Published in final edited form as:

Health Serv Outcomes Res Methodol. 2012 June ; 12(2-3): 219–233. doi:10.1007/s10742-012-0088-8.

Using AIC in Multiple Linear Regression framework with Multiply Imputed Data

Ashok Chaurasia and Ofer Harel

Department of Statistics, University of Connecticut, Storrs, CT, USA

Abstract

Many model selection criteria proposed over the years have become common procedures in applied research. However, these procedures were designed for complete data. Complete data is rare in applied statistics, in particular in medical, public health and health policy settings. Incomplete data, another common problem in applied statistics, introduces its own set of complications in light of which the task of model selection can get quite complicated. Recently, few have suggested model selection procedures for incomplete data with varying degrees of success. In this paper we explore model selection by the Akaike Information Criterion (AIC) in the multivariate regression setting with ignorable missing data accounted for via multiple imputation.

Keywords

Model Selection; AIC; Missing Data; Multiple imputation; Incomplete Data

1. Introduction

Regression analysis is one of the most widely used of all statistical procedures and a common task in regression analysis is that of variable selection; the search for subset(s) of variables that "best" explain the response, where "best" is defined with respect to a specific purpose such as model interpretation or prediction. Selecting variables that "best" explains the data is called variable selection. Variable selection, though well studied in the statistical literature, remains rich in unsolved problems (Rao and Wu, 2001; Miller, 2002) especially as related to missing response data. Complete data is quite rare in practice and the cause of missing data can be attributed to factors such as human error, poorly designed surveys, intrusive/personal questions, respondent refusal, subject attrition, etc. In dealing with missing data some practitioners resort to complete case analysis (CCA) even though it is well documented that CCA is an inefficient way to handle missing data and leads to biased results (Schafer, 1997, pp.24–26).

Alternatives to CCA include, but are not limited to, single imputation, weighting schemes (Rubin, 1987), Maximum Likelihood (ML) (Little and Rubin, 2002), Multiple Imputation (MI) (Rubin, 1987). Each of these methods, with their advantages and disadvantages, rely on different set of missing data assumptions. The missing data assumptions are mostly based on work done by Rubin (1976, 1987) who classified missing data as missing completely at random (MCAR), missing at random (MAR), missing not at random (MNAR), ignorable and non-ignorable. Further discussion of these terms is given in section 3.2. In this paper, we focus exclusively on the MI method for dealing with missing data. In multiple imputation (MI) missing values are replaced with random draws from the posterior predictive distribution to create sets of imputed data. The posterior predictive distribution relies on the assumptions about mechanism of missingness and the imputation model. In practice it is

common to assume ignorable models (and hence MAR) for the missing data and a multivariate normal imputation model. Schafer (1997) has shown that MI is quite forgiving to departures from the imputation model when one is interested in analyses that depend on means and variances as in regression and principal-component methods. There is no clear evidence to describe the impact of departure from an imputation model (for both response and covariates) on variable selection in a linear regression. In regards to model selection with multiply imputed data sets, implementing traditional variable selection methods (such as AIC) separately for each imputation is straightforward but the question remains as to how best to combine the separate model selection results? Recent work in this area includes Yang et al. (2005), Wood et al. (2008) and, Claeskens and Consentino (2008), Consentino and Claeskens (2010).

Yang et al. (2005) propose two methods — *impute then select* (ITS), *simultaneously impute and select* (SIAS), in a Bayesian setting to compute the average posterior probabilities over the imputed data for each candidate model. Their simulations suggest that (i) SIAS slightly outperforms ITS and provides smaller standard errors, (ii) SIAS has higher signal-to-noise ratio than ITS, and (iii) SIAS had a lower number of incorrect variables selected than ITS. Wood et al. (2008) propose averaging imputed data sets (under three weighting schemes) and then performing model selection. Their investigation suggests that their proposed method (i) has higher power, but at a cost of an inflated type 1 error, and (ii) is likely to be adversely affected by incomplete response variables (Y). The authors recommend that the weighting schemes should be applied after excluding observations with missing Y . In Claeskens and Consentino (2008), the authors derive a new model selection criteria for incomplete data, by the use of the EM algorithm and the weighting method of Ibrahim (1990). In Consentino and Claeskens (2010), the authors use the combined likelihood ratio tests for imputed data sets (Meng and Rubin, 1992) in constructing an a version of AIC for multiply imputed data. Their investigation in both papers suggests that their model selection method leads to models that tend to overfit (i.e. models that contain the correct model plus some additional variables.) This is not surprising because it is well known that AIC tends to select over-specified models (Hurvich and Tsai, 1990).

The simulation study in each of these articles presents the present results of their proposed methods when missingness occurs in covariates only. Hence, in this paper we explore the issue of model selection in data with incomplete response variable in the multiple linear regression framework. The objective of this paper is to (i) highlight the complexity of this issue in the multiple regression framework, (ii) elucidate the consequence of using computationally simple, but naïve, methods for model selection in incomplete data, and (iii) the need for generalization of complete data model selection procedures in light of incomplete responses and covariate data. Section 2 describes the regression model of interest for our simulation study. Section 3 gives a brief discussion of variable selection, missing data assumptions, and MI. Section 4 provides the simulation setup and results.

2. Model and Notation

Consider the usual multiple regression model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1)$$

where $\mathbf{Y}$ is a n -dimensional vector, $\mathbf{X}$ is a $n \times p$ matrix, $\boldsymbol{\beta}$ is a p -dimensional vector, and $\boldsymbol{\varepsilon}$ is the n -dimensional (uncorrelated) error term which is normally distributed with zero mean and constant variance (σ^2). It is assumed that n is sufficiently larger than p to ensure estimability of the parameters. Let $\mathcal{D}$ represents the data in rectangular form, which is partitioned as $(\mathcal{D}_{obs}, \mathcal{D}_{mis})$, where $\mathcal{D}_{obs}$ is the observed part of $\mathcal{D}$ and $\mathcal{D}_{mis}$ is the missing part

of $\mathcal{D}$. Consider $\mathbf{R}$, a matrix of the same dimension as $\mathcal{D}$ indicating whether cells in $\mathcal{D}$ are observed ($R_{ij} = 1$) or missing ($R_{ij} = 0$). Furthermore, let ϕ be the missingness parameter and $\theta = (\beta, \sigma^2)$ be the data parameter for the model given by equation (1).

3. Methods

3.1 Variable or Model Selection in complete data using the AIC

A variable selection method intends to select the "best" subset of predictor variables for a specific purpose such as interpretation or prediction. The term "best" usually refers to a balance between the number of explanatory variables and goodness of fit. In the search for the "best" model(s), it is possible for the true data generating model to not be in the collection of models being considered in the analysis phase. Also, there is no guarantee that the subset of predictors deemed as "best" will be unique. Earlier variable selection criteria centered around (adjusted) residual sums of squares, stepwise methods for selecting significant variables, or information theoretic approaches (Rao and Wu, 2001; Miller, 2002).

Akaike (1974) introduced the Akaike Information Criterion, an information theoretic approach for model/variable selection, via Kullback-Leibler divergence. AIC is one of the most common model selection procedures that is available in most statistical software packages. Under the setting of model (1), when $\mathcal{D}$ is fully observed, AIC is given as follows:

$$AIC_p = n \ln(\hat{\sigma}_{mle}^2) + 2p \quad (2)$$

where $\hat{\sigma}_{mle}^2$ is the *maximum likelihood estimate* of σ^2 . Under this selection criterion the model with the smallest AIC value is deemed as "best". The popularity of AIC is due to the fact that it adheres to the concept of parsimony and is easy to implement in models such as (1). Like any model selection criterion, AIC is not foolproof. Many researchers have shown that for small samples, AIC is inconsistent and leads to overfitting (see, e.g., Hurvich and Tsai, 1989). However, AIC has several advantages – it is asymptotically efficient, allows for simultaneous comparison of multiple nested or non-nested models (see Chamberlain, 1890), and allows for model-averaged inference. For further details refer to Hurvich and Tsai (1990), Rao and Wu (2001), Miller (2002), Burnham and Anderson (2004), Claeskens and Consentino (2008).

3.2 Missing Data and Assumptions

The mechanism of missingness relates the probability of missing values to the observed and other missing values in a data set. When the likelihood of a missing response is dependent on the observed data ($\mathcal{D}_{obs}$) and not on the missing part of the data ($\mathcal{D}_{mis}$) then the missing mechanism is MAR; i.e. $Pr(\mathbf{R}|\mathcal{D}, \phi) = Pr(\mathbf{R}|\mathcal{D}_{obs}, \phi)$. When the likelihood of a missing response depends on neither observed nor missing values in the data then the missing mechanism is MCAR; i.e. $Pr(\mathbf{R}|\mathcal{D}, \phi) = Pr(\mathbf{R}|\phi)$. In contrast, when the likelihood of an observed response depends on both observed and unobserved data then the missingness mechanism is NMAR. The concept of ignorability addresses the important question of whether or not to model the missing data process. Missing data mechanism is termed as ignorable if data is MAR, and the parameter of interest (θ) and the parameter describing the missing data (ϕ) are distinct (i.e. $P(\theta, \phi) = P(\theta)P(\phi)$). A missing data mechanism is nonignorable when at least one of the conditions of ignorability is not met. The mechanism of missingness allows for the weakest set of conditions (ignorable) under which the distribution of $\mathbf{R}$ does not need to be taken into account in the inference about θ (Rubin, 1976). In general, the true missing-data mechanism is unknown and assumptions about the missing-data mechanism for the imputation model are untenable. The general rule of thumb for the imputation model is to include all available variables in $\mathcal{D}$ Collins et al. (2001) and

Schafer and Graham (2002) have shown that inclusion of auxiliary variables make this ignorability assumption more plausible.

3.3 Multiple Imputation

Multiple imputation is a technique in which each missing value is replaced by ($m > 1$) simulated values. The simulated values are realizations from the posterior predictive distribution of the missing values given the observed values. These imputed data maintain the relationships observed in the original data. Rubin (1987) introduced the framework of MI and developed its rules and assumptions. Multiple imputation consists of three stages – (i) Imputation, (ii) Analysis, and (iii) Combining. The imputation phase relies on assumptions about the missing-data mechanism. Assuming a classification for the missing-data mechanism results in an imputation model which yields m imputed data sets – $\mathcal{D}_1, \dots, \mathcal{D}_m$. If $\mathbf{Q}$ represents the parameter of interest, then the analysis phase of MI consists of obtaining point and variance estimates for $\mathbf{Q}$ from each of the m imputed data sets. In the combining stage, if $\hat{\mathbf{Q}}$ is (approximately) normally distributed, then m point and variance estimates are combined using Rubin's rules (Rubin, 1987) to make inferences about the unknown quantity $\mathbf{Q}$. When $\hat{\mathbf{Q}}$ is non-normal, a transformation of $\hat{\mathbf{Q}}$ to approximate normality is considered (Rubin, 1976) and the transformed estimates are then combined using Rubin's rules (see, e.g., Harel, 2009).

4. Simulations

4.1 Simulation Setup

We designed a simulation study that considers 64 distinct data configurations. The various data configurations result from varying data parameters – number of observations ($n = 50, 100$), number of predictor variables ($k = 3, 4$), variance (σ^2 ; $\sigma = 0.3, 0.5, 0.75, 1, 2, 4, 6, 8$) of the error term in (1), percentages of missingness ($\gamma = 25\%, 50\%$). With respect to the choices for sample size (n) relative to p , since there is no golden rule for determining if n is sufficiently large, we consider at least 10 cases per predictor as n being sufficiently larger than p . For γ , we chose 25% (one-fourth of the responses are missing) and 50% (one-half of the responses are missing) to reflect a practical situation in favor of appropriately handling missing cases rather than ignoring them in analysis.

Complete data generation—For each combination of (n, k)

- i. Generate $\mathbf{X} \sim \mathbf{N}_k(\mathbf{0}, \mathbf{\Sigma}_\mathbf{X})$ with $(\mathbf{\Sigma}_\mathbf{X})_{ij} = \begin{cases} 1 & : i=j; i, j=1, \dots, k \\ 0 & : (i, j) \in \{(1, 2), (2, 1)\}; i, j=1, \dots, k \\ 0.4 & : \text{otherwise} \end{cases}$
- ii. Generate $\mathbf{Y}$ such that $\mathbf{Y} \sim \mathbf{N}_n(\mathbf{X}\boldsymbol{\beta}, \mathbf{I}\sigma^2)$, where $\boldsymbol{\beta} = (0, 20, 30, 0)'$ for $p = 4$, $\boldsymbol{\beta} = (0, 20, 30, 0, 0)'$ for $p = 5$ with $p = k + 1$. "+" indicates inclusion of intercept term.

Let $\mathbf{x}_k$ represents the k^{th} column of $\mathbf{X}$. These coefficients were chosen to reflect a fairly strong linear relationship between $\mathbf{y}$ and $\mathbf{x}_1, \mathbf{x}_2$. As per the correlation structure $\mathbf{\Sigma}_\mathbf{X}$, this was chosen to reflect some multicollinearity in the data and also to adhere to the assumption of ignorable type of missingness

Missingness mechanism—Missing values in $\mathbf{Y}$ were assumed to be MAR with non-response rates (γ) of 25% and 50%. We introduce this type of missingness as follows: If $\Phi_\gamma^{-1}(x_3)$ represent the $(1-\gamma)100^{\text{th}}$ percentile of x_3 , then Y_i is missing if $\{x_3\}_i \geq \Phi_\gamma^{-1}(x_3)$ for $i = 1, \dots, n$.

Method for handling missing data and Imputation algorithm—In this paper we focus on MI as the method for handling missing data. We use a joint normal imputation model. Imputations are created via the standard approach of using a Bayesian predictive model for the missing data given the observed data (Schafer, 1997, p. 154). We used 30 imputations rather than the usual 5, based on the recommendations in Harel (2007).

Model Selection via AIC and AIC variants—For model selection, we focus on AIC because this criterion is common in practice and is readily available in most statistical software. In the case of partially observed data, the *AIC* values for each of the 2^p models vary across the imputed data sets. The issue is how to combine these results. Suppose the regression model (1) for incomplete data is multiply imputed m times. Let s_j^2 represent the maximum likelihood estimate of σ^2 from the j^{th} imputed data set, $j = 1, \dots, m$. We consider the Arithmetic Mean (AM) and Geometric Mean (M) of these m point estimates (of σ^2) to

construct the AIC variants: $AIC_{AM} = n \ln(m^{-1} \sum_{j=1}^m s_j^2) + 2p$, and $AIC_{GM} = n \ln(\frac{1}{m} \prod_{j=1}^m s_j^2) + 2p$. Since

σ^2 is a scale parameter, the Geometric Mean of s_j^2 's serves as a reasonable estimator. The Arithmetic Mean was considered as a contrast to the Geometric Mean. For the sake of comparison, we also consider AIC_{full} (AIC value for a candidate model before introducing missingness), AIC_{CCA} (AIC value for a candidate model under CCA), and AIC_{SI} (AIC value for a candidate model using a single imputation). For each data configuration, the AIC types (AIC_{full} , AIC_{CCA} , AIC_{SI} , AIC_{AM} , AIC_{GM}) were evaluated and the model where their minimum occurred was also recorded. This was repeated 1000 times to yield the proportion of times a model was selected as "best" under each AIC type.

4.2 Comparing success rate of the AIC types

In the simulation results that follow, correct model refers to the model with x_1 and x_2 ; Imputation model refers to the model with x_1 , x_2 and x_3 when $p = 3$, and x_1 , x_2 , x_3 and x_4 when $p = 4$; and Over-fitted models refers to models that contain x_1 and x_2 along with some other variables. In our discussion we only presents results for $\sigma = 1$ because the results were similar across the levels of σ considered in our simulations; detail results for each level of σ are provided in the appendix section.

The proportion of times a model was selected (over 1000 simulations) by each AIC type for $\sigma = 1$ and $\gamma = 25\%$ are summarized in tables 1 and 2. In table 1, for the case $n = 50$, we find that none of the criteria considered work properly. The AIC types tend to select the model with x_1 and x_3 ; we also observe that the pattern of model selection by AIC_{AM} and AIC_{GM} is very similar to that of AIC_{full} . However, when $n = 100$ the good model selection rate (i.e. correct model and over-fitted model) chosen by each AIC type increases to 100%. The correct model success rate of AIC_{CCA} (83%) is much closer to AIC_{full} (85%). This unexpected finding may be result of 25% rate of missingness not being large enough, when $n = 100$, to reflect the inefficiency of CCA. Similarly, the correct model success rate of AIC_{SI} (57%), though relatively higher than AIC_{AM} and AIC_{GM} , is not advocated because this method assumes that imputed values are the true data values. The results of AIC_{AM} and AIC_{GM} tend to over-fit, which is a known natural tendency of AIC (Hurvich and Tsai, 1989).

In table 2, for the case $n = 50$, we again find that none of the criteria work properly. The AIC types tend to select the model with x_1 and variables other than x_2 , followed by models not containing x_1 , x_2 or both; we also observe that the pattern of model selection by AIC_{AM} and AIC_{GM} is very similar to that of AIC_{full} . The results for the case with $n = 100$ still aren't promising because the none of the criteria select the good models (i.e. correct, imputation, or

over-fitted models). The AIC types almost always select models with x_2 and variables other than x_1 ; we again observe that the pattern of model selection by AIC_{AM} and AIC_{GM} is very similar to that of AIC_{full} .

When comparing results of table 1 to table 2, we notice that the addition of another (correlated) variable in the data drastically affects the success rate of AIC_{full} . Furthermore, the usual assumption of including all predictors in the imputation model results in the imputation model variables competing with the data generating variables. Though, some correlation between variables is good for the imputation model, our simulations show that this has a negative impact on model selection in the analysis phase. Hence, by no means should AIC be considered a golden standard for model selection and this was clearly illustrated in our simulation results. The cases where $\gamma = 50\%$ gave similar results, thus, for brevity these results are not discussed but their results are provided in the appendix section.

5. Conclusion

Model selection is an important problem in applied research. Moreover, the problem of incomplete data occurs in almost all data collection endeavors. Combining these two complications can cause hardship to data analysts. Commonly used model selection procedures such as AIC are assumed to work similarly in complete and incomplete data (there is almost no work published on this topic). But this assumption is far from the truth.

In our simulations we noticed that in some of the data configurations the pattern of AIC_{AM} and AIC_{GM} was almost similar to that of AIC_{full} . However, this does not mean we advocate their use because AIC_{full} itself may be unreliable for good model selection (correct and over-fitted models). For example, in our simulations good model selection rate of AIC_{full} was drastically reduced when the number of predictor variables in the data was increased to $k = 4$ even when $n = 100$. Though, our simulations considered a simplistic regression with only 3 or 4 predictors, we were able to demonstrate the uncertainty of AIC in model selection even when $n = 100$ appears to be sufficiently larger than $p = 4$. The poor performance of AIC in the case with $p = 4$ may be attributed to $n = 50, 100$ not being sufficiently larger than p , but the question remains how large of n should be considered when the focus is on model selection, i.e. the rate at which sample size affects reliability of a model selection procedure. Furthermore, our simulations indicate that the issue of model selection with imputed data is not only a question of how to combine model selection results from imputed data, but is also about the impact of the assumed imputation model on model selection in the analysis phase. The importance of additional variables in the imputation model is exaggerated in the analysis phase when performing model selection and this exaggerated importance increases with rate of missingness. This complexity can't be resolved by merely increasing the number of imputations. Thus, one should not assume that MI will be forgiving when interest lies on model selection in the analysis phase. There is a need for generalization of complete data model selection procedures that addresses these concerns for data with incomplete response and covariate data. We hope that this work entices researchers to further explore this issue and recent work published on this matter.

6. Additional Comments

In our simulation study we also considered model selection using the Bayesian Information Criterion (BIC) and found similar model selection rates (hence, results not included). This is not surprising because AIC and BIC formulae differ only in penalty term. The issue is still about generalizing such model selection procedures to account for impact of imputation model on model selection in the analysis phase.

Acknowledgments

We would like to thank the referees for providing comments that helped in improving the quality of our paper. This project was partially supported by Award Number K01MH087219 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

References

- Akaike H. A new look at statistical model identification. *IEEE Transactions on Automatic Control*. 1974; 19(6):716–723.
- Burnham K, Anderson D. Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research*. 2004; 33:261–304.
- Chamberlain T. The method of multiple working hypotheses. *Science*. 1890; 15:93.
- Claeskens G, Consentino F. Variable selection with incomplete covariate data. *Biometrics*. 2008; 64:1062–1096. [PubMed: 18371121]
- Collins L, Schafer J, Kam C. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*. 2001; 6:330–351. [PubMed: 11778676]
- Consentino F, Claeskens G. Order selection tests with multiply-imputed data. *Computational Statistics and Data Analysis*. 2010; 54(10):2284–2295.
- Harel O. Inferences on missing information under multiple imputation and two-stage multiple imputation. *Stat. Method*. 2007; 4:75.
- Harel O. The estimation of R^2 and adjusted R^2 in incomplete data sets using multiple imputation. *Journal of Applied Statistics*. 2009; 36(10):1109–1118.
- Hurvich C, Tsai C. Regression and time series model selection in small samples. *Biometrika*. 1989; 76:297–307.
- Hurvich C, Tsai C. The impact of model selection on inference in linear regression. *The American Statistician*. 1990; 44:214–217.
- Ibrahim JG. Incomplete data in generalized linear models. *Journal of the American Statistical Association*. 1990; 85:765–769.
- Little, R.; Rubin, D. *Statistical analysis with missing data*. 2 ed.. New York: John Wiley & Sons; 2002.
- Meng XL, Rubin DB. Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika*. 1992; 79:103–111.
- Miller, A. *Subset Selection in Regression*. London: Chapman & Hall; 2002.
- Rao C, Wu Y. Model selection. *Lecture Notes-Monograph Series*. 2001; 38:1–64.
- Rubin D. Inference and missing data. *Biometrika*. 1976; 63:581.
- Rubin, D. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons; 1987.
- Schaffer, J. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall; 1997.
- Schaffer J, Graham J. Missing data: our view of the state of the art. *Psychol. Methods*. 2002; 7:147. [PubMed: 12090408]
- Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Statist. Med*. 2008; 27:3227–3246.
- Yang X, Belin TR, Boscardin WJ. Imputation and variable selection in linear regression models with missing covariates. *Biometrics*. 2005; 61(2):498–506. [PubMed: 16011697]

Appendix

In order to interpret that simulation results that follow, we use certain labels to indicate variables (model) selected by each AIC type as "best". These labels are as follows: $\mathcal{C}$ indicates the variables in the true data generating model i.e. $\{x_1 x_2\}$, $\mathcal{Q}$ indicates the variables in the true imputation model i.e. $\{x_1 x_2 x_3\}$ when $p = 3$ and $\{x_1 x_2 x_3 x_4\}$ when $p = 4$, $\mathcal{Q}_{x_1}$ indicates the model with x_1 and variables other than x_2 , $\mathcal{Q}_{x_2}$ indicates the model with x_2 and variables other than x_1 , $\mathcal{O}$ indicates over-fitted models that contain true data

generating variables and some other variables (for $p = 3$, $\mathcal{Q} = \emptyset$, $\mathcal{W}$ indicates models not containing any of the true data generating variables).

Table 1

Model Selection Rates of each AIC type in data sets with $\sigma = 1$, $k = 3$ and $\gamma = 25\%$

n	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}
50	{ x_1, x_3 }									
	0.030	0.194	0.029	0.029	0.029	0.924	0.706	0.925	0.922	0.922
	{ x_2 }									
	0.011	0.012	0.01	0.011	0.011	0	0.002	0	0.001	0.001
	Models not containing any of the correct model variables ({ x_3 })									
100	0	0	0	0	0	0.035	0.086	0.036	0.037	0.037
	Imputation model ({ x_1, x_2, x_3 })									
	0.853	0.830	0.570	0.442	0.458	0.147	0.170	0.43	0.558	0.542

Table 2

Model Selection Rates of each AIC type in data sets with $\sigma = 1$, $k = 4$ and $\gamma = 25\%$

	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}
n	x_1					Models with x_1 and variables other than x_2				
50	0.027	0.101	0.029	0.028	0.028	0.611	0.504	0.615	0.615	0.615
100	0	0	0	0	0	0.001	0	0	0.001	0.001
n	x_2					Models with x_2 and variables other than x_1				
50	0.004	0.004	0.004	0.004	0.004	0.002	0.006	0.002	0.002	0.002
100	0	0.007	0	0	0	0.999	0.993	1	0.999	0.999
n	Correct, Imputation or Over-fitted models					Models not containing any of the correct model variables				
50	0	0	0	0	0	0.356	0.385	0.350	0.351	0.351
100	0	0	0	0	0	0	0	0	0	0

Table 3

Model selection rates of each AIC type in data sets with $n = 50$, $k = 3$ and $\gamma = 25\%$

σ	x_1					x_2					$\mathcal{G}$					$\mathcal{H}$				
	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}
0.3	0.030	0.195	0.029	0.030	0.030	0.030	0.195	0.029	0.030	0.030	0.036	0.087	0.036	0.036	0.036	0.036	0.087	0.036	0.036	0.036
0.5	0.030	0.195	0.029	0.030	0.030	0.030	0.195	0.029	0.030	0.030	0.035	0.088	0.036	0.036	0.036	0.036	0.088	0.036	0.036	0.036
0.75	0.030	0.196	0.029	0.029	0.029	0.035	0.196	0.029	0.029	0.029	0.035	0.087	0.036	0.036	0.036	0.036	0.087	0.036	0.036	0.036
1	0.030	0.194	0.029	0.029	0.029	0.035	0.194	0.029	0.029	0.029	0.035	0.087	0.036	0.036	0.036	0.036	0.087	0.036	0.036	0.036
2	0.027	0.190	0.028	0.028	0.028	0.031	0.190	0.028	0.028	0.028	0.037	0.087	0.034	0.034	0.035	0.037	0.087	0.034	0.035	0.035
4	0.030	0.193	0.034	0.031	0.031	0.037	0.193	0.034	0.031	0.031	0.041	0.090	0.041	0.041	0.041	0.041	0.090	0.041	0.041	0.041
6	0.034	0.198	0.041	0.037	0.037	0.043	0.198	0.041	0.037	0.037	0.043	0.101	0.047	0.047	0.047	0.047	0.101	0.047	0.047	0.047
8	0.042	0.211	0.054	0.040	0.041	0.049	0.211	0.054	0.040	0.041	0.049	0.104	0.050	0.050	0.050	0.049	0.104	0.050	0.045	0.045
0.3	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
0.5	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
0.75	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
1	0.011	0.012	0.01	0.011	0.011	0.011	0.012	0.01	0.011	0.011	0.011	0.012	0.01	0.011	0.011	0.011	0.012	0.01	0.011	0.011
2	0.011	0.013	0.013	0.014	0.013	0.013	0.013	0.013	0.014	0.013	0.013	0.013	0.013	0.014	0.013	0.013	0.013	0.013	0.014	0.013
4	0.016	0.012	0.018	0.016	0.016	0.016	0.012	0.018	0.016	0.016	0.016	0.012	0.018	0.016	0.016	0.016	0.012	0.018	0.016	0.016
6	0.018	0.012	0.019	0.014	0.014	0.014	0.012	0.019	0.014	0.014	0.014	0.012	0.019	0.014	0.014	0.014	0.012	0.019	0.014	0.014
8	0.020	0.014	0.029	0.025	0.025	0.025	0.014	0.029	0.025	0.025	0.025	0.014	0.029	0.025	0.025	0.025	0.014	0.029	0.025	0.025
0.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4

Model selection rates of each AIC type in data sets with $n = 100$, $k = 3$ and $\gamma = 25\%$

σ	$\mathcal{L}$					$\mathcal{Q}$				
	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}
0.3	0.853	0.830	0.539	0.426	0.443	0.147	0.170	0.461	0.574	0.557
0.5	0.853	0.830	0.553	0.435	0.447	0.147	0.170	0.447	0.565	0.553
0.75	0.853	0.830	0.559	0.443	0.456	0.147	0.170	0.441	0.557	0.544
1	0.853	0.830	0.570	0.442	0.458	0.147	0.170	0.43	0.558	0.542
2	0.853	0.830	0.564	0.430	0.446	0.147	0.170	0.436	0.57	0.554
4	0.853	0.830	0.556	0.456	0.473	0.147	0.170	0.444	0.544	0.527
6	0.853	0.830	0.552	0.442	0.454	0.147	0.170	0.448	0.558	0.546
8	0.853	0.830	0.546	0.436	0.450	0.147	0.170	0.454	0.564	0.55

Table 5

Model selection rates of each AIC type in data sets with $n = 50$, $k = 4$ and $\gamma = 25\%$

σ	x_1					x_2					$\mathcal{G}, \mathcal{Q}, \mathcal{O}$					$\mathcal{W}$				
	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{St}	AIC _{AM}	AIC _{GM}
0.3	0.027	0.100	0.025	0.025	0.025	0.002	0.100	0.025	0.025	0.025	0.002	0.100	0.025	0.025	0.025	0.002	0.100	0.025	0.025	0.025
0.5	0.026	0.100	0.026	0.026	0.026	0.003	0.100	0.026	0.026	0.026	0.003	0.100	0.026	0.026	0.026	0.003	0.100	0.026	0.026	0.026
0.75	0.027	0.102	0.028	0.028	0.028	0.004	0.102	0.028	0.028	0.028	0.004	0.102	0.028	0.028	0.028	0.004	0.102	0.028	0.028	0.028
1	0.027	0.101	0.029	0.029	0.028	0.004	0.101	0.029	0.028	0.028	0.004	0.101	0.029	0.028	0.028	0.004	0.101	0.029	0.028	0.028
2	0.025	0.101	0.03	0.028	0.028	0.004	0.101	0.03	0.028	0.028	0.004	0.101	0.03	0.028	0.028	0.004	0.101	0.03	0.028	0.028
4	0.026	0.102	0.035	0.03	0.03	0.004	0.102	0.035	0.03	0.03	0.004	0.102	0.035	0.03	0.03	0.004	0.102	0.035	0.03	0.03
6	0.030	0.107	0.038	0.028	0.029	0.006	0.107	0.038	0.028	0.029	0.006	0.107	0.038	0.028	0.029	0.006	0.107	0.038	0.028	0.029
8	0.032	0.099	0.033	0.034	0.036	0.007	0.099	0.033	0.034	0.036	0.007	0.099	0.033	0.034	0.036	0.007	0.099	0.033	0.034	0.036
$\mathcal{G}, \mathcal{Q}, \mathcal{O}$																				
0.3	0.004	0.003	0.004	0.003	0.003	0.002	0.003	0.004	0.003	0.003	0.002	0.003	0.004	0.003	0.003	0.002	0.003	0.004	0.003	0.003
0.5	0.003	0.004	0.003	0.003	0.003	0.002	0.004	0.003	0.003	0.003	0.002	0.004	0.003	0.003	0.003	0.002	0.004	0.003	0.003	0.003
0.75	0.003	0.003	0.004	0.004	0.004	0.002	0.003	0.004	0.004	0.004	0.002	0.003	0.004	0.004	0.004	0.002	0.003	0.004	0.004	0.004
1	0.004	0.004	0.004	0.004	0.004	0.002	0.004	0.004	0.004	0.004	0.002	0.004	0.004	0.004	0.004	0.002	0.004	0.004	0.004	0.004
2	0.006	0.004	0.007	0.004	0.004	0.002	0.004	0.007	0.004	0.004	0.002	0.004	0.007	0.004	0.004	0.002	0.004	0.007	0.004	0.004
4	0.007	0.008	0.010	0.010	0.010	0.001	0.008	0.010	0.010	0.010	0.001	0.008	0.010	0.010	0.010	0.001	0.008	0.010	0.010	0.010
6	0.013	0.01	0.011	0.015	0.015	0	0.01	0.011	0.015	0.015	0	0.01	0.011	0.015	0.015	0	0.01	0.011	0.015	0.015
8	0.01	0.012	0.016	0.014	0.014	0.001	0.012	0.016	0.014	0.014	0.001	0.012	0.016	0.014	0.014	0.001	0.012	0.016	0.014	0.014
$\mathcal{W}$																				
0.3	0	0	0	0	0.35	0.383	0	0	0	0.35	0.383	0	0	0.35	0.383	0.349	0.354	0.354	0	0
0.5	0	0	0	0	0.349	0.385	0	0	0	0.349	0.385	0.351	0.352	0.352	0.352	0.351	0.352	0.352	0.352	0
0.75	0	0	0	0	0.353	0.38	0	0	0	0.353	0.38	0.352	0.353	0.353	0.353	0.350	0.353	0.353	0.353	0
1	0	0	0	0	0.356	0.385	0	0	0	0.356	0.385	0.350	0.351	0.351	0.351	0.350	0.351	0.351	0.351	0
2	0	0	0	0	0.356	0.379	0	0	0	0.356	0.379	0.353	0.358	0.358	0.358	0.353	0.358	0.357	0.357	0
4	0	0	0	0	0.343	0.375	0	0	0	0.343	0.375	0.345	0.343	0.343	0.343	0.341	0.343	0.343	0.343	0
6	0	0	0	0	0.349	0.367	0	0	0	0.349	0.367	0.341	0.348	0.348	0.348	0.341	0.348	0.349	0.349	0
8	0	0	0	0	0.340	0.367	0	0	0	0.340	0.367	0.360	0.338	0.338	0.338	0.360	0.338	0.334	0.334	0

Table 6

Model selection rates of each AIC type in data sets with $n = 100$, $k = 4$ and $\gamma = 25\%$

σ	x_1					$\mathcal{Q}_{x_1}$				
	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}
0.3	0	0	0	0	0	0.001	0	0.001	0.001	0.001
0.5	0	0	0	0	0	0.001	0	0	0.001	0.001
0.75	0	0	0	0	0	0.001	0	0.001	0.001	0.001
1	0	0	0	0	0	0.001	0	0	0.001	0.001
2	0	0	0	0	0	0.001	0	0	0	0
4	0	0	0	0	0	0.001	0.001	0.001	0.001	0.001
6	0	0	0	0	0	0.002	0.003	0.001	0.001	0.001
8	0	0	0	0	0	0.003	0.003	0.002	0.001	0.001
σ	x_2					$\mathcal{Q}_{x_2}$				
	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}
0.3	0	0.006	0	0	0	0.999	0.994	0.999	0.999	0.999
0.5	0	0.006	0	0	0	0.999	0.994	1	0.999	0.999
0.75	0	0.007	0	0	0	0.999	0.993	0.999	0.999	0.999
1	0	0.007	0	0	0	0.999	0.993	1	0.999	0.999
2	0	0.005	0	0	0	0.999	0.995	1	1	1
4	0	0.008	0	0	0	0.999	0.991	0.999	0.999	0.999
6	0	0.007	0	0	0	0.998	0.990	0.999	0.999	0.999
8	0	0.007	0	0.002	0.002	0.997	0.990	0.998	0.997	0.997
σ	$\mathcal{L}, \mathcal{Q}, \hat{\mathcal{Q}}$					$\mathcal{M}$				
	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{SI}	AIC _{AM}	AIC _{GM}
All σ	0	0	0	0	0	0	0	0	0	0

Table 7

Model selection rates of each AIC type in data sets with $n = 50$, $k = 3$ and $\gamma = 50\%$

	x_1					x_2					φ_{x_1}					φ_{x_2}					$\mathcal{M}$				
σ	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}
0.3	0.03	0.415	0.03	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.922	0.402	0.921	0.922	0.922	0.922	0.402	0.921	0.922	0.922	0.922	0.922	0.922	0.922	0.922
0.5	0.03	0.413	0.031	0.028	0.028	0.028	0.028	0.028	0.028	0.028	0.924	0.402	0.919	0.922	0.922	0.922	0.402	0.919	0.922	0.922	0.922	0.922	0.922	0.922	0.922
0.75	0.03	0.411	0.031	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.924	0.403	0.919	0.922	0.922	0.922	0.403	0.919	0.922	0.922	0.922	0.922	0.922	0.922	0.922
1	0.03	0.413	0.031	0.029	0.029	0.029	0.029	0.029	0.029	0.029	0.924	0.403	0.919	0.922	0.922	0.922	0.403	0.919	0.922	0.922	0.922	0.922	0.922	0.922	0.922
2	0.027	0.41	0.027	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.925	0.402	0.917	0.923	0.923	0.923	0.4	0.871	0.901	0.901	0.901	0.901	0.901	0.901	0.901
4	0.03	0.407	0.049	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.913	0.392	0.842	0.883	0.883	0.883	0.392	0.842	0.883	0.883	0.883	0.883	0.883	0.883	0.883
6	0.034	0.409	0.061	0.044	0.045	0.045	0.045	0.045	0.045	0.045	0.889	0.386	0.771	0.831	0.831	0.831	0.386	0.771	0.831	0.831	0.831	0.831	0.831	0.831	0.831
8	0.042	0.401	0.104	0.059	0.06	0.06	0.06	0.06	0.06	0.06	0.889	0.386	0.771	0.831	0.831	0.831	0.386	0.771	0.831	0.831	0.831	0.831	0.831	0.831	0.831
σ	x_1					x_2					φ_{x_1}					φ_{x_2}					$\mathcal{M}$				
0.3	0.011	0.019	0.011	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.001	0.006	0.001	0.001	0.001	0.001	0.006	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
0.5	0.011	0.02	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0	0.006	0.001	0.001	0.001	0.001	0.006	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
0.75	0.011	0.02	0.013	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0	0.007	0.001	0.001	0.001	0.001	0.007	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
1	0.011	0.019	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0	0.008	0.001	0.001	0.001	0.001	0.008	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
2	0.011	0.022	0.016	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0	0.009	0	0.001	0.001	0.001	0.009	0	0.001	0.001	0.001	0.001	0.001	0.001	0.001
4	0.016	0.023	0.027	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0	0.011	0.001	0.001	0.001	0.001	0.011	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
6	0.018	0.019	0.034	0.028	0.027	0.027	0.028	0.027	0.027	0.027	0	0.016	0.002	0.002	0	0	0.016	0.002	0.002	0	0	0	0	0	0
8	0.02	0.027	0.044	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0	0.017	0.004	0.004	0.001	0.001	0.017	0.004	0.004	0.001	0.001	0.001	0.001	0.001	0.001
σ	$\mathcal{L}, \mathcal{Q}, \mathcal{O}$					$\mathcal{M}$					$\mathcal{M}$					$\mathcal{M}$					$\mathcal{M}$				
0.3	0	0	0	0	0	0	0	0	0	0	0.036	0.158	0.037	0.036	0.036	0.036	0.158	0.037	0.036	0.036	0.036	0.036	0.036	0.036	0.036
0.5	0	0	0	0	0	0	0	0	0	0	0.035	0.159	0.037	0.037	0.037	0.037	0.159	0.037	0.037	0.037	0.037	0.037	0.037	0.037	0.037
0.75	0	0	0	0	0	0	0	0	0	0	0.035	0.159	0.036	0.036	0.036	0.036	0.159	0.036	0.036	0.036	0.036	0.036	0.036	0.036	0.036
1	0	0	0	0	0	0	0	0	0	0	0.035	0.157	0.038	0.038	0.038	0.038	0.157	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.038
2	0	0	0	0	0	0	0	0	0	0	0.037	0.157	0.04	0.04	0.04	0.04	0.157	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
4	0	0	0	0	0	0	0	0	0	0	0.041	0.159	0.052	0.052	0.052	0.052	0.159	0.052	0.052	0.052	0.052	0.052	0.052	0.052	0.052
6	0	0	0	0	0	0	0	0	0	0	0.043	0.164	0.061	0.061	0.061	0.061	0.164	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061
8	0	0	0	0	0	0	0	0	0	0	0.049	0.169	0.077	0.077	0.077	0.077	0.169	0.077	0.077	0.077	0.077	0.077	0.077	0.077	0.077

Table 8

Model selection rates of each AIC type in data sets with $n = 100$, $k = 3$ and $\gamma = 50\%$

σ	$\mathcal{C}$					$\mathcal{Q}$				
	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}
0.3	0.853	0.829	0.358	0.016	0.018	0.147	0.171	0.642	0.984	0.982
0.5	0.853	0.829	0.349	0.015	0.016	0.147	0.171	0.651	0.985	0.984
0.75	0.853	0.829	0.371	0.022	0.026	0.147	0.171	0.629	0.978	0.974
1	0.853	0.829	0.37	0.014	0.016	0.147	0.171	0.63	0.986	0.984
2	0.853	0.829	0.361	0.022	0.026	0.147	0.171	0.639	0.978	0.974
4	0.853	0.829	0.344	0.016	0.023	0.147	0.171	0.656	0.984	0.977
6	0.853	0.829	0.375	0.018	0.02	0.147	0.171	0.625	0.982	0.98
8	0.853	0.829	0.376	0.02	0.023	0.147	0.171	0.624	0.98	0.977

Table 9

Model selection rates of each AIC type in data sets with $n = 50$, $k = 4$ and $\gamma = 50\%$

	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	$\mathbf{x}_1$					$\mathcal{Q}_{\mathbf{x}_1}$				
σ															
0.3	0.027	0.117	0.026	0.024	0.024	0.617	0.478	0.614	0.622		0.622		0.622		0.622
0.5	0.026	0.116	0.023	0.025	0.025	0.62	0.481	0.617	0.623		0.623		0.623		0.623
0.75	0.027	0.114	0.023	0.025	0.025	0.615	0.478	0.618	0.618		0.618		0.618		0.618
1	0.027	0.112	0.025	0.025	0.025	0.611	0.477	0.606	0.611		0.611		0.611		0.611
2	0.025	0.11	0.029	0.025	0.025	0.611	0.489	0.602	0.615		0.615		0.615		0.614
4	0.026	0.111	0.027	0.023	0.023	0.623	0.481	0.598	0.615		0.615		0.615		0.616
6	0.03	0.114	0.043	0.032	0.031	0.608	0.484	0.576	0.591		0.591		0.591		0.593
8	0.032	0.126	0.045	0.035	0.034	0.617	0.468	0.552	0.591		0.591		0.591		0.593
	$\mathbf{x}_2$					$\mathcal{Q}_{\mathbf{x}_2}$									
σ															
0.3	0.004	0.013	0.003	0.003	0.003	0.002	0.025	0.002	0.002		0.002		0.002		0.002
0.5	0.003	0.014	0.003	0.002	0.002	0.002	0.026	0.002	0.002		0.002		0.002		0.002
0.75	0.003	0.015	0.005	0.002	0.002	0.002	0.025	0.002	0.002		0.002		0.002		0.002
1	0.004	0.013	0.005	0.002	0.002	0.002	0.024	0.002	0.002		0.002		0.002		0.002
2	0.006	0.011	0.005	0.005	0.005	0.002	0.019	0.001	0.002		0.001		0.002		0.002
4	0.007	0.019	0.01	0.011	0.011	0.001	0.018	0.003	0.001		0.003		0.001		0.001
6	0.013	0.026	0.014	0.011	0.012	0	0.02	0.003	0.003		0.003		0.003		0.002
8	0.01	0.024	0.02	0.02	0.02	0.001	0.021	0.006	0.004		0.004		0.004		0.004
	$\mathcal{E}, \mathcal{Q}, \mathcal{O}$					$\mathcal{W}$									
σ															
0.3	0	0	0	0	0	0.35	0.367	0.355	0.349		0.349		0.349		0.349
0.5	0	0	0	0	0	0.349	0.363	0.355	0.348		0.348		0.348		0.348
0.75	0	0	0	0	0	0.353	0.368	0.352	0.353		0.353		0.353		0.353
1	0	0	0	0	0	0.356	0.374	0.362	0.36		0.36		0.36		0.36
2	0	0	0	0	0	0.356	0.371	0.363	0.353		0.353		0.353		0.354
4	0	0	0	0	0	0.343	0.371	0.362	0.35		0.349		0.349		0.349
6	0	0	0	0	0	0.349	0.356	0.364	0.363		0.363		0.363		0.362
8	0	0	0	0	0	0.34	0.361	0.377	0.35		0.35		0.349		0.349

Table 10

Model selection rates of each AIC type in data sets with $n = 100$, $k = 4$ and $\gamma = 50\%$

	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}	AIC _{full}	AIC _{CCA}	AIC _{CD}	AIC _{AM}	AIC _{GM}
σ	x_1					ψ_{x_1}				
0.3	0	0	0	0	0	0.001	0.003	0.001	0.001	0.001
0.5	0	0	0	0	0	0.001	0.004	0.001	0.001	0.001
0.75	0	0	0	0	0	0.001	0.003	0.001	0.001	0.001
1	0	0	0	0	0	0.001	0.004	0.001	0.001	0.001
2	0	0	0	0	0	0.001	0.007	0.001	0.001	0.001
4	0	0	0	0	0	0.001	0.007	0.002	0.001	0.001
6	0	0	0	0	0	0.002	0.008	0.003	0.002	0.002
8	0	0	0	0	0	0.003	0.013	0.004	0.004	0.004
σ	x_2					ψ_{x_2}				
0.3	0	0.032	0	0	0	0.999	0.965	0.999	0.999	0.999
0.5	0	0.029	0	0	0	0.999	0.967	0.999	0.999	0.999
0.75	0	0.026	0	0	0	0.999	0.971	0.999	0.999	0.999
1	0	0.028	0	0	0	0.999	0.968	0.999	0.999	0.999
2	0	0.03	0	0	0	0.999	0.963	0.999	0.999	0.999
4	0	0.033	0	0	0	0.999	0.96	0.998	0.999	0.999
6	0	0.035	0.002	0.001	0.001	0.998	0.957	0.995	0.997	0.997
8	0	0.05	0.001	0.001	0.001	0.997	0.937	0.995	0.995	0.995
σ	$\mathcal{L}, \varphi, \theta$					$\mathcal{W}$				
All_σ	0	0	0	0	0	0	0	0	0	0