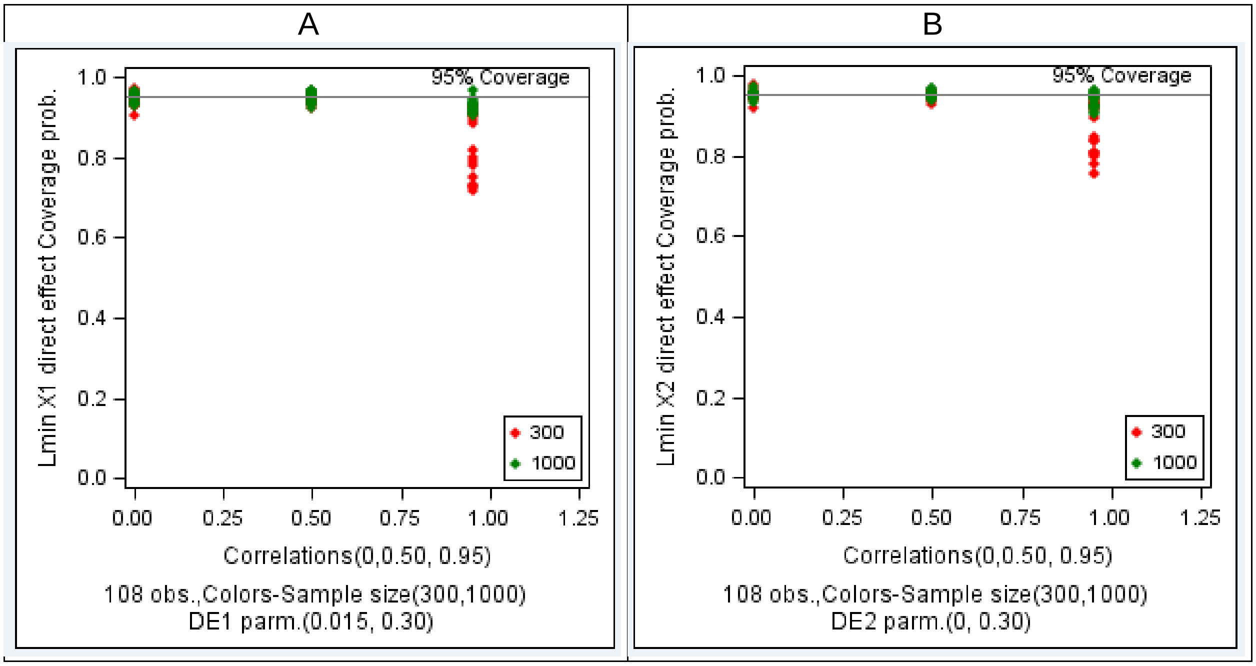


Figure 4.85 Lmin A) X_1 B) X_2 LASSO Indiv. Direct Effect Estimate's Coverage $DE_{1,2}^{L_{\min}\, Cov.prob.}$

The low coverage for the 8 conditions out of 108, is for small sample sizes (N=300) with high correlated predictors ($\rho_{12} = 0.95$), detecting a small effect size (0.15) as shown in Figure 4.85 A for $DE_1^{L_{\min}\, Cov.}$. The low coverage for the 9 conditions out of 108, shown in Figure 4.85 B is for $DE_2^{L_{\min}\, Cov.}$ is for small sample sizes (N=300) with high correlated predictors.

4.3.3 LASSO Method 3-Variable Mediation Analysis

4.3.3.1 Joint Mediated Effect – LASSO Regression *Est., Bias & RMSE*

Figure 4.86 A and B show the influencing variables that determine the joint mediated effect for 3-variable mediation using the LASSO method. The influencing variables for the LASSO regression coefficients determined by cross validation (min, min+1s.e.) are: the gamma parameter $\gamma(0, 0.25, 0.35)$ and the joint theta parameter $\theta_{123}(0.60, 0.55, 0.55)$.

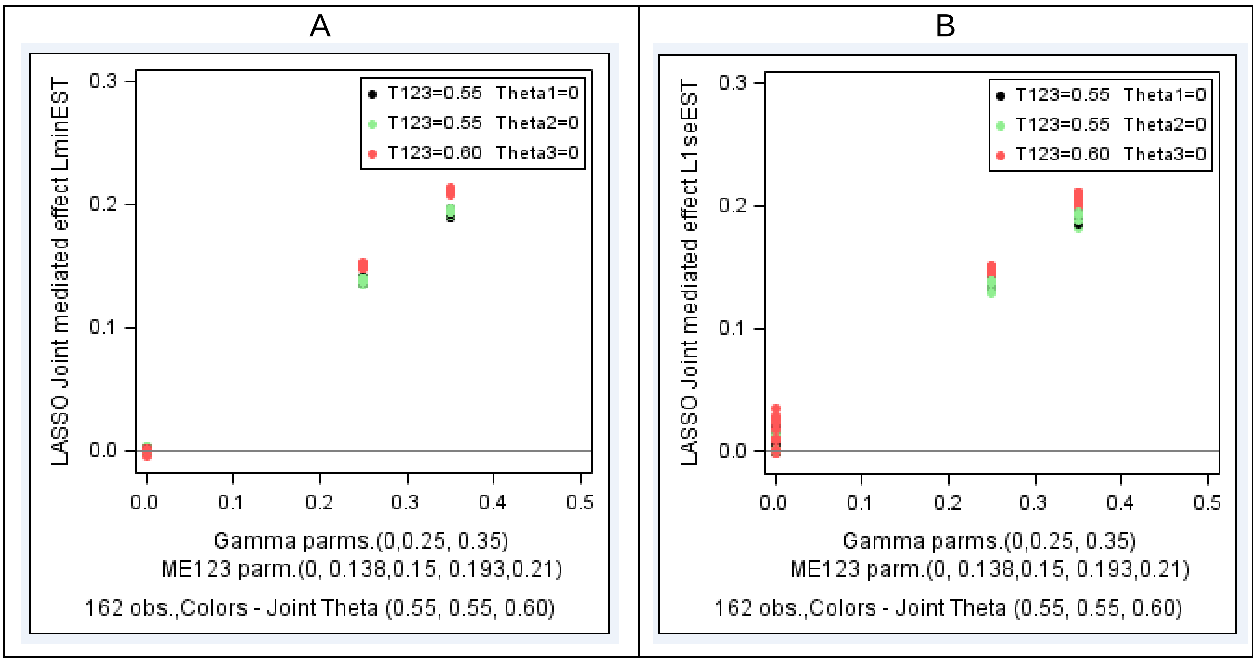


Figure 4.86 Joint Indirect Effect Estimate for ME_{123}^{LASSO} A) L_{min} & B) $L_{min+1s.e.}$ Coefficients

The higher values for γ, θ_{123} increase the joint indirect effect ME_{123}^{LASSO}. The method of selecting the LASSO regression coefficients using the minimum $MSE_{min+1s.e.}$ on the cross validation curve, introduces additional bias and variability to the LASSO estimate as compared to the minimum coefficient selection method which provides a more unbiased estimate with a smaller standard error for the estimate.

The joint mediated effect's estimate *Bias* is shown in Figure 4.87 A and B for $ME_{123}^{L_{\min}^{Bias}}$ and $ME_{123}^{L_{(\min+1s.e.)}^{Bias}}$. As inferred from Figure 4.86 A and B, the estimate bias is larger and sensitive to small sample sizes, when selecting the LASSO regression coefficients using the $MSE_{min+1s.e.}$ on the cross validation curve.

Figure 4.87 A and B show that the LASSO method can provide a near unbiased estimate (±0.005) for the true value's range (0, 0.21). Neither γ, θ_{123} nor ρ_{123}, N affect the estimate bias for $ME_{123}^{L_{\min}^{Bias}}$, as they do for $ME_{123}^{L_{(\min+1s.e.)}^{Bias}}$ shown in Figure 4.87 B and C.

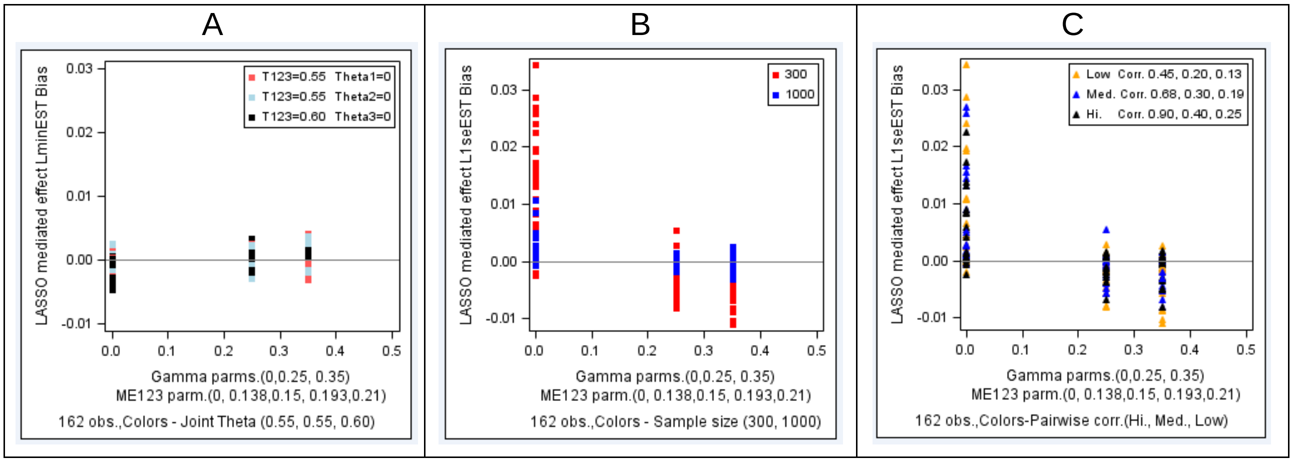


Figure 4.87 Joint Est. *Bias* by γ A) $ME_{123}^{L_{\min}^{Bias}}$ by θ_{123} B&C) $ME_{123}^{L_{(\min+1s.e.)}^{Bias}}$ Grouped by N and ρ_{123}

The LASSO estimate's *RMSE* ranges from 0.02 to 0.04, is influenced by the sample size, which groups the *RMSE* values, and is not influenced by the pairwise correlations between the predictors shown in Figure 4.88. The $ME_{123}^{L_{\min+1s.e}^{RMSE}}$ estimate *RMSE* and standard errors are higher than for $ME_{123}^{L_{\min}^{RMSE}}$ due to a larger bias on the gamma parameter and a higher joint theta parameter value.

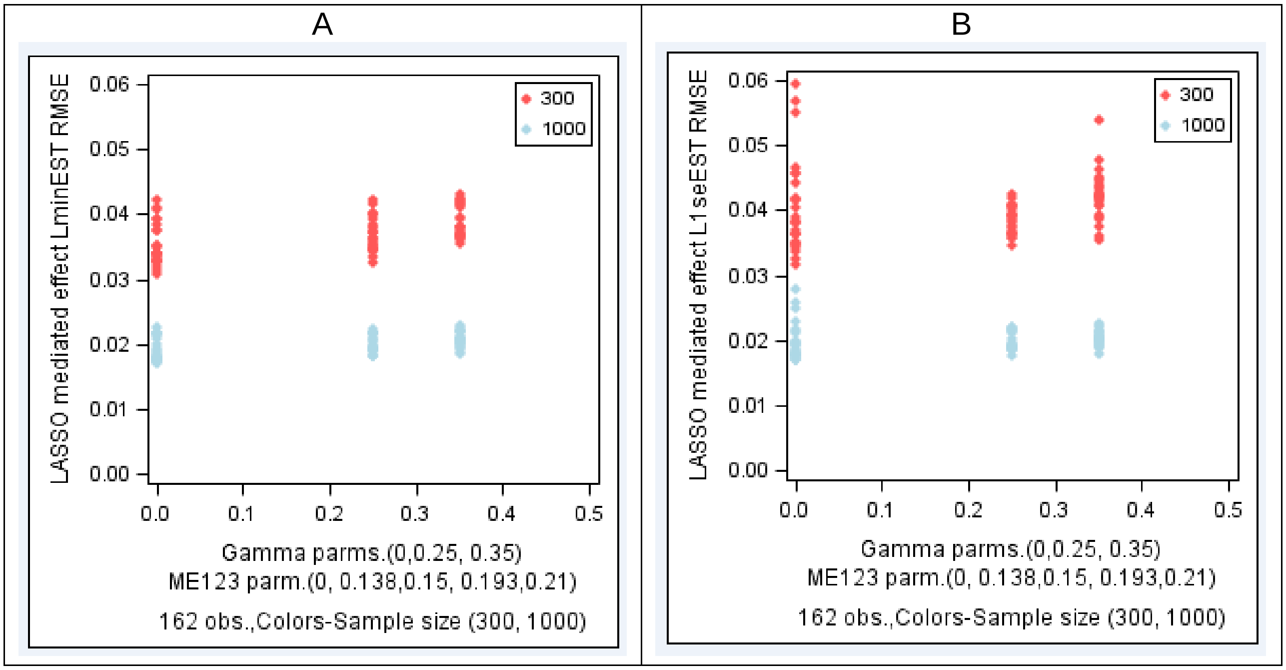


Figure 4.88 Estimate *RMSE* for ME_{123}^{LASSO} by γ A) $ME_{123}^{L_{\min}^{RMSE}}$ and B) $ME_{123}^{L_{\min+1s.e}^{RMSE}}$

4.3.3.2 LASSO Joint 2-Variable Mediated Effect –Coverage, Type1 Error & Power

The LASSO method's coverage probability for the joint mediated effect ME_{123}^{LASSO} ranges 0.91 to 0.97 for 108 conditions of $ME_{123}^{L_{\min}^{Cov}}$ shown in Figure 4.89 A and 0.70 to 0.96 for $ME_{123}^{L_{(\min+1s.e.)}^{Cov}}$ shown in Figure 4.89 B, which reduces coverage for the null mediated effect when the sample size is small. The LASSO method's power for detecting the joint mediated effect is shown in Figure 4.89 C and D. Both LASSO methods have a high power for detecting the joint mediated effect $ME_{123}^{\substack{Estimate\\LASSO}}$.

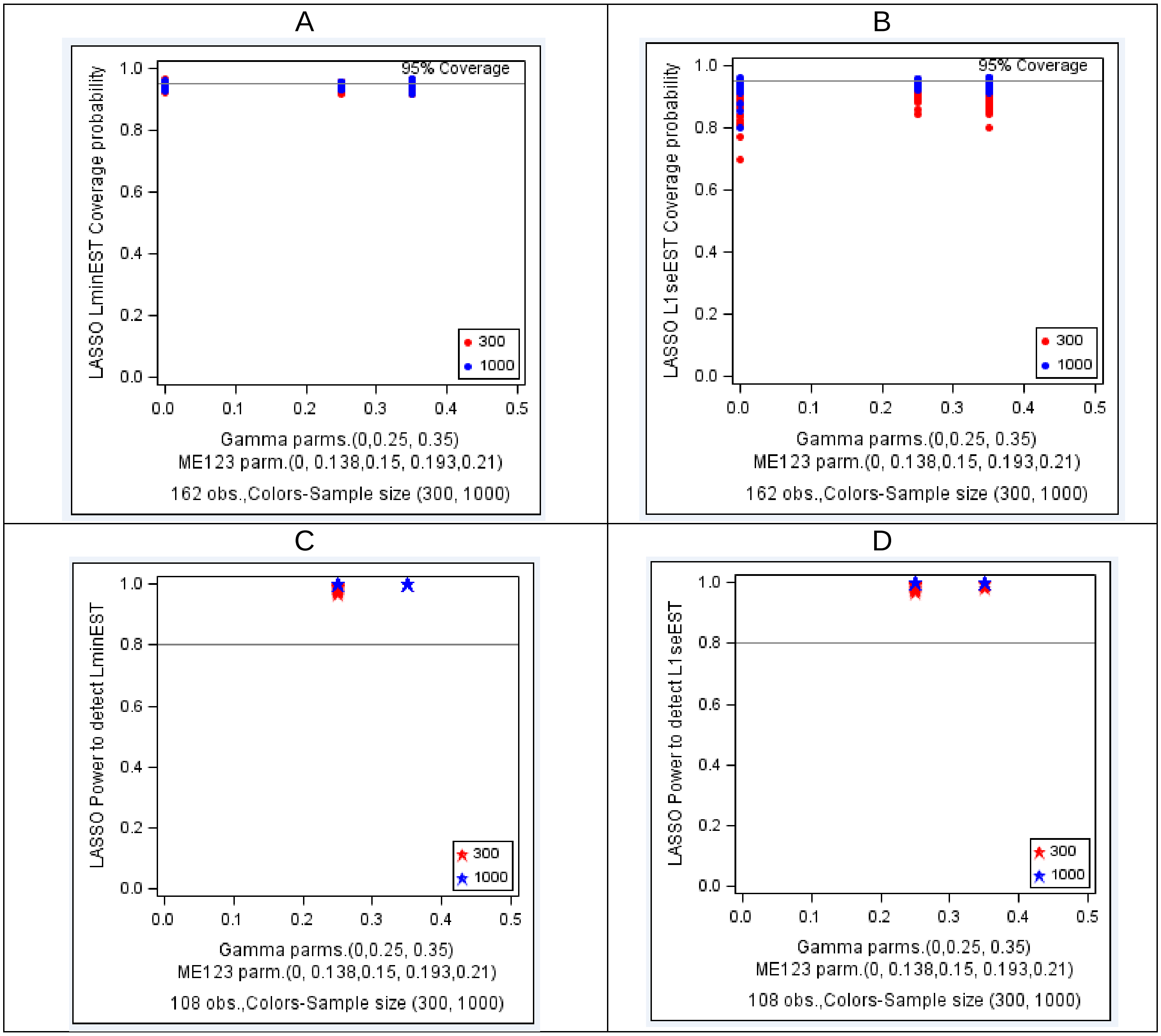


Figure 4.89 LASSO Coverage, Power A&C) $ME_{123}^{L_{\min}^{Cov.}}$, $ME_{123}^{L_{\min}^{PWR.}}$ B&D) $ME_{123}^{L_{\min+1s.e}^{Cov.}}$, $ME_{123}^{L_{\min+1s.e}^{PWR.}}$

The LASSO joint mediated effect's type1 error rate for $ME_{123}^{L_{\min}^{Typ1}}$ and $ME_{123}^{L_{\min+1s.e}^{Typ1}}$ with the joint theta parameter $\theta_{123}(0.60, 0.55, 0.55)$ on the x-axis grouped by sample size are presented in Figure 4.90 A and B.

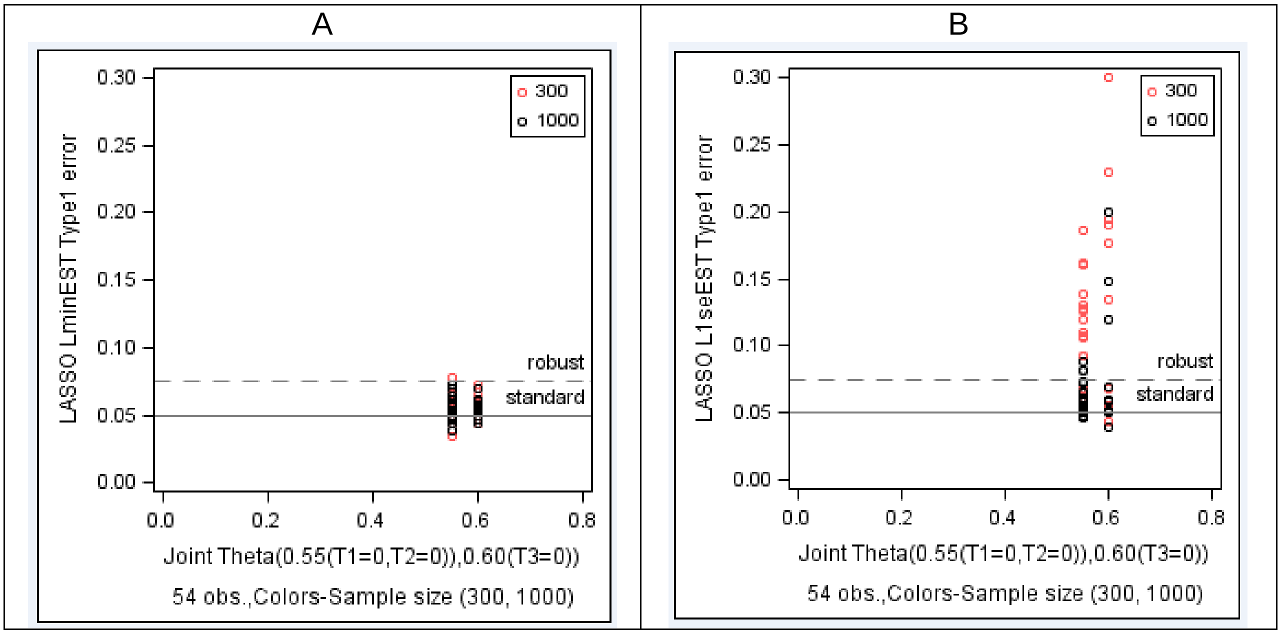


Figure 4.90 LASSO Type1 Error Rate for A) $ME_{123}^{L_{\min}^{Typ1}}$ and B) $ME_{123}^{L_{\min+1s.e}^{Typ1}}$

There was one exception amongst the possible 54 conditions where $ME_{123} = 0$, having a type1 error (0.078) exceeding the *a priori* set limit of 0.075. For this exception shown in Figure 4.90 A, the sample size was small, and the data had low pairwise correlations. Overall, the $ME_{123}^{L_{\min}}$ method utilizes the type1 error limit well and has a high power for detecting the joint indirect effect with high coverage probability. However, the standard error for the smaller sample sized estimates are high especially for $ME_{123}^{L_{\min+1s.e}^{Typ1}}$. The $L_{\text{min+1s.e.}}$ method is not preferred since it has 26 of the 54 conditions that were exceptions (type1 error >0.075), ranging from 0.08 to 0.30.

4.3.3.3 LASSO Individual Mediated effects - *Estimates, Bias & RMSE*

The influencing variables that determine X_1's mediated effect using the LASSO method are: the gamma parameter $\gamma(0, 0.25, 0.35)$ and the joint theta parameter $\theta_{123}(0.55, 0.55, 0.60)$ shown in Figure 4.91

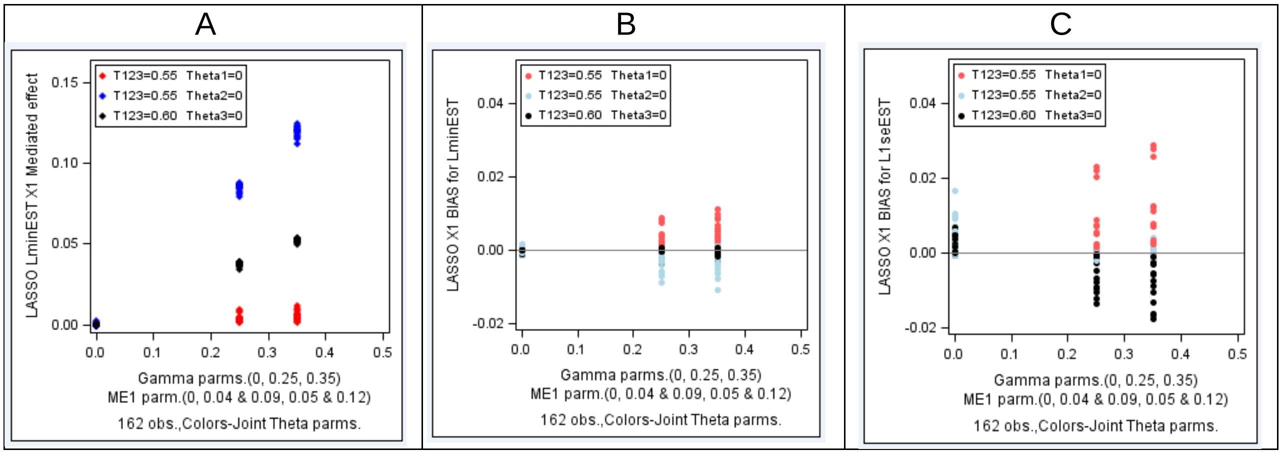


Figure 4.91 LASSO Individual Predictor Estimate, Bias for A) $ME_{1,2,3}^{L_{\min}\,Est.}$ B&C) $ME_{1,2,3}^{L_{\min},L_{\min+1s.e.}\,Bias}$

Higher values for gamma increase the indirect effect $ME_{1\,LASSO}^{Estimate}$ for all nonzero theta parameter values. The individual mediated effect estimates are grouped based on their parameter values for $X_1(\theta_1=0,\text{red}), X_2(\theta_2=0,\text{blue}),\text{and } X_3(\theta_3=0,\text{black})$, which markedly influence the estimate groupings. The individual mediated effect's estimate ($ME_{1,2,3}^{L_{\min}}$ in A, D, G) and the estimate's *Bias* (B, E, H - $ME_{1,2,3}^{L_{\min}Bias}$), and (C, F, I - $ME_{1,2,3}^{L_{1se}Bias}$) are shown in Figure 4.91 A-I. Panel A, D and G are comparable plots for the individual estimates. *ME₁* estimates relate to predictor *X₁*, and its color coded estimates for theta=0, lie along the x-axis. The corresponding plots B, E & H for $ME_{1,2,3}^{L_{\min}Bias}$ show that the LASSO method provides a fairly unbiased estimate (bias range: ±0.01) for the estimates true values ranging from 0 to 0.16. Individual mediated effect estimates for the zero valued theta are positively biased (for $\gamma \neq 0$ i.e. $ME_{1,2,3} \neq 0$), while the other non-zero theta valued mediated estimates (for $\gamma \neq 0$) are negatively biased in the LASSO multiple regressions shown in plots B, E, and H, *i.e.* the null parameter effects are positively biased while non-zero parameter effects are negatively biased regressing towards the mean effect, for all the predictors.

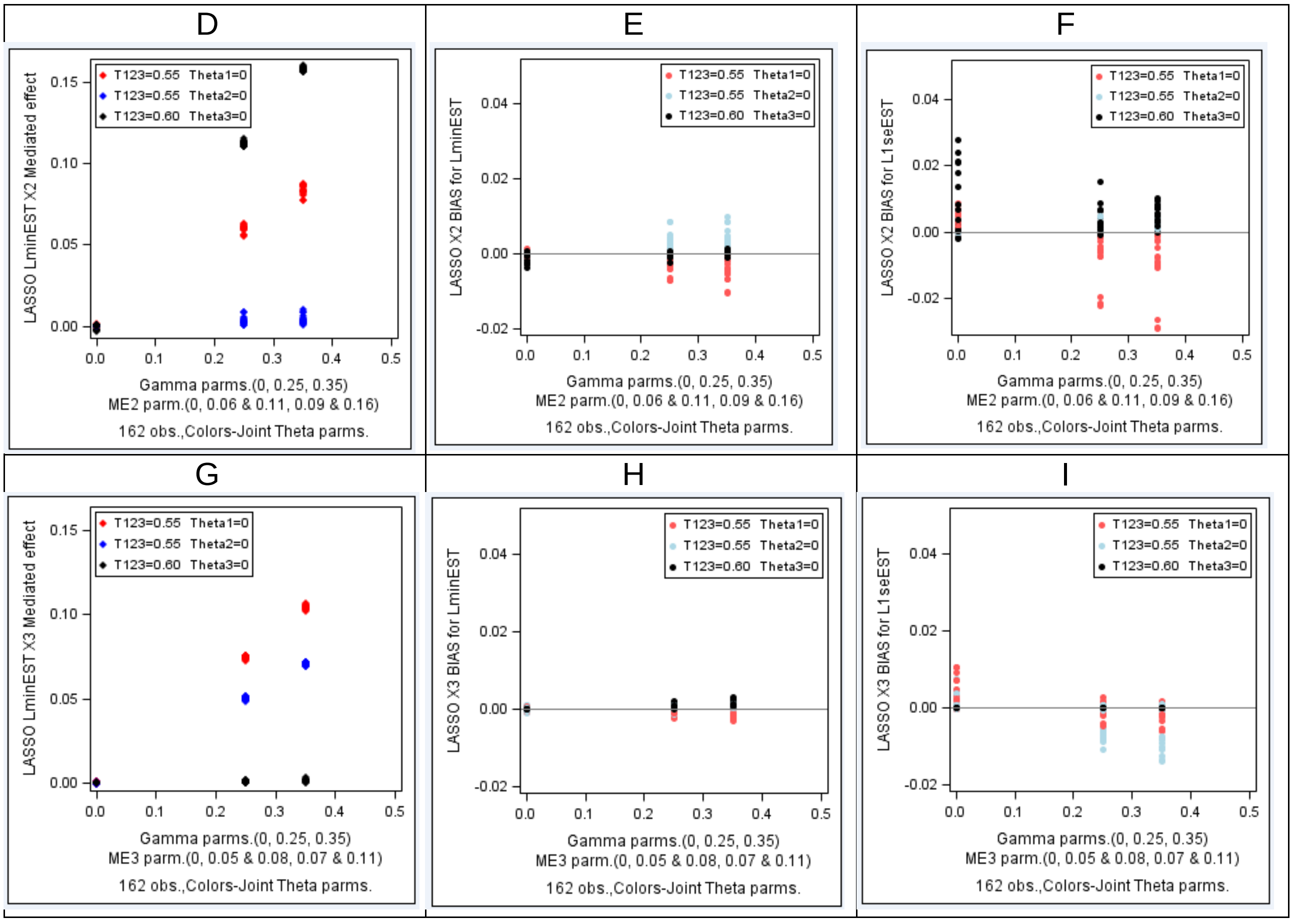
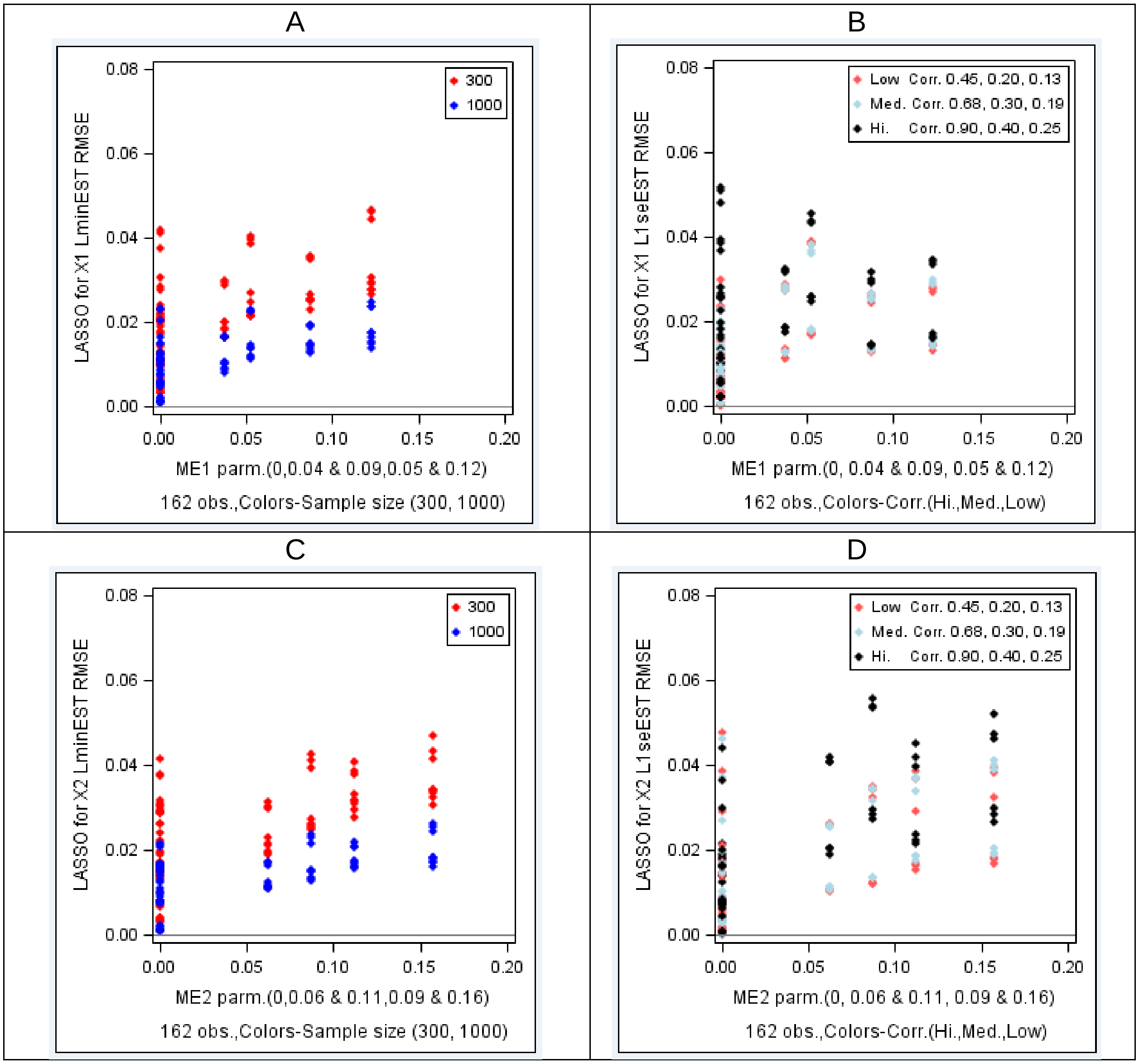


Figure 4.91 Indiv. $ME_{1,2,3}^{L_{\min}}$ A, D, G)- Estimates & Bias B, E, H)- $ME_{1,2,3}^{L_{\min}Bias}$ C, F, I) - $ME_{1,2,3}^{L_{1se}Bias}$

The mediated effects for individual predictors shown in Figure 4.91 A, D and G have standard errors for their estimates which are discussed next in Figure 4.92. The magnitude of the *RMSE* is large for the values of the individual mediated estimate indicating large variability. The low bias indicates an almost unbiased average estimate. A mediated estimate from a single dataset is of no practical value because of the high variability of the LASSO estimate. The factors influencing the estimate's *RMSE* are: correlation between the individual predictors (panels B, D, and F), the estimate's true value, and the sample size.


A
LASSO for X1 LminEST RMSE
300
1000
ME1 parm.(0,0.04 & 0.09,0.05 & 0.12)
162 obs.,Colors-Sample size (300, 1000)
B
LASSO for X1 L1seEST RMSE
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
ME1 parm.(0, 0.04 & 0.09, 0.05 & 0.12)
162 obs.,Colors-Corr.(Hi.,Med.,Low)
C
LASSO for X2 LminEST RMSE
300
1000
ME2 parm.(0,0.06 & 0.11,0.09 & 0.16)
162 obs.,Colors-Sample size (300, 1000)
D
LASSO for X2 L1seEST RMSE
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
ME2 parm.(0, 0.06 & 0.11, 0.09 & 0.16)
162 obs.,Colors-Corr.(Hi.,Med.,Low)

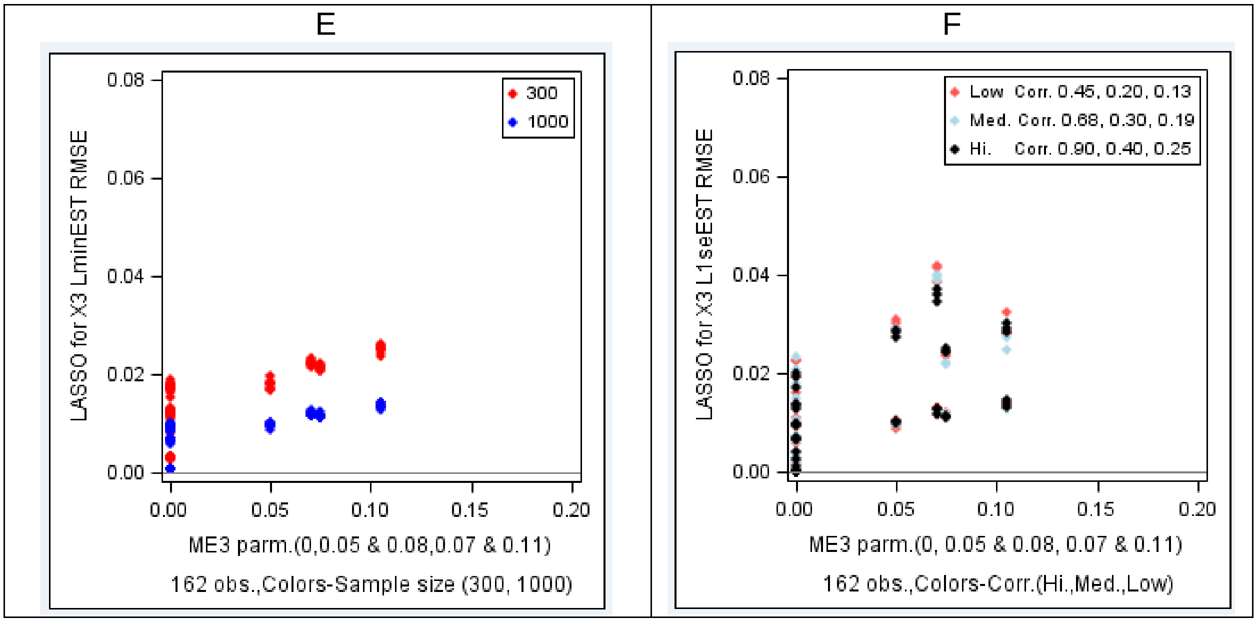
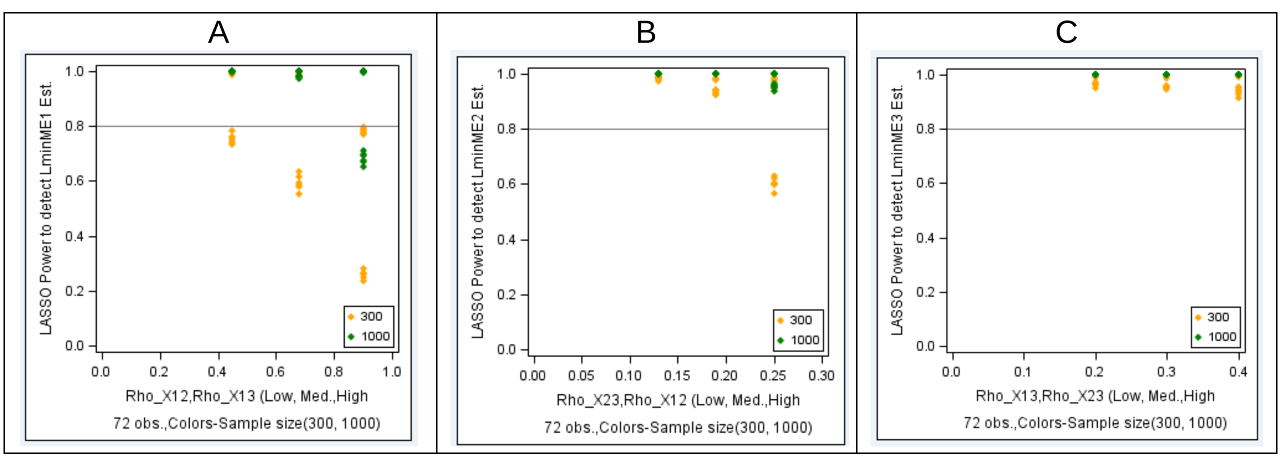


Figure 4.92 $ME_{1,2,3}^{\overset{RMSE}{L_{\min}}}$ A, C, E) grouped by *N* and D, B, F) $ME_{1,2,3}^{\overset{RMSE}{L_{\min+1s.e.}}}$ grouped by *Corr.*

4.3.3.4 LASSO Individual Mediated Effects – Power, Coverage and Type1 Error

The power for LASSO to detect an individual mediated effect is influenced by three parameter values listed in decreasing order of influence; the pairwise correlations (high correlations result in lower power), the sample size (small sample sizes result in lower power), shown in Figure 4.93 A, B, and C, and by the individual mediated effect sizes (smaller values have lower power), shown in D, E, and F.



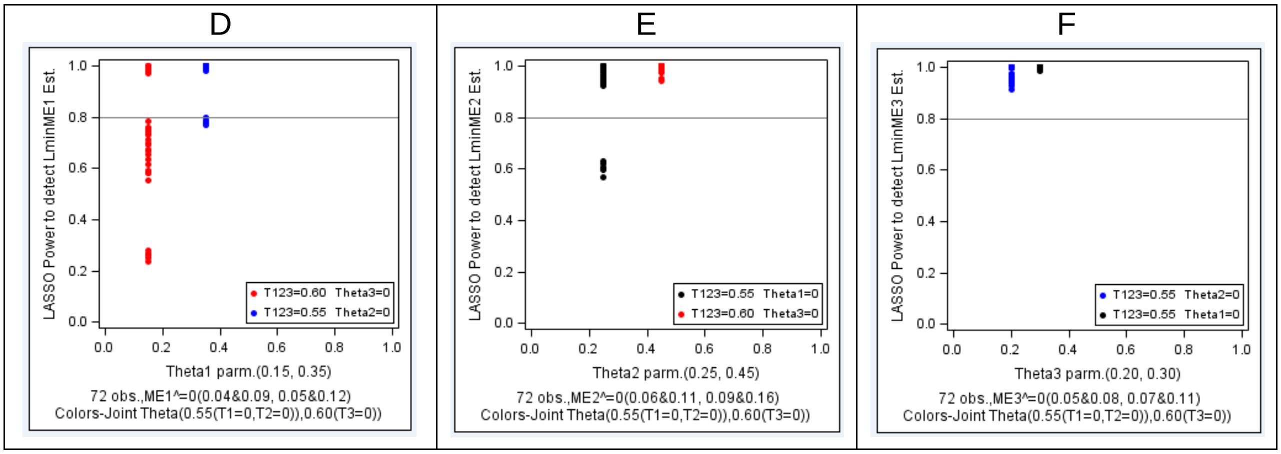
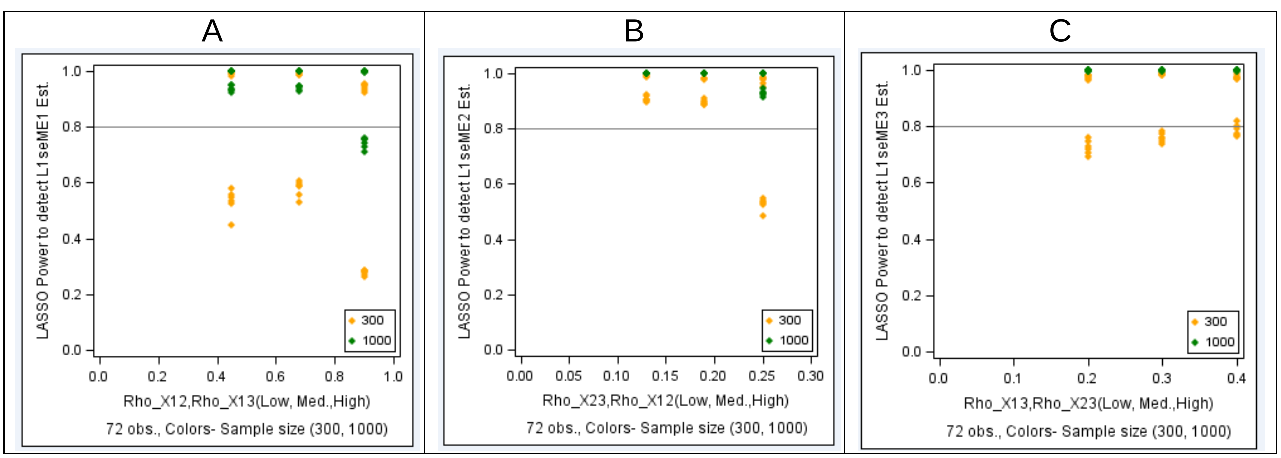


Figure 4.93 $ME^{Power}_{L_{\min}\,1,2,3}$ A, B, C) Power by *Corr.* D, E, F) $ME^{Power}_{L_{\min}\,1,2,3}$ Power by θ_1, θ_2, θ_3

A combination of high correlation, low sample size(300) results and small effect size (0.0375) results in the lowest power (0.21) to detect an individual effect using the L_{min} method are shown in Figure 4.93 A and D The power increases with lower pairwise correlations and larger sample sizes. Figure 4.93 A-F support the conclusion that the LASSO L_{min} method performs better when detecting larger effects in large samples, and having low pairwise predictor correlations.



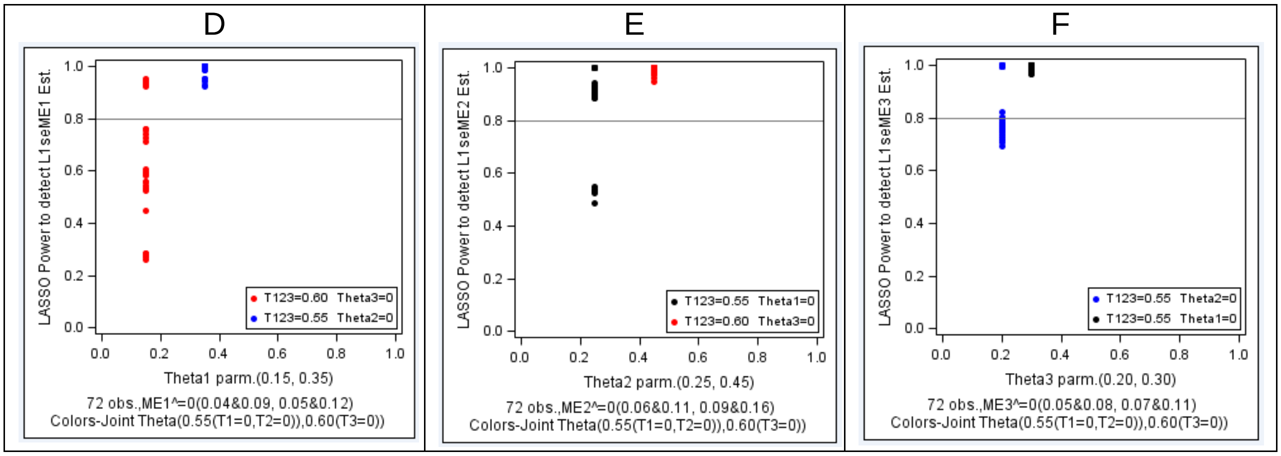


Figure 4.94 $ME_{1,2,3}^{\overset{Power}{L_{\min+1s.e.}}}$ A, B, C) Power by *Corr.* D, E, F) $ME_{1,2,3}^{\overset{Power}{L_{\min+1s.e.}}}$ Power by $\theta_1,\ \theta_2,\ \theta_3$

The corresponding results for $ME_{1,2,3}^{\overset{Power}{L_{\min+1s.e.}}}$ in comparison to $ME_{1,2,3}^{\overset{Power}{L_{\min}}}$ shown in Figure 4.93, have reduced power, and are shown in Figure 4.94 A-F compared to Figure 4.93 A-F. The lower powered conditions have high correlated predictors in small sample sizes detecting small mediated effects (0.04, 0.05), shown for predictor X_1 in Figure 4.94 A and D. This further reinforces the decision to use the $L_{\min}$ and not the $L_{\min+1s.e.}$ method, in LASSO regressions for mediation analysis.

The individual LASSO mediated effect coverage probabilities $ME_{1,2,3}^{\overset{Cov.}{L_{\min}}}$ and $ME_{1,2,3}^{\overset{Cov.}{L_{\min+1s.e.}}}$ with the $L_{\min}$ and $L_{\min+1s.e.}$ methods are discussed using Figure 4.95. The mediated effect's coverage probability using the $ME_{1,2,3}^{\overset{Cov.prob.}{L_{\min}}}$ method has 6 exceptions for the 162 conditions (blue cluster at 0.80) in Figure 4.95 A, having a small sample size and high pairwise correlations, covering a small mediated effect size (0.04) and effect size of 0.05 in Figure 4.95 D. However, the $ME_{1,2,3}^{\overset{Cov.prob.}{L_{\min.+1s.e.}}}$ method shown in Figure 4.95 D-F have worse coverage probability than the $ME_{1,2,3}^{\overset{Cov.prob.}{L_{\min}}}$ shown in Figure 4.95 A-C.

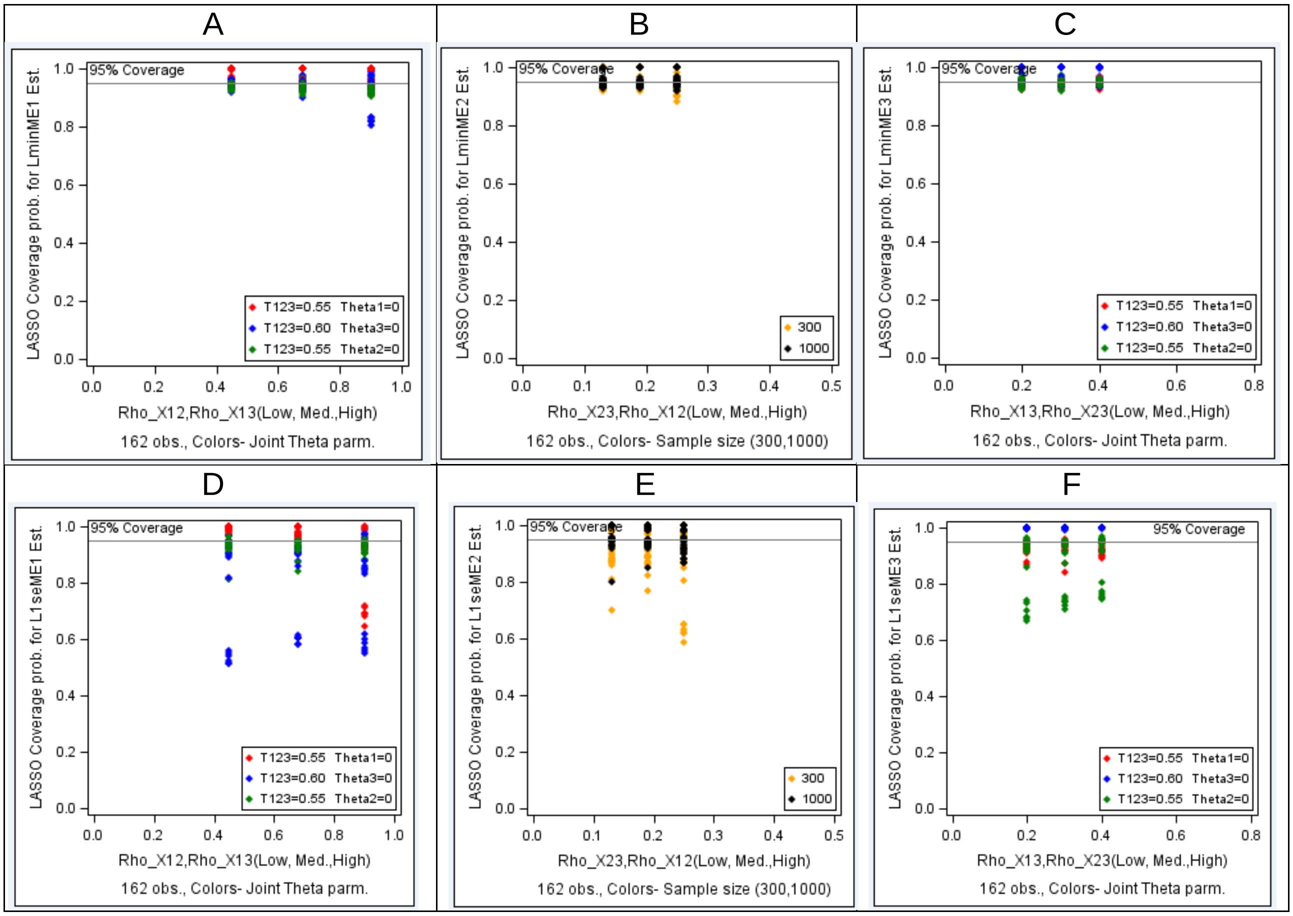


Figure 4.95 Individual coverage A, B, C) $\overset{Cov.prob.}{ME_{1,2,3}^{L_{\min}}}$ and D, E, F) $\overset{Cov.prob.}{ME_{1,2,3}^{L_{\min+1s.e.}}}$ by *Corr.*, θ_{123}, & *N*

The individual mediated effect's type1 errors when using the LASSO regression methods in 3-variable mediation analysis are shown in Figure 4.96 A-F. There were 90 conditions where $ME_1 = 0$ for type1 error consideration in each figure. There were 54 conditions where $\gamma = 0, \theta_1 = (0, 0.15, 0.35)$ with eighteen conditions for each of the theta values and thirty-six additional conditions with non-zero gamma parameters having a null theta parameter. Figure 4.96 A shows four exceptions for the 90 conditions, two exceptions each for $\gamma = (0.25, 0.35)$, with type1 errors (0.08 to 0.09) exceeding the limit of 0.075. The sample sizes were small and the pairwise correlations were high for these exceptions.

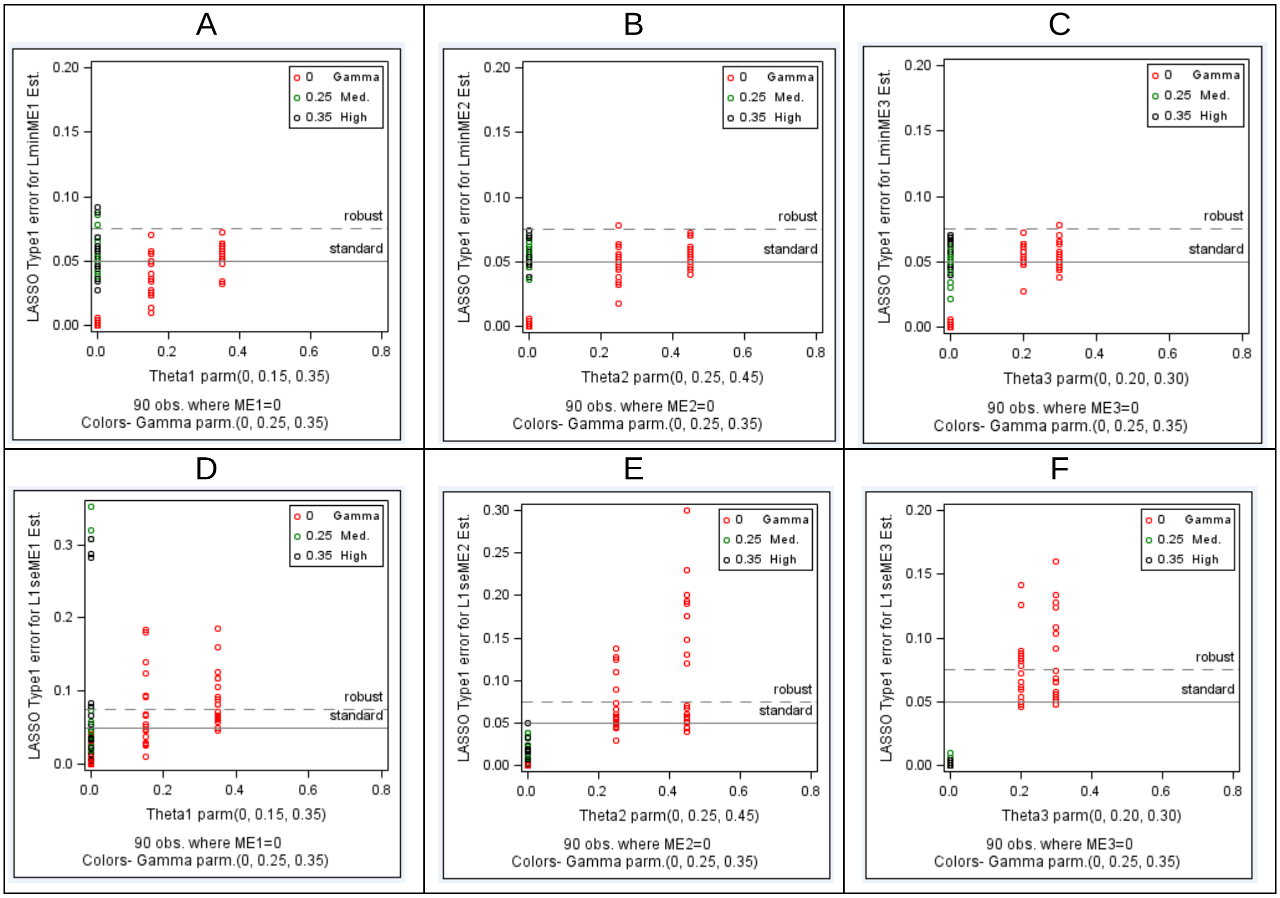


Figure 4.96 Individual Type1 Err. by θ_1, θ_2, θ_3 A, B, C) $ME_{1,2,3}^{\overset{Type1\ err.}{L_{\min.}}}$ D, E, F) $ME_{1,2,3}^{\overset{Type1\ err.}{L_{\min+1s.e.}}}$ by γ

There were 90 conditions where $ME_2 = 0$ for type1 error consideration. One exception (0.078) was noted and shown in Figure 4.96 B. There were 90 conditions where $ME_3 = 0$ for type1 error consideration. One exception (0.078) was noted and shown in Figure 4.96 C. The type1 error rates using the $L_{min+1s.e.}$ method were higher for each individual mediated effect and the plots are shown for comparison in Figure 4.96 D-F. Overall, the L_{min} method for LASSO regression uses the type error rate limits (0.05 to 0.075) efficiently for each of the individual mediated effects.

4.3.3.5 Joint & Individual Direct Effects for 3-Variable Mediation Analysis

The *Estimate*, *Bias* and *RMSE* for joint direct effects are discussed next. LASSO method estimates the joint direct effect with a small bias (-0.005, 0.015) and a standard error for the estimate of 0.03 (large *N*) to 0.09(small *N*), given the true value of the joint direct effect to be 0.35, 0.45, and 0.50 are shown in Figure 4.97 A and D. Plot F shows that larger sample sizes and higher correlations reduce the joint direct estimate's *RMSE*.

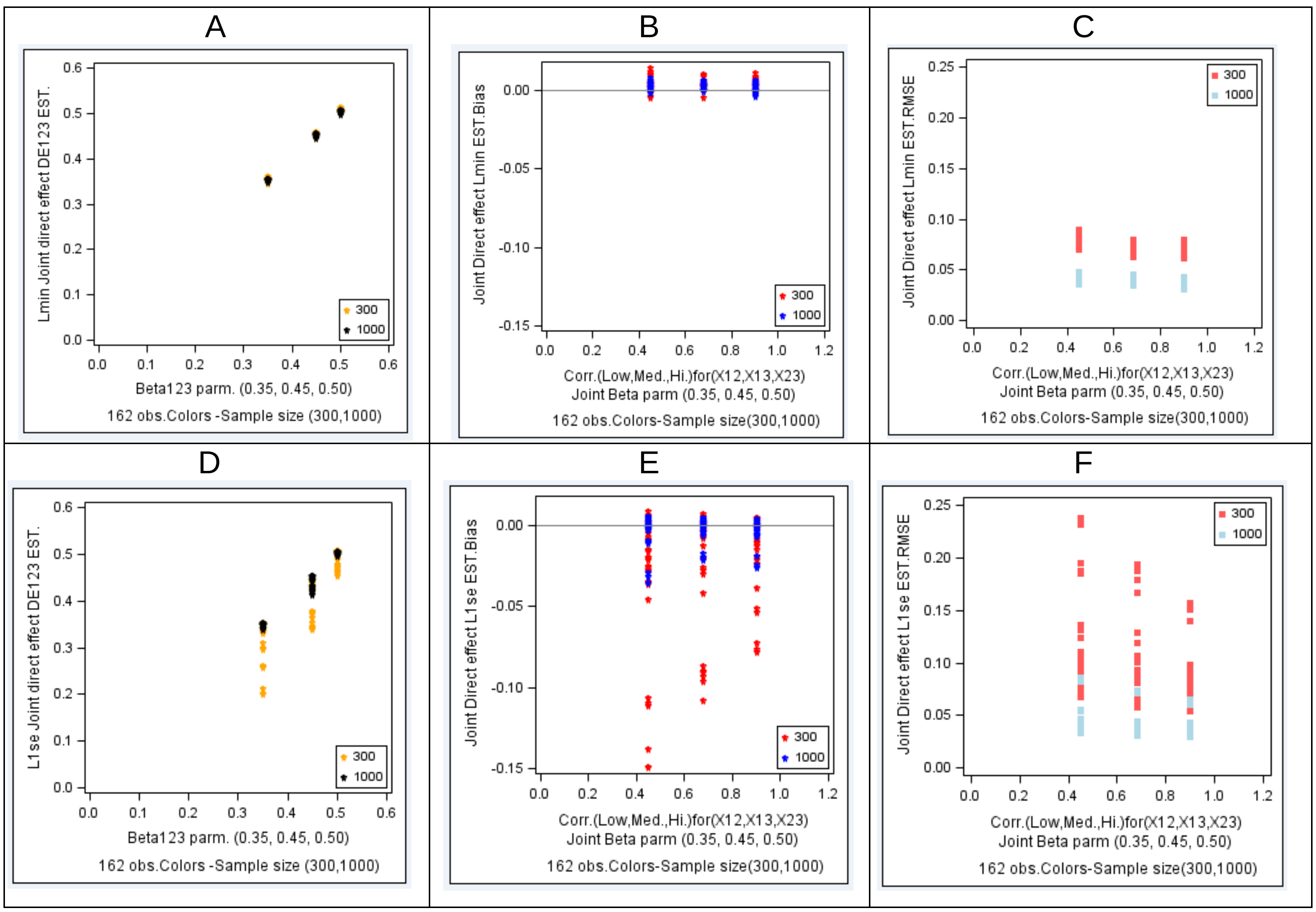


Figure 4.97 LASSO Joint Direct Effect L_{min}(A,B,C) *Estimate,Bias &RMSE*;$L_{min+1se}$(D,E,F)

4.3.3.6 LASSO Joint Direct Effects' Type1 Error, Power & Coverage Probability

The LASSO method's type1 error for estimating a null joint direct effect could not be evaluated since $\beta_{123} = (0.35, 0.45, 0.50)$ and $DE_{123} \neq 0$, for all 162 conditions.

There were 8 conditions (power range: 0.63 to 0.78) out of the possible 162 conditions for the LASSO method detecting a joint direct effect $DE_{123}^{L\min}$ having a low power, and the common influencing variables were: small sample size and low joint beta parameter values (0.35, black) to be detected, shown in Figure 4.98 A and B. The pairwise correlations work differently for the two LASSO methods: $DE_{123}^{L_{\min}}$'s power decreases and for $DE_{123}^{L_{\min+1s.e.}}$'s power increases, with increasing pairwise correlations. The joint direct effect's coverage probability for all 162 conditions exceeded 0.95 for the L_{min} method but the $L_{min+1s.e}$ had a much poorer coverage probability (0.49 to 1).

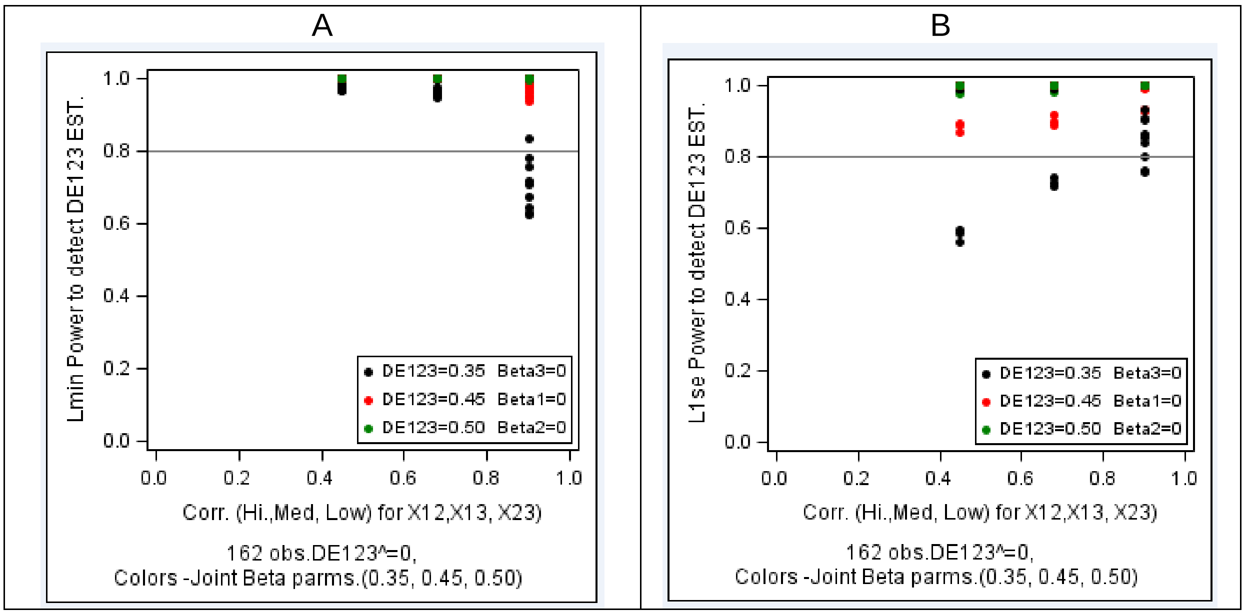


Figure 4.98 LASSO Method's Joint Direct Effect Power A) $DE_{123}^{L\min}$ & B) $DE_{123}^{L\min+1s.e.}$

4.3.2.7 LASSO 3-Variable Individual Mediated Direct Effect's *Estimate*, *Bias* & *RMSE*

The influencing variable that determines X_1's direct effect using the LASSO method is the beta parameter value associated with the independent variable as shown in Figure 4.99. Higher values for gamma and sample size reduce the

variability for the estimated direct effect. Neither sample size nor pairwise correlations of (high, medium, low) for (X_{12}, X_{13}, *and* X_{23}), markedly influence the individual direct estimate. The LASSO estimates have high standard errors, when correlations are high and the sample sizes are small. Comparing the L_{min} vs. $L_{min+1s.e.}$ methods for individual predictors X_1's, X_2's, X_3's estimate, *Bias* and *RMSE* in Figure 4.99 A vs. D, B vs. E, and C vs. F, and Figure 4.100 G vs. J, H vs. K, and I vs. L, and Figure 4.101 M vs. P, N vs. Q, and O vs. R respectively, the conclusion is that the L_{min} method has more precision in its estimate with lesser *Bias* and smaller *RMSE* for the average estimated individual direct effects than the $L_{min+1s.e.}$ method.

The corresponding plots B, H, and N for $DE_{1,2,3}^{\underset{L\min}{Bias}}$ show that the LASSO method provides an estimate bias in the range: -0.03 to 0.03, for the estimate's true value range: 0 to 0.30. The beta parameter value associated with the individual predictor, combined with the predictor's pairwise correlation strength, determines the estimate's bias. The average mediated effects for individual predictors shown in Figures 4.99 to 4.101 A, G & M have standard errors for the average estimates, which are shown in Figures 4.99 to 4.101 C, I, & O. Contrary to the joint direct effect's *RMSE* (Figure 4.97 C) the individual predictor *RMSEs* markedly increase with higher correlation between the predictors (Figures 4.99 and 4.101 C and I. Larger sample sizes consistently reduce the standard error of the estimates.

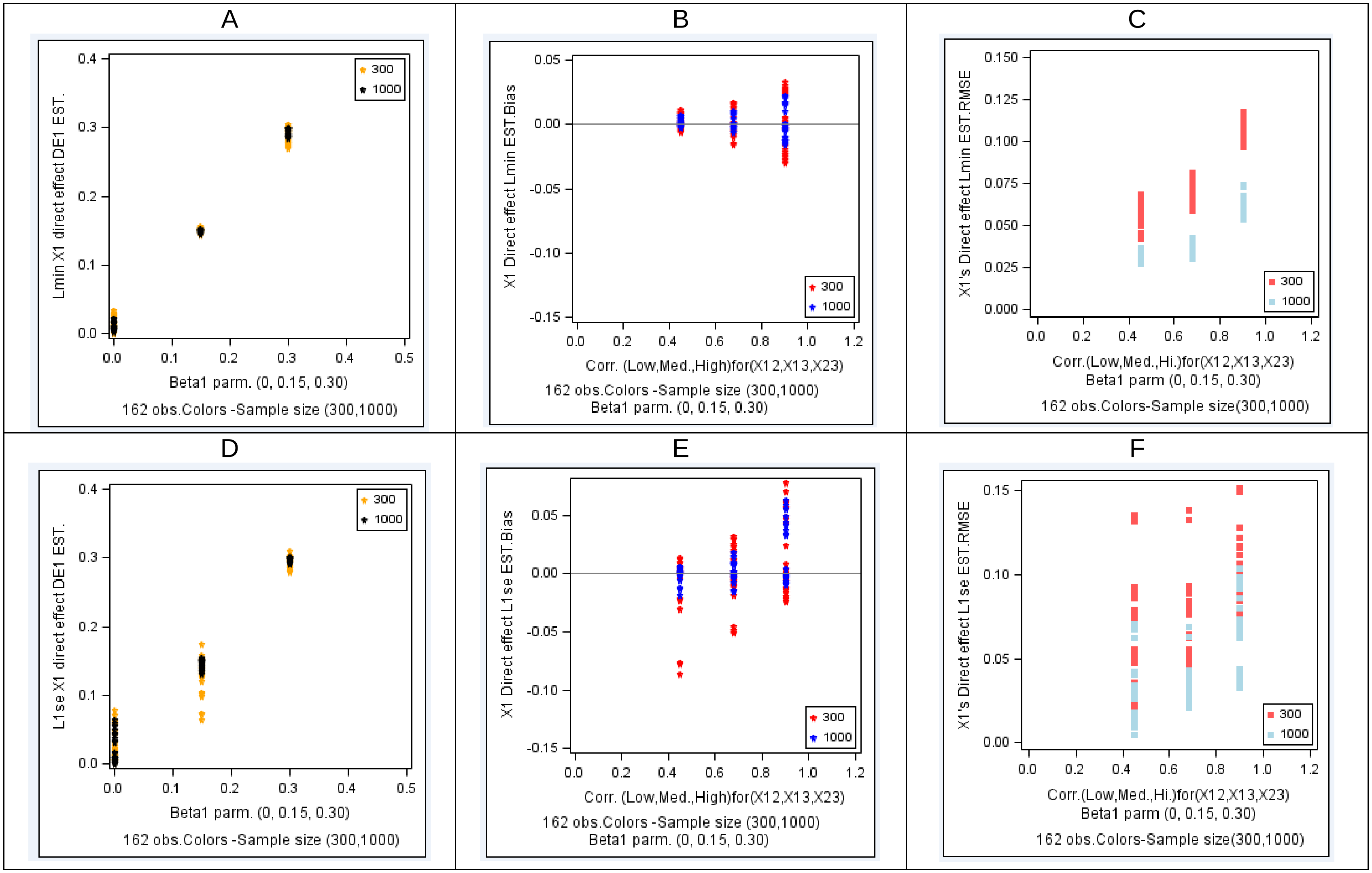
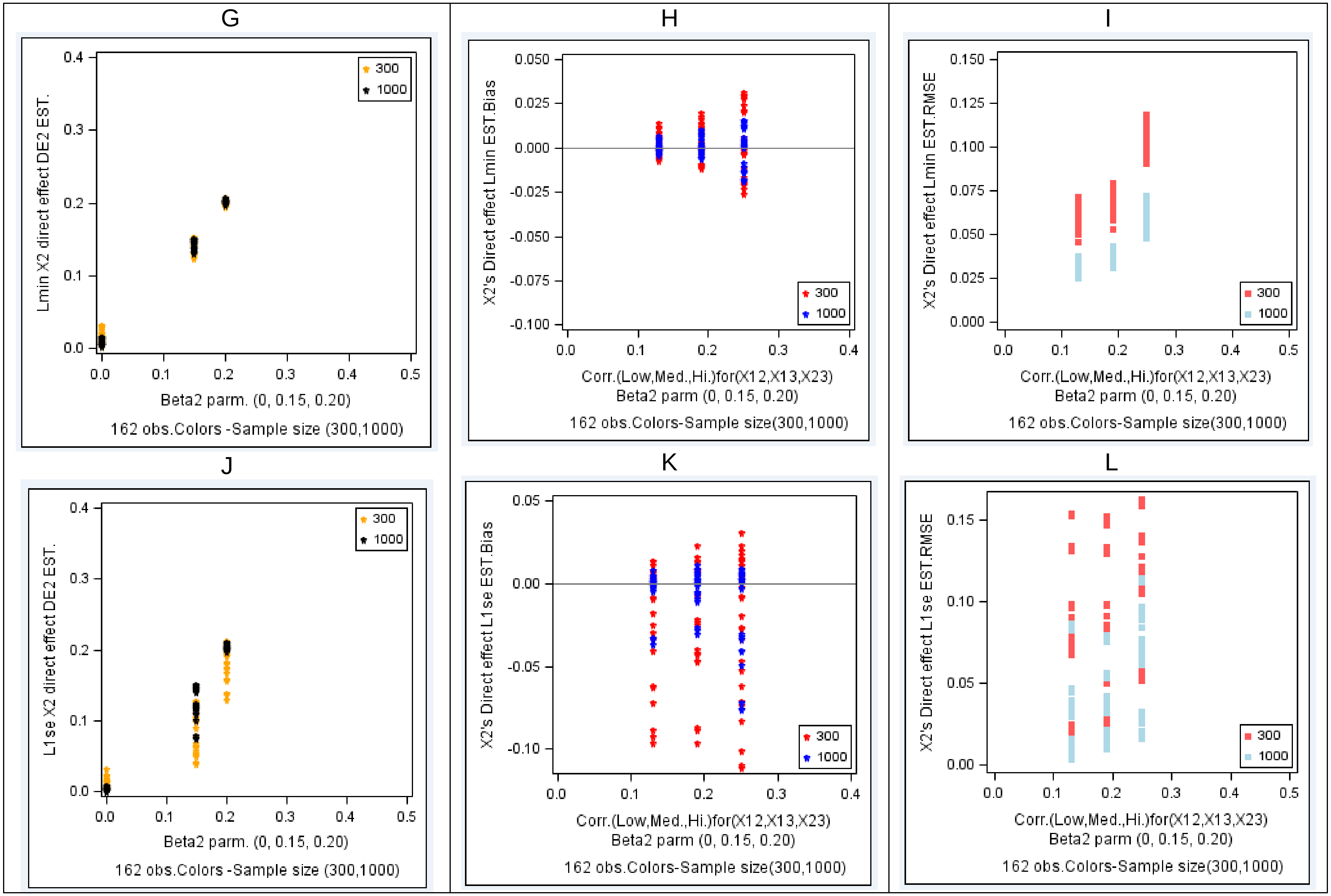
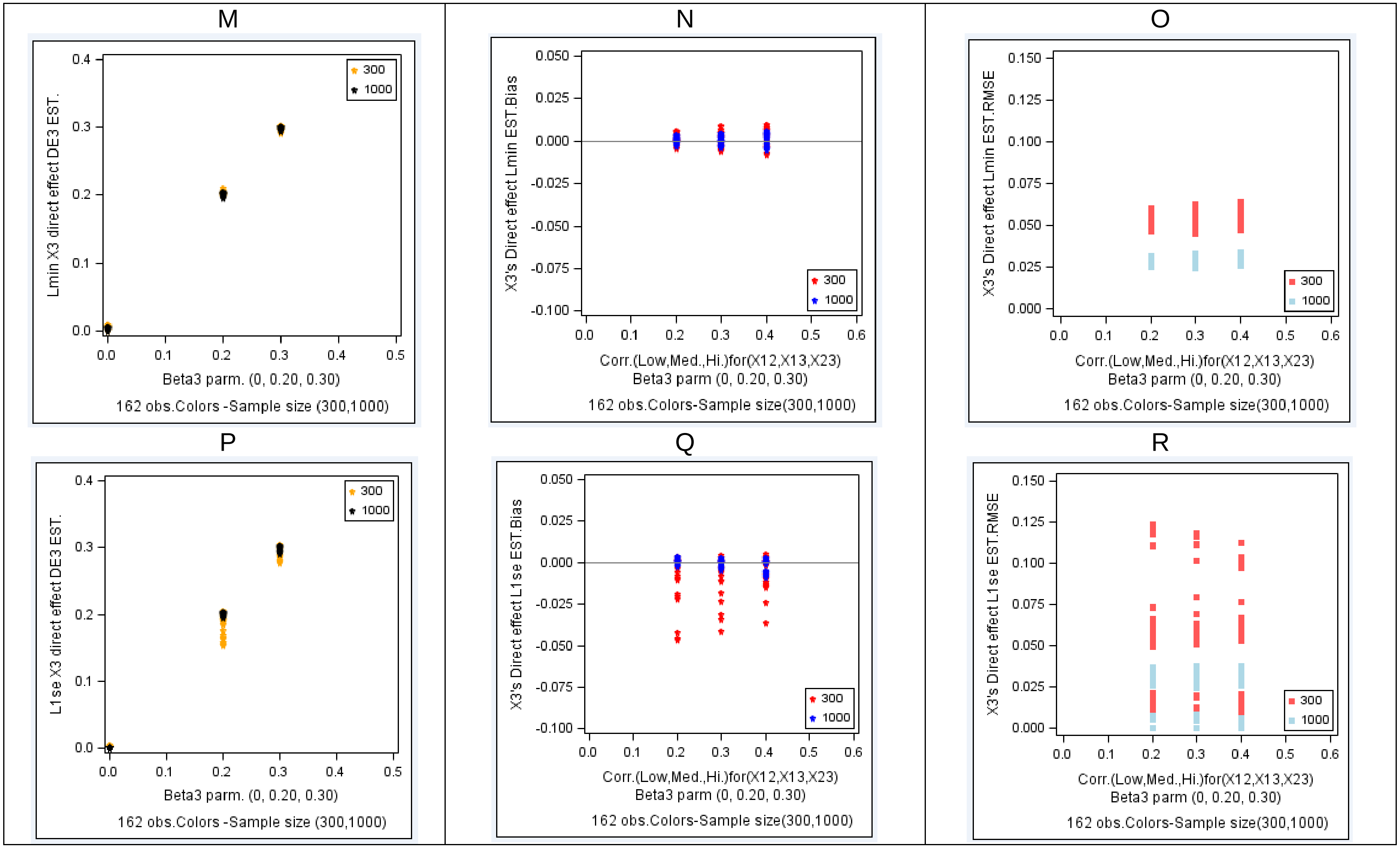


Figure 4.99 LASSO Indiv. Direct Effect L_{min} (A, B, C) *Estimate, Bias & RMSE*; $L_{min+1se}$ (D, E, F) for X_1



Figures 4.100 LASSO Indiv. Direct Effect L_{min} (A, B, C) *Estimate, Bias & RMSE*; $L_{min+1se}$ (D, E, F) for X_2



Figures 4.101 LASSO Indiv. Direct Effect L_{min} (A, B, C) *Estimate, Bias & RMSE*; $L_{min+1se}$ (D, E, F) for X_3

4.3.2.8 LASSO 3-Variable Individual Direct Effect's Type1 Error, Power and Coverage

The LASSO individual direct effect's type1 error rate with the gamma parameter values (0, 0.25, and 0.35) on the x-axis and color coding based on the increasing predictor pairwise correlations are shown in Figure 4.102 for each set of 54 conditions having a type1 error rate by the individual predictor X_1, X_2, and X_3. There were 14 exceptions (type 1 error > 0.075) amongst the 54 conditions where $DE_1 = 0$, having type1 errors ranging between 0.078 and 0.184. The joint beta parameters for the three plots A, B & C were $\beta_{123}(0.45(\beta_1 = 0),\ 0.50(\beta_2 = 0),$ and $0.35(\beta_3 = 0))$. The 14 exceptions were due to high correlations between X_1, X_2 (0.9) and X_1, X_3 (0.4).

There were zero type1 error rate exceptions for the 54 conditions shown in Figure 4.102 B where DE_2=0 and DE_3=0 because of the lower correlations between the predictors (X_3, X_2) 0.25 & (X_3, X_1) 0.4, and the smaller joint direct effect size (0.35). The $DE_{1,2,3}^{L_{\min}^{TYP1}}$ method performs poorly for highly correlated predictors having smaller sample sizes when the effect being detected is small. Figure 4.102 D-F for $DE_{1,2,3}^{L_{\min+1s.e.}^{TYP1}}$ perform worse than for $DE_{1,2,3}^{L_{\min}^{TYP1}}$ in Figure 4.102 A-C, having 20 exceptions for DE_1=0 out of 54 conditions with type1 errors ranging from 0.12 to 0.44.

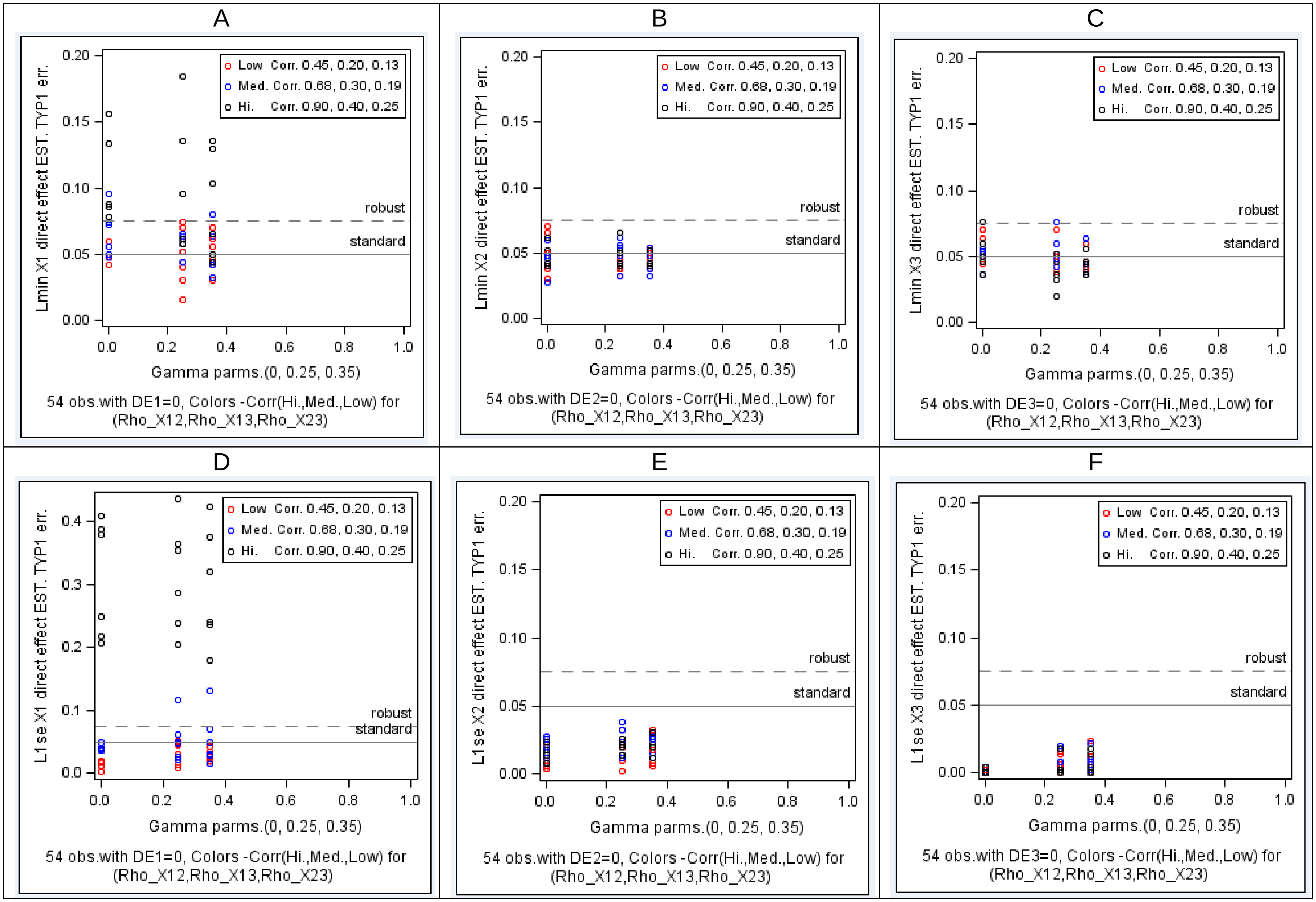


Figure 4.102 LASSO Method's Direct Effect Type1 Error A-C) $DE_{1,2,3}^{L_{\min}^{TYP1}}$ D-F) $DE_{1,2,3}^{L_{\min+1s.e.}^{TYP1}}$

The LASSO method's power to detect each individual predictor's direct effects is shown in Figure 4.103 A-C) and D-F). The influencing factors that determine the power are inferred from plots A and B as: higher pairwise predictor correlations reduce the LASSO method's power to detect an individual direct effect, smaller sample sizes have less power, and lower individual beta parameter non-zero values for the direct effect being investigated have less power. The two clusters for a given sample size, pairwise correlation and different individual beta parameter values $N = 300, \rho_{12} = 0.90\,\&\,\beta_1 = 0.15$ vs. $N = 300, \rho_{23} = 0.90\,\&\,\beta_1 = 0.30$ are shown in Figure 4.103 A with an average power of 0.2 vs. 0.7, respectively. An overall assessment of the LASSO method's power to detect an individual mediated direct effect is that it is poor for small sample sizes, high predictor correlations, and small effect sizes and that $L_{min+1.s.e}$ method is not preferred. The individual mediated direct effects shown in Figure 4.104 A-C with the lowest coverage probability $DE_{1,2,3}^{L_{\min}}{}^{Cov.prob.}$ =0.74 has a small sample size, high predictor correlations between $\rho_{12} = 0.90,\ \rho_{13} = 0.68,\ \rho_{23} = 0.45$ and small effect size (0.15). The best coverage probability is for predictor X_3, having a larger effect size (0.30) and smaller pairwise correlations $\rho_{23} = (0.13, 0.19, 0.25)$ and shown in Figure 4.104 C.

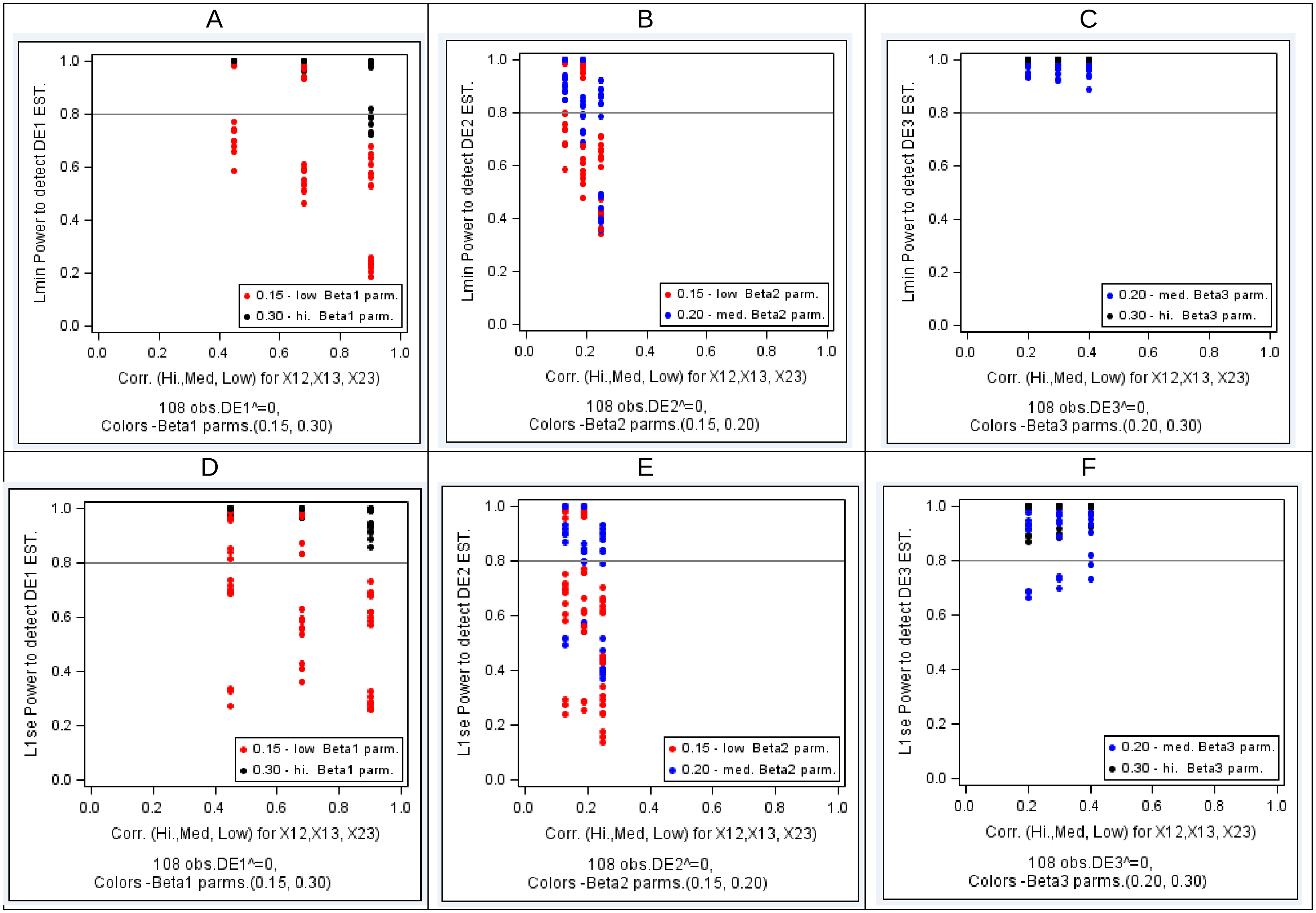
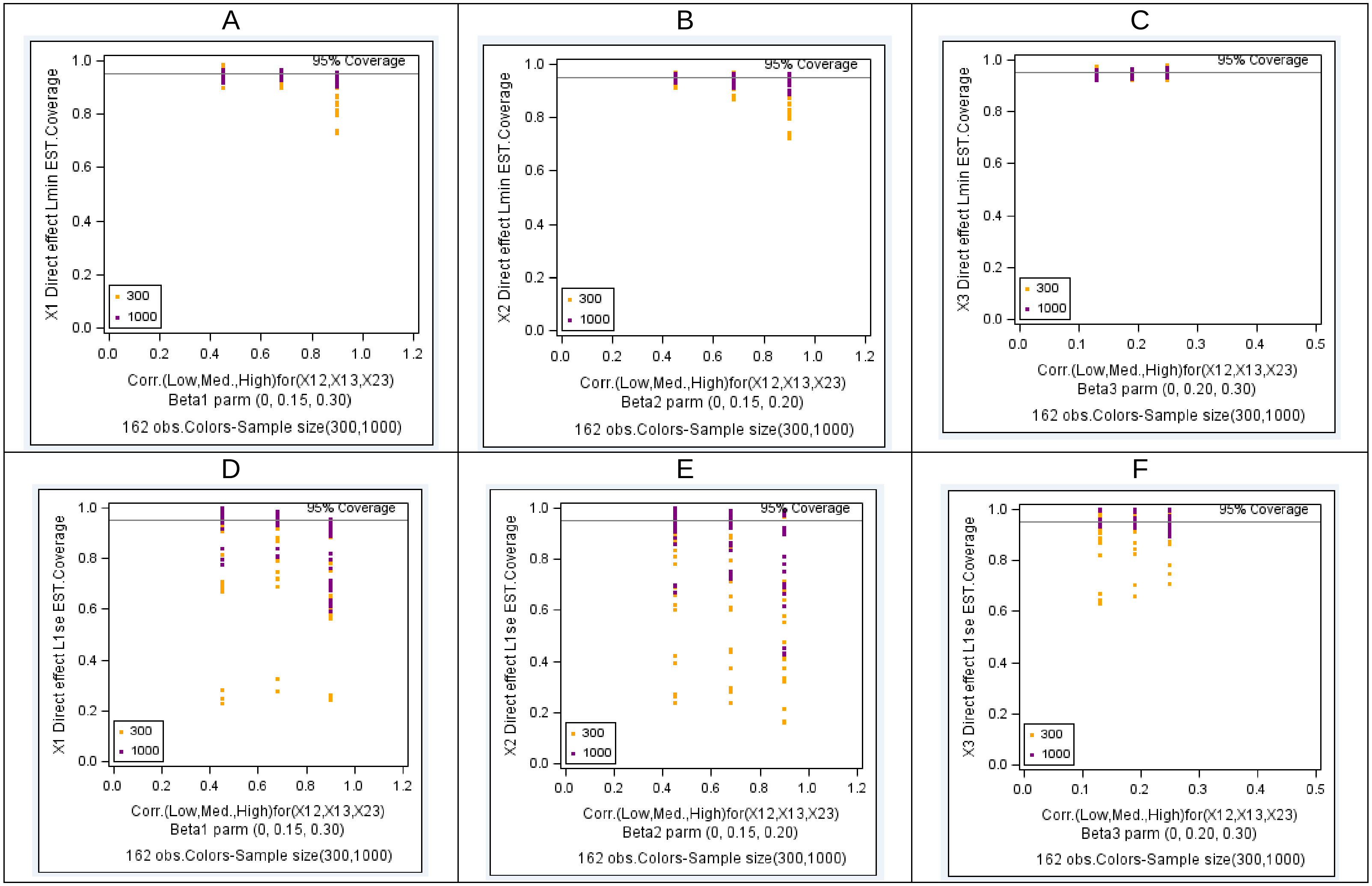


Figure 4.103 LASSO Power for Individual Direct Effects A-C) $DE_{1,2,3}^{L_{\min}^{Power}}$ D-F) $DE_{1,2,3}^{L_{\min+1s.e.}^{Power}}$



A
95% Coverage
X1 Direct effect Lmin EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta1 parm (0, 0.15, 0.30)
162 obs.Colors-Sample size(300,1000)
B
95% Coverage
X2 Direct effect Lmin EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta2 parm (0, 0.15, 0.20)
162 obs.Colors-Sample size(300,1000)
C
95% Coverage
X3 Direct effect Lmin EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta3 parm (0, 0.20, 0.30)
162 obs.Colors-Sample size(300,1000)
D
95% Coverage
X1 Direct effect L1se EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta1 parm (0, 0.15, 0.30)
162 obs.Colors-Sample size(300,1000)
E
95% Coverage
X2 Direct effect L1se EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta2 parm (0, 0.15, 0.20)
162 obs.Colors-Sample size(300,1000)
F
95% Coverage
X3 Direct effect L1se EST.Coverage
300
1000
Corr.(Low,Med.,High)for(X12,X13,X23)
Beta3 parm (0, 0.20, 0.30)
162 obs.Colors-Sample size(300,1000)

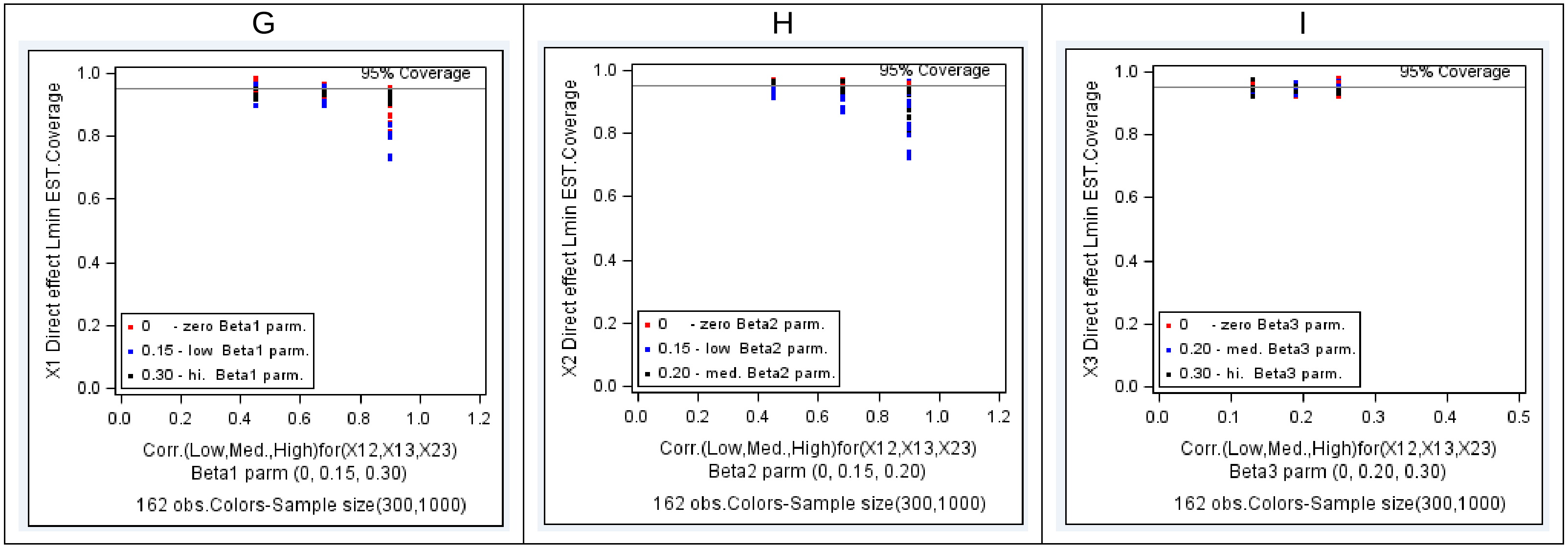


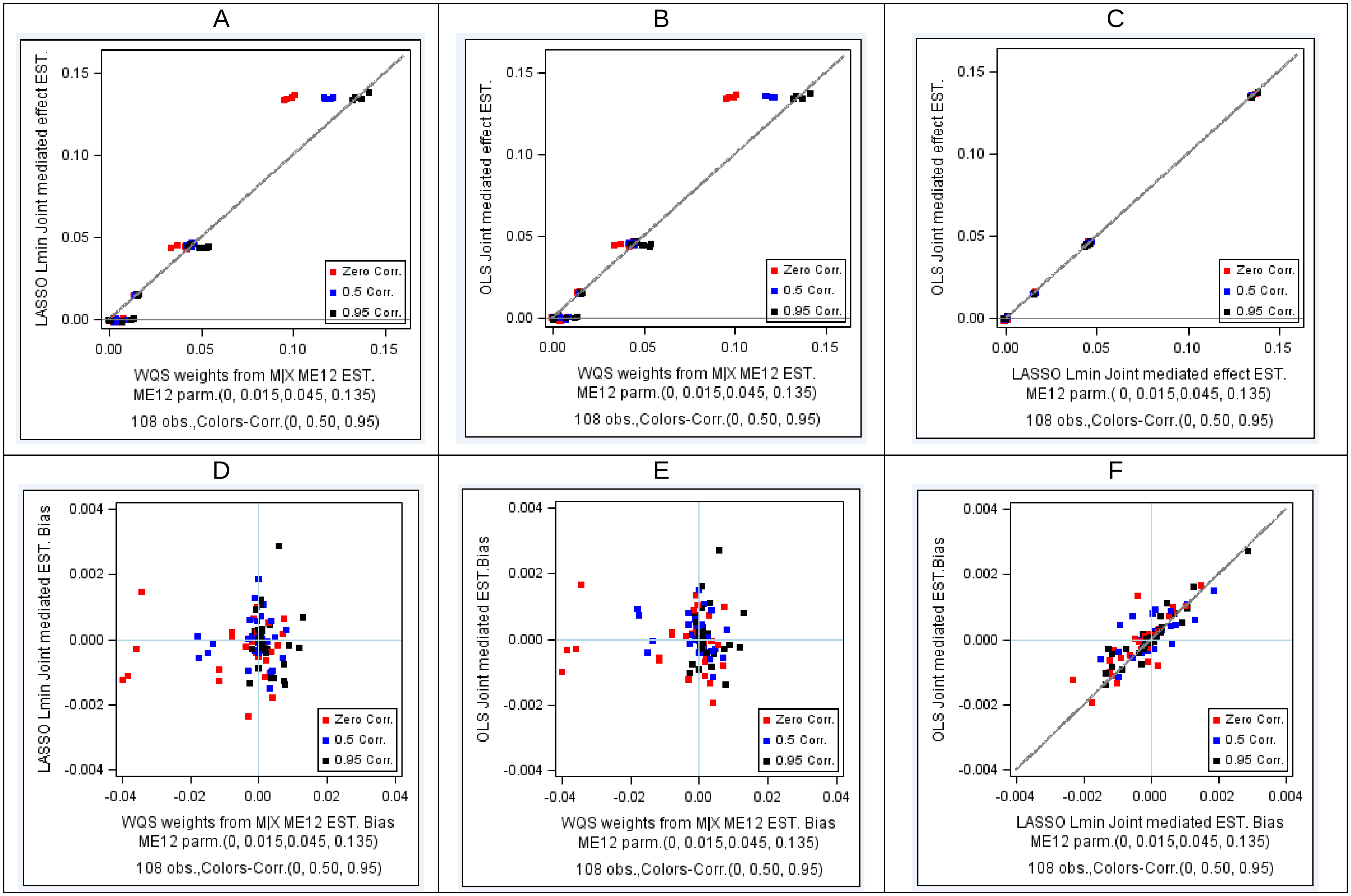
Figure 4.104 Individual Direct Effect A-C, G-I) $DE_{1,2,3}^{L_{\min}\ Cov.prob.}$ by Corr. & (N, $\beta_{1,2,3}$), D-F) $DE_{1,2,3}^{L_{\min+1s.e.}\ Cov.prob.}$

4.4 Pairwise Comparisons of WQS, LASSO, and OLS Method for 2-Var. Mediation

4.4.1.1 Joint 2-Variable Indirect Effect Est. Comparisons: WQS, LASSO & OLS

The 2-variable joint *Estimate*, *Bias*, *RMSE* comparisons for WQS vs. L_{min}, WQS vs. OLS, and L_{min} vs. OLS are discussed next. Two variable mediation analysis using WQS, LASSO and OLS which are shown in Figure 4.105 support the conclusion that OLS and LASSO methods are indistinguishable methods for all the 108 conditions, but the WQS method has positively biased estimates for the null conditions (ME_{12}=0) and negatively biased estimates (-0.04) for mediated effects with high gamma parameter values (0.3) and low pairwise correlations $\rho_{12} = (0, 0.5)$. The same conditions under the WQS method have higher standard errors for the estimates (*RMSE*=0.04 to 0.05). For the other 96 of the 108 conditions, the WQS method traded an increased bias for a reduced standard error of the joint mediated effect estimates, thus reducing the effects of multicollinearity better than for the OLS and LASSO methods.

Pairwise comparisons of 2-variable joint indirect effect coverage probability, type1 error, and power WQS vs. LASSO, WQS vs. OLS, and LASSO vs. OLS are shown in Figure 4.106. The coverage probability for the joint mediated effect is uniformly good for the OLS and LASSO methods, but the WQS method has two breakdown conditions with coverage = 0.2. These exceptions have a high joint mediated effect (0.135) with a large negative bias of -0.04, and a large *RMSE* of 0.04. The estimate was obtained from a large sample sized data set (N=1000), which reduced the width of the interval estimate for the biased joint indirect effect $ME_{12}^{\substack{Estimate\\ WQS_{M|X}}}$, resulting in a percentile confidence interval around the joint indirect effect estimate which excludes the true value 80% of the time resulting in a coverage of 0.2



A
LASSO Lmin Joint mediated effect EST.
WQS weights from M|X ME12 EST.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
B
OLS Joint mediated effect EST.
WQS weights from M|X ME12 EST.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
C
OLS Joint mediated effect EST.
LASSO Lmin Joint mediated effect EST.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
D
LASSO Lmin Joint mediated EST. Bias
WQS weights from M|X ME12 EST. Bias
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
E
OLS Joint mediated EST.Bias
WQS weights from M|X ME12 EST. Bias
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
F
OLS Joint mediated EST.Bias
LASSO Lmin Joint mediated EST. Bias
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
Zero Corr.
0.5 Corr.
0.95 Corr.

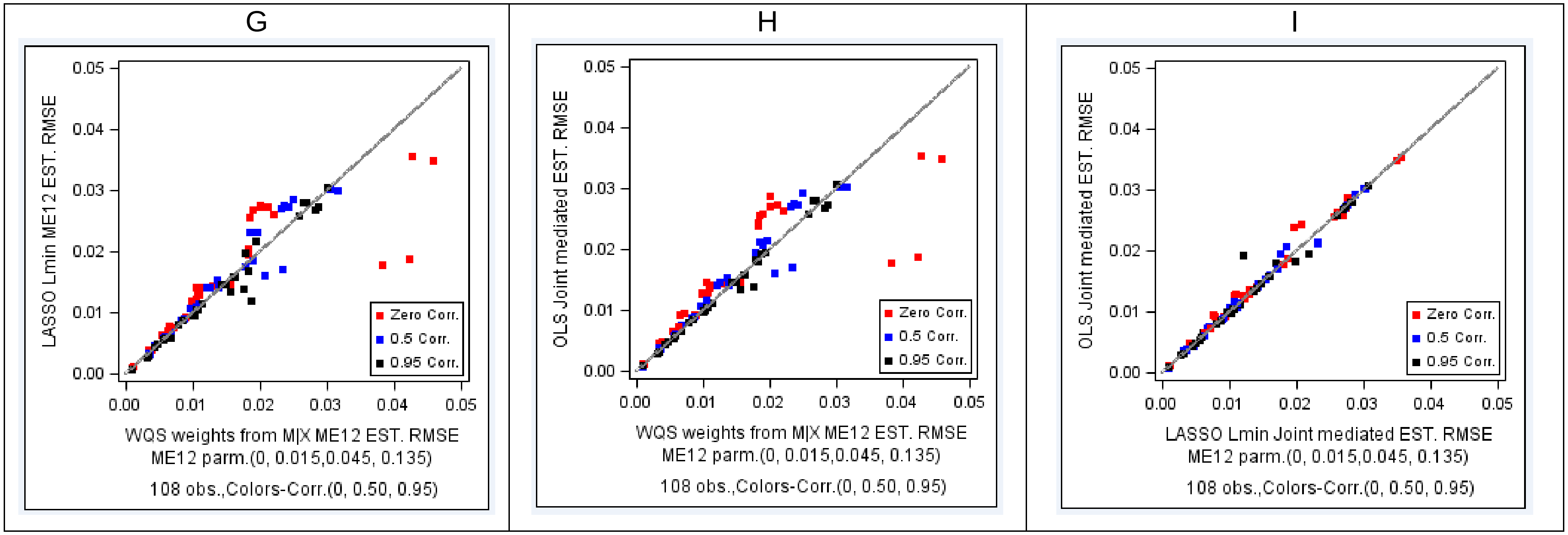
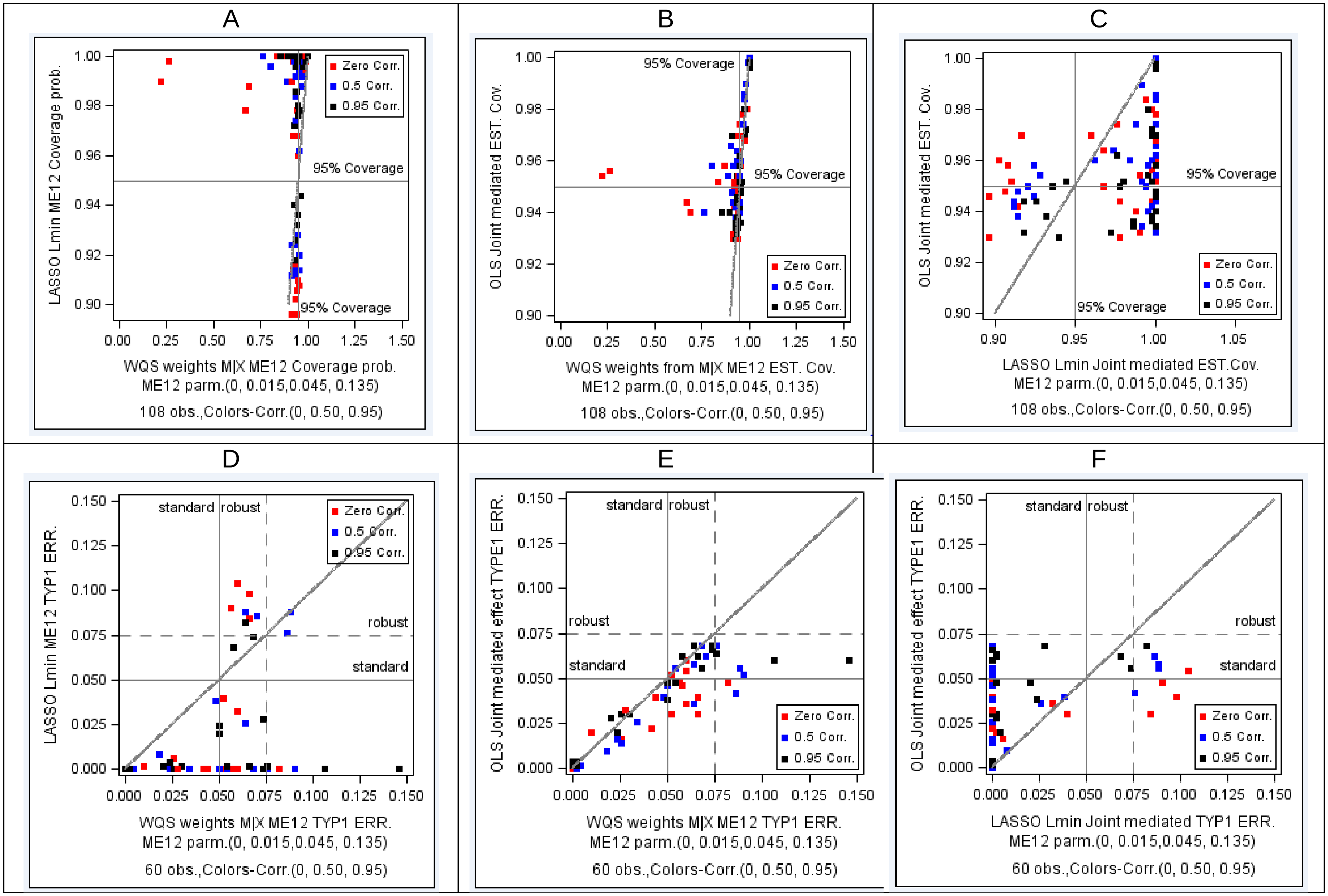


Figure 4.105 Comparison of Joint Indirect Effects A-C) $ME_{12}^{WQS_{M|X}\,Estimate}$ vs. $ME_{12}^{OLS\,Estimate}$ vs. $ME_{12}^{L_{\min}\,Estimate}$, D-F) $ME_{12}^{WQS_{M|X}\,Bias}$ vs. $ME_{12}^{OLS\,Bias}$ vs. $ME_{12}^{L_{\min}\,Bias}$, and G-I) $ME_{12}^{WQS_{M|X}\,RMSE}$ vs. $ME_{12}^{OLS\,RMSE}$ vs. $ME_{12}^{L_{\min}\,RMSE}$

The type1 errors for OLS shown in Figure 4.106 E are below the limit of 0.075 for all 60 conditions, but the LASSO method has 9 exceptions and a type1 error = 0.10, for a small, uncorrelated predictor dataset. The WQS method has 8 exceptions out of 60 conditions and a type1 error = 0.15, for a large dataset exhibiting high multicollinearity. The comparison of statistical power for the three methods is shown in Figure 4.106 G-I, and is better for the WQS method, for all 48 conditions with $ME_{12} \neq 0$ when compared with the OLS and LASSO (L_{min}) methods.

In summary, when detecting the joint mediated effect for 2-variable mediation, the WQS method is preferred over the OLS and LASSO methods, besides the few exceptions discussed above, since the WQS method reduces the effects of multicollinearity better than the OLS and LASSO methods.


A
LASSO Lmin ME12 Coverage prob.
WQS weights M|X ME12 Coverage prob.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
95% Coverage
Zero Corr.
0.5 Corr.
0.95 Corr.
B
OLS Joint mediated EST. Cov.
WQS weights from M|X ME12 EST. Cov.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
C
OLS Joint mediated EST. Cov.
LASSO Lmin Joint mediated EST.Cov.
ME12 parm.(0, 0.015,0.045, 0.135)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
D
LASSO Lmin ME12 TYP1 ERR.
WQS weights M|X ME12 TYP1 ERR.
ME12 parm.(0, 0.015,0.045, 0.135)
60 obs.,Colors-Corr.(0, 0.50, 0.95)
standard
robust
E
OLS Joint mediated effect TYPE1 ERR.
WQS weights M|X ME12 TYP1 ERR.
ME12 parm.(0, 0.015,0.045, 0.135)
60 obs.,Colors-Corr.(0, 0.50, 0.95)
F
OLS Joint mediated effect TYPE1 ERR.
LASSO Lmin Joint mediated TYP1 ERR.
ME12 parm.(0, 0.015,0.045, 0.135)
60 obs.,Colors-Corr.(0, 0.50, 0.95)

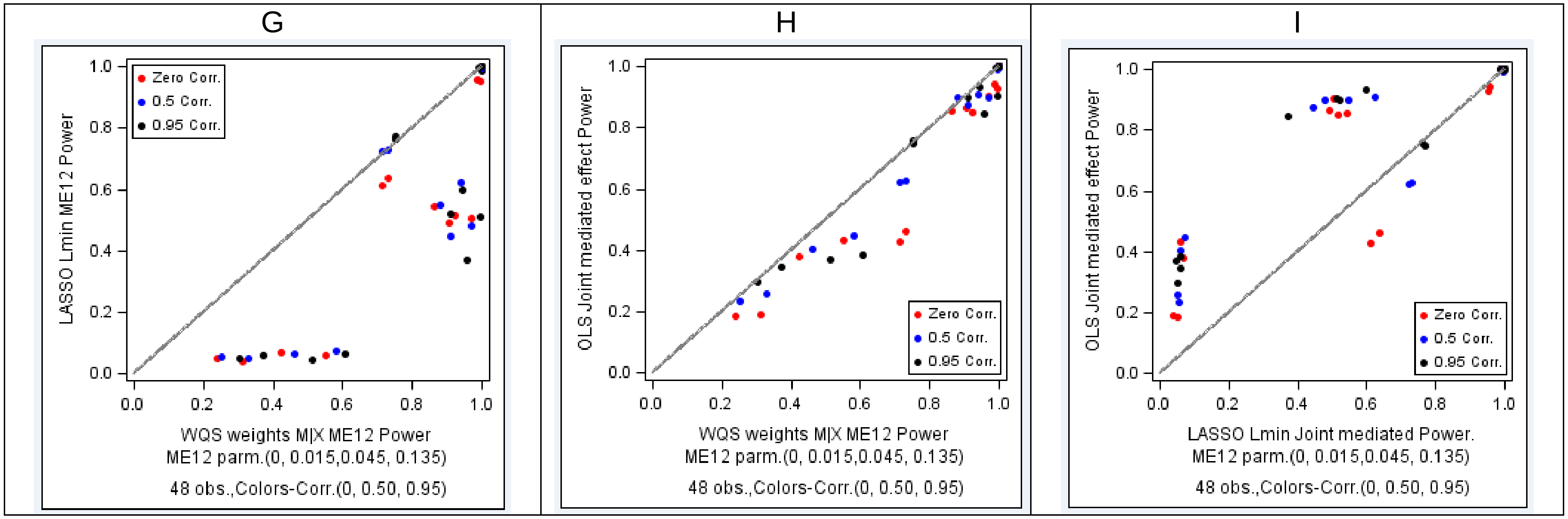
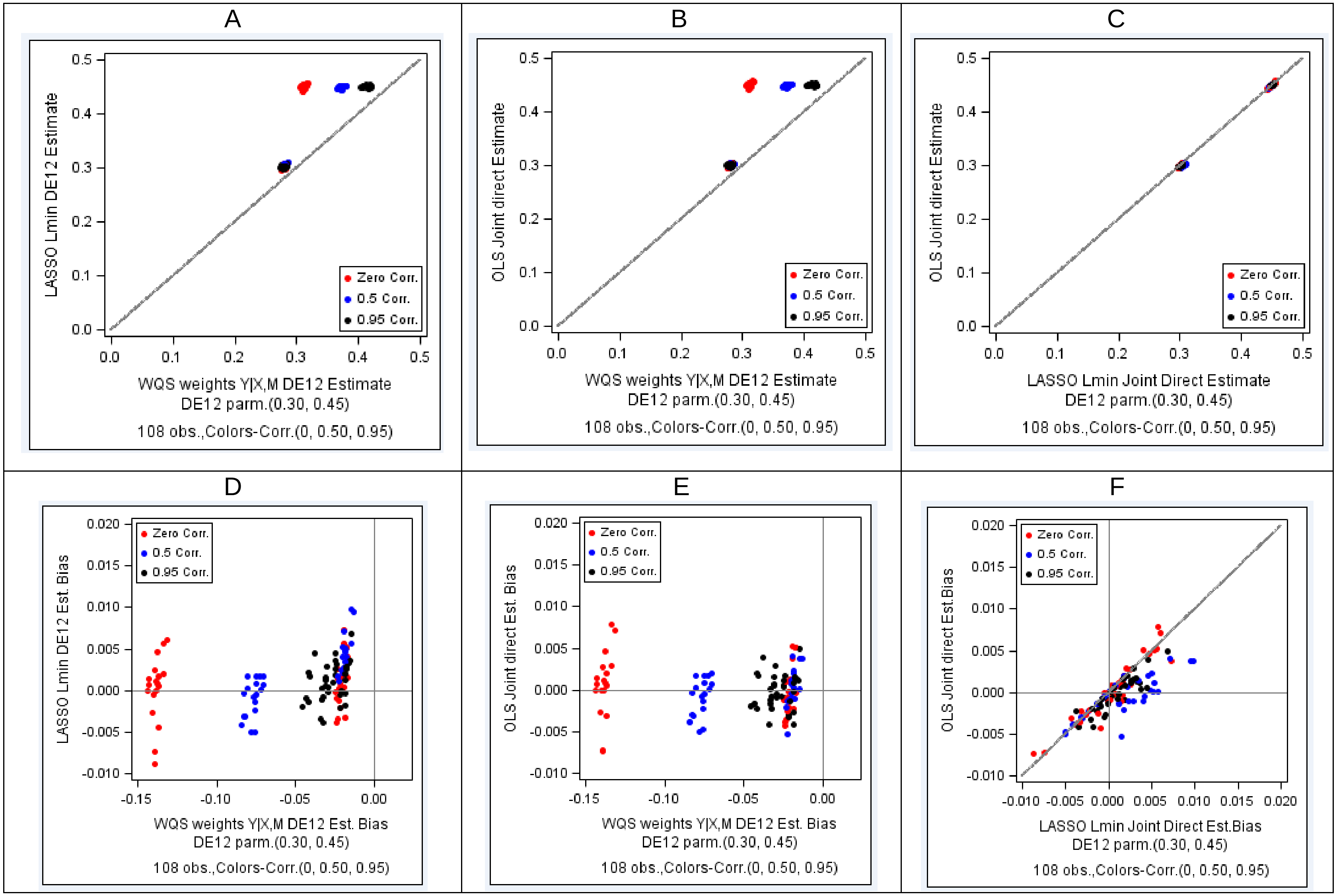


Figure 4.106 Comparison of Joint Indirect Effects A-C) $ME_{12}^{WQS_{M|X}^{Cov.prob.}}$ vs. $ME_{12}^{OLS^{Cov.prob}}$ vs. $ME_{12}^{L_{\min}^{Cov.prob}}$,
D-F) $ME_{12}^{WQS_{M|X}^{Type1\ Err.}}$ vs. $ME_{12}^{OLS^{Type1\ Err.}}$ vs. $ME_{12}^{L_{\min}^{Type1\ Err.}}$, G-I) $ME_{12}^{WQS_{M|X}^{Power}}$ vs. $ME_{12}^{OLS^{Power}}$ vs. $ME_{12}^{L_{\min}^{Power}}$

4.4.1.2 Joint Direct Effect Est. for 2-Variable comparing WQS, LASSO, and OLS

The comparison of 2-variable joint direct *Estimates*, *Bias*, and *RMSE* for WQS vs. LASSO, WQS vs. OLS, and LASSO vs. OLS are shown in Figure 4.107. The WQS method places a negative bias on the larger direct effects (-0.15 for a true effect size of 0.45) shown in Figure 4.107 D-E, whereas the joint direct effect estimates are more accurate for the LASSO and OLS methods shown in Figure 4.107 A-C. The WQS method has higher standard errors for the estimates (0.15 vs. 0.10) as compared to LASSO and OLS methods, when the dataset has weakly correlated or uncorrelated predictors in a large sized dataset which is shown as two red clusters (joint direct effects = 0.35 and 0.45) in Figure 4.107 G and H vs. I. The WQS coverage probability, estimated from large data sets with uncorrelated predictors, is almost zero for 9 joint direct effects shown as a cluster of red dots in Figure 4.108 A and B. This is because of their large biased estimates (-0.14) which have high standard errors (0.14). However, the power of these 9 conditions to detect the joint direct effect is 100%, because of the large effect size (0.45) as shown in Figure 4.108 D and E. The OLS and LASSO methods have over 90% coverage probability for the joint direct effects as shown in Figure 4.108 C, but have varying power to detect them (0.4 to 1.0) as shown in Figure 4.108 F. The OLS method is affected by multicollinearity due to high predictor correlations in small datasets, when detecting a smaller joint direct effect having high standard errors, which resulted in 9 conditions with an almost zero power, shown in Figure 4.108 E and F. These 9 conditions are not the same conditions as the previous 9 discussed for WQS coverage probability in Figure 4.108 A & B.


A
LASSO Lmin DE12 Estimate
WQS weights Y|X,M DE12 Estimate
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
Zero Corr.
0.5 Corr.
0.95 Corr.
B
OLS Joint direct Estimate
WQS weights Y|X,M DE12 Estimate
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
C
OLS Joint direct Estimate
LASSO Lmin Joint Direct Estimate
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
D
LASSO Lmin DE12 Est. Bias
WQS weights Y|X,M DE12 Est. Bias
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
E
OLS Joint direct Est. Bias
WQS weights Y|X,M DE12 Est. Bias
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)
F
OLS Joint direct Est.Bias
LASSO Lmin Joint Direct Est.Bias
DE12 parm.(0.30, 0.45)
108 obs.,Colors-Corr.(0, 0.50, 0.95)

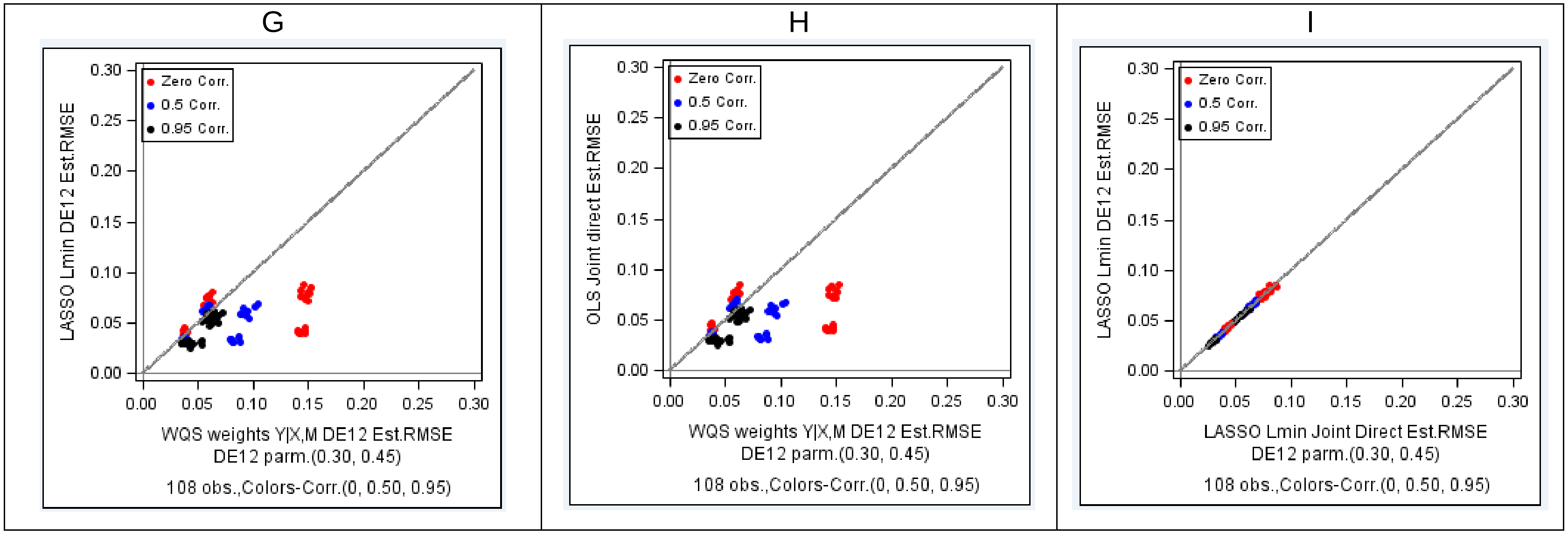


Figure 4.107 Comparison of Joint Direct Effects A-C) $DE_{12}^{WQS_{M|X}\ Estimate}$ vs. $DE_{12}^{OLS\ Estimate}$ vs. $DE_{12}^{L_{\min}\ Estimate}$, D-F) $DE_{12}^{WQS_{M|X}\ Bias}$ vs. $DE_{12}^{OLS\ Bias}$ vs. $DE_{12}^{L_{\min}\ Bias}$, and G-I) $DE_{12}^{WQS_{M|X}\ RMSE}$ vs. $DE_{12}^{OLS\ RMSE}$ vs. $DE_{12}^{L_{\min}\ RMSE}$

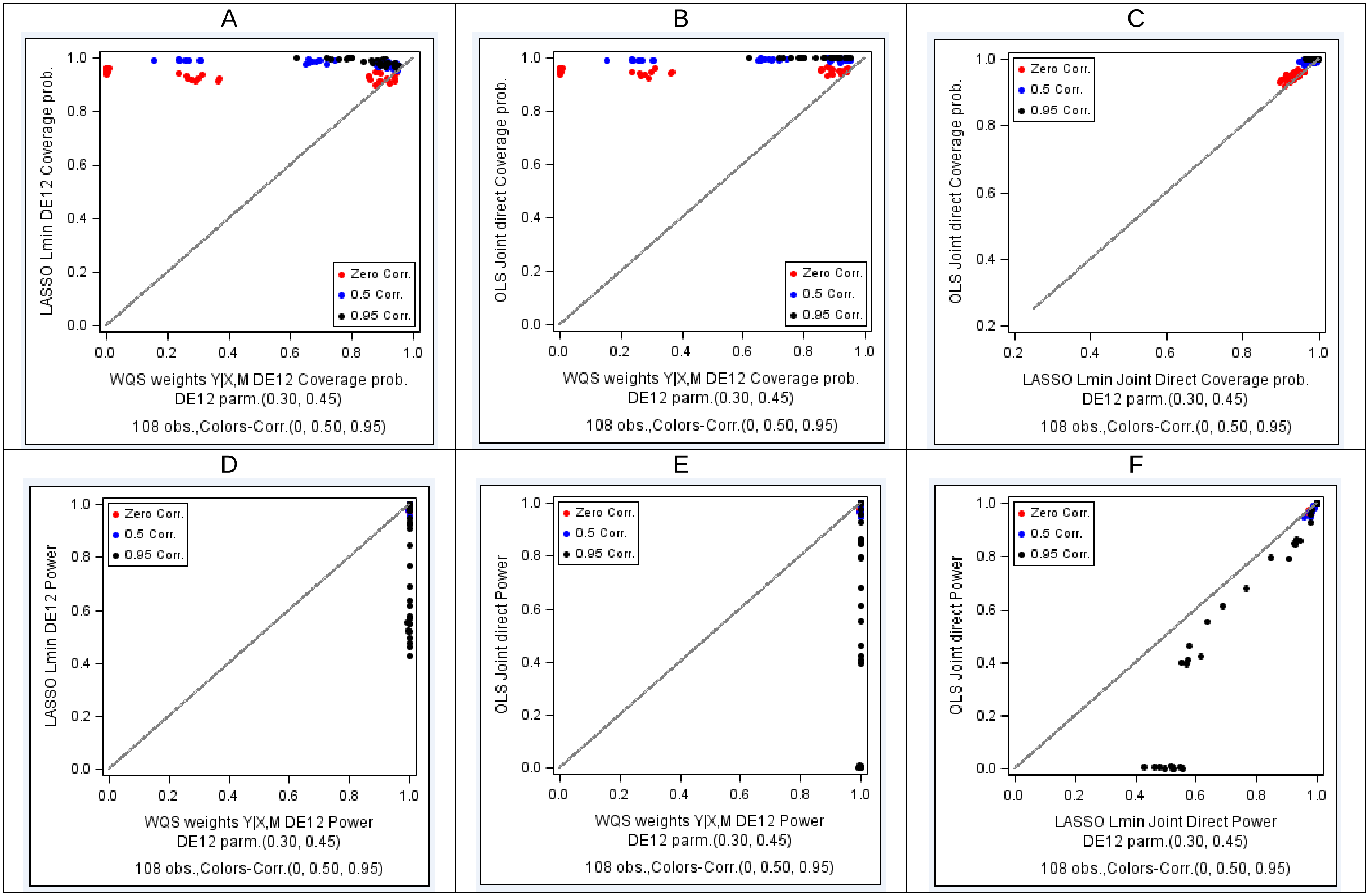


Figure 4.108 Comparison of Joint Direct Effects A-C) $DE_{12}^{WQS_{Y|X,M}\ Cov.prob.}$ vs. $DE_{12}^{OLS\ Cov.prob.}$ vs. $DE_{12}^{L_{\min}\ Cov.prob.}$ D-F) $DE_{12}^{WQS_{Y|X,M}\ PWR.}$ vs. $DE_{12}^{OLS\ PWR.}$ vs. $DE_{12}^{L_{\min}\ PWR.}$

4.4.1.3 Individual 2-Variable Mediated Effects *Est.,Bias,RMSE,*Type1 err., Power

The individual mediated estimates can only be calculated for the OLS and LASSO methods, but the influencing predictor may be identified using the WQS method, if the joint type 1 error or the joint power exists. Figures 4.109 B and C for X_1, and E and F for X_2, show the advantage of the LASSO method which has a higher bias for the individual mediated effect but a reduced standard error for the estimates. The OLS method lacks precision for the individual mediated effects (ME_1^{OLS}, ME_2^{OLS}) when datasets have highly correlated predictors causing the troubling effects of multicollinearity in multiple regression. The LASSO L_{min} method is preferred over OLS when identifying individual mediated effects $ME_1^{L_{min}}, ME_2^{L_{min}}$.

In Figure 4.110 of the 108 total conditions, ME_1=0 for 84 conditions shown in panels A-C and ME_2=0 for 60 conditions shown in panels D-F. The OLS method has zero type1 error exceptions while the LASSO method has 6 exceptions (max type 1 error = 0.20) out of 84 conditions for X_1, and 2 exceptions (max type 1 error = 0.10) out of 60 conditions for X_2. When identifying the influencing individual predictor(s) using the WQS method, provided a significant joint indirect effect is present, the individual predictor type1 error rate for ME_{12}=0 and ME_1=0, had 8 exceptions each, for $ME_{1\,M|X}^{TYP1}$ out of 84 conditions and $ME_{2\,M|X}^{TYP1}$ out of 60 conditions, when the cut-off value for individual WQS weights was set to a zero value.

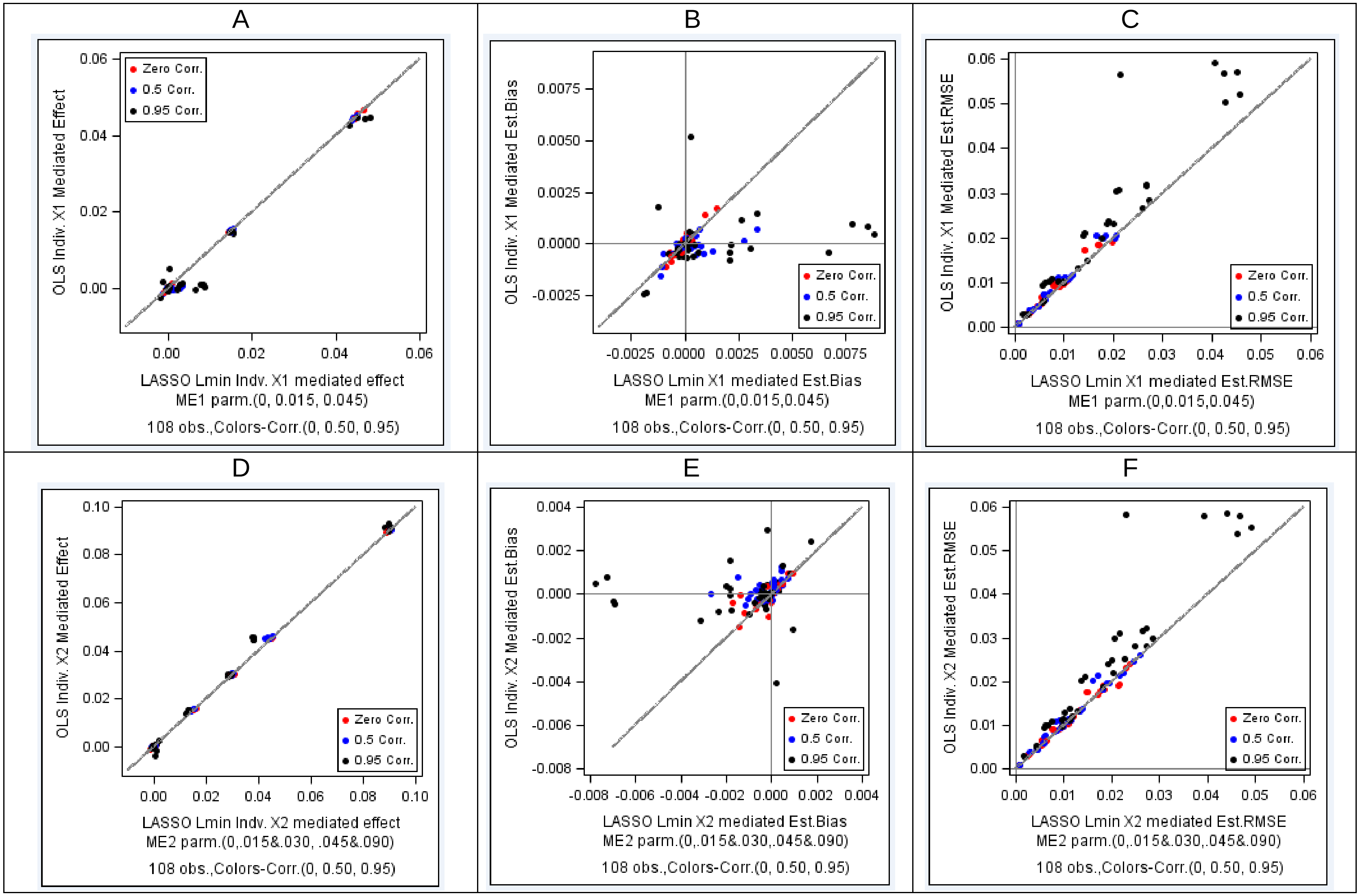


Figure 4.109 Comparison of Indirect Effects LASSO vs. OLS A-C) $ME_1^{Estimate_{L\min}}$ vs. $ME_1^{Estimate_{OLS}}$, $ME_1^{Bias_{L\min}}$ vs. $ME_1^{Bias_{OLS}}$, $ME_1^{RMSE_{L\min}}$ vs. $ME_1^{RMSE_{OLS}}$

D-F) $ME_2^{Estimate_{L\min}}$ vs. $ME_2^{Estimate_{OLS}}$, $ME_2^{Bias_{L\min}}$ vs. $ME_2^{Bias_{OLS}}$, $ME_2^{RMSE_{L\min}}$ vs. $ME_2^{RMSE_{OLS}}$

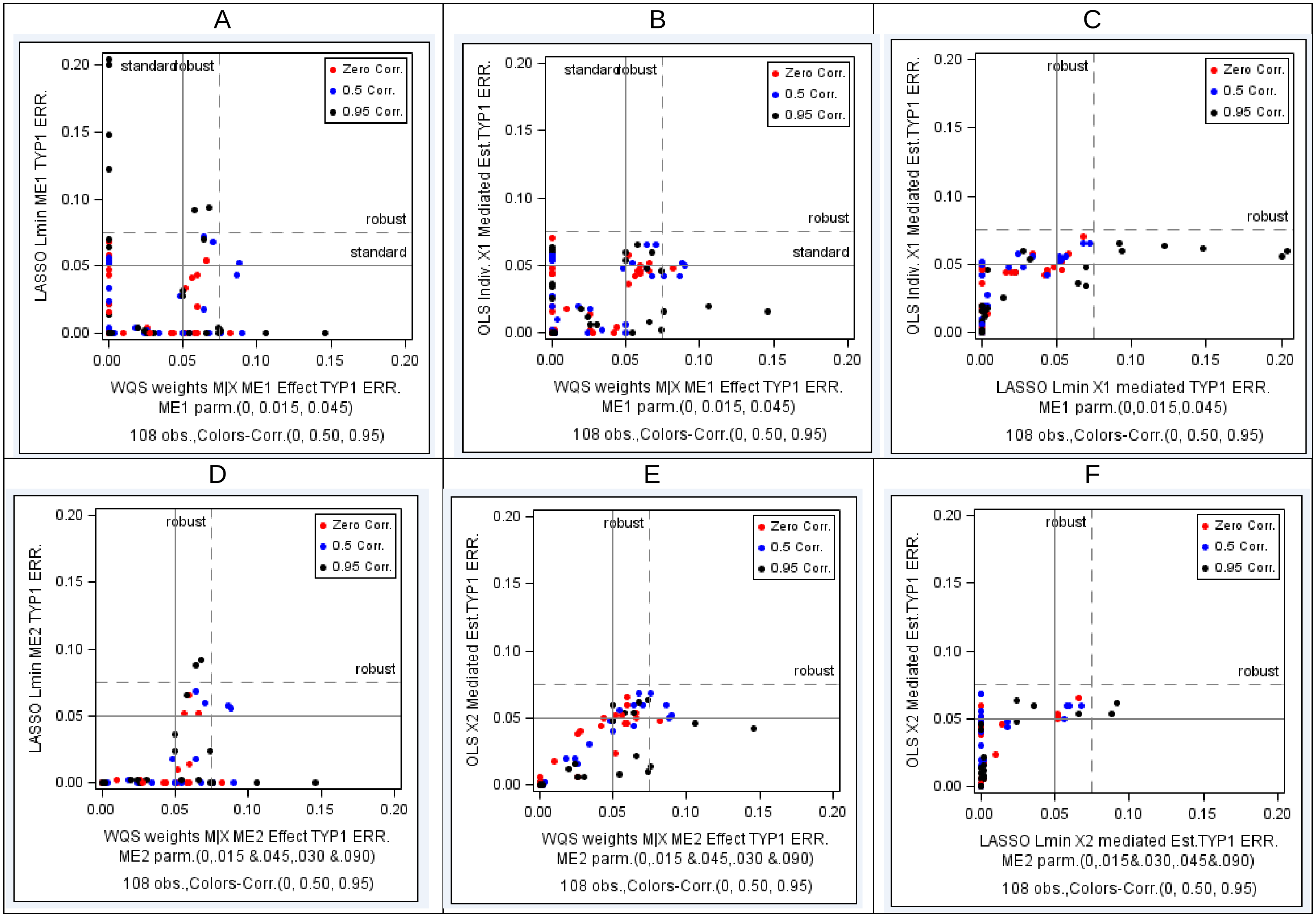


Figure 4.110 Comparison of Indiv. Indirect Effects A-C) $ME_1^{WQS_{M|X}\ Type1\ Err.}$ vs. $ME_1^{OLS\ Type1\ Err.}$ vs. $ME_1^{L\min\ Type1\ Err.}$ D-F) $ME_2^{WQS_{M|X}\ Type1\ Err.}$ vs. $ME_2^{OLS\ Type1\ Err.}$ vs. $ME_2^{L\min\ Type1\ Err.}$

Figure 4.111 A-F compares the type1 error rates between the three methods (WQS, LASSO and OLS) for individual mediated effects X_1, X_2. With the WQS method, a cut-off value for the individual WQS weights of (0, 0.1, 0.2, 0.3 and 0.4) were used to specify zero weights having no influence in determining the joint mediated effect type1 error rate. The effect on the individual WQS type1 error rate for increasing cutoff values for the WQS weights, which contribute to the joint mediated effect type 1 error rates are shown in Figure 4.111 X1.20 to X2.40; e.g. increasing the cut-off rate for WQS weights for X_1 from 0.20 to 0.30 reduces the exceptions where >0.075, from five to two. For 2-variable mediation a cut-off =0.30 might identify the individual influential predictors which contribute to the Type1 error rate of the joint mediated effect well.

Figure 4.112 A-F compare the power to identify influencing predictors between the three methods (WQS, LASSO and OLS) for individual mediated effects X_1, X_2. The WQS method has a power of 1 for both $ME_{1}^{\substack{PWR \\ M|X}} \& ME_{2}^{\substack{PWR \\ M|X}}$ with a cut-off=0, but the type1 error rates from Figure 4.111 had suggested a non-zero cut-off value to contain the type1 errors. The effect of reduced power to detect a WQS joint mediated effect using cut-off values = (0.2, 0.3 and 0.4) for the individual WQS weights is shown in Figure 4.113 X1.20 to X2.40 below; e.g. increasing the cut-off rate for WQS weights for X_1 and X_2 from 0.20 to 0.30 reduces the power to detect their influence on the joint mediated effect, as shown in X1_20 and X1_30, and X2_20 and X2_30 respectively.

The WQS method has a higher power for the individual mediated effects than the LASSO and OLS methods but the type1 error rates are most conservative for the OLS method as shown in Figure 4.113.

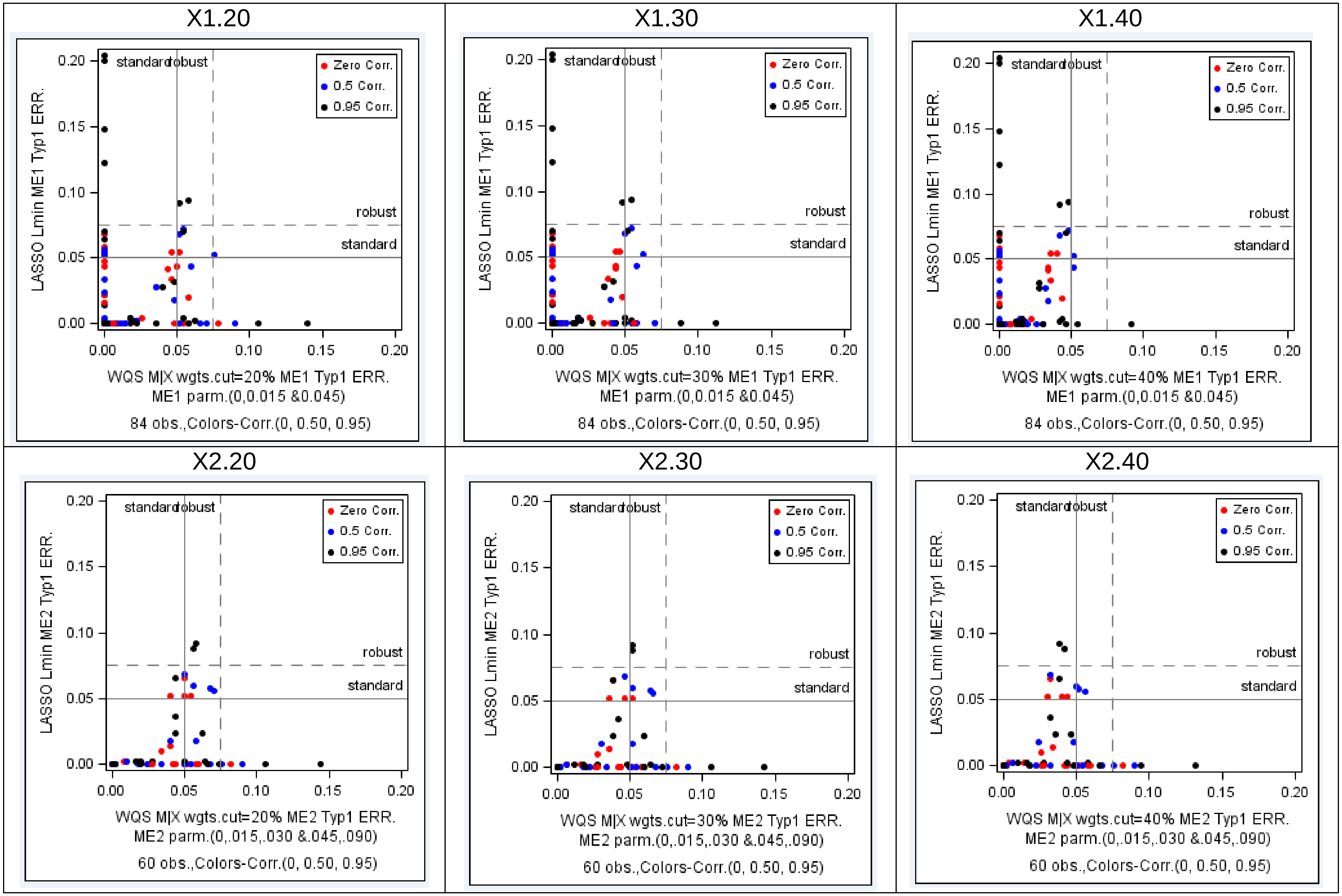
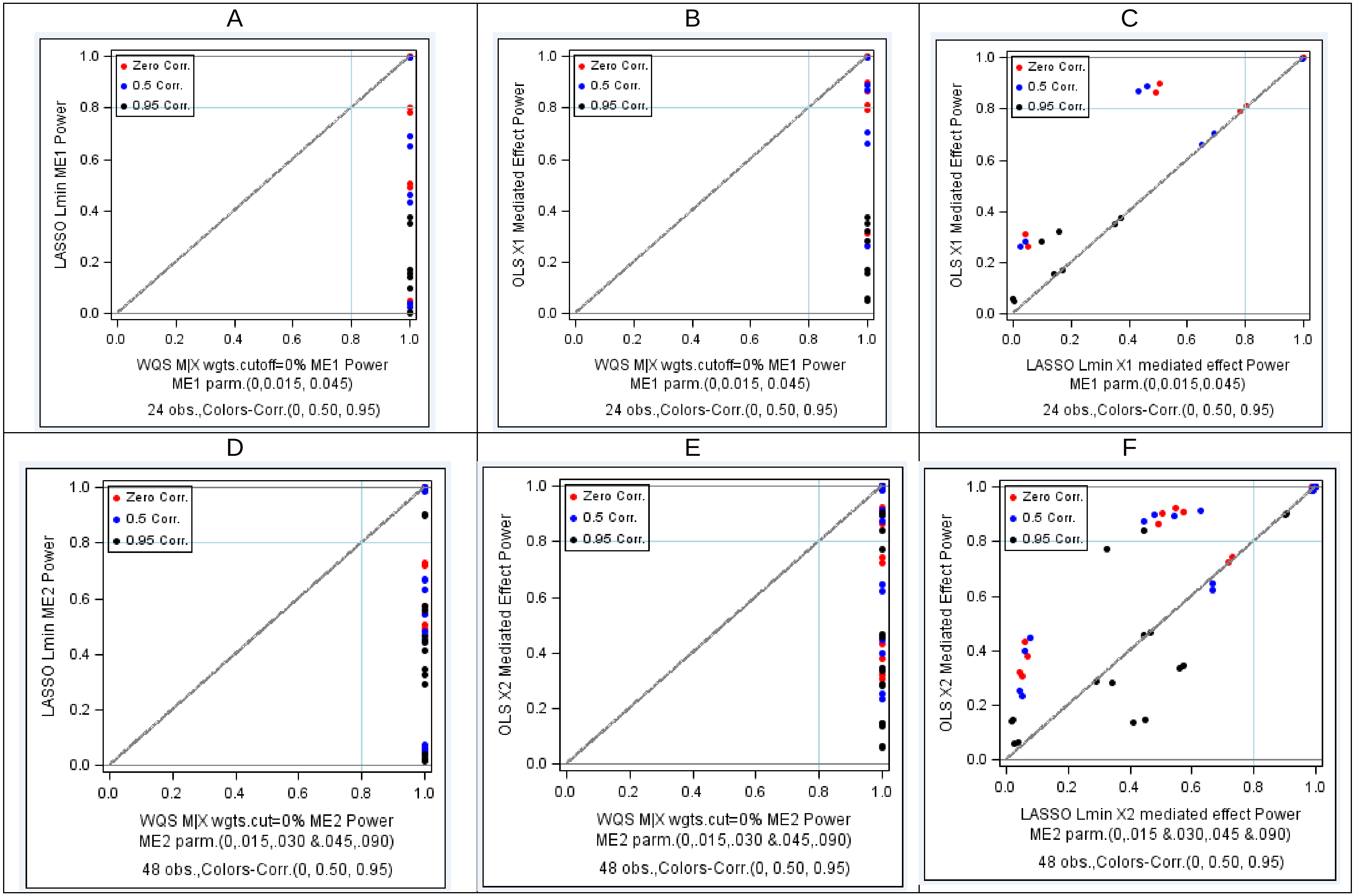


Figure 4.111 Individual Indirect effects with cut-off values 0.2 to 0.4 by 0.1 X1.20 to X2.40) $\overset{Typ1\_20}{ME_{1,2}^{WQS_{M|X}}}$ vs. $\overset{Typ1\_30}{ME_{1,2}^{WQS_{M|X}}}$ vs. $\overset{Typ1\_40}{ME_{1,2}^{WQS_{M|X}}}$



Figures 4.112 Comparison of Indiv. Indirect Effect Power A-C) $ME_1^{Power}{}_{WQS_{M|X}}$ vs. $ME_1^{Power}{}_{OLS}$ vs. $ME_1^{Power}{}_{L_{\min}}$ D-F) $ME_2^{Power}{}_{WQS_{M|X}}$ vs. $ME_2^{Power}{}_{OLS}$ vs. $ME_2^{Power}{}_{L_{\min}}$

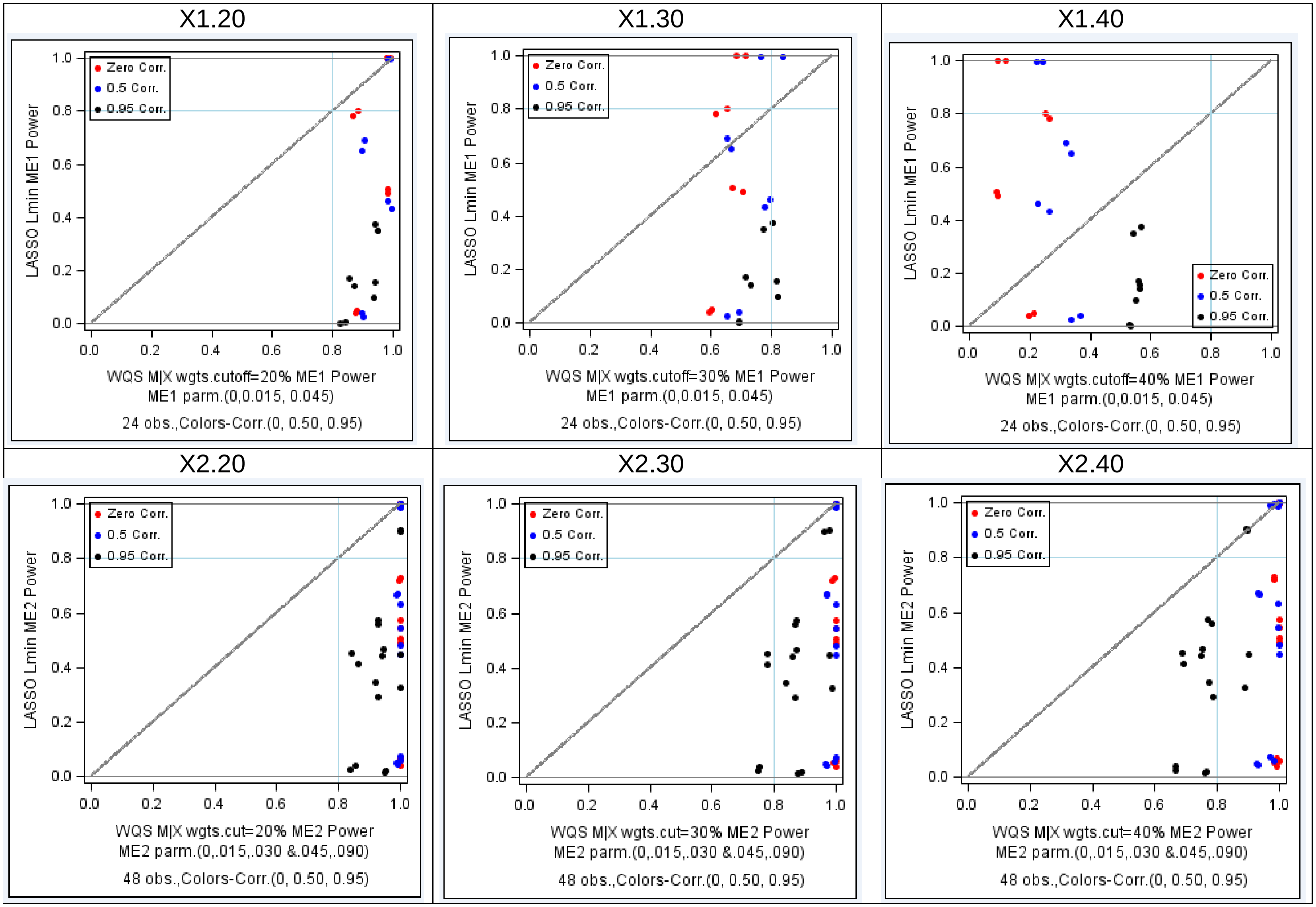


Figure 4.113 Comparison of Power for Individual Indirect Effects for cut-off=0.20 to 0.40 by 0.1 for X_1, X_2

X1_20 – X1_40) $ME_1^{\substack{Power\_20. \\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Power\_30. \\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Power\_40. \\ WQS_{M|X}}}$ X2_20 – X2_40) $ME_2^{\substack{Power\_20. \\ WQS_{M|X}}}$ vs. $ME_2^{\substack{Power\_30. \\ WQS_{M|X}}}$ vs. $ME_2^{\substack{Power\_40. \\ WQS_{M|X}}}$

4.4.1.4 Individual 2-Variable Direct Effect Est., Bias, *RMSE,* Type1 Err., Power

The WQS method only provides the joint indirect effects without estimates for the individual indirect effects. For this reason only the LASSO and OLS methods are compared for the individual direct effect mean estimates, estimate's bias and estimate's *RMSE* in Figure 4.114. The OLS method has unbiased estimates with large standard errors, when the pairwise correlations are high. The LASSO method adds a bias to one predictor while adding a bias of the opposite sign to the other predictor, shown in Figure 4.114 B and E for X_1 & X_2 respectively, and consequently reducing the individual estimate's standard error compared to the OLS method shown in Figure 4.114 C and F.

Since there were no conditions where DE_1=0, Figure 4.115 shows the type1 errors for only DE_2 which are compared between the WQS, LASSO and OLS methods.

The individual type 1 error rate for DE_2=0 is shown in Figure 4.115 A-C. The WQS method has a zero type1 error rate for all 54 conditions where DE_2=0. As observed in the individual mediated effect cases before in Figure 4.110 B and C, E and F, the OLS method has a conservative type1 error rate with zero exceptions (type1 error >0.075), while the LASSO method has 3 exceptions out of the 54 conditions where DE_2=0 shown in panels Figure 4.115 A and C, for a small dataset with high predictor pairwise correlations resulting in the effects of multicollinearity in regression models.

The power to detect the individual direct effects is discussed below in Figure 4.116. The WQS method has a power of 1 for all 108 conditions where $DE_1 \neq 0$ and 54 conditions where $DE_2 \neq 0$. Between OLS and LASSO methods for detecting individual direct effect the latter has a higher power for 8 of the 108 conditions for DE_1, and 1 case out of 54 conditions for DE_2 as shown in Figure 4.116 C and F.

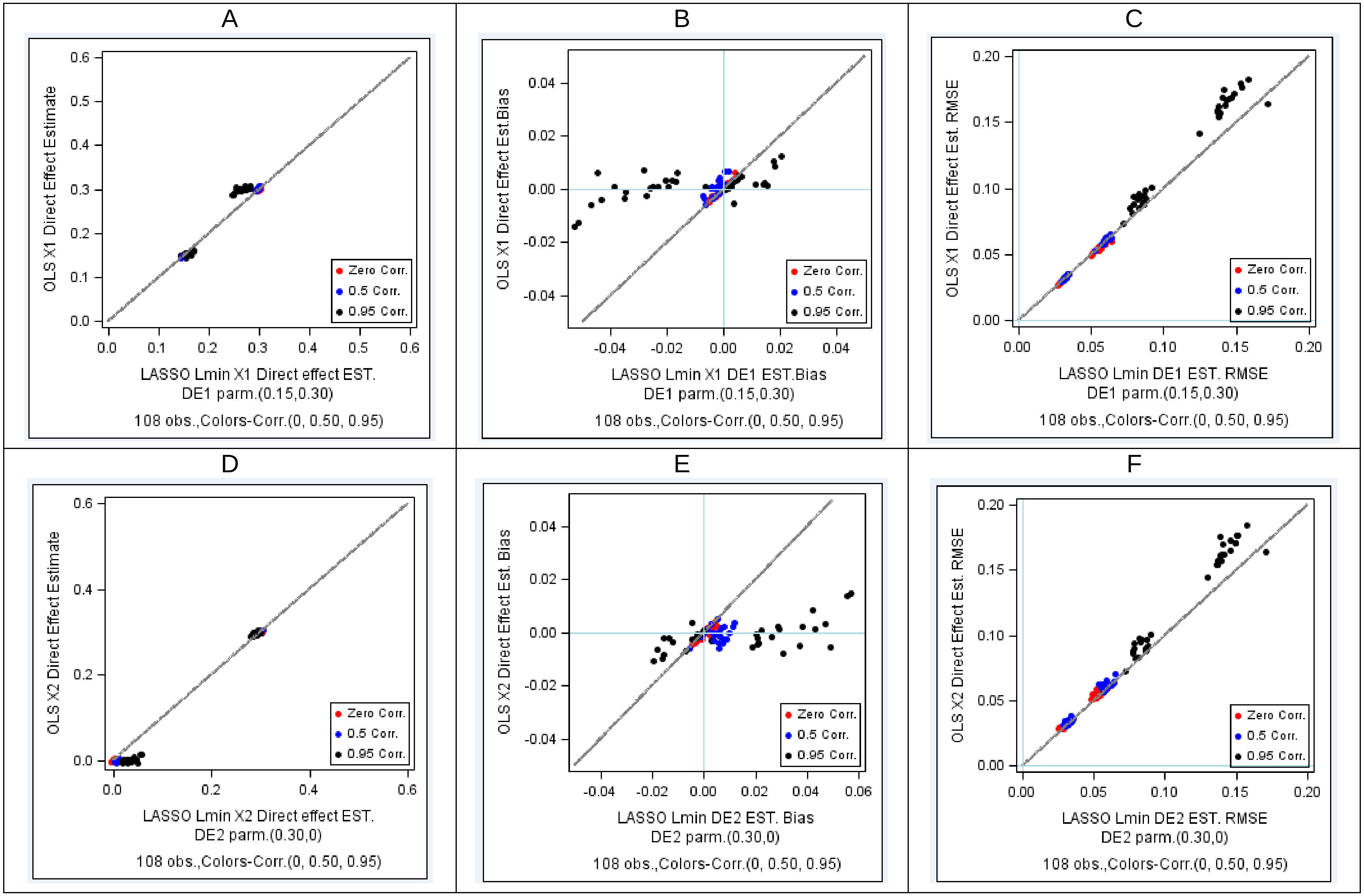


Figure 4.114 Comparison of Indiv. Direct Effect *by* LASSO and OLS A-C) $DE_{1}^{EST.}_{OLS}$ vs. $DE_{1}^{EST.}_{L\min}$, $DE_{1}^{Bias}_{OLS}$ vs. $DE_{1}^{Bias}_{L\min}$, $DE_{1}^{RMSE}_{OLS}$ vs. $DE_{1}^{RMSE}_{L\min}$

D-F) $DE_{2}^{EST.}_{OLS}$ vs. $DE_{2}^{EST.}_{L\min}$, $DE_{2}^{Bias}_{OLS}$ vs. $DE_{2}^{Bias}_{L\min}$, $DE_{2}^{RMSE}_{OLS}$ vs. $DE_{2}^{RMSE}_{L\min}$

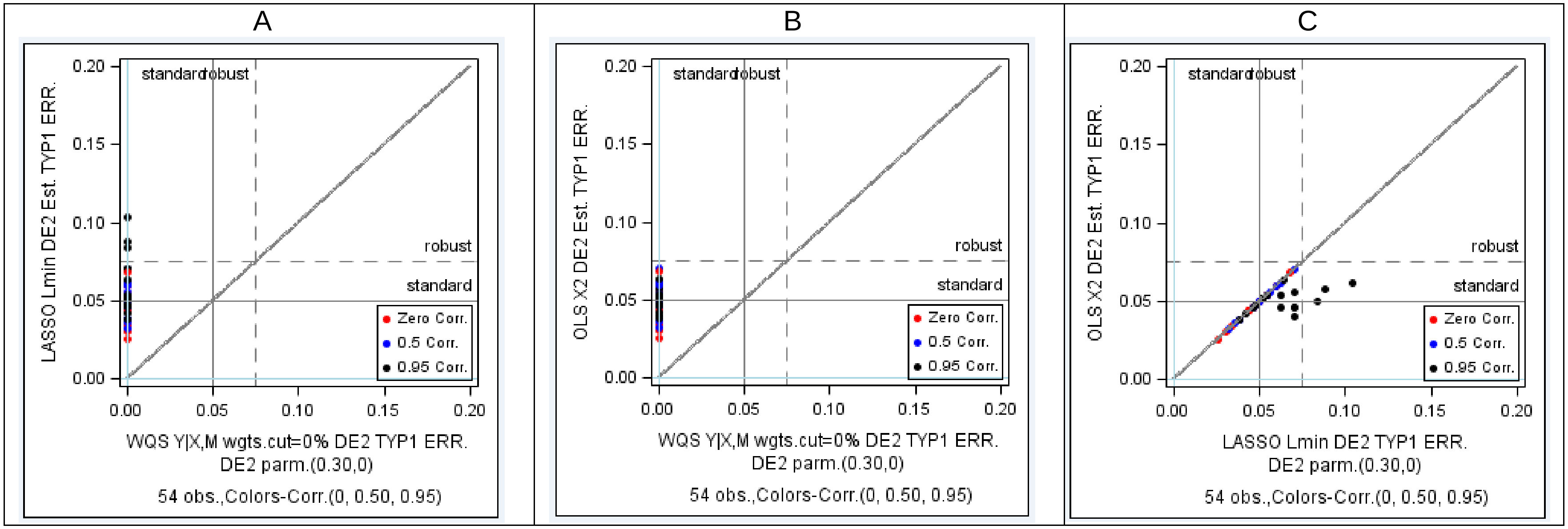


Figure 4.115 Comparison of Indiv. Direct Effects X_2 A-C) $DE_2^{\substack{Typ1.Err.\\WQS_{Y|X,M}}}$ vs. $DE_2^{\substack{Typ1.Err.\\OLS}}$ vs. $DE_2^{\substack{Typ1.Err.\\L_{\min}}}$

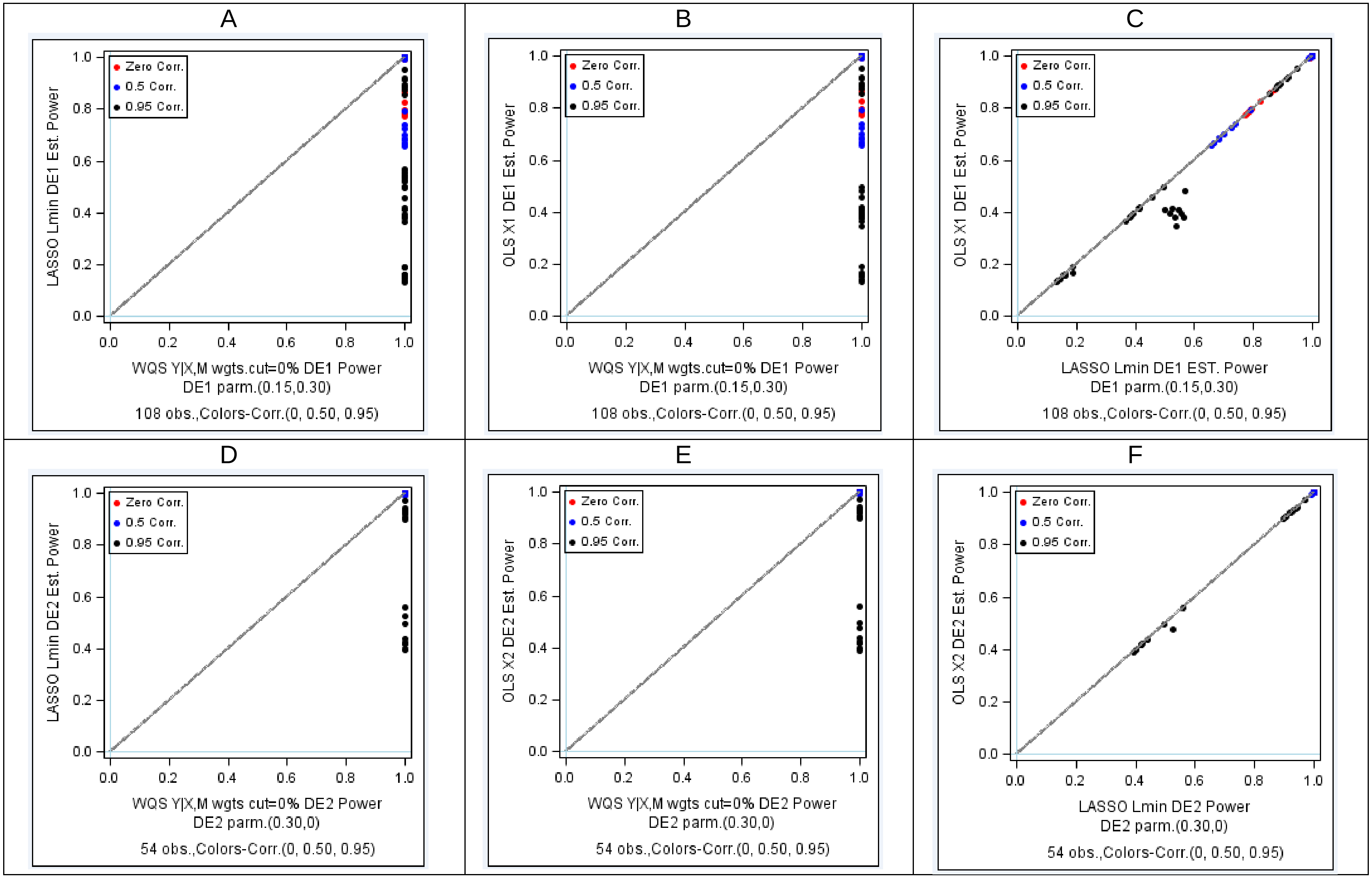


Figure 4.116 Comparison of Indiv. Direct Effects A-C) $DE_1^{Power.}{}_{WQS_{Y|X,M}}$ vs. $DE_1^{Power}{}_{OLS}$ vs. $DE_1^{Power}{}_{L_{\min}}$, D-F) $DE_2^{Power.}{}_{WQS_{Y|X,M}}$ vs. $DE_2^{Power}{}_{OLS}$ vs. $DE_2^{Power}{}_{L_{\min}}$

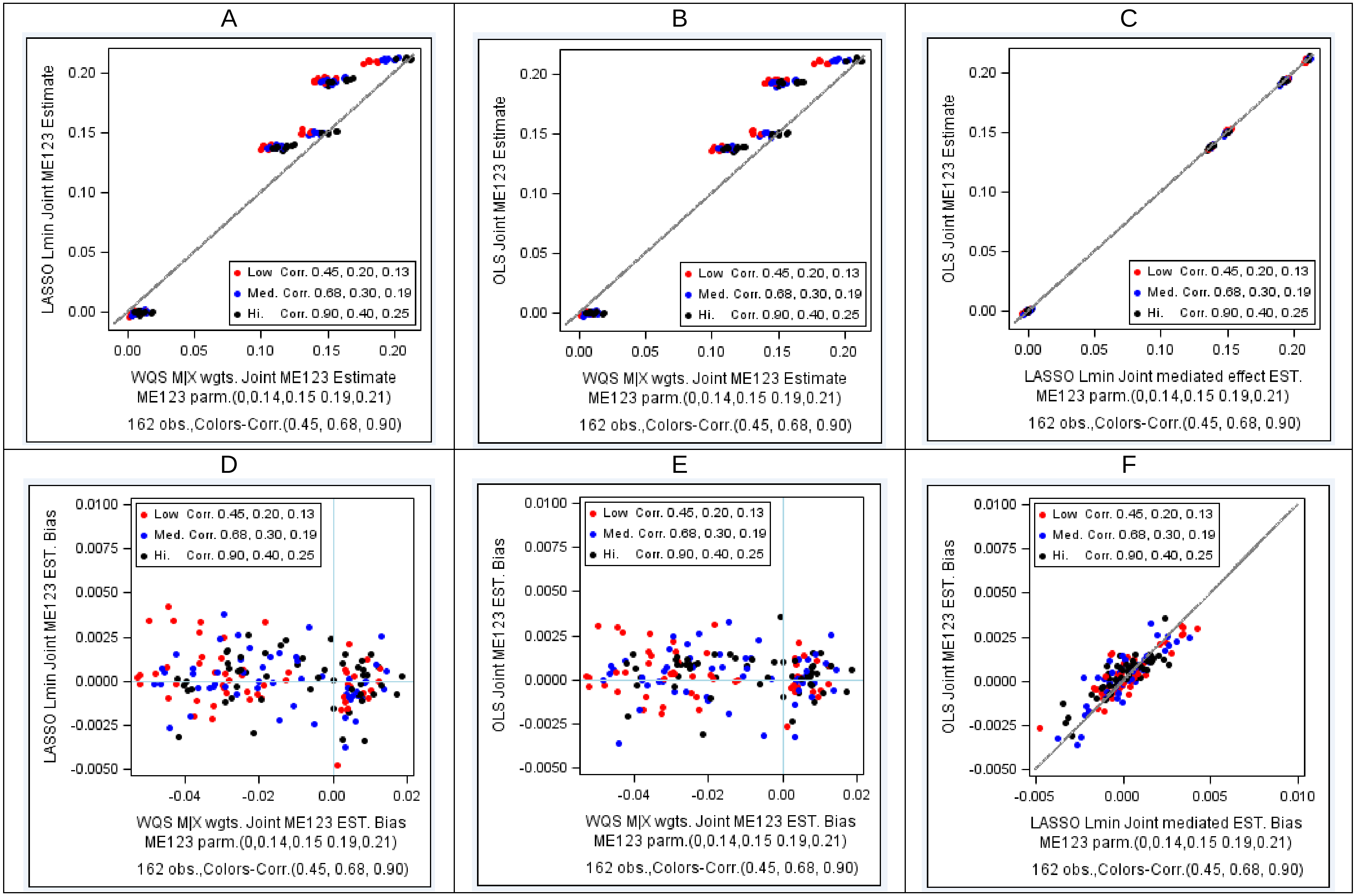
4.5 Pairwise Comparison of WQS, LASSO, and OLS for 3-Variable Mediation

4.5.1 Joint ME_{123} Est., Bias, & RMSE comparison by WQS, LASSO, & OLS

The 3-variable joint indirect effect *Estimates, Bias,* and *RMSE* for WQS vs. LASSO, WQS vs. OLS, and LASSO vs. OLS are shown in Figure 4.117 A-I. In the WQS method, conditions with low predictor correlations and low joint mediated effects have a larger bias (-0.05) as compared to conditions with higher joint mediated effects (-0.02), shown in Figure 4.117 A and D or B and E. The OLS and LASSO methods have low estimate biases (±0.005) as shown in Figure 4.117 C and F. The WQS method has higher standard errors (up to 0.06) as compared to the OLS and LASSO methods (0.02 to 0.04), shown in Figure 4.117 G, H and I.

The coverage, type1 error rates and power for the joint indirect effects using WQS, LASSO, and OLS are compared in Figure 4.118. The WQS coverage for the joint mediated effect is poor but around 0.95 for both the OLS and LASSO methods. There is one type1 error (>0.075) exception in 54 possible conditions with $\gamma = 0$, for both the OLS and LASSO methods, but 16 exceptions out of 54 conditions for the WQS method. This makes the OLS and LASSO methods preferable to the WQS method for joint indirect effects as shown in Figure 4.118 C and I. The power for all three methods is very close to 1.

In summary, using either the OLS or LASSO methods for 3-variable mediation proves to be better for estimating the joint indirect effects as compared to the WQS method (16 of 54 Type1 error exceptions) as shown in Figure 4.118 D and E.


A
LASSO Lmin Joint ME123 Estimate
WQS M|X wgts. Joint ME123 Estimate
ME123 parm.(0,0.14,0.15 0.19,0.21)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
B
OLS Joint ME123 Estimate
WQS M|X wgts. Joint ME123 Estimate
C
OLS Joint ME123 Estimate
LASSO Lmin Joint mediated effect EST.
D
LASSO Lmin Joint ME123 EST. Bias
WQS M|X wgts. Joint ME123 EST. Bias
E
OLS Joint ME123 EST. Bias
WQS M|X wgts. Joint ME123 EST. Bias
F
OLS Joint ME123 EST. Bias
LASSO Lmin Joint mediated EST. Bias

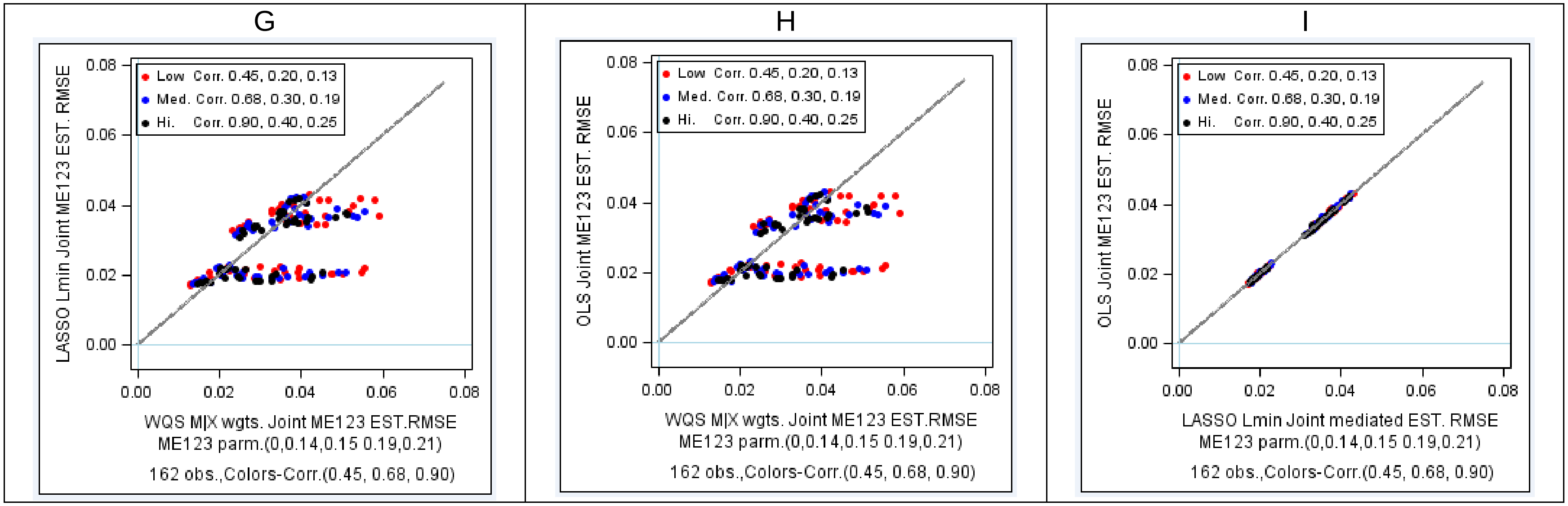
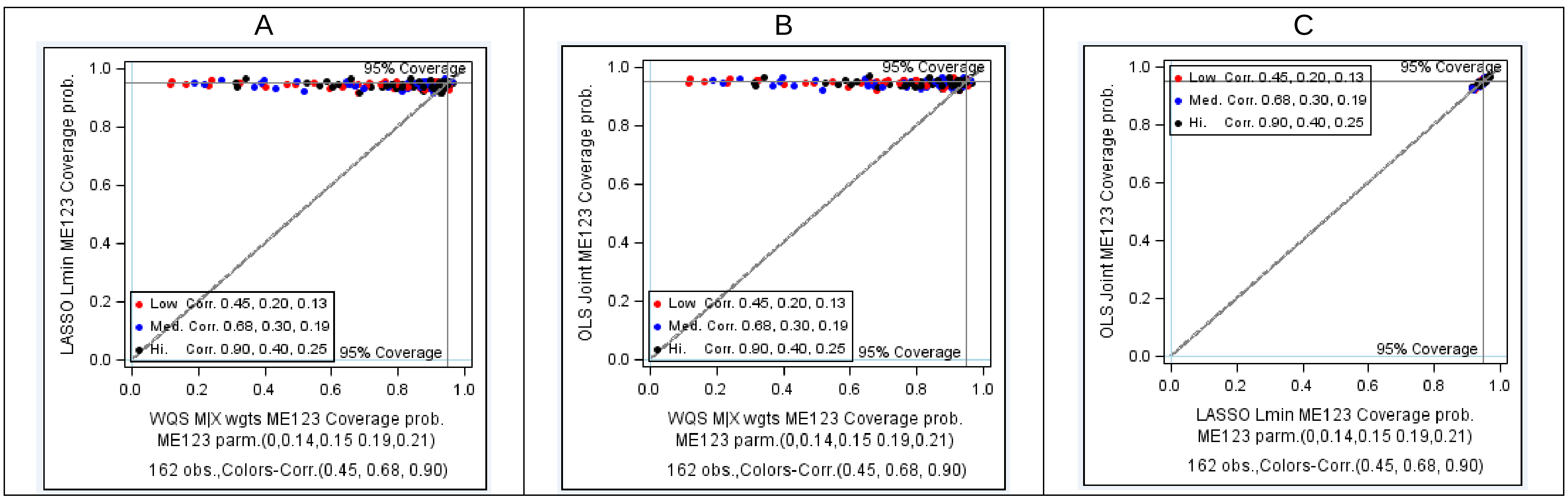


Figure 4.117 Comparison of Joint Indirect Effect *Estimate*, Estimate *Bias*, and Estimate *RMSE* by WQS, LASSO and OLS

A-C) $ME_{123}^{\overset{Estimate}{WQS_{M|X}}}$ vs. $ME_{123}^{\overset{Estimate}{OLS}}$ vs. $ME_{123}^{\overset{Estimate}{L_{\min}}}$, D-F) $ME_{123}^{\overset{Bias}{WQS_{M|X}}}$ vs. $ME_{123}^{\overset{Bias}{OLS}}$ vs. $ME_{123}^{\overset{Bias}{L_{\min}}}$, G-I) $ME_{123}^{\overset{RMSE}{WQS_{M|X}}}$ vs. $ME_{123}^{\overset{RMSE}{OLS}}$ vs. $ME_{123}^{\overset{RMSE}{L_{\min}}}$



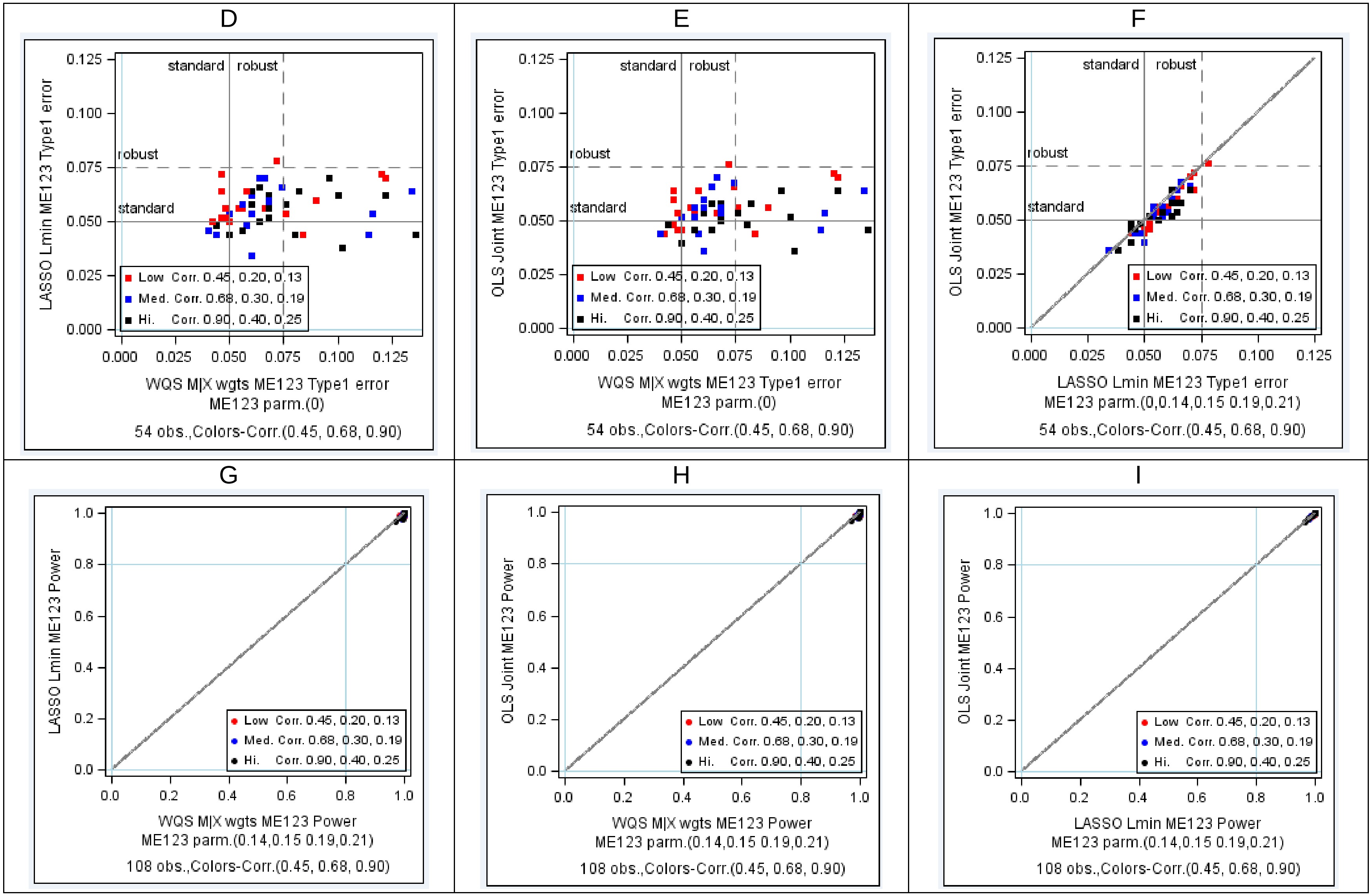


Figure 4.118 Comparison of Joint Indirect Effect Coverage Prob., Type1 Error, and Power by WQS, LASSO and OLS

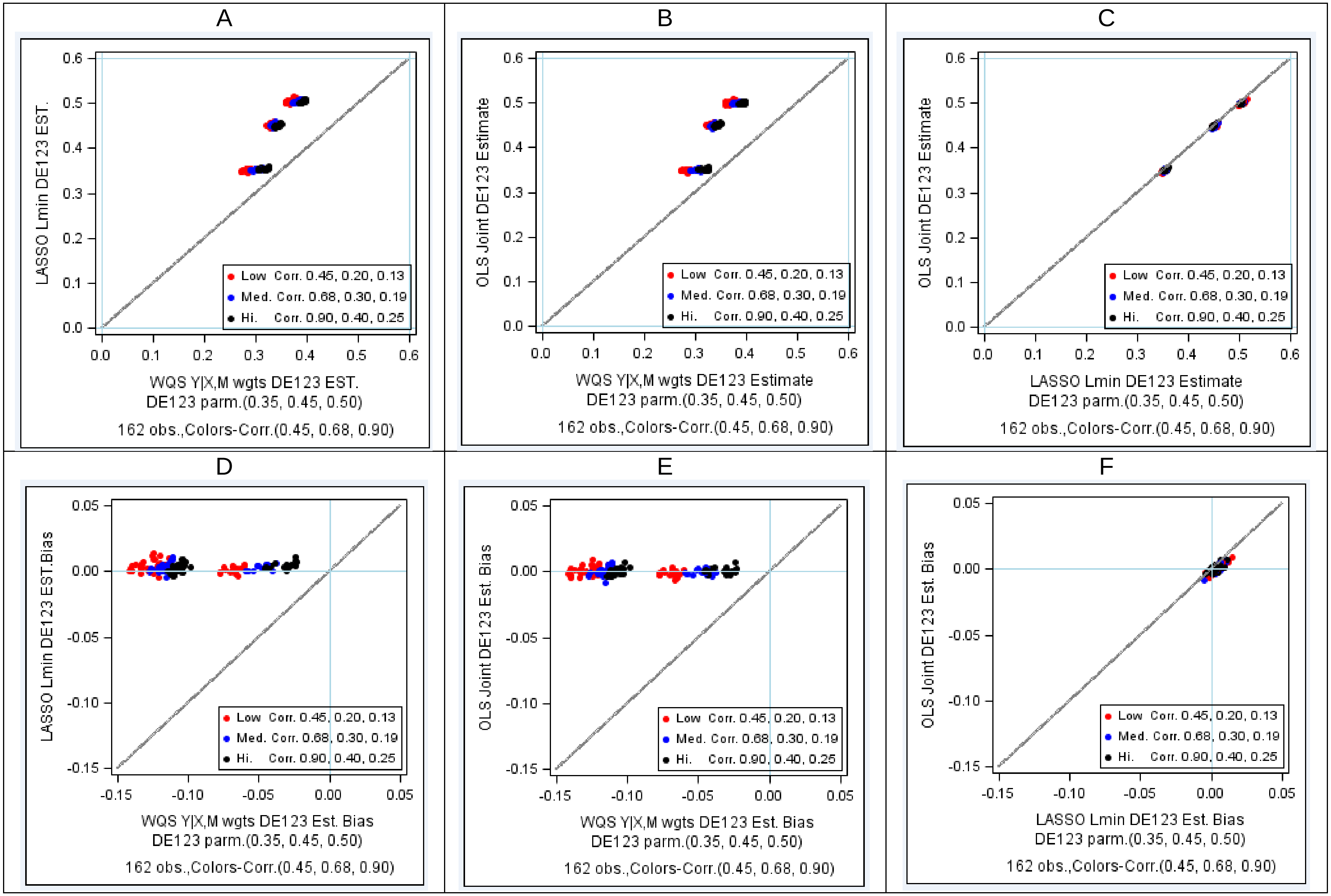
A-C) $ME_{123}^{Cov.prob.\ WQS_{M|X}}$ vs. $ME_{123}^{Cov.prob.\ OLS}$ vs. $ME_{123}^{Cov.prob.\ L_{\min}}$, D-F) $ME_{123}^{Type1err.\ WQS_{M|X}}$ vs. $ME_{123}^{Type1err.\ OLS}$ vs. $ME_{123}^{Type1err.\ L_{\min}}$, G-I) $ME_{123}^{Power\ WQS_{M|X}}$ vs. $ME_{123}^{Power\ OLS}$ vs. $ME_{123}^{Power\ L_{\min}}$

4.5.2 Comparison of 3-Var. Direct Effect Performance for WQS, LASSO & OLS

4.5.2.1 3-Var. Joint Direct *Est., Bias* & *RMSE* compared by WQS, LASSO & OLS.

The 3-variable joint direct effect estimate, estimate's bias and estimate's *RMSE* are shown for the WQS, LASSO and OLS methods in Figure 4.119. In the WQS method, conditions with low pairwise correlations and higher non-zero joint direct effects (0.45, 0.50) have a larger negative bias (up to -0.14) as compared to conditions with low joint direct effects (up to -0.08) as shown in Figure 4.119 A and D or B and E. The OLS and LASSO methods have low estimate biases (±0.01) for the joint direct effect as shown in panels C and F. The WQS method also has higher standard errors (up to 0.15) for uncorrelated or low pairwise correlated predictors in large sized dataset, as compared to the OLS and LASSO methods (0.03 for $\gamma = 0.25$ or 0.09 for $\gamma = 0.35$), and shown in Figure 4.119 G and H. The coverage and power for the joint direct effects using WQS, LASSO and OLS are compared next.

There were no conditions with DE_{123}=0 where type1 error rates for the joint direct effect could be determined. The coverage for the joint direct effect is poor for the WQS method but is above 0.95 for both the OLS and LASSO methods as shown in Figure 4.120 C. The power for detecting the joint direct effect is very close to 1 for WQS since the effect size is non-zero and large (0.35, 0.45, and 0.50). Over the 162 conditions the LASSO method has the higher power over the OLS method as shown in Figure 4.120 F, confirming that the OLS method performs poorly with small datasets, exhibiting effects of high multicollinearity, especially when detecting smaller joint direct effects (0.35).


A
LASSO Lmin DE123 EST.
WQS Y|X,M wgts DE123 EST.
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
B
OLS Joint DE123 Estimate
WQS Y|X,M wgts DE123 Estimate
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
C
OLS Joint DE123 Estimate
LASSO Lmin DE123 Estimate
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
D
LASSO Lmin DE123 EST.Bias
WQS Y|X,M wgts DE123 Est. Bias
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
E
OLS Joint DE123 Est. Bias
WQS Y|X,M wgts DE123 Est. Bias
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
F
OLS Joint DE123 Est. Bias
LASSO Lmin DE123 Est. Bias
DE123 parm.(0.35, 0.45, 0.50)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25

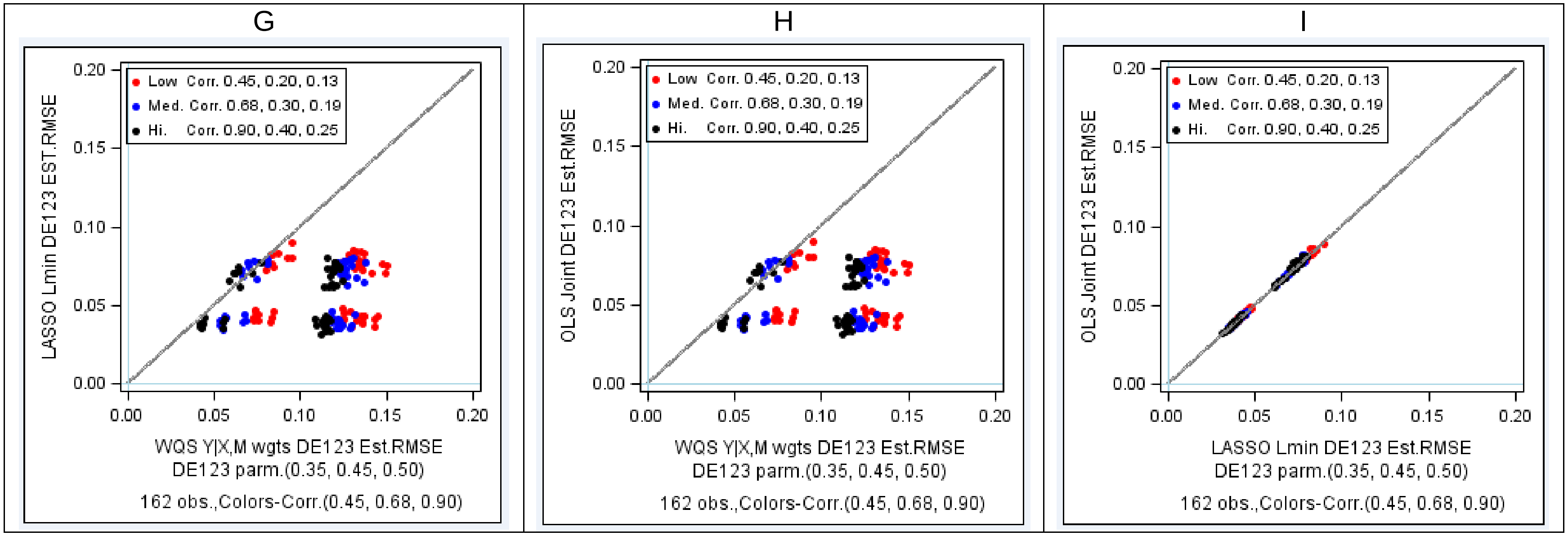


Figure 4.119 Comparison of Joint Direct Effects by WQS, LASSO and OLS A-C) $DE_{123}^{Estimate}{}_{WQS_{Y|X,M}}$ vs. $DE_{123}^{Estimate}{}_{OLS}$ vs. $DE_{123}^{Estimate}{}_{L_{\min.}}$, D-F) $DE_{123}^{Bias}{}_{WQS_{Y|X,M}}$ vs. $DE_{123}^{Bias}{}_{OLS}$ vs. $DE_{123}^{Bias}{}_{L_{\min.}}$, G-I) $DE_{123}^{RMSE}{}_{WQS_{Y|X,M}}$ vs. $DE_{123}^{RMSE}{}_{OLS}$ vs. $DE_{123}^{RMSE}{}_{L_{\min.}}$

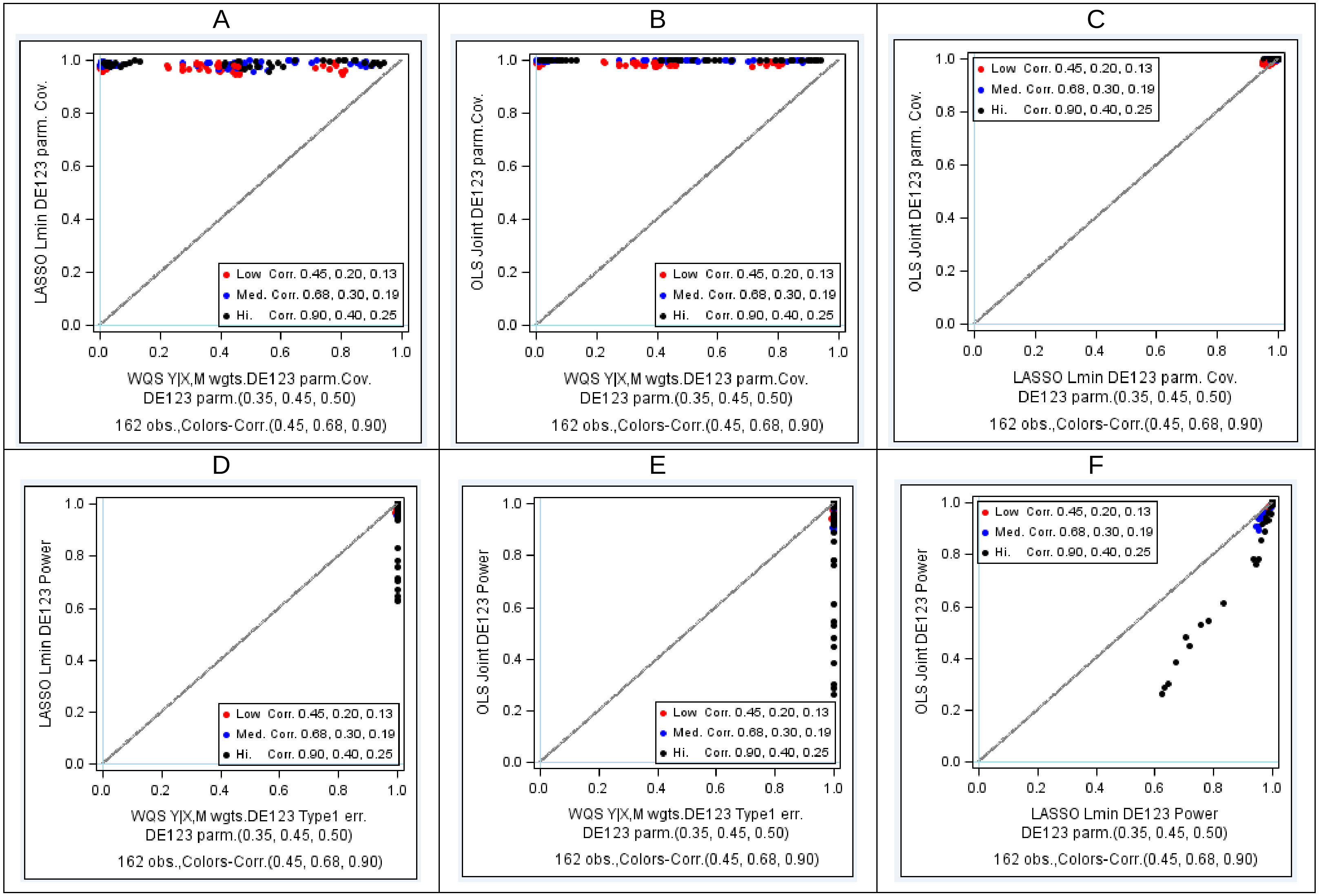


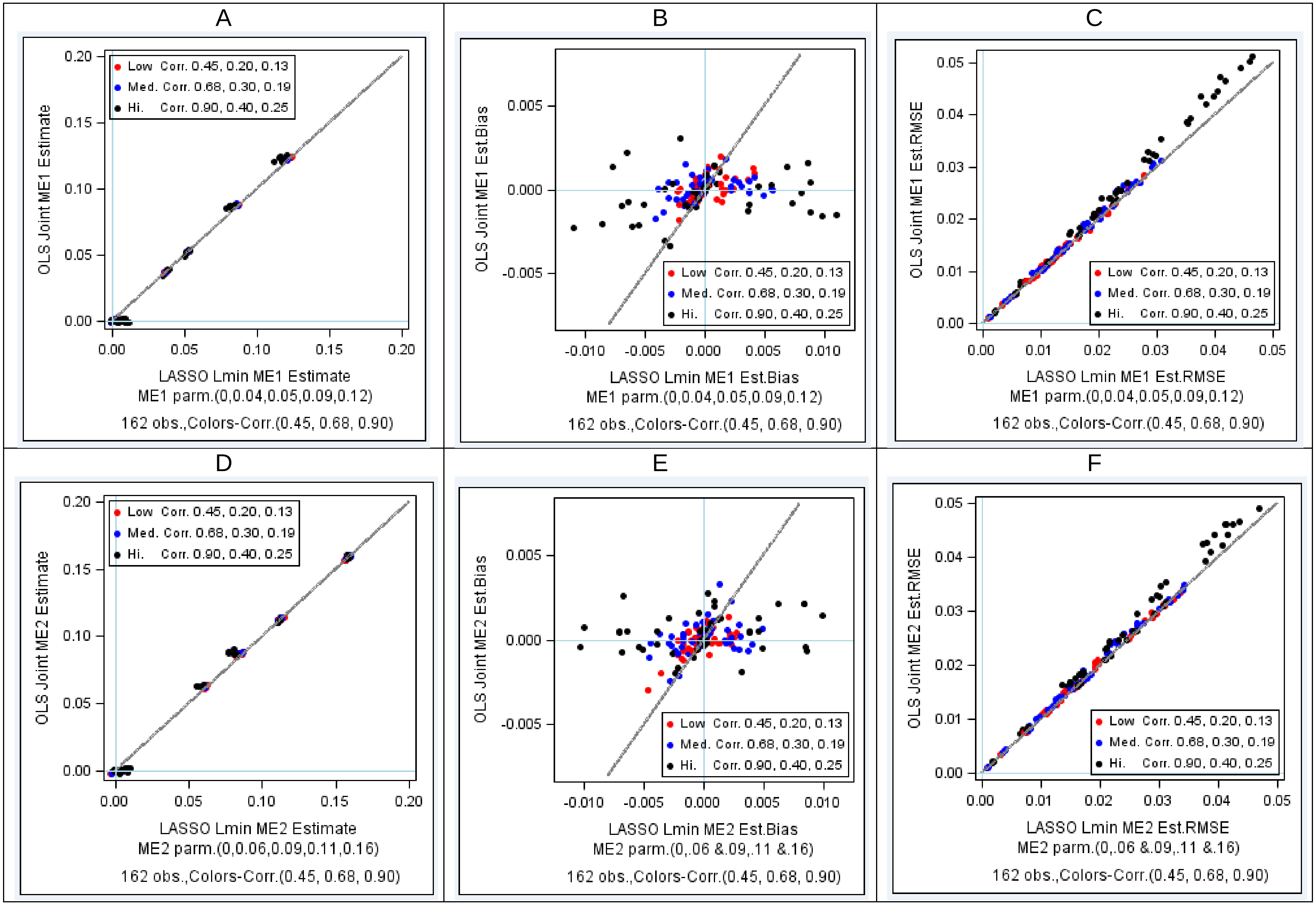
Figure 4.120 Comparison of Joint Direct Effects by WQS, LASSO and OLS A-C) $\overset{Cov.\,prob.}{DE_{123}^{WQS_{Y|X,M}}}$ vs. $\overset{Cov.\,prob.}{DE_{123}^{OLS}}$ vs. $\overset{Cov.\,prob.}{DE_{123}^{L_{\min}.}}$,
D-F) $\overset{Power}{DE_{123}^{WQS_{Y|X,M}}}$ vs. $\overset{Power}{DE_{123}^{OLS}}$ vs. $\overset{Power}{DE_{123}^{L_{\min}.}}$

In summary, the LASSO L_{min} method proves to be a better method for estimating and detecting the joint direct effects as compared to the OLS method, given the results of this simulation study of 162 conditions with pre-set parameter values for *M|X, Y|X, M*, $\rho_{12}, \rho_{13}, \rho_{23}$, and sample size. The WQS method $WQS_{index}^{Y|X,\,M}$ seems preferable to OLS and LASSO in detecting joint direct effect because the direct effect parameters ($\beta_{1,}\ \beta_{2,}\ \beta_{3}$) are large, non-zero parameter values. However, the WQS method is not preferred over LASSO and OLS when estimating the joint direct effect since the coverage probability shown in Figure 4.120 A and B is low and sometimes zero for several of the 162 conditions.

4.5.3 Comparison of 3-Variable Indiv. Indirect Effect for WQS, LASSO & OLS

The individual 3-predictor indirect effect's *Estimate, Bias, RMSE,* and Coverage Probability can only be discussed for the LASSO and OLS methods, since the WQS method only provides a joint indirect effect and no estimates for an individual predictor's estimate, bias, *RMSE* or coverage probability can be obtained. The individual indirect effects using only the OLS and LASSO methods are compared for their average estimates, estimate's bias and estimate's *RMSE* as shown in Figure 4.121. When the pairwise correlations are high (black), the LASSO method adds a (±0.010) bias, four times the bias existing in the OLS method (±0.0025), for each individual indirect effect. As a consequence the LASSO method has slightly reduced *RMSE* for the estimates (black), as shown in Figure 4.121 C and F, as compared to the OLS method. The bias and *RMSE* values are smaller for ME_3, since the correlations associated with X_3 are small $\rho_{13} = 0.20, \rho_{23} = 0.13$.

The Figure 4.122 shows the type1 error rates for the individual indirect effects grouped by predictor correlations in a pairwise comparison between WQS, LASSO, and OLS methods. Panels C, F and I show that the OLS method is conservative on the type1 errors with 0 exceptions for ME_1 and one each for ME_2 and ME_3, while LASSO has 4 exceptions ($ME_1^{TYP1}{}_{L_{\min}}$ type1 error >0.075) shown in Figure 4.122 A and C, caused by small sized datasets with high multicollinearity (black). In Figure 4.122 A and B, D and E, and G and H, where the cut-off value is zero, there are 16 exceptions for $ME_1^{TYP1}{}_{WQS_{M|X}}$ (type1 error > 0.075), and in Figure 4.123 X1.20 showed 10 exceptions for $ME_1^{TYP1}{}_{WQS_{M|X}}$ with a cut-off =0.2, in Figure 4.123 X1.30 showed 6 exceptions for $ME_1^{TYP1}{}_{WQS_{M|X}}$ with a cut-off =0.30 and in Figure 4.123 X1.40 showed 4 exceptions for $ME_1^{TYP1}{}_{WQS_{M|X}}$ with a cut-off- 0.40. When the cut-off value is set to zero there were 16 exceptions for $ME_2^{TYP1}{}_{WQS_{M|X}}$ and $ME_3^{TYP1.}{}_{WQS_{M|X}}$, there were 10 exceptions for $ME_2^{TYP1}{}_{WQS_{M|X}}$ with the cut-off =0.2, 8 exceptions for $ME_2^{TYP1}{}_{WQS_{M|X}}$ with the cut-off =0.30 and 6 exceptions for $ME_2^{TYP1}{}_{WQS_{M|X}}$ with the cut-off= 0.40 as shown in in Figure 4.123 X2.20 through X2.40. There were 10 exceptions for $ME_3^{TYP1.}{}_{WQS_{M|X}}$ with the cut-off =0.2, 8 exceptions for $ME_3^{TYP1.}{}_{WQS_{M|X}}$ with the cut-off =0.30 and 4 exceptions for $ME_3^{TYP1.}{}_{WQS_{M|X}}$ with the cut-off= 0.40 as shown in in Figure 4.123 X3.20 through X3.40.


A
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
OLS Joint ME1 Estimate
LASSO Lmin ME1 Estimate
ME1 parm.(0,0.04,0.05,0.09,0.12)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
B
OLS Joint ME1 Est.Bias
LASSO Lmin ME1 Est.Bias
ME1 parm.(0,0.04,0.05,0.09,0.12)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
C
OLS Joint ME1 Est.RMSE
LASSO Lmin ME1 Est.RMSE
ME1 parm.(0,0.04,0.05,0.09,0.12)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
D
OLS Joint ME2 Estimate
LASSO Lmin ME2 Estimate
ME2 parm.(0,0.06,0.09,0.11,0.16)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
E
OLS Joint ME2 Est.Bias
LASSO Lmin ME2 Est.Bias
ME2 parm.(0,.06 &.09,.11 &.16)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
F
OLS Joint ME2 Est.RMSE
LASSO Lmin ME2 Est.RMSE
ME2 parm.(0,.06 &.09,.11 &.16)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)

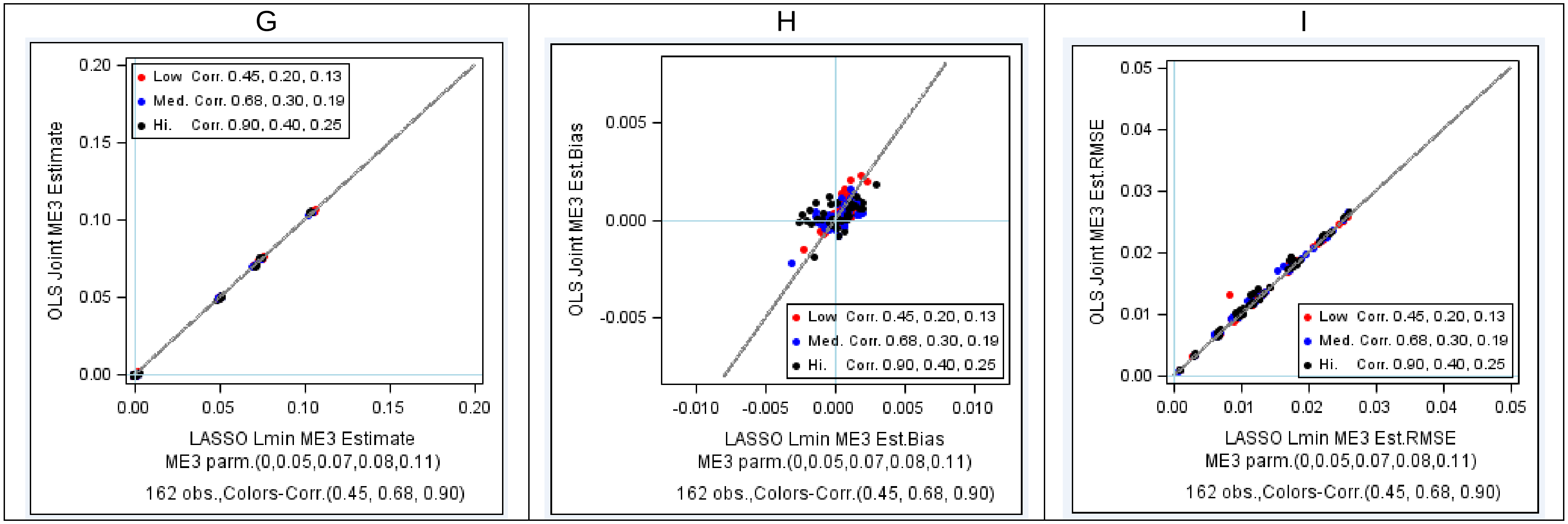
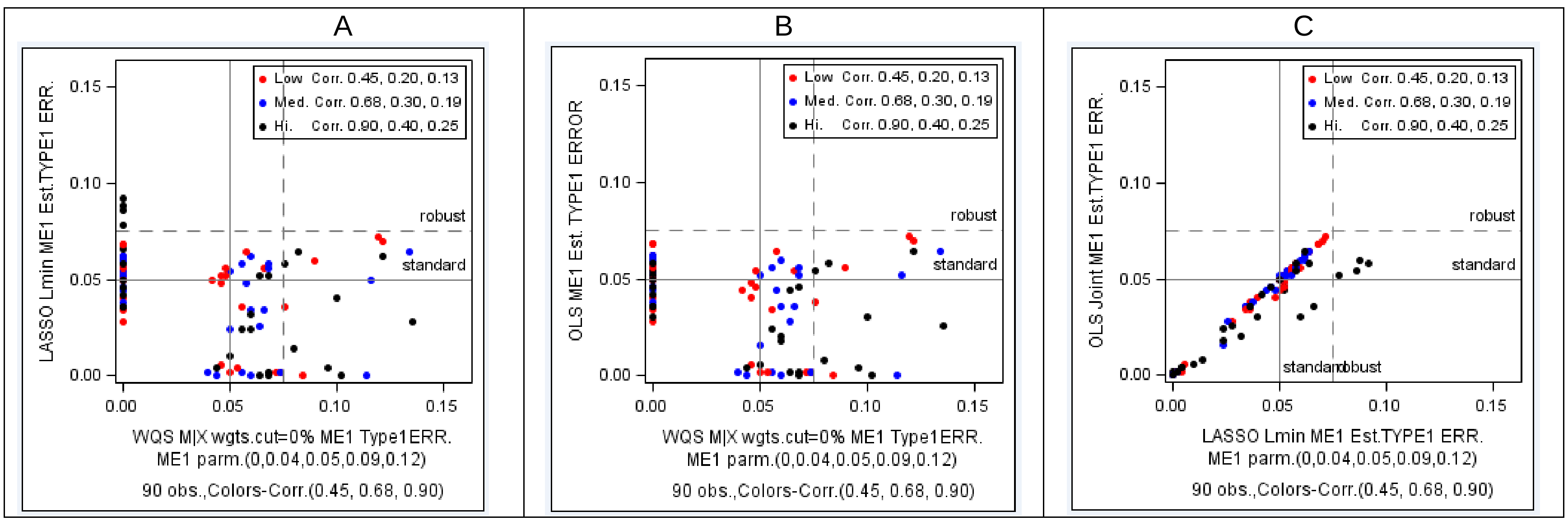


Figure 4.121 Comparison of Individual Indirect Effects LASSO vs. OLS A-I) $ME^{EST.}_{OLS\,1,2,3} vs. ME^{EST.}_{L_{\min}\,1,2,3}, ME^{Bias}_{OLS\,1,2,3} vs. ME^{Bias}_{L_{\min}\,1,2,3}, ME^{RMSE}_{OLS\,1,2,3} vs. ME^{RMSE}_{L_{\min}\,1,2,3}$



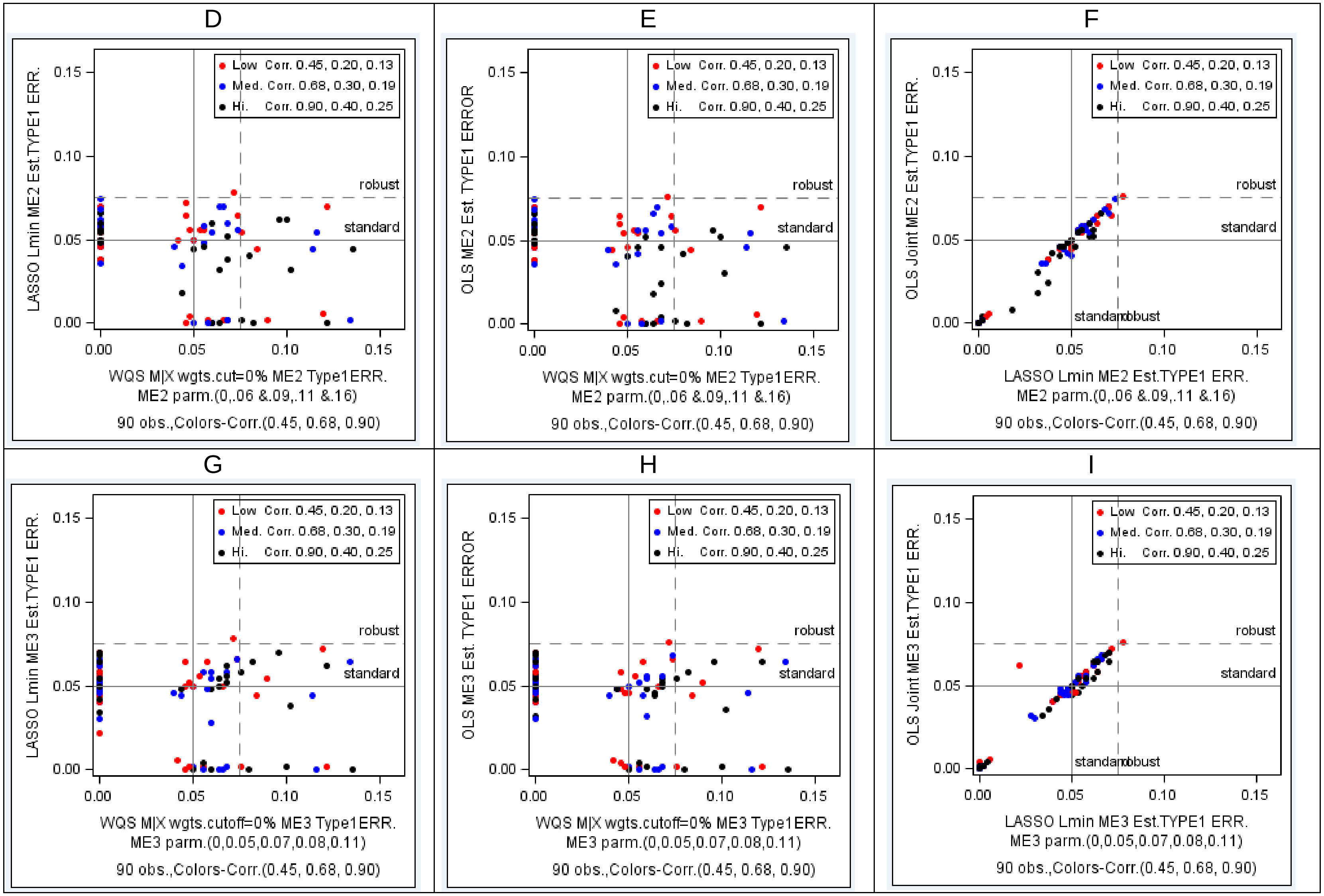
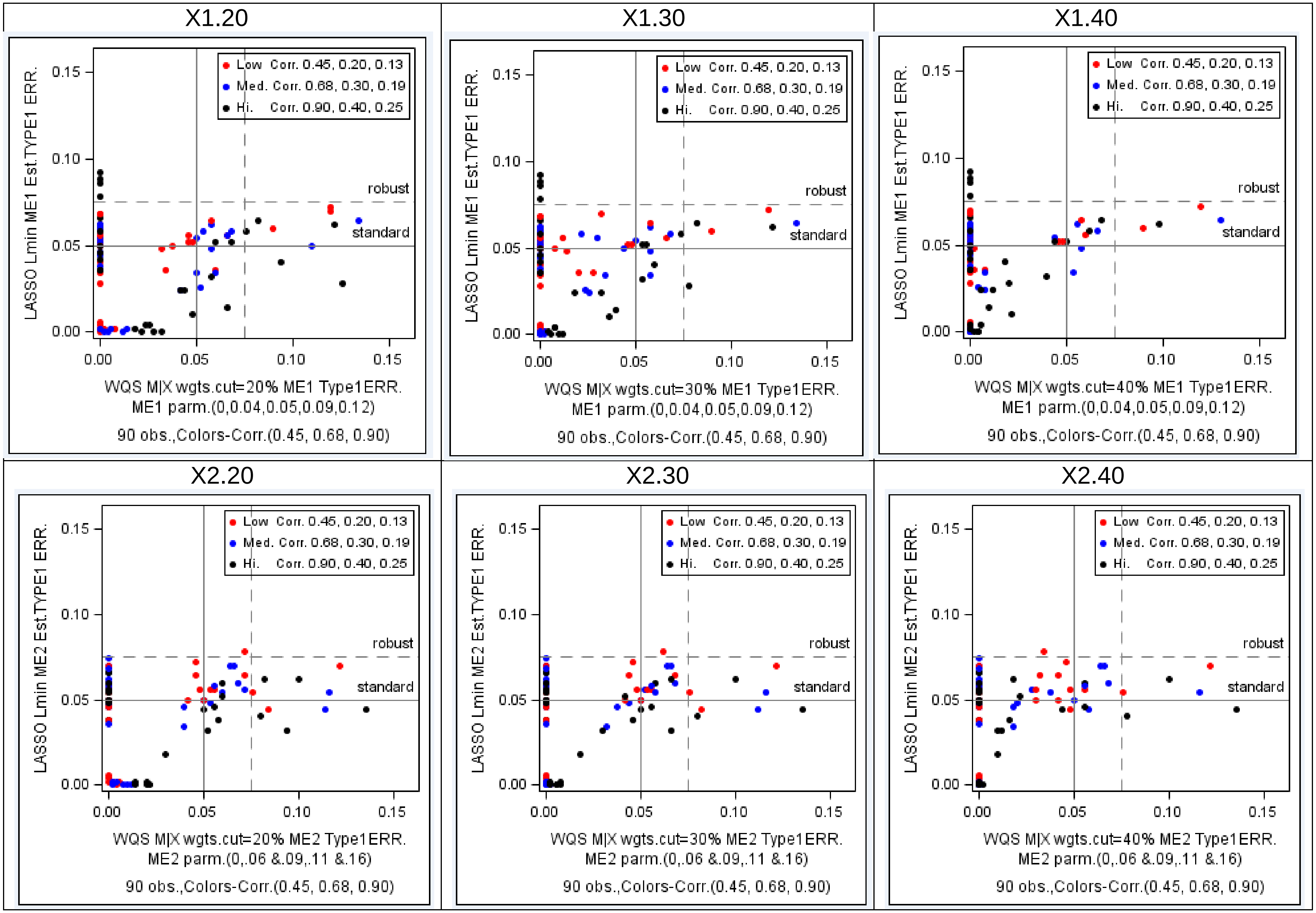


Figure 4.122 Comparison of Indiv. Indirect Type1 Err. by WQS, LASSO, & OLS A-C) $ME_1^{\substack{Type1\ Err.\\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Type1\ Err.\\ OLS}}$ vs. $ME_1^{\substack{Type1\ Err.\\ L_{\min}}}$, D-F) $ME_2^{\substack{Type1\ Err.\\ WQS_{M|X}}}$ vs. $ME_2^{\substack{Type1\ Err.\\ OLS}}$ vs. $ME_2^{\substack{Type1\ Err.\\ L_{\min}}}$, G-I) $ME_1^{\substack{Type1\ Err.\\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Type1\ Err.\\ OLS}}$ vs. $ME_1^{\substack{Type1\ Err.\\ L_{\min}}}$


X1.20
X1.30
X1.40
X2.20
X2.30
X2.40
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
robust
standard
0.00
0.05
0.10
0.15
LASSO Lmin ME1 Est.TYPE1 ERR.
WQS M|X wgts.cut=20% ME1 Type1ERR.
WQS M|X wgts.cut=30% ME1 Type1ERR.
WQS M|X wgts.cut=40% ME1 Type1ERR.
ME1 parm.(0,0.04,0.05,0.09,0.12)
LASSO Lmin ME2 Est.TYPE1 ERR.
WQS M|X wgts.cut=20% ME2 Type1ERR.
WQS M|X wgts.cut=30% ME2 Type1ERR.
WQS M|X wgts.cut=40% ME2 Type1ERR.
ME2 parm.(0,.06 &.09,.11 &.16)
90 obs.,Colors-Corr.(0.45, 0.68, 0.90)

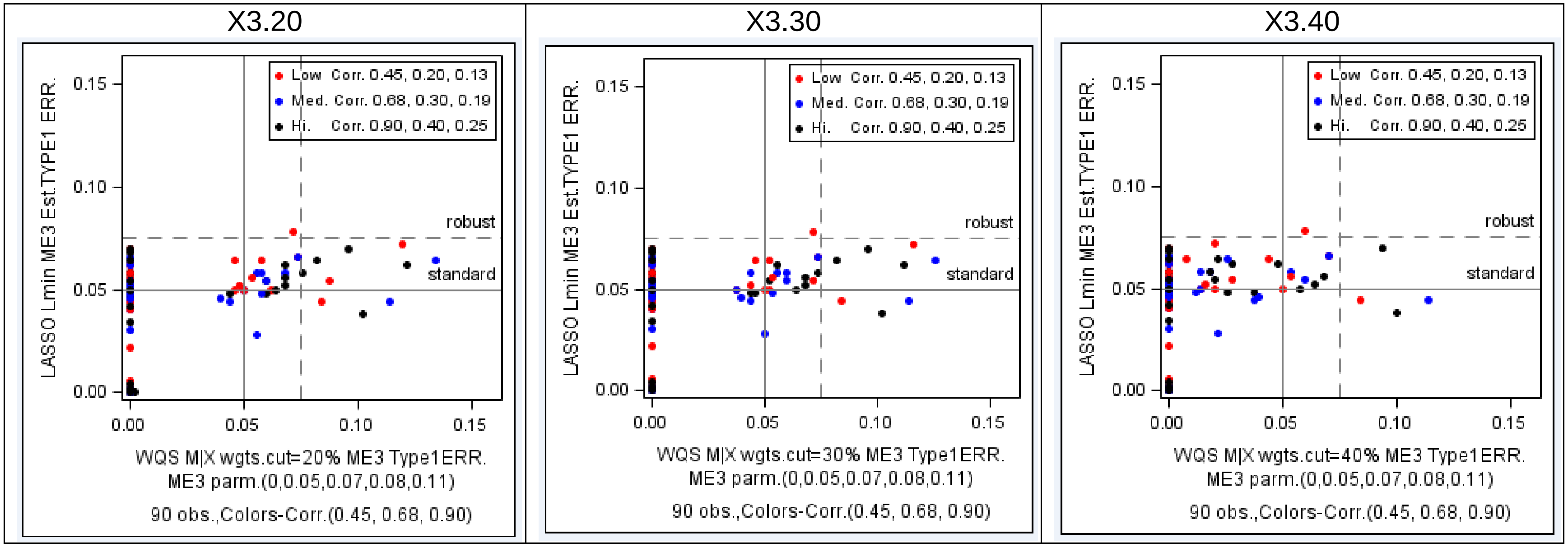
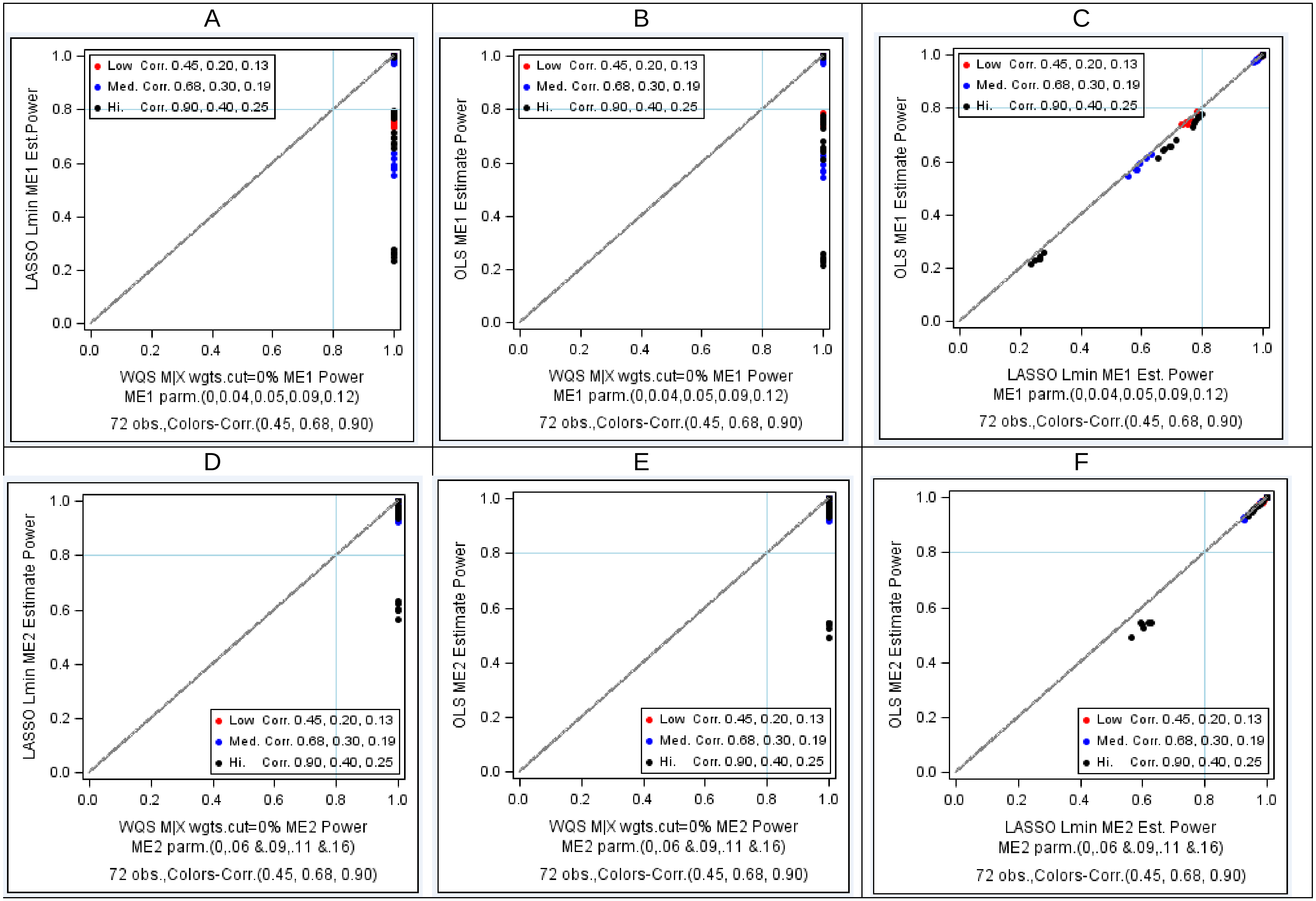


Figure 4.123 Comparison of Indiv. Indirect Type1 Errors LASSO vs. WQS X_{1_20} to X_{1_40}) $ME_1^{\substack{Typ1\_20 \\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Typ1\_30 \\ WQS_{M|X}}}$ vs. $ME_1^{\substack{Typ1\_40 \\ WQS_{M|X}}}$, X_{2_20} to X_{2_40}) $ME_2^{\substack{Typ1\_20 \\ WQS_{M|X}}}$ vs. $ME_2^{\substack{Typ1\_30 \\ WQS_{M|X}}}$ vs. $ME_{2\theta}^{\substack{Typ1\_40 \\ WQS_{M|X}}}$, and X_{3_20} to X_{3_40}) $ME_3^{\substack{Typ1\_20 \\ WQS_{M|X}}}$ vs. $ME_3^{\substack{Typ1\_30 \\ WQS_{M|X}}}$ vs. $ME_3^{\substack{Typ1\_40 \\ WQS_{M|X}}}$

Even when the cut-off value was set to 0.40 as the limit below which an individual predictor's WQS weight is considered to be 0 in the determination of the joint indirect effect, there were (4 exceptions for X_1 + 6 exceptions for X_2 + 4 exceptions for X_3), 14 exceptions in a total of 162 conditions i.e. $6\frac{2}{3}\%$ of the conditions $\theta_{123} = 0.60, \gamma = 0$ with type1 error rates > 0.075. When the WQS method gives a positive bias to the null $\hat{\gamma}$ estimate, the high value for the joint theta parameter amplifies the individual indirect effect causing its deviation from zero and causing a type 1 error rate. Other than for the null effect conditions, the WQS method can be a comparable mediation analysis method, used with the proper selection of a cut-off rate for the individual predictor WQS weights, which will reduce the number of exceptions for individual type1 errors.

The individual indirect effect's statistical power using the WQS, LASSO, and OLS are compared in Figure 4.124. Panels C and E show that for $ME_{1}^{\overset{PWR}{OLS}} vs. ME_{1}^{\overset{PWR}{L_{\min}}}$, the lowest power was evident for 6 of the 72 conditions, having small sized datasets with highly correlated predictors causing multicollinearity effects in the regression estimates and reduced power. By comparison the power for all the individual indirect effect cases for $ME_{1,2,3}^{\overset{PWR}{M|X}}$ were almost 1.

Figure 4.125 shows that as the cut-off rate for the WQS weights is increased from 0.20 to 0.40 in steps of 0.1, the WQS method's power for each set of 72 conditions drops markedly, making LASSO the preferred method at a cut-off=0.30% for individual WQS weights. Figure 4.125 X1.30, X2.30, and X3.30 show that LASSO has a higher power than the WQS method for detecting an individual predictor's contribution to the joint indirect effect in 3-variable mediation analysis.


A
LASSO Lmin ME1 Est.Power
WQS M|X wgts.cut=0% ME1 Power
ME1 parm.(0,0.04,0.05,0.09,0.12)
72 obs.,Colors-Corr.(0.45, 0.68, 0.90)
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
B
OLS ME1 Estimate Power
C
OLS ME1 Estimate Power
LASSO Lmin ME1 Est. Power
D
LASSO Lmin ME2 Estimate Power
WQS M|X wgts.cut=0% ME2 Power
ME2 parm.(0,.06 &.09,.11 &.16)
E
OLS ME2 Estimate Power
F
OLS ME2 Estimate Power
LASSO Lmin ME2 Est. Power

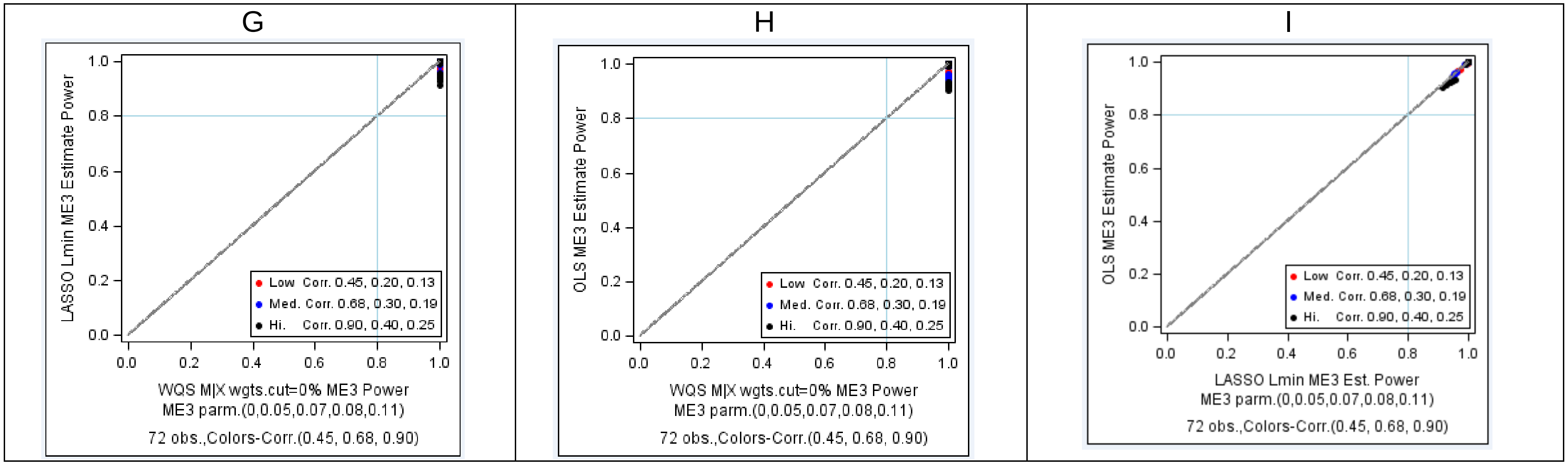
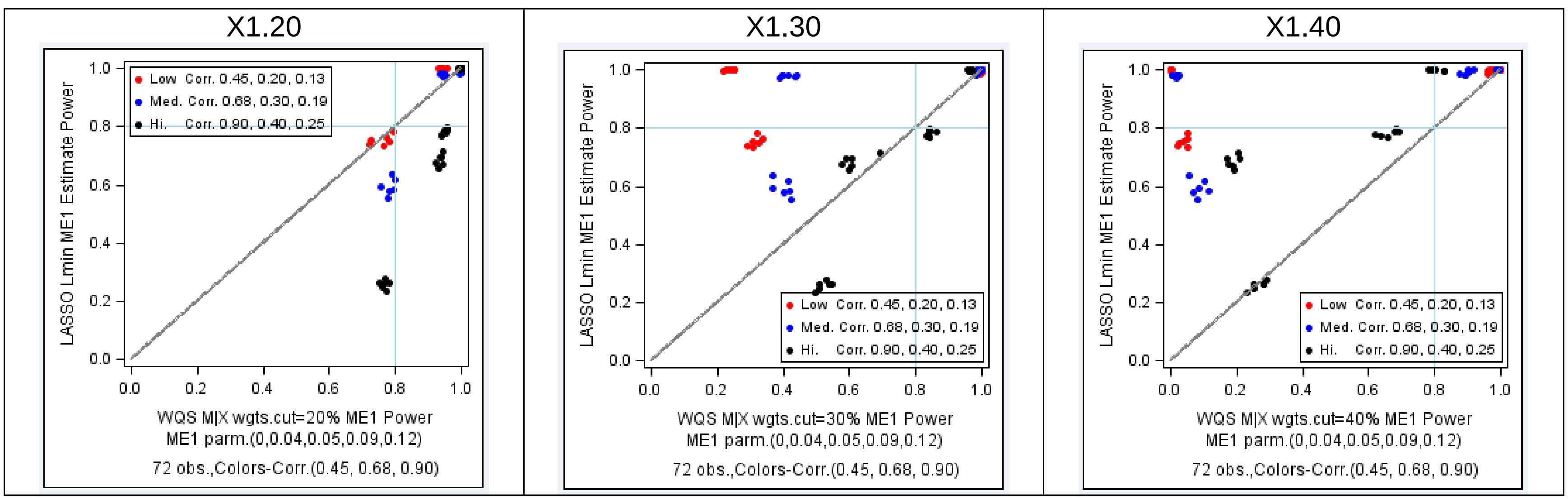


Figure 4.124 Comparison of Power for Indiv. Indirect Effects by WQS, LASSO, and OLS A-C) $ME_1^{\overset{Power}{WQS_{M|X}}}$ vs. $ME_1^{\overset{Power}{OLS}}$ vs. $ME_1^{\overset{Power}{L_{\min}}}$ D-F) $ME_2^{\overset{Power}{WQS_{M|X}}}$ vs. $ME_2^{\overset{Power}{OLS}}$ vs. $ME_2^{\overset{Power}{L_{\min}}}$ G-I) $ME_3^{\overset{Power}{WQS_{M|X}}}$ vs. $ME_3^{\overset{Power}{OLS}}$ vs. $ME_3^{\overset{Power}{L_{\min}}}$



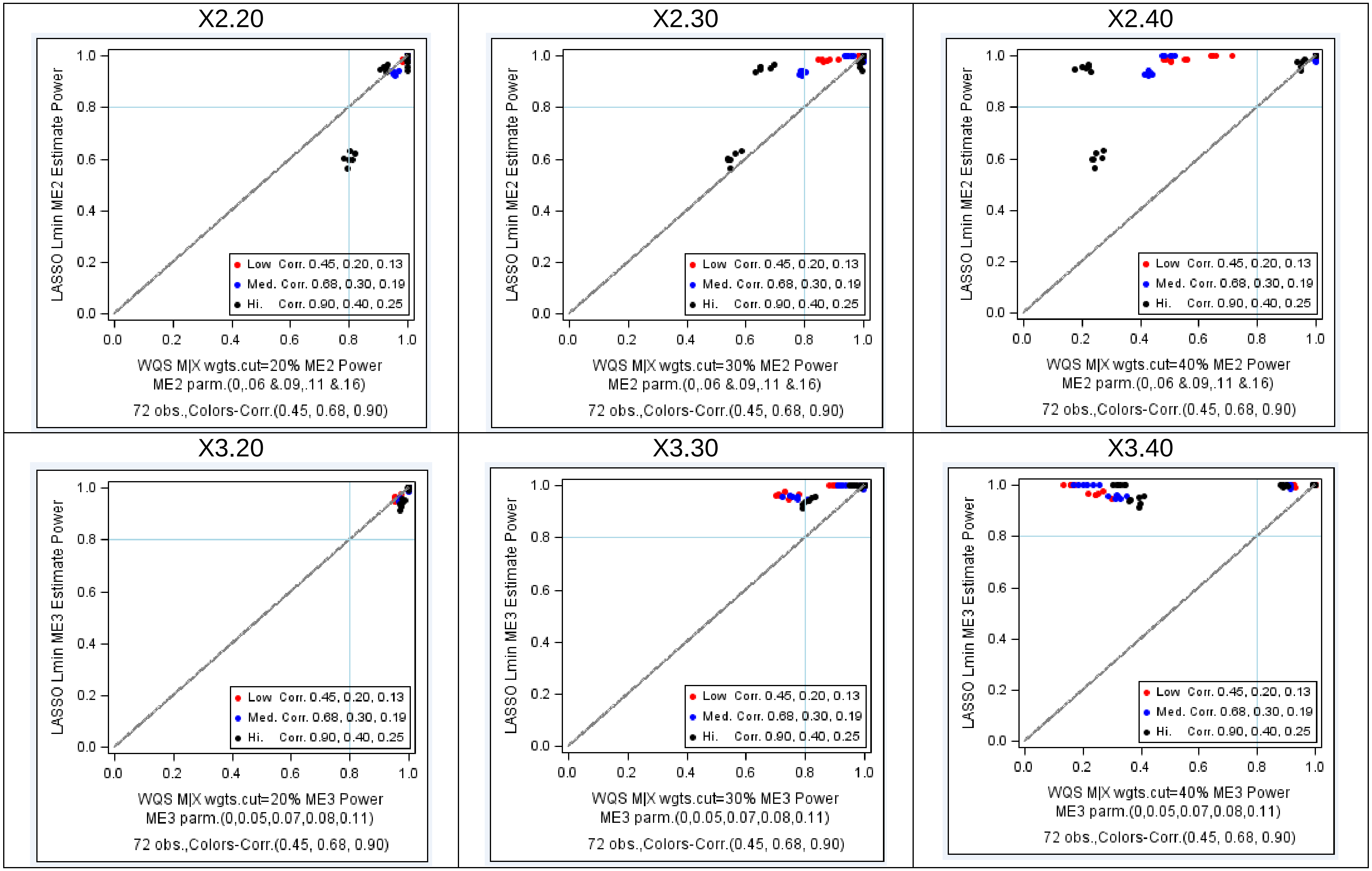


Figure 4.125 Comparison of Indiv. Indirect Effects for WQS, OLS & LASSO, X_{1_20} -X_{3_40} $ME_1^{\overset{PWR\_20}{WQS_{M|X}}}$ vs. $ME_1^{\overset{PWR\_30}{WQS_{M|X}}}$ vs. $ME_1^{\overset{PWR\_40}{WQS_{M|X}}}$, $ME_2^{\overset{PWR\_20}{WQS_{M|X}}}$ vs. $ME_2^{\overset{PWR\_30}{WQS_{M|X}}}$ vs. $ME_2^{\overset{PWR\_40}{WQS_{M|X}}}$, $ME_3^{\overset{PWR\_20}{WQS_{M|X}}}$ vs. $ME_3^{\overset{PWR\_30}{WQS_{M|X}}}$ vs. $ME_3^{\overset{PWR\_40}{WQS_{M|X}}}$

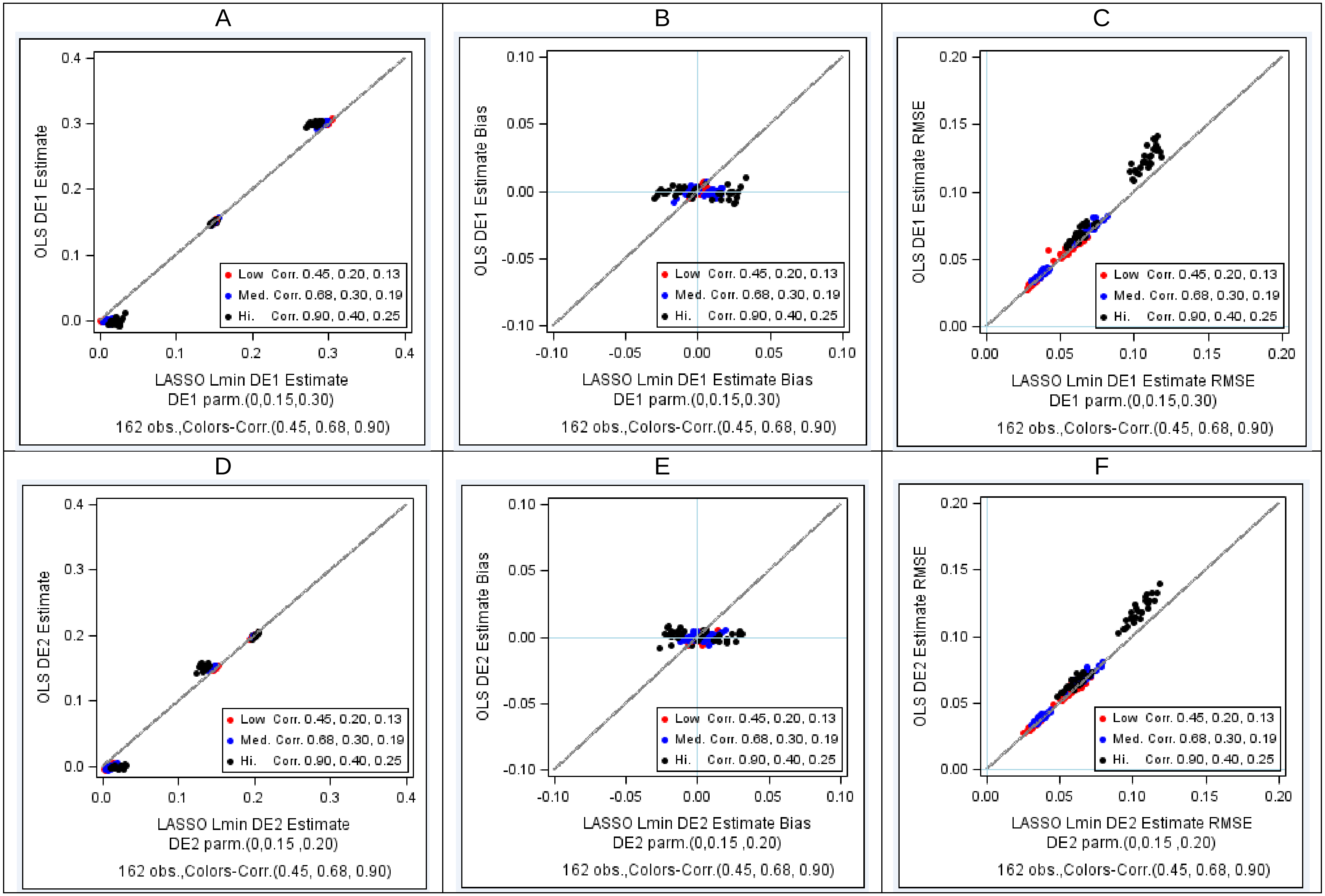
4.5.4 3-Var. Indiv. Direct Effect Est., *Bias*, *RMSE,* Cov., Type1 error & Power

The individual direct effects for X_1, X_2 and X_3 using the OLS and LASSO methods are compared for the average estimate, the estimate's bias and estimate's RMSE. Comparison of the WQS method with LASSO and OLS methods for individual direct effects is not shown because earlier it was found that the WQS method is not preferred over LASSO and OLS when estimating the joint direct effect, because the coverage probability shown in Figure 4.120 A &B is low to zero for several of the conditions, even though the WQS power to estimate the joint direct effect is 1 (Figure 4.120 D & E).

When the pairwise correlations are high (black), the LASSO method adds a bias to each individual direct effect shown in Figure 4.126 B and E. As a consequence of the LASSO method this additional bias is traded for a slightly reduced *RMSE* of the individual direct effect estimates (black cluster), as shown in panels C and F. The bias and *RMSE* values are smaller for ME_3, since the correlations associated with X_3 are small $\rho_{13} = 0.20, \rho_{23} = 0.13$.

The type1 error rate for the individual direct effects using the WQS method $DE_{1,2,3}^{\underset{WQS_{Y|X,M}}{TYP1}}$ is near 0 for all the conditions under considerations. The OLS method has 1 exception for $DE_{1}^{\underset{OLS}{TYP1}} > 0.075$ shown in Figure 4.127 B and C and is the conservative as compared to the LASSO method which has 14 exceptions to the 54 conditions with DE_1=0, making the LASSO method unsuitable to detect the null individual direct effects. The OLS method is more suitable for individual direct effects in the 3-variable mediation analysis, since it utilizes the limit of the type1 errors better than the WQS method but with poor statistical power.

The Figure 4.128 shows that the WQS method has the higher power detecting individual direct effects (since they have large, non-zero beta parameters) as compared to OLS and LASSO shown in panels C and F. However, the coverage probability for the WQS method in estimating the true joint direct effect was low, with several conditions having zero coverage and was not the preferred method as shown in Figure 4.120.
The Figure 4.129 shows that as the cut-off rate for the WQS weights is increased from 0.20 to 0.40 in steps of 0.1, the power for each set of 108 conditions having a non-zero direct effect drops but the WQS method has a higher power than the LASSO method for 3-variable individual predictor mediation analysis of individual direct effects if the cut-off value is limited to 0.30.


A
OLS DE1 Estimate
LASSO Lmin DE1 Estimate
DE1 parm.(0,0.15,0.30)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
B
OLS DE1 Estimate Bias
LASSO Lmin DE1 Estimate Bias
DE1 parm.(0,0.15,0.30)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
C
OLS DE1 Estimate RMSE
LASSO Lmin DE1 Estimate RMSE
DE1 parm.(0,0.15,0.30)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
D
OLS DE2 Estimate
LASSO Lmin DE2 Estimate
DE2 parm.(0,0.15 ,0.20)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
E
OLS DE2 Estimate Bias
LASSO Lmin DE2 Estimate Bias
DE2 parm.(0,0.15 ,0.20)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)
F
OLS DE2 Estimate RMSE
LASSO Lmin DE2 Estimate RMSE
DE2 parm.(0,0.15 ,0.20)
162 obs.,Colors-Corr.(0.45, 0.68, 0.90)

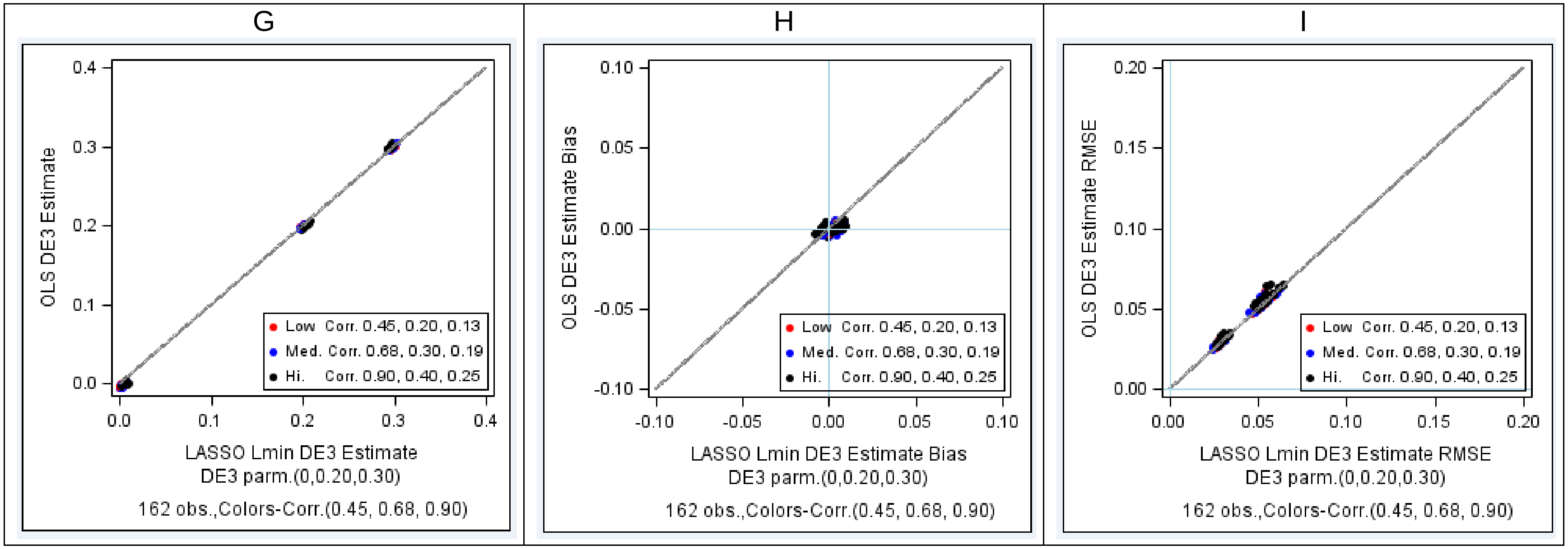
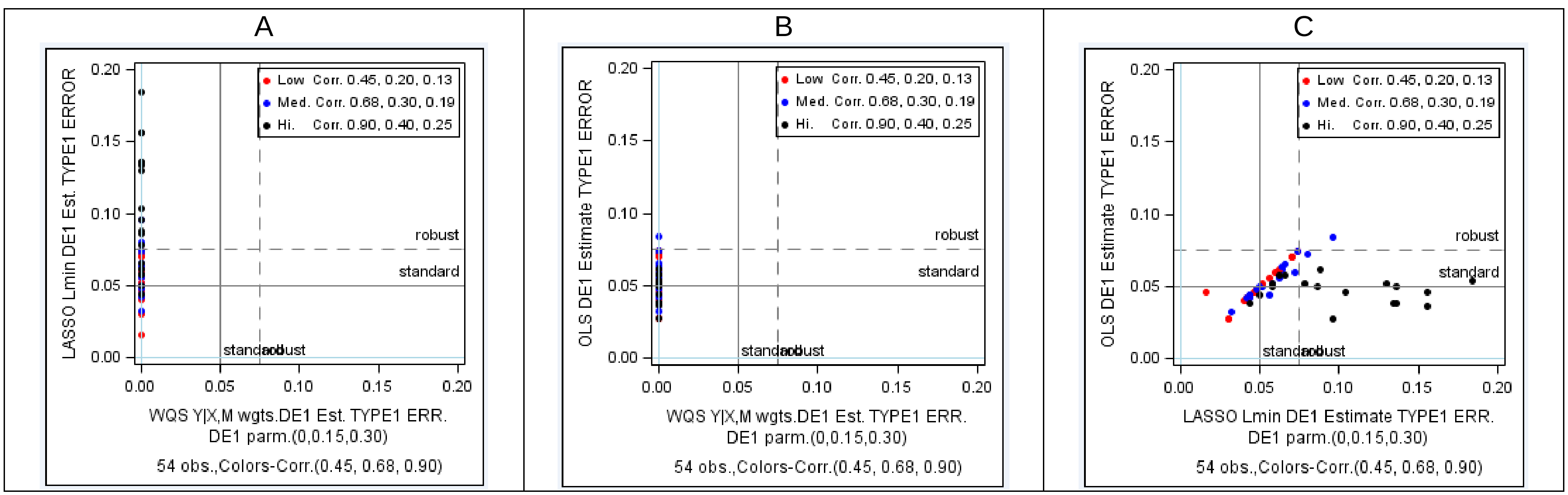


Figure 4.126 Indiv. Direct Effect Comparisons between LASSO & OLS A-I) $DE^{EST.}_{OLS\;1,2,3}$ vs. $DE^{EST.}_{L_{\min}\;1,2,3}$, $DE^{Bias}_{OLS\;1,2,3}$ vs. $DE^{Bias}_{L_{\min}\;1,2,3}$, $DE^{RMSE}_{OLS\;1,2,3}$ vs. $DE^{RMSE}_{L_{\min}\;1,2,3}$



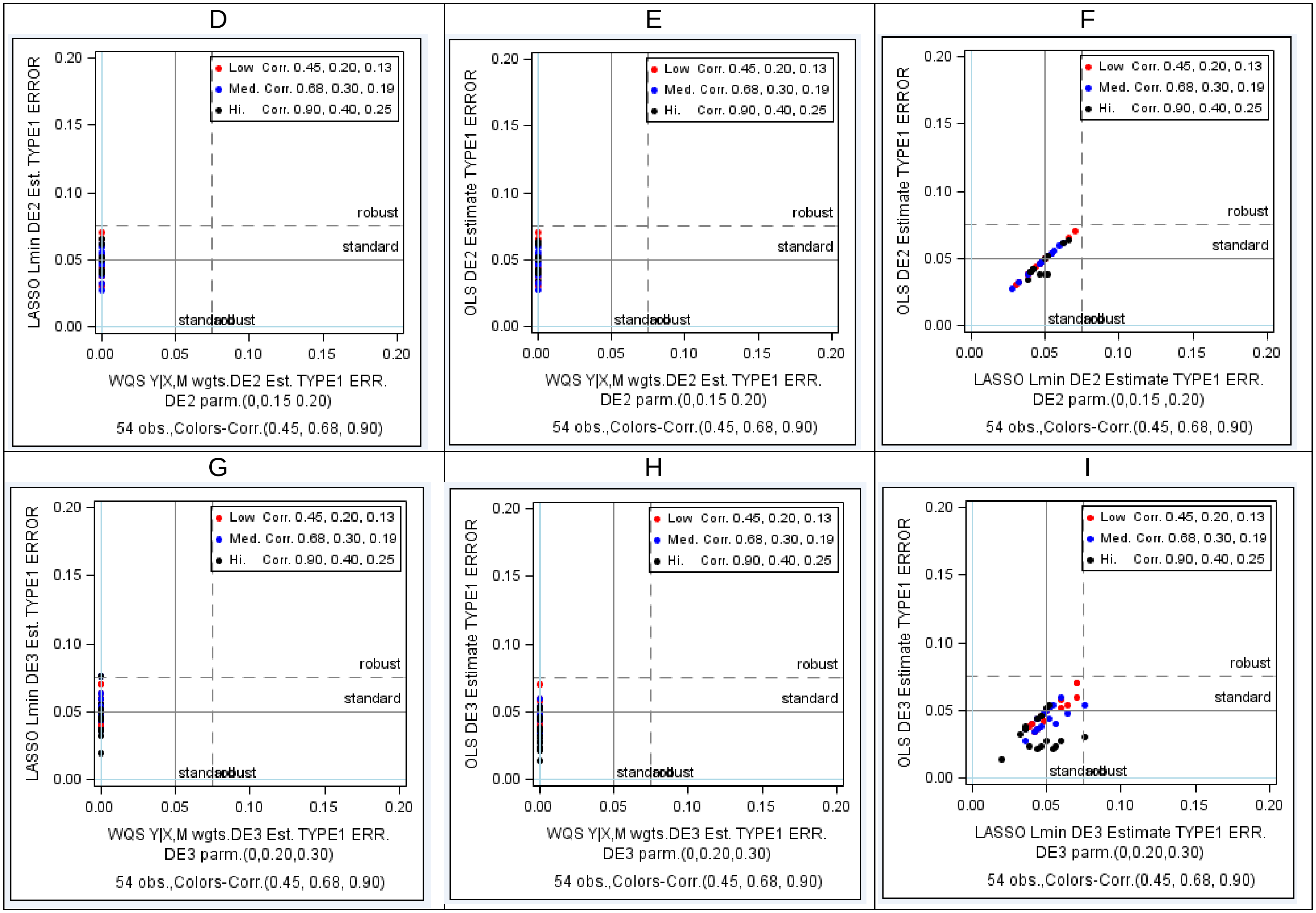
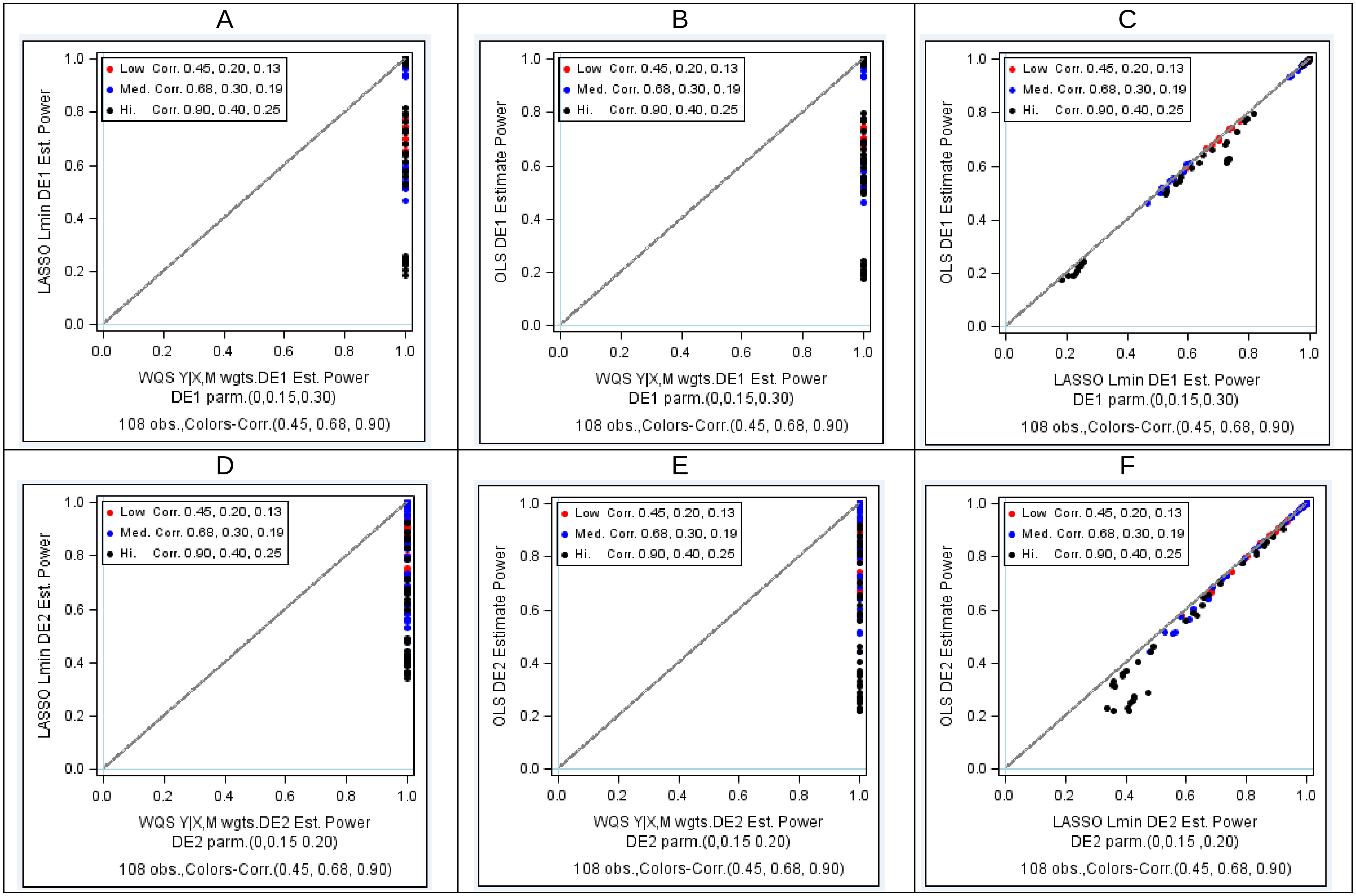


Figure 4.127 Comparison of Indiv. Direct Effect Type1 Err. by WQS, LASSO and OLS A-I) $DE_{1,2,3}^{\substack{TYP1\\WQS_{Y|X,M}}}$ vs. $DE_{1,2,3}^{\substack{TYP1\\OLS}}$ vs. $DE_{1,2,3}^{\substack{TYP1\\L_{\min}}}$


A
LASSO Lmin DE1 Est. Power
WQS Y|X,M wgts.DE1 Est. Power
DE1 parm.(0,0.15,0.30)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
B
OLS DE1 Estimate Power
WQS Y|X,M wgts.DE1 Est. Power
DE1 parm.(0,0.15,0.30)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
C
OLS DE1 Estimate Power
LASSO Lmin DE1 Est. Power
DE1 parm.(0,0.15,0.30)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
D
LASSO Lmin DE2 Est. Power
WQS Y|X,M wgts.DE2 Est. Power
DE2 parm.(0,0.15 0.20)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
E
OLS DE2 Estimate Power
WQS Y|X,M wgts.DE2 Est. Power
DE2 parm.(0,0.15 0.20)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
F
OLS DE2 Estimate Power
LASSO Lmin DE2 Est. Power
DE2 parm.(0,0.15 ,0.20)
108 obs.,Colors-Corr.(0.45, 0.68, 0.90)
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25

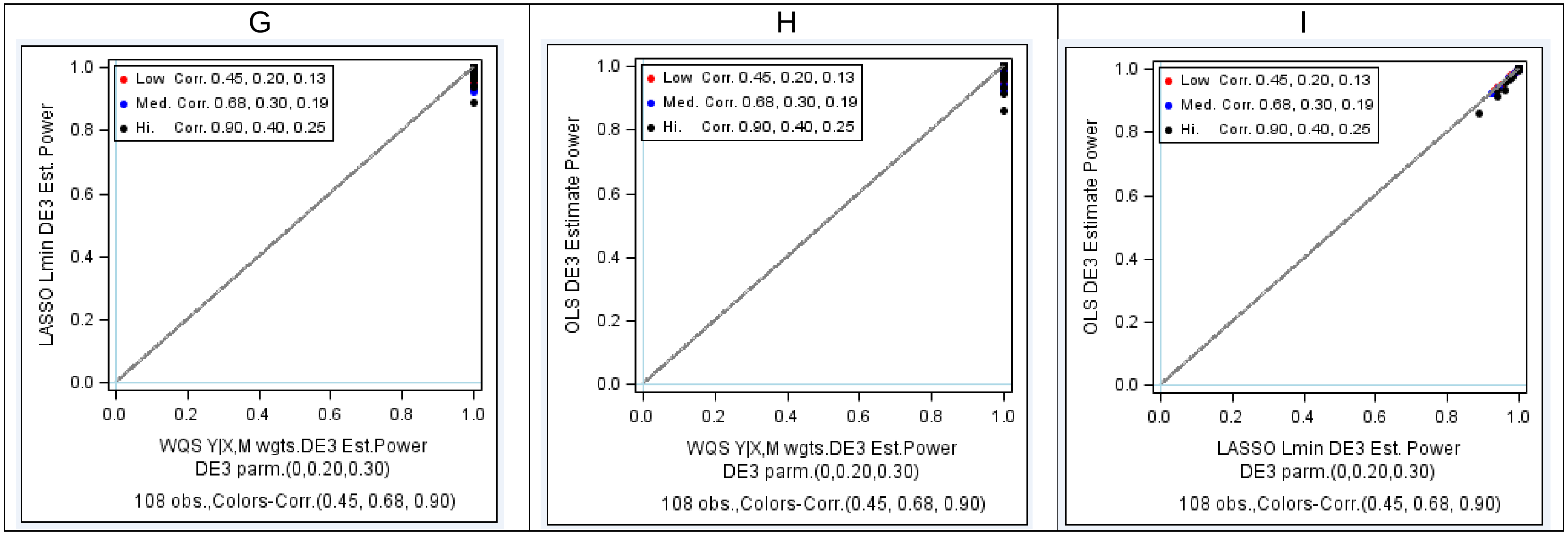
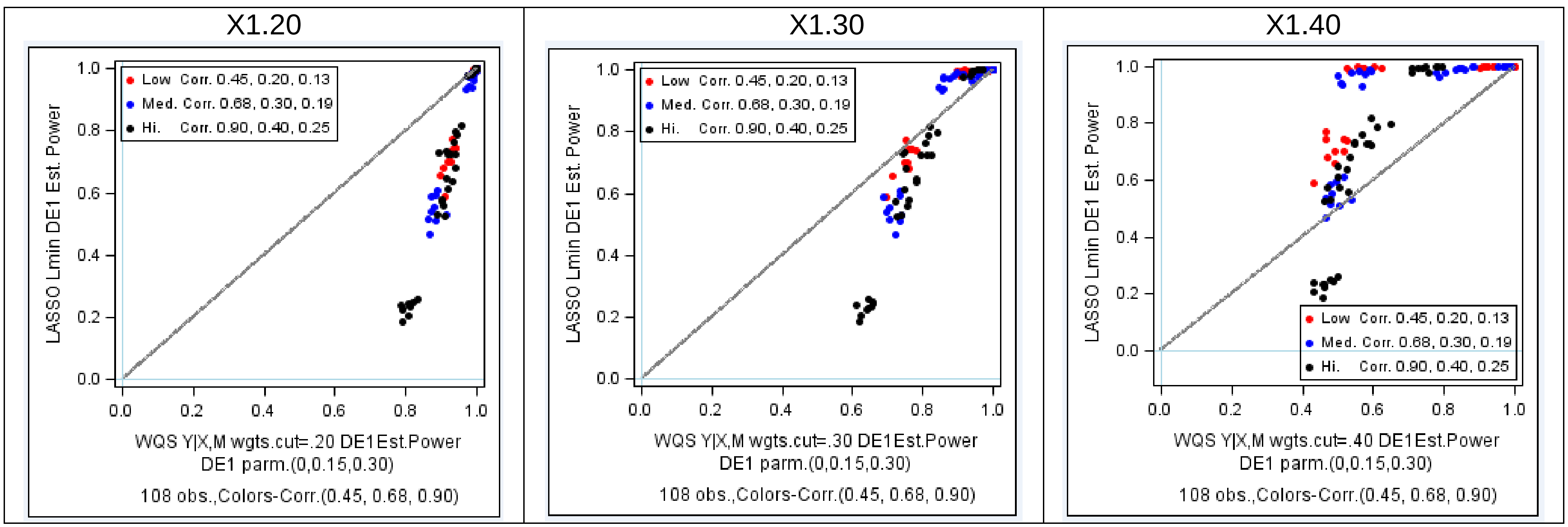


Figure 4.128 Compare of Individual Direct Effect Power by WQS, LASSO and OLS A-I) $DE^{PWR}_{WQS_{M|X}\,1,2,3}$ *vs.* $DE^{PWR}_{OLS\,1,2,3}$ *vs.* $DE^{PWR}_{L_{\min}\,1,2,3}$



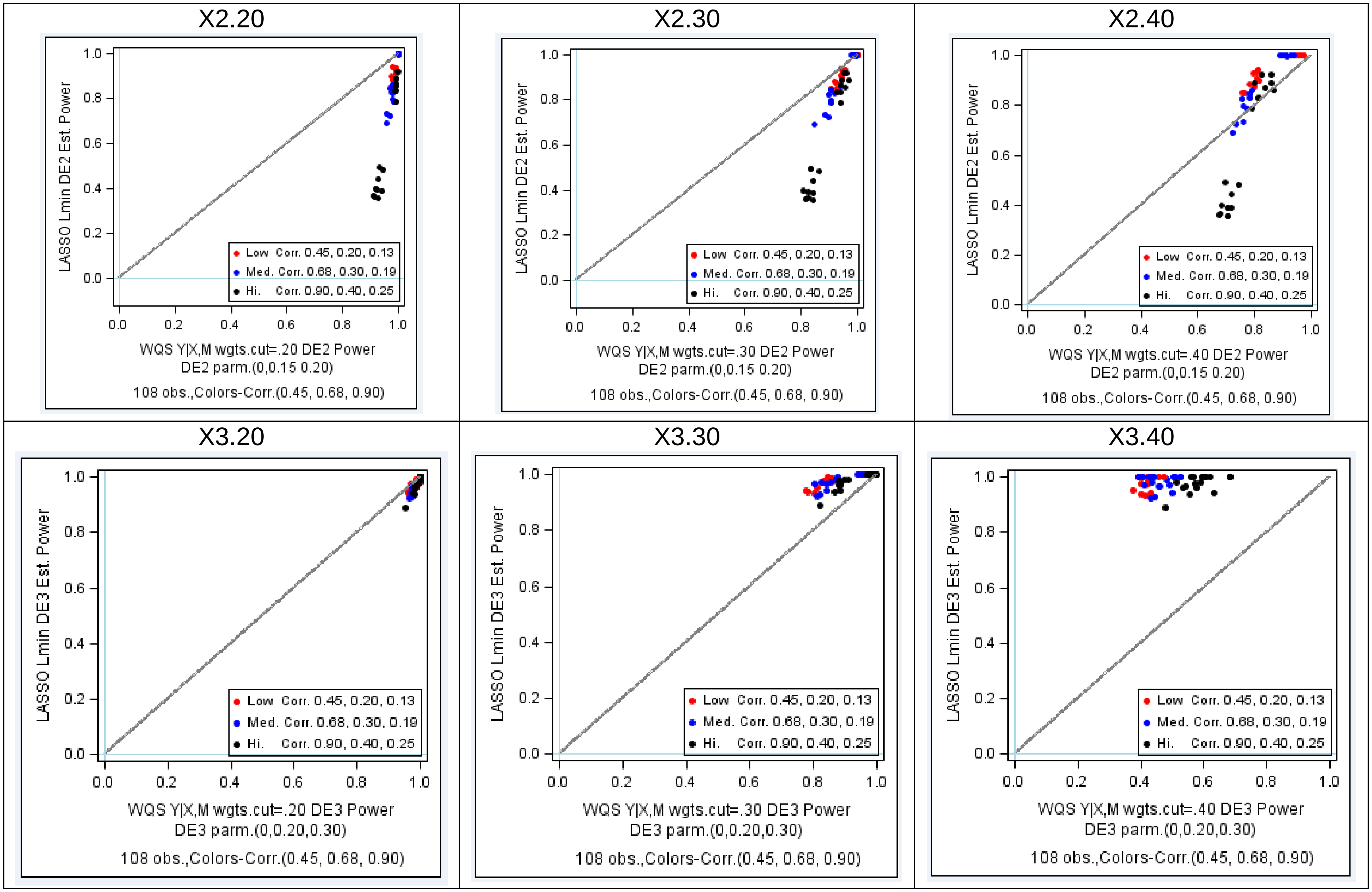


Figure 4.129 Comparison of Indiv. Direct Effect by LASSO vs. WQS Methods X_{1_20} to X_{3_40} $DE_{1,2,3}^{\substack{PWR\_20\\WQS_{M|X}}}$ vs. $DE_{1,2,3}^{\substack{PWR\_30\\WQS_{M|X}}}$ vs. $DE_{1,2,3}^{\substack{PWR\_40\\WQS_{M|X}}}$

The summary for X_1 & X_2 and X_1, X_2 & X_3 predictor mediation analysis for combined and individual mediated effects using estimate bias, estimate *RMSE*, coverage probability, type1 error and power statistics to summarize WQS LASSO and OLS regressions are shown in Table 4.20 A for indirect effects and 4.20B for direct effects. The individual mediated effects use only the type1 error and power statistics to evaluate the method's performance since WQS regression method does not estimate the individual predictor's indirect or direct effects.

The OLS method performed consistently well out of the six methods considered in the simulation studies for 2&3 predictor mediation analysis, when estimating the combined mediated effect using estimate bias, estimate *RMSE*, coverage probability, type1 error and power statistics, and the individual mediated effects using the type1 error and power statistics. The LASSO L_{min} regression method did not perform well for the 2-predictor individual mediated effect estimation for reasons that require further investigation for having high type1 errors and conditions with near zero power. The performance of LASSO L_{min} regression method improves for three predictor mediation models only for the combined indirect and direct effects but the individual effects continue to have high type1 errors and certain conditions with low power. The WQS regression method of using weights from $WQS_{index}^{M|X}$ to estimate the indirect effects and weights from $WQS_{index}^{Y|X,M}$ to estimate the direct effects did not perform well in the two and three predictor mediation models for type1 errors in conditions with a zero or near zero true effect size. Especially for estimating the indirect effect $ME_{12,\,123}^{M|X} = \hat{\theta}_{WQS}\hat{\gamma}_{WQS}$, the type1 error rates for null effects were high. The WQS regression method is a viable Variable Selection method (cut-off=0.3) when the effect size is not zero as shown in Table 4.20B.

Tables 4.20A & 4.20B

Table 4.20 Comparison of WQS, LASSO and OLS regressions for point and interval indirect effect estimates (combined & individual estimates for 2-predictor and 3-predictor mediation models)

	Indirect Effect											
	Two predictor mediation models						Three predictor mediation models					
	Combined indirect effect			Individual indirect effect			Combined indirect effect			Individual indirect effect		
Methods	$WQS^{M\|X}$	LASSO (Lmin)	OLS (REF.)	$WQS^{M\|X}$ 0.30 cut-off	LASSO (Lmin)	OLS (REF.)	$WQS^{M\|X}$	LASSO (Lmin)	OLS (REF.)	$WQS^{M\|X}$ 0.40 cut-off	LASSO (Lmin)	OLS (REF.)
Estimate Bias (point estimate)	>>OLS	OLS	OLS-0.005	Indeterminate	>>OLS	OLS-0.008	>>OLS	OLS	OLS-0.008	Indeterminate	>OLS	OLS-0.007
Estimate RMSE (point estimate)	>OLS	OLS	OLS-0.035	Indeterminate	<OLS	OLS - 0.06	>OLS	OLS	OLS - 0.04	Indeterminate	OLS	OLS - 0.055
Coverage Probability (interval estimate)	X Range (0.2, 1)	Range (0.90, 1)	Range (0.90, 1)	Indeterminate	Range (0.70, 1)	Range (0.92, 1)	X Range (0.1, 1)	Range (0.92, 0.97)	Range (0.92, 0.97)	Indeterminate	Range (0.80, 1)	Range (0.92, 1)
Type1 error (interval estimate)	X 7/60	X 8/60	OLS 0/60	X 2/84, 5/60	X 6/84, 2/60	OLS 0/84, 0/60	X 14/54	1/54	1/54	X 4/90 , 6/90, 4/90	LASSO 4/90 , 1/90, 1/90	OLS 0/90 , 1/90, 1/90
Power (interval estimate)	Range (0.24, 1)	X Range (0, 1)	Range (0.2, 1)	Range (0.2, 1)	X Range (0, 1)	X Range (0.05, 1)	Range (0.97, 1)	Range (0.96, 1)	Range (0.96, 1)	Range (0.55, 1)	Range (0.24, 1)	Range (0.2, 1)

Table 4.21 Comparison of WQS, LASSO and OLS regressions for point and interval direct effect estimates (combined & individual estimates for 2-predictor and 3-predictor mediation models)

	Direct Effect											
	Two predictor mediation models						Three predictor mediation models					
	Combined direct effect			Individual direct effect			Combined direct effect			Individual direct effect		
Methods	$WQS^{Y\|X,M}$	LASSO (Lmin)	OLS (REF.)	$WQS^{Y\|X,M}$ 0.30 cut-off	LASSO (Lmin)	OLS (REF.)	$WQS^{Y\|X,M}$	LASSO (Lmin)	OLS (REF.)	$WQS^{Y\|X,M}$ 0.30 cut-off	LASSO (Lmin)	OLS (REF.)
Estimate Bias (point estimate)	>>OLS	OLS	OLS-0.02	Indeterminate	>OLS	OLS-0.025	>>OLS	OLS	OLS-0.02	Indeterminate	>OLS	OLS-0.02
Estimate RMSE (point estimate)	>OLS	OLS	OLS - 0.09	Indeterminate	<OLS	OLS - 0.20	>OLS	OLS	OLS - 0.09	Indeterminate	<OLS	OLS - 0.14
Coverage Probability (interval estimate)	X Range (0, 1)	(>0.90)	(>0.92)	Indeterminate	Range (0.72, 0.97)	Range (0.92, 0.98)	X Range (0, 0.95)	(>0.95)	(>0.97)	Indeterminate	Range (0.72, 0.98)	Range (0.92, 0.98)
Type1 error (interval estimate)	No data	No data	No data	0/54	X 3/54	0/54	No data	No data	No data	0/54	X 11/54, 0/54, 1/54	OLS 1/54, 0/54, 0/54
Power (interval estimate)	Range (0.99, 1)	Range (0.4, 1)	X Range (0, 1)	Range (0.57, 1)	X Range (0.13, 1.0)	X Range (0.10, 1.0)	Range (0.99, 1)	Range (0.63, 1)	X Range (0.2,1)	Range (0.14, 1)	X Range (0.19, 1.0)	X Range (0.18, 1.0)

5 An Application of the WQS, LASSO, OLS Regression Methods to Mediation

The aim of the study application is to understand the relationship between polychlorinated biphenyls and liver disease such as hepatic cirrhosis and discover if this relationship is mediated by leukocyte telomere lengths (Scinicariello & Buser, 2015). Another interest is to see whether the three analysis methods (WQS, LASSO and OLS) would select the same predictor variables as being important and find similar individual, significant indirect and direct effects in the mediation model.

Telomeres are the ends of chromosomes whose length shortens as a normal consequence of animal aging. When telomeres become critically short, the cell nucleus signals cell non-proliferation resulting in cell death. Mutations to the telomeres result in a loss of genetic information have been linked to a spectrum of *telomere-mediated* diseases such as severe cirrhosis of the liver. Hepatic cirrhosis is attributed to an accelerated telomere attrition (Calado & Young, 2012). The National Health and Nutritional Examination Survey (NHANES) 2001-2002 dataset measured the mean chromosome’s telomere lengths and standard deviation for participating subjects. Liver enzymes such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyltransferase (GGT), and lactate dehydrogenase (LDH), were measured for the same subjects by NHANES, as general biomarkers used to assess the level of liver disease. Thirty-two lipid adjusted PCBs levels in serum were reported for the same subjects in the NHANES 2001-2002 databases.

The outcome variable was chosen as ALT from the set of liver enzyme biomarkers recorded for the participants in the NHANES 2001-2002 study: ALT (alanine aminotransferase), AST (aspartate aminotransferase), GGT (gamma-glutamyltransferase), and LDH (lactate dehydrogenase), since it had the strongest association with telomere length (mediator) in the mediation model. Giannini et al. (2005) suggested alanine aminotransferase to be the most specific enzyme in serum that is associated with liver disease. The other three biomarkers (AST, GGT and LDH) were insignificantly associated with the mediator in a simple regression and in full multiple regression models.

First the core set of covariates important to predict the mediator (telomere length) and the outcome were identified as age (continuous), gender (binary), race (4 categories), alcohol consumption (binary), smoking habit (binary), body-mass index (3 categories), and education level (binary). The selection of a sub-set of 12 PCBs from the 32 PCB predictors was done in the following manner. SAS9.4 PROC REG was used to fit a multiple regression for the continuous response variables mediator M and outcome Y. The core set of relevant covariates identified as C (the vector of common covariates for both mediation regression equations $M|X, C$ and $Y|X, M, C$) were consistently used in the two mediation regression equations to isolate the strongest four associations between the 32 PCB predictors (X) and the telomere lengths (M) or the outcome (Y). Two variables with the strongest, significant associations with the telomere length (M) were identified as PCB-099 and PCB-195 when $M|X, C$. Two different PCBs (PCB-149 and PCB-167) were identified as having the strongest, statistically significant associations with the outcome variable Y in the multiple regression $Y|X, M, C$. The four strongest associated predictors from $M|X, C$ and $Y|X, M, C$ were selected for the study

application so that the results from the simulation studies in Chapter 4 (2-predictors and 3-predictor mediation models) could be extended to the results from this application.

To build in multicollinearity, a set of four different predictors based on the high pairwise correlations (bivariate correlations of 0.8 and above) were added to the predictor set along with four different weakly associated predictors (bivariate correlations between 0.2 to 0.4). The objective was to allow for an assessment of each method's variable selection process based on the predictors selected as being significant to the mediation process and those predictors that were considered to be unimportant. Pairwise correlations of this sub-set of 12 PCBs are shown Table A1.1 in the Appendix 1, which constitutes the set of 12 correlated PCB predictors that were used in the study application.

The final model had ALT liver enzyme as the outcome *Y*, telomere length as the mediator *M*, seven covariates *C* (age, gender, race, BMI, alcohol consumption, smoking and education level) and the correlated cluster of twelve PCBs as predictors *X* (PCB-052, PCB-074, PCB-087, P-099, PCB-116, PCB-146, PCB-149, PCB-151, PCB-157, PCB-167, PCB-189, and PCB-195) . A complete case analysis (with no missing values) for these 21 variables resulted in a cross-sectional observation dataset of 640 subjects, whose laboratory measures and questionnaire responses were recorded in the 2001-2002 NHANES databases.

The summary statistics on the continuous outcome (log_ALT), mediator (telomere length) and the covariates are reported in the tables and figures below. A bivariate pairwise correlation for the 12 PCB congeners is shown in Table A1.1 in Appendix-1 and the hierarchical cluster diagram using (1- correlation) as intercluster distance is

shown in Figure A2.1 in Appendix-2. PCBs shown in the hierarchical cluster diagram that are closer to the x-axis indicate the high pairwise correlations of approximately 1.

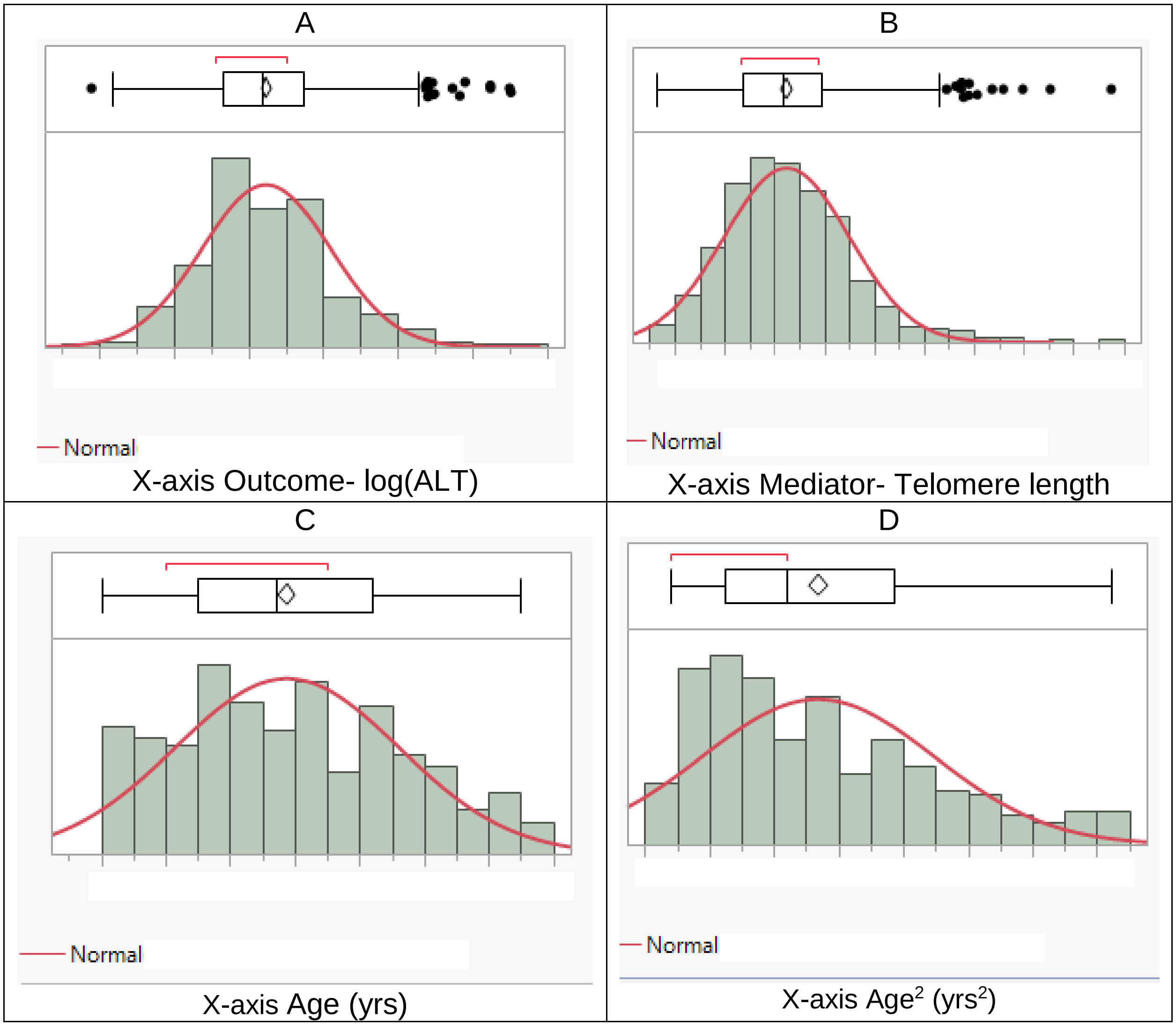


Figure 5.2A-D Histograms, Bar graphs with outliers and overlaid normal distributions for outcome (log_ALT), mediator (Telomere length), covariates (Age, Age^2)

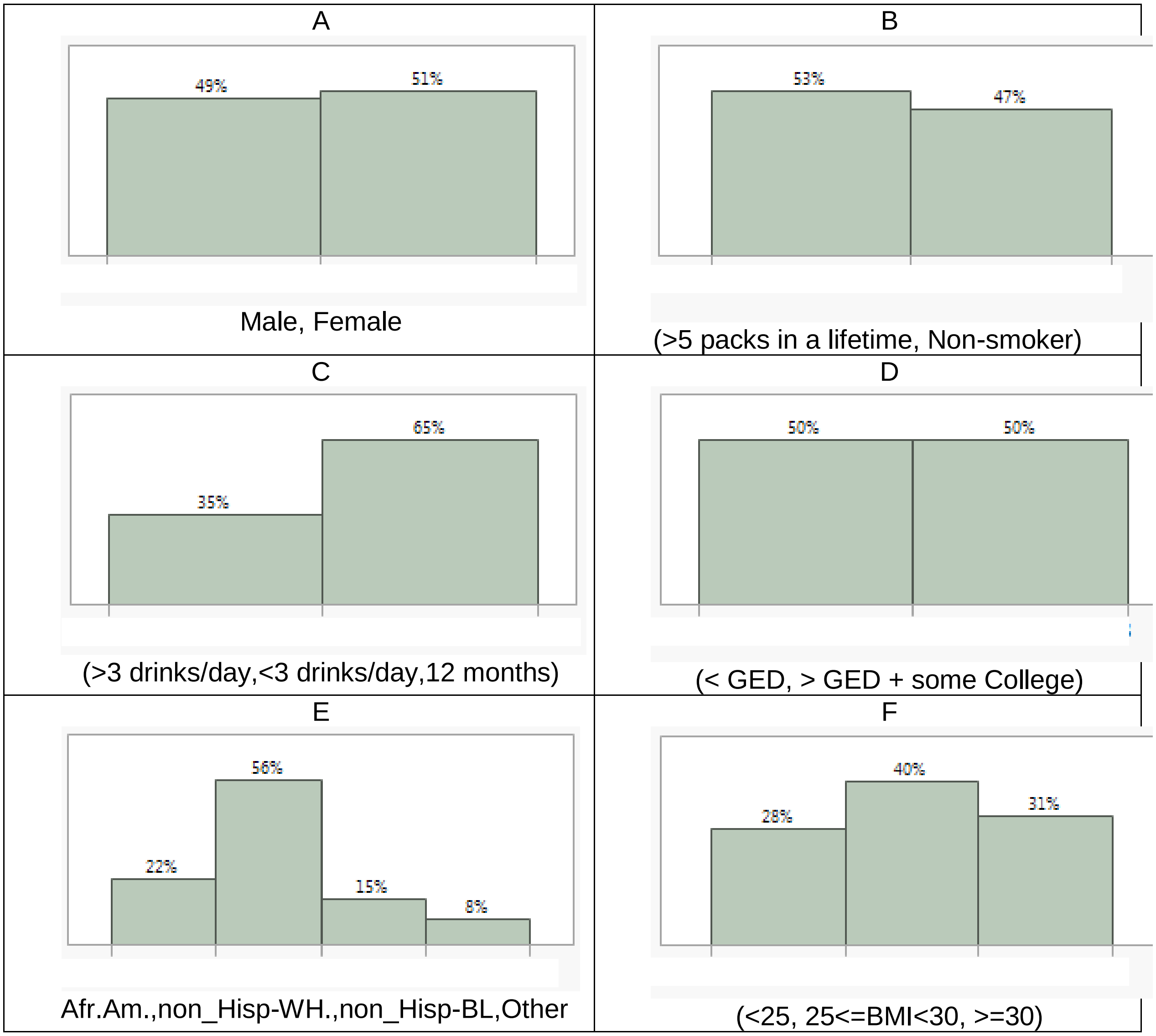


Figure 5.3A-F Frequency histograms for categorical covariates – Gender, Smoking, Alcohol consumption, Education, Race and Body Mass Index (BMI)

The seven covariates described above were consistently used in all regression models to adjust for relevant measured extraneous variables, which could partially explain variation in the outcome (ALT enzyme levels) e.g. the alcohol consumption covariate's effect on cirrhotic liver disease and the consequent elevated levels of the liver enzyme ALT.

As suggested in the scientific literature on telomere length, the scatter plot between mean telomere length and the age covariate shows a negative correlation and an attrition of the leukocyte telomere length due to aging (Hoffmann & Spyridopoulos 2011).

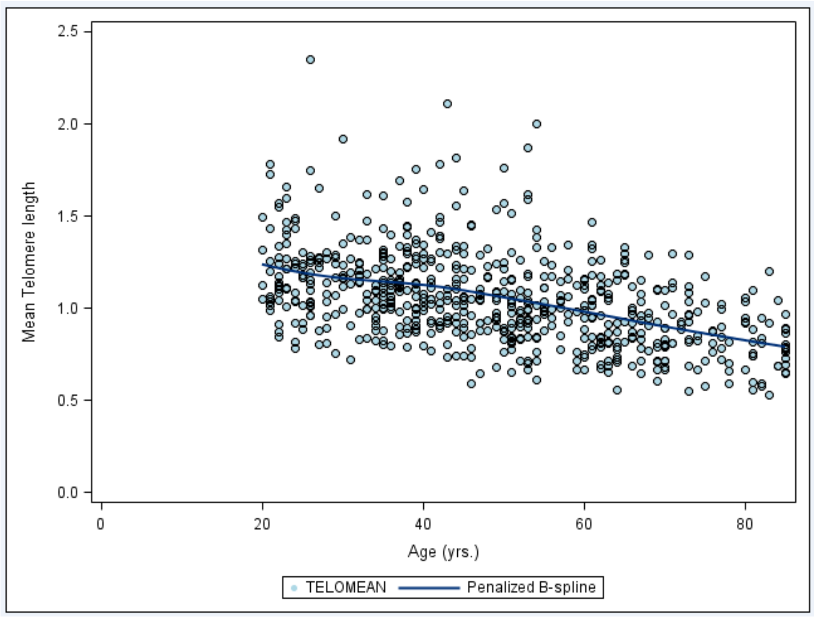


Figure 5.4 Scatter plot of the mediator (Mean Telomere Length) vs. covariate Age

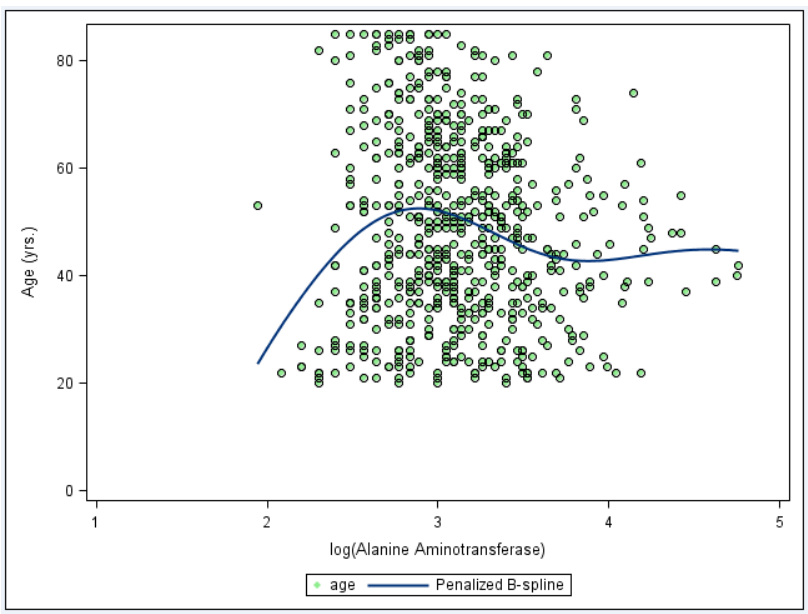


Figure 5.5 Scatter plot Age (yrs.) vs. Outcome variable log_ALT

Figures 5.4 and 5.5 suggest that a linear relationship exists between age and the mediator telomere length and a partial parabolic relationship exists between age and the outcome ALT.

OLS Regression Method Applied to the Mediation Model

Table 5.2

OLS Regression Method *M |X, C* Regression Results N=640

Variable	DF	Estimate	s.e.	t-statistic	p_value
PCB-052	1	0.0077	0.0039	1.99	0.0468
PCB-074	1	0.0025	0.0018	1.37	0.1717
PCB-087	1	-0.0445	0.0717	-0.62	0.5353
PCB-099	1	0.0004	0.0019	0.22	0.8235
PCB-118	1	-0.0012	0.0013	-0.91	0.3630
PCB-146	1	0.0012	0.0032	0.39	0.6941
PCB-149	1	0.0876	0.1945	0.45	0.6527
PCB-151	1	0.0837	0.1107	0.76	0.4500
PCB-157	1	-0.0208	0.0162	-1.28	0.2004
PCB-167	1	0.0133	0.0148	0.90	0.3697
PCB-189	1	-0.2531	0.1988	-1.27	0.2036
PCB-195	1	0.0384	0.0194	1.99	0.0475
Age	1	-0.0073	0.0007	-9.74	<.0001
Alcohol	1	0.0364	0.0207	1.76	0.0790
Smoking	1	0.0231	0.0189	1.22	0.2238
Male	1	-0.0325	0.0195	-1.67	0.0959
African American	1	-0.0333	0.0241	-1.38	0.1687
Non-Hispanic-Black	1	0.0213	0.0278	0.77	0.4442
Race-Other	1	-0.0060	0.0337	-0.18	0.8580
25<BMI<30	1	-0.0171	0.0224	-0.76	0.4462
BMI>=30	1	-0.0078	0.0238	-0.33	0.7434
Education	1	-0.0172	0.0191	-0.90	0.3672

PCB-052 and PCB-195 are the statistically significant predictors ($p \leq 0.05$) while the remaining 10 PCB are non-significant in their association with telomere length. PCB-052 (0.0077/ (ng/g) with a coefficient of error CE of 0.0039/0.0077=51%, and PCB-195 (0.0384/ (ng/g) also with CE=51%), were positively associated with mean telomere

length for a unit change in the PCBs, while aging was negatively associated with the mean telomere length 0.0073/yr. The multicollinearity in the design matrix ***X*** causes instability in the regression estimates resulting in high standard errors for the estimates. The outcome was regressed on the mediator, predictors, and covariates to get the results for *Y |M, X, C* to identify significant indirect and direct effects for predictors in the mediation model.

Table 5.3

OLS Regression Method *Y |M, X, C* Regression Results N=640

Variable	DF	Estimate	s.e.	t- statistic	p_value
Telomere	1	0.0751	0.0687	1.09	0.2748
PCB-052	1	0.0031	0.0066	0.47	0.6377
PCB-074	1	-0.0017	0.0032	-0.54	0.5925
PCB-087	1	-0.3798	0.1225	-3.10	0.0020
PCB-099	1	0.0008	0.0033	0.24	0.8088
PCB-118	1	0.0043	0.0022	1.99	0.0472
PCB-146	1	0.0056	0.0054	1.05	0.2948
PCB-149	1	1.3623	0.3320	4.10	<.0001
PCB-151	1	-0.2138	0.1890	-1.13	0.2583
PCB-157	1	0.0590	0.0278	2.13	0.0339
PCB-167	1	-0.0638	0.0254	-2.51	0.0122
PCB-189	1	-0.8226	0.3389	-2.43	0.0155
PCB-195	1	-0.0005	0.0331	-0.01	0.9892
Age	1	0.0214	0.0052	4.09	<.0001
Age^2	1	-0.0002	0.0001	-4.74	<.0001
Alcohol	1	0.0706	0.0354	2.00	0.0464
Smoking	1	-0.0404	0.0324	-1.25	0.2130
Male	1	0.3031	0.0333	9.11	<.0001
African American	1	0.0751	0.0413	1.82	0.0697
Non-Hispanic-Black	1	-0.0849	0.0477	-1.78	0.0753
Race-Other	1	-0.0106	0.0573	-0.19	0.8531
25<BMI<30	1	0.1032	0.0382	2.70	0.0070
BMI >=30	1	0.2129	0.0405	5.25	<.0001
Education	1	0.0106	0.0327	0.32	0.7463

The association between the mediator (telomere length) and the outcome (liver enzyme ALT) is not significant at $\alpha = 0.05$, indicating that the estimated gamma parameter in the regression model M |Y, X, C is zero $\hat{\gamma} = 0$. The mediated effect is determined by product of $\hat{\theta}_j \hat{\gamma}$; $j \in \{1..12\}$ for the 12 predictors. Bootstrap the data set 500 times and calculate the percentile confidence interval and the mediated effect for each of the predictors.

Table 5.4

OLS Individual Mediated Effect (2.5th, 97.5th) Percentile Confidence Interval

Predictor	**Indirect effect Estimate**	**Mediated effect Percentile CI**
PCB-052	0.0006	(-0.0008, 0.0023)
PCB-074	0.0002	(-0.0002, 0.0009)
PCB-087	-0.0033	(-0.1160, 0.2607)
PCB-099	0.0000	(-0.0003, 0.0005)
PCB-118	-0.0001	(-0.0004, 0.0002)
PCB-146	0.0001	(-0.0006, 0.0009)
PCB-149	0.0066	(-0.6906, 0.4574)
PCB-151	0.0063	(-0.2718, 0.1940)
PCB-157	-0.0016	(-0.0076, 0.0021)
PCB-167	0.0010	(-0.0019, 0.0057)
PCB-189	-0.0190	(-0.2703, 0.9833)
PCB-195	0.0029	(-0.0078, 0.0304)

The OLS method does not find telomere length as a mediator between any of the twelve predictors and the outcome. However, it is important to note that PCB-149 has a strong positive association with elevated levels of the outcome ALT as shown in Table 5.3. PCB-087 has a statistically significant positive association with ALT. PCB-118 and PCB-157 have a significant positive effect on the mean ALT value, while PCB-167 and PCB-189 each have a significant (fractional) positive effect on the mean ALT value, for a unit increase in the lipid adjusted PCB (ng/g) concentration, all shown in Table 5.3.

Analysis of the covariates indicate that males have a mean effect of (1.354/ (ng/g), and obese (BMI>30) individuals have an effect of (1.237/ (ng/g), ref. BMI <25). It is noteworthy that the sub-set of PCBs selected by OLS regression as having a significant effect on the mediator telomere length (PCB-195 and PCB-052), is a different sub-set of PCBs selected by OLS regression as having a significant effect on the outcome ALT (PCB-087, PCB-118, PCB-149, PCB-157, PCB-167 and PCB-189), and the mediator has a non-significant effect on the outcome ALT i.e. $\gamma = 0$. Table 5.4 shows that none of the predictors have an indirect effect through telomere length on the liver enzyme ALT.

LASSO Regression Methods Applied to the Mediation Model

The LASSO method is used to analyze the proposed mediation model to study the variables selected as being significant in their association with the mediator and the outcome, and to estimate the mediated effects if they exist. The LASSO regression method admitted PCB-52 and PCB-74 out of the 12 predictors adjusted for the seven covariates on the mediator telomere length in M | X, C when using n-fold cross-validation to identify the minimum *MSE*. The method admitted predictors PCB-99 and PCB-189 with a gamma coefficient (0.062) as the Lasso regression of the mediator (telomere length), 12 predictors, adjusted for the 7 covariates on the outcome ALT in *Y*|*M, X, C* using n-fold cross-validation to identify coefficients associated with the minimum *MSE*. These results are tabulated Table 5.5 shown below. Referring to the correlation hierarchical cluster diagram in the Appendix-2, note that LASSO admits one predictor from the left correlated cluster (PCB-052) and one predictor from the right correlated cluster (PCB-074) for *M*|*X,C,* and admits one predictor from the left correlated

cluster (PCB-189) and one predictor from the right correlated cluster (PCB-099) for *Y|M, X, C*; thus reducing the dimensionality from 12 to 2 in the two multiple regressions of the mediation model. This reduction in the number of predictors in the final model reduces multicollinearity in the design matrix ***X***.

Table 5.5

Lasso Regression Admitted Predictors' Regression Coefficients for L_{min}

Mediator/ Predictor	Regression Coefficient	Theta MSE (min)	Regression Coefficient	Beta MSE (min)
Telomere			0.0622	L_{min}_M
PCB-52	0.0054	L_{min}_T52	0	L_{min}_B52
PCB-74	0.0006	L_{min}_T74	0	L_{min}_B74
PCB-87	0	L_{min}_T87	0	L_{min}_B87
PCB-99	0	L_{min}_T99	0.0045	L_{min}_B99
PCB-118	0	L_{min}_T118	0	L_{min}_B118
PCB-146	0	L_{min}_T146	0	L_{min}_B146
PCB-149	0	L_{min}_T149	0	L_{min}_B149
PCB-151	0	L_{min}_T151	0	L_{min}_B151
PCB-157	0	L_{min}_T157	0	L_{min}_B157
PCB-167	0	L_{min}_T167	0	L_{min}_B167
PCB-189	0	L_{min}_T189	-0.0393	L_{min}_B189
PCB-195	0	L_{min}_T195	0	L_{min}_B195

When the coefficients were obtained using $MSE_{min.+\ 1s.e.}$ in n-fold cross-validation, PCB-52 was admitted in *M | X, C* and PCB-99 in *Y |M, X, C* with a non-zero gamma coefficient (0.0692) for the Lasso regression of the mediator (telomere length), 12 predictors, adjusted for the 7 covariates on the outcome ALT in *Y|M, X, C*. Referring to the correlation hierarchical cluster diagram in the Appendix-2, note that LASSO admits one predictor from the left correlated cluster (PCB-052) and no predictor from the right correlated cluster for *M|X,C* while admitting no predictor from the left correlated cluster and one predictor from the right correlated cluster (PCB-099) for *Y|M, X, C*; thus reducing the dimensionality from 12 to 1 in the two multiple regressions of the mediation

model. This reduction in the number of predictors in the final model reduces multicollinearity in the design matrix ***X***.

Table 5.6

Lasso Regression Admitted Predictor's Regression Coefficients for $L_{min+1se}$

Mediator/ Predictor	Regression Coefficient	Theta $MSE_{min+1se}$	Regression Coefficient	Beta $MSE_{min+1se}$
Telomere			0.0692	$L_{min+1se}$_M
PCB-52	0.0000	$L_{min+1se}$_T52	0	$L_{min+1se}$_B52
PCB-74	0	$L_{min+1se}$_T74	0	$L_{min+1se}$_B74
PCB-87	0	$L_{min+1se}$_T87	0	$L_{min+1se}$_B87
PCB-99	0	$L_{min+1se}$_T99	0.0000	$L_{min+1se}$_B99
PCB-118	0	$L_{min+1se}$_T118	0	$L_{min+1se}$_B118
PCB-146	0	$L_{min+1se}$_T146	0	$L_{min+1se}$_B146
PCB-149	0	$L_{min+1se}$_T149	0	$L_{min+1se}$_B149
PCB-151	0	$L_{min+1se}$_T151	0	$L_{min+1se}$_B151
PCB-157	0	$L_{min+1se}$_T157	0	$L_{min+1se}$_B157
PCB-167	0	$L_{min+1se}$_T167	0	$L_{min+1se}$_B167
PCB-189	0	$L_{min+1se}$_T189	0	$L_{min+1se}$_B189
PCB-195	0	$L_{min+1se}$_T195	0	$L_{min+1se}$_B195

The $L_{min+1se}$ method is more conservative since the Lasso regression coefficients are further shrunk towards zero as compared to the L_{min} method (Table 5.5 vs. Table 5.6). Post processing the admitted predictors from LASSO regressions shown in Table 5.5 consists of regressing PCB-052, PCB-074,PCB-99,PCB-189 on the outcome variable log(ALT) and the Mediator (Telomere length) using the OLS method.

Table 5.7

$\text{LASSO}_{\text{min}}$ *M* |*X, C* Regression Results N=640

Variable	DF	Estimate	s.e.	t_stat	p_value
PCB-052	1	0.0092	0.0038	2.42	0.0159
PCB-074	1	0.0016	0.0015	1.10	0.2738
PCB-099	1	-0.0006	0.0015	-0.39	0.6941
PCB-189	1	-0.0202	0.0187	-1.08	0.2808
Age	1	-0.0072	0.0007	-10.13	<.0001
Alcohol	1	0.0377	0.0206	1.83	0.0683
Smoking	1	0.0246	0.0188	1.31	0.1911
Male	1	-0.0367	0.0191	-1.93	0.0546
African American	1	-0.0384	0.0240	-1.60	0.1098
Non-Hispanic-Black	1	0.0274	0.0269	1.02	0.3084
Race-Other	1	-0.0038	0.0335	-0.11	0.9094
25<BMI<30	1	-0.0181	0.0221	-0.82	0.4132
BMI>=30	1	-0.0101	0.0232	-0.44	0.6621
Education	1	-0.0535	0.0189	-0.81	0.4170

PCB-052 is significantly associated with the mediator telomere length $\hat{\theta}_{52} = 0.009$ and the age covariate reduces the mean telomere length as age increases (-0.0072). A LASSO regression of *Y* |*M, X, C* produces the following results tabulated below.

Table 5.8

LASSO$_{min}$ *Y* |*M, X, C* Regression Results N=640

Variable	DF	Estimate	s.e.	t_stat	p_value
Telomere length	1	0.0742	0.0691	1.07	0.2833
PCB-052	1	0.0013	0.0066	0.20	0.8399
PCB-074	1	-0.0003	0.0026	-0.13	0.9002
PCB-099	1	0.0070	0.0026	2.66	0.0080
PCB-189	1	-0.0586	0.0323	-1.81	0.0070
Age	1	0.0239	0.0052	4.60	<.0001
Age^2	1	-0.0003	0.0001	-5.23	<.0001
Alcohol	1	0.0716	0.0357	2.01	0.0450
Smoking	1	-0.0400	0.0325	-1.23	0.2177
Male	1	0.2997	0.0329	9.11	<.0001
African American	1	0.0698	0.0415	1.68	0.0930
Non-Hispanic-Black	1	-0.0875	0.0466	-1.88	0.0609
Race-Other	1	-0.0109	0.0576	-0.19	0.8501
25<BMI<30	1	0.0835	0.0380	2.20	0.0284
BMI>=30	1	0.1866	0.0400	4.72	<.0001
Education	1	0.0171	0.0328	0.52	0.6034

PCB-099 and PCB-189 were found to be significantly associated with the outcome log(ALT) $\hat{\beta}_{99} = 1.01, \hat{\beta}_{189} = 0.943$ and the obese, older female raises the mean ALT liver enzyme measured by (0.74 + 1.21 =1.95 in ALT units) as compared to their male counterparts with BMI <25. The LASSO method mediation analysis also shows that *M* |*Y, X, C* yields $\hat{\gamma} = 0$, the regression coefficient for telomere length in *Y*|*M, X, C* shown in Table 5.8. No predictors had a significant mediated effect (Table 5.9) using the LASSO$_{min}$ method.

Table 5.9

$LASSO_{min}$ Individual Mediated Effect (2.5th, 97.5th) Percentile Confidence Interval

Predictor	Indirect effect Estimate	Mediated effect Percentile CI (Lmin)
PCB-052	0.0007	(-0.0005, 0.0019)
PCB-074	0.0001	(-0.0001, 0.0004)

The same post processing of $LASSO_{min+1s.e.}$ selected variables with non-zero regression coefficients at n-fold cross-validation prediction $error_{min+1.s.e.}$ is completed using the OLS regression method for *M*|*X, C* and *Y*|*M, X, C* with results in Table 5.10.

Table 5.10

$LASSO_{min+1s.e.}$ *M* |*X, C* Regression Results N=640

Variable	**DF**	**Estimate**	**s.e.**	**t_stat**	**p_value**
PCB-052	1	0.0073	0.0030	2.43	0.0152
PCB-099	1	0.0006	0.0010	0.57	0.5678
Age	1	-0.0067	0.0006	-11.02	<.0001
Alcohol	1	0.0417	0.0205	2.04	0.0418
Smoking	1	0.0245	0.0186	1.31	0.1901
Male	1	-0.0418	0.0185	-2.25	0.0245
African American	1	-0.0403	0.0239	-1.69	0.0921
Non-Hispanic-Black	1	0.0227	0.0267	0.85	0.3961
Race-Other	1	-0.0048	0.0335	-0.14	0.8862
25<BMI<30	1	-0.0156	0.0218	-0.72	0.4743
BMI>=30	1	-0.0066	0.0230	-0.29	0.7751
Education	1	-0.0175	0.0188	-0.93	0.3522

PCB-52 was found to be significantly associated with the mediator (telomere length) with aging, alcohol drinking males having reduced mean telomere lengths indicating increased cell deaths.

Table 5.11

LASSO$_{min+1s.e.}$ *Y |M, X, C* Regression Results N=640

Variable	DF	Estimate	s.e.	t_stat	p_value
Telomere length	1	0.0788	0.0690	1.14	0.2535
PCB-052	1	-0.0060	0.0052	-1.15	0.2498
PCB-099	1	0.0065	0.0017	3.77	0.0002
Age	1	0.0245	0.0052	4.72	<.0001
Age^2	1	-0.0003	0.0001	-5.42	<.0001
Alcohol	1	0.0752	0.0355	2.12	0.0342
Smoking	1	-0.0354	0.0323	-1.09	0.2742
Male	1	0.2998	0.0321	9.34	<.0001
African American	1	0.0706	0.0415	1.70	0.0892
Non-Hispanic-Black	1	-0.0939	0.0464	-2.02	0.0434
Race-Other	1	-0.0142	0.0577	-0.25	0.8061
25<BMI<30	1	0.0933	0.0377	2.48	0.0135
BMI>=30	1	0.1981	0.0397	4.99	<.0001
Education	1	0.0167	0.0327	0.51	0.6103

PCB-099 was found to be significantly associated with the outcome log(ALT) $\hat{\beta}_{99} = 1.007$ and the obese, alcohol consuming, aging male raises the mean ALT liver enzyme measured by (1.82 in ALT units) as compared to their younger female counterparts with BMI <=25. This LASSO method also shows that *M |Y, X, C* yields $\hat{\gamma} = 0$, the regression coefficient for telomere length in *Y|M, X, C* shown in Table 5.11. No indirect effect was identified when PCB levels in the serum act through telomere lengths to determine the mean liver enzyme ALT measured in the blood of the test subjects, using the LASSO$_{min+1s.e.}$ method, which is summarized in Table 5.12.

Table 5.12

$LASSO_{min+1s.e.}$ Individual mediated effect (2.5th, 97.5th) Percentile Confidence Interval

Predictor	Indirect effect Estimate	Mediated effect Percentile CI (Lmin+1s.e.)
PCB-052	0.0006	(-0.0007, 0.0020)

WQS Regression Methods Applied to the Mediation Model

The WQS methods $WQS_{index}^{M|X}, WQS_{index}^{Y|X,M}, WQS_{index}^{Y|X}$ are used to conduct mediation analysis using 12 PCBs as the predictors, mean leukocyte telomere lengths as the mediator and the log (ALT) lever enzyme levels as the outcome variable.

Table 5.13

WQS Results for 12 PCB Predictors Mediated via Telomere Length on ALT

WQS weights from	**WQS Theta**	**WQS Gamma**	**Mediated Effect (ME)**	**2.5th pcntl. CI**	**97.5th pcntl. CI**	**Direct Effect (DE)**	**DE s.e.**	**DE Lower CI**	**DE Upper CI**
M|X,C	0.043	0.071	0.003	-0.003	0.010	0.037	0.024	-0.010	0.084
Y|M,X,C	0.016	0.090	0.001	-0.004	0.009	-0.021	0.018	-0.056	0.014
Y|X,C	0.017	0.089	0.002	-0.004	0.009	-0.019	0.018	-0.054	0.016

Simulation results from Chapter 4 indicate that the WQS method results for estimating the joint mediated effects is best represented by using the weights from *M|X, C*, mean indirect effect =0.003 units of ALT with CI (-0.003, 0.010) indicating that no indirect effect was found when representing the predictors in the model with $WQS_{index}^{M|X}$. The aggregate direct effect is best estimated by using the weights from *Y|M, X, C* of mean joint direct effect =-0.021 in log_ALT units with CI (-0.056, 0.014) and a standard error for the estimate of 0.018 log_ALT units or *Y|X, C* of mean joint direct effect =

-0.019 in log-ALT units with CI (-0.054, 0.016) with a standard error for the estimate of 0.018 log_ALT units as shown in Table 5.13. The predictors contributing to these joint indirect and joint direct effects can be determined from the WQS weights for each of the 12 predictors and by using an *a priori* chosen cut-off value to representing zero predictor weights in the WQS regression model e.g. if the cut-off value is determined using the heuristic $\frac{1}{1+p} = 0.077$.

Comparison of WQS, LASSO and OLS Regression Methods Applied to the Mediation Model

Table 5.14

WQS Weights for Three Different Indices to Represent the Predictor Set $\text{cut-off} = 0.08$

WQS weights from	PCB 052	PCB 087	PCB 149	PCB 195	PCB 151	PCB 189	PCB 157	PCB 167	PCB 074	PCB 099	PCB 118	PCB 146
M|X,C	0.042	0.003	0.003	0.061	0.006	0.003	0.001	**0.173**	**0.349**	**0.108**	**0.209**	0.042
Y|M,X,C	**0.122**	0.063	0.063	0.064	0.064	**0.320**	0.066	0.039	0.040	0.038	**0.099**	0.022
Y|X,C	**0.114**	0.057	0.058	0.063	0.066	**0.318**	0.073	0.031	0.044	0.032	**0.122**	0.022

The WQS weights below the cut-off value of 0.08 identify the predictors which may be considered to have an unimportant effect on the combined effect. The $WQS_{index}^{M|X}$ method using *M|X, C* weights to detect the indirect and direct effects selects PCB-074, PCB-099, PCB-118, and PCB-167 (accounting for 84% of the WQS weights) as important contributors to the combined mediated effect. The OLS method shown in Table 5.15 found four additional predictors PCB-087 PCB-149, PCB-195 and PCB-157 which were not selected by either WQS or LASSO methods. There were two predictors PCB-146 and PCB-151 that were not found to be important by any of the six regression

methods. LASSO and OLS regression methods identified PCB-052 (singleton in the hierarchical cluster) which was also identified by the WQS regression methods $WQS_{index}^{Y|X,M}$ and $WQS_{index}^{Y|X}$ using weights from *Y|X, M, C* and *Y|X, C* which also included PCB-118 and PCB-189 (Table 5.14). OLS does not find PCB-074 or PCB-099 to be important contributors to the overall mediated effects. The OLS method highlights eight of the twelve predictors as being significant predictors at $p \leq 0.05$ in the multiple regressions for mediation analysis. The LASSO methods select four of the twelve predictors PCB-052, PCB-074, PCB-099 and PCB-189 but the $L_{min+1s.e}$ method selects only PCB-052 and PCB-099 while shrinking the remaining ten predictor's coefficients to zero. The OLS regression method has an unconstrained minimization of squared residuals as its optimization function, unlike WQS and LASSO methods, which deliver a parsimonious model (4), and WQS (3 to 4) of 12 predictors, as being influential in this mediation analysis. All the methods failed to include PCB-146 and PCB-151 as shown in Table 5.15.

Table 5.15

Regression Methods and Selection of Influencing Predictors in Mediation (0.08 cut-off)

Methods	**PCB 52**	**PCB 87**	**PCB 149**	**PCB 195**	**PCB 151**	**PCB 189**	**PCB 157**	**PCB 167**	**PCB 74**	**PCB 99**	**PCB 118**	**PCB 146**
OLS	✓	✓	✓	✓	✗	✓	✓	✓	✗	✗	✓	✗
L_{min}	✓	✗	✗	✗	✗	✓	✗	✗	✓	✓	✗	✗
$L_{min+1s.e.}$	✓	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗
$WQS_{M\|X,C}$	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✗
$WQS_{Y\|M,X,C}$	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗
$WQS_{Y\|X,C}$	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✓	✗
Est. VIF	2	40	208	20	61	196	6	11	5	5	11	11

The hierarchial clustering of the PCB predictors based on a distance defined as (1 – pairwise correlations) is shown in Appendix-2 Figure A2.1 as having a singleton

(PCB-052) and two major clusters of PCB congeners. Analyzing the outomes of the three regression methods (WQS, LASSO and OLS) using this hiearchical clustering diagram sheds light on those individual predictors within the correlated clusters that were determined to be influential in the mediation model by each method.

PCB-052 (VIF=2) is the singleton which was selected by the OLS, both LASSO methods, $WQS^{Y|X}$, and $WQS^{Y|X,M}$ methods as being significantly associated with the outcome. The second cluster of 5 PCB congeners (PCB-087, PCB-149, PCB-195, PCB-151, and PCB-189 having pairwise correlations higher than 0.9) shows that OLS selected 4 of the 5 by excluding PCB-151 which was rejected by all six regression methods, LASSO selected only PCB-189 while WQS exhibits a *grouping effect* for this correlated cluster and reject all 5 in the cluster. The third cluster of six PCB congeners consists of PCB-157 (VIF=6.5) & PCB-167 (VIF=11), PCB-074 (VIF=5) & PCB-099 (VIF=5), and PCB-118 (VIF=11) & PCB146 (VIF=11). The OLS regression method selected both PCB-157 (VIF=6.5) & PCB-167 (VIF=11) but LASSO rejected both as part of the regularization process and admits PCB-074 (VIF=5) and PCB-099 (VIF=5) having reduced the effects of multicollinearity. The OLS method fails to find significance for the effects from PCB-074 (VIF=5) or PCB-099 (VIF=5) due to multicollinearity. At least one of the WQS_{index} methods selected four of the six PCBs as being influenctial in the mediation model from the third cluster. Could this be another grouping effect of admitting the similarly correlated cluster of six (right cluster) while rejecting the second similarly correlated cluster of five PCBs (left cluster) as shown in the hiearchical cluster diagram in Appendix-2? The other pair within this cluster (is PCB-074 and PCB-099 with VIF=5), and $WQS_{index}^{M|X,C}$ and $LASSO_{min}$ selected both, while OLS rejected both PCB

congeners possibly due to the effects of multicollinearity. The final pair within the cluster is (PCB-118 & PCB-146 with VIF=11) and LASSO rejected both, while OLS selected PCB-118 but failed to find PCB-146 as being significant. PCB-146 was not highlighted by any of the six regression methods as being important to mediaiton analysis.

Individual predictor indirect effects for OLS, LASSO and WQS Regressions Applied to the Mediation Model.

The OLS indirect effects were estimated with their 2.5 percentile and 97.5 percentile CI (all 12 PCB congeners) shown in Table 5.4. None were found to have a significant indirect effect with telomere length as the mediator. The LASSO methods had parsimonious solution of PCB-052 for both L_{min} and $L_{min+1s.e.}$ and PCB-074 for L_{min} as shown in Table 5.9 for L_{min} and Table 5.12 for $L_{min+1s.e.}$. There were similar estimates from OLS and LASSO regression methods for PCB-052 and PCB-074. The total indirect effects for the three WQS methods are shown in Table 5.13 but there were no joint indirect effects that were significant in any of the three WQS methods with Telomere length as the mediator. The influencing predictors are (6 of 12) shown in Table 5.14 as being PCB-052, PCB-189, PCB-167, PCB-074, PCB-099, and PCB-118 for *M*|*X,C, Y*|*X,C,* and *Y*|*M,X,C*. The $WQS_{index}^{M|X}$, $WQS_{index}^{Y|X}$, and $WQS_{index}^{Y|M,X}$ methods did not identify any different predictors which were not already found to be important by either OLS or the two LASSO regression methods. This demonstrates one application of the six alternative methods presented in the simulation studies: $WQS_{M|X, C}$, $WQS_{Y|X, C}$, $WQS_{Y|M, X, C}$, $LASSO_{min}$, $LASSO_{min+1se}$, and OLS regression methods for analyzing mediation models.

6 Thesis Summary, Limitations and Future Direction

In recent decades, there has been a growing interest to go beyond association studies to understand the mechanism by which independent variables produce their effects on the outcome. Such knowledge would allow for a systematic control of the outcomes based on the findings of a well-framed research question in a methodological study. In a meta-analysis a summary of the tests for significant mediated effects in different datasets from the same target population would provide evidence-based knowledge of the mechanism of action, which could eventually lead to interventions in clinical practice allowing for additional control on the outcome.

The comparative performance between OLS, LASSO and WQS regression methods in mediation analysis is detailed and summarized pictorially in Appendix -4. With OLS regression, the relative importance of a predictor can to be inferred from the strength of the regression coefficient which is found to be significant at $p \leq 0.05$.

The LASSO regression method addresses the issues with multicollinearity by shrinking the regression coefficients towards zero, and admitting only the predictors with the strongest associations with the outcome into the final solution set. Simulation studies suggested that a post processing of the reduced predictor set using the OLS method yields results similar to the OLS method, for any LASSO admitted independent variables. The LASSO admitted predictors have the unique advantage of having coefficient estimates with increased precision benefiting from reduced multicollinearity but are biased due to the excluded variables.

OLS does not reduce the number of independent variables from a set of correlated predictors to an important sub-set and the multicollinearity could have a destabilizing effect on the individual OLS indirect effect estimates. The WQS method performs differently from the OLS and LASSO regression methods, by weighting individual predictors based on the strength of their association with the outcome, while reducing multicollinearity through a regularization technique of constrained optimization. In the WQS method, the choice of the cut-off value for the WQS weights determines the predictors that enter the subset of predictors that contribute to a significant overall mediated effect. The illustrative example in Chapter 5 shows the WQS predictor-set to be more inclusive (6 PCBs out of 12 predictors in the original set for a cut-off value $\frac{1}{1+(p=12)}$ = 0.077) than the LASSO solution set (2 PCBs out of 12), while OLS found 8 of the 12 PCBs to be statistically significant in their associations with the output but none having a significant indirect effect. A higher Type 1 error rate and poor coverage probability were reported in our simulation studies when using the WQS regression methods for those few conditions with small sample sized datasets, having high multicollinearity, and a parameter set resulting in a zero indirect effect. This was found to be due to the WQS method's bias placed on a zero gamma parameter estimate in conjunction with non-zero theta parameters causing a non-zero indirect effect estimate for those conditions.

A limitation of the study was that it was limited to two (7 parameters with 108 conditions) and three correlated predictors (9 parameters with 162 conditions) within the mediation model. The illustrative example showed that including conditions for a higher numbers of correlated predictors e.g. five with 22-parameters, set to different values in

the simulation design, would have allowed for a pattern to be identified for the performance of the six methods, that could guide in selecting an optimal method from WQS, LASSO, and OLS regression methods. The three sets of results (2-vars, 3-vars and higher-vars) consisting of joint and individual mediated estimates measured by estimated effect values, their standard errors, mean squared error estimates, coverage probability, type1 error rate, and power could establish a pattern for the theoretical results. The two and three independent variable simulation studies did not include conditions where the joint direct effect was zero, to enable the reporting of a type1 error rate for direct effects. The WQS method in this thesis stops at identifying the sub-set of predictors which influences the significant aggregate indirect and direct effects.

In this dissertation, the outcome variables in the regressions M |X and Y |M, X, were sampled from normal Gaussian distributions (M, Y), C representing the vector of covariates. This limitation prohibits the results from being generalized to binary outcomes (diseased / healthy) and ordered categorical outcomes, which are more common in clinical medicine. An extension of the WQS mediation analysis method to process a binary outcome would expand its application to more medical research.

References

Baron, R.M., & Kenny, D.A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Belsley, D. A. (2006). Conditioning Diagnostics. *Encyclopedia of Statistical Sciences*. Retrieved February 3, 2017, from http://onlinelibrary.wiley.com/doi/10.1002/0471667196.ess0275.pub2/full

Bollen, K.A., & Stine, R. (1990). Direct and Indirect Effects: Classical and Bootstrap Estimates of Variability. *Sociological Methodology, 20*, 115–140.

Byrd, R.H., Gilbert, J.C., & Nocedal, J. (2000). A Trust Region based on Interior Point Techniques for Non-linear Programming. *Mathematical Programming, Series B, 89(1)*, 149-185.

Carrico, C. (2013) Characterization of a Weighted Quantile Score Approach for Highly Correlated Data in Risk Analysis Scenarios. PhD dissertation: Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia.

Christensen, K.L.Y., Carrico, C.K., Sanyal, A.J., & Gennings, C. (2013). Multiple Classes of Environmental Chemicals are Associated with Liver Disease: NHANES 2003-2004. *International Journal of Hygiene and Environmental Health, 216(6*), 703-709.

Czarnota, J., Gennings, C, Colt, J.S., De Roos, A.J., Cerhan, J.R., Severson, R.K., … Wheeler, D.J. (2016). Analysis of Environmental Chemical Mixtures and Non-Hodgkin Lymphoma Risk in the NCI-SEER NHL Study. *Environmental Health Perspectives, 123(10)*, 965-970.

Duzan, H., & Shariff, N.S.B.M. (2015). Ridge Regression for Solving the Multicollinearity Problem: Review of Methods and Models. *Journal of Applied Sciences, 15(3),* 392-404.

Efron, B & Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. New York, NY: Chapman & Hall.

Farrar, D.E. & Glauber, R.R. (1967). Multicollinearity in Regression Analysis: The Problem Revisited. *Review of Economics and Statistics, 49(1)*, 92-107.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent*. Journal of Statistical Software, 33(1),* 1-22.

Gennings, C., Sabo, R., & Carney, E. (2010). Identifying Subsets of Complex Mixtures most Associated with Complex Diseases: Polychlorinated Biphenyls and Endometriosis as a Case Study. *Epidemiology, 21 (Suppl. 4)*, S77–S84.

Graham, M.H. (2003). Confronting Multicollinearity in Ecological Multiple Regression. *Ecology, 84(11),* 2809-2815.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer Series in Statistics.

Hoerl, A.E., & Kennard, R.W. (1970). Ridge Regression: Applications to Non-orthogonal Problems. *Technometrics*, *12*, 69-82.

Holland, P.W., & Welsch, R.E. (1977). Regression Using Iteratively Reweighted Least-squares. *Communications in Statistics Theory and Methods, 6(9)*, 813-827.

Land, K.C. (1969). Principles of Path Analysis. *Sociological Methodology.1*, 3-37.

MacKinnon, D.P., Fairchild, A.J., & Fritz, M.S. (2007). Mediation Analysis. *Annual Review of Psychology, 58*, 593-614.

MacKinnon, D.P., Lockwood, C.M., & Williams, J. (2004). Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. *Multivariate Behavioral Research, 39(1)*, 99-128.

MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V.A. (2002). Comparison of Methods to Test the Significance of Mediation and other Intervening Variable Effects. *Psychological Methods, 7*, 83–104.

Mallinckrodt, B., Abraham, W.T., Wei, M., & Russell D.W. (2006). Advances in Testing the Statistical Significance of Mediating Effects. *Journal of Counseling Psychology, 53(3)*, 372-378.

Mansfield, E.R., & Helms, B.P. (1982). Detecting Multicollinearity. *The American Statistician*, *36(3a)*, 158-160.

Meinshausen, N. & Buhlmann, P. (2010). Stability Selection. *Journal of the Royal Statistical Society, 72*, 417-473.

Oishi, S., Tanabe, K., Ogita, T., & Rump, S.M. (2007). Convergence of Rump's Method for Inverting Arbitrarily Ill-conditioned Matrices. *Journal of Computational and Applied Mathematics, 205(1),* 533–544.

R: (Version 3.2.0) [Computer software]. The R Foundation for Statistical Computing.

SAS (Version 9.4 TS Level 1M2) [Computer software]. Cary, NC: SAS Institute Inc.

Sewall, W. (1934). The Method of Path Coefficients. *The Annals of Mathematical Statistics, 5(3),* 161-215.

Shrout, P.E., & Bolger, N. (2002). Mediation in Experimental and Non-experimental Studies: New Procedures and Recommendations. *Psychological methods, 7*, 422-445.

Sobel, M.E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology, 13*, 290–312.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, *58(1),* 267-288.

Zou, H., & Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society, B-67(2),* 301-320.

Appendix 1

Table A1.1 `

Predictor pairwise correlations for 12 PCB congeners and VIFs for predictor's coefficient estimate variance

	PCB52	PCB74	PCB87	PCB99	PCB118	PCB146	PCB149	PCB151	PCB157	PCB167	PCB189	PCB195
PCB52	1	0.248	0.589	0.238	0.163	0.212	0.632	0.627	0.395	0.330	0.639	0.628
PCB74	0.248	1	0.277	0.821	0.765	0.776	0.193	0.143	0.633	0.670	0.167	0.161
PCB87	0.589	0.277	1	0.359	0.316	0.420	**0.962**	**0.892**	0.669	0.602	**0.929**	**0.892**
PCB99	0.238	0.821	0.359	1	0.782	0.798	0.247	0.184	0.658	0.687	0.193	0.183
PCB118	0.163	0.765	0.316	0.782	1	0.826	0.193	0.195	0.642	0.834	0.152	0.145
PCB146	0.212	0.776	0.42	0.798	0.826	1	0.290	0.244	0.773	0.844	0.249	0.235
PCB149	0.632	0.193	**0.962**	0.247	0.193	0.290	1	**0.967**	0.602	0.495	**0.991**	**0.957**
PCB151	0.627	0.143	**0.892**	0.184	0.195	0.244	**0.967**	1	0.574	0.461	**0.980**	**0.937**
PCB157	0.395	0.633	0.669	0.658	0.642	0.773	0.602	0.574	1	0.856	0.576	0.550
PCB167	0.330	0.670	0.602	0.687	0.834	0.844	0.495	0.461	0.856	1	0.455	0.430
PCB189	0.639	0.167	**0.929**	0.193	0.152	0.249	**0.991**	**0.980**	0.576	0.455	1	**0.969**
PCB195	0.628	0.161	**0.892**	0.183	0.145	0.235	**0.957**	**0.937**	0.550	0.430	**0.969**	1
VIF	1.9	5.3	**40.2**	5.2	11.5	6.3	**208.3**	**61.0**	6.4	11.1	**196.0**	**19.5**

Appendix 2

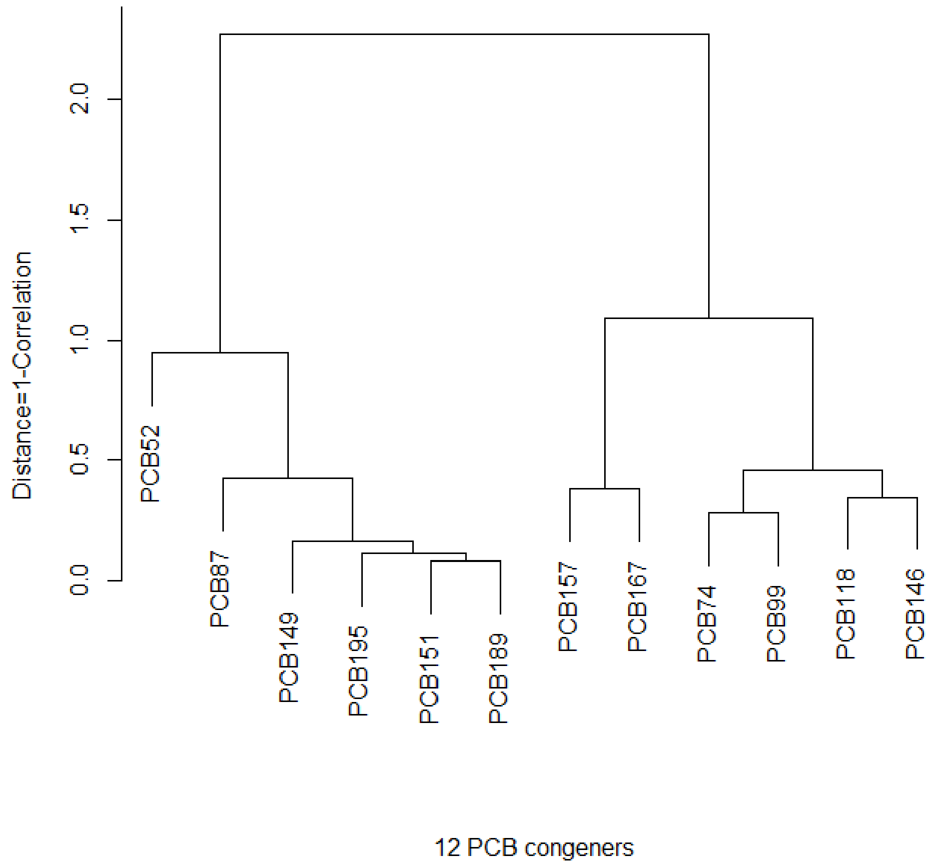
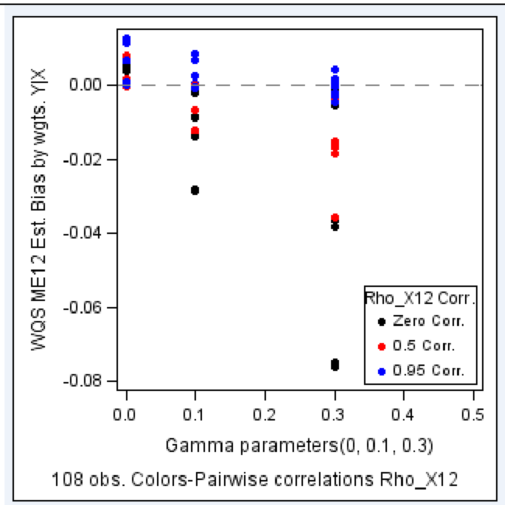


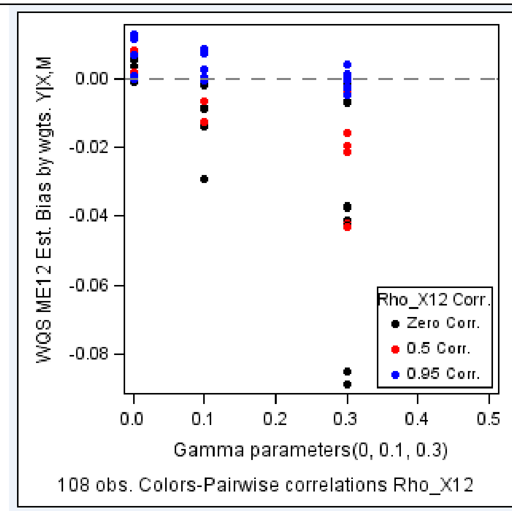
Figure A2.1 Predictor pairwise correlation hierarchical clustering diagram

APPENDIX-3

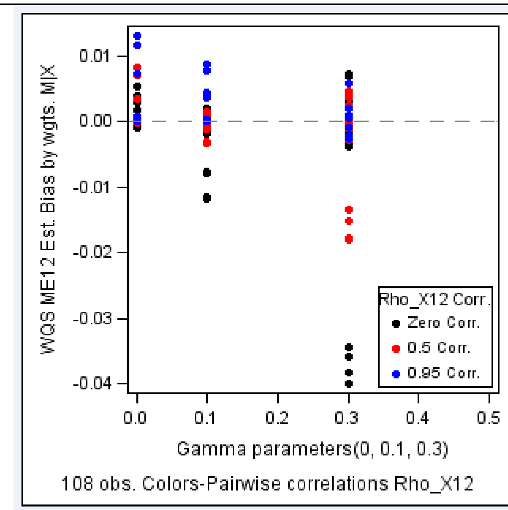
ADDITIONAL PLOTS (PREFERRED vs. LESS PREFERRED METHODS) 2-Predictors Mediation



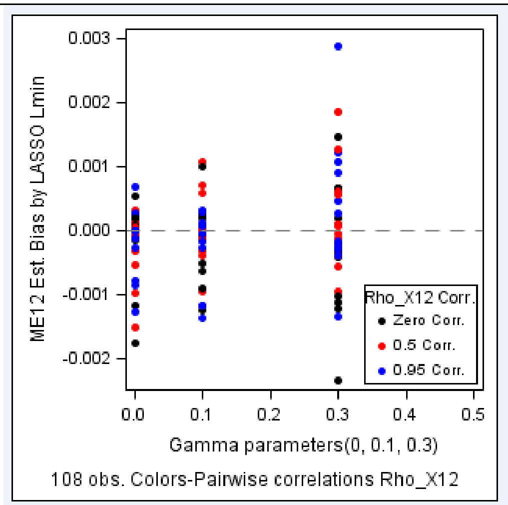
WQS 2V Y|X Estimate Bias



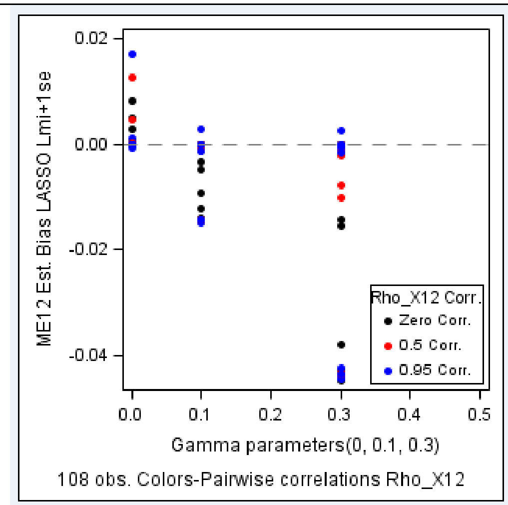
WQS 2V Y|X,M Estimate Bias



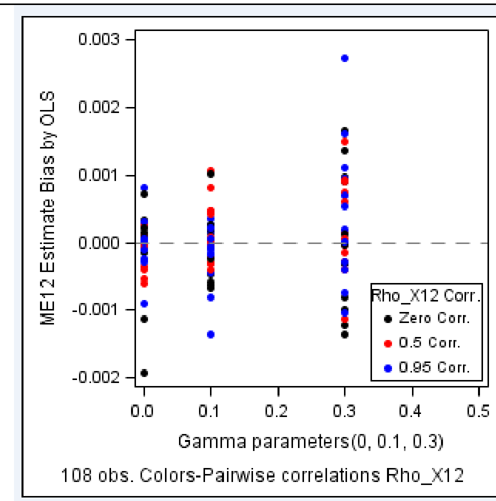
WQS 2V M|X Estimate Bias



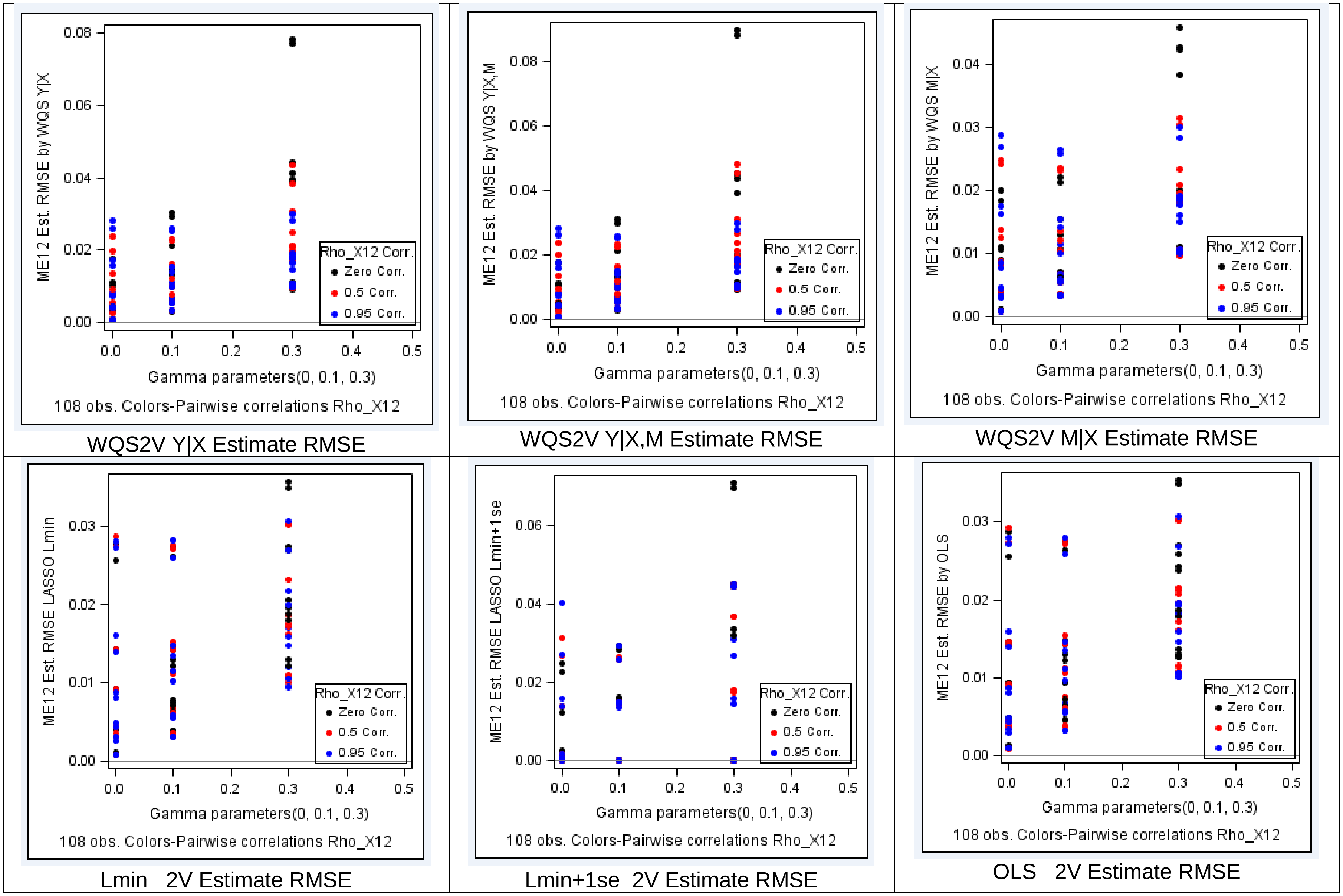
Lmin 2V Estimate Bias

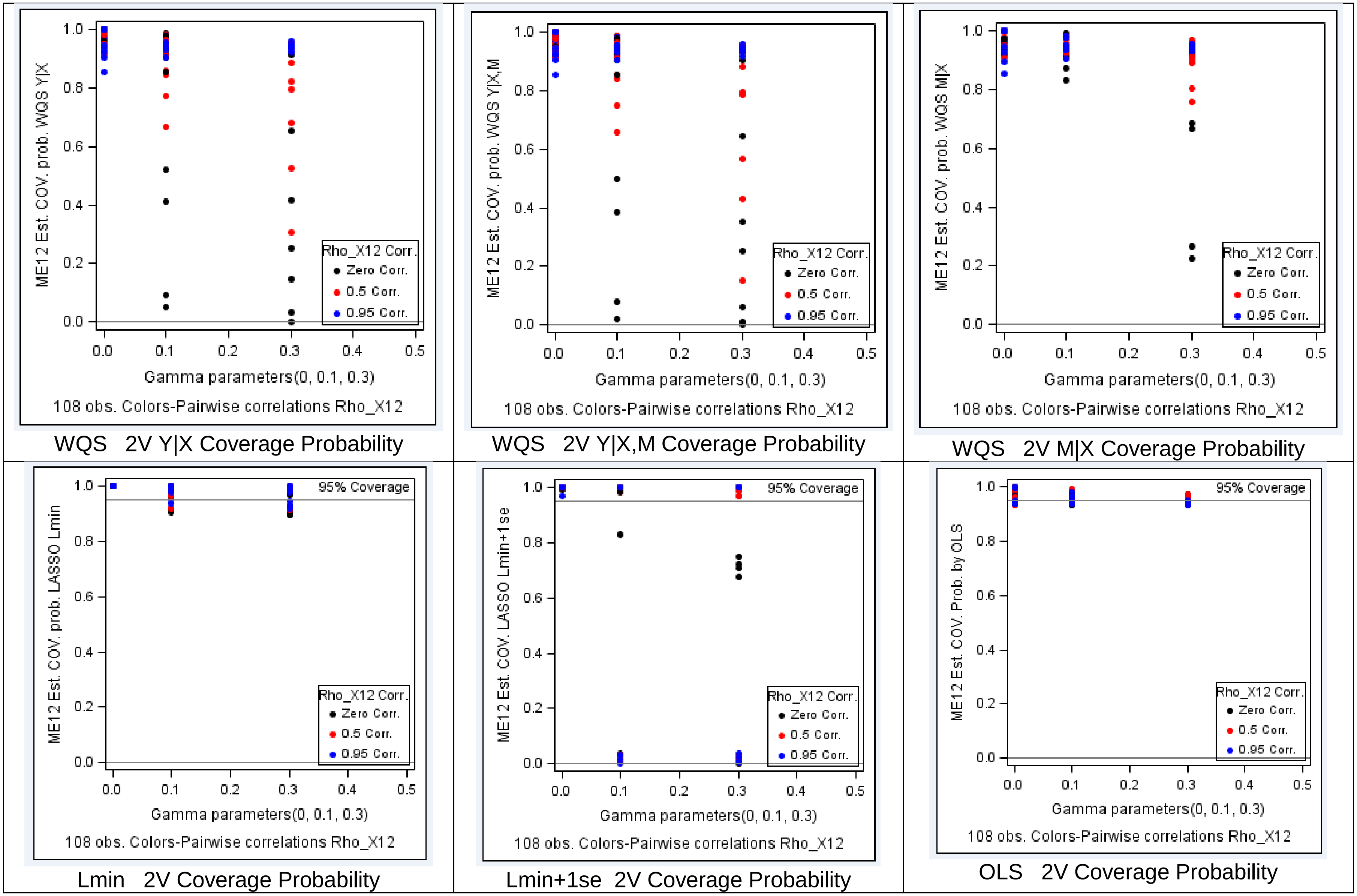


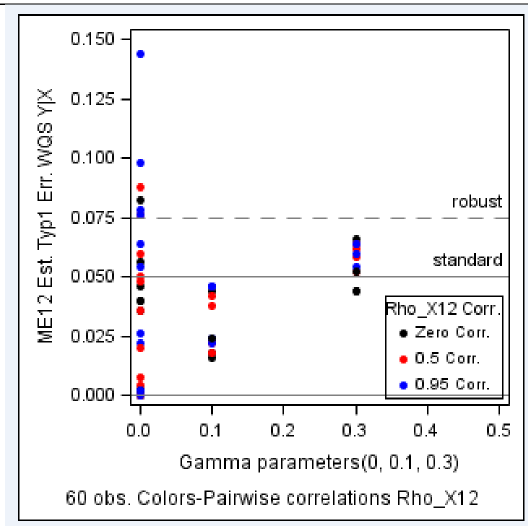
Lmin+1se 2V Estimate Bias



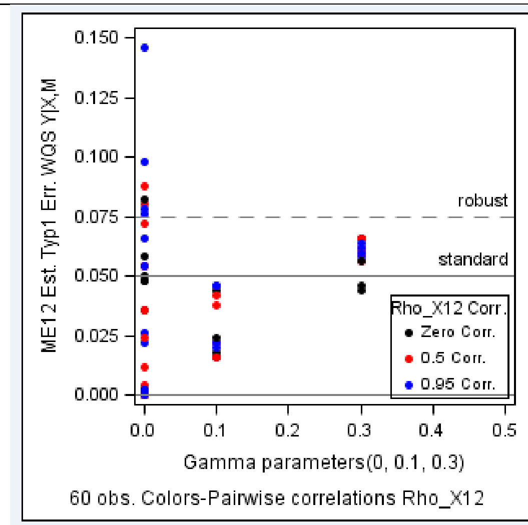
OLS 2V Estimate Bias


ME12 Est. RMSE by WQS Y|X
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
WQS2V Y|X Estimate RMSE
ME12 Est. RMSE by WQS Y|X,M
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
WQS2V Y|X,M Estimate RMSE
ME12 Est. RMSE by WQS M|X
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
WQS2V M|X Estimate RMSE
ME12 Est. RMSE LASSO Lmin
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
Lmin 2V Estimate RMSE
ME12 Est. RMSE LASSO Lmin+1se
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
Lmin+1se 2V Estimate RMSE
ME12 Est. RMSE by OLS
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
OLS 2V Estimate RMSE

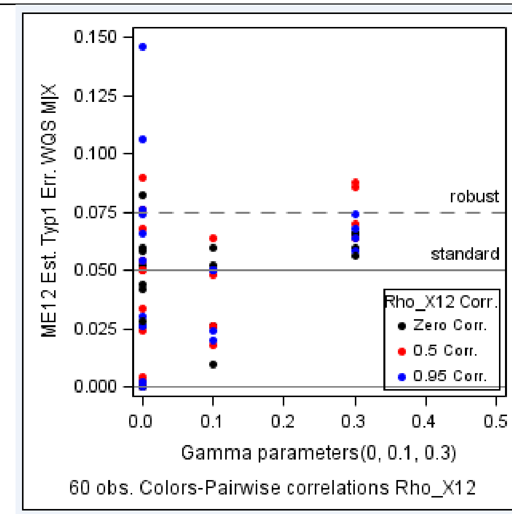

ME12 Est. COV. prob. WQS Y|X
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
WQS 2V Y|X Coverage Probability
ME12 Est. COV. prob. WQS Y|X,M
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
WQS 2V Y|X,M Coverage Probability
ME12 Est. COV. prob. WQS M|X
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
WQS 2V M|X Coverage Probability
ME12 Est. COV. prob. LASSO Lmin
95% Coverage
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Lmin 2V Coverage Probability
ME12 Est. COV. LASSO Lmin+1se
95% Coverage
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
Lmin+1se 2V Coverage Probability
ME12 Est. COV. Prob. by OLS
95% Coverage
Gamma parameters(0, 0.1, 0.3)
108 obs. Colors-Pairwise correlations Rho_X12
OLS 2V Coverage Probability



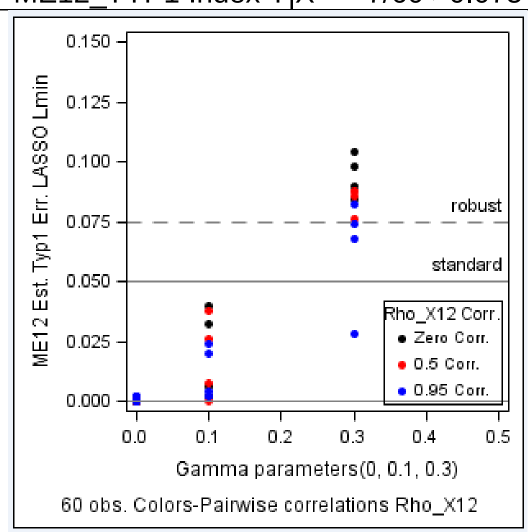
ME12_TYP1 Index Y|X 7/60 >0.075



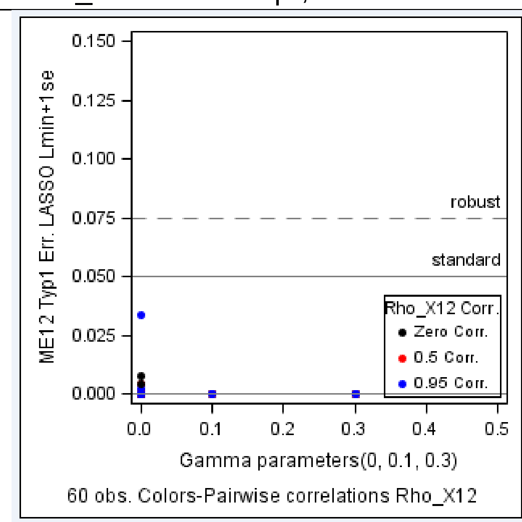
ME12_TYP1 Index Y|X,M 7/60 >0.075



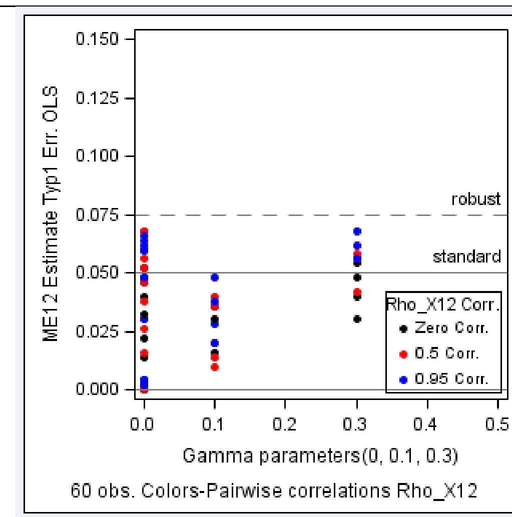
ME12_TYP1 Index M|X 8/60 >0.075



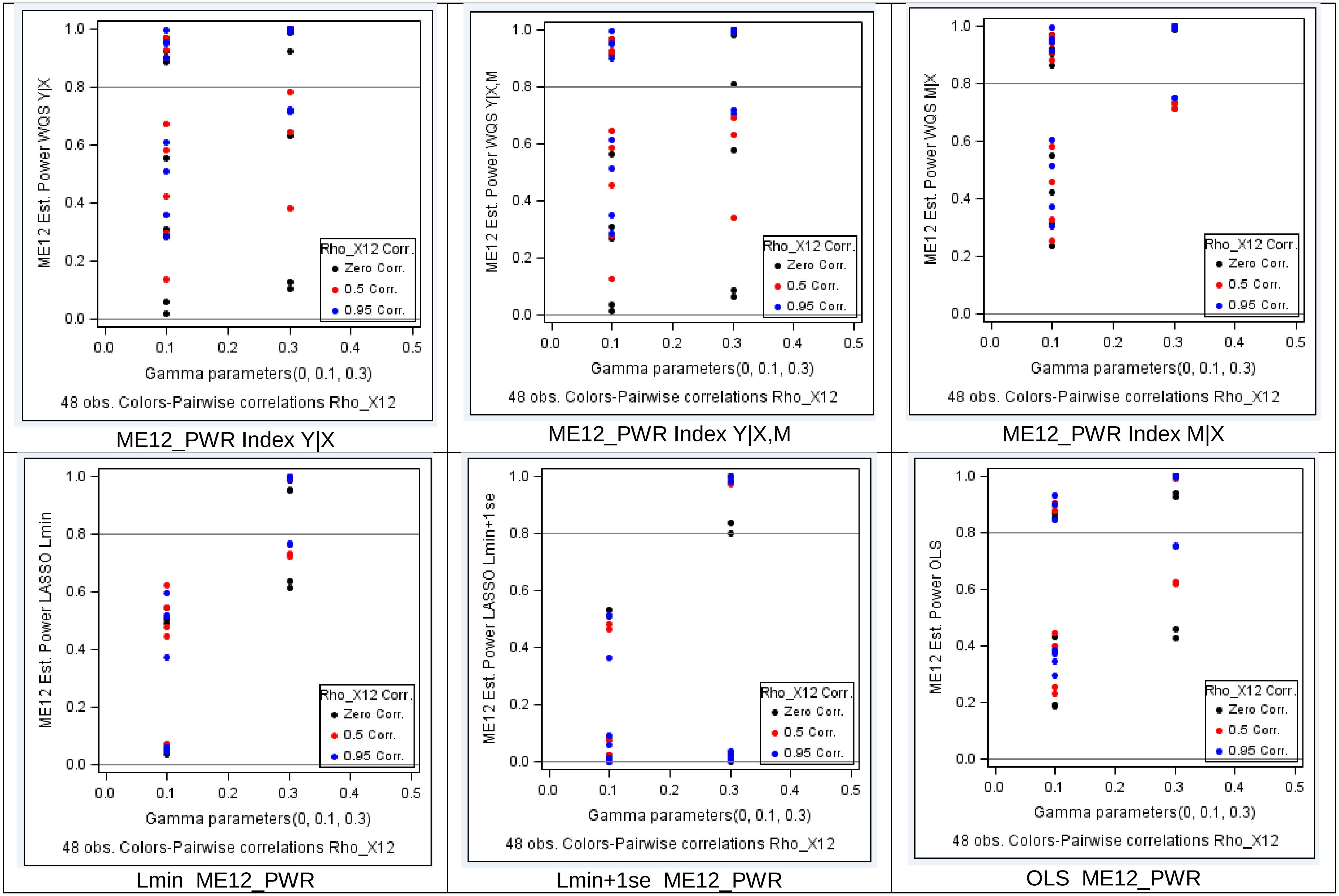
Lmin ME12_TYP1 9/60 >0.075

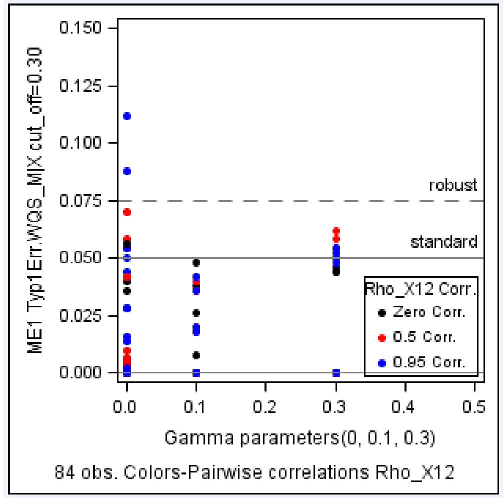


Lmin+1se ME12_TYP1 0/60 >0.075

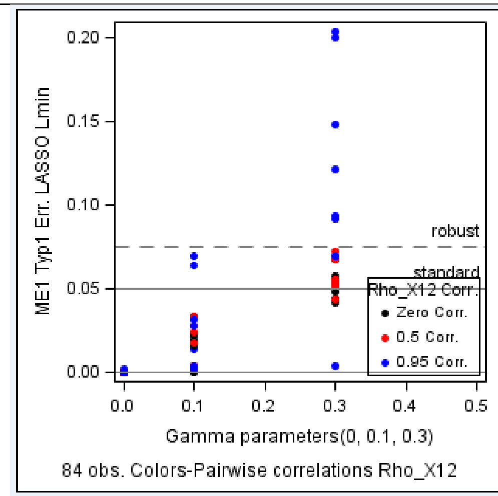


OLS ME12_TYP1 0/60 >0.075

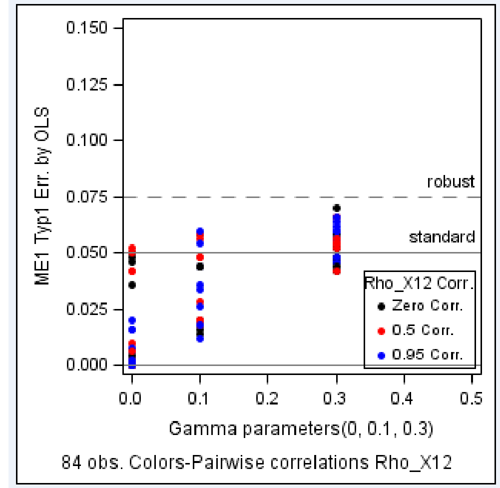

ME12 Est. Power WQS Y|X
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
ME12_PWR Index Y|X
ME12 Est. Power WQS Y|X,M
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
ME12_PWR Index Y|X,M
ME12 Est. Power WQS M|X
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
ME12_PWR Index M|X
ME12 Est. Power LASSO Lmin
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
Lmin ME12_PWR
ME12 Est. Power LASSO Lmin+1se
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
Lmin+1se ME12_PWR
ME12 Est. Power OLS
Gamma parameters(0, 0.1, 0.3)
48 obs. Colors-Pairwise correlations Rho_X12
Rho_X12 Corr.
Zero Corr.
0.5 Corr.
0.95 Corr.
OLS ME12_PWR



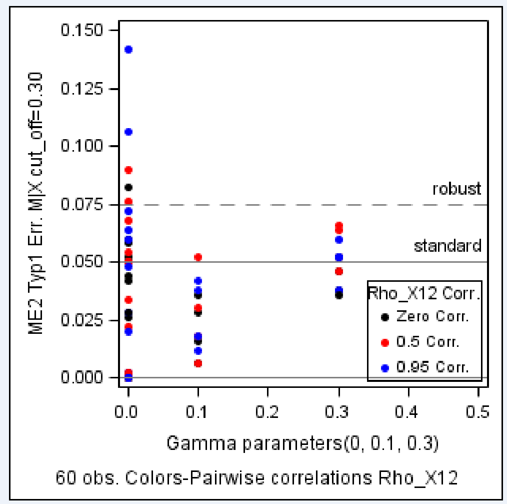
ME1 WQS TYP1 Err. Index M|X 2/60>.075



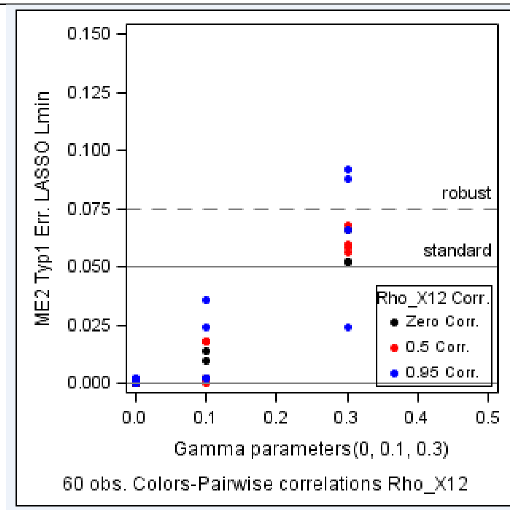
ME1 TYP1 Err. LASSO$_{min}$ 6/84 >0.075



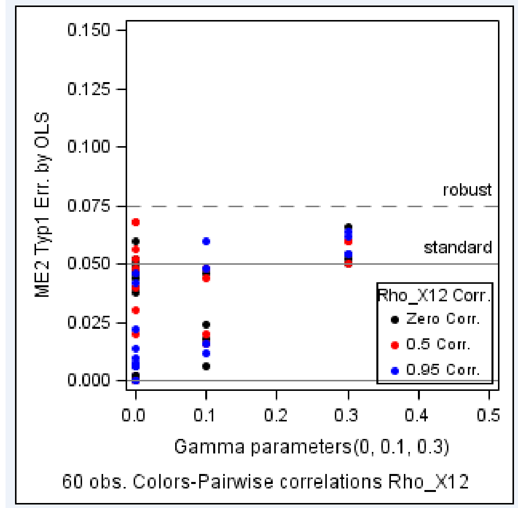
ME1 TYP1 Err. OLS 0/84 >0.075



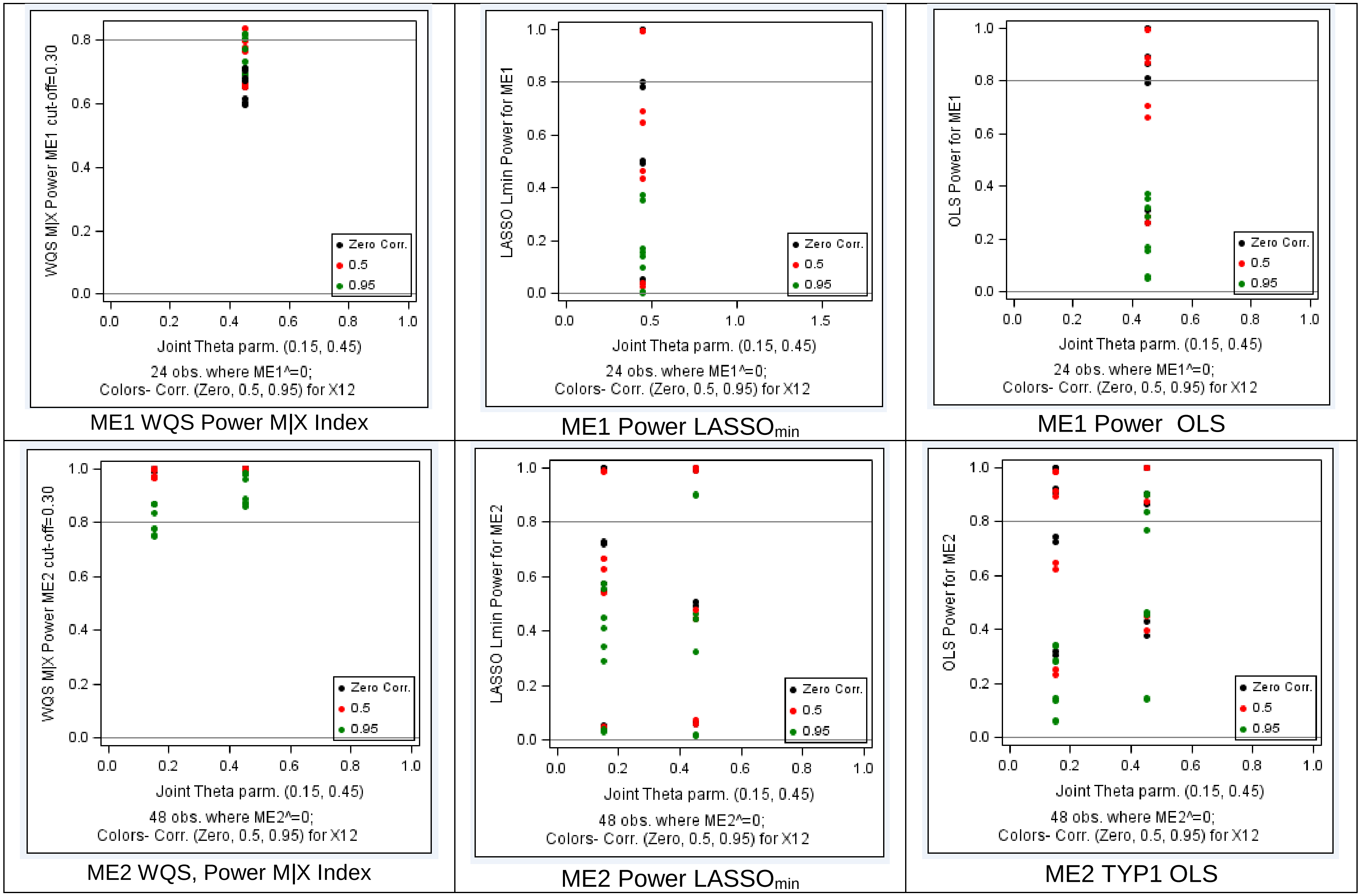
ME2 WQS TYP1 Err. Index M|X 5/60>.075



ME2 TYP1 Err. LASSO$_{min}$ 2/84>0.075

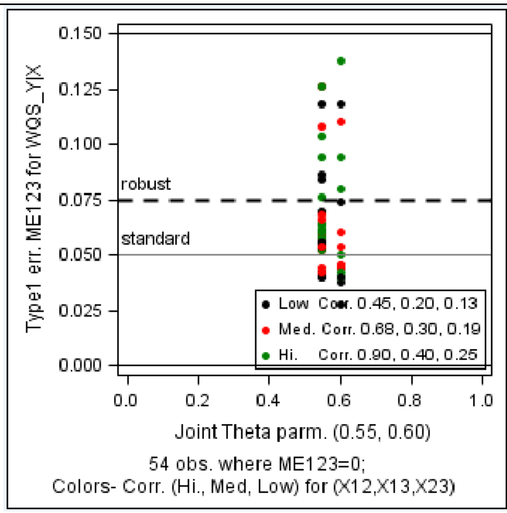


ME2 TYP1 Err. OLS 0/84 >0.075

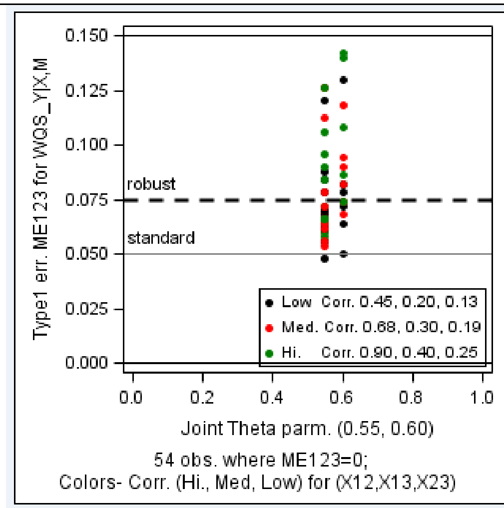

WQS M|X Power ME1 cut-off=0.30
Joint Theta parm. (0.15, 0.45)
24 obs. where ME1^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
Zero Corr.
0.5
0.95
ME1 WQS Power M|X Index
LASSO Lmin Power for ME1
Joint Theta parm. (0.15, 0.45)
24 obs. where ME1^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
ME1 Power LASSOmin
OLS Power for ME1
Joint Theta parm. (0.15, 0.45)
24 obs. where ME1^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
ME1 Power OLS
WQS M|X Power ME2 cut-off=0.30
Joint Theta parm. (0.15, 0.45)
48 obs. where ME2^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
ME2 WQS, Power M|X Index
LASSO Lmin Power for ME2
Joint Theta parm. (0.15, 0.45)
48 obs. where ME2^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
ME2 Power LASSOmin
OLS Power for ME2
Joint Theta parm. (0.15, 0.45)
48 obs. where ME2^=0;
Colors- Corr. (Zero, 0.5, 0.95) for X12
ME2 TYP1 OLS

APPENDIX-4

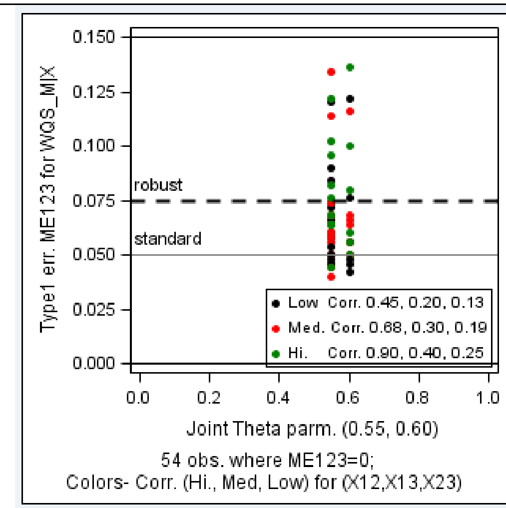
ADDITIONAL PLOTS (PREFERRED vs. LESS PREFERRED METHODS) 3-Predictors Mediation



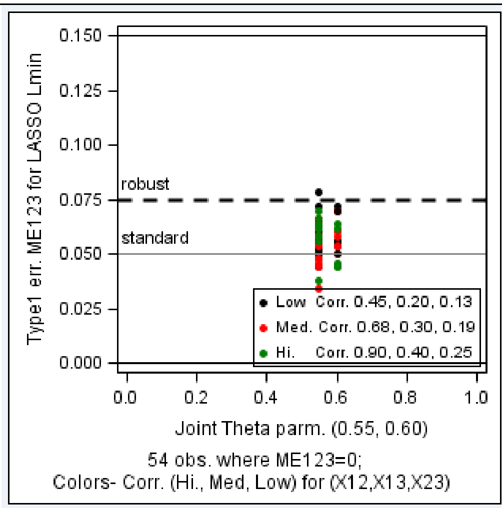
ME123_TYP1 Index Y|X 14/54 >0.075



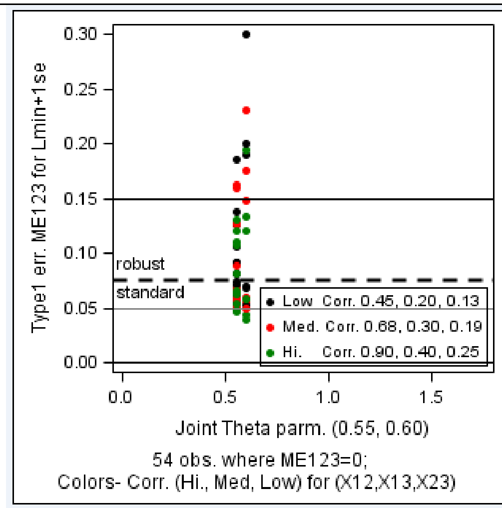
ME123_TYP1 Index Y|X,M 24/54>0.075



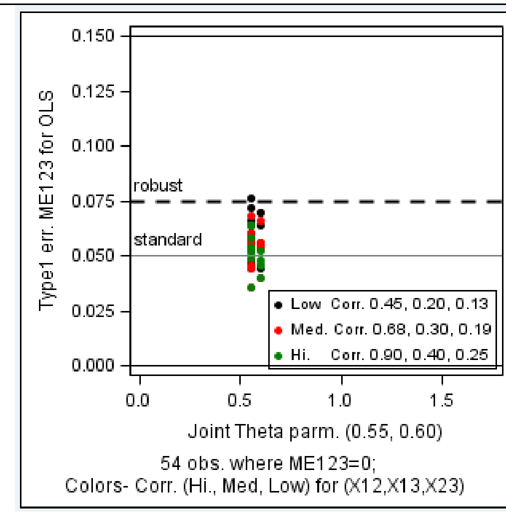
ME123_TYP1 Index M|X 16/54 >0.075



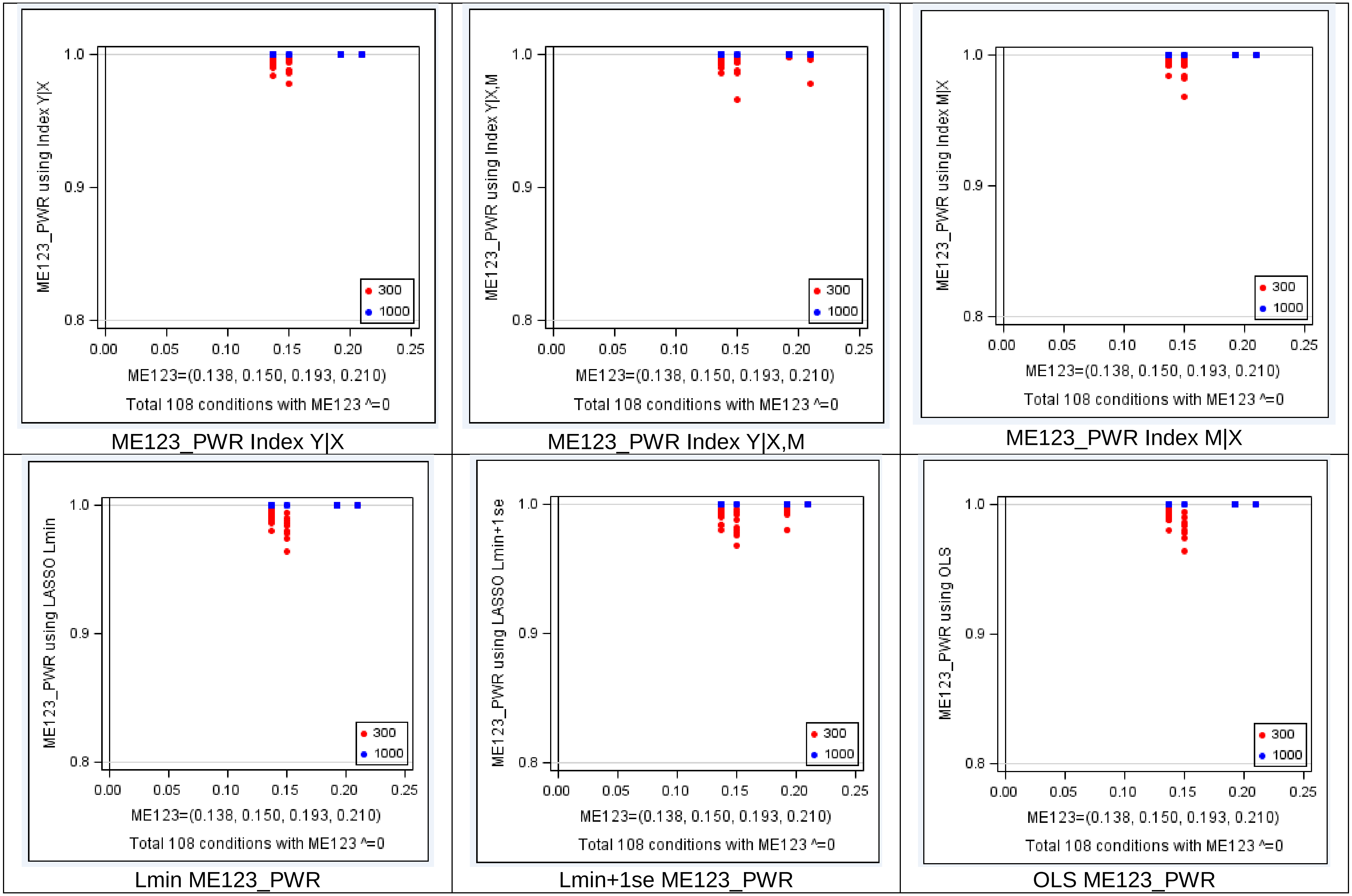
Lmin ME123_TYP1 1/54 > 0.075

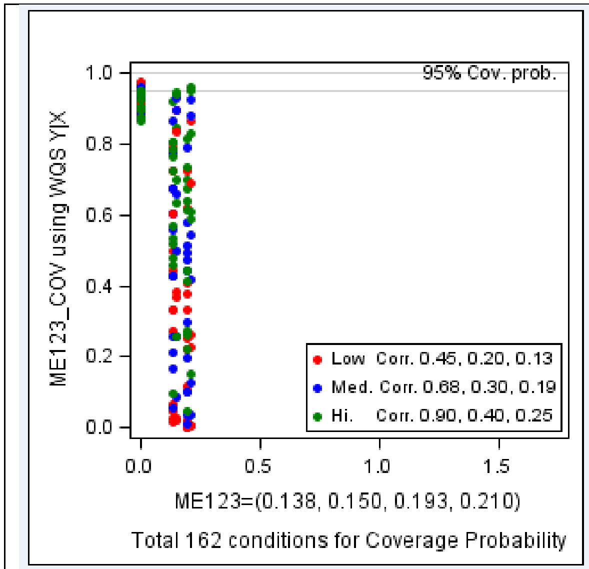


Lmin+1se ME123_TYP1 26/54 > 0.075

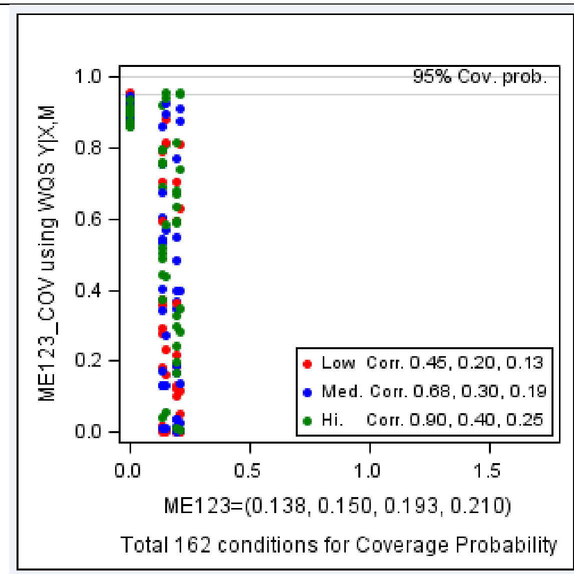


OLS ME123_TYP1 1/54 > 0.075

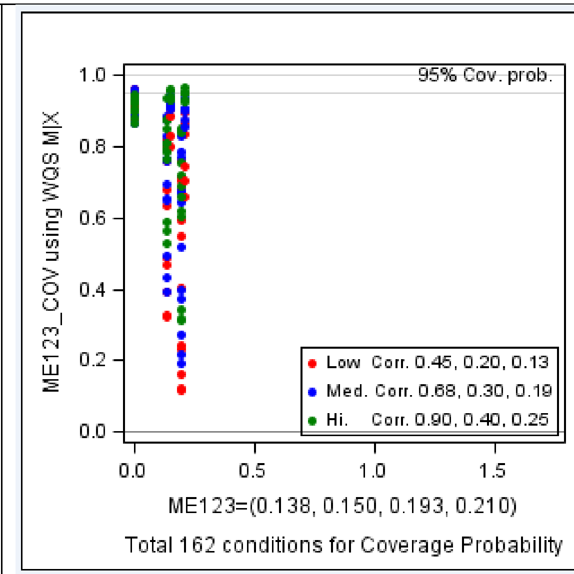

ME123_PWR using Index Y|X
ME123_PWR using Index Y|X,M
ME123_PWR using Index M|X
ME123_PWR using LASSO Lmin
ME123_PWR using LASSO Lmin+1se
ME123_PWR using OLS
1.0
0.9
0.8
0.00
0.05
0.10
0.15
0.20
0.25
300
1000
ME123=(0.138, 0.150, 0.193, 0.210)
Total 108 conditions with ME123 ^=0
ME123_PWR Index Y|X
ME123_PWR Index Y|X,M
ME123_PWR Index M|X
Lmin ME123_PWR
Lmin+1se ME123_PWR
OLS ME123_PWR



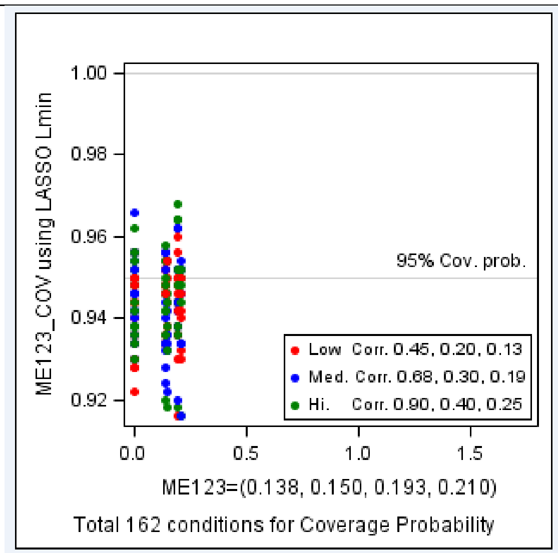
WQS Y|X Coverage Probability



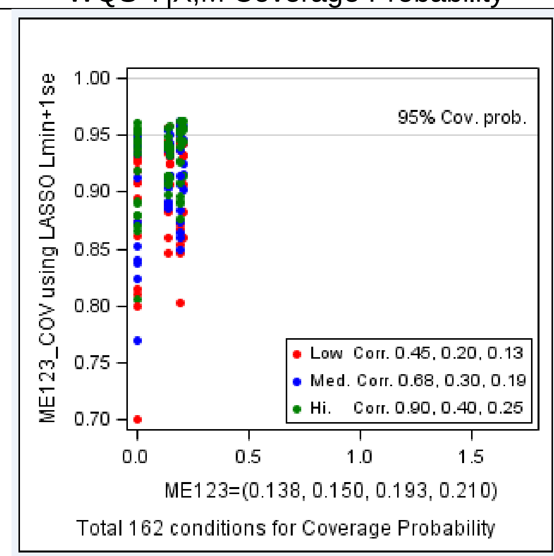
WQS Y|X,M Coverage Probability



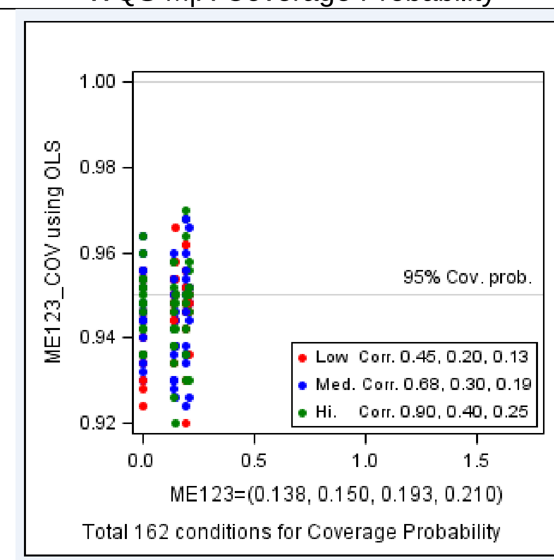
WQS M|X Coverage Probability



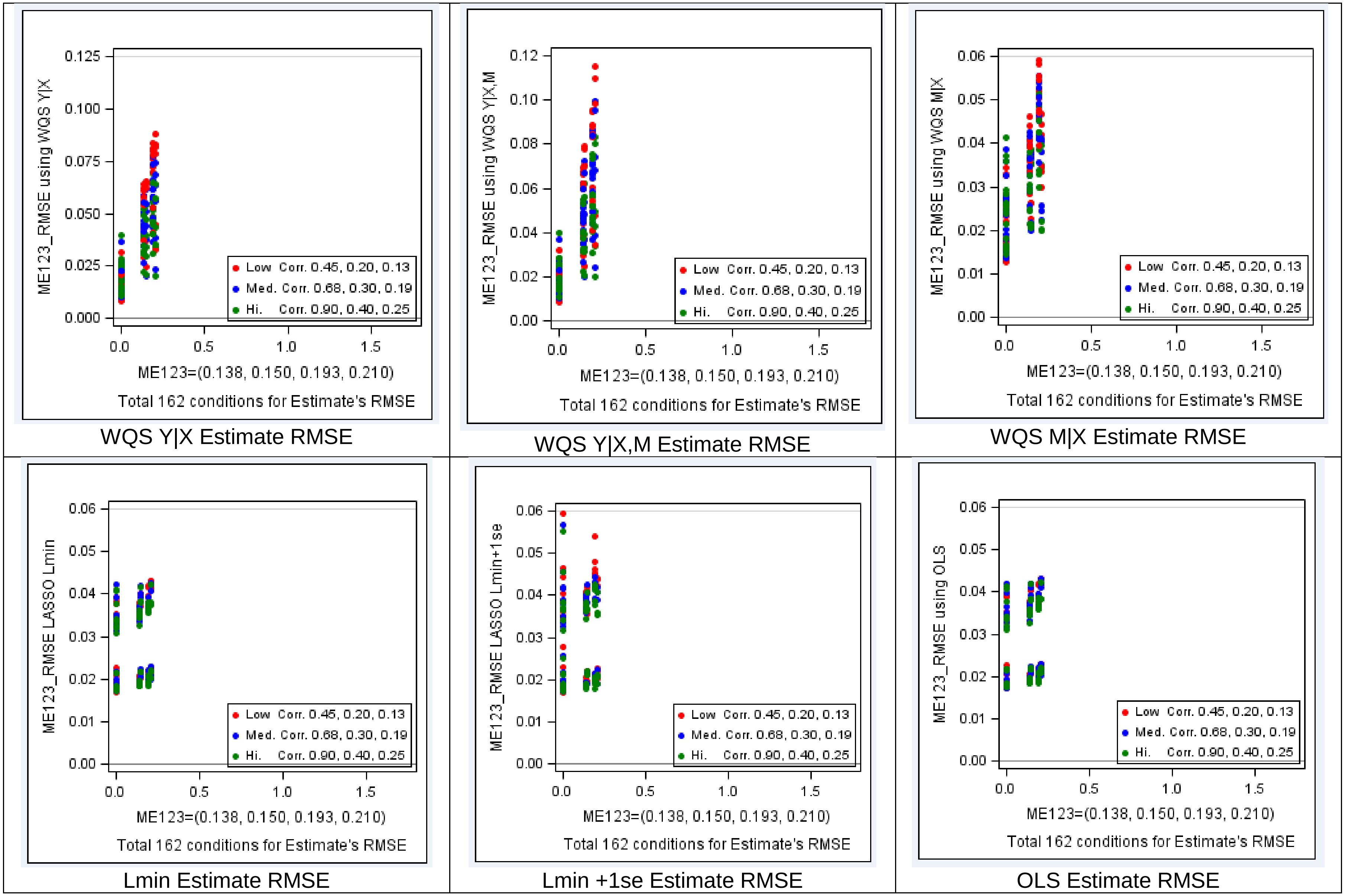
Lmin Coverage Probability

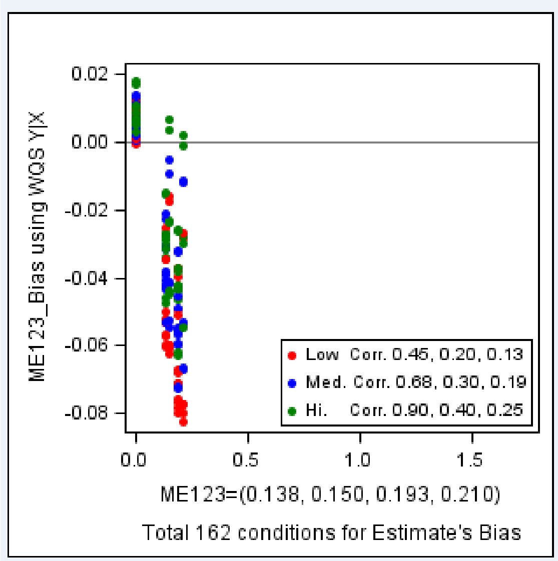


Lmin +1se Coverage Probability

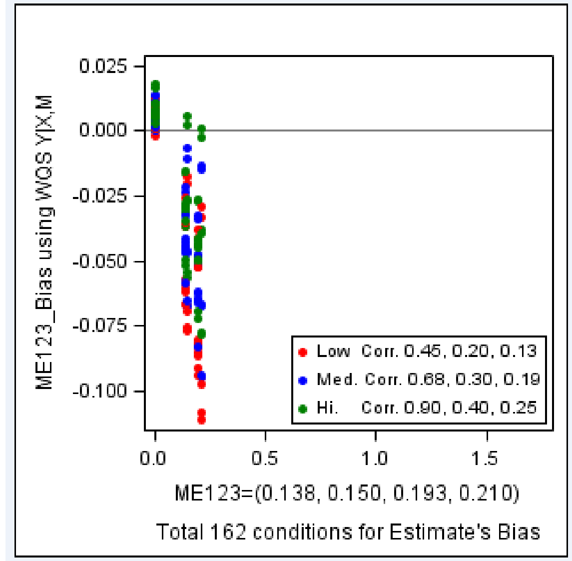


OLS Coverage Probability

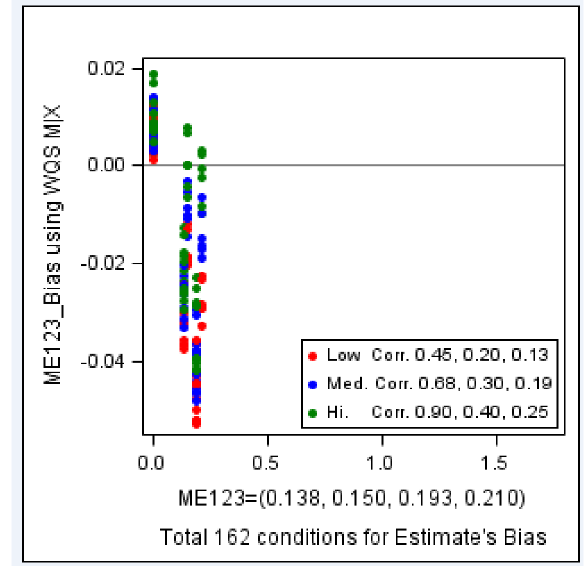

ME123_RMSE using WQS Y|X
0.125
0.100
0.075
0.050
0.025
0.000
0.0
0.5
1.0
1.5
ME123=(0.138, 0.150, 0.193, 0.210)
Total 162 conditions for Estimate's RMSE
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
WQS Y|X Estimate RMSE
ME123_RMSE using WQS Y|X,M
0.12
0.10
0.08
0.06
0.04
0.02
0.00
WQS Y|X,M Estimate RMSE
ME123_RMSE using WQS M|X
0.06
0.05
0.04
0.03
0.02
0.01
0.00
WQS M|X Estimate RMSE
ME123_RMSE LASSO Lmin
Lmin Estimate RMSE
ME123_RMSE LASSO Lmin+1se
Lmin +1se Estimate RMSE
ME123_RMSE using OLS
OLS Estimate RMSE



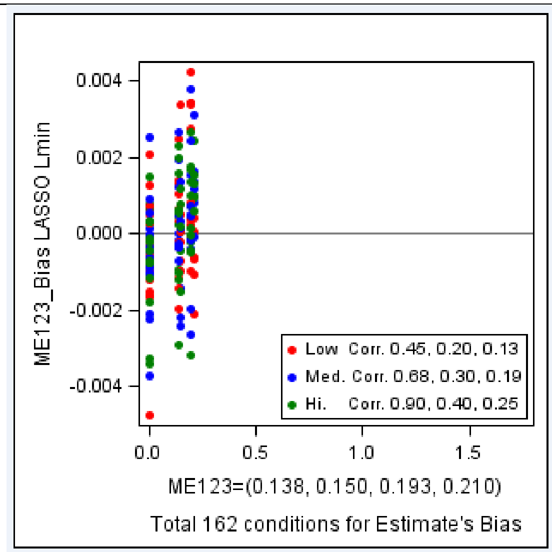
WQS Y|X Estimate Bias



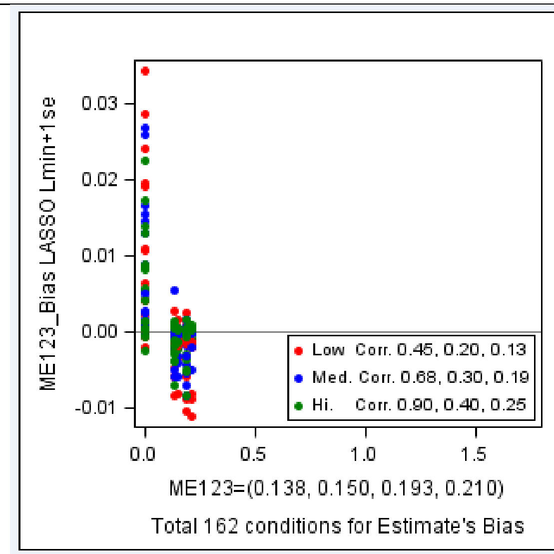
WQS Y|X,M Estimate Bias



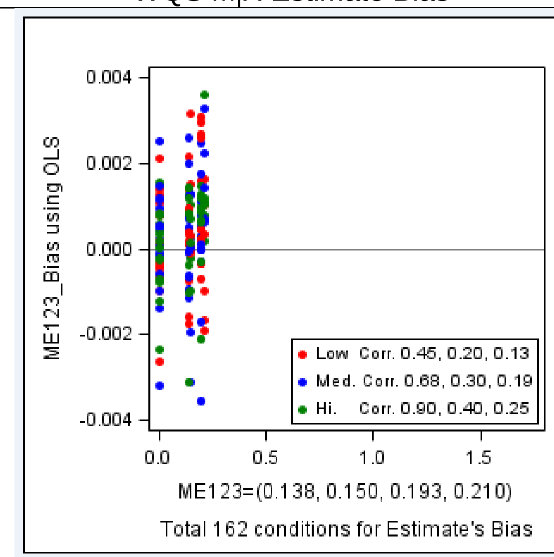
WQS M|X Estimate Bias



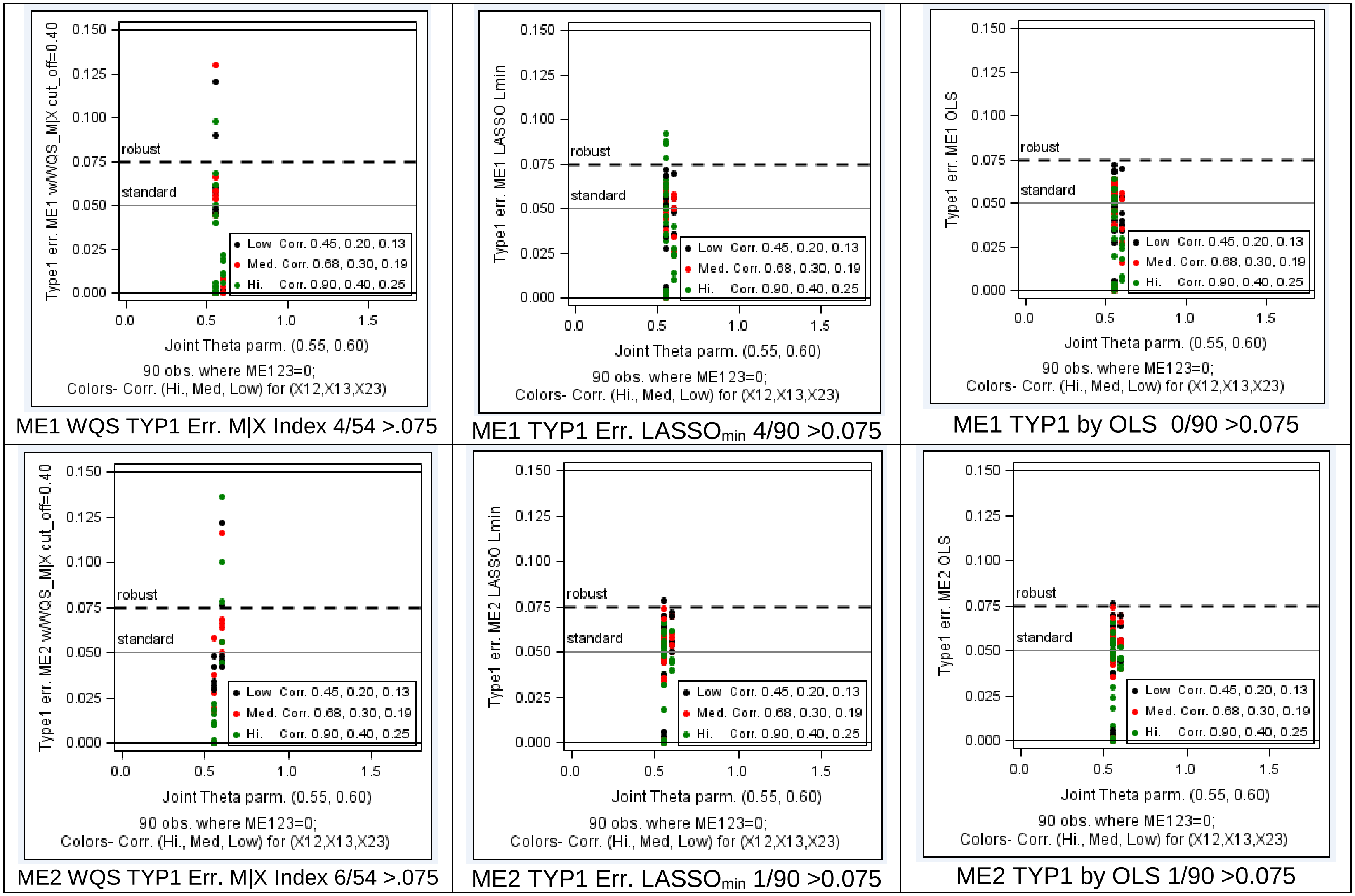
Lmin Estimate Bias



Lmin +1se Estimate Bias



OLS Estimate Bias



ME1 WQS TYP1 Err. M|X Index 4/54 >.075

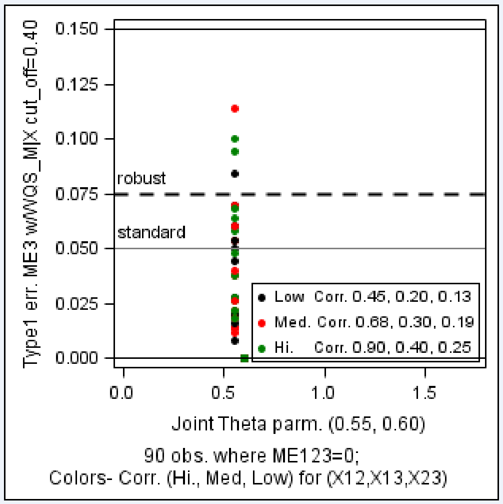
ME1 TYP1 Err. $LASSO_{min}$ 4/90 >0.075

ME1 TYP1 by OLS 0/90 >0.075

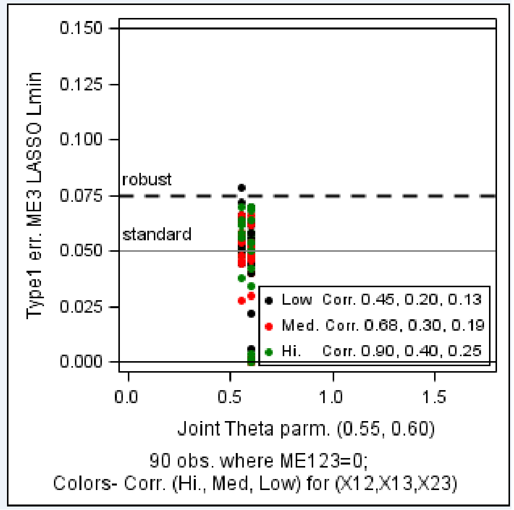
ME2 WQS TYP1 Err. M|X Index 6/54 >.075

ME2 TYP1 Err. $LASSO_{min}$ 1/90 >0.075

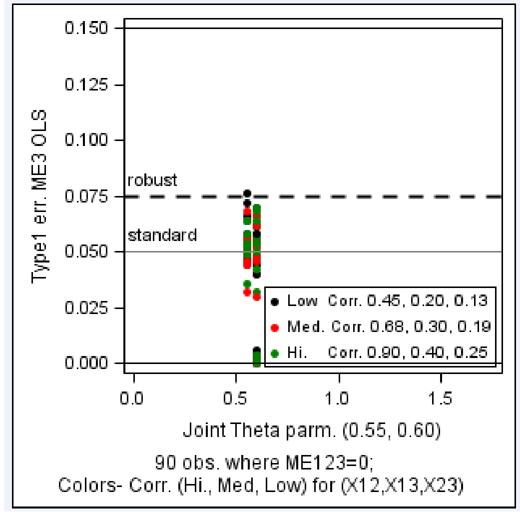
ME2 TYP1 by OLS 1/90 >0.075



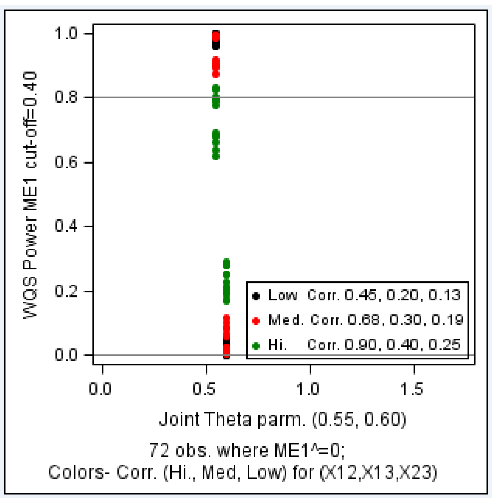
ME3 WQS TYP1 Err. M|X Index 4/54 >.075



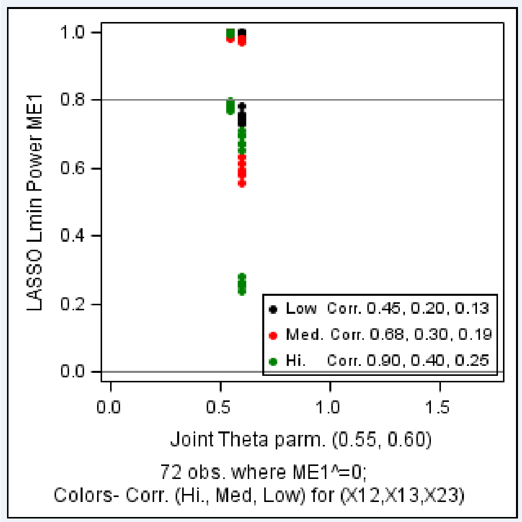
ME3 TYP1 Err. $\text{LASSO}_{\text{min}}$ 1/90 > 0.075



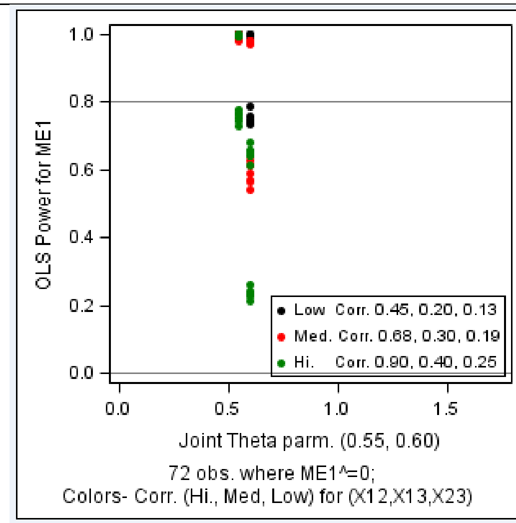
ME3 TYP1 by OLS 1/90 > 0.075



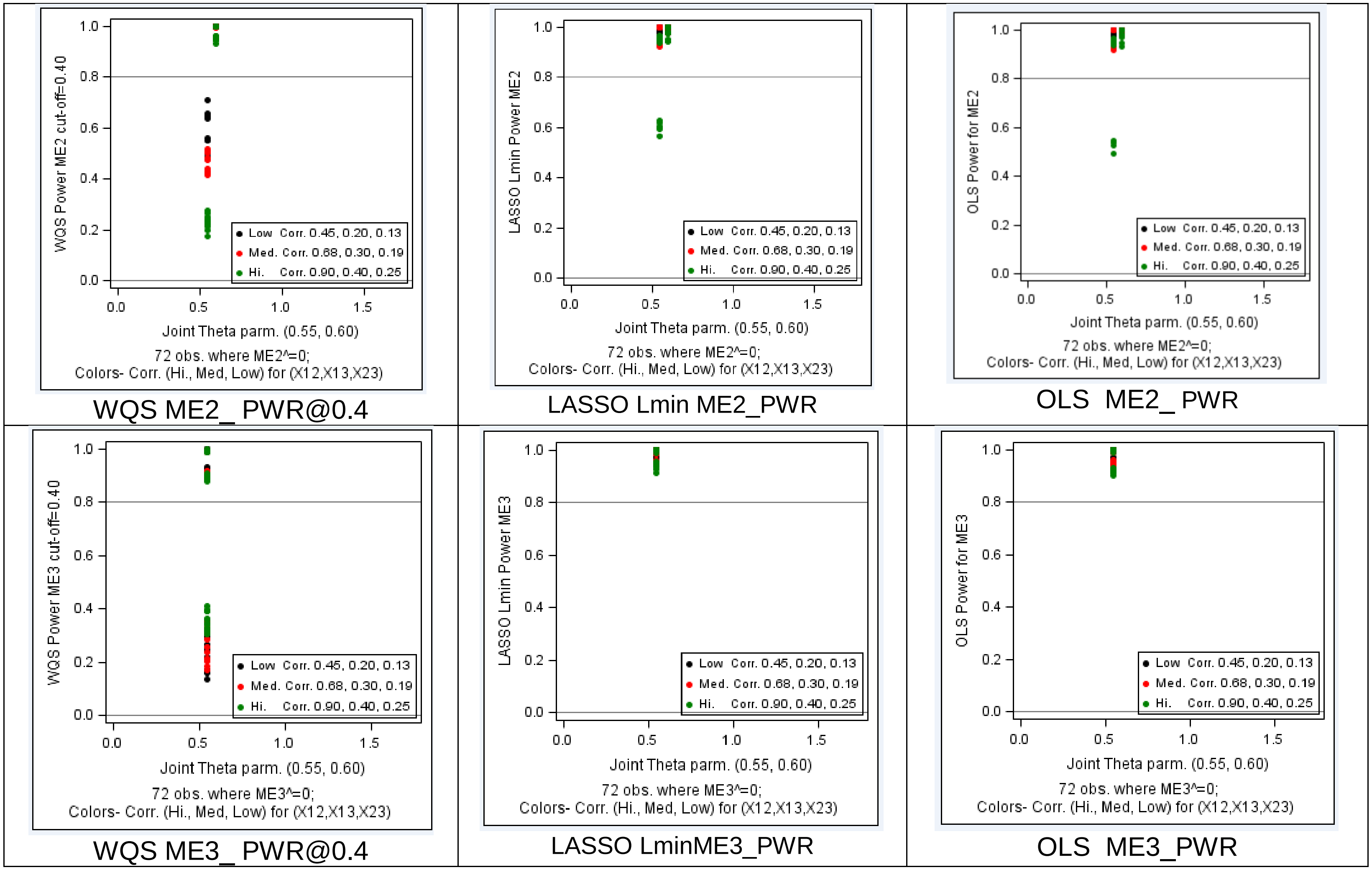
WQS ME1_PWR@0.4



LASSO LminME1_PWR



OLS ME1_PWR


WQS Power ME2 cut-off=0.40
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME2^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
WQS ME2_ PWR@0.4
LASSO Lmin Power ME2
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME2^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
LASSO Lmin ME2_PWR
OLS Power for ME2
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME2^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
OLS ME2_ PWR
WQS Power ME3 cut-off=0.40
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME3^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
WQS ME3_ PWR@0.4
LASSO Lmin Power ME3
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME3^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
LASSO LminME3_PWR
OLS Power for ME3
Low Corr. 0.45, 0.20, 0.13
Med. Corr. 0.68, 0.30, 0.19
Hi. Corr. 0.90, 0.40, 0.25
Joint Theta parm. (0.55, 0.60)
72 obs. where ME3^=0;
Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)
OLS ME3_PWR

APPENDIX-5

PROGRAM CODES (SAS AND R)

GENERATING SIMULATED DATASETS

```
OPTIONS source pagesize=256 linesize=80 nodate replace; OPTIONS FORMCHAR="|----|+|---+=|-/\<>*";
goptions device=jpeg gsfname=pic gsfmode=replace colors= (black) htext=3 ftext=swiss hby=1.5; pattern
c=green v=solid; RUN;

*libname Temp "/home/evanibm/M_A"; /* RUN on the cluster*/

* Simulated predictor DATA using Cholesky decomposition to generate correlated X-DATA. Start with a
correlation matrix for Y X1 X2 M

/*Initialize a standard normal array matrix of size 500 *4(Y X1 X2 M) each of Num=300 DATA. Selected
values for Regression parameters and Correlation between predictors Xjs */

ODS listing;
PROC IML;
DO i=1 to 162;
USE Temp.Corr_Matrix_3Var; /* Make sure these Conditions are available for this program */
READ all into Possible3;
CLOSE Temp.Corr_Matrix_3Var;
Std_YXM = j (Possible3 [i, 2], 5, .);/* Allocate an array for a single DATASET of N=300 or N=1,000 with 5
columns */

DO sampnum=1 to 500;
CALL randseed (100653+sampnum);
CALL randgen (Std_YXM, "Normal"); /* N5 (0, 1) */

IF sampnum=1 then do;
  NAMES= {"Std_Y","Std_X1","Std_X2","Std_X3","Std_M"};

Create Std_YXM from Std_YXM[c=names];
APPEND from Std_YXM;
Close Std_YXM; /* this single DATA SET is used for Lasso Regression */
END;
else do;
    NAMES= {"Std_Y","Std_X1","Std_X2","Std_X3","Std_M"};
   Edit Std_YXM;
   APPEND from Std_YXM;
   Close Std_YXM;/* This 300 DATASET is used for WQS regression */
END;
END; /* DO loop sampnum */

/* calculate the corresponding correlation matrix Rho=SET value*/
   SIZE=Possible3 [i, 2]; /* DATA size is 300 or 1000 */
 Rho12=Possible3 [i, 3]; Rho13=Possible3 [i, 4]; Rho23=Possible3 [i, 5]; Gamma=Possible3 [i,
6];Theta1=Possible3 [i, 7];
Theta2=Possible3 [i, 8]; Theta3=Possible3 [i, 9]; Beta1=Possible [i, 10];   Beta2=Possible3 [i, 11];
         Beta3=Possible3 [i, 12];
```

```
   SL=Possible3 [i, 1]; /* Serial number of the condition being RUN */

Rho_1y= Beta1 + Theta1*Gamma + Rho12*Theta2*Gamma + Rho12*Beta2 + Rho13*Theta3*Gamma +
Rho13*Beta3;
Rho_2y= Beta2 + Theta2*Gamma + Rho12*Theta1*Gamma + Rho12*Beta1 + Rho23*Theta3*Gamma +
Rho23*Beta3;
Rho_3y= Beta3 + Theta3*Gamma + Rho23*Theta2*Gamma + Rho23*Beta2 + Rho13*Theta1*Gamma +
Rho13*Beta1;
Rho_1m=Theta1 + Rho12*Theta2 + Rho13*Theta3;
Rho_2m=Theta2 + Rho12*Theta1 + Rho23*Theta3;
Rho_3m=Theta3 + Rho23*Theta2 + Rho13*Theta1;
Rho_my=Gamma+ Theta1*Beta1+ Theta2*Beta2+ Theta3*Beta3 + Theta1*Rho12*Beta2+
Theta1*Rho13*Beta3 + Theta2*Rho12*Beta1+ Theta2*Rho23*Beta3 +
Theta3*Rho13*Beta1+Theta3*Rho23*Beta2;

Corr_Matrix = shape ({0}, 5, 5, 0); /* Initialize the Correlation Matrix */
   PRINT Corr_Matrix;
/* SET values in the Correlation Matrix */
Corr_Matrix [1, 1] =1;          Corr_Matrix [1, 2] =Rho_1y;     Corr_Matrix [1, 3] =Rho_2y;
         Corr_Matrix [1, 4] =Rho_3y;
Corr_Matrix [1, 5] =Rho_my;     Corr_Matrix [2, 1] =Rho_1y;     Corr_Matrix [2, 2] =1;
         Corr_Matrix [2, 3] =Rho12;
Corr_Matrix [2, 4] =Rho13;      Corr_Matrix [2, 5] =Rho_1m;     Corr_Matrix [3, 1] =Rho_2y;
         Corr_Matrix [3, 2] =Rho12;
Corr_Matrix [3, 3] =1;          Corr_Matrix [3, 4] =Rho23;      Corr_Matrix [3, 5] =Rho_2m;
         Corr_Matrix [4, 1] =Rho_3y;
Corr_Matrix [4, 2] =Rho13;      Corr_Matrix [4, 3] =Rho23;      Corr_Matrix [4, 4] =1;
         Corr_Matrix [4, 5] =Rho_3m;
Corr_Matrix [5, 1] =Rho_my;     Corr_Matrix [5, 2] =Rho_1m;     Corr_Matrix [5, 3] =Rho_2m;
         Corr_Matrix [5, 4] =Rho_3m;
Corr_Matrix [5, 5] =1;
         /* Correlation between Y X1 X2 X3 and M Corr_X1 X2 SET to 0.95  */

rname1 = {"Y" "X1" "X2" "X3" "M"};
cname1 = {"Y" "X1" "X2" "X3" "M"};
MATTRIB Corr_Matrix rowname=rname1 colname=cname1;
PRINT Corr_Matrix;
/* Correlation between Y X1 X2 X3 and M in 500 replications of num=300 original DATASET*/
Use Std_YXM;
Read all into Std_YXM;
Close Std_YXM;  /* 500 DATASETs of N=300 or N=1000 Std Normal YX1 X2 M values */

/* obtain the Cholesky decomposition from the Covariance Matrix for Y X1 X2 and M */
/* create the Rho vector of std_dev for each column of Y X's which is 1 */
D=I (5);
c=1.0; /* Dampening factor for off diagonals in covariance matrix to maintain PSD */

Corr1_c = c*(Corr_Matrix – I (5)) + I(5); /* Using Corr_Matrix (Corr X1X2=0.95) or Corr_Matrix0 (Corr
X1X2=0) */

COV_YXM = D * Corr1_c * D`; /* Obtain the Covariance Matrix from the Correlation Matrix */
                              *names= {"Y" "X1" "X2" "X3" "M"};
/* Generate the Upper Triangle Cholesky matrix for COV_X3 for Replicate 1*/
         U1 = half (COV_YXM); /* A 5x5 upper triangle Cholesky decomposition of Sigma matrix */
```

```
          V1K= (Std_YXM * U1);  /* Gives you the 500 DATASETs of correlated Y-XM simulated DATASET
of size (num *5) */
          names1 = {"Condition" "Sampnum" "Y" "X1" "X2" "X3" "M"};
          snum = j (size*500, 1,  .);
                    cond = j (size*500, 1, .);
                              DO ii=1 to 500;
                                        DO iter=1 to Possible3 [i, 2];
         Cond [(ii-1)*Possible3 [i, 2] + iter] = Possible3 [i, 1];

          Snum [(ii-1)*Possible3[i, 2]+iter]=ii;
          END; /* END do i loop */
                    END;/* END do sampnum */
          V1K=insert (V1K, cond, 0, 1); V1K=insert(V1K,snum, 0, 2);
/* Save this for 500 iterations for WQS Regression */
IF Possible3[i,1]=1 then do;
  Create Yx1x2m from V1K[c=names1];
                              APPEND from V1K;
          Close Yx1x2m; /* SAS DATASET */
                                        END;
ELSE do;
  Edit Yx1x2m;
  APPEND from V1K;
  Close Yx1x2m;
          END;
/* Save the first iteration for Lasso Regression and calculation of the weights*/
END; /* DO i=1 to 162 */

/* Assign quartile ranks for predictors X1 and X2 */
PROC rank DATA=Yx1x2m out=Temp.Q_RV3 groups=4;
             VAR X1 X2 X3;
             RANKS X1_Rank X2_Rank X3_Rank;
             ATTRIB X1_Rank LABEL="QUARTILE_RANK FOR X1";
             ATTRIB X2_Rank LABEL="QUARTILE_RANK FOR X2";
             ATTRIB X3_Rank LABEL="QUARTILE_RANK FOR X3";
RUN;

/*Verification */
PROC CORR DATA=Temp.Q_RV3 out=Corr_43 Spearman;
VAR X1 X2 X3 Y M;
WHERE Condition=43;
RUN;

%macro GENDATASETs;
%do i=1 %to 162;
DATA Temp9.Q_RV3_&i;
SET Temp.Q_RV3;
WHERE Condition=&i;
RUN;
%END;
%mEND GENDATASETs;
%GENDATASETs;
```

LEGEND OF ACRONYNMS USED.

- ME – Mediated Effect (Indirect effect);
- T- Theta
- 0 - Y|X; B - Y|X,M; T - M|X
- G- Gamma
- GZ- Gamma=0
- T1GZ= Theta1=0 and Gamma=0
- T2GZ= Theta2=0 and Gamma=0
- ME1_T = Individual Indirect effect estimate using M|X weighted WQS Index for predictor X1
- TYP1 = Type 1 Error
- PWR - Power
- DE- Direct Effect
- ME1 = Indirect effect for X1
- ME2 = Indirect effect for X2
- ME3 = Indirect effect for X3
- R12 = Rho_X1,X2
- R13=Rho_X1,X3
- R23=Rho_X2,X3
- T1Z= Theta1=0;T2Z=Theta2=0,T3Z=Theta3=0
- T123 / T12 = Joint Theta parameter
- ME123 /ME12 = Joint Indirect effect
- DE123 / DE12 = Joint Direct effect
- T123_0 = Joint Theta estimated from Y|X weighted WQS Index
- T123_T = Joint Theta estimated from M|X weighted WQS Index
- T123_B = Joint Theta estimated from Y|X,M weighted WQS Index
- ME123_0RMSE = Joint Indirect estimate from Y|X weighted WQS Index
- ME1S_TTYP1 = X1 Indirect Effect estimated Type 1 error rate (Sequential Bootstrapped 95th percentile CI)
- ME1N_TTYP1 = X1 Indirect Effect estimated Type 1 error rate (Nested Bootstrapped 95th percentile CI)
- ME2S_TTYP1 = X2 Indirect Effect estimated Type 1 error rate (Sequential Bootstrapped 95th percentile CI)
- STD - Standardized

etc.

WQS METHOD SEQUENTIAL 3V OPTCODE

```
/* Standardized WQS 3-variable statistics Sequential using 500 Replicated DATASETs
    (Sampnum 1 to 500) for each condition (162 conditions)*/

OPTIONS source pagesize=256 linesize=80 nodate replace;
        OPTIONS FORMCHAR="|----|+|---+=|-/\<>*";
        GOPTIONS device=jpeg gsfname=pic gsfmode=replace colors = (black) htext=3 ftext=swiss hby=1.5;
        PATTERN c=green v=solid;
        *libname Temp "/home/evanibm/M_A/20160929"; /* RUN on the cluster*/
        LIBNAME Temp "H:\Thesis\RESULTS";
*ODS noresults; ODS graphics off; ODS html close; *ODS select none; ODS listing;


DATA WQSEST3v; /* Base file 338 columns*/
Condition=.; N=.;R12=.;R23=.;R13=.;DE1=.; DE2=.; DE3=.; DE123=.;Theta1=.; Theta2=.; Theta3=.;
T123=.;Gamma=.;ME1=.;ME2=.;ME3=.;ME123=.;
wXB01=.; wXB02=.; wXB03=.;wXT1=.; wXT2=.;wXT3=.;   wXB1=.;   wXB2=.;  wXB3=.; cME123=.; cDE123=.;

MEAR120=R12;        MEAR130=R13;       MEAR230=R23;       Theta3=. ;          T123_0=. ;           G_0=. ;
ME123_0 =. ;        ME123_0Bias =. ;   ME123_0RMSE =. ;   ME123_0COV =. ;     ME0123=. ;           ME123_0TYP1 =. ;
ME123_0PWR =. ;     ME1R12_0=R12;      ME1R13_0=R13;      ME1_0TYP1=. ;       ME1T1Z_0TYP1=. ;     ME1GZ_0TYP1=. ;
ME1T1GZ_0TYP1=. ;   ME1_0TYP1_10=. ;   ME1_0TYP1_20=. ;   ME1_0TYP1_30=. ;    ME1_0TYP1_40=. ;     ME1_0PWR=. ;
ME1_0PWR_10=. ;     ME1_0PWR_20=. ;    ME1_0PWR_30=. ;    ME1_0PWR_40=. ;     ME2R12_0=R12;        ME2R23_0=R23;
ME2_0TYP1=. ;       ME2T2Z_0TYP1=. ;   ME2GZ_0TYP1=. ;    ME2T2GZ_0TYP1=. ;   ME2_0TYP1_10=. ;     ME2_0TYP1_20=. ;
ME2_0TYP1_30=. ;    ME2_0TYP1_40=. ;   ME2_0PWR=. ;       ME2_0PWR_10=. ;     ME2_0PWR_20=. ;      ME2_0PWR_30=. ;
ME2_0PWR_40=. ;     ME3R13_0=R13;      ME3R23_0=R23;      ME3_0TYP1=. ;       ME3T3Z_0TYP1=. ;     ME3GZ_0TYP1=. ;
ME3T3GZ_0TYP1=. ;   ME3_0TYP1_10=. ;   ME3_0TYP1_20=. ;   ME3_0TYP1_30=. ;    ME3_0TYP1_40=. ;     ME3_0PWR=. ;
ME3_0PWR_10=. ;     ME3_0PWR_20=. ;    ME3_0PWR_30=. ;    ME3_0PWR_40=. ;     MEAR23T=R23;         T123_T=. ; G_T=. ;
ME123_T =. ;        ME123_TTYP1 =. ;   ME123_TPWR =. ;    ME1R12_T=R12;       ME1R13_T=R13;        ME1_TTYP1=. ;
ME1T1Z_TTYP1=. ;    ME1GZ_TTYP1=. ;    ME1T1GZ_TTYP1=. ;  ME1_TTYP1_10=. ;    ME1_TTYP1_20=. ;     ME1_TTYP1_30=. ;
ME1_TTYP1_40=. ;    ME1_TPWR=. ;       ME1_TPWR_10=. ;    ME1_TPWR_20=. ;     ME1_TPWR_30=. ;      ME1_TPWR_40=. ;
ME2R12_T=R12;       ME2R23_T=R23;      ME2_TTYP1=. ;      ME2T2Z_TTYP1=. ;    ME2GZ_TTYP1=. ;      ME2T2GZ_TTYP1=. ;
ME2_TTYP1_10=. ;    ME2_TTYP1_20=. ;   ME2_TTYP1_30=. ;   ME2_TTYP1_40=. ;    ME2_TPWR=. ;         ME2_TPWR_10=. ;
ME2_TPWR_20=. ;     ME2_TPWR_30=. ;    ME2_TPWR_40=. ;    ME3R13_T=R13;       ME3R23_T=R23;        ME3_TTYP1=. ;
ME3T3Z_TTYP1=. ;    ME3GZ_TTYP1=. ;    ME3T3GZ_TTYP1=.;   ME3_TTYP1_10=.;     ME3_TTYP1_20=. ;     ME3_TTYP1_30=. ;
ME3_TTYP1_40=. ;    ME3_TPWR=. ;       ME3_TPWR_10=. ;    ME3_TPWR_20=. ;     ME3_TPWR_30=. ;      ME3_TPWR_40=. ;
MEAR12B=R12;        MEAR13B=R13;       MEAR23B=R23;       T123_B=. ; G_B=. ;  ME123_B =. ;         ME123_BBias =. ;
ME123_BRMSE =. ;    ME123_BCOV =. ;    MEB123=. ;         ME123_BTYP1 =. ;    ME123_BPWR =. ;      ME1R12_B=R12;
ME1R13_B=R13;       ME1_BTYP1=. ;      ME1T1Z_BTYP1=. ;   ME1GZ_BTYP1=. ;     ME1T1GZ_BTYP1=. ;    ME1_BTYP1_10=. ;
ME1_BTYP1_20=. ;    ME1_BTYP1_30=. ;   ME1_BTYP1_40=. ;   ME1_BPWR=. ;        ME1_BPWR_10=. ;      ME1_BPWR_20=. ;
```

```
ME1_BPWR_30=. ;      ME1_BPWR_40=. ;      ME2R12_B=R12;      ME2R23_B=R23;    ME2_BTYP1=. ;        ME2T2Z_BTYP1=. ;
ME2GZ_BTYP1=. ;      ME2T2GZ_BTYP1=. ;  ME2_BTYP1_10=. ;  ME2_BTYP1_20=. ; ME2_BTYP1_30=. ;  ME2_BTYP1_40=. ;
ME2_BPWR=. ;         ME2_BPWR_10=. ;      ME2_BPWR_20=.      ME2_BPWR_30=. ;  ME2_BPWR_40=. ;  ME3R13_B=R13;
ME3R23_B=R23;        ME3_BTYP1=. ;        ME3T3Z_BTYP1=. ;  ME3GZ_BTYP1=. ;  ME3T3GZ_BTYP1=. ; ME3_BTYP1_10=. ;
ME3_BTYP1_20=. ;     ME3_BTYP1_30=. ;     ME3_BTYP1_40=. ;  ME3_BPWR=. ;       ME3_BPWR_10=. ;  ME3_BPWR_20=. ;
ME3_BPWR_30=. ;      ME3_BPWR_40=. ;      DE123_0 =. ;        DE123_0Bias =. ;    DE123_0RMSE =. ;  DE123_0COV =. ;
DE0123=. ;           DEAR120=. ;          DEAR130=. ;          DEAR230=. ;          DE123_0TYP1 =. ;    DE123_0PWR =. ;
DE1R12_0=R12;        DE1R13_0=R13;        DE1_0TYP1 =. ;       DE1_0TYP1_10 =. ;  DE1_0TYP1_20 =. ; DE1_0TYP1_30 =. ;
DE1_0TYP1_40 =. ;    DE1_0PWR =. ;        DE1_0PWR_10 =. ;  DE1_0PWR_20 =. ;  DE1_0PWR_30 =. ;  DE1_0PWR_40 =. ;
DE2R12_0=R12;        DE2R23_0=R23;        DE2_0TYP1 =. ;       DE2_0TYP1_10 =. ;  DE2_0TYP1_20 =. ; DE2_0TYP1_30 =. ;
DE2_0TYP1_40 =. ;    DE2_0PWR =. ;        DE2_0PWR_10 =. ;   DE2_0PWR_20 =. ;   DE2_0PWR_30 =. ; DE2_0PWR_40 =. ;
DE3R13_0=R13;        DE3R23_0=R23;        DE3_0TYP1 =. ;       DE3_0TYP1_10 =. ;  DE3_0TYP1_20 =. ; DE3_0TYP1_30 =. ;
DE3_0TYP1_40 =. ;    DE3_0PWR =. ;        DE3_0PWR_10 =. ;  DE3_0PWR_20 =. ;  DE3_0PWR_30 =.;  DE3_0PWR_40 =. ;
DE123_T =. ;         DE123_TBias =. ;     DE123_TRMSE =. ;  DE123_TCOV =. ;     DET123=. ;           DEAR12T=. ; DEAR13T=. ;
             DEAR23T=. ;              DE123_TTYP1 =. ;  DE123_TPWR =. ;     DE1R12_T=R12;     DE1R13_T=R13; DE1_TTYP1 =. ;
       DE1_TTYP1_10 =. ;     DE1_TTYP1_20 =. ;  DE1_TTYP1_30 =. ;  DE1_TTYP1_40 =.; DE1_TPWR =. ; DE1_TPWR_10 =. ;
       DE1_TPWR_20 =. ;      DE1_TPWR_30 =. ;  DE1_TPWR_40 =. ;   DE2R12_T=R12;     DE2R23_T=R23; DE2_TTYP1 =. ;
       DE2_TTYP1_10 =. ;     DE2_TTYP1_20 =. ;  DE2_TTYP1_30 =. ;  DE2_TTYP1_40 =.; DE2_TPWR =. ; DE2_TPWR_10 =. ;
       DE2_TPWR_20 =. ;      DE2_TPWR_30 =. ;  DE2_TPWR_40 =. ;   DE3R13_T=R13;     DE3R23_T=R23;         DE3_TTYP1 =. ;
       DE3_TTYP1_10 =. ;     DE3_TTYP1_20 =. ;  DE3_TTYP1_30 =. ;  DE3_TTYP1_40 =.; DE3_TPWR =. ;           DE3_TPWR_10 =.
;      DE3_TPWR_20 =. ;      DE3_TPWR_30 =. ;  DE3_TPWR_40 =. ;   DE123_B =. ;         DE123_BBias =. ; DE123_BRMSE =. ;
       DE123_BCOV =. ;       DEB123=. ;           DEAR12B=. ;           DEAR13B=. ;          DEAR23B=. ;
DE123_BTYP1 =. ;     DE123_BPWR =. ;      DE1R12_B=R12;      DE1R13_B=R13;      DE1_BTYP1 =. ;      DE1_BTYP1_10 =. ;
DE1_BTYP1_20 =. ;    DE1_BTYP1_30 =. ;    DE1_BTYP1_40 =.;  DE1_BPWR =. ;       DE1_BPWR_10 =.;  DE1_BPWR_20 =. ;
DE1_BPWR_30 =. ;     DE1_BPWR_40 =. ;     DE2R12_B=R12;      DE2R23_B=R23;      DE2_BTYP1 =. ;      DE2_BTYP1_10 =. ;
DE2_BTYP1_20 =. ;    DE2_BTYP1_30 =. ;    DE2_BTYP1_40 =. ;  DE2_BPWR =. ;       DE2_BPWR_10 =. ; DE2_BPWR_20 =. ;
DE2_BPWR_30 =. ;     DE2_BPWR_40 =. ;     DE3R13_B=R13;      DE3R23_B=R23;      DE3_BTYP1 =. ;      DE3_BTYP1_10 =. ;
DE3_BTYP1_20 =. ;    DE3_BTYP1_30 =. ;    DE3_BTYP1_40 =. ;  DE3_BPWR =. ;       DE3_BPWR_10 =. ; DE3_BPWR_20 =. ;
DE3_BPWR_30 =. ;     DE3_BPWR_40 =. ;
/* ADDED 78 REFRECE POINTS INTO THE TABLE*/
Format        T123_0            T123_T            T123_B            G_0               G_T
              G_B               DE123_0           DE123_T           DE123_B           ME123_0
              ME123_T           ME123_B           DE123_0Bias       DE123_TBias       DE123_BBias
              ME123_0Bias       ME123_TBias       ME123_BBias       DE123_0RMSE       DE123_TRMSE
              DE123_BRMSE       ME123_0RMSE       ME123_TRMSE       ME123_BRMSE       F7.3;
RUN; /* Accumulates 82 Columns of info in 12 variables */

DATA WQSEST3v;
 SET WQSEST3V (OBS=0);
 RUN;
```

```
%macro STATS_3V; /*Temp.WQS3v_STATS_162;*/

%do Cond=1 %to 162;
DATA WQSEST3v_500;/* 500 records for each Condition 1 to 162, each related to a Sampnum */
MERGE Temp.Cond162 (WHERE= (Condition=&Cond)) Temp.stdwqs3v_s&Cond;
                                        BY Condition;
RENAME
          STD_DE123_0   =DE123_0                              STD_DE123_0LCL=DE123_0LCL
          STD_DE123_0UCL=DE123_0UCL                           STD_DE123_T   =DE123_T
          STD_DE123_TLCL=DE123_TLCL                           STD_DE123_TUCL=DE123_TUCL
          STD_DE123_B   =DE123_B                              STD_DE123_BLCL=DE123_BLCL
          STD_DE123_BUCL=DE123_BUCL       `                   STD_G0 = G_0
          STD_GT = G_T                                        STD_GB = G_B
          STD_T123_0 = T123_0                                 STD_T123_T=T123_T
          STD_T123_B = T123_B                                 STD_ME123_0=ME123_0
          STD_P01_ME123_2_5=P01_ME123_2_5                     STD_ME123_T=ME123_T
          STD_P01_ME123_97_5=P01_ME123_97_5                   STD_ME123_B=ME123_B
          STD_PT1_ME123_2_5=PT1_ME123_2_5                     STD_PT1_ME123_97_5=PT1_ME123_97_5
          STD_PB1_ME123_2_5=PB1_ME123_2_5                     STD_PB1_ME123_97_5=PB1_ME123_97_5;
RUN; /* Has Total of 114 COLUMNS -12 PARMS -12 CI =30 + 9 VARS +9 weights variables in 500 rows  */

DATA    WQSEST3v_1;
SET     WQSEST3v_500 (WHERE= (Sampnum=1));
BY Condition;
                    T123 = Theta1 + Theta2 + Theta3;
                    ME0123=ME123;              MET123=ME123;              MEB123=ME123;
                    DE0123=DE123;              DET123=DE123;              DEB123=DE123;
                    DEAR120 = R12;             DEAR130 = R13;             DEAR230 = R23;
                    DE1R12_0 = R12;            DE1R13_0 = R13;            DE2R12_0= R12;
                    DE3R13_0= R13;             DE3R23_0= R23;             DE2R23_0= R23;
                    DEAR12T = R12;             DEAR13T = R13;             DEAR23T = R23;
                    DE1R12_T= R12;             DE1R13_T = R13;            DE2R12_T = R12;
                    DE3R13_T= R13;             DE3R23_T= R23;             DE2R23_T = R23;
                    DEAR12B = R12;             DEAR13B = R13;             DEAR23B = R23;
                    DE1R12_B= R12;             DE1R13_B = R13;            DE2R12_B = R12;
                    DE3R13_B= R13;             DE3R23_B= R23;             DE2R23_B = R23;
                    MEAR120 = R12;             MEAR130 = R13;             MEAR230 = R23;
                    ME1R12_0 = R12;            ME1R13_0 = R13;            ME2R12_0= R12;
                    ME3R13_0= R13;             ME3R23_0= R23;             ME2R23_0= R23;
                    MEAR12T = R12;             MEAR13T = R13;             MEAR23T = R23;
```

```
          ME1R12_T= R12;        ME1R13_T = R13;        ME2R12_T = R12;
          ME3R13_T= R13;        ME3R23_T= R23;         ME2R23_T = R23;
          MEAR12B = R12;        MEAR13B = R13;         MEAR23B = R23;
          ME1R12_B= R12;        ME1R13_B = R13;        ME2R12_B = R12;
          ME3R13_B= R13;        ME3R23_B= R23;         ME2R23_B = R23;
LABEL
          ME0123=ME123          MET123=ME123           MEB123=ME123
          DE0123=DE123          DET123=DE123           DEB123=DE123
          DEAR120 = R12         DEAR130 = R13          DEAR230 = R23
          DE1R12_0 = R12        DE1R13_0 = R13         DE2R12_0= R12
          DE3R23_0= R23         DE2R23_0= R23          DEAR12T = R12
          DEAR23T = R23         DE1R12_T= R12          DE1R13_T = R13
          DE3R13_T= R13         DE3R23_T= R23          DE2R23_T = R23
          DEAR12B = R12         DEAR13B = R13          DEAR23B = R23
          DE1R12_B= R12         DE1R13_B = R13         DE2R12_B = R12
          DE3R13_B= R13         DE3R23_B= R23          DE2R23_B = R23
          MEAR120 = R12         MEAR130 = R13          MEAR230 = R23
          ME1R12_0 = R12        ME1R13_0 = R13         ME2R12_0= R12
          ME3R13_0= R13         ME3R23_0= R23          ME2R23_0= R23
          MEAR12T = R12         MEAR13T = R13          MEAR23T = R23
          ME1R12_T= R12         ME1R13_T = R13         ME2R12_T= R12
          ME3R13_T= R13         ME3R23_T= R23          ME2R23_T= R23
          MEAR12B = R12         MEAR13B = R13          MEAR23B = R23
          ME1R12_B= R12         ME1R13_B = R13         ME2R12_B= R12
          ME3R13_B= R13         ME3R23_B= R23          ME2R23_B= R23;
          DE3R13_0= R13         DEAR13T = R13          DE2R12_T = R12
KEEP  Condition  N         wXB01     wXB02     wXB03     R12       R13       R23
      Theta1     Theta2    Theta3    ME123     T123      Gamma     ME1       ME2
      ME3        DE123     DE1       DE2       DE3       wXT1      wXT2      wXT3
      wXB1       wXB2      wXB3      cME123    cDE123    ME0123    MET123    MEB123
      DE0123     DET123    DEB123    DEAR120   DEAR130   DEAR230   DE1R12_0  DE1R13_0
      DE2R12_0   DE3R13_0  DE3R23_0  DE2R23_0  DEAR12T   DEAR13T   DEAR23T   DE1R12_T
      DE1R13_T   DE2R12_T  DE3R13_T  DE3R23_T  DE2R23_T  DEAR12B   DEAR13B   DEAR23B
      DE1R12_B   DE1R13_B  DE2R12_B  DE3R13_B  DE3R23_B  DE2R23_B  MEAR120   MEAR130
      MEAR230    ME1R12_0  ME1R13_0  ME2R12_0  ME3R13_0  ME3R23_0  ME2R23_0  MEAR12T
      MEAR13T    MEAR23T   ME1R12_T  ME1R13_T  ME2R12_T  ME3R13_T  ME3R23_T  ME2R23_T
      MEAR12B    MEAR13B   MEAR23B   ME1R12_B  ME1R13_B  ME2R12_B  ME3R13_B  ME3R23_B
      ME2R23_B;
RUN; /*89 = 20+12 +24 (12 Variables EST)*/
```

```
PROC UNIVARIATE DATA=WQSEST3v_500; /*114 VARS in 500 records per Condition by Sampnum 24 EST + 48 CL|PCL */
VAR
        DE123_0         DE123_T         DE123_B         T123_0      T123_T      T123_B
        G_0             G_T             G_B             ME123_0     ME123_T     ME123_B;
BY Condition;

OUTPUT OUT= WQS3v_AVGEST            MEAN=

    AVG_DE123_0     AVG_DE123_T     AVG_DE123_B     AVG_T123_0
    AVG_T123_T      AVG_T123_B      AVG_G_0         AVG_G_T
    AVG_G_B         AVG_ME123_0     AVG_ME123_T     AVG_ME123_B;

        RUN;  /* 13 Columns with 12 AVGs Single record for given Condition for Bias Calculations */

DATA      WQSEST3v_1; /* ADD THE MEAN value of the 500 replicates for each of the 12 EST */
MERGE   WQSEST3v_1  WQS3v_AVGEST;
          BY Condition; /* Has 105 variables */

        DE123_0=AVG_DE123_0;     DE123_T=AVG_DE123_T;     DE123_B=AVG_DE123_B;
        T123_0=AVG_T123_0;       T123_T=AVG_T123_T;       T123_B=AVG_T123_B;
        G_0 =AVG_G_0;            G_T =AVG_G_T;            G_B =AVG_G_B;
        ME123_0=AVG_ME123_0;     ME123_T=AVG_ME123_T;     ME123_B=AVG_ME123_B;

FORMAT      DE123_0     DE123_T     DE123_B         T123_0          T123_T          T123_B
            G_0         G_T         G_B             ME123_0         ME123_T         ME123_B F7.3;

DROP        AVG_DE123_0     AVG_DE123_T     AVG_DE123_B     AVG_T123_0      AVG_T123_T
            AVG_T123_B      AVG_G_0         AVG_G_T         AVG_G_B         AVG_ME123_0
            AVG_ME123_T     AVG_ME123_B;
RUN;  /* Has 144 VARS 105 + 39 EST AVG from 500 Sampnums into one record */

DATA WQS3v_BiasSQE;
            MERGE WQSEST3v_500 WQS3v_AVGEST; /* Has 105 Columns 500 Replicated sample DATA */
            BY Condition;

                DE123_0Bias   =         AVG_DE123_0 - DE123;
                DE123_TBias   =         AVG_DE123_T - DE123;
                DE123_BBias   =         AVG_DE123_B - DE123;
                ME123_0Bias   =         AVG_ME123_0 - ME123;
                ME123_TBias   =         AVG_ME123_T - ME123;
```

```
                ME123_BBias   =            AVG_ME123_B - ME123;
                DE123_0SQE    =            (DE123_0 - DE123)**2;
                DE123_TSQE    =            (DE123_T- DE123)**2;
                DE123_BSQE    =            (DE123_B - DE123)**2;
                ME123_0SQE    =            (ME123_0 - ME123)**2;
                ME123_TSQE    =            (ME123_T- ME123)**2;
                ME123_BSQE    =            (ME123_B- ME123)**2;

RUN; /* 117 variables with 105+12 in 500 rows for EST CL|PCL Bias and SQE */

PROC UNIVARIATE DATA=WQS3v_BiasSQE;  /* 500 records for each given Condition */

VAR     DE123_0Bias         DE123_TBias         DE123_BBias
        ME123_0Bias         ME123_TBias         ME123_BBias
        DE123_0SQE          DE123_TSQE          DE123_BSQE
        ME123_0SQE          ME123_TSQE          ME123_BSQE    ;
  BY Condition;
OUTPUT OUT=WQS3v_AVGMS          (KEEP= Condition
    DE123_0Bias                 DE123_TBias         DE123_BBias
    ME123_0Bias                 ME123_TBias         ME123_BBias
    DE123_0MSE                  DE123_TMSE          DE123_BMSE
    ME123_0MSE                  ME123_TMSE          ME123_BMSE)
MEAN=
    DE123_0Bias                 DE123_TBias         DE123_BBias
    ME123_0Bias                 ME123_TBias         ME123_BBias
    DE123_0MSE                  DE123_TMSE          DE123_BMSE
    ME123_0MSE                  ME123_TMSE          ME123_BMSE;
RUN;    /* 12 +1 variables for 6 EST 1 records for each given Condition */

DATA WQS3v_BiasRMSE;   /* Will END up with 49 Columns in 1 record for each given Condition */
        SET WQS3v_AVGMS;    /* 49 Columns in 1 record for each given Condition */
        DE123_0RMSE=SQRT (DE123_0MSE); DE123_TRMSE=SQRT (DE123_TMSE); DE123_BRMSE=SQRT (DE123_BMSE);
        ME123_0RMSE=SQRT (ME123_0MSE); ME123_TRMSE=SQRT(ME123_TMSE); ME123_BRMSE=SQRT (ME123_BMSE);

LABEL
         DE123_0Bias=DE123_0Bias    DE123_TBias=DE123_TBias    DE123_BBias=DE123_BBias
        ME123_0Bias=ME123_0Bias    ME123_TBias=ME123_TBias    ME123_BBias=ME123_BBias;

DROP
        DE123_0MSE   DE123_TMSE   DE123_BMSE   ME123_0MSE   ME123_TMSE   ME123_BMSE;
```

```
FORMAT _ALL_ F7.3; FORMAT Condition F3.0;
RUN; /* 49 Columns in 1 record for each given Condition */

/* UPDATE the single record file WQSEST3v_1*/

DATA WQSEST3v_1;
MERGE   WQSEST3v_1 WQS3v_BiasRMSE; /* ADD 48 Bias and RMSE for 24 VARs */
BY Condition;
RUN;   /* has 144 +48 =192 columns */

/* has one record of 144 from WSEST3V_1 + 48 (Bias and RMSE for 24 EST) = 192 variables */
/* WQSEST3V_500 has 105 Columns with 30 VARS + 48 CIs ++18 + 9 Weights in 105 COLUMNS and 500 rows */
/*******Type 1 Error & Power for 0.20 0.30 0.40 and 0.50 cutoff for individual WQSestimates *********/
/* First do Joint Type1 error and Power for ME123_0 ME123_t and ME123_B then DE123_0 DE123_t and DE123_B */

/**********SEQ 3V ME123 ***************/
DATA ME123_0TYP1 (KEEP=Condition Sampnum ME123_0TYP1);
IF _N_=1 THEN ME123_0TYP1=0;
SET WQSEST3v_500;
IF ((ME123 =0) & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME123_0TYP1=ME123_0TYP1+1/500;
RETAIN ME123_0TYP1;
RUN;

DATA ME123_0PWR (KEEP=Condition Sampnum ME123_0PWR);
IF _N_=1 THEN ME123_0PWR=0;
SET WQSEST3v_500;
IF ((ME123 ^=0) & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME123_0PWR=ME123_0PWR+1/500;
RETAIN ME123_0PWR;
RUN;

DATA ME123_TTYP1 (KEEP=Condition Sampnum ME123_TTYP1);
IF _N_=1 THEN ME123_TTYP1=0;
SET WQSEST3v_500;
IF ((ME123 =0) & ^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME123_TTYP1=ME123_TTYP1+1/500;
RETAIN ME123_TTYP1;
RUN;
```

```
DATA ME123_TPWR (KEEP=Condition Sampnum ME123_TPWR);
IF _N_=1 THEN ME123_TPWR=0;
SET WQSEST3v_500;
IF ((ME123 ^=0) & ^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME123_TPWR=ME123_TPWR+1/500;
RETAIN ME123_TPWR;
RUN;

DATA ME123_BTYP1 (KEEP=Condition Sampnum ME123_BTYP1);
IF _N_=1 THEN ME123_BTYP1=0;
SET WQSEST3v_500;
IF ((ME123 =0) & ^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME123_BTYP1=ME123_BTYP1+1/500;
RETAIN ME123_BTYP1;
RUN;

DATA ME123_BPWR (KEEP=Condition Sampnum ME123_BPWR);
IF _N_=1 THEN ME123_BPWR=0;
SET WQSEST3v_500;
IF ((ME123 ^=0) & ^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME123_BPWR=ME123_BPWR+1/500;
RETAIN ME123_BPWR;
RUN;

/**********SEQ 3V    DE123 ***************/
DATA DE123_0TYP1 (KEEP=Condition Sampnum DE123_0TYP1);
IF _N_=1 THEN DE123_0TYP1=0;
SET WQSEST3v_500;
IF ((DE123 =0) & ^(DE123_0LCL<=0<=DE123_0UCL)) THEN
DE123_0TYP1=DE123_0TYP1+1/500;
RETAIN DE123_0TYP1;
RUN;

DATA DE123_0PWR (KEEP=Condition Sampnum  DE123_0PWR);
IF _N_=1 THEN DE123_0PWR=0;
SET WQSEST3v_500;
IF ((DE123 ^=0) & ^(DE123_0LCL<=0<=DE123_0UCL)) THEN
DE123_0PWR=DE123_0PWR+1/500;
RETAIN DE123_0PWR;
RUN;
```

```
DATA DE123_TTYP1 (KEEP=Condition Sampnum DE123_TTYP1);
IF _N_=1 THEN DE123_TTYP1=0;
SET WQSEST3v_500;
IF ((DE123 =0) & ^(DE123_TLCL<=0<=DE123_TUCL)) THEN
DE123_TTYP1=DE123_TTYP1+1/500;
RETAIN DE123_TTYP1;
RUN;

DATA DE123_TPWR (KEEP=Condition Sampnum DE123_TPWR);
IF _N_=1 THEN DE123_TPWR=0;
SET WQSEST3v_500;
IF ((DE123 ^=0) & ^(DE123_TLCL<=0<=DE123_TUCL)) THEN
DE123_TPWR=DE123_TPWR+1/500;
RETAIN DE123_TPWR;
RUN;

DATA DE123_BTYP1 (KEEP=Condition Sampnum DE123_BTYP1);
IF _N_=1 THEN DE123_BTYP1=0;
SET WQSEST3v_500;
IF ((DE123 =0) & ^(DE123_BLCL<=0<=DE123_BUCL)) THEN
DE123_BTYP1=DE123_BTYP1+1/500;
RETAIN DE123_BTYP1;
RUN;

DATA DE123_BPWR (KEEP=Condition Sampnum DE123_BPWR);
IF _N_=1 THEN DE123_BPWR=0;
SET WQSEST3v_500;
IF ((DE123 ^=0) & ^(DE123_BLCL<=0<=DE123_BUCL)) THEN
DE123_BPWR=DE123_BPWR+1/500;
RETAIN DE123_BPWR;
RUN;

/*Individual Estimates TYPE1 ERR and POWER for 0.10, 0.20, 0.30 , 0.40 and 0.50 cutoff for WQS weigts */
/**************DE1_0***********/
DATA DE1_0TYP1 (KEEP=Condition Sampnum DE1_0TYP1);
IF _N_=1 THEN DE1_0TYP1=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB01>0 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE1_0TYP1=DE1_0TYP1+1/500;
```

```
RETAIN DE1_0TYP1;
RUN;

DATA DE1_0TYP1_10 (KEEP=Condition Sampnum DE1_0TYP1_10);
IF _N_=1 THEN DE1_0TYP1_10=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB01 >0.10 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE1_0TYP1_10=DE1_0TYP1_10+1/500;
RETAIN DE1_0TYP1_10;
RUN;

DATA DE1_0TYP1_20 (KEEP=Condition Sampnum DE1_0TYP1_20);
IF _N_=1 THEN DE1_0TYP1_20=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB01 >0.20 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE1_0TYP1_20=DE1_0TYP1_20+1/500;
RETAIN DE1_0TYP1_20;
RUN;

DATA DE1_0TYP1_30 (KEEP=Condition Sampnum DE1_0TYP1_30);
IF _N_=1 THEN DE1_0TYP1_30=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB01 >0.30 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE1_0TYP1_30=DE1_0TYP1_30+1/500;
RETAIN DE1_0TYP1_30;
RUN;

DATA DE1_0TYP1_40 (KEEP=Condition Sampnum DE1_0TYP1_40);
IF _N_=1 THEN DE1_0TYP1_40=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB01 >0.40 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE1_0TYP1_40=DE1_0TYP1_40+1/500;
RETAIN DE1_0TYP1_40;
RUN;

DATA DE1_0PWR (KEEP=Condition Sampnum DE1_0PWR);
IF _N_=1 THEN DE1_0PWR=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB01>0 ) THEN
DE1_0PWR=DE1_0PWR+1/500;
```

```
RETAIN DE1_0PWR;
RUN;

DATA DE1_0PWR_10 (KEEP=Condition Sampnum DE1_0PWR_10);
IF _N_=1 THEN DE1_0PWR_10=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB01 >0.10) THEN
DE1_0PWR_10=DE1_0PWR_10+1/500;
RETAIN DE1_0PWR_10;
RUN;

DATA DE1_0PWR_20 (KEEP=Condition Sampnum DE1_0PWR_20);
IF _N_=1 THEN DE1_0PWR_20=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB01 >0.20) THEN
DE1_0PWR_20=DE1_0PWR_20+1/500;
RETAIN DE1_0PWR_20;
RUN;

DATA DE1_0PWR_30 (KEEP=Condition Sampnum DE1_0PWR_30);
IF _N_=1 THEN DE1_0PWR_30=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB01 >0.30) THEN
DE1_0PWR_30=DE1_0PWR_30+1/500;
RETAIN DE1_0PWR_30;
RUN;


DATA DE1_0PWR_40 (KEEP=Condition Sampnum DE1_0PWR_40);
IF _N_=1 THEN DE1_0PWR_40=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB01 >0.40) THEN
DE1_0PWR_40=DE1_0PWR_40+1/500;
RETAIN DE1_0PWR_40;
RUN;
/**************DE2_0***********/
DATA DE2_0TYP1 (KEEP=Condition Sampnum DE2_0TYP1);
IF _N_=1 THEN DE2_0TYP1=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB02>0 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
```

```
DE2_0TYP1=DE2_0TYP1+1/500;
RETAIN DE2_0TYP1;
RUN;

DATA DE2_0TYP1_10 (KEEP=Condition Sampnum DE2_0TYP1_10);
IF _N_=1 THEN DE2_0TYP1_10=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB02 >0.10 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE2_0TYP1_10=DE2_0TYP1_10+1/500;
RETAIN DE2_0TYP1_10;
RUN;

DATA DE2_0TYP1_20 (KEEP=Condition Sampnum DE2_0TYP1_20);
IF _N_=1 THEN DE2_0TYP1_20=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB02 >0.20 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE2_0TYP1_20=DE2_0TYP1_20+1/500;
RETAIN DE2_0TYP1_20;
RUN;

DATA DE2_0TYP1_30 (KEEP=Condition Sampnum DE2_0TYP1_30);
IF _N_=1 THEN DE2_0TYP1_30=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB02 >0.30 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE2_0TYP1_30=DE2_0TYP1_30+1/500;
RETAIN DE2_0TYP1_30;
RUN;

DATA DE2_0TYP1_40 (KEEP=Condition Sampnum DE2_0TYP1_40);
IF _N_=1 THEN DE2_0TYP1_40=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB02 >0.40 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE2_0TYP1_40=DE2_0TYP1_40+1/500;
RETAIN DE2_0TYP1_40;
RUN;

DATA DE2_0PWR (KEEP=Condition Sampnum DE2_0PWR);
IF _N_=1 THEN DE2_0PWR=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB02>0) THEN
```

```
DE2_0PWR=DE2_0PWR+1/500;
RETAIN DE2_0PWR;
RUN;

DATA DE2_0PWR_10 (KEEP=Condition Sampnum DE2_0PWR_10);
IF _N_=1 THEN DE2_0PWR_10=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB02 >0.10) THEN
DE2_0PWR_10=DE2_0PWR_10+1/500;
RETAIN DE2_0PWR_10;
RUN;

DATA DE2_0PWR_20 (KEEP=Condition Sampnum DE2_0PWR_20);
IF _N_=1 THEN DE2_0PWR_20=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB02 >0.20) THEN
DE2_0PWR_20=DE2_0PWR_20+1/500;
RETAIN DE2_0PWR_20;
RUN;

DATA DE2_0PWR_30 (KEEP=Condition Sampnum DE2_0PWR_30);
IF _N_=1 THEN DE2_0PWR_30=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB02 >0.30) THEN
DE2_0PWR_30=DE2_0PWR_30+1/500;
RETAIN DE2_0PWR_30;
RUN;


DATA DE2_0PWR_40 (KEEP=Condition Sampnum DE2_0PWR_40);
IF _N_=1 THEN DE2_0PWR_40=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB02 >0.40) THEN
DE2_0PWR_40=DE2_0PWR_40+1/500;
RETAIN DE2_0PWR_40;
RUN;

/*************DE3_0***********/
DATA DE3_0TYP1 (KEEP=Condition Sampnum DE3_0TYP1);
IF _N_=1 THEN DE3_0TYP1=0;
```

```
SET WQSEST3v_500;
IF (DE3 =0 & wXB03>0 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE3_0TYP1=DE3_0TYP1+1/500;
RETAIN DE3_0TYP1;
RUN;

DATA DE3_0TYP1_10 (KEEP=Condition Sampnum DE3_0TYP1_10);
IF _N_=1 THEN DE3_0TYP1_10=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB03 >0.10 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE3_0TYP1_10=DE3_0TYP1_10+1/500;
RETAIN DE3_0TYP1_10;
RUN;

DATA DE3_0TYP1_20 (KEEP=Condition Sampnum DE3_0TYP1_20);
IF _N_=1 THEN DE3_0TYP1_20=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB03 >0.20 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE3_0TYP1_20=DE3_0TYP1_20+1/500;
RETAIN DE3_0TYP1_20;
RUN;

DATA DE3_0TYP1_30 (KEEP=Condition Sampnum DE3_0TYP1_30);
IF _N_=1 THEN DE3_0TYP1_30=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB03 >0.30 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL))  THEN
DE3_0TYP1_30=DE3_0TYP1_30+1/500;
RETAIN DE3_0TYP1_30;
RUN;
DATA DE3_0TYP1_40 (KEEP=Condition Sampnum DE3_0TYP1_40);
IF _N_=1 THEN DE3_0TYP1_40=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB03 >0.40 & DE123=0 & ^(DE123_0LCL<0<DE123_0UCL)) THEN
DE3_0TYP1_40=DE3_0TYP1_40+1/500;
RETAIN DE3_0TYP1_40;
RUN;

DATA DE3_0PWR (KEEP=Condition Sampnum DE3_0PWR);
IF _N_=1 THEN DE3_0PWR=0;
SET WQSEST3v_500;
```

```
IF (DE3 ^=0 & wXB03>0) THEN
DE3_0PWR=DE3_0PWR+1/500;
RETAIN DE3_0PWR;
RUN;

DATA DE3_0PWR_10 (KEEP=Condition Sampnum DE3_0PWR_10);
IF _N_=1 THEN DE3_0PWR_10=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB03 >0.10) THEN
DE3_0PWR_10=DE3_0PWR_10+1/500;
RETAIN DE3_0PWR_10;
RUN;

DATA DE3_0PWR_20 (KEEP=Condition Sampnum DE3_0PWR_20);
IF _N_=1 THEN DE3_0PWR_20=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB03 >0.20) THEN
DE3_0PWR_20=DE3_0PWR_20+1/500;
RETAIN DE3_0PWR_20;
RUN;

DATA DE3_0PWR_30 (KEEP=Condition Sampnum DE3_0PWR_30);
IF _N_=1 THEN DE3_0PWR_30=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB03 >0.30) THEN
DE3_0PWR_30=DE3_0PWR_30+1/500;
RETAIN DE3_0PWR_30;
RUN;

DATA DE3_0PWR_40 (KEEP=Condition Sampnum DE3_0PWR_40);
IF _N_=1 THEN DE3_0PWR_40=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB03 >0.40) THEN
DE3_0PWR_40=DE3_0PWR_40+1/500;
RETAIN DE3_0PWR_40;
RUN;

/**************DE1_T***********/
DATA DE1_TTYP1 (KEEP=Condition Sampnum DE1_TTYP1);
IF _N_=1 THEN DE1_TTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE1 =0 & wXT1>0 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE1_TTYP1=DE1_TTYP1+1/500;
RETAIN DE1_TTYP1;
RUN;

DATA DE1_TTYP1_10 (KEEP=Condition Sampnum DE1_TTYP1_10);
IF _N_=1 THEN DE1_TTYP1_10=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXT1 >0.10 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE1_TTYP1_10=DE1_TTYP1_10+1/500;
RETAIN DE1_TTYP1_10;
RUN;

DATA DE1_TTYP1_20 (KEEP=Condition Sampnum DE1_TTYP1_20);
IF _N_=1 THEN DE1_TTYP1_20=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXT1 >0.20 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE1_TTYP1_20=DE1_TTYP1_20+1/500;
RETAIN DE1_TTYP1_20;
RUN;

DATA DE1_TTYP1_30 (KEEP=Condition Sampnum DE1_TTYP1_30);
IF _N_=1 THEN DE1_TTYP1_30=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXT1 >0.30 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE1_TTYP1_30=DE1_TTYP1_30+1/500;
RETAIN DE1_TTYP1_30;
RUN;
DATA DE1_TTYP1_40 (KEEP=Condition Sampnum DE1_TTYP1_40);
IF _N_=1 THEN DE1_TTYP1_40=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXT1 >0.40 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE1_TTYP1_40=DE1_TTYP1_40+1/500;
RETAIN DE1_TTYP1_40;
RUN;

DATA DE1_TPW(KEEP=Condition Sampnum DE1_TPWR);
IF _N_=1 THEN DE1_TPWR=0;
SET WQSEST3v_500;
```

```
IF (DE1 ^=0 & wXT1>0) THEN
DE1_TPWR=DE1_TPWR+1/500;
RETAIN DE1_TPWR;
RUN;

DATA DE1_TPWR_10 (KEEP=Condition Sampnum DE1_TPWR_10);
IF _N_=1 THEN DE1_TPWR_10=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXT1 >0.10) THEN
DE1_TPWR_10=DE1_TPWR_10+1/500;
RETAIN DE1_TPWR_10;
RUN;

DATA DE1_TPWR_20 (KEEP=Condition Sampnum DE1_TPWR_20);
IF _N_=1 THEN DE1_TPWR_20=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXT1 >0.20) THEN
DE1_TPWR_20=DE1_TPWR_20+1/500;
RETAIN DE1_TPWR_20;
RUN;

DATA DE1_TPWR_30 (KEEP=Condition Sampnum DE1_TPWR_30);
IF _N_=1 THEN DE1_TPWR_30=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXT1 >0.30) THEN
DE1_TPWR_30=DE1_TPWR_30+1/500;
RETAIN DE1_TPWR_30;
RUN;

DATA DE1_TPWR_40 (KEEP=Condition Sampnum DE1_TPWR_40);
IF _N_=1 THEN DE1_TPWR_40=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXT1 >0.40) THEN
DE1_TPWR_40=DE1_TPWR_40+1/500;
RETAIN DE1_TPWR_40;
RUN;

/**************DE2_T***********/
DATA DE2_TTYP1 (KEEP=Condition Sampnum DE2_TTYP1);
IF _N_=1 THEN DE2_TTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE2 =0 & wXT2>0 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE2_TTYP1=DE2_TTYP1+1/500;
RETAIN DE2_TTYP1;
RUN;

DATA DE2_TTYP1_10 (KEEP=Condition Sampnum DE2_TTYP1_10);
IF _N_=1 THEN DE2_TTYP1_10=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXT2 >0.10 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE2_TTYP1_10=DE2_TTYP1_10+1/500;
RETAIN DE2_TTYP1_10;
RUN;

DATA DE2_TTYP1_20 (KEEP=Condition Sampnum DE2_TTYP1_20);
IF _N_=1 THEN DE2_TTYP1_20=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXT2 >0.20 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE2_TTYP1_20=DE2_TTYP1_20+1/500;
RETAIN DE2_TTYP1_20;
RUN;

DATA DE2_TTYP1_30 (KEEP=Condition Sampnum DE2_TTYP1_30);
IF _N_=1 THEN DE2_TTYP1_30=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXT2 >0.30 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE2_TTYP1_30=DE2_TTYP1_30+1/500;
RETAIN DE2_TTYP1_30;
RUN;
DATA DE2_TTYP1_40 (KEEP=Condition Sampnum DE2_TTYP1_40);
IF _N_=1 THEN DE2_TTYP1_40=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXT2 >0.40 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL))  THEN
DE2_TTYP1_40=DE2_TTYP1_40+1/500;
RETAIN DE2_TTYP1_40;
RUN;

DATA DE2_TPWR (KEEP=Condition Sampnum DE2_TPWR);
IF _N_=1 THEN DE2_TPWR=0;
SET WQSEST3v_500;
```

```
IF (DE2 ^=0 & wXT2>0) THEN
DE2_TPWR=DE2_TPWR+1/500;
RETAIN DE2_TPWR;
RUN;

DATA DE2_TPWR_10 (KEEP=Condition Sampnum DE2_TPWR_10);
IF _N_=1 THEN DE2_TPWR_10=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXT2 >0.10) THEN
DE2_TPWR_10=DE2_TPWR_10+1/500;
RETAIN DE2_TPWR_10;
RUN;

DATA DE2_TPWR_20 (KEEP=Condition Sampnum DE2_TPWR_20);
IF _N_=1 THEN DE2_TPWR_20=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXT2 >0.20) THEN
DE2_TPWR_20=DE2_TPWR_20+1/500;
RETAIN DE2_TPWR_20;
RUN;

DATA DE2_TPWR_30 (KEEP=Condition Sampnum DE2_TPWR_30);
IF _N_=1 THEN DE2_TPWR_30=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXT2 >0.30) THEN
DE2_TPWR_30=DE2_TPWR_30+1/500;
RETAIN DE2_TPWR_30;
RUN;

DATA DE2_TPWR_40 (KEEP=Condition Sampnum DE2_TPWR_40);
IF _N_=1 THEN DE2_TPWR_40=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXT2 >0.40) THEN
DE2_TPWR_40=DE2_TPWR_40+1/500;
RETAIN DE2_TPWR_40;
RUN;

/**************DE3_T***********/
DATA DE3_TTYP1 (KEEP=Condition Sampnum DE3_TTYP1);
IF _N_=1 THEN DE3_TTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE3 =0 & wXT3>0 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE3_TTYP1=DE3_TTYP1+1/500;
RETAIN DE3_TTYP1;
RUN;

DATA DE3_TTYP1_10 (KEEP=Condition Sampnum DE3_TTYP1_10);
IF _N_=1 THEN DE3_TTYP1_10=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXT3 >0.10 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE3_TTYP1_10=DE3_TTYP1_10+1/500;
RETAIN DE3_TTYP1_10;
RUN;

DATA DE3_TTYP1_20 (KEEP=Condition Sampnum DE3_TTYP1_20);
IF _N_=1 THEN DE3_TTYP1_20=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXT3 >0.20 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE3_TTYP1_20=DE3_TTYP1_20+1/500;
RETAIN DE3_TTYP1_20;
RUN;

DATA DE3_TTYP1_30 (KEEP=Condition Sampnum DE3_TTYP1_30);
IF _N_=1 THEN DE3_TTYP1_30=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXT3 >0.30 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE3_TTYP1_30=DE3_TTYP1_30+1/500;
RETAIN DE3_TTYP1_30;
RUN;
DATA DE3_TTYP1_40 (KEEP=Condition Sampnum DE3_TTYP1_40);
IF _N_=1 THEN DE3_TTYP1_40=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXT3 >0.40 & DE123=0 & ^(DE123_TLCL<0<DE123_TUCL)) THEN
DE3_TTYP1_40=DE3_TTYP1_40+1/500;
RETAIN DE3_TTYP1_40;
RUN;

DATA DE3_TPWR (KEEP=Condition Sampnum DE3_TPWR);
IF _N_=1 THEN DE3_TPWR=0;
SET WQSEST3v_500;
```

```
IF (DE3 ^=0 & wXT3>0) THEN
DE3_TPWR=DE3_TPWR+1/500;
RETAIN DE3_TPWR;
RUN;

DATA DE3_TPWR_10 (KEEP=Condition Sampnum DE3_TPWR_10);
IF _N_=1 THEN DE3_TPWR_10=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXT3 >0.10) THEN
DE3_TPWR_10=DE3_TPWR_10+1/500;
RETAIN DE3_TPWR_10;
RUN;

DATA DE3_TPWR_20 (KEEP=Condition Sampnum DE3_TPWR_20);
IF _N_=1 THEN DE3_TPWR_20=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXT3 >0.20) THEN
DE3_TPWR_20=DE3_TPWR_20+1/500;
RETAIN DE3_TPWR_20;
RUN;

DATA DE3_TPWR_30 (KEEP=Condition Sampnum DE3_TPWR_30);
IF _N_=1 THEN DE3_TPWR_30=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXT3 >0.30) THEN
DE3_TPWR_30=DE3_TPWR_30+1/500;
RETAIN DE3_TPWR_30;
RUN;

DATA DE3_TPWR_40 (KEEP=Condition Sampnum DE3_TPWR_40);
IF _N_=1 THEN DE3_TPWR_40=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXT3 >0.40) THEN
DE3_TPWR_40=DE3_TPWR_40+1/500;
RETAIN DE3_TPWR_40;
RUN;

/**************DE1_B***********/
DATA DE1_BTYP1 (KEEP=Condition Sampnum DE1_BTYP1);
IF _N_=1 THEN DE1_BTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE1 =0 & wXB1>0 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE1_BTYP1=DE1_BTYP1+1/500;
RETAIN DE1_BTYP1;
RUN;

DATA DE1_BTYP1_10 (KEEP=Condition Sampnum DE1_BTYP1_10);
IF _N_=1 THEN DE1_BTYP1_10=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB1 >0.10 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE1_BTYP1_10=DE1_BTYP1_10+1/500;
RETAIN DE1_BTYP1_10;
RUN;

DATA DE1_BTYP1_20 (KEEP=Condition Sampnum DE1_BTYP1_20);
IF _N_=1 THEN DE1_BTYP1_20=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB1 >0.20 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE1_BTYP1_20=DE1_BTYP1_20+1/500;
RETAIN DE1_BTYP1_20;
RUN;

DATA DE1_BTYP1_30 (KEEP=Condition Sampnum DE1_BTYP1_30);
IF _N_=1 THEN DE1_BTYP1_30=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB1 >0.30 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE1_BTYP1_30=DE1_BTYP1_30+1/500;
RETAIN DE1_BTYP1_30;
RUN;
DATA DE1_BTYP1_40 (KEEP=Condition Sampnum DE1_BTYP1_40);
IF _N_=1 THEN DE1_BTYP1_40=0;
SET WQSEST3v_500;
IF (DE1 =0 & wXB1 >0.40 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE1_BTYP1_40=DE1_BTYP1_40+1/500;
RETAIN DE1_BTYP1_40;
RUN;

DATA DE1_BPWR (KEEP=Condition Sampnum DE1_BPWR);
IF _N_=1 THEN DE1_BPWR=0;
SET WQSEST3v_500;
```

```
IF (DE1 ^=0 & wXB1>0) THEN
DE1_BPWR=DE1_BPWR+1/500;
RETAIN DE1_BPWR;
RUN;

DATA DE1_BPWR_10 (KEEP=Condition Sampnum DE1_BPWR_10);
IF _N_=1 THEN DE1_BPWR_10=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB1 >0.10) THEN
DE1_BPWR_10=DE1_BPWR_10+1/500;
RETAIN DE1_BPWR_10;
RUN;

DATA DE1_BPWR_20 (KEEP=Condition Sampnum DE1_BPWR_20);
IF _N_=1 THEN DE1_BPWR_20=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB1 >0.20) THEN
DE1_BPWR_20=DE1_BPWR_20+1/500;
RETAIN DE1_BPWR_20;
RUN;

DATA DE1_BPWR_30 (KEEP=Condition Sampnum DE1_BPWR_30);
IF _N_=1 THEN DE1_BPWR_30=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB1 >0.30) THEN
DE1_BPWR_30=DE1_BPWR_30+1/500;
RETAIN DE1_BPWR_30;
RUN;

DATA DE1_BPWR_40 (KEEP=Condition Sampnum DE1_BPWR_40);
IF _N_=1 THEN DE1_BPWR_40=0;
SET WQSEST3v_500;
IF (DE1 ^=0 & wXB1 >0.40) THEN
DE1_BPWR_40=DE1_BPWR_40+1/500;
RETAIN DE1_BPWR_40;
RUN;

/**************DE2_B***********/
DATA DE2_BTYP1 (KEEP=Condition Sampnum DE2_BTYP1);
IF _N_=1 THEN DE2_BTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE2 =0 & wXB2>0 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE2_BTYP1=DE2_BTYP1+1/500;
RETAIN DE2_BTYP1;
RUN;

DATA DE2_BTYP1_10 (KEEP=Condition Sampnum DE2_BTYP1_10);
IF _N_=1 THEN DE2_BTYP1_10=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB2 >0.10 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL))  THEN
DE2_BTYP1_10=DE2_BTYP1_10+1/500;
RETAIN DE2_BTYP1_10;
RUN;

DATA DE2_BTYP1_20 (KEEP=Condition Sampnum DE2_BTYP1_20);
IF _N_=1 THEN DE2_BTYP1_20=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB2 >0.20 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE2_BTYP1_20=DE2_BTYP1_20+1/500;
RETAIN DE2_BTYP1_20;
RUN;

DATA DE2_BTYP1_30 (KEEP=Condition Sampnum DE2_BTYP1_30);
IF _N_=1 THEN DE2_BTYP1_30=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB2 >0.30 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE2_BTYP1_30=DE2_BTYP1_30+1/500;
RETAIN DE2_BTYP1_30;
RUN;
DATA DE2_BTYP1_40 (KEEP=Condition Sampnum DE2_BTYP1_40);
IF _N_=1 THEN DE2_BTYP1_40=0;
SET WQSEST3v_500;
IF (DE2 =0 & wXB2 >0.40 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL))  THEN
DE2_BTYP1_40=DE2_BTYP1_40+1/500;
RETAIN DE2_BTYP1_40;
RUN;

DATA DE2_BPWR (KEEP=Condition Sampnum DE2_BPWR);
IF _N_=1 THEN DE2_BPWR=0;
SET WQSEST3v_500;
```

```
IF (DE2 ^=0 & wXB2>0) THEN
DE2_BPWR=DE2_BPWR+1/500;
RETAIN DE2_BPWR;
RUN;

DATA DE2_BPWR_10 (KEEP=Condition Sampnum DE2_BPWR_10);
IF _N_=1 THEN DE2_BPWR_10=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB2 >0.10) THEN
DE2_BPWR_10=DE2_BPWR_10+1/500;
RETAIN DE2_BPWR_10;
RUN;

DATA DE2_BPWR_20 (KEEP=Condition Sampnum DE2_BPWR_20);
IF _N_=1 THEN DE2_BPWR_20=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB2 >0.20) THEN
DE2_BPWR_20=DE2_BPWR_20+1/500;
RETAIN DE2_BPWR_20;
RUN;

DATA DE2_BPWR_30 (KEEP=Condition Sampnum DE2_BPWR_30);
IF _N_=1 THEN DE2_BPWR_30=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB2 >0.30) THEN
DE2_BPWR_30=DE2_BPWR_30+1/500;
RETAIN DE2_BPWR_30;
RUN;

DATA DE2_BPWR_40 (KEEP=Condition Sampnum DE2_BPWR_40);
IF _N_=1 THEN DE2_BPWR_40=0;
SET WQSEST3v_500;
IF (DE2 ^=0 & wXB2 >0.40) THEN
DE2_BPWR_40=DE2_BPWR_40+1/500;
RETAIN DE2_BPWR_40;
RUN;

/**************DE3_B***********/
DATA DE3_BTYP1 (KEEP=Condition Sampnum DE3_BTYP1);
IF _N_=1 THEN DE3_BTYP1=0;
```

```
SET WQSEST3v_500;
IF (DE3 =0 & wXB3>0 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE3_BTYP1=DE3_BTYP1+1/500;
RETAIN DE3_BTYP1;
RUN;

DATA DE3_BTYP1_10 (KEEP=Condition Sampnum DE3_BTYP1_10);
IF _N_=1 THEN DE3_BTYP1_10=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB3 >0.10 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE3_BTYP1_10=DE3_BTYP1_10+1/500;
RETAIN DE3_BTYP1_10;
RUN;

DATA DE3_BTYP1_20 (KEEP=Condition Sampnum DE3_BTYP1_20);
IF _N_=1 THEN DE3_BTYP1_20=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB3 >0.20 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE3_BTYP1_20=DE3_BTYP1_20+1/500;
RETAIN DE3_BTYP1_20;
RUN;

DATA DE3_BTYP1_30 (KEEP=Condition Sampnum DE3_BTYP1_30);
IF _N_=1 THEN DE3_BTYP1_30=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB3 >0.30 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE3_BTYP1_30=DE3_BTYP1_30+1/500;
RETAIN DE3_BTYP1_30;
RUN;
DATA DE3_BTYP1_40 (KEEP=Condition Sampnum DE3_BTYP1_40);
IF _N_=1 THEN DE3_BTYP1_40=0;
SET WQSEST3v_500;
IF (DE3 =0 & wXB3 >0.40 & DE123=0 & ^(DE123_BLCL<0<DE123_BUCL)) THEN
DE3_BTYP1_40=DE3_BTYP1_40+1/500;
RETAIN DE3_BTYP1_40;
RUN;

DATA DE3_BPWR (KEEP=Condition Sampnum DE3_BPWR);
IF _N_=1 THEN DE3_BPWR=0;
SET WQSEST3v_500;
```

```
IF (DE3 ^=0 & wXB3>0) THEN
DE3_BPWR=DE3_BPWR+1/500;
RETAIN DE3_BPWR;
RUN;

DATA DE3_BPWR_10 (KEEP=Condition Sampnum DE3_BPWR_10);
IF _N_=1 THEN DE3_BPWR_10=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB3 >0.10) THEN
DE3_BPWR_10=DE3_BPWR_10+1/500;
RETAIN DE3_BPWR_10;
RUN;

DATA DE3_BPWR_20 (KEEP=Condition Sampnum DE3_BPWR_20);
IF _N_=1 THEN DE3_BPWR_20=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB3 >0.20) THEN
DE3_BPWR_20=DE3_BPWR_20+1/500;
RETAIN DE3_BPWR_20;
RUN;

DATA DE3_BPWR_30 (KEEP=Condition Sampnum DE3_BPWR_30);
IF _N_=1 THEN DE3_BPWR_30=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB3 >0.30) THEN
DE3_BPWR_30=DE3_BPWR_30+1/500;
RETAIN DE3_BPWR_30;
RUN;

DATA DE3_BPWR_40 (KEEP=Condition Sampnum DE3_BPWR_40);
IF _N_=1 THEN DE3_BPWR_40=0;
SET WQSEST3v_500;
IF (DE3 ^=0 & wXB3 >0.40) THEN
DE3_BPWR_40=DE3_BPWR_40+1/500;
RETAIN DE3_BPWR_40;
RUN;

/*Individual Estimates ME SEQ & NEST, TYPE1 ERR and POWER for 0.10, 0.20, 0.30 , 0.40 and 0.50 cutoff for WQS weigts */
/******************ME1_0****************/
DATA ME1_0TYP1 (KEEP=Condition Sampnum ME1_0TYP1);
```

```
IF _N_=1 THEN ME1_0TYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB01>0 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1_0TYP1=ME1_0TYP1+(1/500);
RETAIN ME1_0TYP1;
RUN;

/*ME1T1Z_0TYP1                    ME1GZ_0TYP1            ME1T1GZ_0TYP1 */
DATA ME1T1Z_0TYP1 (KEEP=Condition Sampnum ME1T1Z_0TYP1);
IF _N_=1 THEN           ME1T1Z_0TYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB01>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1T1Z_0TYP1=ME1T1Z_0TYP1+(1/500);
RETAIN ME1T1Z_0TYP1;
RUN;

DATA ME1GZ_0TYP1 (KEEP=Condition Sampnum ME1GZ_0TYP1);
IF _N_=1 THEN           ME1GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB01>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1GZ_0TYP1=ME1GZ_0TYP1+(1/500);
RETAIN ME1GZ_0TYP1;
RUN;

DATA ME1T1GZ_0TYP1 (KEEP=Condition Sampnum ME1T1GZ_0TYP1);
IF _N_=1 THEN           ME1T1GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB01>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1T1GZ_0TYP1=ME1T1GZ_0TYP1+(1/500);
RETAIN ME1T1GZ_0TYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB01 wXB02 wXB03 WXB01 wXB02 and wXB03 values below are =0 weights */
DATA ME1_0TYP1_10 (KEEP=Condition Sampnum   ME1_0TYP1_10);
IF _N_=1 THEN ME1_0TYP1_10=0;
SET WQSEST3v_500;
```

```
IF (ME1=0 & wXB01>0.10 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1_0TYP1_10=ME1_0TYP1_10+(1/500);
RETAIN ME1_0TYP1_10;
RUN;

DATA ME1_0TYP1_20 (KEEP=Condition Sampnum ME1_0TYP1_20);
IF _N_=1 THEN ME1_0TYP1_20=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB01>0.20 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1_0TYP1_20=ME1_0TYP1_20 + (1/500);
RETAIN ME1_0TYP1_20;
RUN;

DATA ME1_0TYP1_30 (KEEP=Condition Sampnum ME1_0TYP1_30);
IF _N_=1 THEN ME1_0TYP1_30=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB01>0.30 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1_0TYP1_30=ME1_0TYP1_30 + (1/500);
RETAIN ME1_0TYP1_30;
RUN;

DATA ME1_0TYP1_40 (KEEP=Condition Sampnum ME1_0TYP1_40);
IF _N_=1 THEN ME1_0TYP1_40=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB01>0.40 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME1_0TYP1_40=ME1_0TYP1_40 + (1/500);
RETAIN ME1_0TYP1_40;
RUN;

DATA ME1_0PWR (KEEP= Condition Sampnum ME1_0PWR);
IF _N_=1 THEN ME1_0PWR=0;
SET WQSEST3v_500;
IF ((ME1^=0) & wXB01>0) THEN
ME1_0PWR=ME1_0PWR + (1/500);
RETAIN ME1_0PWR;
RUN;
```

```
DATA ME1_0PWR_10 (KEEP= Condition Sampnum ME1_0PWR_10);
IF _N_=1 THEN ME1_0PWR_10=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB01>0.10) THEN
ME1_0PWR_10=ME1_0PWR_10 + (1/500);
RETAIN ME1_0PWR_10;
RUN;

DATA ME1_0PWR_20 (KEEP= Condition Sampnum ME1_0PWR_20);
IF _N_=1 THEN ME1_0PWR_20=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB01>0.20) THEN
ME1_0PWR_20=ME1_0PWR_20 + (1/500);
RETAIN ME1_0PWR_20;
RUN;

DATA ME1_0PWR_3 (KEEP= Condition Sampnum ME1_0PWR_30);
IF _N_=1 THEN ME1_0PWR_30=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB01>0.30) THEN
ME1_0PWR_30=ME1_0PWR_30 + (1/500);
RETAIN ME1_0PWR_30;
RUN;
DATA ME1_0PWR_40 (KEEP= Condition Sampnum ME1_0PWR_40);
IF _N_=1 THEN ME1_0PWR_40=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB01>0.40) THEN
ME1_0PWR_40=ME1_0PWR_40 + (1/500);
RETAIN ME1_0PWR_40;
RUN;
/*********ME2_0***********/
DATA ME2_0TYP1 (KEEP=Condition Sampnum ME2_0TYP1);
IF _N_=1 THEN ME2_0TYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB02>0 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2_0TYP1=ME2_0TYP1 + (1/500);
RETAIN ME2_0TYP1;
RUN;
```

```
/*ME2T2Z_0TYP1                    ME2GZ_0TYP1            ME2T2GZ_0TYP1 */
DATA ME2T2Z_0TYP1 (KEEP=Condition Sampnum ME2T2Z_0TYP1);
IF _N_=1 THEN           ME2T2Z_0TYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXB02>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2T2Z_0TYP1=ME2T2Z_0TYP1 + (1/500);
RETAIN ME2T2Z_0TYP1;
RUN;

DATA ME2GZ_0TYP1 (KEEP=Condition Sampnum ME2GZ_0TYP1);
IF _N_=1 THEN           ME2GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB02>0   &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2GZ_0TYP1=ME2GZ_0TYP1 + (1/500);
RETAIN ME2GZ_0TYP1;
RUN;

DATA ME2T2GZ_0TYP1 (KEEP=Condition Sampnum ME2T2GZ_0TYP1);
IF _N_=1 THEN           ME2T2GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB02>0   &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2T2GZ_0TYP1=ME2T2GZ_0TYP1 + (1/500);
RETAIN ME2T2GZ_0TYP1;
RUN;

DATA ME2_0TYP1_10 (KEEP=Condition Sampnum ME2_0TYP1_10);
IF _N_=1 THEN ME2_0TYP1_10=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB02>0.10   &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2_0TYP1_10=ME2_0TYP1_10 + (1/500);
RETAIN ME2_0TYP1_10;
RUN;

DATA ME2_0TYP1_20 (KEEP=Condition Sampnum   ME2_0TYP1_20);
IF _N_=1 THEN ME2_0TYP1_20=0;
```

```
SET WQSEST3v_500;
IF (ME2=0 & wXB02>0.20 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2_0TYP1_20=ME2_0TYP1_20 + (1/500);
RETAIN ME2_0TYP1_20;
RUN;

DATA ME2_0TYP1_30 (KEEP=Condition Sampnum   ME2_0TYP1_30);
IF _N_=1 THEN ME2_0TYP1_30=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB02>0.30 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2_0TYP1_30=ME2_0TYP1_30 + (1/500);
RETAIN ME2_0TYP1_30;
RUN;

DATA ME2_0TYP1_40 (KEEP=Condition Sampnum ME2_0TYP1_40);
IF _N_=1 THEN ME2_0TYP1_40=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB02>0.40 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME2_0TYP1_40=ME2_0TYP1_40 + (1/500);
RETAIN ME2_0TYP1_40;
RUN;

DATA ME2_0PW (KEEP= Condition Sampnum ME2_0PWR);
IF _N_=1 THEN ME2_0PWR=0;
SET WQSEST3v_500;
IF ((ME2^=0) & wXB02>0) THEN
ME2_0PWR=ME2_0PWR + (1/500);
RETAIN ME2_0PWR;
RUN;

DATA ME2_0PWR_10 (KEEP= Condition Sampnum ME2_0PWR_10);
IF _N_=1 THEN ME2_0PWR_10=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB02>0.10) THEN
ME2_0PWR_10=ME2_0PWR_10 + (1/500);
RETAIN ME2_0PWR_10;
RUN;
```

```
DATA ME2_0PWR_20 (KEEP= Condition Sampnum ME2_0PWR_20);
IF _N_=1 THEN ME2_0PWR_20=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB02>0.20) THEN
ME2_0PWR_20=ME2_0PWR_20 + (1/500);
RETAIN ME2_0PWR_20;
RUN;

DATA ME2_0PWR_30 (KEEP= Condition Sampnum ME2_0PWR_30);
IF _N_=1 THEN ME2_0PWR_30=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB02>0.30) THEN
ME2_0PWR_30=ME2_0PWR_30 + (1/500);
RETAIN ME2_0PWR_30;
RUN;

DATA ME2_0PWR_40 (KEEP= Condition Sampnum ME2_0PWR_40);
IF _N_=1 THEN ME2_0PWR_40=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB02>0.40) THEN
ME2_0PWR_40=ME2_0PWR_40 + (1/500);
RETAIN ME2_0PWR_40;
RUN;
/*********ME3_0***********/
DATA ME3_0TYP1 (KEEP=Condition Sampnum   ME3_0TYP1);
IF _N_=1 THEN ME3_0TYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB03>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3_0TYP1=ME3_0TYP1+(1/500);
RETAIN ME3_0TYP1;
RUN;

/*ME3T3Z_0TYP1                    ME3GZ_0TYP1            ME3T3GZ_0TYP1 */
DATA ME3T3Z_0TYP1 (KEEP=Condition Sampnum ME3T3Z_0TYP1);
IF _N_=1 THEN           ME3T3Z_0TYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma^=0 & wXB03>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3T3Z_0TYP1=ME3T3Z_0TYP1 + (1/500);
```

```
RETAIN ME3T3Z_0TYP1;
RUN;

DATA ME3GZ_0TYP1 (KEEP=Condition Sampnum ME3GZ_0TYP1);
IF _N_=1 THEN           ME3GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3^=0 & Gamma=0 & wXB03>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3GZ_0TYP1=ME3GZ_0TYP1 + (1/500);
RETAIN ME3GZ_0TYP1;
RUN;

DATA ME3T3GZ_0TYP1 (KEEP=Condition Sampnum ME3T3GZ_0TYP1);
IF _N_=1 THEN           ME3T3GZ_0TYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma=0 & wXB03>0 &
ME123=0 & ^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3T3GZ_0TYP1=ME3T3GZ_0TYP1 + (1/500);
RETAIN ME3T3GZ_0TYP1;
RUN;

DATA ME3_0TYP1_10 (KEEP=Condition Sampnum ME3_0TYP1_10);
IF _N_=1 THEN ME3_0TYP1_10=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB03>0.10 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3_0TYP1_10=ME3_0TYP1_10 + (1/500);
RETAIN ME3_0TYP1_10;
RUN;

DATA ME3_0TYP1_20 (KEEP=Condition Sampnum ME3_0TYP1_20);
IF _N_=1 THEN ME3_0TYP1_20=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB03>0.20 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3_0TYP1_20=ME3_0TYP1_20 + (1/500);
RETAIN ME3_0TYP1_20;
RUN;

DATA ME3_0TYP1_30 (KEEP=Condition Sampnum   ME3_0TYP1_30);
```

```
IF _N_=1 THEN ME3_0TYP1_30=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB03>0.30 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3_0TYP1_30=ME3_0TYP1_30 + (1/500);
RETAIN ME3_0TYP1_30;
RUN;

DATA ME3_0TYP1_40 (KEEP=Condition Sampnum ME3_0TYP1_40);
IF _N_=1 THEN ME3_0TYP1_40=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB03>0.40 & ME123=0 &
^(P01_ME123_2_5<=0<=P01_ME123_97_5)) THEN
ME3_0TYP1_40=ME3_0TYP1_40 + (1/500);
RETAIN ME3_0TYP1_40;
RUN;

DATA ME3_0PWR (KEEP= Condition Sampnum ME3_0PWR);
IF _N_=1 THEN ME3_0PWR=0;
SET WQSEST3v_500;
IF ((ME3^=0) & wXB03>0) THEN
ME3_0PWR=ME3_0PWR + (1/500);
RETAIN ME3_0PWR;
RUN;

DATA ME3_0PWR_10 (KEEP= Condition Sampnum ME3_0PWR_10);
IF _N_=1 THEN ME3_0PWR_10=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB03>0.10) THEN
ME3_0PWR_10=ME3_0PWR_10 + (1/500);
RETAIN ME3_0PWR_10;
RUN;

DATA ME3_0PWR_20 (KEEP= Condition Sampnum ME3_0PWR_20);
IF _N_=1 THEN ME3_0PWR_20=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB03>0.20) THEN
ME3_0PWR_20=ME3_0PWR_20 + (1/500);
RETAIN ME3_0PWR_20;
RUN;
```

```
DATA ME3_0PWR_30 (KEEP= Condition Sampnum ME3_0PWR_30);
IF _N_=1 THEN ME3_0PWR_30=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB03>0.30) THEN
ME3_0PWR_30=ME3_0PWR_30 + (1/500);
RETAIN ME3_0PWR_30;
RUN;

DATA ME3_0PWR_40 (KEEP= Condition Sampnum ME3_0PWR_40);
IF _N_=1 THEN ME3_0PWR_40=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB03>0.40) THEN
ME3_0PWR_40=ME3_0PWR_40 + (1/500);
RETAIN ME3_0PWR_40;
RUN;

/******************ME1_T****************/
DATA ME1_TTYP1 (KEEP=Condition Sampnum ME1_TTYP1);
IF _N_=1 THEN ME1_TTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & wXT1>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1_TTYP1=ME1_TTYP1 + (1/500);
RETAIN ME1_TTYP1;
RUN;

/*ME1T1Z_TTYP1                    ME1GZ_TTYP1           ME1T1GZ_TTYP1 */
DATA ME1T1Z_TTYP1 (KEEP=Condition Sampnum ME1T1Z_TTYP1);
IF _N_=1 THEN           ME1T1Z_TTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXT1>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1T1Z_TTYP1=ME1T1Z_TTYP1 + (1/500);
RETAIN ME1T1Z_TTYP1;
RUN;

DATA ME1GZ_TTYP1 (KEEP=Condition Sampnum ME1GZ_TTYP1);
IF _N_=1 THEN           ME1GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXT1>0 & ME123=0 &
```

```
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1GZ_TTYP1=ME1GZ_TTYP1 + (1/500);
RETAIN ME1GZ_TTYP1;
RUN;

DATA ME1T1GZ_TTYP1 (KEEP=Condition Sampnum ME1T1GZ_TTYP1);
IF _N_=1 THEN            ME1T1GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXT1>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1T1GZ_TTYP1=ME1T1GZ_TTYP1 + (1/500);
RETAIN ME1T1GZ_TTYP1;
RUN;

DATA ME1_TTYP1_10 (KEEP=Condition Sampnum   ME1_TTYP1_10);
IF _N_=1 THEN ME1_TTYP1_10=0;
SET WQSEST3v_500;
IF (ME1=0 & wXT1>0.10 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1_TTYP1_10=ME1_TTYP1_10 + (1/500);
RETAIN ME1_TTYP1_10;
RUN;

DATA ME1_TTYP1_20 (KEEP=Condition Sampnum ME1_TTYP1_20);
IF _N_=1 THEN ME1_TTYP1_20=0;
SET WQSEST3v_500;
IF (ME1=0 & wXT1>0.20 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1_TTYP1_20=ME1_TTYP1_20 + (1/500);
RETAIN ME1_TTYP1_20;
RUN;

DATA ME1_TTYP1_30 (KEEP=Condition Sampnum ME1_TTYP1_30);
IF _N_=1 THEN ME1_TTYP1_30=0;
SET WQSEST3v_500;
IF (ME1=0 & wXT1>0.30 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1_TTYP1_30=ME1_TTYP1_30 + (1/500);
RETAIN ME1_TTYP1_30;
RUN;
```

```
DATA ME1_TTYP1_40 (KEEP=Condition Sampnum ME1_TTYP1_40);
IF _N_=1 THEN ME1_TTYP1_40=0;
SET WQSEST3v_500;
IF (ME1=0 & wXT1>0.40 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME1_TTYP1_40=ME1_TTYP1_40 + (1/500);
RETAIN ME1_TTYP1_40;
RUN;

DATA ME1_TPWR (KEEP= Condition Sampnum ME1_TPWR);
IF _N_=1 THEN ME1_TPWR=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXT1>0) THEN
ME1_TPWR=ME1_TPWR + (1/500);
RETAIN ME1_TPWR;
RUN;



DATA ME1_TPWR_10 (KEEP= Condition Sampnum ME1_TPWR_10);
IF _N_=1 THEN ME1_TPWR_10=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXT1>0.10) THEN
ME1_TPWR_10=ME1_TPWR_10 + (1/500);
RETAIN ME1_TPWR_10;
RUN;

DATA ME1_TPWR_20 (KEEP= Condition Sampnum ME1_TPWR_20);
IF _N_=1 THEN ME1_TPWR_20=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXT1>0.20) THEN
ME1_TPWR_20=ME1_TPWR_20 + (1/500);
RETAIN ME1_TPWR_20;
RUN;

DATA ME1_TPWR_30 (KEEP= Condition Sampnum ME1_TPWR_30);
IF _N_=1 THEN ME1_TPWR_30=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXT1>0.30) THEN
ME1_TPWR_30=ME1_TPWR_30 + (1/500);
```

```
RETAIN ME1_TPWR_30;
RUN;
DATA ME1_TPWR_40 (KEEP= Condition Sampnum ME1_TPWR_40);
IF _N_=1 THEN ME1_TPWR_40=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXT1>0.40) THEN
ME1_TPWR_40=ME1_TPWR_40 + (1/500);
RETAIN ME1_TPWR_40;
RUN;

/*********ME2_T***********/
DATA ME2_TTYP1 (KEEP=Condition Sampnum ME2_TTYP1);
IF _N_=1 THEN ME2_TTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & wXT2>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME2_TTYP1=ME2_TTYP1 + (1/500);
RETAIN ME2_TTYP1;
RUN;

/*ME2T2Z_TTYP1                    ME2GZ_TTYP1            ME2T2GZ_TTYP1 */
DATA ME2T2Z_TTYP1 (KEEP=Condition Sampnum ME2T2Z_TTYP1);
IF _N_=1 THEN           ME2T2Z_TTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXT2>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5))   THEN
ME2T2Z_TTYP1=ME2T2Z_TTYP1 + (1/500);
RETAIN ME2T2Z_TTYP1;
RUN;

DATA ME2GZ_TTYP1 (KEEP=Condition Sampnum ME2GZ_TTYP1);
IF _N_=1 THEN           ME2GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXT2>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5))   THEN
ME2GZ_TTYP1=ME2GZ_TTYP1 + (1/500);
RETAIN ME2GZ_TTYP1;
RUN;
```

```
DATA ME2T2GZ_TTYP1 (KEEP=Condition Sampnum ME2T2GZ_TTYP1);
IF _N_=1 THEN           ME2T2GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXT2>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5))  THEN
ME2T2GZ_TTYP1=ME2T2GZ_TTYP1 + (1/500);
RETAIN ME2T2GZ_TTYP1;
RUN;

DATA ME2_TTYP1_10 (KEEP=Condition Sampnum ME2_TTYP1_10);
IF _N_=1 THEN ME2_TTYP1_10=0;
SET WQSEST3v_500;
IF (ME2=0 & wXT2>0.10 & ME123=0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME2_TTYP1_10=ME2_TTYP1_10 + (1/500);
RETAIN ME2_TTYP1_10;
RUN;

DATA ME2_TTYP1_20 (KEEP=Condition Sampnum ME2_TTYP1_20);
IF _N_=1 THEN ME2_TTYP1_20=0;
SET WQSEST3v_500;
IF (ME2=0 & wXT2>0.20 & ME123=0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME2_TTYP1_20=ME2_TTYP1_20 + (1/500);
RETAIN ME2_TTYP1_20;
RUN;

DATA ME2_TTYP1_30 (KEEP=Condition Sampnum ME2_TTYP1_30);
IF _N_=1 THEN ME2_TTYP1_30=0;
SET WQSEST3v_500;
IF (ME2=0 & wXT2>0.30 & ME123=0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME2_TTYP1_30=ME2_TTYP1_30 + (1/500);
RETAIN ME2_TTYP1_30;
RUN;

DATA ME2_TTYP1_40 (KEEP=Condition Sampnum ME2_TTYP1_40);
IF _N_=1 THEN ME2_TTYP1_40=0;
SET WQSEST3v_500;
IF (ME2=0 & wXT2>0.40 & ME123=0 & ME123=0 &
```

```
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME2_TTYP1_40=ME2_TTYP1_40 + (1/500);
RETAIN ME2_TTYP1_40;
RUN;

DATA ME2_TPWR (KEEP= Condition Sampnum ME2_TPWR);
IF _N_=1 THEN ME2_TPWR=0;
SET WQSEST3v_500;
IF ((ME2^=0) & wXT2>0) THEN
ME2_TPWR=ME2_TPWR + (1/500);
RETAIN ME2_TPWR;
RUN;

DATA ME2_TPWR_10 (KEEP= Condition Sampnum ME2_TPWR_10);
IF _N_=1 THEN ME2_TPWR_10=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXT2>0.10) THEN
ME2_TPWR_10=ME2_TPWR_10 + (1/500);
RETAIN ME2_TPWR_10;
RUN;

DATA ME2_TPWR_20 (KEEP= Condition Sampnum ME2_TPWR_20);
IF _N_=1 THEN ME2_TPWR_20=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXT2>0.20) THEN
ME2_TPWR_20=ME2_TPWR_20 + (1/500);
RETAIN ME2_TPWR_20;
RUN;

DATA ME2_TPWR_30 (KEEP= Condition Sampnum ME2_TPWR_30);
IF _N_=1 THEN ME2_TPWR_30=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXT2>0.30) THEN
ME2_TPWR_30=ME2_TPWR_30 + (1/500);
RETAIN ME2_TPWR_30;
RUN;

DATA ME2_TPWR_40 (KEEP= Condition Sampnum ME2_TPWR_40);
IF _N_=1 THEN ME2_TPWR_40=0;
SET WQSEST3v_500;
```

```
IF (ME2^=0 & wXT2>0.40) THEN
ME2_TPWR_40=ME2_TPWR_40 + (1/500);
RETAIN ME2_TPWR_40;
RUN;

/*********ME3_T***********/
DATA ME3_TTYP1 (KEEP=Condition Sampnum ME3_TTYP1);
IF _N_=1 THEN ME3_TTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & wXT3>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3_TTYP1=ME3_TTYP1 + (1/500);
RETAIN ME3_TTYP1;
RUN;

/*ME3T3Z_TTYP1                    ME3GZ_TTYP1            ME3T3GZ_TTYP1 */
DATA ME3T3Z_TTYP1 (KEEP=Condition Sampnum ME3T3Z_TTYP1);
IF _N_=1 THEN           ME3T3Z_TTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma^=0 & wXT3>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3T3Z_TTYP1=ME3T3Z_TTYP1 + (1/500);
RETAIN ME3T3Z_TTYP1;
RUN;

DATA ME3GZ_TTYP1 (KEEP=Condition Sampnum ME3GZ_TTYP1);
IF _N_=1 THEN           ME3GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3^=0 & Gamma=0 & wXT3>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3GZ_TTYP1=ME3GZ_TTYP1 + (1/500);
RETAIN ME3GZ_TTYP1;
RUN;

DATA ME3T3GZ_TTYP1 (KEEP=Condition Sampnum ME3T3GZ_TTYP1);
IF _N_=1 THEN           ME3T3GZ_TTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma=0 & wXT3>0 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3T3GZ_TTYP1=ME3T3GZ_TTYP1 + (1/500);
```

```
RETAIN ME3T3GZ_TTYP1;
RUN;

DATA ME3_TTYP1_10 (KEEP=Condition Sampnum   ME3_TTYP1_10);
IF _N_=1 THEN ME3_TTYP1_10=0;
SET WQSEST3v_500;
IF (ME3=0 & wXT3>0.10 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3_TTYP1_10=ME3_TTYP1_10 + (1/500);
RETAIN ME3_TTYP1_10;
RUN;

DATA ME3_TTYP1_20 (KEEP=Condition Sampnum ME3_TTYP1_20);
IF _N_=1 THEN ME3_TTYP1_20=0;
SET WQSEST3v_500;
IF (ME3=0 & wXT3>0.20 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3_TTYP1_20=ME3_TTYP1_20 + (1/500);
RETAIN ME3_TTYP1_20;
RUN;

DATA ME3_TTYP1_30 (KEEP=Condition Sampnum ME3_TTYP1_30);
IF _N_=1 THEN ME3_TTYP1_30=0;
SET WQSEST3v_500;
IF (ME3=0 & wXT3>0.30 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3_TTYP1_30=ME3_TTYP1_30 + (1/500);
RETAIN ME3_TTYP1_30;
RUN;

DATA ME3_TTYP1_40 (KEEP=Condition Sampnum ME3_TTYP1_40);
IF _N_=1 THEN ME3_TTYP1_40=0;
SET WQSEST3v_500;
IF (ME3=0 & wXT3>0.40 & ME123=0 &
^(PT1_ME123_2_5<=0<=PT1_ME123_97_5)) THEN
ME3_TTYP1_40=ME3_TTYP1_40 + (1/500);
RETAIN ME3_TTYP1_40;
RUN;
```

```
DATA ME3_TPWR (KEEP= Condition Sampnum ME3_TPWR);
IF _N_=1 THEN ME3_TPWR=0;
SET WQSEST3v_500;
IF ((ME3^=0) & wXT3>0) THEN
ME3_TPWR=ME3_TPWR + (1/500);
RETAIN ME3_TPWR;
RUN;

DATA ME3_TPWR_10 (KEEP= Condition Sampnum ME3_TPWR_10);
IF _N_=1 THEN ME3_TPWR_10=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXT3>0.10) THEN
ME3_TPWR_10=ME3_TPWR_10 + (1/500);
RETAIN ME3_TPWR_10;
RUN;

DATA ME3_TPWR_20 (KEEP= Condition Sampnum ME3_TPWR_20);
IF _N_=1 THEN ME3_TPWR_20=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXT3>0.20) THEN
ME3_TPWR_20=ME3_TPWR_20 + (1/500);
RETAIN ME3_TPWR_20;
RUN;

DATA ME3_TPWR_30 (KEEP= Condition Sampnum ME3_TPWR_30);
IF _N_=1 THEN ME3_TPWR_30=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXT3>0.30) THEN
ME3_TPWR_30=ME3_TPWR_30 + (1/500);
RETAIN ME3_TPWR_30;
RUN;

DATA ME3_TPWR_40 (KEEP= Condition Sampnum ME3_TPWR_40);
IF _N_=1 THEN ME3_TPWR_40=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXT3>0.40) THEN
ME3_TPWR_40=ME3_TPWR_40 + (1/500);
RETAIN ME3_TPWR_40;
RUN;
```

```
/******************ME1_B****************/
DATA ME1_BTYP1 (KEEP=Condition Sampnum ME1_BTYP1);
IF _N_=1 THEN ME1_BTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB1>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1_BTYP1=ME1_BTYP1 + (1/500);
RETAIN ME1_BTYP1;
RUN;

/*ME1T1Z_BTYP1                    ME1GZ_BTYP1            ME1T1GZ_BTYP1 */
DATA ME1T1Z_BTYP1 (KEEP=Condition Sampnum ME1T1Z_BTYP1);
IF _N_=1 THEN           ME1T1Z_BTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB1>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1T1Z_BTYP1=ME1T1Z_BTYP1 + (1/500);
RETAIN ME1T1Z_BTYP1;
RUN;

DATA ME1GZ_BTYP1 (KEEP=Condition Sampnum ME1GZ_BTYP1);
IF _N_=1 THEN           ME1GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB1>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1GZ_BTYP1=ME1GZ_BTYP1 + (1/500);
RETAIN ME1GZ_BTYP1;
RUN;

DATA ME1T1GZ_BTYP1 (KEEP=Condition Sampnum ME1T1GZ_BTYP1);
IF _N_=1 THEN           ME1T1GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB1>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1T1GZ_BTYP1=ME1T1GZ_BTYP1 + (1/500);
RETAIN ME1T1GZ_BTYP1;
RUN;

DATA ME1_BTYP1_10 (KEEP=Condition Sampnum   ME1_BTYP1_10);
IF _N_=1 THEN ME1_BTYP1_10=0;
```

```
SET WQSEST3v_500;
IF (ME1=0 & wXB1>0.10 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1_BTYP1_10=ME1_BTYP1_10 + (1/500);
RETAIN ME1_BTYP1_10;
RUN;

DATA ME1_BTYP1_20 (KEEP=Condition Sampnum   ME1_BTYP1_20);
IF _N_=1 THEN ME1_BTYP1_20=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB1>0.20 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1_BTYP1_20=ME1_BTYP1_20 + (1/500);
RETAIN ME1_BTYP1_20;
RUN;

DATA ME1_BTYP1_30 (KEEP=Condition Sampnum ME1_BTYP1_30);
IF _N_=1 THEN ME1_BTYP1_30=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB1>0.30 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1_BTYP1_30=ME1_BTYP1_30 + (1/500);
RETAIN ME1_BTYP1_30;
RUN;

DATA ME1_BTYP1_40 (KEEP=Condition Sampnum ME1_BTYP1_40);
IF _N_=1 THEN ME1_BTYP1_40=0;
SET WQSEST3v_500;
IF (ME1=0 & wXB1>0.40 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME1_BTYP1_40=ME1_BTYP1_40 + (1/500);
RETAIN ME1_BTYP1_40;
RUN;

DATA ME1_BPWR (KEEP= Condition Sampnum ME1_BPWR);
IF _N_=1 THEN ME1_BPWR=0;
SET WQSEST3v_500;
IF ((ME1^=0) & wXB1>0) THEN
ME1_BPWR=ME1_BPWR + (1/500);
RETAIN ME1_BPWR;
```

```
RUN;

DATA ME1_BPWR_10 (KEEP= Condition Sampnum ME1_BPWR_10);
IF _N_=1 THEN ME1_BPWR_10=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB1>0.10) THEN
ME1_BPWR_10=ME1_BPWR_10 + (1/500);
RETAIN ME1_BPWR_10;
RUN;

DATA ME1_BPWR_20 (KEEP= Condition Sampnum ME1_BPWR_20);
IF _N_=1 THEN ME1_BPWR_20=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB1>0.20) THEN
ME1_BPWR_20=ME1_BPWR_20 + (1/500);
RETAIN ME1_BPWR_20;
RUN;

DATA ME1_BPWR_30 (KEEP= Condition Sampnum ME1_BPWR_30);
IF _N_=1 THEN ME1_BPWR_30=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB1>0.30) THEN
ME1_BPWR_30=ME1_BPWR_30 + (1/500);
RETAIN ME1_BPWR_30;
RUN;

DATA ME1_BPWR_40 (KEEP= Condition Sampnum ME1_BPWR_40);
IF _N_=1 THEN ME1_BPWR_40=0;
SET WQSEST3v_500;
IF (ME1^=0 & wXB1>0.40) THEN
ME1_BPWR_40=ME1_BPWR_40 + (1/500);
RETAIN ME1_BPWR_40;
RUN;

/*********ME2_B***********/
DATA ME2_BTYP1 (KEEP=Condition Sampnum ME2_BTYP1);
IF _N_=1 THEN ME2_BTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB2>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
```

```
ME2_BTYP1=ME2_BTYP1+ (1/500);
RETAIN ME2_BTYP1;
RUN;

/*ME2T2Z_BTYP1                  ME2GZ_BTYP1           ME2T2GZ_BTYP1 */
DATA ME2T2Z_BTYP1 (KEEP=Condition Sampnum ME2T2Z_BTYP1);
IF _N_=1 THEN           ME2T2Z_BTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXB2>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5))   THEN
ME2T2Z_BTYP1=ME2T2Z_BTYP1 + (1/500);
RETAIN ME2T2Z_BTYP1;
RUN;

DATA ME2GZ_BTYP1 (KEEP=Condition Sampnum ME2GZ_BTYP1);
IF _N_=1 THEN           ME2GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB2>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2GZ_BTYP1=ME2GZ_BTYP1 + (1/500);
RETAIN ME2GZ_BTYP1;
RUN;

DATA ME2T2GZ_BTYP1 (KEEP=Condition Sampnum ME2T2GZ_BTYP1);
IF _N_=1 THEN           ME2T2GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB2>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2T2GZ_BTYP1=ME2T2GZ_BTYP1 + (1/500);
RETAIN ME2T2GZ_BTYP1;
RUN;

DATA ME2_BTYP1_10 (KEEP=Condition Sampnum ME2_BTYP1_10);
IF _N_=1 THEN ME2_BTYP1_10=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB2>0.10 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2_BTYP1_10=ME2_BTYP1_10 + (1/500);
RETAIN ME2_BTYP1_10;
RUN;
```

```
DATA ME2_BTYP1_20 (KEEP=Condition Sampnum ME2_BTYP1_20);
IF _N_=1 THEN ME2_BTYP1_20=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB2>0.20 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2_BTYP1_20=ME2_BTYP1_20 + (1/500);
RETAIN ME2_BTYP1_20;
RUN;

DATA ME2_BTYP1_30 (KEEP=Condition Sampnum ME2_BTYP1_30);
IF _N_=1 THEN ME2_BTYP1_30=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB2>0.30 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2_BTYP1_30=ME2_BTYP1_30+ (1/500);
RETAIN ME2_BTYP1_30;
RUN;
DATA ME2_BTYP1_40 (KEEP=Condition Sampnum ME2_BTYP1_40);
IF _N_=1 THEN ME2_BTYP1_40=0;
SET WQSEST3v_500;
IF (ME2=0 & wXB2>0.40 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME2_BTYP1_40=ME2_BTYP1_40 + (1/500);
RETAIN ME2_BTYP1_40;
RUN;

DATA ME2_BPWR (KEEP= Condition Sampnum ME2_BPWR);
IF _N_=1 THEN ME2_BPWR=0;
SET WQSEST3v_500;
IF ((ME2^=0) & wXB2>0) THEN
ME2_BPWR=ME2_BPWR + (1/500);
RETAIN ME2_BPWR;
RUN;

DATA ME2_BPWR_10 (KEEP= Condition Sampnum ME2_BPWR_10);
IF _N_=1 THEN ME2_BPWR_10=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB2>0.10) THEN
ME2_BPWR_10=ME2_BPWR_10 + (1/500);
```

```
RETAIN ME2_BPWR_10;
RUN;

DATA ME2_BPWR_20 (KEEP= Condition Sampnum ME2_BPWR_20);
IF _N_=1 THEN ME2_BPWR_20=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB2>0.20) THEN
ME2_BPWR_20=ME2_BPWR_20 + (1/500);
RETAIN ME2_BPWR_20;
RUN;

DATA ME2_BPWR_30 (KEEP= Condition Sampnum ME2_BPWR_30);
IF _N_=1 THEN ME2_BPWR_30=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB2>0.30) THEN
ME2_BPWR_30=ME2_BPWR_30 + (1/500);
RETAIN ME2_BPWR_30;
RUN;

DATA ME2_BPWR_40 (KEEP= Condition Sampnum ME2_BPWR_40);
IF _N_=1 THEN ME2_BPWR_40=0;
SET WQSEST3v_500;
IF (ME2^=0 & wXB2>0.40) THEN
ME2_BPWR_40=ME2_BPWR_40 + (1/500);
RETAIN ME2_BPWR_40;
RUN;

/*********ME3_B***********/
DATA ME3_BTYP1 (KEEP=Condition Sampnum ME3_BTYP1);
IF _N_=1 THEN ME3_BTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB3>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3_BTYP1=ME3_BTYP1 + (1/500);
RETAIN ME3_BTYP1;
RUN;

/*ME3T3Z_BTYP1                    ME3GZ_BTYP1           ME3T3GZ_BTYP1 */
DATA ME3T3Z_BTYP1 (KEEP=Condition Sampnum ME3T3Z_BTYP1);
IF _N_=1 THEN ME3T3Z_BTYP1=0;
```

```
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma^=0 & wXB3>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3T3Z_BTYP1=ME3T3Z_BTYP1 + (1/500);
RETAIN ME3T3Z_BTYP1;
RUN;

DATA ME3GZ_BTYP1 (KEEP=Condition Sampnum ME3GZ_BTYP1);
IF _N_=1 THEN ME3GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3^=0 & Gamma=0 & wXB3>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3GZ_BTYP1=ME3GZ_BTYP1 + (1/500);
RETAIN ME3GZ_BTYP1;
RUN;

DATA ME3T3GZ_BTYP1 (KEEP=Condition Sampnum ME3T3GZ_BTYP1);
IF _N_=1 THEN ME3T3GZ_BTYP1=0;
SET WQSEST3v_500;
IF (ME3=0 & Theta3=0 & Gamma=0 & wXB3>0 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3T3GZ_BTYP1=ME3T3GZ_BTYP1 + (1/500);
RETAIN ME3T3GZ_BTYP1;
RUN;

DATA ME3_BTYP1_10 (KEEP=Condition Sampnum ME3_BTYP1_10);
IF _N_=1 THEN ME3_BTYP1_10=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB3>0.10 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3_BTYP1_10=ME3_BTYP1_10+ (1/500);
RETAIN ME3_BTYP1_10;
RUN;

DATA ME3_BTYP1_20 (KEEP=Condition Sampnum ME3_BTYP1_20);
IF _N_=1 THEN ME3_BTYP1_20=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB3>0.20 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3_BTYP1_20=ME3_BTYP1_20 + (1/500);
```

```
RETAIN ME3_BTYP1_20;
RUN;

DATA ME3_BTYP1_30 (KEEP=Condition Sampnum ME3_BTYP1_30);
IF _N_=1 THEN ME3_BTYP1_30=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB3>0.30 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3_BTYP1_30=ME3_BTYP1_30 + (1/500);
RETAIN ME3_BTYP1_30;
RUN;

DATA ME3_BTYP1_40 (KEEP=Condition Sampnum ME3_BTYP1_40);
IF _N_=1 THEN ME3_BTYP1_40=0;
SET WQSEST3v_500;
IF (ME3=0 & wXB3>0.40 & ME123=0 &
^(PB1_ME123_2_5<=0<=PB1_ME123_97_5)) THEN
ME3_BTYP1_40=ME3_BTYP1_40 + (1/500);
RETAIN ME3_BTYP1_40;
RUN;

DATA ME3_BPWR (KEEP= Condition Sampnum ME3_BPWR);
IF _N_=1 THEN ME3_BPWR=0;
SET WQSEST3v_500;
IF ((ME3^=0) & wXB3>0) THEN
ME3_BPWR=ME3_BPWR + (1/500);
RETAIN ME3_BPWR;
RUN;

DATA ME3_BPWR_10 (KEEP= Condition Sampnum ME3_BPWR_10);
IF _N_=1 THEN ME3_BPWR_10=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB3>0.10) THEN
ME3_BPWR_10=ME3_BPWR_10 + (1/500);
RETAIN ME3_BPWR_10;
RUN;

DATA ME3_BPWR_20 (KEEP= Condition Sampnum ME3_BPWR_20);
IF _N_=1 THEN ME3_BPWR_20=0;
SET WQSEST3v_500;
```

```
IF (ME3^=0 & wXB3>0.20) THEN
ME3_BPWR_20=ME3_BPWR_20 + (1/500);
RETAIN ME3_BPWR_20;
RUN;

DATA ME3_BPWR_30 (KEEP= Condition Sampnum ME3_BPWR_30);
IF _N_=1 THEN ME3_BPWR_30=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB3>0.30) THEN
ME3_BPWR_30=ME3_BPWR_30 + (1/500);
RETAIN ME3_BPWR_30;
RUN;

DATA ME3_BPWR_40 (KEEP= Condition Sampnum ME3_BPWR_40);
IF _N_=1 THEN ME3_BPWR_40=0;
SET WQSEST3v_500;
IF (ME3^=0 & wXB3>0.40) THEN
ME3_BPWR_40=ME3_BPWR_40 + (1/500);
RETAIN ME3_BPWR_40;
RUN;

DATA WQS3V_TYP1PWR;   /* ADD TO 144 VARS IN WQSEST3V_1 */
MERGE Wqsest3v_1
ME123_0TYP1 (WHERE= (Sampnum=500))
ME123_0PWR (WHERE= (Sampnum=500))
ME123_TTYP1 (WHERE= (Sampnum=500))
ME123_TPWR (WHERE= (Sampnum=500))
ME123_BTYP1 (WHERE= (Sampnum=500))
ME123_BPWR (WHERE= (Sampnum=500))

/************************/
ME1_0TYP1 (WHERE= (Sampnum=500))
ME1T1Z_0TYP1 (WHERE= (Sampnum=500))
ME1GZ_0TYP1 (WHERE= (Sampnum=500))
ME1T1GZ_0TYP1 (WHERE= (Sampnum=500))
ME1_0PWR (WHERE= (Sampnum=500))
ME1_0TYP1_10 (WHERE= (Sampnum=500))
ME1_0PWR_10 (WHERE= (Sampnum=500))
ME1_0TYP1_20 (WHERE= (Sampnum=500))
ME1_0PWR_20 (WHERE= (Sampnum=500))
```

```
ME1_0TYP1_30 (WHERE= (Sampnum=500))
ME1_0PWR_30 (WHERE= (Sampnum=500))
ME1_0TYP1_40 (WHERE= (Sampnum=500))
ME1_0PWR_40 (WHERE= (Sampnum=500))

ME2_0TYP1 (WHERE= (Sampnum=500))
ME2T2Z_0TYP1 (WHERE= (Sampnum=500))
ME2GZ_0TYP1 (WHERE= (Sampnum=500))
ME2T2GZ_0TYP1 (WHERE= (Sampnum=500))
ME2_0PWR (WHERE= (Sampnum=500))
ME2_0TYP1_10 (WHERE= (Sampnum=500))
ME2_0PWR_10 (WHERE= (Sampnum=500))
ME2_0TYP1_20 (WHERE= (Sampnum=500))
ME2_0PWR_20 (WHERE= (Sampnum=500))
ME2_0TYP1_30 (WHERE= (Sampnum=500))
ME2_0PWR_30 (WHERE= (Sampnum=500))
ME2_0TYP1_40 (WHERE= (Sampnum=500))
ME2_0PWR_40 (WHERE= (Sampnum=500))

ME3_0TYP1 (WHERE= (Sampnum=500))
ME3T3Z_0TYP1 (WHERE= (Sampnum=500))
ME3GZ_0TYP1 (WHERE= (Sampnum=500))
ME3T3GZ_0TYP1 (WHERE= (Sampnum=500))
ME3_0PWR (WHERE= (Sampnum=500))
ME3_0TYP1_10 (WHERE= (Sampnum=500))
ME3_0PWR_10 (WHERE= (Sampnum=500))
ME3_0TYP1_20 (WHERE= (Sampnum=500))
ME3_0PWR_20 (WHERE= (Sampnum=500))
ME3_0TYP1_30 (WHERE= (Sampnum=500))
ME3_0PWR_30 (WHERE= (Sampnum=500))
ME3_0TYP1_40 (WHERE= (Sampnum=500))
ME3_0PWR_40 (WHERE= (Sampnum=500))
/************************/


ME1_TTYP1 (WHERE= (Sampnum=500))
ME1T1Z_TTYP1 (WHERE= (Sampnum=500))
ME1GZ_TTYP1 (WHERE= (Sampnum=500))
ME1T1GZ_TTYP1 (WHERE= (Sampnum=500))
ME1_TPWR (WHERE= (Sampnum=500))
```

```
ME1_TTYP1_10 (WHERE= (Sampnum=500))
ME1_TPWR_10 (WHERE= (Sampnum=500))
ME1_TTYP1_20 (WHERE= (Sampnum=500))
ME1_TPWR_20 (WHERE= (Sampnum=500))
ME1_TTYP1_30 (WHERE= (Sampnum=500))
ME1_TPWR_30 (WHERE= (Sampnum=500))
ME1_TTYP1_40 (WHERE= (Sampnum=500))
ME1_TPWR_40 (WHERE=(Sampnum=500))

ME2_TTYP1 (WHERE= (Sampnum=500))
ME2T2Z_TTYP1 (WHERE= (Sampnum=500))
ME2GZ_TTYP1 (WHERE= (Sampnum=500))
ME2T2GZ_TTYP1 (WHERE= (Sampnum=500))
ME2_TPWR (WHERE= (Sampnum=500))
ME2_TTYP1_10 (WHERE= (Sampnum=500))
ME2_TPWR_10 (WHERE= (Sampnum=500))
ME2_TTYP1_20 (WHERE= (Sampnum=500))
ME2_TPWR_20 (WHERE= (Sampnum=500))
ME2_TTYP1_30 (WHERE= (Sampnum=500))
ME2_TPWR_30 (WHERE= (Sampnum=500))
ME2_TTYP1_40 (WHERE= (Sampnum=500))
ME2_TPWR_40 (WHERE= (Sampnum=500))

ME3_TTYP1 (WHERE= (Sampnum=500))
ME3T3Z_TTYP1 (WHERE= (Sampnum=500))
ME3GZ_TTYP1 (WHERE= (Sampnum=500))
ME3T3GZ_TTYP1 (WHERE= (Sampnum=500))
ME3_TPWR (WHERE= (Sampnum=500))
ME3_TTYP1_10 (WHERE= (Sampnum=500))
ME3_TPWR_10 (WHERE= (Sampnum=500))
ME3_TTYP1_20 (WHERE= (Sampnum=500))
ME3_TPWR_20 (WHERE= (Sampnum=500))
ME3_TTYP1_30 (WHERE= (Sampnum=500))
ME3_TPWR_30 (WHERE= (Sampnum=500))
ME3_TTYP1_40 (WHERE= (Sampnum=500))
ME3_TPWR_40 (WHERE= (Sampnum=500))

/************************/
ME1_BTYP1 (WHERE= (Sampnum=500))
ME1T1Z_BTYP1 (WHERE= (Sampnum=500))
```

```
ME1GZ_BTYP1 (WHERE= (Sampnum=500))
ME1T1GZ_BTYP1 (WHERE= (Sampnum=500))
ME1_BPWR (WHERE= (Sampnum=500))
ME1_BTYP1_10 (WHERE= (Sampnum=500))
ME1_BPWR_10 (WHERE= (Sampnum=500))
ME1_BTYP1_20 (WHERE= (Sampnum=500))
ME1_BPWR_20 (WHERE= (Sampnum=500))
ME1_BTYP1_30 (WHERE= (Sampnum=500))
ME1_BPWR_30 (WHERE= (Sampnum=500))
ME1_BTYP1_40 (WHERE= (Sampnum=500))
ME1_BPWR_40 (WHERE= (Sampnum=500))

ME2_BTYP1 (WHERE= (Sampnum=500))
ME2T2Z_BTYP1 (WHERE= (Sampnum=500))
ME2GZ_BTYP1 (WHERE= (Sampnum=500))
ME2T2GZ_BTYP1 (WHERE Sampnum=500))
ME2_BPWR (WHERE= (Sampnum=500))
ME2_BTYP1_10 (WHERE= (Sampnum=500))
ME2_BPWR_10 (WHERE= (Sampnum=500))
ME2_BTYP1_20 (WHERE= (Sampnum=500))
ME2_BPWR_20 (WHERE= (Sampnum=500))
ME2_BTYP1_30 (WHERE= (Sampnum=500))
ME2_BPWR_30 (WHERE= (Sampnum=500))
ME2_BTYP1_40 (WHERE= (Sampnum=500))
ME2_BPWR_40 (WHERE= (Sampnum=500))

ME3_BTYP1 (WHERE= (Sampnum=500))
ME3T3Z_BTYP1 (WHERE= (Sampnum=500))
ME3GZ_BTYP1 (WHERE= (Sampnum=500))
ME3T3GZ_BTYP1 (WHERE= (Sampnum=500))
ME3_BPWR (WHERE= (Sampnum=500))
ME3_BTYP1_10 (WHERE= (Sampnum=500))
ME3_BPWR_10 (WHERE= (Sampnum=500))
ME3_BTYP1_20 (WHERE= (Sampnum=500))
ME3_BPWR_20 (WHERE= (Sampnum=500))
ME3_BTYP1_30 (WHERE= (Sampnum=500))
ME3_BPWR_30 (WHERE= (Sampnum=500))
ME3_BTYP1_40 (WHERE= (Sampnum=500))
ME3_BPWR_40 (WHERE= (Sampnum=500))
/*************************/
```

```
DE123_0TYP1 (WHERE= (Sampnum=500))
DE123_0PWR (WHERE= (Sampnum=500))
DE123_TTYP1 (WHERE= (Sampnum=500))
DE123_TPWR (WHERE= (Sampnum=500))
DE123_BTYP1 (WHERE= (Sampnum=500))
DE123_BPWR (WHERE= (Sampnum=500))
/*************************/
DE1_0TYP1 (WHERE= (Sampnum=500))
DE1_0TYP1_10 (WHERE= (Sampnum=500))
DE1_0TYP1_20 (WHERE= (Sampnum=500))
DE1_0TYP1_30 (WHERE= (Sampnum=500))
DE1_0TYP1_40 (WHERE= (Sampnum=500))

DE1_0PWR (WHERE= (Sampnum=500))
DE1_0PWR_10 (WHERE= (Sampnum=500))
DE1_0PWR_20 (WHERE= (Sampnum=500))
DE1_0PWR_30 (WHERE= (Sampnum=500))
DE1_0PWR_40 (WHERE= (Sampnum=500))

DE2_0TYP1 (WHERE= (Sampnum=500))
DE2_0TYP1_10 (WHERE= (Sampnum=500))
DE2_0TYP1_20 (WHERE= (Sampnum=500))
DE2_0TYP1_30 (WHERE= (Sampnum=500))
DE2_0TYP1_40 (WHERE= (Sampnum=500))

DE2_0PWR (WHERE= (Sampnum=500))
DE2_0PWR_10 (WHERE= (Sampnum=500))
DE2_0PWR_20 (WHERE= (Sampnum=500))
DE2_0PWR_30 (WHERE= (Sampnum=500))
DE2_0PWR_40 (WHERE= (Sampnum=500))

DE3_0TYP1 (WHERE= (Sampnum=500))
DE3_0TYP1_10 (WHERE= (Sampnum=500))
DE3_0TYP1_20 (WHERE= (Sampnum=500))
DE3_0TYP1_30 (WHERE= (Sampnum=500))
DE3_0TYP1_40 (WHERE= (Sampnum=500))

DE3_0PWR (WHERE= (Sampnum=500))
DE3_0PWR_10 (WHERE= (Sampnum=500))
```

```
DE3_0PWR_20 (WHERE= (Sampnum=500))
DE3_0PWR_30 (WHERE= (Sampnum=500))
DE3_0PWR_40 (WHERE= (Sampnum=500))
/****************************/
DE1_TTYP1 (WHERE= (Sampnum=500))
DE1_TTYP1_10 (WHERE= (Sampnum=500))
DE1_TTYP1_20 (WHERE= (Sampnum=500))
DE1_TTYP1_30 (WHERE= (Sampnum=500))
DE1_TTYP1_40 (WHERE= (Sampnum=500))

DE1_TPWR (WHERE= (Sampnum=500))
DE1_TPWR_10 (WHERE= (Sampnum=500))
DE1_TPWR_20 (WHERE= (Sampnum=500))
DE1_TPWR_30 (WHERE= (Sampnum=500))
DE1_TPWR_40 (WHERE= (Sampnum=500))

DE2_TTYP1 (WHERE= (Sampnum=500))
DE2_TTYP1_10 (WHERE= (Sampnum=500))
DE2_TTYP1_20 (WHERE= (Sampnum=500))
DE2_TTYP1_30 (WHERE= (Sampnum=500))
DE2_TTYP1_40 (WHERE= (Sampnum=500))

DE2_TPWR (WHERE= (Sampnum=500))
DE2_TPWR_10 (WHERE= (Sampnum=500))
DE2_TPWR_20 (WHERE= (Sampnum=500))
DE2_TPWR_30 (WHERE= (Sampnum=500))
DE2_TPWR_40 (WHERE= (Sampnum=500))

DE3_TTYP1 (WHERE= (Sampnum=500))
DE3_TTYP1_10 (WHERE= (Sampnum=500))
DE3_TTYP1_20 (WHERE= (Sampnum=500))
DE3_TTYP1_30 (WHERE= (Sampnum=500))
DE3_TTYP1_40 (WHERE= (Sampnum=500))

DE3_TPWR (WHERE= (Sampnum=500))
DE3_TPWR_10 (WHERE= (Sampnum=500))
DE3_TPWR_20 (WHERE= (Sampnum=500))
DE3_TPWR_30 (WHERE= (Sampnum=500))
DE3_TPWR_40 (WHERE= (Sampnum=500))
/*************************/
```

```
DE1_BTYP1 (WHERE= (Sampnum=500))
DE1_BTYP1_10 (WHERE= (Sampnum=500))
DE1_BTYP1_20 (WHERE= (Sampnum=500))
DE1_BTYP1_30 (WHERE= (Sampnum=500))
DE1_BTYP1_40 (WHERE= (Sampnum=500))

DE1_BPWR (WHERE= (Sampnum=500))
DE1_BPWR_10 (WHERE= (Sampnum=500))
DE1_BPWR_20 (WHERE= (Sampnum=500))
DE1_BPWR_30 (WHERE= (Sampnum=500))
DE1_BPWR_40 (WHERE= (Sampnum=500))

DE2_BTYP1 (WHERE= (Sampnum=500))
DE2_BTYP1_10 (WHERE= (Sampnum=500))
DE2_BTYP1_20 (WHERE= (Sampnum=500))
DE2_BTYP1_30 (WHERE= (Sampnum=500))
DE2_BTYP1_40 (WHERE= (Sampnum=500))

DE2_BPWR (WHERE= (Sampnum=500))
DE2_BPWR_10 (WHERE= (Sampnum=500))
DE2_BPWR_20 (WHERE= (Sampnum=500))
DE2_BPWR_30 (WHERE= (Sampnum=500))
DE2_BPWR_40 (WHERE= (Sampnum=500))

DE3_BTYP1 (WHERE= (Sampnum=500))
DE3_BTYP1_10 (WHERE= (Sampnum=500))
DE3_BTYP1_20 (WHERE= (Sampnum=500))
DE3_BTYP1_30 (WHERE= (Sampnum=500))
DE3_BTYP1_40 (WHERE= (Sampnum=500))

DE3_BPWR (WHERE= (Sampnum=500))
DE3_BPWR_10 (WHERE= (Sampnum=500))
DE3_BPWR_20 (WHERE= (Sampnum=500))
DE3_BPWR_30 (WHERE= (Sampnum=500))
DE3_BPWR_40 (WHERE= (Sampnum=500));
BY Condition;


IF DE123^=0    & DE123_0TYP1 =0     THEN DE123_0TYP1=.;
IF DE123=0     & DE123_0PWR =0      THEN DE123_0PWR=.;
```

```
IF DE123^=0    & DE123_TTYP1 =0    THEN DE123_TTYP1=.;
IF DE123=0     & DE123_TPWR =0     THEN DE123_TPWR=.;
IF DE123^=0    & DE123_BTYP1 =0    THEN DE123_BTYP1=.;
IF DE123=0     & DE123_BPWR =0     THEN DE123_BPWR=.;

IF DE1^=0      & DE1_0TYP1 =0      THEN DE1_0TYP1=.;
IF DE1^=0      & DE1_0TYP1_10 =0   THEN DE1_0TYP1_10=.;
IF DE1^=0      & DE1_0TYP1_20 =0   THEN DE1_0TYP1_20=.;
IF DE1^=0      & DE1_0TYP1_30 =0   THEN DE1_0TYP1_30=.;
IF DE1^=0      & DE1_0TYP1_40 =0   THEN DE1_0TYP1_40=.;

IF DE1=0       & DE1_0PWR =0       THEN DE1_0PWR=.;
IF DE1=0       & DE1_0PWR_10 =0    THEN DE1_0PWR_10=.;
IF DE1=0       & DE1_0PWR_20 =0    THEN DE1_0PWR_20=.;
IF DE1=0       & DE1_0PWR_30 =0    THEN DE1_0PWR_30=.;
IF DE1=0       & DE1_0PWR_40 =0    THEN DE1_0PWR_40=.;

IF DE2^=0      & DE2_0TYP1 =0      THEN DE2_0TYP1=.;
IF DE2^=0      & DE2_0TYP1_10 =0   THEN DE2_0TYP1_10=.;
IF DE2^=0      & DE2_0TYP1_20 =0   THEN DE2_0TYP1_20=.;
IF DE2^=0      & DE2_0TYP1_30 =0   THEN DE2_0TYP1_30=.;
IF DE2^=0      & DE2_0TYP1_40 =0   THEN DE2_0TYP1_40=.;

IF DE2=0       & DE2_0PWR =0       THEN DE2_0PWR=.;
IF DE2=0       & DE2_0PWR_10 =0    THEN DE2_0PWR_10=.;
IF DE2=0       & DE2_0PWR_20 =0    THEN DE2_0PWR_20=.;
IF DE2=0       & DE2_0PWR_30 =0    THEN DE2_0PWR_30=.;
IF DE2=0       & DE2_0PWR_40 =0    THEN DE2_0PWR_40=.;

IF DE3^=0      & DE3_0TYP1 =0      THEN DE3_0TYP1=.;
IF DE3^=0      & DE3_0TYP1_10 =0   THEN DE3_0TYP1_10=.;
IF DE3^=0      & DE3_0TYP1_20 =0   THEN DE3_0TYP1_20=.;
IF DE3^=0      & DE3_0TYP1_30 =0   THEN DE3_0TYP1_30=.;
IF DE3^=0      & DE3_0TYP1_40 =0   THEN DE3_0TYP1_40=.;

IF DE3=0       & DE3_0PWR =0       THEN DE3_0PWR=.;
IF DE3=0       & DE3_0PWR_10 =0    THEN DE3_0PWR_10=.;
IF DE3=0       & DE3_0PWR_20 =0    THEN DE3_0PWR_20=.;
IF DE3=0       & DE3_0PWR_30 =0    THEN DE3_0PWR_30=.;
IF DE3=0       & DE3_0PWR_40 =0    THEN DE3_0PWR_40=.;
```

```
IF DE1^=0      & DE1_TTYP1 =0       THEN DE1_TTYP1=.;
IF DE1^=0      & DE1_TTYP1_10 =0    THEN DE1_TTYP1_10=.;
IF DE1^=0      & DE1_TTYP1_20 =0    THEN DE1_TTYP1_20=.;
IF DE1^=0      & DE1_TTYP1_30 =0    THEN DE1_TTYP1_30=.;
IF DE1^=0      & DE1_TTYP1_40 =0    THEN DE1_TTYP1_40=.;

IF DE1=0       & DE1_TPWR =0        THEN DE1_TPWR=.;
IF DE1=0       & DE1_TPWR_10 =0     THEN DE1_TPWR_10=.;
IF DE1=0       & DE1_TPWR_20 =0     THEN DE1_TPWR_20=.;
IF DE1=0       & DE1_TPWR_30 =0     THEN DE1_TPWR_30=.;
IF DE1=0       & DE1_TPWR_40 =0     THEN DE1_TPWR_40=.;

IF DE2^=0      & DE2_TTYP1 =0       THEN DE2_TTYP1=.;
IF DE2^=0      & DE2_TTYP1_10 =0    THEN DE2_TTYP1_10=.;
IF DE2^=0      & DE2_TTYP1_20 =0    THEN DE2_TTYP1_20=.;
IF DE2^=0      & DE2_TTYP1_30 =0    THEN DE2_TTYP1_30=.;
IF DE2^=0      & DE2_TTYP1_40 =0    THEN DE2_TTYP1_40=.;

IF DE2=0       & DE2_TPWR =0        THEN DE2_TPWR=.;
IF DE2=0       & DE2_TPWR_10 =0     THEN DE2_TPWR_10=.;
IF DE2=0       & DE2_TPWR_20 =0     THEN DE2_TPWR_20=.;
IF DE2=0       & DE2_TPWR_30 =0     THEN DE2_TPWR_30=.;
IF DE2=0       & DE2_TPWR_40 =0     THEN DE2_TPWR_40=.;

IF DE3^=0      & DE3_TTYP1 =0       THEN DE3_TTYP1=.;
IF DE3^=0      & DE3_TTYP1_10 =0    THEN DE3_TTYP1_10=.;
IF DE3^=0      & DE3_TTYP1_20 =0    THEN DE3_TTYP1_20=.;
IF DE3^=0      & DE3_TTYP1_30 =0    THEN DE3_TTYP1_30=.;
IF DE3^=0      & DE3_TTYP1_40 =0    THEN DE3_TTYP1_40=.;

IF DE3=0       & DE3_TPWR =0        THEN DE3_TPWR=.;
IF DE3=0       & DE3_TPWR_10 =0     THEN DE3_TPWR_10=.;
IF DE3=0       & DE3_TPWR_20 =0     THEN DE3_TPWR_20=.;
IF DE3=0       & DE3_TPWR_30 =0     THEN DE3_TPWR_30=.;
IF DE3=0       & DE3_TPWR_40 =0     THEN DE3_TPWR_40=.;

IF DE1^=0      & DE1_BTYP1 =0       THEN DE1_BTYP1=.;
IF DE1^=0      & DE1_BTYP1_10 =0    THEN DE1_BTYP1_10 =.;
IF DE1^=0      & DE1_BTYP1_20 =0    THEN DE1_BTYP1_20 =.;
```

```
IF DE1^=0       & DE1_BTYP1_30 =0    THEN DE1_BTYP1_30 =.;
IF DE1^=0       & DE1_BTYP1_40 =0    THEN DE1_BTYP1_40 =.;

IF DE1=0        & DE1_BPWR   =0      THEN DE1_BPWR =.;
IF DE1=0        & DE1_BPWR_10 =0     THEN DE1_BPWR_10 =.;
IF DE1=0        & DE1_BPWR_20 =0     THEN DE1_BPWR_20 =.;
IF DE1=0        & DE1_BPWR_30 =0     THEN DE1_BPWR_30 =.;
IF DE1=0        & DE1_BPWR_40 =0     THEN DE1_BPWR_40 =.;

IF DE2^=0       & DE2_BTYP1 =0       THEN DE2_BTYP1 =.;
IF DE2^=0       & DE2_BTYP1_10 =0    THEN DE2_BTYP1_10 =.;
IF DE2^=0       & DE2_BTYP1_20 =0    THEN DE2_BTYP1_20 =.;
IF DE2^=0       & DE2_BTYP1_30 =0    THEN DE2_BTYP1_30 =.;
IF DE2^=0       & DE2_BTYP1_40 =0    THEN DE2_BTYP1_40 =.;

IF DE2=0        & DE2_BPWR   =0      THEN DE2_BPWR =.;
IF DE2=0        & DE2_BPWR_10 =0     THEN DE2_BPWR_10 =.;
IF DE2=0        & DE2_BPWR_20 =0     THEN DE2_BPWR_20 =.;
IF DE2=0        & DE2_BPWR_30 =0     THEN DE2_BPWR_30 =.;
IF DE2=0        & DE2_BPWR_40 =0     THEN DE2_BPWR_40 =.;

IF DE3^=0       & DE3_BTYP1 =0       THEN DE3_BTYP1 =.;
IF DE3^=0       & DE3_BTYP1_10 =0    THEN DE3_BTYP1_10 =.;
IF DE3^=0       & DE3_BTYP1_20 =0    THEN DE3_BTYP1_20 =.;
IF DE3^=0       & DE3_BTYP1_30 =0    THEN DE3_BTYP1_30 =.;
IF DE3^=0       & DE3_BTYP1_40 =0    THEN DE3_BTYP1_40 =.;

IF DE3=0        & DE3_BPWR   =0      THEN DE3_BPWR =.;
IF DE3=0        & DE3_BPWR_10 =0     THEN DE3_BPWR_10 =.;
IF DE3=0        & DE3_BPWR_20 =0     THEN DE3_BPWR_20 =.;
IF DE3=0        & DE3_BPWR_30 =0     THEN DE3_BPWR_30 =.;
IF DE3=0        & DE3_BPWR_40 =0     THEN DE3_BPWR_40 =.;

IF ME123^=0     & ME123_0TYP1 =0     THEN ME123_0TYP1 =.;
IF ME123=0      & ME123_0PWR =0      THEN ME123_0PWR =.;
IF ME123^=0     & ME123_TTYP1 =0     THEN ME123_TTYP1 =.;
IF ME123=0      & ME123_TPWR =0      THEN ME123_TPWR =.;
IF ME123^=0     & ME123_BTYP1 =0     THEN ME123_BTYP1 =.;
IF ME123=0      & ME123_BPWR =0      THEN ME123_BPWR =.;
IF ME1^=0       & ME1_0TYP1 =0       THEN ME1_0TYP1 =.;
```

```
/*ME1T1Z_0TYP1        ME1GZ_0TYP1          ME1T1GZ_0TYP1        */

IF ME1^=0       & ME1T1Z_0TYP1 =0   THEN ME1T1Z_0TYP1 =.;
IF ME1^=0       & ME1GZ_0TYP1 =0    THEN ME1GZ_0TYP1 =.;
IF ME1^=0       & ME1T1GZ_0TYP1=0   THEN ME1T1GZ_0TYP1 =.;
IF ME1=0        & ME1_0PWR =0       THEN ME1_0PWR=.;
IF ME1^=0       & ME1_0TYP1_10 =0   THEN ME1_0TYP1_10 =.;
IF ME1=0        & ME1_0PWR_10 =0    THEN ME1_0PWR_10 =.;
IF ME1^=0       & ME1_0TYP1_20 =0   THEN ME1_0TYP1_20 =.;
IF ME1=0        & ME1_0PWR_20 =0    THEN ME1_0PWR_20 =.;
IF ME1^=0       & ME1_0TYP1_30 =0   THEN ME1_0TYP1_30=.;
IF ME1=0        & ME1_0PWR_30 =0    THEN ME1_0PWR_30 =.;
IF ME1^=0       & ME1_0TYP1_40 =0   THEN ME1_0TYP1_40 =.;
IF ME1=0        & ME1_0PWR_40 =0    THEN ME1_0PWR_40 =.;

IF ME2^=0       & ME2_0TYP1 =0      THEN ME2_0TYP1 =.;
/*ME2T2Z_0TYP1        ME2GZ_0TYP1          ME2T2GZ_0TYP1        */
IF ME2^=0       & ME2T2Z_0TYP1=0    THEN ME2T2Z_0TYP1 =.;
IF ME2^=0       & ME2GZ_0TYP1=0     THEN ME2GZ_0TYP1 =.;
IF ME2^=0       & ME2T2GZ_0TYP1=0   THEN ME2T2GZ_0TYP1 =.;
IF ME2=0        & ME2_0PWR =0       THEN ME2_0PWR=.;
IF ME2^=0       & ME2_0TYP1_10 =0   THEN ME2_0TYP1_10 =.;
IF ME2=0        & ME2_0PWR_10 =0    THEN ME2_0PWR_10 =.;
IF ME2^=0       & ME2_0TYP1_20 =0   THEN ME2_0TYP1_20 =.;
IF ME2=0        & ME2_0PWR_20 =0    THEN ME2_0PWR_20 =.;
IF ME2^=0       & ME2_0TYP1_30 =0   THEN ME2_0TYP1_30=.;
IF ME2=0        & ME2_0PWR_30 =0    THEN ME2_0PWR_30 =.;
IF ME2^=0       & ME2_0TYP1_40 =0   THEN ME2_0TYP1_40 =.;
IF ME2=0        & ME2_0PWR_40 =0    THEN ME2_0PWR_40 =.;
IF ME3^=0       & ME3_0TYP1 =0      THEN ME3_0TYP1 =.;
/*ME3T2Z_0TYP1        ME3GZ_0TYP1          ME3T2GZ_0TYP1        */

IF ME3^=0       & ME3T3Z_0TYP1=0    THEN ME3T3Z_0TYP1 =.;
IF ME3^=0       & ME3GZ_0TYP1=0     THEN ME3GZ_0TYP1 =.;
IF ME3^=0       & ME3T3GZ_0TYP1=0   THEN ME3T3GZ_0TYP1 =.;
IF ME3=0        & ME3_0PWR =0       THEN ME3_0PWR=.;
IF ME3^=0       & ME3_0TYP1_10 =0   THEN ME3_0TYP1_10 =.;
IF ME3=0        & ME3_0PWR_10 =0    THEN ME3_0PWR_10 =.;
IF ME3^=0       & ME3_0TYP1_20 =0   THEN ME3_0TYP1_20 =.;
IF ME3=0        & ME3_0PWR_20 =0    THEN ME3_0PWR_20 =.;
```

```
IF ME3^=0       & ME3_0TYP1_30 =0    THEN ME3_0TYP1_30=.;
IF ME3=0        & ME3_0PWR_30 =0     THEN ME3_0PWR_30 =.;
IF ME3^=0       & ME3_0TYP1_40 =0    THEN ME3_0TYP1_40 =.;
IF ME3=0        & ME3_0PWR_40 =0     THEN ME3_0PWR_40 =.;
IF ME1^=0               & ME1_TTYP1 =0 THEN ME1_TTYP1 =.;
/*ME1T1Z_TTYP1          ME1GZ_TTYP1            ME1T1GZ_TTYP1         */

IF ME1^=0       & ME1T1Z_TTYP1=0     THEN ME1T1Z_TTYP1 =.;
IF ME1^=0       & ME1GZ_TTYP1=0      THEN ME1GZ_TTYP1 =.;
IF ME1^=0       & ME1T1GZ_TTYP1=0    THEN ME1T1GZ_TTYP1 =.;
IF ME1=0        & ME1_TPWR =0        THEN ME1_TPWR=.;
IF ME1^=0       & ME1_TTYP1_10 =0    THEN ME1_TTYP1_10 =.;
IF ME1=0        & ME1_TPWR_10 =0     THEN ME1_TPWR_10 =.;
IF ME1^=0       & ME1_TTYP1_20 =0    THEN ME1_TTYP1_20 =.;
IF ME1=0        & ME1_TPWR_20 =0     THEN ME1_TPWR_20 =.;
IF ME1^=0       & ME1_TTYP1_30 =0    THEN ME1_TTYP1_30=.;
IF ME1=0        & ME1_TPWR_30 =0     THEN ME1_TPWR_30 =.;
IF ME1^=0       & ME1_TTYP1_40 =0    THEN ME1_TTYP1_40 =.;
IF ME1=0        & ME1_TPWR_40 =0     THEN ME1_TPWR_40 =.;
IF ME2^=0       & ME2_TTYP1 =0       THEN ME2_TTYP1 =.;
/*ME2T2Z_TTYP1          ME2GZ_TTYP1            ME2T2GZ_TTYP1         */

IF ME2^=0       & ME2T2Z_TTYP1=0     THEN ME2T2Z_TTYP1 =.;
IF ME2^=0       & ME2GZ_TTYP1=0      THEN ME2GZ_TTYP1 =.;
IF ME2^=0       & ME2T2GZ_TTYP1=0    THEN ME2T2GZ_TTYP1 =.;
IF ME2=0        & ME2_TPWR =0        THEN ME2_TPWR=.;
IF ME2^=0       & ME2_TTYP1_10 =0    THEN ME2_TTYP1_10 =.;
IF ME2=0        & ME2_TPWR_10 =0     THEN ME2_TPWR_10 =.;
IF ME2^=0       & ME2_TTYP1_20 =0    THEN ME2_TTYP1_20 =.;
IF ME2=0        & ME2_TPWR_20 =0     THEN ME2_TPWR_20 =.;
IF ME2^=0       & ME2_TTYP1_30 =0    THEN ME2_TTYP1_30=.;
IF ME2=0        & ME2_TPWR_30 =0     THEN ME2_TPWR_30 =.;
IF ME2^=0       & ME2_TTYP1_40 =0    THEN ME2_TTYP1_40 =.;
IF ME2=0        & ME2_TPWR_40 =0     THEN ME2_TPWR_40 =.;
IF ME3^=0       & ME3_TTYP1 =0       THEN ME3_TTYP1 =.;
/*ME3T2Z_TTYP1          ME3GZ_TTYP1            ME3T2GZ_TTYP1         */

IF ME3^=0       & ME3T3Z_TTYP1=0     THEN ME3T3Z_TTYP1 =.;
IF ME3^=0       & ME3GZ_TTYP1=0      THEN ME3GZ_TTYP1 =.;
IF ME3^=0       & ME3T3GZ_TTYP1=0    THEN ME3T3GZ_TTYP1 =.;
```

```
IF ME3=0      & ME3_TPWR =0        THEN ME3_TPWR=.;
IF ME3^=0     & ME3_TTYP1_10 =0    THEN ME3_TTYP1_10 =.;
IF ME3=0      & ME3_TPWR_10 =0     THEN ME3_TPWR_10 =.;
IF ME3^=0     & ME3_TTYP1_20 =0    THEN ME3_TTYP1_20 =.;
IF ME3=0      & ME3_TPWR_20 =0     THEN ME3_TPWR_20 =.;
IF ME3^=0     & ME3_TTYP1_30 =0    THEN ME3_TTYP1_30=.;
IF ME3=0      & ME3_TPWR_30 =0     THEN ME3_TPWR_30 =.;
IF ME3^=0     & ME3_TTYP1_40 =0    THEN ME3_TTYP1_40 =.;
IF ME3=0      & ME3_TPWR_40 =0     THEN ME3_TPWR_40 =.;
IF ME1^=0     & ME1_BTYP1 =0       THEN ME1_BTYP1 =.;
/*ME1T1Z_BTYP1      ME1GZ_BTYP1        ME1T1GZ_BTYP1      */

IF ME1^=0     & ME1T1Z_BTYP1=0     THEN ME1T1Z_BTYP1 =.;
IF ME1^=0     & ME1GZ_BTYP1=0      THEN ME1GZ_BTYP1 =.;
IF ME1^=0     & ME1T1GZ_BTYP1=0    THEN ME1T1GZ_BTYP1 =.;
IF ME1=0      & ME1_BPWR =0        THEN ME1_BPWR=.;
IF ME1^=0     & ME1_BTYP1_10 =0    THEN ME1_BTYP1_10 =.;
IF ME1=0      & ME1_BPWR_10 =0     THEN ME1_BPWR_10 =.;
IF ME1^=0     & ME1_BTYP1_20 =0    THEN ME1_BTYP1_20 =.;
IF ME1=0      & ME1_BPWR_20 =0     THEN ME1_BPWR_20 =.;
IF ME1^=0     & ME1_BTYP1_30 =0    THEN ME1_BTYP1_30=.;
IF ME1=0      & ME1_BPWR_30 =0     THEN ME1_BPWR_30 =.;
IF ME1^=0     & ME1_BTYP1_40 =0    THEN ME1_BTYP1_40 =.;
IF ME1=0      & ME1_BPWR_40 =0     THEN ME1_BPWR_40 =.;
IF ME2^=0     & ME2_BTYP1 =0       THEN ME2_BTYP1 =.;
/*ME2T2Z_BTYP1      ME2GZ_BTYP1        ME2T2GZ_BTYP1      */

IF ME2^=0     & ME2T2Z_BTYP1=0     THEN ME2T2Z_BTYP1 =.;
IF ME2^=0     & ME2GZ_BTYP1=0      THEN ME2GZ_BTYP1 =.;
IF ME2^=0     & ME2T2GZ_BTYP1=0    THEN ME2T2GZ_BTYP1 =.;
IF ME2=0      & ME2_BPWR =0        THEN ME2_BPWR=.;
IF ME2^=0     & ME2_BTYP1_10 =0    THEN ME2_BTYP1_10 =.;
IF ME2=0      & ME2_BPWR_10 =0     THEN ME2_BPWR_10 =.;
IF ME2^=0     & ME2_BTYP1_20 =0    THEN ME2_BTYP1_20 =.;
IF ME2=0      & ME2_BPWR_20 =0     THEN ME2_BPWR_20 =.;
IF ME2^=0     & ME2_BTYP1_30 =0    THEN ME2_BTYP1_30=.;
IF ME2=0      & ME2_BPWR_30 =0     THEN ME2_BPWR_30 =.;
IF ME2^=0     & ME2_BTYP1_40 =0    THEN ME2_BTYP1_40 =.;
IF ME2=0      & ME2_BPWR_40 =0     THEN ME2_BPWR_40 =.;
IF ME3^=0     & ME3_BTYP1 =0       THEN ME3_BTYP1 =.;
```

```
/*ME3T2Z_BTYP1       ME3GZ_BTYP1         ME3T2GZ_BTYP1       */

IF ME3^=0      & ME3T3Z_BTYP1=0    THEN ME3T3Z_BTYP1 =.;
IF ME3^=0      & ME3GZ_BTYP1=0     THEN ME3GZ_BTYP1 =.;
IF ME3^=0      & ME3T3GZ_BTYP1=0 THEN ME3T3GZ_BTYP1 =.;
IF ME3=0       & ME3_BPWR =0       THEN ME3_BPWR=.;
IF ME3^=0      & ME3_BTYP1_10 =0   THEN ME3_BTYP1_10 =.;
IF ME3=0       & ME3_BPWR_10 =0    THEN ME3_BPWR_10 =.;
IF ME3^=0      & ME3_BTYP1_20 =0   THEN ME3_BTYP1_20 =.;
IF ME3=0       & ME3_BPWR_20 =0    THEN ME3_BPWR_20 =.;
IF ME3^=0      & ME3_BTYP1_30 =0   THEN ME3_BTYP1_30=.;
IF ME3=0       & ME3_BPWR_30 =0    THEN ME3_BPWR_30 =.;
IF ME3^=0      & ME3_BTYP1_40 =0   THEN ME3_BTYP1_40 =.;
IF ME3=0       & ME3_BPWR_40 =0    THEN ME3_BPWR_40 =.;

KEEP Condition
DE123_0TYP1     DE123_0PWR      DE123_TTYP1     DE123_TPWR      DE3_BTYP1_40    DE3_BPWR_40
DE123_BTYP1     DE123_BPWR      DE1_0TYP1       DE1_0PWR        DE2_0TYP1       DE2_0PWR
DE3_0TYP1       DE3_0PWR        DE1_0TYP1_10    DE1_0PWR_10     DE2_0TYP1_10    DE2_0PWR_10
DE3_0TYP1_10    DE3_0PWR_10     DE1_0TYP1_20    DE1_0PWR_20     DE2_0TYP1_20    DE2_0PWR_20
DE3_0TYP1_20    DE3_0PWR_20     DE1_0TYP1_30    DE1_0PWR_30     DE2_0TYP1_30    DE2_0PWR_30
DE3_0TYP1_30    DE3_0PWR_30     DE1_0TYP1_40    DE1_0PWR_40     DE2_0TYP1_40    DE2_0PWR_40
DE3_0TYP1_40    DE3_0PWR_40     DE1_BTYP1_40    DE1_BPWR_40     DE2_BTYP1_40    DE2_BPWR_40
DE1_TTYP1       DE1_TPWR        DE2_TTYP1       DE2_TPWR        DE3_TTYP1       DE3_TPWR
DE1_TTYP1_10    DE1_TPWR_10     DE2_TTYP1_10    DE2_TPWR_10     DE3_TTYP1_10    DE3_TPWR_10
DE1_TTYP1_20    DE1_TPWR_20     DE2_TTYP1_20    DE2_TPWR_20     DE3_TTYP1_20    DE3_TPWR_20
DE1_TTYP1_30    DE1_TPWR_30     DE2_TTYP1_30    DE2_TPWR_30     DE3_TTYP1_30    DE3_TPWR_30
DE1_TTYP1_40    DE1_TPWR_40     DE2_TTYP1_40    DE2_TPWR_40     DE3_TTYP1_40    DE3_TPWR_40
DE1_BTYP1       DE1_BPWR        DE2_BTYP1       DE2_BPWR        DE3_BTYP1       DE3_BPWR
DE1_BTYP1_10    DE1_BPWR_10     DE2_BTYP1_10    DE2_BPWR_10     DE3_BTYP1_10    DE3_BPWR_10
DE1_BTYP1_20    DE1_BPWR_20     DE2_BTYP1_20    DE2_BPWR_20     DE3_BTYP1_20    DE3_BPWR_20
DE1_BTYP1_30    DE1_BPWR_30     DE2_BTYP1_30    DE2_BPWR_30     DE3_BTYP1_30    DE3_BPWR_30

ME123_0TYP1     ME123_0PWR      ME123_TTYP1     ME123_TPWR      ME123_BTYP1     ME123_BPWR
ME3_BTYP1_30    ME1T1Z_0TYP1    ME1GZ_0TYP1     ME1T1GZ_0TYP1   ME2T2Z_0TYP1    ME1_TTYP1
ME1T1Z_TTYP1    ME1GZ_TTYP1     ME1T1GZ_TTYP1   ME2_BTYP1_40    ME2_BTYP1_10    ME2_0TYP1
ME1T1Z_BTYP1    ME1GZ_BTYP1     ME1T1GZ_BTYP1   ME2GZ_0TYP1     ME2T2GZ_0TYP1   ME1_BPWR
ME2_0PWR        ME2_TTYP1       ME2T2Z_TTYP1    ME2GZ_TTYP1     ME1_BTYP1_40    ME2_BTYP1
ME2T2GZ_TTYP1   ME2T2Z_BTYP1    ME2GZ_BTYP1     ME2T2GZ_BTYP1   ME1_BPWR_40     ME3_0TYP1
ME3T3Z_0TYP1    ME3GZ_0TYP1     ME3T3GZ_0TYP1   ME2_0TYP1_40    ME3T3Z_TTYP1    ME3_TTYP1
```

```
ME3GZ_TTYP1      ME3T3GZ_TTYP1    ME2_0TYP1_10     ME1_0TYP1_20     ME3T3Z_BTYP1     ME3_BTYP1
ME3GZ_BTYP1      ME3T3GZ_BTYP1    ME1_0TYP1_10     ME1_0PWR_10      ME1_0PWR_20      ME3_BPWR
ME1_0TYP1_30     ME1_0PWR_30      ME1_0TYP1_40     ME1_0PWR_40      ME2_0PWR_10      ME3_TPWR
ME2_0TYP1_20     ME2_0PWR_20      ME2_0TYP1_30     ME2_0PWR_30      ME2_0PWR_40      ME2_TPWR
ME3_0TYP1_10     ME3_0PWR_10      ME3_0TYP1_20     ME3_0PWR_20      ME3_0TYP1_30     ME1_0PWR
ME3_0PWR_30      ME3_0TYP1_40     ME3_0PWR_40      ME1_TTYP1_10     ME1_TPWR_10      ME3_0PWR
ME1_TTYP1_20     ME1_TPWR_20      ME1_TTYP1_30     ME1_TPWR_30      ME1_TTYP1_40     ME1_TPWR_40
ME2_TTYP1_10     ME2_TPWR_10      ME2_TTYP1_20     ME2_TPWR_20      ME2_TTYP1_30     ME2_TPWR_30
ME2_TTYP1_40     ME2_TPWR_40      ME3_TTYP1_10     ME3_TPWR_10      ME3_TTYP1_20     ME3_TPWR_20
ME3_TTYP1_30     ME3_TPWR_30      ME3_TTYP1_40     ME3_TPWR_40      ME1_BTYP1_10     ME2_BPWR
ME1_BPWR_10      ME1_BTYP1_20     ME1_BPWR_20      ME1_BTYP1_30     ME1_BPWR_30      ME1_BTYP1
ME2_BPWR_10      ME2_BTYP1_20     ME2_BPWR_20      ME2_BTYP1_30     ME2_BPWR_30      ME1_TPWR
ME2_BPWR_40      ME3_BTYP1_10     ME3_BPWR_10      ME3_BTYP1_20     ME3_BPWR_20      ME1_0TYP1
ME3_BPWR_30      ME3_BTYP1_40     ME3_BPWR_40;
RUN;
/* 219 VARS RETAINED  in 1  record of all the Type 1 and Power for ME123_0BT ME1_0BT ME2_0BT ME3_0BT for 0.10, 0.20, 0.30 , 0.40
cut-offs  and DE123_0BT DE1_0BT DE2_0BT DE3_0BT  for 0.10, 0.20, 0.30 , 0.40 cut-offs  for a given condition*/

/***Coverage probability only for Joint efffects ME123_0BT and DE123_0BT***/

DATA ME123_0COV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
IF Sampnum=1 THEN CountCV=0;
IF (P01_ME123_2_5<= ME123 <=P01_ME123_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME123_0COV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME123_0COV;
RUN;

DATA ME123_TCOV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PT1_ME123_2_5<= ME123 <=PT1_ME123_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME123_TCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME123_TCOV;
RUN;

DATA ME123_BCOV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
```

```
IF Sampnum=1 THEN CountCV=0;
IF (PB1_ME123_2_5<= ME123 <=PB1_ME123_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME123_BCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME123_BCOV;
RUN;

/**********SEQ 3V DE123  ***************/
DATA DE123_0COV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE123_0LCL<= DE123 <=DE123_0UCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE123_0COV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE123_0COV;
RUN;

DATA DE123_TCOV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE123_TLCL<= DE123 <=DE123_TUCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE123_TCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE123_TCOV;
RUN;

DATA DE123_BCOV (WHERE= (Sampnum=500));
SET WQSEST3v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE123_BLCL<= DE123 <=DE123_BUCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE123_BCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE123_BCOV;
RUN;

DATA WQS3V_COV;
MERGE ME123_0COV   ME123_TCOV   ME123_BCOV          DE123_0COV             DE123_TCOV             DE123_BCOV   ;
BY Condition;
KEEP Condition   ME123_0COV           ME123_TCOV            ME123_BCOV            DE123_0COV   DE123_TCOV    DE123_BCOV;
RUN; /* 6 Vars Single record of Coverage Probabilities for ME123 and DE123 for given Condition*/
```

```
DATA WQSEST3v_1;
MERGE   WQSEST3v_1 WQS3V_TYP1PWR WQS3V_COV;
  BY Condition;
RUN;   /* Has 192 + 24+ 219 =435 variables */

PROC APPEND BASE=WQSEST3V DATA=WQSEST3v_1 FORCE; RUN; /* APPEND DATA to the base file */
DM LOG 'CLEAR';

%END;
%MEND STATS_3V;
%STATS_3V; /* Has 162 records with 354 columns of STATS on 3-variable mediation analysis */

DATA Temp.WQS3V_STATS_162;
SET WQSEST3V;
FORMAT
        DE123_0          DE123_T          DE123_B          ME123_0          ME123_T
        ME123_B          DE123_0Bias      DE123_TBias      DE123_BBias      ME123_0Bias
        ME123_TBias      ME123_BBias      DE123_0RMSE      DE123_TRMSE      DE123_BRMSE
        ME123_0RMSE      ME123_TRMSE      ME123_BRMSE      wXB01            wXB02
        wXB03            wXT1             wXT2             wXT3             wXB1
        wXB2             wXB3             F7.3;
RUN; /* Has 162 records with 354 columns of STATS on 3-variable mediation analysis */
```

SUMMARIZATION CODE

```
/* Standardized WQS 2-ariable statistics Sequential and Nested using 500 Replicated DATASETs (Sampnum 1 to 500) for each condition
(108 conditions )*/

OPTIONS source pagesize=256 linesize=80 nodate replace;
OPTIONS FORMCHAR="|----|+|---+=|-/\<>*";
GOPTIONS device=jpeg gsfname=pic gsfmode=replace colors= (black) htext=3 ftext=swiss hby=1.5;
PATTERN c=green v=solid;
*libname Temp "/home/evanibm/M_A/20160929"; /* RUN on the cluster */
*ODS noresults; ODS graphics off; ODS html close; *ods select none; ods listing;

DATA WQSEST2V;   /* Base file 358 columns*/
Condition=.;  N=.; Rho_X12=.; DE12=.;  DE1=.;  DE2=.; Theta1=.;  Theta2=.; T12=.;  Gamma=.;  ME12=.; ME1=.; ME2=.;
wXB01_S=.;  wXB01_N=.;   wXB01N_SE=.;  wXB02_S=.;  wXB02_N=.;  wXB02N_SE=.;
wXT1_S=.;   wXT1_N=.;    wXT1N_SE=.;   wXT2_S=.;   wXT2_N=.;   wXT2N_SE=.;
wXB1_S=.;   wXB1_N=.;   wXB1N_SE=.;   wXB2_S=.;   wXB2_N=.;   wXB2N_SE=.;

DE12R120=Rho_X12;
DE12_0 =.; DE12_0Bias =.; DE12_0RMSE =.;  DE12_0COV =.; DE012=.;  DE12_0TYP1 =.; DE12_0PWR =.;

DE1R12_0=Rho_X12;  DE1_0TYP1 =.;  _0TYP1_10 =.;  DE1_0TYP1_20 =.;  DE1_0TYP1_30 =.; DE1_0TYP1_40 =.; DE1_0TYP1_50 =.;
DE1_0PWR =.;  DE1_0PWR_10 =.;  DE1_0PWR_20 =.;  DE1_0PWR_30 =.;  DE1_0PWR_40 =.; DE1_0PWR_50 =.;

DE2R12_0=Rho_X12; DE2_0TYP1 =.; DE2_0TYP1_10 =.; DE2_0TYP1_20 =.; DE2_0TYP1_30 =.; DE2_0TYP1_40 =.; DE2_0TYP1_50 =.;
DE2_0PWR =.;  DE2_0PWR_10 =.; DE2_0PWR_20 =.; DE2_0PWR_30 =.; DE2_0PWR_40 =.; DE2_0PWR_50 =.;

DE12R12T=Rho_X12; DE12_T =.; DE12_TBias =.; DE12_TRMSE =.;  DE12_TCOV =.;   DET12=.;  DE12_TTYP1 =.; DE12_TPWR =.;

DE1R12_T=Rho_X12; DE1_TTYP1 =.; DE1_TTYP1_10 =.; DE1_TTYP1_20 =.; DE1_TTYP1_30 =.; DE1_TTYP1_40 =.; DE1_TTYP1_50 =.;
DE1_TPWR =.; DE1_TPWR_10 =.; DE1_TPWR_20 =.; DE1_TPWR_30 =.; DE1_TPWR_40 =.; DE1_TPWR_50 =.;

DE2R12_T=Rho_X12; DE2_TTYP1 =.; DE2_TTYP1_10 =.; DE2_TTYP1_20 =.; DE2_TTYP1_30 =.; DE2_TTYP1_40 =.; DE2_TTYP1_50 =.;
DE2_TPWR =.; DE2_TPWR_10 =.; DE2_TPWR_20 =.; DE2_TPWR_30 =.; DE2_TPWR_40 =.; DE2_TPWR_50 =.;

DE12R12B=Rho_X12; DE12_B =.; DE12_BBias =.; DE12_BRMSE =.; DE12_BCOV =.; DEB12=.;  DE12_BTYP1 =.; DE12_BPWR =.;

DE1R12_B=Rho_X12; DE1_BTYP1 =.; DE1_BTYP1_10 =.; DE1_BTYP1_20 =.; DE1_BTYP1_30 =.; DE1_BTYP1_40 =.; DE1_BTYP1_50 =.;
DE1_BPWR =.; DE1_BPWR_10 =.; DE1_BPWR_20 =.; DE1_BPWR_30 =.; DE1_BPWR_40 =.; DE1_BPWR_50 =.;

DE2R12_B=Rho_X12;
```

DE2_BTYP1 =.; DE2_BTYP1_10 =.; DE2_BTYP1_20 =.; DE2_BTYP1_30 =.; DE2_BTYP1_40 =.; DE2_BTYP1_50 =.; DE2_BPWR =.;
DE2_BPWR_10 =.; DE2_BPWR_20 =.; DE2_BPWR_30 =.; DE2_BPWR_40 =.; DE2_BPWR_50 =.;

ME12R120=Rho_X12; T12_0=.; G_0=.;
ME12_0 =.; ME12_0Bias =.; ME12_0RMSE =.; ME12S_0COV =.; ME12N_0COV =.; ME012=.; ME12S_0TYP1 =.; ME12N_0TYP1 =.;
ME12S_0PWR =.; ME12N_0PWR =.;

ME1R12_0=Rho_X12; ME1S_0TYP1=.; ME1ST1Z_0TYP1=.; ME1SGZ_0TYP1=.; ME1ST1GZ_0TYP1=.; ME1N_0TYP1 =.;
ME1NT1Z_0TYP1=.; ME1NGZ_0TYP1=.; ME1NT1GZ_0TYP1=.;
ME1S_0TYP1_10 =.; ME1N_0TYP1_10 =.; ME1S_0TYP1_20 =.; ME1N_0TYP1_20 =.;
ME1S_0TYP1_30 =.; ME1N_0TYP1_30 =.; ME1S_0TYP1_40 =.; ME1N_0TYP1_40 =.;
ME1S_0TYP1_50 =.; ME1N_0TYP1_50 =.; ME1S_0PWR =.; ME1N_0PWR =.; ME1S_0PWR_10 =.; ME1N_0PWR_10 =.;

ME1S_0PWR_20 =.; ME1N_0PWR_20 =.; ME1S_0PWR_30 =.; ME1N_0PWR_30 =.;
ME1S_0PWR_40 =.; ME1N_0PWR_40 =.; ME1S_0PWR_50 =.; ME1N_0PWR_50 =.;

ME2R12_0=Rho_X12; ME2S_0TYP1=.; ME2ST2Z_0TYP1=.; ME2SGZ_0TYP1=.; ME2ST2GZ_0TYP1=.; ME2N_0TYP1 =.;
ME2NT2Z_0TYP1=.; ME2NGZ_0TYP1=.; ME2NT2GZ_0TYP1=.; ME2S_0TYP1_10 =.; ME2N_0TYP1_10 =.; ME2S_0TYP1_20 =.;
ME2N_0TYP1_20 =.;
ME2S_0TYP1_30 =.; ME2N_0TYP1_30 =.; ME2S_0TYP1_40 =.; ME2N_0TYP1_40 =.;
ME2S_0TYP1_50 =.; ME2N_0TYP1_50 =.; ME2S_0PWR =.; ME2N_0PWR =.; ME2S_0PWR_10 =.; ME2N_0PWR_10 =.; ME2S_0PWR_20
=.; ME2N_0PWR_20 =.; ME2S_0PWR_30 =.; ME2N_0PWR_30 =.; ME2S_0PWR_40 =.; ME2N_0PWR_40 =.; ME2S_0PWR_50 =.;
ME2N_0PWR_50 =.;

ME12R12T=Rho_X12; T12_T=.; G_T=.;
ME12_T =.; ME12_TBias =.; ME12_TRMSE =.; ME12S_TCOV =.; ME12N_TCOV =.;
MET12=.; ME12S_TTYP1 =.; ME12N_TTYP1 =.; ME12S_TPWR =.; ME12N_TPWR =.;

ME1R12_T=Rho_X12; ME1S_TTYP1=.; ME1ST1Z_TTYP1=.; ME1SGZ_TTYP1=.; ME1ST1GZ_TTYP1=.; ME1N_TTYP1 =.;
ME1NT1Z_TTYP1=.; ME1NGZ_TTYP1=.; ME1NT1GZ_TTYP1=.;
ME1S_TTYP1_10 =.; ME1N_TTYP1_10 =.; ME1S_TTYP1_20 =.; ME1N_TTYP1_20 =.;
ME1S_TTYP1_30 =.; ME1N_TTYP1_30 =.; ME1S_TTYP1_40 =.; ME1N_TTYP1_40 =.;
ME1S_TTYP1_50 =.; ME1N_TTYP1_50 =.; ME1S_TPWR =.; ME1N_TPWR =.; ME1S_TPWR_10 =.; ME1N_TPWR_10 =.; ME1S_TPWR_20
=.; ME1N_TPWR_20 =.; ME1S_TPWR_30 =.; ME1N_TPWR_30 =.; ME1S_TPWR_40 =.; ME1N_TPWR_40 =.; ME1S_TPWR_50 =.;
ME1N_TPWR_50 =.;

ME2R12_T=Rho_X12; ME2S_TTYP1=.; ME2ST2Z_TTYP1=.; ME2SGZ_TTYP1=.; ME2ST2GZ_TTYP1=.; ME2N_TTYP1 =.;
ME2NT2Z_TTYP1=.; ME2NGZ_TTYP1=.; ME2NT2GZ_TTYP1=.; ME2S_TTYP1_10 =.; ME2N_TTYP1_10 =.; ME2S_TTYP1_20 =.;
ME2N_TTYP1_20 =.;
ME2S_TTYP1_30 =.; ME2N_TTYP1_30 =.; ME2S_TTYP1_40 =.; ME2N_TTYP1_40 =.;

```
ME2S_TTYP1_50 =.; ME2N_TTYP1_50 =.; ME2S_TPWR =.; ME2N_TPWR =.; ME2S_TPWR_10 =.; ME2N_TPWR_10 =.; ME2S_TPWR_20
=.; ME2N_TPWR_20 =.; ME2S_TPWR_30 =.; ME2N_TPWR_30 =.; ME2S_TPWR_40 =.; ME2N_TPWR_40 =.; ME2S_TPWR_50 =.;
ME2N_TPWR_50 =.; ME12R12B=Rho_X12; T12_B=.;  G_B=.; ME12_B =.; ME12_BBias =.; ME12_BRMSE =.; ME12S_BCOV =.;
ME12N_BCOV =.; MEB12=.;  ME12S_BTYP1 =.; ME12N_BTYP1 =.; ME12S_BPWR =.; ME12N_BPWR =.;

ME1R12_B=Rho_X12; ME1S_BTYP1=.; ME1ST1Z_BTYP1=.; ME1SGZ_BTYP1=.; ME1ST1GZ_BTYP1=.;  ME1N_BTYP1 =.;
ME1NT1Z_BTYP1=.; ME1NGZ_BTYP1=.; ME1NT1GZ_BTYP1=.;
ME1S_BTYP1_10 =.; ME1N_BTYP1_10 =.; ME1S_BTYP1_20 =.; ME1N_BTYP1_20 =.;
ME1S_BTYP1_30 =.; ME1N_BTYP1_30 =.;       ME1S_BTYP1_40 =.; ME1N_BTYP1_40 =.;
ME1S_BTYP1_50 =.; ME1N_BTYP1_50 =.; ME1S_BPWR =.; ME1N_BPWR =.; ME1S_BPWR_10 =.; ME1N_BPWR_10 =.; ME1S_BPWR_20
=.; ME1N_BPWR_20 =.; ME1S_BPWR_30 =.; ME1N_BPWR_30 =.; ME1S_BPWR_40 =.; ME1N_BPWR_40 =.; ME1S_BPWR_50 =.;
ME1N_BPWR_50 =.;

ME2R12_B=Rho_X12; ME2S_BTYP1=.; ME2ST2Z_BTYP1=.; ME2SGZ_BTYP1=.; ME2ST2GZ_BTYP1=.;  ME2N_BTYP1 =.;
ME2NT2Z_BTYP1=.; ME2NGZ_BTYP1=.; ME2NT2GZ_BTYP1=.;
ME2S_BTYP1_10 =.; ME2N_BTYP1_10 =.; ME2S_BTYP1_20 =.; ME2N_BTYP1_20 =.; ME2S_BTYP1_30 =.; ME2N_BTYP1_30 =.;
ME2S_BTYP1_40 =.; ME2N_BTYP1_40 =.;
ME2S_BTYP1_50 =.; ME2N_BTYP1_50 =.; ME2S_BPWR =.; ME2N_BPWR =.; ME2S_BPWR_10 =.; ME2N_BPWR_10 =.; ME2S_BPWR_20
=.; ME2N_BPWR_20 =.; ME2S_BPWR_30 =.; ME2N_BPWR_30 =.; ME2S_BPWR_40 =.; ME2N_BPWR_40 =.; ME2S_BPWR_50 =.;
ME2N_BPWR_50 =.;

LABEL
        DE012=DE12                  DET12=DE12                  DEB12=DE12
        ME012=ME12                  MET12=ME12                  MEB12=ME12
        DE12R120 = Rho_12           DE1R12_0 = Rho_12           DE2R12_0 = Rho_12
        DE1R12_T = Rho_12           DE2R12_T = Rho_12           DE12R12T = Rho_12
        DE12R12B = Rho_12           DE1R12_B = Rho_12           DE2R12_B = Rho_12

        ME12R120 = Rho_12           ME1R12_0 = Rho_12           ME2R12_0 = Rho_12
        ME12R12T = Rho_12           ME1R12_T = Rho_12           ME2R12_T = Rho_12
        ME12R12B = Rho_12           ME1R12_B = Rho_12           ME2R12_B = Rho_12;

FORMAT
        DE12_0                      DE12_0Bias                  DE12_0RMSE
        ME12_0                      ME12_0Bias                  ME12_0RMSE
        DE12_T                      DE12_TBias                  DE12_TRMSE
        ME12_T                      ME12_TBias                  ME12_TRMSE
        DE12_B                      DE12_BBias                  DE12_BRMSE
        ME12_B                      ME12_BBias                  ME12_BRMSE   F7.3;
```

```
FORMAT        wXB01_S       wXB01_N       wXB01N_SE     wXB02_S       wXB02_N       wXB02N_SE   wXT1_S
              wXT1_N        wXT1N_SE       wXT2_S       wXT2_N        wXT2N_SE      wXB1_S      wXB1_N
              wXB1N_SE      wXB2_S        wXB2_N        wXB2N_SE      T12_0         T12_T        T12_B
              G_0           G_T           G_B           F7.3;

RUN; /* 54 (18 *3 EST) +(87 DE +210 ME) CI STATS+ 36+12 Extra variables with parameters  =429 */

DATA WQSEST2V;
 SET WQSEST2V (OBS=0);
 RUN;

%macro STATS_2V; /*Temp.WQS2v_STATS_108SN;*/

%do Cond=1 %to 108;

DATA WQS2v500_SN1;/* 500 records for each Condition 1 to 108, each related to a Sampnum */

        MERGE Temp.Cond108 (WHERE= (Condition=&Cond)) Temp.WQS2v500_SN&Cond;
                      BY Condition;
                      DROP Beta1 Beta2;
RUN; /* Has 126 variables done + 12 +Sampnum = 138 columns  in 500 rows */


DATA WQSEST2v_1;   /* Single record for one Sampnum for given condition */
SET     WQS2v500_SN1 (WHERE= (Sampnum=1));

              ME012=ME12;                 MET12=ME12;                   MEB12=ME12;
              DE012=DE12;                 DET12=DE12;                   DEB12=DE12;
              DE12R120 = Rho_X12;         DE1R12_0 = Rho_X12;           DE2R12_0 = Rho_X12;
              DE12R12T = Rho_X12;         DE1R12_T = Rho_X12;           DE2R12_T = Rho_X12;
              DE12R12B = Rho_X12;         DE1R12_B = Rho_X12;           DE2R12_B = Rho_X12;
              ME12R120 = Rho_X12;         ME1R12_0 = Rho_X12;           ME2R12_0 = Rho_X12;
              ME12R12T = Rho_X12;         ME1R12_T = Rho_X12;           ME2R12_T = Rho_X12;
              ME12R12B = Rho_X12;         ME1R12_B = Rho_x12;           ME2R12_B = Rho_X12;
              T12=Theta1+Theta2;

KEEP          Condition     N             Rho_X12       DE1 DE2       DE12          Theta1        Theta2
              T12           Gamma         ME1           ME2           ME12          ME012
              MET12         MEB12         DE012         DET12         DEB12         DE12R120      DE12R12T
```

```
                DE12R12B        DE1R12_0    DE2R12_0    DE1R12_T    DE2R12_T    DE1R12_B    DE2R12_B
                           ME12R120    ME12R12T    ME12R12B    ME1R12_0    ME2R12_0
                ME1R12_T        ME2R12_T    ME1R12_B    ME2R12_B    wXB01_S     wXB01_N
                wXB01N_SE       wXB02_S     wXB02_N     wXB02N_SE   wXT1_S      wXT1_N
                wXT1N_SE        wXT2_S      wXT2_N      wXT2N_SE    wXB1_S      wXB1_N
                wXB1N_SE        wXB2_S      wXB2_N      wXB2N_SE;

RUN; /*18 +12 VARS from WQS2v500_SN1 +36 NEW =66 Variables EST*/


PROC UNIVARIATE DATA=WQS2v500_SN1; /*139 columns in 500 rows for given Condition */
VAR     DE12_0      DE12_T      DE12_B      ME12_0      ME12_T      ME12_B
        T12_0       G_0         T12_T       G_T         T12_B       G_B;
BY Condition;
OUTPUT OUT= WQS2v_AVGEST                MEAN=
                /* Average the 500 Replicate estimates */
AVG_DE12_0      AVG_DE12_T          AVG_DE12_B
AVG_ME12_0      AVG_ME12_T          AVG_ME12_B
AVG_T12_0       AVG_G_0             AVG_T12_T
AVG_G_T         AVG_T12_B           AVG_G_B;

RUN;   /* 12 Avgs +Condition in a Single record of EST Averages for Bias Calculations */

DATA        WQSEST2v_AVGEST;        /* Single record   of 78 variables*/
MERGE       WQSEST2v_1   WQS2v_AVGEST;
BY Condition;                           /*WQSEST2v_1 48+18 +12 VARS   WS2v_AVGEST   30+1VARS */

DE12_0=AVG_DE12_0;      DE12_T=AVG_DE12_T;      DE12_B=AVG_DE12_B;      ME12_0=AVG_ME12_0;
ME12_T=AVG_ME12_T;      ME12_B=AVG_ME12_B;      T12_0=AVG_T12_0;        G_0=AVG_G_0;
T12_T=AVG_T12_T;        G_T=AVG_G_T;            T12_B=AVG_T12_B;        G_B=AVG_G_B ;

DROP
AVG_DE12_0   AVG_DE12_T   AVG_DE12_B   AVG_ME12_0   AVG_ME12_T   AVG_ME12_B
AVG_T12_0    AVG_G_0      AVG_T12_T    AVG_G_T      AVG_T12_B    AVG_G_B;
RUN;
/* has 67 columns 12 EST averages 80columns from 500 Sampnums into one record */


DATA WQSEST2v_1;    /*48 +12   VARS from WQS2v500_SN1 +36 NEW =96 Variables EST*/
```

```
MERGE WQSEST2v_1 WQSEST2v_AVGEST (KEEP=Condition
        DE12_0          DE12_T          DE12_B          ME12_0          ME12_T    ME12_B      T12_0
        G_0             T12_T           G_T             T12_B           G_B);
        BY Condition;
        RUN; /* this step adds the Average of 500 estimates to 30 variables */

DATA WQS2v_BiasSQE;    /* 205 Variables in 500 Replicated sample DATA */
MERGE   WQS2v500_SN1        WQS2v_AVGEST;    /* ADD 30 AVGs*/
BY Condition; /*WQS2v500_SN1 126 variables done +13 = 139 columns in 500 rows */

DE12_0Bias=AVG_DE12_0 - DE12;           DE12_TBias=AVG_DE12_T - DE12;           DE12_BBias=AVG_DE12_B - DE12;
ME12_0Bias=AVG_ME12_0 - ME12;           ME12_TBias=AVG_ME12_T - ME12;           ME12_BBias=AVG_ME12_B - ME12;
DE12_0SQE=(DE12_0 - DE12)**2;           DE12_TSQE=   (DE12_T- DE12)**2;         DE12_BSQE=(DE12_B  - DE12)**2;
ME12_0SQE=(ME12_0  - ME12)**2;          ME12_TSQE=(ME12_T - ME12)**2 ;          ME12_BSQE=(ME12_B  - ME12)**2 ;
RUN; /* 12 NEW +151 =163 variables with 500 rows for EST CL|PCL Bias and SQE */


PROC UNIVARIATE DATA=WQS2v_BiasSQE;

VAR     DE12_0Bias          DE12_TBias          DE12_BBias
        ME12_0Bias          ME12_TBias          ME12_BBias
        DE12_0SQE           DE12_TSQE           DE12_BSQE
        ME12_0SQE           ME12_TSQE           ME12_BSQE;
  BY Condition;

OUTPUT OUT=WQS2v_AVGMS (KEEP=  Condition
        DE12_0Bias          DE12_TBias          DE12_BBias
        ME12_0Bias          ME12_TBias          ME12_BBias          DE12_0MSE           DE12_TMSE
        DE12_BMSE           ME12_0MSE           ME12_TMSE           ME12_BMSE)

MEAN=
        DE12_0Bias          DE12_TBias          DE12_BBias
        ME12_0Bias          ME12_TBias          ME12_BBias
        DE12_0MSE           DE12_TMSE           DE12_BMSE
        ME12_0MSE           ME12_TMSE           ME12_BMSE;

        RUN;  /* 12+1 variables for 18 EST one row for single condition*/


DATA WQS2v_BiasRMSE; /* 36 VARS in 1 record for given Condition */
```

```
SET WQS2v_AVGMS;   /* 36 VARS in 1 record for given Condition */

DE12_0RMSE=SQRT(DE12_0MSE);        DE12_TRMSE=SQRT(DE12_TMSE);        DE12_BRMSE=SQRT(DE12_BMSE);
ME12_0RMSE=SQRT(ME12_0MSE);        ME12_TRMSE=SQRT(ME12_TMSE);        ME12_BRMSE=SQRT(ME12_BMSE);

Label
        DE12_0Bias=DE12_0Bias          DE12_TBias= DE12_TBias          DE12_BBias=DE12_BBias
        ME12_0Bias=ME12_0Bias          ME12_TBias=ME12_TBias           ME12_BBias=ME12_BBias;

DROP  DE12_0MSE   DE12_TMSE   DE12_BMSE   ME12_0MSE   ME12_TMSE   ME12_BMSE;

FORMAT _ALL_ F7.3; Format Condition F3.0;
RUN;
/* UPDATE the single record file WQSEST2v_1 */

DATA   WQSEST2v_1; /* Total 178 = 172 variables ADD 3 *2 =6 Total 211 columns */
MERGE        WQSEST2v_1         WQS2v_BiasRMSE;           by Condition;
RUN;

/* has 178 variables   (81EST +30 Parms +3)
                        27 EST +27 Bias+ 27RMSE estimates one per Condition */

DATA WQSEST2v_500;   /* 500 records for each Condition 1 to 108, each related to a Sampnum */
MERGE Temp.Cond108 (WHERE= (Condition=&Cond)) Temp.WQS2v500_SN&Cond;
                        BY Condition;
RUN;   /* Has 141 variables */

/*******Type 1 Error & Power for 0.20 0.30 0.40 and 0.50 cutoff for individual WQSestimates *********/
/* First do Joint Type1 error and Power for ME12_0 ME12_t and ME12_B then DE12_0 DE12_t and DE12_B */

/**********SEQ 2V ME12 ***************/

DATA ME12S_0TYP1 (KEEP=Condition Sampnum ME12S_0TYP1);
IF _N_=1 THEN ME12S_0TYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME12S_0TYP1=ME12S_0TYP1+1/500;
RETAIN ME12S_0TYP1;
RUN;
```

```
DATA ME12S_0PWR (KEEP=Condition Sampnum ME12S_0PWR);
IF _N_=1 THEN ME12S_0PWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME12S_0PWR=ME12S_0PWR+1/500;
RETAIN ME12S_0PWR;
RUN;

DATA ME12S_TTYP1 (KEEP=Condition Sampnum ME12S_TTYP1);
IF _N_=1 THEN ME12S_TTYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME12S_TTYP1=ME12S_TTYP1+1/500;
RETAIN ME12S_TTYP1;
RUN;

DATA ME12S_TPWR (KEEP=Condition Sampnum ME12S_TPWR);
IF _N_=1 THEN ME12S_TPWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME12S_TPWR=ME12S_TPWR+1/500;
RETAIN ME12S_TPWR;
RUN;

DATA ME12S_BTYP1 (KEEP=Condition Sampnum ME12S_BTYP1);
IF _N_=1 THEN ME12S_BTYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME12S_BTYP1=ME12S_BTYP1+1/500;
RETAIN ME12S_BTYP1;
RUN;

DATA ME12S_BPWR(KEEP=Condition Sampnum ME12S_BPWR);
IF _N_=1 THEN ME12S_BPWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME12S_BPWR=ME12S_BPWR+1/500;
RETAIN ME12S_BPWR;
RUN;
/**********SEQ 2V  DE12 ***************/
```

```
DATA DE12_0TYP1 (KEEP=Condition Sampnum DE12_0TYP1);
IF _N_=1 THEN DE12_0TYP1=0;
SET WQSEST2v_500;
IF ((DE12 =0) & ^(DE12_0LCL<=0<=DE12_0UCL)) THEN
DE12_0TYP1=DE12_0TYP1+1/500;
RETAIN DE12_0TYP1;
RUN;

DATA DE12_0PWR (KEEP=Condition Sampnum DE12_0PWR);
IF _N_=1 THEN DE12_0PWR=0;
SET WQSEST2v_500;
IF ((DE12 ^=0) & ^(DE12_0LCL<=0<=DE12_0UCL)) THEN
DE12_0PWR=DE12_0PWR+1/500;
RETAIN DE12_0PWR;
RUN;

DATA DE12_TTYP1 (KEEP=Condition Sampnum DE12_TTYP1);
IF _N_=1 THEN DE12_TTYP1=0;
SET WQSEST2v_500;
IF ((DE12 =0) & ^(DE12_TLCL<=0<=DE12_TUCL)) THEN
DE12_TTYP1=DE12_TTYP1+1/500;
RETAIN DE12_TTYP1;
RUN;

DATA DE12_TPWR (KEEP=Condition Sampnum DE12_TPWR);
IF _N_=1 THEN DE12_TPWR=0;
SET WQSEST2v_500;
IF ((DE12 ^=0) & ^(DE12_TLCL<=0<=DE12_TUCL)) THEN
DE12_TPWR=DE12_TPWR+1/500;
RETAIN DE12_TPWR;
RUN;

DATA DE12_BTYP1(KEEP=Condition Sampnum DE12_BTYP1);
IF _N_=1 THEN DE12_BTYP1=0;
SET WQSEST2v_500;
IF ((DE12 =0) & ^(DE12_BLCL<=0<=DE12_BUCL)) THEN
DE12_BTYP1=DE12_BTYP1+1/500;
RETAIN DE12_BTYP1;
RUN;
```

```
DATA DE12_BPWR (KEEP=Condition Sampnum DE12_BPWR);
IF _N_=1 THEN DE12_BPWR=0;
SET WQSEST2v_500;
IF ((DE12 ^=0) & ^(DE12_BLCL<=0<=DE12_BUCL)) THEN
DE12_BPWR=DE12_BPWR+1/500;
RETAIN DE12_BPWR;
RUN;

/*Individual Estimates TYPE1 ERR and POWER for 0.10, 0.20, 0.30, 0.40 and 0.50 cutoff for WQS weigts */
DATA DE1_0TYP1 (KEEP=Condition Sampnum DE1_0TYP1);
IF _N_=1 THEN DE1_0TYP1=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S>0 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL))  THEN
DE1_0TYP1=DE1_0TYP1+1/500;
RETAIN DE1_0TYP1;
RUN;

DATA DE1_0TYP1_10 (KEEP=Condition Sampnum DE1_0TYP1_10);
IF _N_=1 THEN DE1_0TYP1_10=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S >0.10 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE1_0TYP1_10=DE1_0TYP1_10+1/500;
RETAIN DE1_0TYP1_10;
RUN;

DATA DE1_0TYP1_20 (KEEP=Condition Sampnum DE1_0TYP1_20);
IF _N_=1 THEN DE1_0TYP1_20=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S >0.20 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE1_0TYP1_20=DE1_0TYP1_20+1/500;
RETAIN DE1_0TYP1_20;
RUN;

DATA DE1_0TYP1_30 (KEEP=Condition Sampnum DE1_0TYP1_30);
IF _N_=1 THEN DE1_0TYP1_30=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S >0.30 & DE12=0 &
```

```
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE1_0TYP1_30=DE1_0TYP1_30+1/500;
RETAIN DE1_0TYP1_30;
RUN;

DATA DE1_0TYP1_40 (KEEP=Condition Sampnum DE1_0TYP1_40);
IF _N_=1 THEN DE1_0TYP1_40=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S >0.40 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE1_0TYP1_40=DE1_0TYP1_40+1/500;
RETAIN DE1_0TYP1_40;
RUN;

DATA DE1_0TYP1_50 (KEEP=Condition Sampnum DE1_0TYP1_50);
IF _N_=1 THEN DE1_0TYP1_50=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB01_S >0.50 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE1_0TYP1_50=DE1_0TYP1_50+1/500;
RETAIN DE1_0TYP1_50;
RUN;

DATA DE1_0PWR (KEEP=Condition Sampnum DE1_0PWR);
IF _N_=1 THEN DE1_0PWR=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0) THEN
DE1_0PWR=DE1_0PWR+1/500;
RETAIN DE1_0PWR;
RUN;

DATA DE1_0PWR_10 (KEEP=Condition Sampnum DE1_0PWR_10);
IF _N_=1 THEN DE1_0PWR_10=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0.10) THEN
DE1_0PWR_10=DE1_0PWR_10+1/500;
RETAIN DE1_0PWR_10;
RUN;

DATA DE1_0PWR_20 (KEEP=Condition Sampnum DE1_0PWR_20);
```

```
IF _N_=1 THEN DE1_0PWR_20=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0.20) THEN
DE1_0PWR_20=DE1_0PWR_20+1/500;
RETAIN DE1_0PWR_20;
RUN;

DATA DE1_0PWR_30 (KEEP=Condition Sampnum DE1_0PWR_30);
IF _N_=1 THEN DE1_0PWR_30=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0.30) THEN
DE1_0PWR_30=DE1_0PWR_30+1/500;
RETAIN DE1_0PWR_30;
RUN;

DATA DE1_0PWR_40 (KEEP=Condition Sampnum DE1_0PWR_40);
IF _N_=1 THEN DE1_0PWR_40=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0.40) THEN
DE1_0PWR_40=DE1_0PWR_40+1/500;
RETAIN DE1_0PWR_40;
RUN;

DATA DE1_0PWR_50 (KEEP=Condition Sampnum DE1_0PWR_50);
IF _N_=1 THEN DE1_0PWR_50=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB01_S >0.50) THEN
DE1_0PWR_50=DE1_0PWR_50+1/500;
RETAIN DE1_0PWR_50;
RUN;

DATA DE1_TTYP1 (KEEP=Condition Sampnum DE1_TTYP1);
IF _N_=1 THEN DE1_TTYP1=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S>0 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1=DE1_TTYP1+1/500;
RETAIN DE1_TTYP1;
RUN;
```

```
DATA DE1_TTYP1_10 (KEEP=Condition Sampnum DE1_TTYP1_10);
IF _N_=1 THEN DE1_TTYP1_10=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S >0.10 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1_10=DE1_TTYP1_10+1/500;
RETAIN DE1_TTYP1_10;
RUN;

DATA DE1_TTYP1_20 (KEEP=Condition Sampnum DE1_TTYP1_20);
IF _N_=1 THEN DE1_TTYP1_20=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S >0.20 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1_20=DE1_TTYP1_20+1/500;
RETAIN DE1_TTYP1_20;
RUN;

DATA DE1_TTYP1_30 (KEEP=Condition Sampnum DE1_TTYP1_30);
IF _N_=1 THEN DE1_TTYP1_30=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S >0.30 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1_30=DE1_TTYP1_30+1/500;
RETAIN DE1_TTYP1_30;
RUN;

DATA DE1_TTYP1_40 (KEEP=Condition Sampnum DE1_TTYP1_40);
IF _N_=1 THEN DE1_TTYP1_40=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S >0.40 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1_40=DE1_TTYP1_40+1/500;
RETAIN DE1_TTYP1_40;
RUN;

DATA DE1_TTYP1_50 (KEEP=Condition Sampnum DE1_TTYP1_50);
IF _N_=1 THEN DE1_TTYP1_50=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXT1_S >0.50 & DE12=0 &
```

```
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE1_TTYP1_50=DE1_TTYP1_50+1/500;
RETAIN DE1_TTYP1_50;
RUN;

DATA DE1_TPWR (KEEP=Condition Sampnum DE1_TPWR);
IF _N_=1 THEN DE1_TPWR=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXT1_S>0) THEN
DE1_TPWR=DE1_TPWR+1/500;
RETAIN DE1_TPWR;
RUN;

DATA DE1_TPWR_10 (KEEP=Condition Sampnum DE1_TPWR_10);
IF _N_=1 THEN DE1_TPWR_10=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXT1_S >0.10) THEN
DE1_TPWR_10=DE1_TPWR_10+1/500;
RETAIN DE1_TPWR_10;
RUN;

DATA DE1_TPWR_20 (KEEP=Condition Sampnum DE1_TPWR_20);
IF _N_=1 THEN DE1_TPWR_20=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXT1_S >0.20) THEN
DE1_TPWR_20=DE1_TPWR_20+1/500;
RETAIN DE1_TPWR_20;
RUN;

DATA DE1_TPWR_30 (KEEP=Condition Sampnum DE1_TPWR_30);
IF _N_=1 THEN DE1_TPWR_30=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXT1_S >0.30) THEN
DE1_TPWR_30=DE1_TPWR_30+1/500;
RETAIN DE1_TPWR_30;
RUN;

DATA DE1_TPWR_40 (KEEP=Condition Sampnum DE1_TPWR_40);
IF _N_=1 THEN DE1_TPWR_40=0;
SET WQSEST2v_500;
```

```
IF (DE1 ^=0 & wXT1_S >0.40) THEN
DE1_TPWR_40=DE1_TPWR_40+1/500;
RETAIN DE1_TPWR_40;
RUN;

DATA DE1_TPWR_50 (KEEP=Condition Sampnum DE1_TPWR_50);
IF _N_=1 THEN DE1_TPWR_50=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXT1_S >0.50) THEN
DE1_TPWR_50=DE1_TPWR_50+1/500;
RETAIN DE1_TPWR_50;
RUN;

DATA DE1_BTYP1 (KEEP=Condition Sampnum DE1_BTYP1);
IF _N_=1 THEN DE1_BTYP1=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S>0 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1=DE1_BTYP1+1/500;
RETAIN DE1_BTYP1;
RUN;

DATA DE1_BTYP1_10 (KEEP=Condition Sampnum DE1_BTYP1_10);
IF _N_=1 THEN DE1_BTYP1_10=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S >0.10 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1_10=DE1_BTYP1_10+1/500;
RETAIN DE1_BTYP1_10;
RUN;

DATA DE1_BTYP1_20 (KEEP=Condition Sampnum DE1_BTYP1_20);
IF _N_=1 THEN DE1_BTYP1_20=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S >0.20 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1_20=DE1_BTYP1_20+1/500;
RETAIN DE1_BTYP1_20;
RUN;
```

```
DATA DE1_BTYP1_30 (KEEP=Condition Sampnum DE1_BTYP1_30);
IF _N_=1 THEN DE1_BTYP1_30=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S >0.30 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1_30=DE1_BTYP1_30+1/500;
RETAIN DE1_BTYP1_30;
RUN;

DATA DE1_BTYP1_40 (KEEP=Condition Sampnum DE1_BTYP1_40);
IF _N_=1 THEN DE1_BTYP1_40=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S >0.40 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1_40=DE1_BTYP1_40+1/500;
RETAIN DE1_BTYP1_40;
RUN;

DATA DE1_BTYP1_50 (KEEP=Condition Sampnum DE1_BTYP1_50);
IF _N_=1 THEN DE1_BTYP1_50=0;
SET WQSEST2v_500;
IF (DE1 =0 & wXB1_S >0.50 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE1_BTYP1_50=DE1_BTYP1_50+1/500;
RETAIN DE1_BTYP1_50;
RUN;

DATA DE1_BPWR (KEEP=Condition Sampnum DE1_BPWR);
IF _N_=1 THEN DE1_BPWR=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S>0) THEN
DE1_BPWR=DE1_BPWR+1/500;
RETAIN DE1_BPWR;
RUN;

DATA DE1_BPWR_10 (KEEP=Condition Sampnum DE1_BPWR_10);
IF _N_=1 THEN DE1_BPWR_10=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S >0.10) THEN
DE1_BPWR_10=DE1_BPWR_10+1/500;
```

```
RETAIN DE1_BPWR_10;
RUN;

DATA DE1_BPWR_20 (KEEP=Condition Sampnum DE1_BPWR_20);
IF _N_=1 THEN DE1_BPWR_20=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S >0.20) THEN
DE1_BPWR_20=DE1_BPWR_20+1/500;
RETAIN DE1_BPWR_20;
RUN;

DATA DE1_BPWR_30 (KEEP=Condition Sampnum DE1_BPWR_30);
IF _N_=1 THEN DE1_BPWR_30=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S >0.30) THEN
DE1_BPWR_30=DE1_BPWR_30+1/500;
RETAIN DE1_BPWR_30;
RUN;

DATA DE1_BPWR_40 (KEEP=Condition Sampnum DE1_BPWR_40);
IF _N_=1 THEN DE1_BPWR_40=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S >0.40) THEN
DE1_BPWR_40=DE1_BPWR_40+1/500;
RETAIN DE1_BPWR_40;
RUN;

DATA DE1_BPWR_50 (KEEP=Condition Sampnum DE1_BPWR_50);
IF _N_=1 THEN DE1_BPWR_50=0;
SET WQSEST2v_500;
IF (DE1 ^=0 & wXB1_S >0.50) THEN
DE1_BPWR_50=DE1_BPWR_50+1/500;
RETAIN DE1_BPWR_50;
RUN;

/*****************DE2**********************/
DATA DE2_0TYP1 (KEEP=Condition Sampnum DE2_0TYP1);
IF _N_=1 THEN DE2_0TYP1=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S>0 & DE12=0 &
```

```
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1=DE2_0TYP1+1/500;
RETAIN DE2_0TYP1;
RUN;

DATA DE2_0TYP1_10 (KEEP=Condition Sampnum DE2_0TYP1_10);
IF _N_=1 THEN DE2_0TYP1_10=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S >0.10 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1_10=DE2_0TYP1_10+1/500;
RETAIN DE2_0TYP1_10;
RUN;

DATA DE2_0TYP1_20 (KEEP=Condition Sampnum DE2_0TYP1_20);
IF _N_=1 THEN DE2_0TYP1_20=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S >0.20 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1_20=DE2_0TYP1_20+1/500;
RETAIN DE2_0TYP1_20;
RUN;

DATA DE2_0TYP1_30 (KEEP=Condition Sampnum DE2_0TYP1_30);
IF _N_=1 THEN DE2_0TYP1_30=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S >0.30 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1_30=DE2_0TYP1_30+1/500;
RETAIN DE2_0TYP1_30;
RUN;

DATA DE2_0TYP1_40 (KEEP=Condition Sampnum DE2_0TYP1_40);
IF _N_=1 THEN DE2_0TYP1_40=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S >0.40 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1_40=DE2_0TYP1_40+1/500;
RETAIN DE2_0TYP1_40;
RUN;
```

```
DATA DE2_0TYP1_50 (KEEP=Condition Sampnum DE2_0TYP1_50);
IF _N_=1 THEN DE2_0TYP1_50=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB02_S >0.50 & DE12=0 &
^(DE12_0LCL< 0 < DE12_0UCL)) THEN
DE2_0TYP1_50=DE2_0TYP1_50+1/500;
RETAIN DE2_0TYP1_50;
RUN;

DATA DE2_0PWR (KEEP=Condition Sampnum DE2_0PWR);
IF _N_=1 THEN DE2_0PWR=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S>0) THEN
DE2_0PWR=DE2_0PWR+1/500;
RETAIN DE2_0PWR;
RUN;

DATA DE2_0PWR_10 (KEEP=Condition Sampnum DE2_0PWR_10);
IF _N_=1 THEN DE2_0PWR_10=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S >0.10) THEN
DE2_0PWR_10=DE2_0PWR_10+1/500;
RETAIN DE2_0PWR_10;
RUN;

DATA DE2_0PWR_20 (KEEP=Condition Sampnum DE2_0PWR_20);
IF _N_=1 THEN DE2_0PWR_20=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S >0.20) THEN
DE2_0PWR_20=DE2_0PWR_20+1/500;
RETAIN DE2_0PWR_20;
RUN;

DATA DE2_0PWR_30 (KEEP=Condition Sampnum DE2_0PWR_30);
IF _N_=1 THEN DE2_0PWR_30=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S >0.30) THEN
DE2_0PWR_30=DE2_0PWR_30+1/500;
RETAIN DE2_0PWR_30;
RUN;
```

```
DATA DE2_0PWR_40 (KEEP=Condition Sampnum DE2_0PWR_40);
IF _N_=1 THEN DE2_0PWR_40=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S >0.40) THEN
DE2_0PWR_40=DE2_0PWR_40+1/500;
RETAIN DE2_0PWR_40;
RUN;

DATA DE2_0PWR_50 (KEEP=Condition Sampnum DE2_0PWR_50);
IF _N_=1 THEN DE2_0PWR_50=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB02_S >0.50) THEN
DE2_0PWR_50=DE2_0PWR_50+1/500;
RETAIN DE2_0PWR_50;
RUN;

DATA DE2_TTYP1 (KEEP=Condition Sampnum DE2_TTYP1);
IF _N_=1 THEN DE2_TTYP1=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S>0 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE2_TTYP1=DE2_TTYP1+1/500;
RETAIN DE2_TTYP1;
RUN;

DATA DE2_TTYP1_10 (KEEP=Condition Sampnum DE2_TTYP1_10);
IF _N_=1 THEN DE2_TTYP1_10=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S >0.10 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE2_TTYP1_10=DE2_TTYP1_10+1/500;
RETAIN DE2_TTYP1_10;
RUN;

DATA DE2_TTYP1_20 (KEEP=Condition Sampnum DE2_TTYP1_20);
IF _N_=1 THEN DE2_TTYP1_20=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S >0.20 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
```

```
DE2_TTYP1_20=DE2_TTYP1_20+1/500;
RETAIN DE2_TTYP1_20;
RUN;

DATA DE2_TTYP1_30 (KEEP=Condition Sampnum DE2_TTYP1_30);
IF _N_=1 THEN DE2_TTYP1_30=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S >0.30 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE2_TTYP1_30=DE2_TTYP1_30+1/500;
RETAIN DE2_TTYP1_30;
RUN;

DATA DE2_TTYP1_40 (KEEP=Condition Sampnum DE2_TTYP1_40);
IF _N_=1 THEN DE2_TTYP1_40=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S >0.40 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE2_TTYP1_40=DE2_TTYP1_40+1/500;
RETAIN DE2_TTYP1_40;
RUN;

DATA DE2_TTYP1_50 (KEEP=Condition Sampnum DE2_TTYP1_50);
IF _N_=1 THEN DE2_TTYP1_50=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXT2_S >0.50 & DE12=0 &
^(DE12_TLCL< 0 < DE12_TUCL)) THEN
DE2_TTYP1_50=DE2_TTYP1_50+1/500;
RETAIN DE2_TTYP1_50;
RUN;

DATA DE2_TPWR (KEEP=Condition Sampnum DE2_TPWR);
IF _N_=1 THEN DE2_TPWR=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S>0) THEN
DE2_TPWR=DE2_TPWR+1/500;
RETAIN DE2_TPWR;
RUN;
DATA DE2_TPWR_10 (KEEP=Condition Sampnum DE2_TPWR_10);
IF _N_=1 THEN DE2_TPWR_10=0;
```

```
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S >0.10) THEN
DE2_TPWR_10=DE2_TPWR_10+1/500;
RETAIN DE2_TPWR_10;
RUN;

DATA DE2_TPWR_20 (KEEP=Condition Sampnum DE2_TPWR_20);
IF _N_=1 THEN DE2_TPWR_20=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S >0.20) THEN
DE2_TPWR_20=DE2_TPWR_20+1/500;
RETAIN DE2_TPWR_20;
RUN;
DATA DE2_TPWR_30 (KEEP=Condition Sampnum DE2_TPWR_30);
IF _N_=1 THEN DE2_TPWR_30=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S >0.30) THEN
DE2_TPWR_30=DE2_TPWR_30+1/500;
RETAIN DE2_TPWR_30;
RUN;

DATA DE2_TPWR_40 (KEEP=Condition Sampnum DE2_TPWR_40);
IF _N_=1 THEN DE2_TPWR_40=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S >0.40) THEN
DE2_TPWR_40=DE2_TPWR_40+1/500;
RETAIN DE2_TPWR_40;
RUN;

DATA DE2_TPWR_50 (KEEP=Condition Sampnum DE2_TPWR_50);
IF _N_=1 THEN DE2_TPWR_50=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXT2_S >0.50) THEN
DE2_TPWR_50=DE2_TPWR_50+1/500;
RETAIN DE2_TPWR_50;
RUN;

DATA DE2_BTYP1 (KEEP=Condition Sampnum DE2_BTYP1);
IF _N_=1 THEN DE2_BTYP1=0;
SET WQSEST2v_500;
```

```
IF (DE2 =0   & wXB2_S >0 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1=DE2_BTYP1+1/500;
RETAIN DE2_BTYP1;
RUN;

DATA DE2_BTYP1_10 (KEEP=Condition Sampnum DE2_BTYP1_10);
IF _N_=1 THEN DE2_BTYP1_10=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB2_S >0.10 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1_10=DE2_BTYP1_10+1/500;
RETAIN DE2_BTYP1_10;
RUN;

DATA DE2_BTYP1_20 (KEEP=Condition Sampnum DE2_BTYP1_20);
IF _N_=1 THEN DE2_BTYP1_20=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB2_S >0.20 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1_20=DE2_BTYP1_20+1/500;
RETAIN DE2_BTYP1_20;
RUN;

DATA DE2_BTYP1_30 (KEEP=Condition Sampnum DE2_BTYP1_30);
IF _N_=1 THEN DE2_BTYP1_30=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB2_S >0.30 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1_30=DE2_BTYP1_30+1/500;
RETAIN DE2_BTYP1_30;
RUN;

DATA DE2_BTYP1_40 (KEEP=Condition Sampnum DE2_BTYP1_40);
IF _N_=1 THEN DE2_BTYP1_40=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB2_S >0.40 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1_40=DE2_BTYP1_40+1/500;
RETAIN DE2_BTYP1_40;
```

```
RUN;

DATA DE2_BTYP1_50 (KEEP=Condition Sampnum DE2_BTYP1_50);
IF _N_=1 THEN DE2_BTYP1_50=0;
SET WQSEST2v_500;
IF (DE2 =0 & wXB2_S >0.50 & DE12=0 &
^(DE12_BLCL< 0 < DE12_BUCL)) THEN
DE2_BTYP1_50=DE2_BTYP1_50+1/500;
RETAIN DE2_BTYP1_50;
RUN;

DATA DE2_BPWR (KEEP=Condition Sampnum DE2_BPWR);
IF _N_=1 THEN DE2_BPWR=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S>0) THEN
DE2_BPWR=DE2_BPWR+1/500;
RETAIN DE2_BPWR;
RUN;

DATA DE2_BPWR_10 (KEEP=Condition Sampnum DE2_BPWR_10);
IF _N_=1 THEN DE2_BPWR_10=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S >0.10) THEN
DE2_BPWR_10=DE2_BPWR_10+1/500;
RETAIN DE2_BPWR_10;
RUN;

DATA DE2_BPWR_20 (KEEP=Condition Sampnum DE2_BPWR_20);
IF _N_=1 THEN DE2_BPWR_20=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S >0.20) THEN
DE2_BPWR_20=DE2_BPWR_20+1/500;
RETAIN DE2_BPWR_20;
RUN;

DATA DE2_BPWR_30 (KEEP=Condition Sampnum DE2_BPWR_30);
IF _N_=1 THEN DE2_BPWR_30=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S >0.30) THEN
DE2_BPWR_30=DE2_BPWR_30+1/500;
```

```
RETAIN DE2_BPWR_30;
RUN;

DATA DE2_BPWR_40 (KEEP=Condition Sampnum DE2_BPWR_40);
IF _N_=1 THEN DE2_BPWR_40=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S >0.40) THEN
DE2_BPWR_40=DE2_BPWR_40+1/500;
RETAIN DE2_BPWR_40;
RUN;

DATA DE2_BPWR_50 (KEEP=Condition Sampnum DE2_BPWR_50);
IF _N_=1 THEN DE2_BPWR_50=0;
SET WQSEST2v_500;
IF (DE2 ^=0 & wXB2_S >0.50) THEN
DE2_BPWR_50=DE2_BPWR_50+1/500;
RETAIN DE2_BPWR_50;
RUN;

/**********NEST 2V   only for ME ***************/
DATA ME12N_0TYP1 (KEEP=Condition Sampnum ME12N_0TYP1);
IF _N_=1 THEN ME12N_0TYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME12N_0TYP1=ME12N_0TYP1+1/500;
RETAIN ME12N_0TYP1;
RUN;

DATA ME12N_0PWR (KEEP=Condition Sampnum ME12N_0PWR);
IF _N_=1 THEN ME12N_0PWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME12N_0PWR=ME12N_0PWR+1/500;
RETAIN ME12N_0PWR;
RUN;

DATA ME12N_TTYP1 (KEEP=Condition Sampnum ME12N_TTYP1);
IF _N_=1 THEN ME12N_TTYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
```

```
ME12N_TTYP1=ME12N_TTYP1+1/500;
RETAIN ME12N_TTYP1;
RUN;

DATA ME12N_TPWR (KEEP=Condition Sampnum ME12N_TPWR);
IF _N_=1 THEN ME12N_TPWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME12N_TPWR=ME12N_TPWR+1/500;
RETAIN ME12N_TPWR;
RUN;

DATA ME12N_BTYP1 (KEEP=Condition Sampnum ME12N_BTYP1);
IF _N_=1 THEN ME12N_BTYP1=0;
SET WQSEST2v_500;
IF ((ME12 =0) & ^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME12N_BTYP1=ME12N_BTYP1+1/500;
RETAIN ME12N_BTYP1;
RUN;

DATA ME12N_BPWR (KEEP=Condition Sampnum ME12N_BPWR);
IF _N_=1 THEN ME12N_BPWR=0;
SET WQSEST2v_500;
IF ((ME12 ^=0) & ^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME12N_BPWR=ME12N_BPWR+1/500;
RETAIN ME12N_BPWR;
RUN;

/*Individual Estimates ME SEQ & NEST, TYPE1 ERR and POWER for 0.10, 0.20, 0.30 , 0.40 and 0.50 cutoff for WQS weigts */
DATA ME1S_0TYP1 (KEEP=Condition Sampnum ME1S_0TYP1);
IF _N_=1 THEN ME1S_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1=ME1S_0TYP1 + (1/500);
RETAIN ME1S_0TYP1;
RUN;

/*ME1ST1Z_0TYP1                ME1SGZ_0TYP1                    ME1ST1GZ_0TYP1 */
DATA ME1ST1Z_0TYP1 (KEEP=Condition Sampnum ME1ST1Z_0TYP1);
```

```
IF _N_=1 THEN ME1ST1Z_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB01_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1ST1Z_0TYP1=ME1ST1Z_0TYP1 + (1/500);
RETAIN ME1ST1Z_0TYP1;
RUN;

DATA ME1SGZ_0TYP1 (KEEP=Condition Sampnum ME1SGZ_0TYP1);
IF _N_=1 THEN ME1SGZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB01_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1SGZ_0TYP1=ME1SGZ_0TYP1 + (1/500);
RETAIN ME1SGZ_0TYP1;
RUN;

DATA ME1ST1GZ_0TYP1 (KEEP=Condition Sampnum ME1ST1GZ_0TYP1);
IF _N_=1 THEN ME1ST1GZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB01_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1ST1GZ_0TYP1=ME1ST1GZ_0TYP1 + (1/500);
RETAIN ME1ST1GZ_0TYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB01_S wXB01_S wXB01_N WXT2 wXB01_S and wXB01_S values below are =0 weights */
DATA ME1S_0TYP1_10 (KEEP=Condition Sampnum   ME1S_0TYP1_10);
IF _N_=1 THEN ME1S_0TYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0.10 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1_10=ME1S_0TYP1_10 + (1/500);
RETAIN ME1S_0TYP1_10;
RUN;

DATA ME1S_0TYP1_20 (KEEP=Condition Sampnum   ME1S_0TYP1_20);
IF _N_=1 THEN ME1S_0TYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0.20 & ME12=0 &
```

```
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1_20=ME1S_0TYP1_20 + (1/500);
RETAIN ME1S_0TYP1_20;
RUN;

DATA ME1S_0TYP1_30 (KEEP=Condition Sampnum   ME1S_0TYP1_30);
IF _N_=1 THEN ME1S_0TYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0.30 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1_30=ME1S_0TYP1_30 + (1/500);
RETAIN ME1S_0TYP1_30;
RUN;

DATA ME1S_0TYP1_40 (KEEP=Condition Sampnum   ME1S_0TYP1_40);
IF _N_=1 THEN ME1S_0TYP1_40=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0.40 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1_40=ME1S_0TYP1_40 + (1/500);
RETAIN ME1S_0TYP1_40;
RUN;

DATA ME1S_0TYP1_50 (KEEP=Condition Sampnum   ME1S_0TYP1_50);
IF _N_=1 THEN ME1S_0TYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_S>0.50 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME1S_0TYP1_50=ME1S_0TYP1_50 + (1/500);
RETAIN ME1S_0TYP1_50;
RUN;

DATA ME1S_0PWR (KEEP= Condition Sampnum ME1S_0PWR);
IF _N_=1 THEN ME1S_0PWR=0;
SET WQSEST2v_500;
IF ((ME1^=0) & wXB01_S>0) THEN
ME1S_0PWR=ME1S_0PWR + (1/500);
RETAIN ME1S_0PWR;
RUN;
```

```
DATA ME1S_0PWR_10 (KEEP= Condition Sampnum ME1S_0PWR_10);
IF _N_=1 THEN ME1S_0PWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_S>0.10) THEN
ME1S_0PWR_10=ME1S_0PWR_10 + (1/500);
RETAIN ME1S_0PWR_10;
RUN;

DATA ME1S_0PWR_20 (KEEP= Condition Sampnum ME1S_0PWR_20);
IF _N_=1 THEN ME1S_0PWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_S>0.20) THEN
ME1S_0PWR_20=ME1S_0PWR_20 + (1/500);
RETAIN ME1S_0PWR_20;
RUN;
DATA ME1S_0PWR_30 (KEEP= Condition Sampnum ME1S_0PWR_30);
IF _N_=1 THEN ME1S_0PWR_30=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_S>0.30) THEN
ME1S_0PWR_30=ME1S_0PWR_30 + (1/500);
RETAIN ME1S_0PWR_30;
RUN;

DATA ME1S_0PWR_40 (KEEP= Condition Sampnum ME1S_0PWR_40);
IF _N_=1 THEN ME1S_0PWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_S>0.40) THEN
ME1S_0PWR_40=ME1S_0PWR_40 + (1/500);
RETAIN ME1S_0PWR_40;
RUN;

DATA ME1S_0PWR_50 (KEEP= Condition Sampnum ME1S_0PWR_50);
IF _N_=1 THEN ME1S_0PWR_50=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_S>0.50) THEN
ME1S_0PWR_50=ME1S_0PWR_50 + (1/500);
RETAIN ME1S_0PWR_50;
RUN;
```

```
/*********ME2_S0***********/
DATA ME2S_0TYP1 (KEEP=Condition Sampnum ME2S_0TYP1);
IF _N_=1 THEN ME2S_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2S_0TYP1=ME2S_0TYP1 + (1/500);
RETAIN ME2S_0TYP1;
RUN;

/*ME2ST2Z_0TYP1                ME2SGZ_0TYP1                  ME2ST2GZ_0TYP1 */
DATA ME2ST2Z_0TYP1 (KEEP=Condition Sampnum ME2ST2Z_0TYP1);
IF _N_=1 THEN          ME2ST2Z_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXB02_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2ST2Z_0TYP1=ME2ST2Z_0TYP1+ (1/500);
RETAIN ME2ST2Z_0TYP1;
RUN;

DATA ME2SGZ_0TYP1 (KEEP=Condition Sampnum ME2SGZ_0TYP1);
IF _N_=1 THEN          ME2SGZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB02_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2SGZ_0TYP1=ME2SGZ_0TYP1 + (1/500);
RETAIN ME2SGZ_0TYP1;
RUN;

DATA ME2ST2GZ_0TYP1 (KEEP=Condition Sampnum ME2ST2GZ_0TYP1);
IF _N_=1 THEN          ME2ST2GZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB02_S>0 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2ST2GZ_0TYP1=ME2ST2GZ_0TYP1 + (1/500);
RETAIN ME2ST2GZ_0TYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB02_S wXB02_S wXB02_N WXT2 wXB02_S and wXB02_S values below are =0 weights */
DATA ME2S_0TYP1_10 (KEEP=Condition Sampnum   ME2S_0TYP1_10);
```

```
IF _N_=1 THEN ME2S_0TYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0.10 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2S_0TYP1_10=ME2S_0TYP1_10 + (1/500);
RETAIN ME2S_0TYP1_10;
RUN;

DATA ME2S_0TYP1_20 (KEEP=Condition Sampnum   ME2S_0TYP1_20);
IF _N_=1 THEN ME2S_0TYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0.20 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2S_0TYP1_20=ME2S_0TYP1_20 + (1/500);
RETAIN ME2S_0TYP1_20;
RUN;

DATA ME2S_0TYP1_30 (KEEP=Condition Sampnum   ME2S_0TYP1_30);
IF _N_=1 THEN ME2S_0TYP1_30=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0.30 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2S_0TYP1_30=ME2S_0TYP1_30 + (1/500);
RETAIN ME2S_0TYP1_30;
RUN;

DATA ME2S_0TYP1_40 (KEEP=Condition Sampnum   ME2S_0TYP1_40);
IF _N_=1 THEN ME2S_0TYP1_40=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0.40 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
ME2S_0TYP1_40=ME2S_0TYP1_40 + (1/500);
RETAIN ME2S_0TYP1_40;
RUN;

DATA ME2S_0TYP1_50 (KEEP=Condition Sampnum   ME2S_0TYP1_50);
IF _N_=1 THEN ME2S_0TYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_S>0.50 & ME12=0 &
^(PS0_ME12_2_5<=0<=PS0_ME12_97_5)) THEN
```

```
ME2S_0TYP1_50=ME2S_0TYP1_50 + (1/500);
RETAIN ME2S_0TYP1_50;
RUN;

DATA ME2S_0PWR (KEEP= Condition Sampnum ME2S_0PWR);
IF _N_=1 THEN ME2S_0PWR=0;
SET WQSEST2v_500;
IF ((ME2^=0) & wXB02_S>0) THEN
ME2S_0PWR=ME2S_0PWR + (1/500);
RETAIN ME2S_0PWR;
RUN;

DATA ME2S_0PWR_10 (KEEP= Condition Sampnum ME2S_0PWR_10);
IF _N_=1 THEN ME2S_0PWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_S>0.10) THEN
ME2S_0PWR_10=ME2S_0PWR_10 + (1/500);
RETAIN ME2S_0PWR_10;
RUN;

DATA ME2S_0PWR_20 (KEEP= Condition Sampnum ME2S_0PWR_20);
IF _N_=1 THEN ME2S_0PWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_S>0.20) THEN
ME2S_0PWR_20=ME2S_0PWR_20 + (1/500);
RETAIN ME2S_0PWR_20;
RUN;

DATA ME2S_0PWR_30 (KEEP= Condition Sampnum ME2S_0PWR_30);
IF _N_=1 THEN ME2S_0PWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_S>0.30) THEN
ME2S_0PWR_30=ME2S_0PWR_30 + (1/500);
RETAIN ME2S_0PWR_30;
RUN;

DATA ME2S_0PWR_40 (KEEP= Condition Sampnum ME2S_0PWR_40);
IF _N_=1 THEN ME2S_0PWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_S>0.40) THEN
```

```
ME2S_0PWR_40=ME2S_0PWR_40 + (1/500);
RETAIN ME2S_0PWR_40;
RUN;

DATA ME2S_0PWR_50 (KEEP= Condition Sampnum ME2S_0PWR_50);
IF _N_=1 THEN ME2S_0PWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_S>0.50) THEN
ME2S_0PWR_50=ME2S_0PWR_50 + (1/500);
RETAIN ME2S_0PWR_50;
RUN;

/*****************ME1_T & ME2_T**************/
DATA ME1S_TTYP1 (KEEP=Condition Sampnum ME1S_TTYP1);
IF _N_=1 THEN ME1S_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1=ME1S_TTYP1 + (1/500);
RETAIN ME1S_TTYP1;
RUN;

/*ME1ST1Z_TTYP1                 ME1SGZ_TTYP1                     ME1ST1GZ_TTYP1 */
DATA ME1ST1Z_TTYP1 (KEEP=Condition Sampnum ME1ST1Z_TTYP1);
IF _N_=1 THEN           ME1ST1Z_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0   & wXT1_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1ST1Z_TTYP1=ME1ST1Z_TTYP1 + (1/500);
RETAIN ME1ST1Z_TTYP1;
RUN;

DATA ME1SGZ_TTYP1 (KEEP=Condition Sampnum ME1SGZ_TTYP1);
IF _N_=1 THEN           ME1SGZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXT1_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1SGZ_TTYP1=ME1SGZ_TTYP1 + (1/500);
RETAIN ME1SGZ_TTYP1;
RUN;
```

```
DATA ME1ST1GZ_TTYP1 (KEEP=Condition Sampnum ME1ST1GZ_TTYP1);
IF _N_=1 THEN ME1ST1GZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0   & wXT1_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1ST1GZ_TTYP1=ME1ST1GZ_TTYP1 + (1/500);
RETAIN ME1ST1GZ_TTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXT1_S wXT1_S wXT1_N WXT2 wXT1_S and wXT1_S values below are =0 weights */
DATA ME1S_TTYP1_10 (KEEP=Condition Sampnum   ME1S_TTYP1_10);
IF _N_=1 THEN ME1S_TTYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0.10 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1_10=ME1S_TTYP1_10 + (1/500);
RETAIN ME1S_TTYP1_10;
RUN;

DATA ME1S_TTYP1_20 (KEEP=Condition Sampnum   ME1S_TTYP1_20);
IF _N_=1 THEN ME1S_TTYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0.20 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1_20=ME1S_TTYP1_20 + (1/500);
RETAIN ME1S_TTYP1_20;
RUN;

DATA ME1S_TTYP1_30 (KEEP=Condition Sampnum   ME1S_TTYP1_30);
IF _N_=1 THEN ME1S_TTYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0.30 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1_30=ME1S_TTYP1_30 + (1/500);
RETAIN ME1S_TTYP1_30;
RUN;

DATA ME1S_TTYP1_40 (KEEP=Condition Sampnum ME1S_TTYP1_40);
IF _N_=1 THEN ME1S_TTYP1_40=0;
```

```
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0.40 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1_40=ME1S_TTYP1_40 + (1/500);
RETAIN ME1S_TTYP1_40;
RUN;

DATA ME1S_TTYP1_50 (KEEP=Condition Sampnum   ME1S_TTYP1_50);
IF _N_=1 THEN ME1S_TTYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_S>0.50 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME1S_TTYP1_50=ME1S_TTYP1_50 + (1/500);
RETAIN ME1S_TTYP1_50;
RUN;

DATA ME1S_TPWR (KEEP= Condition Sampnum ME1S_TPWR);
IF _N_=1 THEN ME1S_TPWR=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0) THEN
ME1S_TPWR=ME1S_TPWR + (1/500);
RETAIN ME1S_TPWR;
RUN;

DATA ME1S_TPWR_10 (KEEP= Condition Sampnum ME1S_TPWR_10);
IF _N_=1 THEN ME1S_TPWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0.10) THEN
ME1S_TPWR_10=ME1S_TPWR_10 + (1/500);
RETAIN ME1S_TPWR_10;
RUN;

DATA ME1S_TPWR_20 (KEEP= Condition Sampnum ME1S_TPWR_20);
IF _N_=1 THEN ME1S_TPWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0.20) THEN
ME1S_TPWR_20=ME1S_TPWR_20 + (1/500);
RETAIN ME1S_TPWR_20;
RUN;
```

```
DATA ME1S_TPWR_30 (KEEP= Condition Sampnum ME1S_TPWR_30);
IF _N_=1 THEN ME1S_TPWR_30=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0.30) THEN
ME1S_TPWR_30=ME1S_TPWR_30 + (1/500);
RETAIN ME1S_TPWR_30;
RUN;

DATA ME1S_TPWR_40 (KEEP= Condition Sampnum ME1S_TPWR_40);
IF _N_=1 THEN ME1S_TPWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0.40) THEN
ME1S_TPWR_40=ME1S_TPWR_40 + (1/500);
RETAIN ME1S_TPWR_40;
RUN;

DATA ME1S_TPWR_50 (KEEP= Condition Sampnum ME1S_TPWR_50);
IF _N_=1 THEN ME1S_TPWR_50=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_S>0.50) THEN
ME1S_TPWR_50=ME1S_TPWR_50 + (1/500);
RETAIN ME1S_TPWR_50;
RUN;

/*********ME2_ST***********/
DATA ME2S_TTYP1 (KEEP=Condition Sampnum ME2S_TTYP1);
IF _N_=1 THEN ME2S_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0   & wXT2_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5))  THEN
ME2S_TTYP1=ME2S_TTYP1 + (1/500);
RETAIN ME2S_TTYP1;
RUN;

/*ME2ST2Z_TTYP1                 ME2SGZ_TTYP1                     ME2ST2GZ_TTYP1 */
DATA ME2ST2Z_TTYP1 (KEEP=Condition Sampnum ME2ST2Z_TTYP1);
IF _N_=1 THEN ME2ST2Z_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXT2_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
```

```
ME2ST2Z_TTYP1=ME2ST2Z_TTYP1 + (1/500);
RETAIN ME2ST2Z_TTYP1;
RUN;

DATA ME2SGZ_TTYP1 (KEEP=Condition Sampnum ME2SGZ_TTYP1);
IF _N_=1 THEN            ME2SGZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXT2_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2SGZ_TTYP1=ME2SGZ_TTYP1 + (1/500);
RETAIN ME2SGZ_TTYP1;
RUN;

DATA ME2ST2GZ_TTYP1 (KEEP=Condition Sampnum ME2ST2GZ_TTYP1);
IF _N_=1 THEN ME2ST2GZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXT2_S>0 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2ST2GZ_TTYP1=ME2ST2GZ_TTYP1 + (1/500);
RETAIN ME2ST2GZ_TTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXT2_S wXT2_S wXT2_N WXT2 wXT2_S and wXT2_S values below are =0 weights */
DATA ME2S_TTYP1_10 (KEEP=Condition Sampnum   ME2S_TTYP1_10);
IF _N_=1 THEN ME2S_TTYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_S>0.10 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2S_TTYP1_10=ME2S_TTYP1_10 + (1/500);
RETAIN ME2S_TTYP1_10;
RUN;

DATA ME2S_TTYP1_20 (KEEP=Condition Sampnum   ME2S_TTYP1_20);
IF _N_=1 THEN ME2S_TTYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_S>0.20 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2S_TTYP1_20=ME2S_TTYP1_20 + (1/500);
RETAIN ME2S_TTYP1_20;
RUN;
```

```
DATA ME2S_TTYP1_30 (KEEP=Condition Sampnum   ME2S_TTYP1_30);
IF _N_=1 THEN ME2S_TTYP1_30=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_S>0.30 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2S_TTYP1_30=ME2S_TTYP1_30 + (1/500);
RETAIN ME2S_TTYP1_30;
RUN;

DATA ME2S_TTYP1_40 (KEEP=Condition Sampnum ME2S_TTYP1_40);
IF _N_=1 THEN ME2S_TTYP1_40=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_S>0.40 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2S_TTYP1_40=ME2S_TTYP1_40 + (1/500);
RETAIN ME2S_TTYP1_40;
RUN;

DATA ME2S_TTYP1_50 (KEEP=Condition Sampnum   ME2S_TTYP1_50);
IF _N_=1 THEN ME2S_TTYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_S>0.50 & ME12=0 &
^(PST_ME12_2_5<=0<=PST_ME12_97_5)) THEN
ME2S_TTYP1_50=ME2S_TTYP1_50 + (1/500);
RETAIN ME2S_TTYP1_50;
RUN;

DATA ME2S_TPWR (KEEP= Condition Sampnum ME2S_TPWR);
IF _N_=1 THEN ME2S_TPWR=0;
SET WQSEST2v_500;
IF ((ME2^=0) & wXT2_S>0) THEN
ME2S_TPWR=ME2S_TPWR + (1/500);
RETAIN ME2S_TPWR;
RUN;

DATA ME2S_TPWR_10 (KEEP= Condition Sampnum ME2S_TPWR_10);
IF _N_=1 THEN ME2S_TPWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_S>0.10) THEN
```

```
ME2S_TPWR_10=ME2S_TPWR_10 + (1/500);
RETAIN ME2S_TPWR_10;
RUN;

DATA ME2S_TPWR_20 (KEEP= Condition Sampnum ME2S_TPWR_20);
IF _N_=1 THEN ME2S_TPWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_S>0.20) THEN
ME2S_TPWR_20=ME2S_TPWR_20 + (1/500);
RETAIN ME2S_TPWR_20;
RUN;

DATA ME2S_TPWR_30 (KEEP= Condition Sampnum ME2S_TPWR_30);
IF _N_=1 THEN ME2S_TPWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_S>0.30) THEN
ME2S_TPWR_30=ME2S_TPWR_30 + (1/500);
RETAIN ME2S_TPWR_30;
RUN;

DATA ME2S_TPWR_40 (KEEP= Condition Sampnum ME2S_TPWR_40);
IF _N_=1 THEN ME2S_TPWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_S>0.40) THEN
ME2S_TPWR_40=ME2S_TPWR_40 + (1/500);
RETAIN ME2S_TPWR_40;
RUN;

DATA ME2S_TPWR_50 KEEP= Condition Sampnum ME2S_TPWR_50);
IF _N_=1 THEN ME2S_TPWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_S>0.50) THEN
ME2S_TPWR_50=ME2S_TPWR_50 + (1/500);
RETAIN ME2S_TPWR_50;
RUN;

/*****************ME1B***************/
DATA ME1S_BTYP1 (KEEP=Condition Sampnum ME1S_BTYP1);
IF _N_=1 THEN ME1S_BTYP1=0;
SET WQSEST2v_500;
```

```
IF (ME1=0 & wXB1_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1S_BTYP1=ME1S_BTYP1 + (1/500);
RETAIN ME1S_BTYP1;
RUN;

/*ME1ST1Z_BTYP1                ME1SGZ_BTYP1                    ME1ST1GZ_BTYP1 */
DATA ME1ST1Z_BTYP1 (KEEP=Condition Sampnum ME1ST1Z_BTYP1);
IF _N_=1 THEN ME1ST1Z_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB1_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1ST1Z_BTYP1=ME1ST1Z_BTYP1 + (1/500);
RETAIN ME1ST1Z_BTYP1;
RUN;

DATA ME1SGZ_BTYP1 (KEEP=Condition Sampnum ME1SGZ_BTYP1);
IF _N_=1 THEN ME1SGZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB1_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1SGZ_BTYP1=ME1SGZ_BTYP1 + (1/500);
RETAIN ME1SGZ_BTYP1;
RUN;

DATA ME1ST1GZ_BTYP1 (KEEP=Condition Sampnum ME1ST1GZ_BTYP1);
IF _N_=1 THEN ME1ST1GZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB1_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1ST1GZ_BTYP1=ME1ST1GZ_BTYP1 + (1/500);
RETAIN ME1ST1GZ_BTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB01_S wXB01_S wXB01_N WXT2 wXB01_S and wXB01_S values below are =0 weights */
DATA ME1S_BTYP1_10 (KEEP=Condition Sampnum ME1S_BTYP1_10);
IF _N_=1 THEN ME1S_BTYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_S>0.10 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
```

```
ME1S_BTYP1_10=ME1S_BTYP1_10 + (1/500);
RETAIN ME1S_BTYP1_10;
RUN;

DATA ME1S_BTYP1_20 (KEEP=Condition Sampnum   ME1S_BTYP1_20);
IF _N_=1 THEN ME1S_BTYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_S>0.20 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1S_BTYP1_20=ME1S_BTYP1_20 + (1/500);
RETAIN ME1S_BTYP1_20;
RUN;

DATA ME1S_BTYP1_30 (KEEP=Condition Sampnum   ME1S_BTYP1_30);
IF _N_=1 THEN ME1S_BTYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_S>0.30 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1S_BTYP1_30=ME1S_BTYP1_30 + (1/500);
RETAIN ME1S_BTYP1_30;
RUN;

DATA ME1S_BTYP1_40 (KEEP=Condition Sampnum   ME1S_BTYP1_40);
IF _N_=1 THEN ME1S_BTYP1_40=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_S>0.40 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1S_BTYP1_40=ME1S_BTYP1_40 + (1/500);
RETAIN ME1S_BTYP1_40;
RUN;

DATA ME1S_BTYP1_50 (KEEP=Condition Sampnum   ME1S_BTYP1_50);
IF _N_=1 THEN ME1S_BTYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_S>0.50 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME1S_BTYP1_50=ME1S_BTYP1_50 + (1/500);
RETAIN ME1S_BTYP1_50;
RUN;
```

```
DATA ME1S_BPWR (KEEP= Condition Sampnum ME1S_BPWR);
IF _N_=1 THEN ME1S_BPWR=0;
SET WQSEST2v_500;
IF ((ME1^=0) & wXB1_S>0) THEN
ME1S_BPWR=ME1S_BPWR + (1/500);
RETAIN ME1S_BPWR;
RUN;

DATA ME1S_BPWR_10 (KEEP= Condition Sampnum ME1S_BPWR_10);
IF _N_=1 THEN ME1S_BPWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_S>0.10) THEN
ME1S_BPWR_10=ME1S_BPWR_10 + (1/500);
RETAIN ME1S_BPWR_10;
RUN;

DATA ME1S_BPWR_20 (KEEP= Condition Sampnum ME1S_BPWR_20);
IF _N_=1 THEN ME1S_BPWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_S>0.20) THEN
ME1S_BPWR_20=ME1S_BPWR_20 + (1/500);
RETAIN ME1S_BPWR_20;
RUN;

DATA ME1S_BPWR_30 (KEEP= Condition Sampnum ME1S_BPWR_30);
IF _N_=1 THEN ME1S_BPWR_30=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_S>0.30) THEN
ME1S_BPWR_30=ME1S_BPWR_30 + (1/500);
RETAIN ME1S_BPWR_30;
RUN;

DATA ME1S_BPWR_40 (KEEP= Condition Sampnum ME1S_BPWR_40);
IF _N_=1 THEN ME1S_BPWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_S>0.40) THEN
ME1S_BPWR_40=ME1S_BPWR_40 + (1/500);
RETAIN ME1S_BPWR_40;
RUN;
```

```
DATA ME1S_BPWR_50 (KEEP= Condition Sampnum ME1S_BPWR_50);
IF _N_=1 THEN ME1S_BPWR_50=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_S>0.50) THEN
ME1S_BPWR_50=ME1S_BPWR_50 + (1/500);
RETAIN ME1S_BPWR_50;
RUN;



/*********ME2_SB***********/
DATA ME2S_BTYP1 (KEEP=Condition Sampnum ME2S_BTYP1);
IF _N_=1 THEN ME2S_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1=ME2S_BTYP1 + (1/500);
RETAIN ME2S_BTYP1;
RUN;

/*ME2ST2Z_BTYP1          ME2SGZ_BTYP1          ME2ST2GZ_BTYP1 */
DATA ME2ST2Z_BTYP1 (KEEP=Condition Sampnum ME2ST2Z_BTYP1);
IF _N_=1 THEN ME2ST2Z_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXB2_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2ST2Z_BTYP1=ME2ST2Z_BTYP1 + (1/500);
RETAIN ME2ST2Z_BTYP1;
RUN;

DATA ME2SGZ_BTYP1 (KEEP=Condition Sampnum ME2SGZ_BTYP1);
IF _N_=1 THEN      ME2SGZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB2_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2SGZ_BTYP1=ME2SGZ_BTYP1 + (1/500);
RETAIN ME2SGZ_BTYP1;
RUN;

DATA ME2ST2GZ_BTYP1 (KEEP=Condition Sampnum ME2ST2GZ_BTYP1);
```

```
IF _N_=1 THEN           ME2ST2GZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB2_S>0 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2ST2GZ_BTYP1=ME2ST2GZ_BTYP1 + (1/500);
RETAIN ME2ST2GZ_BTYP1;
RUN;



/* Cut-off value of 0.10 for WQS weights wXB2_S wXB2_S wXB2_N WXT2 wXB2_S and wXB2_S values below are =0 weights */
DATA ME2S_BTYP1_10 (KEEP=Condition Sampnum ME2S_BTYP1_10);
IF _N_=1 THEN ME2S_BTYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0.10 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1_10=ME2S_BTYP1_10 + (1/500);
RETAIN ME2S_BTYP1_10;
RUN;

DATA ME2S_BTYP1_20 (KEEP=Condition Sampnum ME2S_BTYP1_20);
IF _N_=1 THEN ME2S_BTYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0.20 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1_20=ME2S_BTYP1_20 + (1/500);
RETAIN ME2S_BTYP1_20;
RUN;

DATA ME2S_BTYP1_30 (KEEP=Condition Sampnum ME2S_BTYP1_30);
IF _N_=1 THEN ME2S_BTYP1_30=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0.30 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1_30=ME2S_BTYP1_30 + (1/500);
RETAIN ME2S_BTYP1_30;
RUN;

DATA ME2S_BTYP1_40 (KEEP=Condition Sampnum ME2S_BTYP1_40);
IF _N_=1 THEN ME2S_BTYP1_40=0;
```

```
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0.40 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1_40=ME2S_BTYP1_40 + (1/500);
RETAIN ME2S_BTYP1_40;
RUN;

DATA ME2S_BTYP1_50 (KEEP=Condition Sampnum ME2S_BTYP1_50);
IF _N_=1 THEN ME2S_BTYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_S>0.50 & ME12=0 &
^(PSB_ME12_2_5<=0<=PSB_ME12_97_5)) THEN
ME2S_BTYP1_50=ME2S_BTYP1_50 + (1/500);
RETAIN ME2S_BTYP1_50;
RUN;

DATA ME2S_BPWR (KEEP= Condition Sampnum ME2S_BPWR);
IF _N_=1 THEN ME2S_BPWR=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0) THEN
ME2S_BPWR=ME2S_BPWR + (1/500);
RETAIN ME2S_BPWR;
RUN;

DATA ME2S_BPWR_10 (KEEP= Condition Sampnum ME2S_BPWR_10);
IF _N_=1 THEN ME2S_BPWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0.10) THEN
ME2S_BPWR_10=ME2S_BPWR_10 + (1/500);
RETAIN ME2S_BPWR_10;
RUN;

DATA ME2S_BPWR_20 (KEEP= Condition Sampnum ME2S_BPWR_20);
IF _N_=1 THEN ME2S_BPWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0.20) THEN
ME2S_BPWR_20=ME2S_BPWR_20 + (1/500);
RETAIN ME2S_BPWR_20;
RUN;
```

```
DATA ME2S_BPWR_30 (KEEP= Condition Sampnum ME2S_BPWR_30);
IF _N_=1 THEN ME2S_BPWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0.30) THEN
ME2S_BPWR_30=ME2S_BPWR_30 + (1/500);
RETAIN ME2S_BPWR_30;
RUN;

DATA ME2S_BPWR_40 (KEEP= Condition Sampnum ME2S_BPWR_40);
IF _N_=1 THEN ME2S_BPWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0.40) THEN
ME2S_BPWR_40=ME2S_BPWR_40 + (1/500);
RETAIN ME2S_BPWR_40;
RUN;

DATA ME2S_BPWR_50 (KEEP= Condition Sampnum ME2S_BPWR_50);
IF _N_=1 THEN ME2S_BPWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_S>0.50) THEN
ME2S_BPWR_50=ME2S_BPWR_50 + (1/500);
RETAIN ME2S_BPWR_50;
RUN;

/**************NESTED INDIVIUAL TYPE1 and POWER fro 0.10, 0.20, 0.30, 0.40 and 0.50 cutoff for WQS weights */
DATA ME1N_0TYP1 (KEEP=Condition Sampnum ME1N_0TYP1);
IF _N_=1 THEN ME1N_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1=ME1N_0TYP1+ (1/500);
RETAIN ME1N_0TYP1;
RUN;

/*ME1NT1Z_0TYP1                ME1NGZ_0TYP1                ME1NT1GZ_0TYP1 */
DATA ME1NT1Z_0TYP1 (KEEP=Condition Sampnum ME1NT1Z_0TYP1);
IF _N_=1 THEN ME1NT1Z_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB01_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
```

```
ME1NT1Z_0TYP1=ME1NT1Z_0TYP1 + (1/500);
RETAIN ME1NT1Z_0TYP1;
RUN;

DATA ME1NGZ_0TYP1 (KEEP=Condition Sampnum ME1NGZ_0TYP1);
IF _N_=1 THEN ME1NGZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB01_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1NGZ_0TYP1=ME1NGZ_0TYP1 + (1/500);
RETAIN ME1NGZ_0TYP1;
RUN;

DATA ME1NT1GZ_0TYP1 (KEEP=Condition Sampnum ME1NT1GZ_0TYP1);
IF _N_=1 THEN ME1NT1GZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB01_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1NT1GZ_0TYP1=ME1NT1GZ_0TYP1 + (1/500);
RETAIN ME1NT1GZ_0TYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB01_N wXB01_N wXB01_N WXT2 wXB01_N and wXB01_N values below are =0 weights */
DATA ME1N_0TYP1_10 (KEEP=Condition Sampnum ME1N_0TYP1_10);
IF _N_=1 THEN ME1N_0TYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0.10 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1_10=ME1N_0TYP1_10 + (1/500);
RETAIN ME1N_0TYP1_10;
RUN;

DATA ME1N_0TYP1_20 (KEEP=Condition Sampnum ME1N_0TYP1_20);
IF _N_=1 THEN ME1N_0TYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0.20 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1_20=ME1N_0TYP1_20 + (1/500);
RETAIN ME1N_0TYP1_20;
RUN;
```

```
DATA ME1N_0TYP1_30 (KEEP=Condition Sampnum ME1N_0TYP1_30);
IF _N_=1 THEN ME1N_0TYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0.30 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1_30=ME1N_0TYP1_30 + (1/500);
RETAIN ME1N_0TYP1_30;
RUN;

DATA ME1N_0TYP1_40 (KEEP=Condition Sampnum ME1N_0TYP1_40);
IF _N_=1 THEN ME1N_0TYP1_40=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0.40 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1_40=ME1N_0TYP1_40 + (1/500);
RETAIN ME1N_0TYP1_40;
RUN;

DATA ME1N_0TYP1_50 (KEEP=Condition Sampnum ME1N_0TYP1_50);
IF _N_=1 THEN ME1N_0TYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB01_N>0.50 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME1N_0TYP1_50=ME1N_0TYP1_50 + (1/500);
RETAIN ME1N_0TYP1_50;
RUN;

DATA ME1N_0PWR (KEEP= Condition Sampnum ME1N_0PWR);
IF _N_=1 THEN ME1N_0PWR=0;
SET WQSEST2v_500;
IF ((ME1^=0) & wXB01_N>0) THEN
ME1N_0PWR=ME1N_0PWR + (1/500);
RETAIN ME1N_0PWR;
RUN;

DATA ME1N_0PWR_10 (KEEP= Condition Sampnum ME1N_0PWR_10);
IF _N_=1 THEN ME1N_0PWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_N>0.10) THEN
```

```
ME1N_0PWR_10=ME1N_0PWR_10 + (1/500);
RETAIN ME1N_0PWR_10;
RUN;

DATA ME1N_0PWR_20 (KEEP= Condition Sampnum ME1N_0PWR_20);
IF _N_=1 THEN ME1N_0PWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_N>0.20) THEN
ME1N_0PWR_20=ME1N_0PWR_20 + (1/500);
RETAIN ME1N_0PWR_20;
RUN;

DATA ME1N_0PWR_30 (KEEP= Condition Sampnum ME1N_0PWR_30);
IF _N_=1 THEN ME1N_0PWR_30=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_N>0.30) THEN
ME1N_0PWR_30=ME1N_0PWR_30 + (1/500);
RETAIN ME1N_0PWR_30;
RUN;

DATA ME1N_0PWR_40 (KEEP= Condition Sampnum ME1N_0PWR_40);
IF _N_=1 THEN ME1N_0PWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_N>0.40) THEN
ME1N_0PWR_40=ME1N_0PWR_40 + (1/500);
RETAIN ME1N_0PWR_40;
RUN;

DATA ME1N_0PWR_50 (KEEP= Condition Sampnum ME1N_0PWR_50);
IF _N_=1 THEN ME1N_0PWR_50=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB01_N>0.50) THEN
ME1N_0PWR_50=ME1N_0PWR_50 + (1/500);
RETAIN ME1N_0PWR_50;
RUN;

/*********ME2_N0***********/
DATA ME2N_0TYP1 (KEEP=Condition Sampnum ME2N_0TYP1);
IF _N_=1 THEN ME2N_0TYP1=0;
SET WQSEST2v_500;
```

```
IF (ME2=0 & wXB02_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2N_0TYP1=ME2N_0TYP1 + (1/500);
RETAIN ME2N_0TYP1;
RUN;

/*ME2NT2Z_0TYP1                  ME2NGZ_0TYP1                      ME2NT2GZ_0TYP1 */
DATA ME2NT2Z_0TYP1 (KEEP=Condition Sampnum ME2NT2Z_0TYP1);
IF _N_=1 THEN ME2NT2Z_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXB02_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2NT2Z_0TYP1=ME2NT2Z_0TYP1 + (1/500);
RETAIN ME2NT2Z_0TYP1;
RUN;

DATA ME2NGZ_0TYP1 (KEEP=Condition Sampnum ME2NGZ_0TYP1);
IF _N_=1 THEN ME2NGZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB02_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2NGZ_0TYP1=ME2NGZ_0TYP1 + (1/500);
RETAIN ME2NGZ_0TYP1;
RUN;

DATA ME2NT2GZ_0TYP1 (KEEP=Condition Sampnum ME2NT2GZ_0TYP1);
IF _N_=1 THEN ME2NT2GZ_0TYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB02_N>0 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2NT2GZ_0TYP1=ME2NT2GZ_0TYP1 + (1/500);
RETAIN ME2NT2GZ_0TYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB02_N wXB02_N wXB02_N WXT2 wXB02_N and wXB02_N values below are =0 weights */
DATA ME2N_0TYP1_10 (KEEP=Condition Sampnum   ME2N_0TYP1_10);
IF _N_=1 THEN ME2N_0TYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_N>0.10 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
```

```
ME2N_0TYP1_10=ME2N_0TYP1_10 + (1/500);
RETAIN ME2N_0TYP1_10;
RUN;

DATA ME2N_0TYP1_20 (KEEP=Condition Sampnum   ME2N_0TYP1_20);
IF _N_=1 THEN ME2N_0TYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_N>0.20 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2N_0TYP1_20=ME2N_0TYP1_20 + (1/500);
RETAIN ME2N_0TYP1_20;
RUN;

DATA ME2N_0TYP1_30 (KEEP=Condition Sampnum ME2N_0TYP1_30);
IF _N_=1 THEN ME2N_0TYP1_30=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_N>0.30 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2N_0TYP1_30=ME2N_0TYP1_30 + (1/500);
RETAIN ME2N_0TYP1_30;
RUN;

DATA ME2N_0TYP1_40 (KEEP=Condition Sampnum ME2N_0TYP1_40);
IF _N_=1 THEN ME2N_0TYP1_40=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_N>0.40 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2N_0TYP1_40=ME2N_0TYP1_40 + (1/500);
RETAIN ME2N_0TYP1_40;
RUN;

DATA ME2N_0TYP1_50 (KEEP=Condition Sampnum ME2N_0TYP1_50);
IF _N_=1 THEN ME2N_0TYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB02_N>0.50 & ME12=0 &
^(PN0_ME12_2_5<=0<=PN0_ME12_97_5)) THEN
ME2N_0TYP1_50=ME2N_0TYP1_50 + (1/500);
RETAIN ME2N_0TYP1_50;
RUN;
```

```
DATA ME2N_0PWR (KEEP= Condition Sampnum ME2N_0PWR);
IF _N_=1 THEN ME2N_0PWR=0;
SET WQSEST2v_500;
IF ((ME2^=0) & wXB02_N>0) THEN
ME2N_0PWR=ME2N_0PWR + (1/500);
RETAIN ME2N_0PWR;
RUN;

DATA ME2N_0PWR_10 (KEEP= Condition Sampnum ME2N_0PWR_10);
IF _N_=1 THEN ME2N_0PWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_N>0.10) THEN
ME2N_0PWR_10=ME2N_0PWR_10 + (1/500);
RETAIN ME2N_0PWR_10;
RUN;

DATA ME2N_0PWR_20 (KEEP= Condition Sampnum ME2N_0PWR_20);
IF _N_=1 THEN ME2N_0PWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_N>0.20) THEN
ME2N_0PWR_20=ME2N_0PWR_20 + (1/500);
RETAIN ME2N_0PWR_20;
RUN;

DATA ME2N_0PWR_30 (KEEP= Condition Sampnum ME2N_0PWR_30);
IF _N_=1 THEN ME2N_0PWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_N>0.30) THEN
ME2N_0PWR_30=ME2N_0PWR_30 + (1/500);
RETAIN ME2N_0PWR_30;
RUN;

DATA ME2N_0PWR_40 (KEEP= Condition Sampnum ME2N_0PWR_40);
IF _N_=1 THEN ME2N_0PWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_N>0.40) THEN
ME2N_0PWR_40=ME2N_0PWR_40 + (1/500);
RETAIN ME2N_0PWR_40;
RUN;
```

```
DATA ME2N_0PWR_50 (KEEP= Condition Sampnum ME2N_0PWR_50);
IF _N_=1 THEN ME2N_0PWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB02_N>0.50) THEN
ME2N_0PWR_50=ME2N_0PWR_50 + (1/500);
RETAIN ME2N_0PWR_50;
RUN;



/*****************ME1_T & ME2_T**************/
DATA ME1N_TTYP1 (KEEP=Condition Sampnum ME1N_TTYP1);
IF _N_=1 THEN ME1N_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1=ME1N_TTYP1 + (1/500);
RETAIN ME1N_TTYP1;
RUN;

/*ME1NT1Z_TTYP1                ME1NGZ_TTYP1                  ME1NT1GZ_TTYP1 */
DATA ME1NT1Z_TTYP1 (KEEP=Condition Sampnum ME1NT1Z_TTYP1);
IF _N_=1 THEN ME1NT1Z_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0   & wXT1_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1NT1Z_TTYP1=ME1NT1Z_TTYP1 + (1/500);
RETAIN ME1NT1Z_TTYP1;
RUN;

DATA ME1NGZ_TTYP1 (KEEP=Condition Sampnum ME1NGZ_TTYP1);
IF _N_=1 THEN ME1NGZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXT1_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1NGZ_TTYP1=ME1NGZ_TTYP1 + (1/500);
RETAIN ME1NGZ_TTYP1;
RUN;

DATA ME1NT1GZ_TTYP1 (KEEP=Condition Sampnum ME1NT1GZ_TTYP1);
```

```
IF _N_=1 THEN ME1NT1GZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0   & wXT1_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1NT1GZ_TTYP1=ME1NT1GZ_TTYP1 + (1/500);
RETAIN ME1NT1GZ_TTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXT1_N wXT1_N wXT1_N WXT2 wXT1_N and wXT1_N values below are =0 weights */
DATA ME1N_TTYP1_10 (KEEP=Condition Sampnum ME1N_TTYP1_10);
IF _N_=1 THEN ME1N_TTYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0.10 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1_10=ME1N_TTYP1_10 + (1/500);
RETAIN ME1N_TTYP1_10;
RUN;

DATA ME1N_TTYP1_20 (KEEP=Condition Sampnum ME1N_TTYP1_20);
IF _N_=1 THEN ME1N_TTYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0.20 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1_20=ME1N_TTYP1_20 + (1/500);
RETAIN ME1N_TTYP1_20;
RUN;

DATA ME1N_TTYP1_30 (KEEP=Condition Sampnum   ME1N_TTYP1_30);
IF _N_=1 THEN ME1N_TTYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0.30 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1_30=ME1N_TTYP1_30 + (1/500);
RETAIN ME1N_TTYP1_30;
RUN;

DATA ME1N_TTYP1_40 (KEEP=Condition Sampnum ME1N_TTYP1_40);
IF _N_=1 THEN ME1N_TTYP1_40=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0.40 & ME12=0 &
```

```
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1_40=ME1N_TTYP1_40 + (1/500);
RETAIN ME1N_TTYP1_40;
RUN;

DATA ME1N_TTYP1_50 (KEEP=Condition Sampnum ME1N_TTYP1_50);
IF _N_=1 THEN ME1N_TTYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXT1_N>0.50 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME1N_TTYP1_50=ME1N_TTYP1_50 + (1/500);
RETAIN ME1N_TTYP1_50;
RUN;

DATA ME1N_TPWR (KEEP= Condition Sampnum ME1N_TPWR);
IF _N_=1 THEN ME1N_TPWR=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0) THEN
ME1N_TPWR=ME1N_TPWR + (1/500);
RETAIN ME1N_TPWR;
RUN;

DATA ME1N_TPWR_10 (KEEP= Condition Sampnum ME1N_TPWR_10);
IF _N_=1 THEN ME1N_TPWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0.10) THEN
ME1N_TPWR_10=ME1N_TPWR_10 + (1/500);
RETAIN ME1N_TPWR_10;
RUN;

DATA ME1N_TPWR_20 (KEEP= Condition Sampnum ME1N_TPWR_20);
IF _N_=1 THEN ME1N_TPWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0.20) THEN
ME1N_TPWR_20=ME1N_TPWR_20 + (1/500);
RETAIN ME1N_TPWR_20;
RUN;

DATA ME1N_TPWR_30 (KEEP= Condition Sampnum ME1N_TPWR_30);
IF _N_=1 THEN ME1N_TPWR_30=0;
```

```
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0.30) THEN
ME1N_TPWR_30=ME1N_TPWR_30 + (1/500);
RETAIN ME1N_TPWR_30;
RUN;

DATA ME1N_TPWR_40 (KEEP= Condition Sampnum ME1N_TPWR_40);
IF _N_=1 THEN ME1N_TPWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0.40) THEN
ME1N_TPWR_40=ME1N_TPWR_40 + (1/500);
RETAIN ME1N_TPWR_40;
RUN;

DATA ME1N_TPWR_50 (KEEP= Condition Sampnum ME1N_TPWR_50);
IF _N_=1 THEN ME1N_TPWR_50=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXT1_N>0.50) THEN
ME1N_TPWR_50=ME1N_TPWR_50 + (1/500);
RETAIN ME1N_TPWR_50;
RUN;

/*********ME2_NT***********/
DATA ME2N_TTYP1 (KEEP=Condition Sampnum ME2N_TTYP1);
IF _N_=1 THEN ME2N_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1=ME2N_TTYP1 + (1/500);
RETAIN ME2N_TTYP1;
RUN;

/*ME2NT2Z_TTYP1                 ME2NGZ_TTYP1                    ME2NT2GZ_TTYP1 */
DATA ME2NT2Z_TTYP1 (KEEP=Condition Sampnum ME2NT2Z_TTYP1);
IF _N_=1 THEN ME2NT2Z_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma^=0 & wXT2_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2NT2Z_TTYP1=ME2NT2Z_TTYP1+(1/500);
RETAIN ME2NT2Z_TTYP1;
```

```
RUN;

DATA ME2NGZ_TTYP1 (KEEP=Condition Sampnum ME2NGZ_TTYP1);
IF _N_=1 THEN ME2NGZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXT2_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2NGZ_TTYP1=ME2NGZ_TTYP1 + (1/500);
RETAIN ME2NGZ_TTYP1;
RUN;
DATA ME2NT2GZ_TTYP1 (KEEP=Condition Sampnum ME2NT2GZ_TTYP1);
IF _N_=1 THEN ME2NT2GZ_TTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXT2_N>0 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2NT2GZ_TTYP1=ME2NT2GZ_TTYP1 + (1/500);
RETAIN ME2NT2GZ_TTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXT2_N wXT2_N wXT2_N WXT2 wXT2_N and wXT2_N values below are =0 weights */
DATA ME2N_TTYP1_10 (KEEP=Condition Sampnum   ME2N_TTYP1_10);
IF _N_=1 THEN ME2N_TTYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0.10 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1_10=ME2N_TTYP1_10 + (1/500);
RETAIN ME2N_TTYP1_10;
RUN;

DATA ME2N_TTYP1_20 (KEEP=Condition Sampnum ME2N_TTYP1_20);
IF _N_=1 THEN ME2N_TTYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0.20 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1_20=ME2N_TTYP1_20 + (1/500);
RETAIN ME2N_TTYP1_20;
RUN;

DATA ME2N_TTYP1_30 (KEEP=Condition Sampnum ME2N_TTYP1_30);
IF _N_=1 THEN ME2N_TTYP1_30=0;
```

```
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0.30 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1_30=ME2N_TTYP1_30 + (1/500);
RETAIN ME2N_TTYP1_30;
RUN;

DATA ME2N_TTYP1_40 (KEEP=Condition Sampnum ME2N_TTYP1_40);
IF _N_=1 THEN ME2N_TTYP1_40=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0.40 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1_40=ME2N_TTYP1_40 + (1/500);
RETAIN ME2N_TTYP1_40;
RUN;

DATA ME2N_TTYP1_50 (KEEP=Condition Sampnum ME2N_TTYP1_50);
IF _N_=1 THEN ME2N_TTYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXT2_N>0.50 & ME12=0 &
^(PNT_ME12_2_5<=0<=PNT_ME12_97_5)) THEN
ME2N_TTYP1_50=ME2N_TTYP1_50 + (1/500);
RETAIN ME2N_TTYP1_50;
RUN;

DATA ME2N_TPWR (KEEP= Condition Sampnum ME2N_TPWR);
IF _N_=1 THEN ME2N_TPWR=0;
SET WQSEST2v_500;
IF ((ME2^=0) & wXT2_N>0) THEN
ME2N_TPWR=ME2N_TPWR + (1/500);
RETAIN ME2N_TPWR;
RUN;

DATA ME2N_TPWR_10 (KEEP= Condition Sampnum ME2N_TPWR_10);
IF _N_=1 THEN ME2N_TPWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_N>0.10) THEN
ME2N_TPWR_10=ME2N_TPWR_10 + (1/500);
RETAIN ME2N_TPWR_10;
RUN;
```

```
DATA ME2N_TPWR_20 (KEEP= Condition Sampnum ME2N_TPWR_20);
IF _N_=1 THEN ME2N_TPWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_N>0.20) THEN
ME2N_TPWR_20=ME2N_TPWR_20 + (1/500);
RETAIN ME2N_TPWR_20;
RUN;

DATA ME2N_TPWR_30 (KEEP= Condition Sampnum ME2N_TPWR_30);
IF _N_=1 THEN ME2N_TPWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_N>0.30) THEN
ME2N_TPWR_30=ME2N_TPWR_30 + (1/500);
RETAIN ME2N_TPWR_30;
RUN;

DATA ME2N_TPWR_40 (KEEP= Condition Sampnum ME2N_TPWR_40);
IF _N_=1 THEN ME2N_TPWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_N>0.40) THEN
ME2N_TPWR_40=ME2N_TPWR_40 + (1/500);
RETAIN ME2N_TPWR_40;
RUN;

DATA ME2N_TPWR_50 (KEEP= Condition Sampnum ME2N_TPWR_50);
IF _N_=1 THEN ME2N_TPWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXT2_N>0.50) THEN
ME2N_TPWR_50=ME2N_TPWR_50 + (1/500);
RETAIN ME2N_TPWR_50;
RUN;

/*****************ME1B***************/
DATA ME1N_BTYP1 (KEEP=Condition Sampnum ME1N_BTYP1);
IF _N_=1 THEN ME1N_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1=ME1N_BTYP1 + (1/500);
```

```
RETAIN ME1N_BTYP1;
RUN;

/*ME1NT1Z_BTYP1           ME1NGZ_BTYP1           ME1NT1GZ_BTYP1 */
DATA ME1NT1Z_BTYP1 (KEEP=Condition Sampnum ME1NT1Z_BTYP1);
IF _N_=1 THEN ME1NT1Z_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma^=0 & wXB1_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1NT1Z_BTYP1=ME1NT1Z_BTYP1 + (1/500);
RETAIN ME1NT1Z_BTYP1;
RUN;

DATA ME1NGZ_BTYP1 (KEEP=Condition Sampnum ME1NGZ_BTYP1);
IF _N_=1 THEN       ME1NGZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1^=0 & Gamma=0 & wXB1_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1NGZ_BTYP1=ME1NGZ_BTYP1 + (1/500);
RETAIN ME1NGZ_BTYP1;
RUN;

DATA ME1NT1GZ_BTYP1 (KEEP=Condition Sampnum ME1NT1GZ_BTYP1);
IF _N_=1 THEN       ME1NT1GZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME1=0 & Theta1=0 & Gamma=0 & wXB1_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1NT1GZ_BTYP1=ME1NT1GZ_BTYP1 + (1/500);
RETAIN ME1NT1GZ_BTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB01_N wXB01_N wXB01_N WXT2 wXB01_N and wXB01_N values below are =0 weights */
DATA ME1N_BTYP1_10 (KEEP=Condition Sampnum ME1N_BTYP1_10);
IF _N_=1 THEN ME1N_BTYP1_10=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0.10 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1_10=ME1N_BTYP1_10 + (1/500);
RETAIN ME1N_BTYP1_10;
RUN;
```

```
DATA ME1N_BTYP1_20 (KEEP=Condition Sampnum   ME1N_BTYP1_20);
IF _N_=1 THEN ME1N_BTYP1_20=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0.20 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1_20=ME1N_BTYP1_20 + (1/500);
RETAIN ME1N_BTYP1_20;
RUN;

DATA ME1N_BTYP1_30 (KEEP=Condition Sampnum ME1N_BTYP1_30);
IF _N_=1 THEN ME1N_BTYP1_30=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0.30 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1_30=ME1N_BTYP1_30 + (1/500);
RETAIN ME1N_BTYP1_30;
RUN;

DATA ME1N_BTYP1_40 (KEEP=Condition Sampnum ME1N_BTYP1_40);
IF _N_=1 THEN ME1N_BTYP1_40=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0.40 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1_40=ME1N_BTYP1_40 + (1/500);
RETAIN ME1N_BTYP1_40;
RUN;

DATA ME1N_BTYP1_50 (KEEP=Condition Sampnum ME1N_BTYP1_50);
IF _N_=1 THEN ME1N_BTYP1_50=0;
SET WQSEST2v_500;
IF (ME1=0 & wXB1_N>0.50 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME1N_BTYP1_50=ME1N_BTYP1_50 + (1/500);
RETAIN ME1N_BTYP1_50;
RUN;

DATA ME1N_BPWR (KEEP= Condition Sampnum ME1N_BPWR);
IF _N_=1 THEN ME1N_BPWR=0;
SET WQSEST2v_500;
```

```
IF ((ME1^=0) & wXB1_N>0) THEN
ME1N_BPWR=ME1N_BPWR + (1/500);
RETAIN ME1N_BPWR;
RUN;

DATA ME1N_BPWR_10 (KEEP= Condition Sampnum ME1N_BPWR_10);
IF _N_=1 THEN ME1N_BPWR_10=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_N>0.10) THEN
ME1N_BPWR_10=ME1N_BPWR_10 + (1/500);
RETAIN ME1N_BPWR_10;
RUN;

DATA ME1N_BPWR_20 (KEEP= Condition Sampnum ME1N_BPWR_20);
IF _N_=1 THEN ME1N_BPWR_20=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_N>0.20) THEN
ME1N_BPWR_20=ME1N_BPWR_20 + (1/500);
RETAIN ME1N_BPWR_20;
RUN;

DATA ME1N_BPWR_30 (KEEP= Condition Sampnum ME1N_BPWR_30);
IF _N_=1 THEN ME1N_BPWR_30=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_N>0.30) THEN
ME1N_BPWR_30=ME1N_BPWR_30 + (1/500);
RETAIN ME1N_BPWR_30;
RUN;

DATA ME1N_BPWR_40 (KEEP= Condition Sampnum ME1N_BPWR_40);
IF _N_=1 THEN ME1N_BPWR_40=0;
SET WQSEST2v_500;
IF (ME1^=0 & wXB1_N>0.40) THEN
ME1N_BPWR_40=ME1N_BPWR_40 + (1/500);
RETAIN ME1N_BPWR_40;
RUN;

DATA ME1N_BPWR_50 (KEEP= Condition Sampnum ME1N_BPWR_50);
IF _N_=1 THEN ME1N_BPWR_50=0;
SET WQSEST2v_500;
```

```
IF (ME1^=0 & wXB1_N>0.50) THEN
ME1N_BPWR_50=ME1N_BPWR_50 + (1/500);
RETAIN ME1N_BPWR_50;
RUN;

/*********ME2_NB***********/

DATA ME2N_BTYP1 (KEEP=Condition Sampnum ME2N_BTYP1);
IF _N_=1 THEN ME2N_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1=ME2N_BTYP1 + (1/500);
RETAIN ME2N_BTYP1;
RUN;

/*ME2NT2Z_BTYP1                ME2NGZ_BTYP1                    ME2NT2GZ_BTYP1 */
DATA ME2NT2Z_BTYP1 (KEEP=Condition Sampnum ME2NT2Z_BTYP1);
IF _N_=1 THEN ME2NT2Z_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0  & Theta2=0 & Gamma^=0   & wXB2_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2NT2Z_BTYP1=ME2NT2Z_BTYP1 + (1/500);
RETAIN ME2NT2Z_BTYP1;
RUN;

DATA ME2NGZ_BTYP1 (KEEP=Condition Sampnum ME2NGZ_BTYP1);
IF _N_=1 THEN ME2NGZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2^=0 & Gamma=0 & wXB2_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2NGZ_BTYP1=ME2NGZ_BTYP1 + (1/500);
RETAIN ME2NGZ_BTYP1;
RUN;

DATA ME2NT2GZ_BTYP1 (KEEP=Condition Sampnum ME2NT2GZ_BTYP1);
IF _N_=1 THEN ME2NT2GZ_BTYP1=0;
SET WQSEST2v_500;
IF (ME2=0 & Theta2=0 & Gamma=0 & wXB2_N>0 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
```

```
ME2NT2GZ_BTYP1=ME2NT2GZ_BTYP1 + (1/500);
RETAIN ME2NT2GZ_BTYP1;
RUN;

/* Cut-off value of 0.10 for WQS weights wXB2_N wXB2_N wXB2_N WXT2 wXB2_N and wXB2_N values below are =0 weights */
DATA ME2N_BTYP1_10 (KEEP=Condition Sampnum   ME2N_BTYP1_10);
IF _N_=1 THEN ME2N_BTYP1_10=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0.10 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1_10=ME2N_BTYP1_10 + (1/500);
RETAIN ME2N_BTYP1_10;
RUN;

DATA ME2N_BTYP1_20 (KEEP=Condition Sampnum ME2N_BTYP1_20);
IF _N_=1 THEN ME2N_BTYP1_20=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0.20 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1_20=ME2N_BTYP1_20 + (1/500);
RETAIN ME2N_BTYP1_20;
RUN;

DATA ME2N_BTYP1_30 (KEEP=Condition Sampnum ME2N_BTYP1_30);
IF _N_=1 THEN ME2N_BTYP1_30=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0.30 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1_30=ME2N_BTYP1_30 + (1/500);
RETAIN ME2N_BTYP1_30;
RUN;

DATA ME2N_BTYP1_40 (KEEP=Condition Sampnum ME2N_BTYP1_40);
IF _N_=1 THEN ME2N_BTYP1_40=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0.40 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1_40=ME2N_BTYP1_40 + (1/500);
RETAIN ME2N_BTYP1_40;
RUN;
```

```
DATA ME2N_BTYP1_50 (KEEP=Condition Sampnum ME2N_BTYP1_50);
IF _N_=1 THEN ME2N_BTYP1_50=0;
SET WQSEST2v_500;
IF (ME2=0 & wXB2_N>0.50 & ME12=0 &
^(PNB_ME12_2_5<=0<=PNB_ME12_97_5)) THEN
ME2N_BTYP1_50=ME2N_BTYP1_50 + (1/500);
RETAIN ME2N_BTYP1_50;
RUN;

DATA ME2N_BPWR (KEEP= Condition Sampnum ME2N_BPWR);
IF _N_=1 THEN ME2N_BPWR=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0) THEN
ME2N_BPWR=ME2N_BPWR + (1/500);
RETAIN ME2N_BPWR;
RUN;

DATA ME2N_BPWR_10 (KEEP= Condition Sampnum ME2N_BPWR_10);
IF _N_=1 THEN ME2N_BPWR_10=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0.10) THEN
ME2N_BPWR_10=ME2N_BPWR_10 + (1/500);
RETAIN ME2N_BPWR_10;
RUN;

DATA ME2N_BPWR_20 (KEEP= Condition Sampnum ME2N_BPWR_20);
IF _N_=1 THEN ME2N_BPWR_20=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0.20) THEN
ME2N_BPWR_20=ME2N_BPWR_20 + (1/500);
RETAIN ME2N_BPWR_20;
RUN;

DATA ME2N_BPWR_30 (KEEP= Condition Sampnum ME2N_BPWR_30);
IF _N_=1 THEN ME2N_BPWR_30=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0.30) THEN
ME2N_BPWR_30=ME2N_BPWR_30 + (1/500);
RETAIN ME2N_BPWR_30;
```

```
RUN;

DATA ME2N_BPWR_40 (KEEP= Condition Sampnum ME2N_BPWR_40);
IF _N_=1 THEN ME2N_BPWR_40=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0.40) THEN
ME2N_BPWR_40=ME2N_BPWR_40 + (1/500);
RETAIN ME2N_BPWR_40;
RUN;

DATA ME2N_BPWR_50 (KEEP= Condition Sampnum ME2N_BPWR_50);
IF _N_=1 THEN ME2N_BPWR_50=0;
SET WQSEST2v_500;
IF (ME2^=0 & wXB2_N>0.50) THEN
ME2N_BPWR_50=ME2N_BPWR_50 + (1/500);
RETAIN ME2N_BPWR_50;
RUN;

DATA WQS2V_TYP1PWR;   /* Has 132 + 270 variables =481*/
MERGE   WQSEST2v_1
ME12S_0TYP1 (WHERE= (Sampnum=500))
ME12S_0PWR (WHERE= (Sampnum=500))
ME12N_0TYP1 (WHERE= (Sampnum=500))
ME12N_0PWR (WHERE= (Sampnum=500))

ME12S_TTYP1 (WHERE= (Sampnum=500))
ME12S_TPWR (WHERE= (Sampnum=500))
ME12N_TTYP1 (WHERE= (Sampnum=500))
ME12N_TPWR (WHERE= (Sampnum=500))

ME12S_BTYP1 (WHERE= (Sampnum=500))
ME12S_BPWR (WHERE= (Sampnum=500))
ME12N_BTYP1 (WHERE= (Sampnum=500))
ME12N_BPWR (WHERE= (Sampnum=500))
/*************************/
ME1S_0TYP1 (WHERE= (Sampnum=500))
ME1ST1Z_0TYP1 (WHERE= (Sampnum=500))
ME1SGZ_0TYP1 (WHERE= (Sampnum=500))
ME1ST1GZ_0TYP1 (WHERE= (Sampnum=500))
ME1S_0PWR (WHERE= (Sampnum=500))
```

```
ME1S_0TYP1_10 (WHERE= (Sampnum=500))
ME1S_0PWR_10 (WHERE= (Sampnum=500))
ME1S_0TYP1_20 (WHERE= (Sampnum=500))
ME1S_0PWR_20 (WHERE= (Sampnum=500))
ME1S_0TYP1_30 (WHERE= (Sampnum=500))
ME1S_0PWR_30 (WHERE= (Sampnum=500))
ME1S_0TYP1_40 (WHERE= (Sampnum=500))
ME1S_0PWR_40 (WHERE= (Sampnum=500))
ME1S_0TYP1_50 (WHERE= (Sampnum=500))
ME1S_0PWR_50 (WHERE= (Sampnum=500))

ME2S_0PWR (WHERE= (Sampnum=500))
ME2S_0TYP1 (WHERE= (Sampnum=500))
ME2ST2Z_0TYP1 (WHERE= (Sampnum=500))
ME2SGZ_0TYP1 (WHERE= (Sampnum=500))
ME2ST2GZ_0TYP1 (WHERE= (Sampnum=500))
ME2S_0TYP1_10 (WHERE= (Sampnum=500))
ME2S_0PWR_10 (WHERE= (Sampnum=500))
ME2S_0TYP1_20 (WHERE= (Sampnum=500))
ME2S_0PWR_20 (WHERE= (Sampnum=500))
ME2S_0TYP1_30 (WHERE= (Sampnum=500))
ME2S_0PWR_30 (WHERE= (Sampnum=500))
ME2S_0TYP1_40 (WHERE= (Sampnum=500))
ME2S_0PWR_40 (WHERE= (Sampnum=500))
ME2S_0TYP1_50 (WHERE= (Sampnum=500))
ME2S_0PWR_50 (WHERE= (Sampnum=500))
/*************************/
ME1N_0TYP1 (WHERE= (Sampnum=500))
ME1NT1Z_0TYP1 (WHERE= (Sampnum=500))
ME1NGZ_0TYP1 (WHERE= (Sampnum=500))
ME1NT1GZ_0TYP1 (WHERE= (Sampnum=500))
ME1N_0PWR (WHERE= (Sampnum=500))
ME1N_0TYP1_10 (WHERE= (Sampnum=500))
ME1N_0PWR_10 (WHERE= (Sampnum=500))
ME1N_0TYP1_20 (WHERE= (Sampnum=500))
ME1N_0PWR_20 (WHERE= (Sampnum=500))
ME1N_0TYP1_30 (WHERE= (Sampnum=500))
ME1N_0PWR_30 (WHERE= (Sampnum=500))
ME1N_0TYP1_40 (WHERE= (Sampnum=500))
ME1N_0PWR_40 (WHERE= (Sampnum=500))
```

```
ME1N_0TYP1_50 (WHERE= (Sampnum=500))
ME1N_0PWR_50 (WHERE= (Sampnum=500))

ME2N_0TYP1 (WHERE= (Sampnum=500))
ME2NT2Z_0TYP1 (WHERE= (Sampnum=500))
ME2NGZ_0TYP1 (WHERE= (Sampnum=500))
ME2NT2GZ_0TYP1 (WHERE= (Sampnum=500))
ME2N_0PWR (WHERE= (Sampnum=500))
ME2N_0TYP1_10 (WHERE= (Sampnum=500))
ME2N_0PWR_10 (WHERE= (Sampnum=500))
ME2N_0TYP1_20 (WHERE= (Sampnum=500))
ME2N_0PWR_20 (WHERE= (Sampnum=500))
ME2N_0TYP1_30 (WHERE= (Sampnum=500))
ME2N_0PWR_30 (WHERE= (Sampnum=500))
ME2N_0TYP1_40 (WHERE= (Sampnum=500))
ME2N_0PWR_40 (WHERE= (Sampnum=500))
ME2N_0TYP1_50 (WHERE= (Sampnum=500))
ME2N_0PWR_50 (WHERE= (Sampnum=500))
/*************************/

ME1S_TTYP1 (WHERE= (Sampnum=500))
ME1ST1Z_TTYP1 (WHERE= (Sampnum=500))
ME1SGZ_TTYP1 (WHERE= (Sampnum=500))
ME1ST1GZ_TTYP1 (WHERE= (Sampnum=500))
ME1S_TPWR (WHERE= (Sampnum=500))
ME1S_TTYP1_10 (WHERE= (Sampnum=500))
ME1S_TPWR_10 (WHERE= (Sampnum=500))
ME1S_TTYP1_20 (WHERE= (Sampnum=500))
ME1S_TPWR_20 (WHERE= (Sampnum=500))
ME1S_TTYP1_30 (WHERE= (Sampnum=500))
ME1S_TPWR_30 (WHERE= (Sampnum=500))
ME1S_TTYP1_40 (WHERE= (Sampnum=500))
ME1S_TPWR_40 (WHERE= (Sampnum=500))
ME1S_TTYP1_50 (WHERE= (Sampnum=500))
ME1S_TPWR_50 (WHERE= (Sampnum=500))

ME2S_TPWR (WHERE= (Sampnum=500))
ME2S_TTYP1 (WHERE= (Sampnum=500))
ME2ST2Z_TTYP1 (WHERE= (Sampnum=500))
ME2SGZ_TTYP1 (WHERE= (Sampnum=500))
```

```
ME2ST2GZ_TTYP1 (WHERE= (Sampnum=500))
ME2S_TTYP1_10 (WHERE= (Sampnum=500))
ME2S_TPWR_10 (WHERE= (Sampnum=500))
ME2S_TTYP1_20 (WHERE= (Sampnum=500))
ME2S_TPWR_20 (WHERE= (Sampnum=500))
ME2S_TTYP1_30 (WHERE= (Sampnum=500))
ME2S_TPWR_30 (WHERE= (Sampnum=500))
ME2S_TTYP1_40 (WHERE= (Sampnum=500))
ME2S_TPWR_40 (WHERE= (Sampnum=500))
ME2S_TTYP1_50 (WHERE= (Sampnum=500))
ME2S_TPWR_50 (WHERE= (Sampnum=500))
/*************************/
ME1N_TTYP1 (WHERE= (Sampnum=500))
ME1NT1Z_TTYP1 (WHERE= (Sampnum=500))
ME1NGZ_TTYP1 (WHERE= (Sampnum=500))
ME1NT1GZ_TTYP1 (WHERE= (Sampnum=500))
ME1N_TPWR (WHERE= (Sampnum=500))
ME1N_TTYP1_10 (WHERE= (Sampnum=500))
ME1N_TPWR_10 (WHERE= (Sampnum=500))
ME1N_TTYP1_20 (WHERE= (Sampnum=500))
ME1N_TPWR_20 (WHERE= (Sampnum=500))
ME1N_TTYP1_30 (WHERE= (Sampnum=500))
ME1N_TPWR_30 (WHERE= (Sampnum=500))
ME1N_TTYP1_40 (WHERE= (Sampnum=500))
ME1N_TPWR_40 (WHERE= (Sampnum=500))
ME1N_TTYP1_50 (WHERE= (Sampnum=500))
ME1N_TPWR_50 (WHERE= (Sampnum=500))

ME2N_TTYP1 (WHERE= (Sampnum=500))
ME2NT2Z_TTYP1 (WHERE= (Sampnum=500))
ME2NGZ_TTYP1 (WHERE= (Sampnum=500))
ME2NT2GZ_TTYP1 (WHERE= (Sampnum=500))
ME2N_TPWR (WHERE= (Sampnum=500))
ME2N_TTYP1_10 (WHERE= (Sampnum=500))
ME2N_TPWR_10 (WHERE= (Sampnum=500))
ME2N_TTYP1_20 (WHERE= (Sampnum=500))
ME2N_TPWR_20 (WHERE= (Sampnum=500))
ME2N_TTYP1_30 (WHERE= (Sampnum=500))
ME2N_TPWR_30 (WHERE= (Sampnum=500))
ME2N_TTYP1_40 (WHERE= (Sampnum=500))
```

```
ME2N_TPWR_40 (WHERE= (Sampnum=500))
ME2N_TTYP1_50 (WHERE= (Sampnum=500))
ME2N_TPWR_50 (WHERE= (Sampnum=500))
/*************************/
ME1S_BTYP1 (WHERE= (Sampnum=500))
ME1ST1Z_BTYP1 (WHERE= (Sampnum=500))
ME1SGZ_BTYP1 (WHERE= (Sampnum=500))
ME1ST1GZ_BTYP1 (WHERE= (Sampnum=500))
ME1S_BPWR (WHERE= (Sampnum=500))
ME1S_BTYP1_10 (WHERE= (Sampnum=500))
ME1S_BPWR_10 (WHERE= (Sampnum=500))
ME1S_BTYP1_20 (WHERE= (Sampnum=500))
ME1S_BPWR_20 (WHERE= (Sampnum=500))
ME1S_BTYP1_30 (WHERE= (Sampnum=500))
ME1S_BPWR_30 (WHERE= (Sampnum=500))
ME1S_BTYP1_40 (WHERE= (Sampnum=500))
ME1S_BPWR_40 (WHERE= (Sampnum=500))
ME1S_BTYP1_50 (WHERE= (Sampnum=500))
ME1S_BPWR_50 (WHERE= (Sampnum=500))

ME2S_BPWR (WHERE= (Sampnum=500))
ME2S_BTYP1 (WHERE= (Sampnum=500))
ME2ST2Z_BTYP1 (WHERE= (Sampnum=500))
ME2SGZ_BTYP1 (WHERE= (Sampnum=500))
ME2ST2GZ_BTYP1 (WHERE= (Sampnum=500))
ME2S_BTYP1_10 (WHERE= (Sampnum=500))
ME2S_BPWR_10 (WHERE= (Sampnum=500))
ME2S_BTYP1_20 (WHERE= (Sampnum=500))
ME2S_BPWR_20 (WHERE= (Sampnum=500))
ME2S_BTYP1_30 (WHERE= (Sampnum=500))
ME2S_BPWR_30 (WHERE= (Sampnum=500))
ME2S_BTYP1_40 (WHERE= (Sampnum=500))
ME2S_BPWR_40 (WHERE= (Sampnum=500))
ME2S_BTYP1_50 (WHERE= (Sampnum=500))
ME2S_BPWR_50 (WHERE= (Sampnum=500))
/*************************/
ME1N_BTYP1 (WHERE= (Sampnum=500))
ME1NT1Z_BTYP1 (WHERE= (Sampnum=500))
ME1NGZ_BTYP1 (WHERE= (Sampnum=500))
ME1NT1GZ_BTYP1 (WHERE= (Sampnum=500))
```

```
ME1N_BPWR (WHERE= (Sampnum=500))
ME1N_BTYP1_10 (WHERE= (Sampnum=500))
ME1N_BPWR_10 (WHERE= (Sampnum=500))
ME1N_BTYP1_20 (WHERE= (Sampnum=500))
ME1N_BPWR_20 (WHERE= (Sampnum=500))
ME1N_BTYP1_30 (WHERE= (Sampnum=500))
ME1N_BPWR_30 (WHERE= (Sampnum=500))
ME1N_BTYP1_40 (WHERE= (Sampnum=500))
ME1N_BPWR_40 (WHERE= (Sampnum=500))
ME1N_BTYP1_50 (WHERE= (Sampnum=500))
ME1N_BPWR_50 (WHERE= (Sampnum=500))

ME2N_BTYP1 (WHERE= (Sampnum=500))
ME2NT2Z_BTYP1 (WHERE= (Sampnum=500))
ME2NGZ_BTYP1 (WHERE= (Sampnum=500))
ME2NT2GZ_BTYP1 (WHERE= (Sampnum=500))
ME2N_BPWR (WHERE= (Sampnum=500))
ME2N_BTYP1_10 (WHERE= (Sampnum=500))
ME2N_BPWR_10 (WHERE= (Sampnum=500))
ME2N_BTYP1_20 (WHERE= (Sampnum=500))
ME2N_BPWR_20 (WHERE= (Sampnum=500))
ME2N_BTYP1_30 (WHERE= (Sampnum=500))
ME2N_BPWR_30 (WHERE= (Sampnum=500))
ME2N_BTYP1_40 (WHERE= (Sampnum=500))
ME2N_BPWR_40 (WHERE= (Sampnum=500))
ME2N_BTYP1_50 (WHERE= (Sampnum=500))
ME2N_BPWR_50 (WHERE= (Sampnum=500))
/*************************/
DE12_0TYP1 (WHERE= (Sampnum=500))
DE12_0PWR (WHERE= (Sampnum=500))
DE12_TTYP1 (WHERE= (Sampnum=500))
DE12_TPWR (WHERE= (Sampnum=500))
DE12_BTYP1 (WHERE= (Sampnum=500))
DE12_BPWR (WHERE= (Sampnum=500))
/*************************/
DE1_0TYP1 (WHERE= (Sampnum=500))
DE1_0TYP1_10 (WHERE= (Sampnum=500))
DE1_0TYP1_20 (WHERE= (Sampnum=500))
DE1_0TYP1_30 (WHERE= (Sampnum=500))
DE1_0TYP1_40 (WHERE= (Sampnum=500))
```

```
DE1_0TYP1_50 (WHERE= (Sampnum=500))

DE1_0PWR (WHERE= (Sampnum=500))
DE1_0PWR_10 (WHERE= (Sampnum=500))
DE1_0PWR_20 (WHERE= (Sampnum=500))
DE1_0PWR_30 (WHERE= (Sampnum=500))
DE1_0PWR_40 (WHERE= (Sampnum=500))
DE1_0PWR_50 (WHERE= (Sampnum=500))

DE2_0TYP1 (WHERE= (Sampnum=500))
DE2_0TYP1_10 (WHERE= (Sampnum=500))
DE2_0TYP1_20 (WHERE= (Sampnum=500))
DE2_0TYP1_30 (WHERE= (Sampnum=500))
DE2_0TYP1_40 (WHERE= (Sampnum=500))
DE2_0TYP1_50 (WHERE= (Sampnum=500))

DE2_0PWR (WHERE= (Sampnum=500))
DE2_0PWR_10 (WHERE= (Sampnum=500))
DE2_0PWR_20 (WHERE= (Sampnum=500))
DE2_0PWR_30 (WHERE= (Sampnum=500))
DE2_0PWR_40 (WHERE= (Sampnum=500))
DE2_0PWR_50 (WHERE= (Sampnum=500))
/****************************/
DE1_TTYP1 (WHERE= (Sampnum=500))
DE1_TTYP1_10 (WHERE= (Sampnum=500))
DE1_TTYP1_20 (WHERE= (Sampnum=500))
DE1_TTYP1_30 (WHERE= (Sampnum=500))
DE1_TTYP1_40 (WHERE= (Sampnum=500))
DE1_TTYP1_50 (WHERE= (Sampnum=500))

DE1_TPWR (WHERE= (Sampnum=500))
DE1_TPWR_10 (WHERE= (Sampnum=500))
DE1_TPWR_20 (WHERE= (Sampnum=500))
DE1_TPWR_30 (WHERE= (Sampnum=500))
DE1_TPWR_40 (WHERE= (Sampnum=500))
DE1_TPWR_50 (WHERE= (Sampnum=500))

DE2_TTYP1 (WHERE= (Sampnum=500))
DE2_TTYP1_10 (WHERE= (Sampnum=500))
DE2_TTYP1_20 (WHERE= (Sampnum=500))
```

```
DE2_TTYP1_30 (WHERE= (Sampnum=500))
DE2_TTYP1_40 (WHERE= (Sampnum=500))
DE2_TTYP1_50 (WHERE= (Sampnum=500))

DE2_TPWR (WHERE= (Sampnum=500))
DE2_TPWR_10 (WHERE= (Sampnum=500))
DE2_TPWR_20 (WHERE= (Sampnum=500))
DE2_TPWR_30 (WHERE= (Sampnum=500))
DE2_TPWR_40 (WHERE= (Sampnum=500))
DE2_TPWR_50 (WHERE= (Sampnum=500))
/*************************/
DE1_BTYP1 (WHERE= (Sampnum=500))
DE1_BTYP1_10 (WHERE= (Sampnum=500))
DE1_BTYP1_20 (WHERE= (Sampnum=500))
DE1_BTYP1_30 (WHERE= (Sampnum=500))
DE1_BTYP1_40 (WHERE= (Sampnum=500))
DE1_BTYP1_50 (WHERE= (Sampnum=500))

DE1_BPWR (WHERE= (Sampnum=500))
DE1_BPWR_10 (WHERE= (Sampnum=500))
DE1_BPWR_20 (WHERE= (Sampnum=500))
DE1_BPWR_30 (WHERE= (Sampnum=500))
DE1_BPWR_40 (WHERE= (Sampnum=500))
DE1_BPWR_50 (WHERE= (Sampnum=500))

DE2_BTYP1 (WHERE= (Sampnum=500))
DE2_BTYP1_10 (WHERE= (Sampnum=500))
DE2_BTYP1_20 (WHERE= (Sampnum=500))
DE2_BTYP1_30 (WHERE= (Sampnum=500))
DE2_BTYP1_40 (WHERE= (Sampnum=500))
DE2_BTYP1_50 (WHERE= (Sampnum=500))

DE2_BPWR (WHERE= (Sampnum=500))
DE2_BPWR_10 (WHERE= (Sampnum=500))
DE2_BPWR_20 (WHERE= (Sampnum=500))
DE2_BPWR_30 (WHERE= (Sampnum=500))
DE2_BPWR_40 (WHERE= (Sampnum=500))
DE2_BPWR_50 (WHERE= (Sampnum=500)) ;
BY Condition;
```

```
IF DE12^=0     & DE12_0TYP1 =0     THEN DE12_0TYP1=.;
IF DE12=0      & DE12_0PWR =0      THEN DE12_0PWR=.;
IF DE12^=0     & DE12_TTYP1 =0     THEN DE12_TTYP1=.;
IF DE12=0      & DE12_TPWR =0      THEN DE12_TPWR=.;
IF DE12^=0     & DE12_BTYP1 =0     THEN DE12_BTYP1=.;
IF DE12=0      & DE12_BPWR =0      THEN DE12_BPWR=.;
IF DE1^=0      & DE1_0TYP1 =0      THEN DE1_0TYP1=.;
IF DE1^=0      & DE1_0TYP1_10 =0   THEN DE1_0TYP1_10=.;
IF DE1^=0      & DE1_0TYP1_20 =0   THEN DE1_0TYP1_20=.;
IF DE1^=0      & DE1_0TYP1_30 =0   THEN DE1_0TYP1_30=.;
IF DE1^=0      & DE1_0TYP1_40 =0   THEN DE1_0TYP1_40=.;
IF DE1^=0      & DE1_0TYP1_50 =0   THEN DE1_0TYP1_50=.;
IF DE1=0       & DE1_0PWR =0       THEN DE1_0PWR=.;
IF DE1=0       & DE1_0PWR_10 =0    THEN DE1_0PWR_10=.;
IF DE1=0       & DE1_0PWR_20 =0    THEN DE1_0PWR_20=.;
IF DE1=0       & DE1_0PWR_30 =0    THEN DE1_0PWR_30=.;
IF DE1=0       & DE1_0PWR_40 =0    THEN DE1_0PWR_40=.;
IF DE1=0       & DE1_0PWR_50 =0    THEN DE1_0PWR_50=.;
IF DE2^=0      & DE2_0TYP1 =0      THEN DE2_0TYP1=.;
IF DE2^=0      & DE2_0TYP1_10 =0   THEN DE2_0TYP1_10=.;
IF DE2^=0      & DE2_0TYP1_20 =0   THEN DE2_0TYP1_20=.;
IF DE2^=0      & DE2_0TYP1_30 =0   THEN DE2_0TYP1_30=.;
IF DE2^=0      & DE2_0TYP1_40 =0   THEN DE2_0TYP1_40=.;
IF DE2^=0      & DE2_0TYP1_50 =0   THEN DE2_0TYP1_50=.;
IF DE2=0       & DE2_0PWR =0       THEN DE2_0PWR=.;
IF DE2=0       & DE2_0PWR_10 =0    THEN DE2_0PWR_10=.;
IF DE2=0       & DE2_0PWR_20 =0    THEN DE2_0PWR_20=.;
IF DE2=0       & DE2_0PWR_30 =0    THEN DE2_0PWR_30=.;
IF DE2=0       & DE2_0PWR_40 =0    THEN DE2_0PWR_40=.;
IF DE2=0       & DE2_0PWR_50 =0    THEN DE2_0PWR_50=.;
IF DE1^=0      & DE1_TTYP1 =0      THEN DE1_TTYP1=.;
IF DE1^=0      & DE1_TTYP1_10 =0   THEN DE1_TTYP1_10=.;
IF DE1^=0      & DE1_TTYP1_20 =0   THEN DE1_TTYP1_20=.;
IF DE1^=0      & DE1_TTYP1_30 =0   THEN DE1_TTYP1_30=.;
IF DE1^=0      & DE1_TTYP1_40 =0   THEN DE1_TTYP1_40=.;
IF DE1^=0      & DE1_TTYP1_50 =0   THEN DE1_TTYP1_50=.;
IF DE1=0       & DE1_TPWR =0       THEN DE1_TPWR=.;
IF DE1=0       & DE1_TPWR_10 =0    THEN DE1_TPWR_10=.;
IF DE1=0       & DE1_TPWR_20 =0    THEN DE1_TPWR_20=.;
IF DE1=0       & DE1_TPWR_30 =0    THEN DE1_TPWR_30=.;
```

```
IF DE1=0      & DE1_TPWR_40 =0    THEN DE1_TPWR_40=.;
IF DE1=0      & DE1_TPWR_50 =0    THEN DE1_TPWR_50=.;
IF DE2^=0     & DE2_TTYP1 =0      THEN DE2_TTYP1=.;
IF DE2^=0     & DE2_TTYP1_10 =0   THEN DE2_TTYP1_10=.;
IF DE2^=0     & DE2_TTYP1_20 =0   THEN DE2_TTYP1_20=.;
IF DE2^=0     & DE2_TTYP1_30 =0   THEN DE2_TTYP1_30=.;
IF DE2^=0     & DE2_TTYP1_40 =0   THEN DE2_TTYP1_40=.;
IF DE2^=0     & DE2_TTYP1_50 =0   THEN DE2_TTYP1_50=.;
IF DE2=0      & DE2_TPWR =0       THEN DE2_TPWR=.;
IF DE2=0      & DE2_TPWR_10 =0    THEN DE2_TPWR_10=.;
IF DE2=0      & DE2_TPWR_20 =0    THEN DE2_TPWR_20=.;
IF DE2=0      & DE2_TPWR_30 =0    THEN DE2_TPWR_30=.;
IF DE2=0      & DE2_TPWR_40 =0    THEN DE2_TPWR_40=.;
IF DE2=0      & DE2_TPWR_50 =0    THEN DE2_TPWR_50=.;
IF DE1^=0     & DE1_BTYP1 =0      THEN DE1_BTYP1=.;
IF DE1^=0     & DE1_BTYP1_10 =0   THEN DE1_BTYP1_10 =.;
IF DE1^=0     & DE1_BTYP1_20 =0   THEN DE1_BTYP1_20 =.;
IF DE1^=0     & DE1_BTYP1_30 =0   THEN DE1_BTYP1_30 =.;
IF DE1^=0     & DE1_BTYP1_40 =0   THEN DE1_BTYP1_40 =.;
IF DE1^=0     & DE1_BTYP1_50 =0   THEN DE1_BTYP1_50=.;
IF DE1=0      & DE1_BPWR   =0     THEN DE1_BPWR =.;
IF DE1=0      & DE1_BPWR_10 =0    THEN DE1_BPWR_10 =.;
IF DE1=0      & DE1_BPWR_20 =0    THEN DE1_BPWR_20 =.;
IF DE1=0      & DE1_BPWR_30 =0    THEN DE1_BPWR_30 =.;
IF DE1=0      & DE1_BPWR_40 =0    THEN DE1_BPWR_40 =.;
IF DE1=0      & DE1_BPWR_50 =0    THEN DE1_BPWR_50=.;
IF DE2^=0     & DE2_BTYP1   =0    THEN DE2_BTYP1=.;
IF DE2^=0     & DE2_BTYP1_10 =0   THEN DE2_BTYP1_10 =.;
IF DE2^=0     & DE2_BTYP1_20 =0   THEN DE2_BTYP1_20 =.;
IF DE2^=0     & DE2_BTYP1_30 =0   THEN DE2_BTYP1_30 =.;
IF DE2^=0     & DE2_BTYP1_40 =0   THEN DE2_BTYP1_40 =.;
IF DE2^=0     & DE2_BTYP1_50 =0   THEN DE2_BTYP1_50=.;
IF DE2=0      & DE2_BPWR   =0     THEN DE2_BPWR=.;
IF DE2=0      & DE2_BPWR_10 =0    THEN DE2_BPWR_10 =.;
IF DE2=0      & DE2_BPWR_20 =0    THEN DE2_BPWR_20 =.;
IF DE2=0      & DE2_BPWR_30 =0    THEN DE2_BPWR_30 =.;
IF DE2=0      & DE2_BPWR_40 =0    THEN DE2_BPWR_40 =.;
IF DE2=0      & DE2_BPWR_50=0     THEN DE2_BPWR_50=.;
IF ME12^=0    & ME12S_0TYP1=0     THEN ME12S_0TYP1 =.;
IF ME12=0     & ME12S_0PWR=0      THEN ME12S_0PWR =.;
```

```
IF ME12^=0     & ME12N_0TYP1=0     THEN ME12N_0TYP1 =.;
IF ME12=0      & ME12N_0PWR=0      THEN ME12N_0PWR=.;
IF ME12^=0     & ME12S_TTYP1=0     THEN ME12S_TTYP1 =.;
IF ME12=0      & ME12S_TPWR=0      THEN ME12S_TPWR =.;
IF ME12^=0     & ME12N_TTYP1=0     THEN ME12N_TTYP1 =.;
IF ME12=0      & ME12N_TPWR=0      THEN ME12N_TPWR=.;
IF ME12^=0     & ME12S_BTYP1=0     THEN ME12S_BTYP1 =.;
IF ME12=0      & ME12S_BPWR=0      THEN ME12S_BPWR =.;
IF ME12^=0     & ME12N_BTYP1=0     THEN ME12N_BTYP1 =.;
IF ME12=0      & ME12N_BPWR=0      THEN ME12N_BPWR=.;

IF ME1^=0      & ME1S_0TYP1=0             THEN ME1S_0TYP1 =.;
/*ME1ST1Z_0TYP1      ME1SGZ_0TYP1         ME1ST1GZ_0TYP1      */
IF ME1^=0       & ME1ST1Z_0TYP1=0         THEN ME1ST1Z_0TYP1 =.;
IF ME1^=0      & ME1SGZ_0TYP1=0           THEN ME1SGZ_0TYP1 =.;
IF ME1^=0      & ME1ST1GZ_0TYP1=0         THEN ME1ST1GZ_0TYP1 =.;
IF ME1=0       & ME1S_0PWR=0              THEN ME1S_0PWR=.;
IF ME1^=0      & ME1S_0TYP1_10=0          THEN ME1S_0TYP1_10 =.;
IF ME1=0       & ME1S_0PWR_10=0           THEN ME1S_0PWR_10 =.;
IF ME1^=0      & ME1S_0TYP1_20=0          THEN ME1S_0TYP1_20 =.;
IF ME1=0       & ME1S_0PWR_20=0           THEN ME1S_0PWR_20 =.;
IF ME1^=0      & ME1S_0TYP1_30=0          THEN ME1S_0TYP1_30=.;
IF ME1=0       & ME1S_0PWR_30=0           THEN ME1S_0PWR_30 =.;
IF ME1^=0      & ME1S_0TYP1_40=0          THEN ME1S_0TYP1_40 =.;
IF ME1=0       & ME1S_0PWR_40=0           THEN ME1S_0PWR_40 =.;
IF ME1^=0      & ME1S_0TYP1_50=0          THEN ME1S_0TYP1_50 =.;
IF ME1=0       & ME1S_0PWR_50=0           THEN ME1S_0PWR_50=.;

IF ME2^=0      & ME2S_0TYP1=0             THEN ME2S_0TYP1 =.;
/*ME2ST2Z_0TYP1      ME2SGZ_0TYP1         ME2ST2GZ_0TYP1      */
IF ME2^=0      & ME2ST2Z_0TYP1=0          THEN ME2ST2Z_0TYP1 =.;
IF ME2^=0      & ME2SGZ_0TYP1=0           THEN ME2SGZ_0TYP1 =.;
IF ME2^=0      & ME2ST2GZ_0TYP1=0         THEN ME2ST2GZ_0TYP1 =.;
IF ME2=0       & ME2S_0PWR =0             THEN ME2S_0PWR=.;
IF ME2^=0      & ME2S_0TYP1_10=0          THEN ME2S_0TYP1_10 =.;
IF ME2=0       & ME2S_0PWR_10=0           THEN ME2S_0PWR_10 =.;
IF ME2^=0      & ME2S_0TYP1_20=0          THEN ME2S_0TYP1_20 =.;
IF ME2=0       & ME2S_0PWR_20=0           THEN ME2S_0PWR_20 =.;
IF ME2^=0      & ME2S_0TYP1_30=0          THEN ME2S_0TYP1_30=.;
IF ME2=0       & ME2S_0PWR_30=0           THEN ME2S_0PWR_30 =.;
```

```
IF ME2^=0       & ME2S_0TYP1_40=0           THEN ME2S_0TYP1_40 =.;
IF ME2=0        & ME2S_0PWR_40=0            THEN ME2S_0PWR_40 =.;
IF ME2^=0       & ME2S_0TYP1_50=0           THEN ME2S_0TYP1_50 =.;
IF ME2=0        & ME2S_0PWR_50=0            THEN ME2S_0PWR_50=.;

IF ME1^=0       & ME1N_0TYP1 =0             THEN ME1N_0TYP1=.;
/*ME1NT1Z_0TYP1      ME1NGZ_0TYP1           ME1NT1GZ_0TYP1      */
IF ME1^=0       & ME1NT1Z_0TYP1=0           THEN ME1NT1Z_0TYP1 =.;
IF ME1^=0       & ME1NGZ_0TYP1=0            THEN ME1NGZ_0TYP1 =.;
IF ME1^=0       & ME1NT1GZ_0TYP1=0          THEN ME1NT1GZ_0TYP1 =.;
IF ME1=0        & ME1N_0PWR=0               THEN ME1N_0PWR=.;
IF ME1^=0       & ME1N_0TYP1_10=0           THEN ME1N_0TYP1_10 =.;
IF ME1=0        & ME1N_0PWR_10=0            THEN ME1N_0PWR_10 =.;
IF ME1^=0       & ME1N_0TYP1_20=0           THEN ME1N_0TYP1_20 =.;
IF ME1=0        & ME1N_0PWR_20=0            THEN ME1N_0PWR_20 =.;
IF ME1^=0       & ME1N_0TYP1_30=0           THEN ME1N_0TYP1_30=.;
IF ME1=0        & ME1N_0PWR_30=0            THEN ME1N_0PWR_30 =.;
IF ME1^=0       & ME1N_0TYP1_40=0           THEN ME1N_0TYP1_40=.;
IF ME1=0        & ME1N_0PWR_40=0            THEN ME1N_0PWR_40 =.;
IF ME1^=0       & ME1N_0TYP1_50=0           THEN ME1N_0TYP1_50=.;
IF ME1=0        & ME1N_0PWR_50=0            THEN ME1N_0PWR_50=.;

IF ME2^=0       & ME2N_0TYP1=0              THEN ME2N_0TYP1 =.;
/*ME2NT2Z_0TYP1      ME2NGZ_0TYP1           ME2NT2GZ_0TYP1      */
IF ME2^=0       & ME2NT2Z_0TYP1=0  T        HEN ME2NT2Z_0TYP1 =.;
IF ME2^=0       & ME2NGZ_0TYP1=0            THEN ME2NGZ_0TYP1 =.;
IF ME2^=0       & ME2NT2GZ_0TYP1=0          THEN ME2NT2GZ_0TYP1 =.;
IF ME2=0        & ME2N_0PWR=0               THEN ME2N_0PWR=.;
IF ME2^=0       & ME2N_0TYP1_10=0           THEN ME2N_0TYP1_10 =.;
IF ME2=0        & ME2N_0PWR_10=0            THEN ME2N_0PWR_10 =.;
IF ME2^=0       & ME2N_0TYP1_20=0           THEN ME2N_0TYP1_20 =.;
IF ME2=0        & ME2N_0PWR_20=0            THEN ME2N_0PWR_20 =.;
IF ME2^=0       & ME2N_0TYP1_30=0           THEN ME2N_0TYP1_30=.;
IF ME2=0        & ME2N_0PWR_30=0            THEN ME2N_0PWR_30 =.;
IF ME2^=0       & ME2N_0TYP1_40=0           THEN ME2N_0TYP1_40=.;
IF ME2=0        & ME2N_0PWR_40=0            THEN ME2N_0PWR_40 =.;
IF ME2^=0       & ME2N_0TYP1_50=0           THEN ME2N_0TYP1_50=.;
IF ME2=0        & ME2N_0PWR_50=0            THEN ME2N_0PWR_50=.;

IF ME1^=0       & ME1S_TTYP1=0              THEN ME1S_TTYP1 =.;
```

```
/*ME1ST1Z_TTYP1     ME1SGZ_TTYP1        ME1ST1GZ_TTYP1     */
IF ME1^=0       & ME1ST1Z_TTYP1=0       THEN ME1ST1Z_TTYP1 =.;
IF ME1^=0       & ME1SGZ_TTYP1=0        THEN ME1SGZ_TTYP1 =.;
IF ME1^=0       & ME1ST1GZ_TTYP1=0      THEN ME1ST1GZ_TTYP1 =.;
IF ME1=0        & ME1S_TPWR=0           THEN ME1S_TPWR=.;
IF ME1^=0       & ME1S_TTYP1_10=0       THEN ME1S_TTYP1_10 =.;
IF ME1=0        & ME1S_TPWR_10=0        THEN ME1S_TPWR_10 =.;
IF ME1^=0       & ME1S_TTYP1_20=0       THEN ME1S_TTYP1_20 =.;
IF ME1=0        & ME1S_TPWR_20=0        THEN ME1S_TPWR_20 =.;
IF ME1^=0       & ME1S_TTYP1_30=0       THEN ME1S_TTYP1_30=.;
IF ME1=0        & ME1S_TPWR_30=0        THEN ME1S_TPWR_30 =.;
IF ME1^=0       & ME1S_TTYP1_40=0       THEN ME1S_TTYP1_40 =.;
IF ME1=0        & ME1S_TPWR_40=0        THEN ME1S_TPWR_40 =.;
IF ME1^=0       & ME1S_TTYP1_50=0       THEN ME1S_TTYP1_50 =.;
IF ME1=0        & ME1S_TPWR_50=0        THEN ME1S_TPWR_50=.;

IF ME2^=0       & ME2S_TTYP1=0          THEN ME2S_TTYP1 =.;
/*ME2ST2Z_TTYP1     ME2SGZ_TTYP1        ME2ST2GZ_TTYP1     */
IF ME2^=0       & ME2ST2Z_TTYP1=0       THEN ME2ST2Z_TTYP1 =.;
IF ME2^=0       & ME2SGZ_TTYP1=0        THEN ME2SGZ_TTYP1 =.;
IF ME2^=0       & ME2ST2GZ_TTYP1=0      THEN ME2ST2GZ_TTYP1 =.;
IF ME2=0        & ME2S_TPWR =0          THEN ME2S_TPWR=.;
IF ME2^=0       & ME2S_TTYP1_10=0       THEN ME2S_TTYP1_10 =.;
IF ME2=0        & ME2S_TPWR_10=0        THEN ME2S_TPWR_10 =.;
IF ME2^=0       & ME2S_TTYP1_20=0       THEN ME2S_TTYP1_20 =.;
IF ME2=0        & ME2S_TPWR_20=0        THEN ME2S_TPWR_20 =.;
IF ME2^=0       & ME2S_TTYP1_30=0       THEN ME2S_TTYP1_30=.;
IF ME2=0        & ME2S_TPWR_30=0        THEN ME2S_TPWR_30 =.;
IF ME2^=0       & ME2S_TTYP1_40=0       THEN ME2S_TTYP1_40 =.;
IF ME2=0        & ME2S_TPWR_40=0        THEN ME2S_TPWR_40 =.;
IF ME2^=0       & ME2S_TTYP1_50=0       THEN ME2S_TTYP1_50 =.;
IF ME2=0        & ME2S_TPWR_50=0        THEN ME2S_TPWR_50=.;

IF ME1^=0       & ME1N_TTYP1 =0         THEN ME1N_TTYP1=.;
/*ME1NT1Z_TTYP1     ME1NGZ_TTYP1        ME1NT1GZ_TTYP1     */
IF ME1^=0       & ME1NT1Z_TTYP1=0       THEN ME1NT1Z_TTYP1 =.;
IF ME1^=0       & ME1NGZ_TTYP1=0        THEN ME1NGZ_TTYP1 =.;
IF ME1^=0       & ME1NT1GZ_TTYP1=0      THEN ME1NT1GZ_TTYP1 =.;
IF ME1=0        & ME1N_TPWR=0           THEN ME1N_TPWR=.;
IF ME1^=0       & ME1N_TTYP1_10=0       THEN ME1N_TTYP1_10 =.;
```

```
IF ME1=0       & ME1N_TPWR_10=0        THEN ME1N_TPWR_10 =.;
IF ME1^=0      & ME1N_TTYP1_20=0       THEN ME1N_TTYP1_20 =.;
IF ME1=0       & ME1N_TPWR_20=0        THEN ME1N_TPWR_20 =.;
IF ME1^=0      & ME1N_TTYP1_30=0       THEN ME1N_TTYP1_30=.;
IF ME1=0       & ME1N_TPWR_30=0        THEN ME1N_TPWR_30 =.;
IF ME1^=0      & ME1N_TTYP1_40=0       THEN ME1N_TTYP1_40=.;
IF ME1=0       & ME1N_TPWR_40=0        THEN ME1N_TPWR_40 =.;
IF ME1^=0      & ME1N_TTYP1_50=0       THEN ME1N_TTYP1_50=.;
IF ME1=0       & ME1N_TPWR_50=0        THEN ME1N_TPWR_50=.;

IF ME2^=0      & ME2N_TTYP1=0          THEN ME2N_TTYP1 =.;
/*ME2NT2Z_TTYP1    ME2NGZ_TTYP1        ME2NT2GZ_TTYP1    */
IF ME2^=0      & ME2NT2Z_TTYP1=0       THEN ME2NT2Z_TTYP1 =.;
IF ME2^=0      & ME2NGZ_TTYP1=0        THEN ME2NGZ_TTYP1 =.;
IF ME2^=0      & ME2NT2GZ_TTYP1=0      THEN ME2NT2GZ_TTYP1 =.;
IF ME2=0       & ME2N_TPWR=0           THEN ME2N_TPWR=.;
IF ME2^=0      & ME2N_TTYP1_10=0       THEN ME2N_TTYP1_10 =.;
IF ME2=0       & ME2N_TPWR_10=0        THEN ME2N_TPWR_10 =.;
IF ME2^=0      & ME2N_TTYP1_20=0       THEN ME2N_TTYP1_20 =.;
IF ME2=0       & ME2N_TPWR_20=0        THEN ME2N_TPWR_20 =.;
IF ME2^=0      & ME2N_TTYP1_30=0       THEN ME2N_TTYP1_30=.;
IF ME2=0       & ME2N_TPWR_30=0        THEN ME2N_TPWR_30 =.;
IF ME2^=0      & ME2N_TTYP1_40=0       THEN ME2N_TTYP1_40=.;
IF ME2=0       & ME2N_TPWR_40=0        THEN ME2N_TPWR_40 =.;
IF ME2^=0      & ME2N_TTYP1_50=0       THEN ME2N_TTYP1_50=.;
IF ME2=0       & ME2N_TPWR_50=0        THEN ME2N_TPWR_50=.;


IF ME1^=0      & ME1S_BTYP1=0          THEN ME1S_BTYP1 =.;
/*ME1ST1Z_BTYP1    ME1SGZ_BTYP1        ME1ST1GZ_BTYP1    */
IF ME1^=0      & ME1ST1Z_BTYP1=0       THEN ME1ST1Z_BTYP1 =.;
IF ME1^=0      & ME1SGZ_BTYP1=0        THEN ME1SGZ_BTYP1 =.;
IF ME1^=0      & ME1ST1GZ_BTYP1=0      THEN ME1ST1GZ_BTYP1 =.;
IF ME1=0       & ME1S_BPWR=0           THEN ME1S_BPWR=.;
IF ME1^=0      & ME1S_BTYP1_10=0       THEN ME1S_BTYP1_10 =.;
IF ME1=0       & ME1S_BPWR_10=0        THEN ME1S_BPWR_10 =.;
IF ME1^=0      & ME1S_BTYP1_20=0       HEN ME1S_BTYP1_20 =.;
IF ME1=0       & ME1S_BPWR_20=0        THEN ME1S_BPWR_20 =.;
IF ME1^=0      & ME1S_BTYP1_30=0       THEN ME1S_BTYP1_30=.;
IF ME1=0       & ME1S_BPWR_30=0        THEN ME1S_BPWR_30 =.;
```

```
IF ME1^=0      & ME1S_BTYP1_40=0          THEN ME1S_BTYP1_40 =.;
IF ME1=0       & ME1S_BPWR_40=0           THEN ME1S_BPWR_40 =.;
IF ME1^=0      & ME1S_BTYP1_50=0          THEN ME1S_BTYP1_50 =.;
IF ME1=0       & ME1S_BPWR_50=0           THEN ME1S_BPWR_50=.;

IF ME2^=0      & ME2S_BTYP1=0             THEN ME2S_BTYP1 =.;
/*ME2ST2Z_BTYP1     ME2SGZ_BTYP1          ME2ST2GZ_BTYP1     */
IF ME2^=0      & ME2ST2Z_BTYP1=0          THEN ME2ST2Z_BTYP1 =.;
IF ME2^=0      & ME2SGZ_BTYP1=0           THEN ME2SGZ_BTYP1 =.;
IF ME2^=0      & ME2ST2GZ_BTYP1=0         THEN ME2ST2GZ_BTYP1 =.;
IF ME2=0       & ME2S_BPWR =0             THEN ME2S_BPWR=.;
IF ME2^=0      & ME2S_BTYP1_10=0          THEN ME2S_BTYP1_10 =.;
IF ME2=0       & ME2S_BPWR_10=0           THEN ME2S_BPWR_10 =.;
IF ME2^=0      & ME2S_BTYP1_20=0          THEN ME2S_BTYP1_20 =.;
IF ME2=0       & ME2S_BPWR_20=0           THEN ME2S_BPWR_20 =.;
IF ME2^=0      & ME2S_BTYP1_30=0          THEN ME2S_BTYP1_30=.;
IF ME2=0       & ME2S_BPWR_30=0           THEN ME2S_BPWR_30 =.;
IF ME2^=0      & ME2S_BTYP1_40=0          THEN ME2S_BTYP1_40 =.;
IF ME2=0       & ME2S_BPWR_40=0           THEN ME2S_BPWR_40 =.;
IF ME2^=0      & ME2S_BTYP1_50=0          THEN ME2S_BTYP1_50 =.;
IF ME2=0       & ME2S_BPWR_50=0           THEN ME2S_BPWR_50=.;

IF ME1^=0      & ME1N_BTYP1 =0            THEN ME1N_BTYP1=.;
/*ME1NT1Z_BTYP1     ME1NGZ_BTYP1          ME1NT1GZ_BTYP1     */
IF ME1^=0      & ME1NT1Z_BTYP1=0          THEN ME1NT1Z_BTYP1 =.;
IF ME1^=0      & ME1NGZ_BTYP1=0           THEN ME1NGZ_BTYP1 =.;
IF ME1^=0      & ME1NT1GZ_BTYP1=0         THEN ME1NT1GZ_BTYP1 =.;
IF ME1=0       & ME1N_BPWR=0              THEN ME1N_BPWR=.;
IF ME1^=0      & ME1N_BTYP1_10=0          THEN ME1N_BTYP1_10 =.;
IF ME1=0       & ME1N_BPWR_10=0           THEN ME1N_BPWR_10 =.;
IF ME1^=0      & ME1N_BTYP1_20=0          THEN ME1N_BTYP1_20 =.;
IF ME1=0       & ME1N_BPWR_20=0           THEN ME1N_BPWR_20 =.;
IF ME1^=0      & ME1N_BTYP1_30=0          THEN ME1N_BTYP1_30=.;
IF ME1=0       & ME1N_BPWR_30=0           THEN ME1N_BPWR_30 =.;
IF ME1^=0      & ME1N_BTYP1_40=0          THEN ME1N_BTYP1_40=.;
IF ME1=0       & ME1N_BPWR_40=0           THEN ME1N_BPWR_40 =.;
IF ME1^=0      & ME1N_BTYP1_50=0          THEN ME1N_BTYP1_50=.;
IF ME1=0       & ME1N_BPWR_50=0           THEN ME1N_BPWR_50=.;

IF ME2^=0      & ME2N_BTYP1=0             THEN ME2N_BTYP1 =.;
```

```
/*ME2NT2Z_BTYP1      ME2NGZ_BTYP1         ME2NT2GZ_BTYP1      */
IF ME2^=0       & ME2NT2Z_BTYP1=0  T        HEN ME2NT2Z_BTYP1 =.;
IF ME2^=0       & ME2NGZ_BTYP1=0            THEN ME2NGZ_BTYP1 =.;
IF ME2^=0       & ME2NT2GZ_BTYP1=0          THEN ME2NT2GZ_BTYP1 =.;
IF ME2=0        & ME2N_BPWR=0               THEN ME2N_BPWR=.;
IF ME2^=0       & ME2N_BTYP1_10=0           THEN ME2N_BTYP1_10 =.;
IF ME2=0        & ME2N_BPWR_10=0            THEN ME2N_BPWR_10 =.;
IF ME2^=0       & ME2N_BTYP1_20=0           THEN ME2N_BTYP1_20 =.;
IF ME2=0        & ME2N_BPWR_20=0            THEN ME2N_BPWR_20 =.;
IF ME2^=0       & ME2N_BTYP1_30=0           THEN ME2N_BTYP1_30=.;
IF ME2=0        & ME2N_BPWR_30=0            THEN ME2N_BPWR_30 =.;
IF ME2^=0       & ME2N_BTYP1_40=0           THEN ME2N_BTYP1_40=.;
IF ME2=0        & ME2N_BPWR_40=0            THEN ME2N_BPWR_40 =.;
IF ME2^=0       & ME2N_BTYP1_50=0           THEN ME2N_BTYP1_50=.;
IF ME2=0        & ME2N_BPWR_50=0            THEN ME2N_BPWR_50=.;


KEEP Condition
DE12_0TYP1   DE12_0PWR        DE12_TTYP1       DE12_TPWR        DE12_BTYP1       DE12_BPWR

DE1_0TYP1    DE1_0TYP1_10     DE1_0TYP1_20     DE1_0TYP1_30     DE1_0TYP1_40     DE1_0TYP1_50
DE1_0PWR     DE1_0PWR_10      DE1_0PWR_20      DE1_0PWR_30      DE1_0PWR_40      DE1_0PWR_50
DE2_0TYP1    DE2_0TYP1_10     DE2_0TYP1_20     DE2_0TYP1_30     DE2_0TYP1_40     DE2_0TYP1_50
DE2_0PWR     DE2_0PWR_10      DE2_0PWR_20      DE2_0PWR_30      DE2_0PWR_40      DE2_0PWR_50

DE1_TTYP1    DE1_TTYP1_10     DE1_TTYP1_20     DE1_TTYP1_30     DE1_TTYP1_40     DE1_TTYP1_50
DE1_TPWR     DE1_TPWR_10      DE1_TPWR_20      DE1_TPWR_30      DE1_TPWR_40      DE1_TPWR_50
DE2_TTYP1    DE2_TTYP1_10     DE2_TTYP1_20     DE2_TTYP1_30     DE2_TTYP1_40     DE2_TTYP1_50
DE2_TPWR     DE2_TPWR_10      DE2_TPWR_20      DE2_TPWR_30      DE2_TPWR_40      DE2_TPWR_50

DE1_BTYP1    DE1_BTYP1_10     DE1_BTYP1_20     DE1_BTYP1_30     DE1_BTYP1_40     DE1_BTYP1_50
DE1_BPWR     DE1_BPWR_10      DE1_BPWR_20      DE1_BPWR_30      DE1_BPWR_40      DE1_BPWR_50
DE2_BTYP1    DE2_BTYP1_10     DE2_BTYP1_20     DE2_BTYP1_30     DE2_BTYP1_40     DE2_BTYP1_50
DE2_BPWR     DE2_BPWR_10      DE2_BPWR_20      DE2_BPWR_30      DE2_BPWR_40      DE2_BPWR_50

ME12S_0TYP1         ME12S_0PWR          ME12N_0TYP1        ME12N_0PWR          ME12S_TTYP1       ME12S_TPWR
ME12N_TTYP1         ME12N_TPWR          ME12S_BTYP1        ME12S_BPWR          ME12N_BTYP1       ME12N_BPWR
ME1S_0TYP1          ME1ST1Z_0TYP1       ME1SGZ_0TYP1       ME1ST1GZ_0TYP1      ME1S_0PWR         ME1S_0TYP1_10
ME1S_0PWR_10        ME1S_0TYP1_20       ME1S_0PWR_20       ME1S_0TYP1_30       ME1S_0PWR_30      ME1S_0TYP1_40
ME1S_0PWR_40        ME1S_0TYP1_50       ME1S_0PWR_50       ME2S_0TYP1          ME2ST2Z_0TYP1     ME2SGZ_0TYP1
```

```
ME2ST2GZ_0TYP1    ME2S_0PWR         ME2S_0TYP1_10     ME2S_0PWR_10      ME2S_0TYP1_20     ME2S_0PWR_20
ME2S_0TYP1_30     ME2S_0PWR_30      ME2S_0TYP1_40     ME2S_0PWR_40      ME2S_0TYP1_50     ME2S_0PWR_50

ME1N_0TYP1        ME1NT1Z_0TYP1     ME1NGZ_0TYP1      ME1NT1GZ_0TYP1    ME1N_0PWR         ME1N_0TYP1_10
ME1N_0PWR_10      ME1N_0TYP1_20     ME1N_0PWR_20      ME1N_0TYP1_30     ME1N_0PWR_30      ME1N_0TYP1_40
ME1N_0PWR_40      ME1N_0TYP1_50     ME1N_0PWR_50      ME2N_0TYP1        ME2NT2Z_0TYP1     ME2NGZ_0TYP1
ME2NT2GZ_0TYP1    ME2N_0PWR         ME2N_0TYP1_10     ME2N_0PWR_10      ME2N_0TYP1_20     ME2N_0PWR_20
ME2N_0TYP1_30     ME2N_0PWR_30      ME2N_0TYP1_40     ME2N_0PWR_40      ME2N_0TYP1_50     ME2N_0PWR_50

ME1S_TTYP1        ME1ST1Z_TTYP1     ME1SGZ_TTYP1      ME1ST1GZ_TTYP1    ME1S_TPWR         ME1S_TTYP1_10
ME1S_TPWR_10      ME1S_TTYP1_20     ME1S_TPWR_20      ME1S_TTYP1_30     ME1S_TPWR_30      ME1S_TTYP1_40
ME1S_TPWR_40      ME1S_TTYP1_50     ME1S_TPWR_50      ME2S_TTYP1        ME2ST2Z_TTYP1     ME2SGZ_TTYP1
ME2ST2GZ_TTYP1    ME2S_TPWR         ME2S_TTYP1_10     ME2S_TPWR_10      ME2S_TTYP1_20     ME2S_TPWR_20
ME2S_TTYP1_30     ME2S_TPWR_30      ME2S_TTYP1_40     ME2S_TPWR_40      ME2S_TTYP1_50     ME2S_TPWR_50

ME1N_TTYP1        ME1NT1Z_TTYP1     ME1NGZ_TTYP1      ME1NT1GZ_TTYP1    ME1N_TPWR         ME1N_TTYP1_10
ME1N_TPWR_10      ME1N_TTYP1_20     ME1N_TPWR_20      ME1N_TTYP1_30     ME1N_TPWR_30      ME1N_TTYP1_40
ME1N_TPWR_40      ME1N_TTYP1_50     ME1N_TPWR_50      ME2N_TTYP1        ME2NT2Z_TTYP1     ME2NGZ_TTYP1
ME2NT2GZ_TTYP1    ME2N_TPWR         ME2N_TTYP1_10     ME2N_TPWR_10      ME2N_TTYP1_20     ME2N_TPWR_20
ME2N_TTYP1_30     ME2N_TPWR_30      ME2N_TTYP1_40     ME2N_TPWR_40      ME2N_TTYP1_50     ME2N_TPWR_50

ME1S_BTYP1        ME1ST1Z_BTYP1     ME1SGZ_BTYP1      ME1ST1GZ_BTYP1    ME1S_BPWR         ME1S_BTYP1_10
ME1S_BPWR_10      ME1S_BTYP1_20     ME1S_BPWR_20      ME1S_BTYP1_30     ME1S_BPWR_30      ME1S_BTYP1_40
ME1S_BPWR_40      ME1S_BTYP1_50     ME1S_BPWR_50      ME2S_BTYP1        ME2ST2Z_BTYP1     ME2SGZ_BTYP1
ME2ST2GZ_BTYP1    ME2S_BPWR         ME2S_BTYP1_10     ME2S_BPWR_10      ME2S_BTYP1_20     ME2S_BPWR_20
ME2S_BTYP1_30     ME2S_BPWR_30      ME2S_BTYP1_40     ME2S_BPWR_40      ME2S_BTYP1_50     ME2S_BPWR_50

ME1N_BTYP1        ME1NT1Z_BTYP1     ME1NGZ_BTYP1      ME1NT1GZ_BTYP1    ME1N_BPWR         ME1N_BTYP1_10
ME1N_BPWR_10      ME1N_BTYP1_20     ME1N_BPWR_20      ME1N_BTYP1_30     ME1N_BPWR_30      ME1N_BTYP1_40
ME1N_BPWR_40      ME1N_BTYP1_50     ME1N_BPWR_50      ME2N_BTYP1        ME2NT2Z_BTYP1     ME2NGZ_BTYP1
ME2NT2GZ_BTYP1    ME2N_BPWR         ME2N_BTYP1_10     ME2N_BPWR_10      ME2N_BTYP1_20     ME2N_BPWR_20
ME2N_BTYP1_30     ME2N_BPWR_30      ME2N_BTYP1_40     ME2N_BPWR_40      ME2N_BTYP1_50     ME2N_BPWR_50;
RUN;   /* has 270 +1 variables */

/* One record of all the Type 1 and Power for ME12S_0BT ME12N_0BT ME1S_0BT ME2S_0BT ME1N_0BT ME2N_0BT for 0.10, 0.20, 0.30,
0.40 and 0.50 cut-offs and DE12_0BT DE1_0BT DE2_0BT a given condition*/
/***Coverage probability only for Joint effects ME12S_0BT ME12N_0BT and DE12_0BT***/

DATA ME12S_0COV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
```

```
IF Sampnum=1 THEN CountCV=0;
IF (PS0_ME12_2_5<= ME12 <=PS0_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12S_0COV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME12S_0COV;
RUN;

DATA ME12N_0COV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PN0_ME12_2_5<= ME12 <=PN0_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12N_0COV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME12N_0COV;
RUN;

DATA ME12S_TCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PST_ME12_2_5<= ME12 <=PST_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12S_TCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME12S_TCOV;
RUN;

DATA ME12N_TCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PNT_ME12_2_5<= ME12 <=PNT_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12N_TCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME12N_TCOV;
RUN;

DATA ME12S_BCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PSB_ME12_2_5<= ME12 <=PSB_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12S_BCOV=CountCV/500;
RETAIN CountCV;
```

```
KEEP Condition Sampnum ME12S_BCOV;
RUN;

DATA ME12N_BCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (PNB_ME12_2_5<= ME12 <=PNB_ME12_97_5) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN ME12N_BCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum ME12N_BCOV;
RUN;

/**********SEQ 2V   DE12 ***************/
DATA DE12_0COV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE12_0LCL<= DE12 <=DE12_0UCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE12_0COV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE12_0COV;
RUN;

DATA DE12_TCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE12_TLCL<= DE12 <=DE12_TUCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE12_TCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE12_TCOV;
RUN;

DATA DE12_BCOV (WHERE= (Sampnum=500));
SET WQSEST2v_500;
IF Sampnum=1 THEN CountCV=0;
IF (DE12_BLCL<= DE12 <=DE12_BUCL) THEN CountCV= (CountCV+1);
IF Sampnum=500 THEN DE12_BCOV=CountCV/500;
RETAIN CountCV;
KEEP Condition Sampnum DE12_BCOV;
RUN;
```

```
DATA WQS2V_COV;
MERGE            ME12S_0COV      ME12N_0COV      ME12S_TCOV      ME12N_TCOV      ME12S_BCOV
                 ME12N_BCOV      DE12_0COV       DE12_TCOV       DE12_BCOV;
BY Condition;

KEEP    Condition          ME12S_0COV      ME12N_0COV      ME12S_TCOV      ME12N_TCOV
        ME12S_BCOV         ME12N_BCOV      DE12_0COV       DE12_TCOV       DE12_BCOV        ;

RUN; /* 27 Vars Single record of Coverage Probabilities for ME12 and DE12 for given Condition*/
DATA WQSEST2v_1;
MERGE   WQSEST2v_1 WQS2V_TYP1PWR WQS2V_COV;
  BY Condition;
RUN;   /* has 429 variables = 132 +27+270=428 variables */

PROC APPEND base=WQSEST2V DATA=WQSEST2v_1 force; RUN; /* APPEND DATA to the base file */

DM LOG 'CLEAR';
%END;
%mEND STATS_2V;
%STATS_2V;

DATA Temp.WQS2V_STATS_108;   /*KEEP the 500 records one for each Sampnum in Database for later use */
SET   WQSEST2V;   /* Has 108 records with 429 columns of STATS on 2-variable mediation analysis */
RUN;
```

LASSO CODE IN R

```
 setwd ("C:/Evani/Lasso")
 getwd ()
 require (graphics)
 require (glmnet)
 infile <- array (1:162, dim= c (162))
 outfile <- array (1:162, dim= c (162))
   for (ii in seq(2,162,2)) {
             infile [ii]<- paste("QrV3_",ii,".csv",sep="")
            outfile [ii]<- paste("L3_",ii,".csv",sep='')
            yx123m<- read.csv (infile [ii], header=TRUE)
   attach (yx123m[3:7])
```

```
for (i in 1:500) {
        j<- (i-1)*1000+1   #or alternatively odd DATASETs have 300 sample size
        k<- (i*1000)         #or alternatively odd DATASETs have 300 sample size
        x <- as.matrix (cbind(X1 [j: k], X2 [j: k], X3 [j: k])) # matrix of predictors
        y <- as.matrix (Y [j: k])   #column response vector of Y
        m <- as.matrix (M [j: k])   #column response vector of M
        x_m<- as.matrix (cbind(X1[j: k], X2[j: k], X3[j: k], M [j: k])) # matrix of predictors
  #Prepare GLMNET Coefficients

  glmnet_T1<-glmnet(x, m, penalty.factor= c (1, 1, 1), alpha = 1)  # lambda multiplier for differential shrinkage
  cv_T1<-cv.glmnet(x, m, nfolds=1000, grouped=FALSE, alpha = 1)

  model_T1se=glmnet(x, m, lambda=cv_T1$lambda.1se, alpha = 1)
  model_T1min=glmnet(x, m, lambda=cv_T1$lambda.min, alpha = 1)

  coef_X1<-as.matrix (predict (model_T1se, type="coefficients")) ["V1",]
  coef_X2<-as.matrix (predict (model_T1se, type="coefficients")) ["V2",]
  coef_X3<-as.matrix (predict (model_T1se, type="coefficients")) ["V3",]
  Result_T1se<-cbind (coef_X1, coef_X2, coef_X3)

  coef_X1<-as.matrix (predict (model_T1min, type="coefficients")) ["V1",]
  coef_X2<-as.matrix (predict (model_T1min, type="coefficients")) ["V2",]
  coef_X3<-as.matrix (predict (model_T1min, type="coefficients")) ["V3",]
  Result_T1min<-cbind (coef_X1, coef_X2, coef_X3)

  Results_T<-cbind (Result_T1se, Result_T1min)
  colnames(Results_T)<-c("L1se_T1", "L1se_T2", "L1se_T3","Lmin_T1","Lmin_T2","Lmin_T3")
  ### extract the coefficient vector with L1 norm=1.0

  #Prepare GLMNET
  glmnet_B1<- glmnet (x_m, y, penalty.factor = c (1,1,1,0),alpha = 1)  # lambda multiplier for differential shrinkage
  # obtain GLMNET Coefficients and Plot the admittance functions also print out the GLMNET Solution

  cv_B1<-cv.glmnet (x_m, y, nfolds=1000, grouped=FALSE, alpha = 1)
  model_B1se=glmnet (x_m, y, lambda=cv_B1$lambda.1se, alpha = 1)
  model_B1min=glmnet (x_m, y, lambda=cv_B1$lambda.min, alpha = 1)

  coef_M <-as.matrix (predict (model_B1se, type="coefficients")) ["V4",]
  coef_X1<-as.matrix (predict (model_B1se, type="coefficients")) ["V1",]
  coef_X2<-as.matrix (predict (model_B1se, type="coefficients")) ["V2",]
```

```
    coef_X3<-as.matrix (predict (model_B1se, type="coefficients")) ["V3",]
    Result_B1se<-cbind (coef_M, coef_X1, coef_X2, coef_X3)

    coef_M<-as.matrix (predict (model_B1min, type="coefficients")) ["V4",]
    coef_X1<-as.matrix (predict (model_B1min, type="coefficients")) ["V1",]
    coef_X2<-as.matrix (predict (model_B1min, type="coefficients")) ["V2",]
    coef_X3<-as.matrix (predict (model_B1min, type="coefficients")) ["V3",]
    Result_B1min<-cbind (coef_M, coef_X1, coef_X2, coef_X3)

    Results_B<-cbind (Result_B1se, Result_B1min)
    colnames(Results_B) <- c("L1se_M", "L1se_B1", "L1se_B2", "L1se_B3","Lmin_M", "Lmin_B1", "Lmin_B2", "Lmin_B3")
    Results_B
    Results<- cbind (ii, i, Results_B, Results_T)
    colnames(Results)<- ("Condition","Sampnum","L1se_M","L1se_B1","L1se_B2","L1se_B3","Lmin_M","Lmin_B1","Lmin_B2","Lmin_B3",
                                  "L1se_T1","L1se_T2","L1se_T3","Lmin_T1","Lmin_T2","Lmin_T3")

    write.table (Results, outfile [ii], sep ="\n", col.names=FALSE, APPEND=TRUE)

  } # for loop sampnum Five hundred times
       Detach (yx123m [3:7])

 } # Loop 81 times
```

PLOTS

```
%PUT %sysfunc (getoption (work));
ODS HTML body= 'ANOVA.html' style=HTMLBlue options (pagebreak='no');
ODS HTML close; ODS listing; ODS html; ODS graphics on; ODS listing style= OLS_WQS_LASSO;

PROC format;
VALUE legfmt12ME /*cME12 */
0 = '0       -Zero theta*gamma parm'
1 = '0.015-Low theta*gamma parm'
2 = '0.045-Med. theta*gamma parm'
3 = '0.135-Lrg. theta*gamma parm';

VALUE legfmt1ME /*cME1 */
0 = '0       -zero theta1*gamma param'
1 = '0.015-low theta1*gamma param'
2 = '0.045-high theta1*gamma param';

VALUE legfmt2ME /*cME2 */
0 = '0       -zero theta2*gamma param'
1 = '0.015-low theta2*gamma param'
2 = '0.03   -medium theta2*gamma param'
3 = '0.045 -high theta2*gamma param'
4 = '0.09 -large theta2*gamma param';

VALUE legfmt123ME /*cME123 */
0 = '0       -Zero theta*gamma param'
1 = '0.135-Low theta*gamma param'
2 = '0.15-Medium theta*gamma param'
3 = '0.1925-High theta*gamma param'
4 = '0.21-Large theta*gamma param';

VALUE legfmtN   /*cN */
1 = "300 Sample size"
2 = "1000 Sample size";

VALUE legfmtRho_12X   /* cR12 */
1 = "Zero Correlation"
2 = "0.5 Correlation"
3 = "0.95 Correlation";
```

```
VALUE legfmtBeta   /* cB12 */
1 = "0.30   Low Beta parameter"
2 = "0.45   High Beta parameter";

VALUE legfmt3cTheta   /* cT123 */
1 = "T123=0.55   Theta1=0"
2 = "T123=0.55   Theta2=0"
3 = "T123=0.60   Theta3=0";

VALUE legfmtTheta    /* cT12*/
1 = "0   Null"
2 = "0.15   Low Theta parameter"
3 = "0.45   High Theta parameter";

VALUE legfmtGamma   /* cG */
1 = "0   Null"
2 = "0.10   Low Gamma parameter"
3 = "0.30   High Gamma parameter";

VALUE legfmt3Gamma    /* cG3 */
0 = "0   Null"
1 = "0.25   Mid Gamma parameter"
2 = "0.35   High Gamma parameter";

VALUE legfmtRho_123X    /* cR123 */
3 = "Low   Corr. 0.45, 0.20, 0.13"
2 = "Med. Corr. 0.68, 0.30, 0.19"
1 = "Hi.    Corr. 0.90, 0.40, 0.25";

VALUE legfmt12DE   /*cDE12 */
1= '0.30 - med. Beta params.'
2= '0.45 - high Beta params.'  ;

VALUE legfmt1DE    /*cDE1 */
1= '0.15 - low Beta1 parameter'
2= '0.30 - high Beta1 parameter';

VALUE legfmt2DE   /*cDE2 */
0= '0       -zero Beta2 parameter'
1= '0.30 - high Beta2 parameter';
```

```
VALUE legfmt123B   /*cDE123 */
1= 'DE123=0.50    Beta2=0'
2= 'DE123=0.45    Beta1=0'
3= 'DE123=0.35    Beta3=0';

VALUE legfmt0TB   /*c0TB*/
1 = "(OLS_X1X2, WQS Y|X1X2)"
2= "(OLS_X1X2, WQS M|X1, X2)"
3 = "(OLS_X1X2, WQS Y|X1X2, M)";

VALUE legfmtmin1se   /*cmin1se*/
1 = "(OLS_X1X2, LASSO_min)"
2= "(OLS_X1X2, LASSO_1se)";

VALUE legfmtminvs1se   /*cmin1se*/
1 = "LASSO_min vs. LASSO_min+1se";
RUN;

/**** ME123 Estimate, Bias and RMSE ****/
DATA Wqs3v_stats_162;
MERGE   Temp.Wqs3v_stats_162 Temp.Cond162;
BY Condition;
RUN;

PROC SORT DATA=Wqs2v_stats_108 out=WQS_ME12;
 BY N Rho_X12 T12 Gamma ME12 DE12;
 RUN;
PROC SORT DATA=Wqs3v_stats_162 out=WQS_ME123;
 BY N cR123 T123 cG3 Gamma ME123;
 RUN;

PROC sql noprint;
 CREATE table minmax as
 SELECT min (ME123_T) as minv, /*, min (ME123_TBias), min (ME123_TRMSE)) as minv,*/
 MAX (ME123_T) as maxv /*, max (ME123_TBias), max (ME123_TRMSE)) as maxv */
 FROM WQS_ME123;
 QUIT;   /*-0. 075 to 0.225 */
DATA WQS_ESTME123 (keep=N cR123 T123 cT123 cG3 Gamma ME123 cME123 cDE123 ME123_T ME123_TBias ME123_TRMSE
ME123_0 ME123_0Bias ME123_0RMSE ME123_B ME123_BBias ME123_BRMSE my_x_lineval my_y_lineval);
```

```
RETAIN minv maxv;
 IF _n_ = 1 then SET minmax;
 SET WQS_ME123;
 BY N cR123 T123 Gamma ME123;

 IF first.N then do;
 my_x_lineval = minv;
 my_y_lineval= minv;
 END;
 IF last.N then do;
 my_x_lineval = maxv;
 my_y_lineval = maxv;
 END;

 IF first.cR123 then do;
 my_x_lineval = minv;
 my_y_lineval= minv;
 END;
 IF last.cR123 then do;
 my_x_lineval = maxv;
 my_y_lineval = maxv;
 END;

 IF first.T123 then do;
 my_x_lineval = minv;
 my_y_lineval= minv;
 END;
 IF last.T123 then do;
 my_x_lineval = maxv;
 my_y_lineval = maxv;
 END;

 IF first.Gamma then do;
 my_x_lineval = minv;
 my_y_lineval= minv;
 END;
 IF last.Gamma then do;
 my_x_lineval = maxv;
 my_y_lineval = maxv;
 END;
```

```
 IF first.ME123 then do;
 my_x_lineval = minv;
 my_y_lineval= minv;
 END;
 IF last.ME123 then do;
 my_x_lineval = maxv;
 my_y_lineval = maxv;
 END;
RUN;

ODS HTML close;
ODS GRAPHICS on / width= 3.25in height =3.25in; ODS listing gpath="H:\Thesis\RESULTS";
PROC SGPLOT DATA= WQS_ESTME123 unIForm=group;
STYLEATTRS DATASYMBOLS= (trianglefilled star plus asterisk)
            DATACONTRASTCOLORS= (green red magenta black lightgreen red);
 XAXIS values= (0.00 TO 0.3 BY 0.05);
 XAXIS LABEL="Gamma parm. (0, 0.25, 0.35)";
 YAXIS values= (0.00 TO 0.25 BY 0.01);
 YAXIS LABEL="ME123 Estimate RMSE using wgts. M|X";
 FOOTNOTE1 "162 obs. ME123 (0, 0.138 & 0.15, 0.193 & 0.21);          Colors- sample size";

*REFLINE 0.05 /AXIS=Y LABEL="standard" LABELLOC=INSIDE LINEATTRS= (thickness=1 color=grey pattern=SOLID);
REFLINE 0.08 /AXIS=Y LABELloc=OUTSIDE LABELPOS=MIN  LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
REFLINE 0 0.60 /AXIS=X LABELLOC=INSIDE LINEATTRS= (thickness=0 color=white pattern=solid) NOCLIP;
REFLINE 0 /AXIS=Y LABELLOC=INSIDE LINEATTRS= (thickness=1 color=black pattern=solid) NOCLIP;
*REFLINE 0.075 / AXIS=Y LABEL="robust" LABELLOC=INSIDE LINEATTRS= (thickness=2 color=black pattern=DASH);

*WHERE DE12=0;
 SCATTER x= Gamma y= ME123_TRMSE/group=N MARKERATTRS=(size=7 symbol= trianglefilled) name ="scp";
*SERIES x= my_x_lineval y= my_y_lineval / LINEATTRS= (color=black);

LABEL cT123="legfmt3cTheta" N="LegfmtN" cR123='legfmtRho_123X' cG3='legfmt3Gamma';
FORMAT cT123 legfmt3cTheta. N legfmtN. cR123 legfmtRho_123X. cG3 legfmt3Gamma. ;
 KEYLEGEND "scp" / border across= 1 down= 4 location= inside position= bottomright;
RUN;

 /******Comparative plots *****/
 ODS html close;
 ODS graphics on / width= 3.25in height =3.25in; ODS listing gpath="H:\Thesis\RESULTS";
PROC sgplot DATA= WQS_ESTME123 unIForm=group;
```

```
STYLEATTRS DATAsymbols= (circlefilled diamond triangle square star plus asterisk)
           DATAcontrastcolors= (red lightblue black);

 XAXIS values= (0.00 TO 0.3 BY 0.05);
 XAXIS LABEL="ME123 Estimate using wgts. Y|X, M";
 YAXIS values= (0.00 TO 0.25 BY 0.01);
 YAXIS LABEL="ME123 Estimate using wgts. Y|X";
 FOOTNOTE1 "162 obs. ME123 (0, 0.138 & 0.15, 0.193 & 0.21); Colors- Corr. (Hi. Med. & Low) for (X12, X13, X23)";
 *WHERE DE12=0;
 SCATTER x= ME123_B y= ME123_0/group=cR123 MARKERATTRS= (size=7 symbol= circlefilled) name ="scp";
 SERIES x= my_x_lineval y= my_y_lineval / LINEATTRS= (color=black);
*REFLINE 0.05 /AXIS=Y LABEL="standard" LABELLOC=INSIDE LINEATTRS= (thickness=1 color=grey pattern=SOLID);
REFLINE 0.25 0 /AXIS=Y LABELLOC=OUTSIDE LABELPOS=MIN LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
REFLINE 0 0.25 /AXIS=X LABELLOC=INSIDE LINEATTRS= (thickness=0 color=white pattern=solid) NOCLIP;
*REFLINE 0 /AXIS=Y LABELLOC=INSIDE LINEATTRS= (thickness=1 color=lightgrey pattern=solid) NOCLIP;
*REFLINE 0.075 / AXIS=Y LABEL="robust" LABELLOC=INSIDE LINEATTRS= (thickness=2 color=black pattern=DASH);
 LABEL cT123="legfmt3cTheta" N="LegfmtN" cR123='legfmtRho_123X' cG3='legfmt3Gamma';
 FORMAT cT123 legfmt3cTheta. N legfmtN. cR123 legfmtRho_123X. cG3 legfmt3Gamma. ;
 KEYLEGEND "scp" / border across= 1 down= 4 location= inside position= bottomright;
 RUN;

/*********************ME123 TYP1 ******/
PROC SORT DATA=Wqs3v_stats_162 out=WQS_ME123;
BY N cR123 T123 Gamma ME123;
RUN;

PROC sql noprint;
CREATE table minmax as
SELECT min (ME123_TTYP1) as minv,
MAX (ME123_TTYP1) as maxv
FROM WQS_ME123;
QUIT;   /* 0 to 0.136 (Preferred) */

DATA Minmax;
SET Minmax;
Minv = 0; maxv = 0.15;
RUN;

DATA WQS_ESTME123 (keep=N cR123 T123 cT123 cG3 R12 R13 R23 Gamma ME123 cME123 cDE123 ME123_T ME123_TTYP1
ME123_0 ME123_0TYP1 ME123_B ME123_BTYP1 my_x_lineval my_y_lineval);
```

```
RETAIN minv maxv;
 IF _n_ = 1 then SET minmax;
 SET WQS_ME123;
 BY N cR123 T123 Gamma ME123;

ODS html close;
ODS graphics on / width= 3.25in height =3.25in; ODS listing gpath="H:\Thesis\RESULTS";
PROC sgplot DATA= WQS_ESTME123 unIForm=group;
STYLEATTRS DATASYMBOLS= (circlefilled diamond triangle square star plus asterisk)
        DATACONTRASTCOLORS= (black red green maroon olive purple black);
 XAXIS values= (0.00 TO 0.3 BY 0.05);
 XAXIS LABEL="Joint Theta parm. (0.55, 0.60)";
 YAXIS values= (0.00 TO 0.15 BY 0.01);
 YAXIS LABEL="Type1 err. ME123 for WQS_M|X";
 footnote1 "54 obs. where ME123=0;
                                        Colors- Corr. (Hi., Med, Low) for (X12,X13,X23)";
REFLINE 0.05 /AXIS=Y LABEL="standard" LABELLOC=INSIDE LABELPOS=MIN      LINEATTRS= (thickness=1 color=grey
pattern=SOLID);
REFLINE 0.075 / AXIS=Y LABEL="robust" LABELLOC=INSIDE LABELPOS=MIN      LINEATTRS= (thickness=2 color=black
pattern=DASH);
REFLINE 0 0.15/AXIS=Y LABELLOC=INSIDE LINEATTRS= (thickness=1 color=black pattern=SOLID) NOCLIP;
REFLINE 1.0 /AXIS=X LABELLOC=INSIDE LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
REFLINE 0 /AXIS=X LABELLOC=INSIDE LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
*REFLINE 0.075 /AXIS=X LABEL="robust" LABELLOC=INSIDE  LINEATTRS= (thickness=2 color=black pattern=DASH);

WHERE ME123=0;
 SCATTER x= T123 y= ME123_TTYP1/group=cR123   MARKERATTRS=(size=7 symbol= circlefilled) name ="scp";
*SERIES x= my_x_lineval y= my_y_lineval / LINEATTRS=(color=black);

 LABEL cT123="legfmt3cTheta" N="LegfmtN" cR123='legfmtRho_123X' cG3='legfmt3Gamma';
 FORMAT cT123 legfmt3cTheta. N legfmtN. cR123 legfmtRho_123X. cG3 legfmt3Gamma. ;
 KEYLEGEND "scp" / border across= 1 down= 4 location= inside position= bottomright;
 RUN;

 /********************ME123 Power *********************/
PROC SORT DATA=WQS_ME123;
BY N cR123 cT123 cG3 ME123;
RUN;
 PROC sql noprint;
 CREATE table minmax as
```

```
 SELECT min (min (ME123_TPWR), min (ME123_0PWR)) as minv,
 MAX (max (ME123_TPWR), max (ME123_0PWR)) as maxv
FROM WQS_ME123;
QUIT; /* min 0.95 max 1 */

 DATA Minmax;
 SET Minmax;
 Minv = 0.95; maxv=1;
 RUN;

 DATA WQS_ME123PWR (keep=N cR123 R12 ME1 ME2 ME3 ME123 DE123 T123 cT123 Gamma cG3 ME123_TPWR ME123_0PWR
ME123_BPWR ME1_TPWR ME1_TPWR_10 ME1_TPWR_20 ME1_TPWR_30 ME1_TPWR_40
 ME2_TPWR ME2_TPWR_10 ME2_TPWR_20 ME2_TPWR_30 ME2_TPWR_40
 ME3_TPWR ME3_TPWR_10 ME3_TPWR_20 ME3_TPWR_30 ME3_TPWR_40 my_x_lineval my_y_lineval);
RETAIN minv maxv;
IF _n_ = 1 then SET minmax;
SET WQS_ME123;
BY N cR123 cT123 cG3 ME123;


ODS graphics / width= 3.25in height= 3.25in; ODS listing gpath="H:\Thesis\RESULTS";
PROC sgplot DATA=WQS_ME123PWR unIForm=group;
STYLEATTRS DATAsymbols= (circlefilled) DATAcontrastcolors= (red blue green black orange blue);
 XAXIS VALUES= (0 to 1 BY 0.2);
 XAXIS LABEL="ME123= (0.05 0.07, 0.075, 0.105)          WQS weights cut-off=0.4";

 YAXIS VALUES= (0.95 TO 1 BY 0.01);
 YAXIS LABEL="ME3_PWR using Index M|X";

*REFLINE 0 1.3 /axis=y LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
*REFLINE 0 /axis=x LINEATTRS= (thickness=0 color=white pattern=SOLID) NOCLIP;
*REFLINE 0 0.50 /axis=x LABELLOC=outside LABELPOS=MIN LINEATTRS= (thickness=2 color=white pattern=SOLID) NOCLIP;
REFLINE 1.0/axis=y LABELLOC=outside LABELPOS=MIN LINEATTRS= (thickness=1 color=lightgrey pattern=SOLID) NOCLIP;
REFLINE 0 0.15/axis=x LABELLOC=outside LABELPOS=MIN LINEATTRS= (thickness=1 color=lightgrey pattern=SOLID) NOCLIP;
REFLINE 0 /axis=x LABELLOC=outside LABELPOS=MIN LINEATTRS= (thickness=1 color=black pattern=SOLID) NOCLIP;
REFLINE 0.80 /axis=y LABELLOC=outside LABELPOS=MIN LINEATTRS= (thickness=1 color=lightgrey pattern=SOLID) NOCLIP;

 WHERE ME3 ^=0;
 SCATTER x=ME3 y= ME3_TPWR_40 /group=N MARKERATTRS= (size=7 symbol= circlefilled) name ="scp";
*SERIES x= my_x_lineval y= my_y_lineval/ LINEATTRS= (thickness=1 color=black);
```

```
FOOTNOTE1 "Total 72 conditions with ME3 ^=0";
LABEL cT123="legfmt3cTheta" N="LegfmtN" cR123='legfmtRho_123X' cG3='legfmt3Gamma';
FORMAT cT123 legfmt3cTheta. N legfmtN. cR123 legfmtRho_123X. cG3 legfmt3Gamma. ;
KEYLEGEND "scp" / border across= 1 down= 4 location= inside position= bottomright;
RUN;
```

*****