Addressing Challenges of Hierarchical Structural Equation Modeling in Animal Agriculture

by

Kessinee Chitakasempornkul

B.BA., Mahidol University International Colleges, 2008
M.S., California State University Long Beach, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

Abstract

Feeding a world population of 9 billion people by 2050 is a fundamental challenge of our generation. Animal agriculture is positioned to play a major role on this challenge by ensuring a safe and secure supply of animal protein for food. Thus motivated, an understanding of the mechanistic interconnections between multiple outcomes in agricultural production systems is critical. Structural equation models (SEM) are being increasingly used for investigating directionality in the associations between outcomes in the system. Agricultural data pose peculiar challenges to the implementation of SEM, among them its structured architecture and its multidimensional heterogeneity. For example, observations on a given outcome collected at the animal level are often not mutually independent but rather likely to have correlation patterns due to clustering within pens or cohorts, which in turn may be subjected to common management or business practices defined at the level of a commercial operation. Also, agricultural outcomes of interest are often correlated and at multiple levels. Furthermore, a key assumption underlying SEM is that of a causal homogeneity, whereby the structural coefficients defining functional links in a network are assumed homogeneous and impervious to environmental conditions or management factors. This assumption seems particularly questionable in the context of animal agriculture, where production systems are regularly subjected to explicit interventions intended to optimize the necessary trade-offs between efficacy and efficiency of production.

Despite the recent extension of SEM to a mixed-models framework, inferential issues related to hierarchical data architecture in the context of designed experiments and observational studies are not well understood. Hence, my dissertation work investigates the importance of properly specifying data structure in a hierarchical SEM. My research further develops methodological extensions to SEM to account for heterogeneity in the structural coefficients of a

network and characterizes problems to be expected otherwise. Throughout my PhD dissertation, I implemented SEM in a hierarchical Bayesian framework, and I used as motivator an observational dataset in beef cattle feedlot production and a dataset from a designed experiment in swine reproduction.

I first evaluated the inferential implications of properly accounting for (or ignoring) existent correlation structure due to data architecture when modeling feedlot data with SEM. Results indicated impaired model fit, biased estimation and precision loss for SEM parameters when data architecture was mispecified or ignored. I then investigated potential causal interconnections between reproductive performance outcomes in swine, for which I leveraged the mixed-models adapted inductive causation algorithm to search for and infer upon causal links. Results indicated reproductive networks distinctive by parity groups, thereby suggesting potential network heterogeneity; this finding was in direct conflict with the standard SEM assumption of causal homogeneity. Therefore, I proposed and developed a methodological extension to hierarchical SEM that explicitly specifies structural coefficients as functions of systematic and non-systematic sources of variation, thus allowing for hierarchical heterogeneity of network links in structured data. I validated the proposed method using a simulation study and applied it to assess heterogeneity of functional reproductive links in swine.

Overall, this dissertation characterizes problems of SEM-based modeling and develops a general approach to hierarchical SEM for the joint network-type analysis of multiple outcomes with potential heterogeneity in functional links.

Addressing Challenges of Hierarchical Structural Equation Modeling in Animal Agriculture

by

Kessinee Chitakasempornkul

B.BA., Mahidol University International Colleges, 2008
M.S., California State University Long Beach, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2019

Approved by:

Major Professor
Nora M. Bello, DVM, PhD

Copyright

Abstract

Feeding a world population of 9 billion people by 2050 is a fundamental challenge of our generation. Animal agriculture is positioned to play a major role on this challenge by ensuring a safe and secure supply of animal protein for food. Thus motivated, an understanding of the mechanistic interconnections between multiple outcomes in agricultural production systems is critical. Structural equation models (SEM) are being increasingly used for investigating directionality in the associations between outcomes in the system. Agricultural data pose peculiar challenges to the implementation of SEM, among them its structured architecture and its multidimensional heterogeneity. For example, observations on a given outcome collected at the animal level are often not mutually independent but rather likely to have correlation patterns due to clustering within pens or cohorts, which in turn may be subjected to common management or business practices defined at the level of a commercial operation. Also, agricultural outcomes of interest are often correlated and at multiple levels. Furthermore, a key assumption underlying SEM is that of a causal homogeneity, whereby the structural coefficients defining functional links in a network are assumed homogeneous and impervious to environmental conditions or management factors. This assumption seems particularly questionable in the context of animal agriculture, where production systems are regularly subjected to explicit interventions intended to optimize the necessary trade-offs between efficacy and efficiency of production.

Despite the recent extension of SEM to a mixed-models framework, inferential issues related to hierarchical data architecture in the context of designed experiments and observational studies are not well understood. Hence, my dissertation work investigates the importance of properly specifying data structure in a hierarchical SEM. My research further develops methodological extensions to SEM to account for heterogeneity in the structural coefficients of a

network and characterizes problems to be expected otherwise. Throughout my PhD dissertation, I implemented SEM in a hierarchical Bayesian framework, and I used as motivator an observational dataset in beef cattle feedlot production and a dataset from a designed experiment in swine reproduction.

I first evaluated the inferential implications of properly accounting for (or ignoring) existent correlation structure due to data architecture when modeling feedlot data with SEM. Results indicated impaired model fit, biased estimation and precision loss for SEM parameters when data architecture was mispecified or ignored.

I then investigated potential causal interconnections between reproductive performance outcomes in swine, for which I leveraged the mixed-models adapted inductive causation algorithm to search for and infer upon causal links. Results indicated reproductive networks distinctive by parity groups, thereby suggesting potential network heterogeneity; this finding was in direct conflict with the standard SEM assumption of causal homogeneity. Therefore, I proposed and developed a methodological extension to hierarchical SEM that explicitly specifies structural coefficients as functions of systematic and non-systematic sources of variation, thus allowing for hierarchical heterogeneity of network links in structured data. I validated the proposed method using a simulation study and applied it to assess heterogeneity of functional reproductive links in swine.

Overall, this dissertation characterizes problems of SEM-based modeling and develops a general approach to hierarchical SEM for the joint network-type analysis of multiple outcomes with potentially heterogeneity in functional links.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my major professor, Dr. Nora M. Bello, for her thoughtful guidance over the years, for her unconditional support and encouragement, and her willingness to give her time and endless patience over many things that we have gone through together. I cannot be where I have been so far without you. Indeed, I am very much grateful for your tremendous help and support as the best of mentor. I truly appreciate every moment with her. Once again, thank you Dr. Bello!

I would like to express my sincere thank you to my supervisory committee members, Dr. Guilherme Rosa, Dr. Abigail Jager, and Dr. Trevor Hefley. I cannot thank enough for your tremendous help and support over the years. I truly appreciate your time and dedication. I am also grateful to Dr. Marianne Korten, my outside Chairperson, for your time and support. Dr. James Neill and Dr. Christopher Vahl, former and current department heads, respectively, I am very much thankful for the unconditional support I have accomplished over my graduate training at Kansas State University.

My most sincere thanks and appreciation also goes to my faculty members, collaborative CORE teams and fellow graduate students at Kansas State University. I also would like to thank faculty members and fellow graduate students at the department of Animal Sciences, University of Wisconsin-Madison, I am very much grateful to be part of your teams. Further, I am very much grateful to Messmer family for your unconditional love and support. Messmer!, you are a big part of who I am today.

My mother and father, your love and support are my foundation of my success. Thank you very much from the bottom of my heart.

Dedication

To Mom & Dad

For your endless love and support

"I did it!"

Preface

Chapter 2 and 4 in this dissertation were written in the style required for publication in *Journal of Agricultural Biological and Environmental Statistics*. Chapter 3 was written in the style required for publication in *Journal of Animal Science*.

Chapter 2 corresponds to the following article: Chitakasempornkul, K., Sanderson, M., Cha, E., Renter, D. R., Jager, A., and Bello, N. M. 2018. "Accounting for Data Architecture on Structural Equation Modeling of Feedlot Cattle Performance," *Journal of Agricultural Biological and Environmental Statistics*, 23(4), 529-549.

Chapter 3 is currently in print at *Journal of Animal Science*: Chitakasempornkul, K., M. B. Menegat, G. J. M. Rosa, F. B. Lopes, A. Jager, M. A. D. Gonçalve, S. S. Dritz, M. D. Tokach, R. D. Goodband, and N. M. Bello. 2019. Investigating causal biological relationships between reproductive performance traits in high-performing gilts and sows. J. Anim. Sci.

Chapter 4 corresponds to the pre-peer reviewed version (in preparation) of an invited article for a special issue of *Journal of Agricultural Biological and Environmental Statistics* on "Recent Advances in Design and Analysis of Experiments and Observational Studies in Agriculture" to be published in 2020.

Chapter 1 - Introduction

The challenge of feeding an estimated world population of 9 billion people by 2050 is a major and very tangible concern of our generation. As the global population continues to grow, we face the challenge of growing, managing and sustaining a safe and secure food supply (Godfray et al. 2010). Confronting this global challenge will require truly integrated strategies that cut across scientific disciplines and provide multifaceted system-level solutions to manage and further develop global food production system (Foley et al. 2005, Foley 2011). Animal agriculture serves a major role in providing nutritious wholesome food, particularly animal protein in the form of milk, meat, and eggs. Pivotal to a systems approach to animal agriculture is an in-depth understanding of the mechanistic functional connections that underlie interconnections between animal physiology, management practices, environmental conditions, and performance outcomes of agricultural productions.

Data from animal agriculture is often analyzed by linear mixed models fitted separately to individual outcome variable. Yet, this approach does not allow for claiming how these outcomes might affect each other. Tailored multivariate statistical modeling strategies are considered to tackle and address the complexity of animal production systems. Structural equation models (SEM) are being increasingly used for investigating directionality in the associations between outcome variables. The overall objective of this dissertation is to characterize and address challenges associated with hierarchical SEM in animal agriculture with the overarching goal of providing guidelines and recommendations for practical implementation for integrated systems management.

1.1. Classical multiple-trait model

Classical multiple-trait models (MTM) are typically considered to simultaneously evaluate the joint behavior of multiple outcome variables measured on an individual and study their probabilistic relationships expressed as correlations and covariances (Henderson and Quaas 1976, Van Vleck and Edlin 1984) or functions thereof (Bello et al. 2010). In a mixed-models context, MTM can be used to investigate these relationships at multiple levels of data architecture, for example animal (i.e. residual) or pen (i.e. random effect) levels. However, classical MTM are limited in that they cannot assess directionality of correlations, thereby ignoring potential causal connections among outcome variables (Rosa et al. 2011). For an illustration, for two correlated outcomes y_1 and y_2, the nature of their correlation could be explained by (i) an effect of y_1 on y_2, (ii) an effect of y_2 on y_1, (iii) an effect of a third variable, say y_3, on both y_1 and y_2. That is, classical MTM can be used to estimate correlations but cannot discriminate between plausible mechanisms generating such correlations. Indeed, this limitation of classical MTM hinders their use for exploring and inferring causal connections within a network, as might be of interest in complex systems with multiple interconnected outcomes.

1.2. Structural equation modeling framework

Structural Equation Models (SEM) (Haavelmo 1943) are a special type of multiple-trait models that convey functional links among outcomes. Under certain assumptions (Pearl 2009), these links can be interpreted for causal directionality in the association between outcome variables. This, in turn, enables a system approach to the study of functional interconnections between outcomes and facilitates more comprehensive network-type strategies to evaluate the system as a whole, as opposed to individual outcomes, one at a time. In an SEM, links between outcomes are represented by structural coefficient parameters, thus allowing for one or more

outcome variables to act as predictors of other outcomes in a non-cyclic manner (Gianola and Sorensen 2004). Structural coefficients can then be interpreted as rates of change in a regression-like fashion (Gianola and Sorensen 2004). Under some non-trivial assumptions (Pearl 2009), structural coefficients can further be interpreted as causal effects and discriminate direct and indirect components of the effects that accrue to the observed correlations between outcomes (Wright 1934).

Recently, SEM were adapted to a mixed models framework (Gianola and Sorensen 2004) and have since been increasingly used in agricultural applications, especially in animal genetics and genomics across multiple species. Examples include dairy goats (de los Campos et al. 2006a), dairy cattle (de los Campos et al. 2006b, Wu et al. 2007, Konig et al. 2008, de Maturana et al. 2009), quail (Valente et al. 2011), swine (Varona et al. 2014), and beef cattle (Inoue et al. 2016). However, application of SEM in the context of animal agriculture beyond genetics and genomics has been limited. We are particularly interested in SEM applications in agricultural data from designed experiments and observational studies since agricultural data poses peculiar challenges to the implementation of SEM.

For an illustration of hierarchical SEM modeling, consider 4 outcome variables namely y_1, y_2, y_3, y_4. Each outcome y_j is affected by an independent residual e_j and by the corresponding random effect b_j, whereby these random effects are assumed to be correlated. A hierarchical SEM following Gianola and Sorensen (2004), such that

$$\boldsymbol{y}_i = \boldsymbol{\Lambda}\boldsymbol{y}_i + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b} + \boldsymbol{e}_i \tag{1.1}$$

where $\boldsymbol{y}_i' = [y_{i1}' \; y_{i2}' \; y_{i3}' \; y_{i4}']$ is a vector of $J = 4$ outcome variables ($j = 1, \ldots, J$) observed on the i^{th} subject ($i = 1, 2, \ldots, n$); in turn, $\boldsymbol{\Lambda}$ is a $J \times J$ matrix of zeroes and non-zero structural

coefficient parameters that describe the functional relationships or causal structure amongst outcome variables, such that

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & \lambda_{32} & 0 & 0 \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & 0 \end{bmatrix} \tag{1.2}$$

The structural coefficients $\lambda_{jj'}$ $(j' < j)$ expresses a functional link whereby outcome j' has a direct effect on outcome j. For instance, here, λ_{21} represents the magnitude of the direct effect of $y_1$on y_2 and represents the expected change in y_2 per unit increase in y_1. In Equation (1.2), this is a fully-recursive SEM with a complete lower triangle matrix $\boldsymbol{\Lambda}$ such that all the $\lambda_{jj'}$ indicative of recursive effects are estimated. Notably, the fully recursive SEM specification can be shown to be statistically equivalent to a classical MTM (Gianola and Sorensen 2004). Further, some of those $\lambda_{jj'}$ in matrix $\boldsymbol{\Lambda}$ can be set to zero to elicit specific causal structure.

From Equation (1.1), $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2' \ \boldsymbol{\beta}_3' \ \boldsymbol{\beta}_4']$ is a vector of unknown "fixed-effect" location parameters associated with factors and covariates (e.g. treatment, demographics) through the corresponding incidence matrix $\boldsymbol{X}_i = diag[\boldsymbol{x}_{i1}' \ \boldsymbol{x}_{i2}' \ \boldsymbol{x}_{i3}' \ \boldsymbol{x}_{i4}']$ unique to each subject *i*. Also, $\boldsymbol{b}' = [\boldsymbol{b}_1' \ \boldsymbol{b}_2' \ \boldsymbol{b}_3' \ \boldsymbol{b}_4']$ is a vector of unknown random effects associated with blocking factors or other components of the data architecture through the corresponding design matrix $\boldsymbol{Z}_i = diag[\boldsymbol{z}_{i1}' \ \boldsymbol{z}_{i2}' \ \boldsymbol{z}_{i3}' \ \boldsymbol{z}_{i4}']$. Random effects $\boldsymbol{b}$ are assumed multivariate normally distributed as:

$$\boldsymbol{b} \sim MVN\left(\boldsymbol{0}, \ \boldsymbol{B} \otimes \boldsymbol{I}_q\right) \tag{1.3}$$

where $\boldsymbol{B} = \begin{bmatrix} \sigma^2_{b_1} & \sigma_{b_{12}} & \sigma_{b_{13}} & \sigma_{b_{14}} \\ \sigma_{b_{21}} & \sigma^2_{b_2} & \sigma_{b_{23}} & \sigma_{b_{24}} \\ \sigma_{b_{31}} & \sigma_{b_{32}} & \sigma^2_{b_3} & \sigma_{b_{34}} \\ \sigma_{b_{41}} & \sigma_{b_{42}} & \sigma_{b_{43}} & \sigma^2_{b_4} \end{bmatrix}$, and q is the number of levels of the random effect factor.

Finally, $\boldsymbol{e_i'} = [e_{i1}\ \ e_{i2}\ \ e_{i3}\ \ e_{i4}]$ is the corresponding set of residuals for subject i and is assumed multivariate normal, as $\boldsymbol{e}_i \sim MVN(\boldsymbol{0}, \boldsymbol{R})$, where

$$\boldsymbol{R} = \begin{bmatrix} \sigma^2_{e_1} & 0 & 0 & 0 \\ 0 & \sigma^2_{e_2} & 0 & 0 \\ 0 & 0 & \sigma^2_{e_3} & 0 \\ 0 & 0 & 0 & \sigma^2_{e_4} \end{bmatrix} \tag{1.4}$$

The underlying assumption of a diagonal residual covariance matrix $\boldsymbol{R}$ (i.e. $\sigma_{e_{jj'}} = 0$) is standard in the context of SEM to ensure parameter identifiability in non-cyclic causal framework (Gianola and Sorensen 2004). Any correlations between outcomes above and beyond those due to random effects in the design structure, are attributed to the causal effects among outcomes that are represented by structural coefficients $\lambda_{jj'}$ (Gianola and Sorensen 2004).

1.3. Elicitation of causal structure

Elicitation of hypothesized causal structures in SEM is often based on expert knowledge and prior beliefs (de los Campos et al. 2006a, 2006b, Wu et al. 2007, Konig et al. 2008, de Maturana et al. 2010). In this way, multiple scientifically plausible causal networks can be fitted and compared on the basis of data fit. However, this approach can result in poor exploration of the structure space, since it focuses on a pre-selected subset of casual networks. One should keep in mind that the number of possible causal structures can be very large even when only a few outcome variables are studied.

Instead, search algorithms can be used to explore the structure spaces without applying prior substantive knowledge. One such search algorithm is the inductive causation (IC) algorithm

(Verma and Pearl 1991, 1992, Pearl 2009), which was recently extended to be used with structured data in a hierarchical Bayesian framework (Valente et al. 2010). For the adapted IC algorithm (Valente et al. 2010), the search was conducted at the residual level of a MTM, that is, after conditioning on the random effects that characterize the experimental design (described later). This enhanced IC algorithm has been successfully applied in many applications in animal genetics (Valente et al. 2011, Peñagarican et al. 2015, Inoue et al. 2016). The adapted IC search the data for structures is solely based on information about conditional independent relationships between residuals, assuming that those independencies reflect *d*-separation in the causal structure at residual level.

Making causal interpretation to the structure learned with the adapted IC algorithm does require additional non-trivial assumptions that need cautious attention (Pearl 2009). These critical assumptions includes the Markov condition, the assumption of faithfulness or stability, and causal sufficiency. Firstly, Markov condition is an important causal assumption for observational data; given available information on its parents, a node is conditionally independent of all nondescendent variables in the network (Pearl 2009). Secondly, the assumption of faithfulness or stability ensure that the causal effects (i.e. direct and indirect) are not completely canceled out by opposing causal paths of opposite sign. For example, a direct path cancels out an indirect path, thus failing to recognize the causal mechanistic information. Finally, causal sufficiency is the most delicate assumption since it implies that all relevant confounders in the system are recognized and either have been measured or can be managed by other measured variables. These non-trivial causal assumptions cannot be tested directly from data; in turn, they are the basis of inferring causal claims (Robins 1999).

1.4. Implementation in Bayesian framework

My implementation of a hierarchical SEM was based on the Bayesian statistical framework, which is suitable to handle structured data with multiple layers (Sorensen 2002). Moreover, the Bayesian framework provides a fundamental way of combining prior information with data, thereby allowing for incorporating past information about a parameter (Sorensen and Gianola 2002). More importantly, the results of a Bayesian analysis are interpreted probabilistically and can be represented by exhibiting the entire posterior distribution and/or posterior summaries such as the mean, median, or percentile depending on the research question posed (Sorensen 2002). The implementation of Bayesian inference based on Markov Chain Monte Carlo (MCMC) make computations tractable, (Gilks et al. 1996), thus facilitating fitting and estimation of multiple parameter models.

1.5. A few challenges associated with SEM

Implementation of a hierarchical SEM in the context of animal agriculture is not without challenges. Agricultural data often display a well-defined hierarchical structure, defined as data structure or data architecture (Stroup 2013). For example, in feedlot production, animals are regularly managed in clusters defined by pens or cohorts, and they are more likely to share a common background. Besides, pens or cohorts of animals may be subjected to common management or business practices defined at the level of the commercial operation. Hence, this clustered structure indicates that observations at the level of animal are not mutually independent but likely to have correlation patterns. Another source of data architecture is that of design components which are prevalent in agricultural data include blocking, nested effects or subsampling, among others. The data architecture of designed experiments and observational epidemiological studies can be accounted for using mixed models (Littell et al. 2006, Milliken

and Johnson 2009, Gbur 2012, Stroup 2013, Dohoo et al. 2014). During the past few decades, a considerable amount of research has been devoted to issues of hierarchical data structure in data analysis (Littell et al. 2006, Milliken and Johnson 2009, Gbur et al. 2012, Stroup 2013). Therefore, these aspects of data architecture in the arena of SEM for designed experiments and observational studies have not been comprehensively studied.

Data architecture can pose peculiar challenges in the context of SEM, as correlations are apparent not only between clustered observations of an outcome (i.e. due to experimental design), but also as correlations between outcomes (i.e. due to the multiple-trait model). Further, these correlations can take multiple levels, namely the level of observations (i.e. residuals) and level of clustering or other design components (i.e. random effects). Therefore, correlations might be expected between observations from individual animals within pens or between pens within an operation, and also between outcomes at both animal and pen levels (Dohoo, Martin, and Stryhn 2014).

Another challenge of SEM implementation is the key assumption of homogeneous network (Gianola and Sorensen 2004), whereby the structural coefficients defining functional links in a network are assumed homogeneous and impervious to environmental conditions or management factors. This assumption seems especially questionable in the context of animal agriculture as functionality in the system may differ between subpopulations and is often subjected to explicit interventions to optimize the necessary trade-offs between outcomes. For example, early work by Wu et al. (2007) showed that direct effects obtained from a simultaneous and recursive modeling framework to describe the associations between milk yield and udder health in dairy cows were yield-dependent and changed during a lactation period. Therefore,

methodological extensions to SEM are guarantees to accommodate potential differences in functionality in a production system.

1.6. Rational and significance

The study of causal networks and associated modeling strategies, such as SEM, can provide insightful information in helping us understand 'how' and 'why' intermediate production outcomes affect each other, as well as their overall effect on system output. An enhanced understanding of the inner workings of a system can then be used in decision making and in the development and implementation of interventions for effective and efficient management. Inferring a causal network may be thought of as designing a blueprint that explicitly describes how outcome variables are connected, structured and organized in a system: what are the causes, how does information flow downstream and/or upstream, and what are the effects? These questions can be addressed in the context of designed experiments and observational studies in animal agriculture using SEM adapted to a mixed models framework (Gianola and Sorensen 2004).

1.7. Objectives specific goals of this dissertation

The specific objectives of this dissertation are 1) to evaluate the inferential implications of mispecified data architecture in a hierarchical SEM framework. Research Chapter 2 in this dissertation addresses this specific goal. Further, 2) to investigate potential causal biological relationships between reproductive performance traits in high-producing gilts and sows in the context of data from a designed experiment. Research Chapter 3 in this dissertation provides details of achieving this specific goal. Lastly, 3) to develop methodological extensions to hierarchical SEM to accommodate explicit specification of heterogeneous structural coefficients

as a function of systematic and non-systematic sources of variation. Research Chapter 4 in this dissertation addresses this objective.

REFERENCES

Bello, N. M., Steibel, J. P., and Tempelman, R. J. 2010. "Hierarchical Bayesian Modeling of Random and Residual Variance–covariance Matrices in Bivariate Mixed Effects Models," *Biometrical Journal*, 52(3), 297-313.

de los Campos, G., Gianola, D., Boettcher, P., and Moroni, P. 2006a. "A Structural Equation Model for Describing Relationships between Somatic Cell Score and Milk Yield in Dairy Goats," *Journal of Animal Science*, 84, 2934-2941.

de los Campos, G., Gianola, D., and Heringstad, B. 2006b. "A Structural Equation Model for Describing Relationships between Somatic Cell Score and Milk Yield in First-lactation Dairy Cows," *Journal of Dairy Science*, 89(11), 4445-4455.

de Maturana, E. L., Wu, X. L., Gianola, D., Weigel, K. A., and Rosa, G. J. M. 2009. "Exploring Biological Relationships between Calving Traits in Primiparous Cattle with a Bayesian Recursive Model," *Genetics*, 181, 277-287.

de Maturana, E. L., de los Campos, G., Wu, X. L., Gianola, D., Weigel, K. A., and Rosa, G. L. M. 2010. "Modeling Relationships between Calving Traits: A Comparison between Standard and Recursive Mixed Models," *Genetics Selection Evolution*, 42, 1.

Dohoo, I., Martin, W., and Stryhn, H. 2014. *Veterinary Epidemiologic Research* (2nd ed.), Canada: VER Inc.

Foley, J. A. 2011. "Can We Feed the World and Sustain the Planet? A five-step Global Plan Could Double Food Production by 2050 While Greatly Reducing Environmental Damage," *Science American*, 24, 84-89.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, S. F., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K.. 2005. "Global Consequences of Land Use," *Science*, 309(5734), 570.

Gbur, E. E., Stroup, W. W., McCarter, K. S., Durham, S., Young, L. J., Christman, M., West, M., and Kramer, M. 2012. *Analysis of Generalized Linear Mixed Models in the Agricultural and Natural Resources Sciences*, Madison, WI, USA: American Society of Agronomy, Soil Science Society of America, Crop Science Society of America, Inc.

Gianola, D., and Sorensen, D. 2004. "Quantitative Genetic Models for Describing Simultaneous and Recursive Relationships between Phenotypes," *Genetics*, 167(3), 1407-1424.

Gilks, W. R. , Richardson, S., and Spiegelhalter, D. J. 1996. *Makov Chain Mente Carlo in Practice*, London, U.K.: Chapman & Hall.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. 2010. "Food Security: The Challenge of Feeding 9 Billion People," *Science*, 327(5967), 812.

Haavelmo, T. 1943. "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11(1), 1-12.

Henderson, C. R., and Quaas, R. L. 1976. "Multiple Trait Evaluation Using Relatives Records," *Journal of Animal Science*, 43(6), 1188-1197.

Inoue, K., Valente, B. D., Shoji, N., Honda, T., Oyama, K., and Rosa, G. J. M. 2016. "Inferring Phenotypic Causal Structures among Meat Quality Traits and the Application of a Structural Equation Model in Japanese Black Cattle," *Journal of Animal Science*, 94(10), 4133-4142.

Konig, S., Wu, X. L., Gianola, D., Heringstad, B., and Simianer, H. 2008. "Exploration of Relationships between Claw Disorders and Milk Yield in Holstein Cows via Recursive Linear and Threshold Models," *Journal of Dairy Science*, 91(1), 395-406.

Littell, R. C., Milliken, G. A., Stroup, W. W., Russell, W. D., and Oliver, S. 2006. *SAS for Mixed Models* (2nd ed.), Cary, NC: SAS Institute Inc.

Milliken, G. A., and Johnson, D. E. 2009. *Analysis of Messy Data - Volume 1: Designed Experiments* (2nd ed.), Boca Raton, Florida, USA: Chapman and Hall/CRC Press.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inference* (2nd ed.), USA: Cambridge University Press.

Peñagarican, F., Valente, B. D., Steibel, J. P., Bates, R. O., Ernst, C. W., Khatib, H., and Rosa, G. J. M. 2015. "Searching for Causal Networks Involving Latent Variables in Complex Traits: Application to Growth, Carcass, and Meat Quality Traits in Pigs," *Journal of Animal Science*, 93(10), 4617-4623.

Robins, J. M. 1999. "Association, Causation, and Marginal Structural Models," *Synthese*, 121(1/2), 151-179.

Rosa, G. J. M., Valente, B. D., de los Campos, G., Wu, X. L., Gianola, D., and Silva, M. A. 2011. "Inferring Causal Phenotype Networks Using Structural Equation Models," *Genetics Selection Evolution*, 43, 13.

Sorensen, D., and Gianola, D. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*, New Your: Springer-Verlag.

Stroup, W. W. 2013. *Generalized Linear Mixed Models*, Boca Raton, FL: CRC Press Taylor & Francis Group.

Valente, B. D., Rosa, G. J. M., Silva, M. A., Teixeira, R. B., and Torres, R. A. 2011. "Searching for Phenotypic Causal Networks Involving Complex Traits: An Application to European Quail," *Genetics Selection Evolution*, 43, 37.

Valente, B. D., Rosa, G. J. M., de Los Campos, G., Gianola, D., and Silva, M. A. 2010. "Searching for Recursive Causal Structures in Multivariate Quantitative Genetics Mixed Models," *Genetics*, 185(2) ,633-644.

Van Vleck, L. D., and Edlin, K. M. 1984. "Multiple Trait Evaluation of Bulls for Calving Ease," *Journal of Dairy Science*, 67(12), 3025-3033.

Varona, L., and Sorensen, D. 2014. "Joint Analysis of Binomial and Continuous Traits with a Recursive Model: A Case Study Using Mortality and Litter Size of Pigs," *Genetics*, 196(3), 643-651.

Verma, T., and Pearl, J. 1991. "Equivalence and Synthesis of Causal Models," *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, Elsevier Science Inc.

Verma, T., and Pearl, J. 1992. "An Algorithm for Deciding If a Set of Observed Independencies Has a Causal Explanation," *Uncertainty in Artificial Intelligence*, Elsevier Science Inc.

Wright, S. 1934. "An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs," *Genetics*, 19, 506-536.

Wu, X. L., Heringstad, B., Chang, Y. M., de Los Campos, G., and Gianola, D. 2007. "Inferring Relationships between Somatic Cell Score and Milk Yield Using Simultaneous and Recursive Models," *Journal of Dairy Science*, 90(7), 3508-3521.

Wu, X. L., Heringstad, B., and Gianola, D. 2010. "Bayesian Structural Equation Models for Inferring Relationships between Phenotypes: A Review of Methodology, Identifiability, and Applications," *Journal of Animal Breeding and Genetics*, 127(1), 3-15.

Chapter 2 - Accounting for Data Architecture on Structural Equation Modeling of Feedlot Cattle Performance

ABSTRACT

Structural Equation Models (SEM) are a type of multi-trait model increasingly being used for inferring functional relationships between multiple outcomes using operational data from livestock production systems. These data often present a hierarchical architecture given by clustering of observations at multiple levels including animals, cohorts and farms. A hierarchical data architecture introduces correlation patterns that, if ignored, can have detrimental effects on parameter estimation and inference. Here, we evaluate the inferential implications of accounting for, or conversely, mispecifying data architecture in the context of SEM. Motivated by beef cattle feedlot data, we designed simulation scenarios consisting of multiple responses in a clustered architecture. Competing fitted SEMs differed in their model specification so that data architecture was explicitly accounted for (M1; true model) or mispecified due to disregarding either the cluster-level correlation between responses (M2) or the correlation between observations of a response within a cluster (M3), or ignored all together (M4). Model fit was increasingly impaired when data architecture was mispecified or ignored. Both accuracy and precision of estimation were also negatively affected when data architecture was disregarded. Our findings are further illustrated using data from feedlot operations from the US Great Plains. Standing statistical recommendations that call for proper model specification capturing relevant hierarchical levels in data structure extend to the multivariate context of structural equation modeling.

Key Words: Hierarchical modeling; Multi-level correlation; Structural equation models; Beef cattle.

2.1. INTRODUCTION

Understanding of functional interrelationships between multiple response variables is becoming increasingly important for efficient management of agricultural production systems. Classical multiple-trait models (MTM) can be used to simultaneously evaluate the joint behavior of multiple responses and study their probabilistic relationships expressed as correlations and covariances (Johnson and Wichern 2007) or functions thereof (Bello et al. 2010). However, classical MTM cannot assess directionality of correlations neither discriminate the plausible causal mechanisms generating such correlations (Rosa et al. 2011). This limitation of classical MTM hinders their use for exploring causal relationships between the multiple responses involved in complex systems.

Structural Equation Models (SEM) (Haavelmo 1943) are a special type of multivariate models that, under certain assumptions (Pearl 2009), can recognize directional causality in the association between responses, thereby enabling the study of functional networks in a systems approach. Following from a long tradition in biology, sociometrics and econometrics, SEM date back to path analysis (Wright 1934, Duncan 1966) and simultaneous equation systems (Haavelmo 1943). Additional developments adapted the SEM framework to a multivariate linear context (Joreskog 1973), introduced ideas from mixed models methodology (Gianola and Sorensen 2004) and graphical models (Lauritzen 1996), and incorporated a causality focus (Pearl 2009).

Briefly, SEM introduce structural coefficient parameters that allow for one or more response variables to also serve as predictors of other responses in the model in a non-cyclic manner. Functional links between responses can then be interpreted from structural coefficients in a linear regression-type fashion. Fitting a SEM requires specification of a causal structure, as

represented by a directed acyclic graph (DAG) (Pearl 2009) characterized by nodes (i.e. variables) connected by directed edges (i.e. functional links). It is only recently that SEMs started to be utilized in the context of animal agriculture, most notably in applications of quantitative genetics and genomics. Examples include dairy goats (de los Campos et al. 2006), dairy cattle (Konig et al. 2008, Lopez de Maturana et al. 2009), quails (Valente et al. 2011), swine (Varona et al. 2014, Peñagaricano et al. 2015), and beef cattle (Inoue et al. 2016).

By contrast, applications of mixed-model SEM in the realm of designed agricultural experiments and observational studies in veterinary epidemiology have been limited. Data from these domains usually exhibit a well-defined hierarchical structure, also referred to as data architecture (Stroup 2013). Multi-level data architecture often includes animal, cohort and farm levels, which are arranged in some design structure (e.g. blocks, nested effects, subsamples). Particularly common to these agricultural applications is a clustered data structure (Dohoo, Martin, and Stryhn 2014), by which observations at the animal level share a common background and are thus not mutually independent, but rather correlated within pens or cohorts; similarly, pens or cohorts within an operation or farm can also be expected to yield correlated responses.

In this article, we explore the consequences of accounting for (or ignoring) existent correlation structure due to data architecture in the specific context of SEMs. This topic has been extensively studied in single-response mixed models for designed experiments (Littell et al. 2006, Milliken and Johnson 2009, Stroup 2013) and epidemiological studies (Dohoo, Martin, and Stryhn 2014), but little has permeated to the multivariate arena of hierarchical SEMs. Therefore, our objective in this study was to evaluate the inferential implications of mispecified data architecture in a SEM framework. We used a simulation approach, followed by a data

application in the area of animal production, specifically commercial beef cattle feedlots. Section 2 describes our methodological approach, including a brief introduction to SEMs in the context of mixed models and our hierarchical Bayesian implementation. The design of the simulation study and the data application are also described in Section 2. We present results in Section 3, provide further discussion in Section 4 and outline conclusions in Section 5.

2.2. METHODS

2.2.1. Structural equation models

Consider the DAG in *Figure 1* depicting putative causal relationships between response variables y_j (j = 1, 2, …, k) for k = 4. Here, y_1 has a direct effect on y_2 and on y_3. Also, y_2 directly affects y_3 and y_4. Also from Figure 2.1, it is apparent that each response variable y_j is directly affected by a corresponding cluster-specific effect $\boldsymbol{u}_j$, assumed mutually correlated among clusters, and by a corresponding residual $\boldsymbol{e}_j$. Implicit in Figure 2.1 are response-specific location parameters for systematic effects on the means, as these are typically not included in a DAG. Following Gianola and Sorensen (2004), a mixed-models SEM representation of the DAG in Figure 2.1 is

$$\boldsymbol{y}_i = \boldsymbol{\Lambda}\boldsymbol{y}_i + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u} + \boldsymbol{e}_i\,; \qquad i = 1,2,\ldots,n \tag{2.1}$$

where $\boldsymbol{y}_i' = [y_{i1}\ \ y_{i2}\ \ y_{i3}\ \ y_{i4}]'$ is a vector of response variables $y_1,\ y_2,\ y_3$ and y_4 observed on the i^{th} subject, and $\boldsymbol{\Lambda}$ is a $k \times k$ matrix of zeroes and unknown structural coefficient parameters $\lambda_{jj'}$ that describe the direct effect of variable j on variable j'. Here,

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & \lambda_{32} & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 \end{bmatrix} \tag{2.2}$$

For instance, here, λ_{21} represents the expected change in y_2 per unit increase in y_1 and represents the direct effect of y_1 on y_2, as in Figure 2.1. In Equation (2.1), $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1' \ \boldsymbol{\beta}_2' \ \boldsymbol{\beta}_3' \ \boldsymbol{\beta}_4']$ is a vector of unknown "fixed-effect" location parameters that, when associated with factors and covariates (e.g. treatment, demographics) in $\boldsymbol{X}_i = diag(\boldsymbol{x}_{i1}' \ \boldsymbol{x}_{i2}' \ \boldsymbol{x}_{i3}' \ \boldsymbol{x}_{i4}')$ unique to each subject *i*, define response-specific means implicit in Figure 2.1. Next, $\boldsymbol{u}' = [\boldsymbol{u}_1' \ \boldsymbol{u}_2' \ \boldsymbol{u}_3' \ \boldsymbol{u}_4']$ is a vector of unknown random effects associated with clustering factors or other components of data architecture in $\boldsymbol{Z}_i = diag(\boldsymbol{z}_{i1}' \ \boldsymbol{z}_{i2}' \ \boldsymbol{z}_{i3}' \ \boldsymbol{z}_{i4}')$. Random effects in $\boldsymbol{u}$ are assumed multivariate normally distributed with null mean and covariance matrix $\boldsymbol{G}$, such that:

$$\boldsymbol{u} \sim MVN\left(\boldsymbol{0},\ \boldsymbol{G} \otimes \boldsymbol{I}_q = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_{12}} & \sigma_{u_{13}} & \sigma_{u_{14}} \\ \sigma_{u_{21}} & \sigma_{u_2}^2 & \sigma_{u_{23}} & \sigma_{u_{24}} \\ \sigma_{u_{31}} & \sigma_{u_{32}} & \sigma_{u_3}^2 & \sigma_{u_{34}} \\ \sigma_{u_{41}} & \sigma_{u_{42}} & \sigma_{u_{43}} & \sigma_{u_4}^2 \end{bmatrix} \otimes \boldsymbol{I}_q\right) \tag{2.3}$$

where *q* is the number of levels of the random effect factor. Finally, $\boldsymbol{e_i'} = [e_{i1} \ e_{i2} \ e_{i3} \ e_{i4}]$ is the set of residuals for the i^{th} subject assumed multivariate normally distributed with null mean and a diagonal residual covariance matrix $\boldsymbol{R}$, such that:

$$\boldsymbol{e}_i \sim MVN\left(\boldsymbol{0}, \boldsymbol{R} = \begin{bmatrix} \sigma_{e_1}^2 & 0 & 0 & 0 \\ 0 & \sigma_{e_2}^2 & 0 & 0 \\ 0 & 0 & \sigma_{e_3}^2 & 0 \\ 0 & 0 & 0 & \sigma_{e_4}^2 \end{bmatrix}\right) \tag{2.4}$$

The assumption of a diagonal $\boldsymbol{R}$ (i.e. all $\sigma_{e_{jj'}} = 0$) is a standard reparameterization in the specification of SEM to ensure parameter identifiability (Gianola and Sorensen 2004, Wu et al. 2007). This constraint implies that the only additional source of covariance (beyond the random effects) are the functional relationships between the response variables, as captured by structural coefficients $\lambda_{jj'}$.

2.2.2. Simulation study

Motivated by our data application (see later), we designed plausible simulation scenarios in which observations were generated for a set of $k = 4$ response variables that assumed a clustered data architecture as in Figure 2.1and expressed as in Equation (2.1). Specifically, simulation scenarios consisted of four combinations of (i) two cluster sizes and (ii) two relative magnitudes of cluster-level variability. Cluster sizes were either balanced (i.e. observations available on 10 subjects for each of 100 clusters) or, more realistically, unbalanced for 100 clusters with size determined by a random draw from a discrete uniform distribution bounded between 1 and 20. Scenarios of balanced and unbalanced cluster sizes were designed to ensure a similar number of subjects. In addition, scenarios for relative magnitudes of cluster-level variability were defined by moderate (i.e. ~10%) or large (i.e. ~30% to 50%) intraclass correlation (ICC), defined for each j^{th} response as $ICC_j = \sigma_{u_j}^2 / \left(\sigma_{u_j}^2 + \sigma_{e_j}^2\right)$.

For each simulation scenario, 10 datasets were generated, thus yielding a total of 40 simulated datasets. In all cases, the choice of parameter values used for data simulation reflected those of the data application. Specifically, we simulated location parameters in $\boldsymbol{\beta}$ akin to fixed effects with 2 levels using a set-to-zero parameterization (Miliken and Johnson 2009) such that $\boldsymbol{\beta}_1' = [100 \;\; -30]$, $\boldsymbol{\beta}_2' = [30 \;\; -10], \boldsymbol{\beta}_3' = [81 \;\; -20]$, $\boldsymbol{\beta}_4' = [100 \;\; -25]$, whereby in each case the first vector element indicates the mean response for one factor level, and the second vector element indicates the differential effect of the second factor level. Further, we specified $\boldsymbol{R} = diag(900 \;\; 36 \;\; 81 \;\; 100)$, $\lambda_{21} = 0.5$, $\lambda_{31} = 0.2$, $\lambda_{32} = -0.4$, and $\lambda_{42} = 0.3$. The cluster-level correlations between responses were specified as:

$$\boldsymbol{\rho_u} = \begin{bmatrix} 1 & -0.7 & 0.4 & 0.1 \\ -0.7 & 1 & -0.5 & -0.2 \\ 0.4 & -0.5 & 1 & 0 \\ 0.1 & -0.2 & 0 & 1 \end{bmatrix}$$

Within each scenario of relative magnitudes of cluster-level variability, the (co)variance matrix $\boldsymbol{G}$ was obtained by pre- and post-multiplying matrix $\boldsymbol{\rho_u}$ by a diagonal matrix containing the square root of the corresponding variance parameters $\sigma^2_{u_j}$ (j = 1, …,4), whereby the values used for $\sigma^2_{u_j}$ were defined by the scenarios of moderate and large relative magnitude of the cluster-level variability, as described above.

2.2.3. Alternative model specifications

To evaluate the inferential implication of mispecified data architecture, we designed alternative specifications of an SEM. Model 1 (M1) was specified as in Equation (2.1), thus exactly reflecting the true model used for data generation whereby data architecture was fully and explicitly accounted for by way of a cluster-level covariance matrix $\boldsymbol{G}$ analogous to that in Equation (2.3). As such, M1 recognized cluster-level correlations between responses (i.e. functions of off-diagonal elements of $\boldsymbol{G}$), as well as within-cluster correlations between observations of a given response (i.e. functions of the corresponding j^{th} diagonal elements in $\boldsymbol{G}$ and $\boldsymbol{R}$, as expressed by ICC_j). The other SEMs mispecified one or more aspects of the data architecture. Specifically, model 2 (M2) followed from Bae et al. (2016) and disregarded any cluster-level correlations between responses by explicitly setting all off-diagonal elements of $\boldsymbol{G}$ to zero, thereby incorrectly assuming that the responses were mutually independent at the level of clusters. Model 3 (M3) treated clusters $\boldsymbol{u}$ as fixed effects, thereby effectively ignoring any correlation between observations of a response within a cluster. We note that, implicitly, M3 made the same assumption as M2 and ignored cluster-level correlations between responses. Lastly, Model 4 (M4) did not specify any cluster effects, either as fixed or random effects, on any of the responses, thus completely ignoring data architecture. Here, M2, M3, and M4

represent models with increasingly mispecified data architecture in the context of SEMs. Alternative SEM specifications M1 – M4 were fitted to each of the 40 simulated datasets.

2.2.4. Application to feedlot data

Data consisted of observations collected on 1430 lots of cattle representing a total of 178,983 animals from 9 feedlots located across the US Great Plains during 2014 and 2015. A lot can be loosely defined as a cohort of cattle that are managed similarly and concomitantly in a feedlot operation. For each lot, $k = 3$ response variables were observed, namely arrival weight (AW; kg per head), Bovine Respiratory Disease-related treatment costs (BRD$; $ per head) and average daily body weight gain (ADG; g per head per day). BRD$ and ADG were selected as indicators of health and production performance, respectively. Table 2.1 shows raw descriptive statistics for each response variable. A complete description of the data is available in Cha et al. (2017).

Of the lots observed, 544 consisted of all steers, whereas 886 consisted of all heifers; lots with mixed genders were excluded from the data. Contemporary groups were defined as the combination of month and year of arrival to a given feedlot, thereby defining clusters of animals managed concurrently within a feedlot operation. The data consisted of 110 contemporary groups, ranging in size from 1 to 57 lots. In turn, lots varied in size from 20 to 607 head, with approximately 50% of the lots consisting of at least 100 head.

Using expert knowledge of the epidemiology of BRD in feedlot cattle, we elicited the DAG in Figure 2.2 to represent a hypothesized causal process between health and performance outcomes in feedlot cattle. Specifically, we hypothesized a direct effect of AW on health (i.e., AW → BRD$) and of AW on production performance (i.e., AW → ADG) based on Sanderson, Dargatz, and Wagner (2008) and Hay et al. (2014). Further, we hypothesized a direct effect of

health on production performance (i.e., BRD$ → ADG) (Cernicchiaro 2013, Wittum1996, Babcock 2009). This DAG arrangement is further consistent with the temporality of the outcomes considered in this study.

Alternative SEM specifications M1-M4 were fitted to the data. In all cases, the linear predictor for each response variable included a fixed effect of gender in $\boldsymbol{\beta}$. Clustering effects of contemporary group were specified differentially for each model M1 – M4, as described in a previous section, so that data architecture due to lots clustered within contemporary group would be explicitly accommodated (M1), mispecified (M2, M3), or ignored all together (M4).

2.2.5. Hierarchical Bayesian implementation and posterior inference

Alternative SEMs M1 - M4 were fitted in a hierarchical Bayesian framework implemented with Markov Chain Monte Carlo (MCMC). In general terms, the joint posterior distribution of all unknowns is then

$$p(\boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{G}, \boldsymbol{R}, \boldsymbol{\Lambda}, |\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{u}, \boldsymbol{R}, \boldsymbol{\Lambda}) p(\boldsymbol{\beta}) p(\boldsymbol{u}|\boldsymbol{G}) p(\boldsymbol{G}) p(\boldsymbol{R}) p(\boldsymbol{\Lambda}) \quad (2.6)$$

We selected conditionally conjugate prior distributions to facilitate recognition of fully conditional densities in closed form, thereby implementing Gibbs sampling. For all M1-M4, prior distributions for structural coefficients $\lambda_{jj'}$ and for fixed effects location parameters $\boldsymbol{\beta}_j$ were assumed flat such that $p(\lambda_{jj'}) \propto$ constant and $p(\boldsymbol{\beta}_j) \propto$ constant for all j and j'. Our estimation approach implemented a set-to-zero restricted parameterization on the fixed effects (Milliken and Johnson 2009), to ensure parameter identifiability and to facilitate scalability of mixed-models SEM to multiple-factorial treatment structures in $\boldsymbol{X}_i\boldsymbol{\beta}$. Also for M1-M4, residual variance parameters $\sigma^2_{e_j}$along the diagonal of $\boldsymbol{R}$ were each specified a flat, though improper, scaled-inverse chi-square prior distribution with degrees of freedom v_{e_j}= -1 and scale parameter

$s_{e_j}^2$= 0. This prior can be shown to be equivalent to $\sqrt{\sigma_{e_j}^2} \sim U(0, A)$, for any finite but sufficiently large value of *A*, as recommended for variance components by Gelman (2006).

Priors on $\boldsymbol{u}$ and $\boldsymbol{G}$ differed among M1-M4 to reflect alternative SEM specifications of data architecture. For M1 and M2, we assigned a structural prior on $\boldsymbol{u}$ such that $p(\boldsymbol{u}|\boldsymbol{G}) \sim N(\boldsymbol{0}, \boldsymbol{G})$ to allow for recovery and borrowing of information across clusters, as is standard for random effects in a mixed models framework (Robinson 1991). Models M1 and M2 differed in that the latter explicitly specified all cluster-level covariances (i.e. off-diagonal elements of $\boldsymbol{G}$) to equate to zero. In M1 we implemented a flat, improper inverse-Wishart (IW) prior on $\boldsymbol{G}$ (i.e. specified to have degrees of freedom v_G = -(*k*+1) and a $k \times k$ scale matrix of zeroes). In turn, in M2 we assigned a univariate flat, though improper, prior to each diagonal element of $\boldsymbol{G}$, similar to the priors indicated above for residual variances following Gelman (2006). For M3, the clustering effect $\boldsymbol{u}$ was assumed to have a subjective flat prior distribution (i.e. $p(\boldsymbol{u}) \propto$ constant), thereby effectively specifying cluster as equivalent to a frequentist fixed effect (Sorensen and Gianola 2002). This specification in M3 effectively ignores any between-cluster correlation between responses (i.e. analogous to M2) as well as any within-cluster correlation between observations of a response. Lastly, M4 effectively removed $\boldsymbol{u}$ from the model, thus leaving clustering effects unaccounted for.

Data simulation and MCMC implementation was programmed using R software (R Development Core Team 2016). For each of the simulated data sets, each of the SEM specifications was fitted using a single MCMC chain run for 20,000 iterations after a burn-in period of 5000 cycles; one of every 5 samples was saved and used for posterior inference. For the data application, a single MCMC chain of 155,000 iterations was run for each alternative SEM. The initial 5000 iterations of the chain were discarded as burn-in and inference was based

on posterior samples thinned every third iteration, yielding 50,000 MCMC samples for posterior inference. In all cases, convergence diagnostics were implemented using the R package CODA (Plummer et al. 2006). Specifically, parameters $\lambda_{jj'}, \beta_{1,j}, \beta_{2,j}, \sigma^2_{e_j}, \sigma^2_{u_j}, \sigma_{u_{jj'}}$ were monitored for convergence using trace plots and autocorrelation plots, as well as the diagnostic approach proposed by Raftery and Lewis (1992). Also, effective sample size (ESS) was estimated to assess the number of effectively independent samples among the autocorrelated MCMC samples for each parameter indicated above (Sorensen et al. 1995). Length of MCMC chains was tuned to ensure that ESS was greater than 100 for all parameters. Posterior distributions are summarized using posterior means and 95% highest posterior density (HPD) intervals.

2.2.6. Model comparison

Data fit of SEM specifications M2, M3 and M4 were compared to that of the true data generating model (M1) using the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002). Better fitting models are characterized by smaller DIC values and differences of 7 or greater are considered as an arbitrarily-selected threshold indicative of improved model fit.

In the simulation study, accuracy of parameter estimation under SEM models M1 – M4 was assessed based on coverage, defined as how often the 95% HPD interval of a parameter contained the true value used for simulation. Thus, this approach assumed a 95% estimation accuracy when the estimation process was working appropriately. We further assessed precision of parameter estimation by evaluating the width of the 95% HPD interval.

2.3. RESULTS

2.3.1. Simulation study

Figure 2.3 illustrates global model fit of alternative SEM specifications M2, M3 and M4, expressed as DIC differences relative to the true data-generating model (M1), for scenarios of

cluster-level variability of moderate and large relative magnitude under conditions of balanced cluster size. Overall, DIC differences relative to M1 were above the threshold of 7 points specified to declare differences in model fit, thereby indicating substantially impaired global fit of M2, M3 and M4 in a manner directly proportional to the severity of misspecification of data architecture. Even the presumably mild assumption of ignored cluster-level correlations between responses (M2: all $\sigma_{u_{jj'}}$ set to zero) was detrimental to model fit in most cases evaluated. An additional impairment in global fit (i.e. increase in DIC difference relative to M1) was apparent when the effect of clusters was incorrectly fitted as a fixed effect (M3) as opposed to a random effect (M2), thereby ignoring any within-cluster correlation between observations of a response. Also, the relative magnitude of cluster-level variances was as a key modifier of global fit when data architecture was misspecified. When cluster-level variances were large, SEMs that ignored data architecture (M4) produced an even greater loss in global model fit related to scenarios with moderate cluster-level variances. Results for simulations scenarios of unbalanced cluster size (Appendix A - Figure A.1) were similar to those described here for balanced clusters and are thus not discussed further.

Next, we evaluate the effect of mispecified data architecture on the bias and variance of estimators of SEM parameters. Specifically, Figure 2.4 illustrates overall accuracy of parameter estimation, expressed as percent coverage of 95% HPD intervals, of alternative SEM specifications under simulation scenarios of balanced cluster size. Table 2.2 further details coverage for different types of parameters, namely $\beta_{1,j},\ \beta_{2,j},\ \lambda_{jj'},\ \sigma^2_{e_j},\ \sigma^2_{u_j},\ \sigma_{u_{jj'}}$ and Table 2.3 shows their estimation precision expressed as width of their 95% HPD intervals. As can be expected, estimation accuracy under a model with properly specified data architecture (M1) was close to the nominal 95% supported by the HPD interval and coverage was on target both overall

(Figure 2.4) and separately for each parameter type (Table 2.2), random variability notwithstanding. In turn, as SEMs increasingly mispecified data architecture (i.e. M2 < M3 < M4), estimation accuracy progressively declined relative to M1 (Figure 2.4). Even a presumably mild mispecification of data architecture, as in M2 (i.e. all $\sigma_{u_{jj'}}$ set to zero), proved detrimental to estimation accuracy, specifically for the structural coefficients $\lambda_{jj'}$, as indicated by a drop in coverage relative to M1 (Table 2.2). In particular, M2 seemed to underestimate structural coefficients $\lambda_{jj'}$, therefore distorting the functional links between outcomes. Yet, changes in estimation precision of $\lambda_{jj'}$ were relatively minor, with slightly narrower HPD intervals under M2 compared to M1 (Table 2.3). For cluster-level variances $\sigma^2_{u_j}$, estimation accuracy under M2 was on target but precision seemed inflated relative to M1, as indicated by narrower HPD intervals (Table2.3).

When within-cluster correlations between observations were disregarded (M3), the coverage of fixed-effect location parameters $\beta_{1,j}$ was only ~28 % and ~30% for moderate and large cluster-level variances, respectively, thereby indicating estimation bias (Table 2.2). This loss in accuracy may be partially explained by artifactually high estimation precision under M3, as indicated by narrower 95% HPD intervals relative to M1 and M2 (Table 2.3). Conversely, greater estimation accuracy (i.e. less bias) for location parameters $\beta_{2,j}$ under M3 reflected larger estimation variance and thus, lower precision (Tables 2.2 and 2.3).

If data architecture was ignored and observations were assumed mutually independent (i.e. without any clustering effects, as in M4), losses in estimation accuracy spanned across all parameter types and were most dramatic for scenarios of large cluster-level variance (Table 2.2). Especially estimates of residual variances $\sigma^2_{e_j}$ seemed to be consistently biased downwards

(Table 2.2) despite lower estimation precision reflected in correspondingly broader 95% HPD intervals (Table 2.3).

Taken together, simulation results support the expected trade-off between estimation bias (i.e. accuracy) and estimation variance (i.e. precision) for SEM parameters. A similar pattern of results was observed for simulation scenarios of unbalanced cluster size and when cluster-level variability was of large relative magnitude (Appendix A - Figure A.1 and A.2 and Table A.2 – A.4).

2.3.2. Application to feedlot data

We further illustrate the findings from the simulation study using data from commercial feedlot cattle as a proof-of-concept application. In order to align the specific characteristics of the data application to the simulation scenarios considered, we first obtained variance component estimates for contemporary groups (cluster-level variance) and for lots (residual-level variance) and used them to evaluate the corresponding ICC for each response variable. The estimated ICC were approximately 62%, 57% and 14% for the responses AW, BRD$ and ADG, respectively, thereby indicating large relative magnitude of the cluster-level variance for AW and BRD$, and a moderate magnitude for ADG.

Table 2.4 shows DIC of alternative SEMs M1 – M4 fitted to the feedlot data. M1 and M2 had substantially smaller DIC than M3 and M4, thereby supporting specification of contemporary group as a random effect into the SEM to account for clustered data architecture. We further note a small DIC difference between M1 and M2 (~4 points; Table 2.4), the magnitude of which is inconclusive to support selection of one SEM as better fitting over the other. This is consistent with posterior inference on cluster-level covariances between AW, BRD$ and ADG under M1, namely $\sigma_{u_{BRD\$,AW}}$, $\sigma_{u_{ADG,AW}}$ and $\sigma_{u_{ADG,BRD\$}}$, whereby the

corresponding 95% HPD intervals included zero (Table 2.5). Recall that such was the assumption made by M2 (i.e. $\sigma_{u_{jj'}} = 0$). Thus, it is not surprising that the DIC difference between M1 and M2 is fully explained by their differential penalty for model complexity, expressed as effective number of parameters pD (Table 2.4).

Table 2.5 summarizes the posterior distribution of selected parameters of interest based on SEM specifications M1 – M4 fitted to the feedlot data. As with any application to real data, the true state of nature is unknown so it is not possible to ascertain accuracy of parameter estimation. However, we did notice differences in point estimates (i.e. posterior means) between M1-M4, and these differences were consistent with findings from the simulation study. Specifically, estimates of $\lambda_{BRD\$,AW}$ based on M4 were slightly smaller than those obtained by fitting the other models. Furthermore, M4-, and sometimes M3-, based estimates of $\beta_{1,j}$ and $\beta_{2,j}$ for AW, BRD$ and ADG seemed to differ from those of M1 and M2. Similarly, M4-based estimates of $\sigma^2_{e_j}$ were typically larger than those of other SEM specifications, consistent with simulations results.

Table 2.6 shows precision of parameter estimates on the feedlot dataset, expressed as width of the 95% HPD interval under M1 – M4. Again, these results were consistent with those from our simulation study and suggested a bias-variance trade-off in the estimation process. Most notably, estimation precision for location parameters of AW and BRD$ (i.e. $\beta_{1,j}$, $\beta_{2,j}$) was decreased (i.e. broader 95% HPD intervals) when contemporary groups were fitted as fixed effects (M3) as opposed to random effects (M1 and M2), thus disregarding within-cluster correlations between observations. Also consistent with results from the simulation study, estimation precision of residual variances $\sigma^2_{e_j}$ for AW, BRD$ and ADG was decreased in M4 relative to other SEMs, as indicated by wider 95% HPD intervals.

We chose SEM M1 for final inference on functional links between AW, BRD$ and ADG because, as previously discussed, M1 had the most general specification (i.e. no restrictive assumptions made on cluster-level covariance parameters $\sigma_{u_{jj'}}$) among alternative SEMs with comparable DIC-based model fit. We focus inference on the direct effects of AW on BRD$, of BRD$ on ADG and of AW on ADG, as characterized by structural coefficients $\lambda_{BRD\$,AW}$, $\lambda_{ADG,BRD\$}$ and $\lambda_{ADG,AW}$, respectively (Table 2.5). Specifically, every AW increase of 1 kg/head was estimated to have a positive direct effect on ADG of 0.0020 kg/head*day with 95% HPD interval [0.0017, 0.0022], and also a favorable effect on health by reducing BRD$ (i.e. -0.0785 USD per head with 95% HPD interval [-0.0904, -0.0667]). On average, worsening BRD-related health, as indicated by increases of $1 USD/head in the proxy BRD$, had a direct effect of decreasing ADG (i.e. -0.0039 kg/head*day with 95% HPD interval [-0.0050, -0.0027]). These results are consistent with those of other studies (Sanderson et al. 2008, Cernicchiaro et al. 2013) reporting that increased entry weights and ADG were associated with decreased treatment costs in feedlot cattle. All direct effects between AW, BRD$ and ADG were considered of significant magnitude as their corresponding 95% HPD intervals did not include zero (Table 2.5).

Furthermore, we interpret indirect and overall effects of AW on ADG following classical path analysis, as in Wright (1921) and Shipley (2002). The indirect, BRD$-mediated, effects of AW on ADG were estimated as the product of the corresponding direct effects, namely ($\lambda_{BRD\$,AW} \times \lambda_{ADG,BRD\$}$). The posterior mean of the indirect BRD$-mediated effect of AW on ADG was 0.0003 kg/head*day with 95% HPD interval = [0.0002, 0.0004]. Direct and indirect effects of AW on ADG were then summed to yield an estimate of the overall effect of AW on ADG, that is $\lambda_{ADG,AW} + (\lambda_{BRD\$,AW} \times \lambda_{ADG,BRD\$})$, for which the posterior distribution had a mean of 0.0023 kg/head/day with 95% HPD boundaries of [0.0020, 0.0025].

2.4. DISCUSSION

We evaluated the inferential implications of incorrectly specifying data architecture and ensuing covariance structure in a structural equation modelling framework. Our evaluation was based on both a simulation and a data application in the area of animal agriculture, specifically in the beef cattle feedlot industry. More specifically, we assessed the impact of ignoring correlations between traits and correlations between observations within clusters (i.e. contemporary group in our data application) on relative global model fit as well as on accuracy and precision of parameter estimation. By leveraging recent methodological developments extending SEMs to a mixed models framework (Gianola and Sorensen 2004), we specified alternative SEMs that either accommodated data architecture fully (M1) or failed to do so in some way. Model misspecification was represented by models that either disregarded any cluster-level correlations between responses (M2), overlooked the inherently correlated nature of observations of a response within a cluster (M3) or completely ignored any effect of clustering (M4).

Perhaps not surprisingly, simulation results indicated that proper accounting for data architecture enhanced quality of data fit, as well as accuracy and precision of parameter estimation, with structural coefficients being an important inferential target. In turn, ignoring correlation patterns lead to impaired global fit to data, and this impairment was proportional to the severity of model mispecification: the greater the SEM mispecification, the poorer the global fit to data. Furthermore, disregard for the inherent structure of the data, even if as simple as clustering, resulted in impaired inference due to biased estimates and loss of precision. The latter may have ramifications for losses in statistical power (Tempelman 2009). In addition, these inferential problems may compound, particularly when dealing with experimental designs

characterized by one or multiple random variance components of relatively large magnitude (Dohoo 2008); for example in studies designed to thoughtfully account for sources of variability in the data.

For the data application, we used alternative SEMs to infer upon the interrelationships between health and production performance in feedlot cattle. While intended as proof-of-concept, the type of hierarchical data architecture in this application is quite common to agricultural applications. Specifically to livestock production systems, animals are commonly arranged in temporo-spatial clusters defined as contemporary groups, whereby observations within clusters are likely correlated due to shared environmental conditions and management background (Dohoo, Martin, and Stryhn 2014).

Our study builds on previous work relevant to public health by Bae et al. (2016), who assessed inference on correlated data using a special type of SEM known as Bayesian networks. Consistent with our results, Bae et al. (2016) showed that ignoring the correlation between subjects within a cluster resulted in biased parameter estimates and a substantially inflated rate of false positives associations. Notably, though, the data correlation patterns recognized by Bae et al. (2016) models were limited and did not account for any cluster-level correlations between responses. Specifically, Bae constrained $\sigma_{u_{jj'}} = 0$ for all $j \neq j'$, analogous to our SEM specification M2. Here, our SEM M1 relaxed such assumption and allowed cluster-level covariance parameters to be estimated and accounted for when inferring other model parameters of interest. As shown by simulation, estimation of cluster-level covariances (M1), as opposed to arbitrarily setting these parameters to zero (M2), was indeed relevant for inference, provided that the data were informative for cluster-level covariances.

In the feedlot data application, SEMs M1 and M2 fitted substantially better than M3 and M4, thereby supporting the specification of contemporary groups as random clustering effects and recognizing correlations between lot-level observations within a contemporary group. Unexpectedly, M2 was found to have a DIC value slightly smaller than M1, though the small numerical magnitude of the difference (~4 points; Table 2.4) was deemed inconclusive to choose between M1 and M2 based on relative goodness of fit. As a result, choice of a final model for inference was not clear. The observed numerical difference in DIC was primarily explained by the increased model complexity of M1 over M2 (Table 2.4), consistent with the difference between models in number of covariance parameters to be estimated. Given that the cluster-level covariance parameters were of minimal magnitude in this data application (i.e. all corresponding 95% HPD intervals included zero; Table 2.5), the complexity-driven difference in DIC values between M1 and M2 was not surprising. It would then seem that, at least for this data application, the assumption of $\sigma_{u_{jj'}} = 0$ for all $j \neq j'$ (as implied by our M2 and also by Bae et al. (2016)) was not particularly harmful for inference. Notably, though, the restrictive assumption of M2 should not be taken lightly, as it is in direct conflict with the multivariate nature of hierarchically-structured data. Rather, we chose M1 for final inference because it allowed cluster-level covariances $\sigma_{u_{jj'}}$ to vary within their parameter space, rather than forcing them to a constant value, thus enabling propagation of uncertainty through the analysis. Doing so is arguably a more desirable statistical practice, particularly in cases like this one in which model choice decisions rely on random-level parameters for which the data are often least informative.

Proper modeling of data architecture is directly relevant to applications in animal agriculture. For example, in quantitative genetics, contemporary groups defined as herd-year-season combinations are often treated as fixed effects (Visscher and Goddard 1993, Babot et al.

2003), as in our SEM M3. Instead, a random effect specification would allow for recovery of interblock (i.e. inter-cluster) information, and this is well known to improve the efficiency of estimates of fixed effects such as treatments (Yates 1940). Further, treating contemporary groups as random in SEM applications to quantitative genetics would allow for a broader scope of inference, thus enabling genetic predictions and potentially causal inference to be applicable beyond the specific groups represented in the data (Robinson 1991, Valente et al. 2015)

Our assessment of data architecture in this study is limited to clusters, which can be considered analogous to blocks in experimental design settings. Additional, more complex, components of data architecture that often characterize agricultural data include nested effects, split plots, repeated measures and subsampling, among others (Gbur 2012, Stroup 2013). Given recent methodological developments that extend SEM to a mixed models framework (Sorensen and Gianola 2004) and the availability of efficient software tools for implementation (e.g. R function gibbsREC) (Valente and Rosa 2013), proper accounting for data architecture and ensuing correlation patterns in the multivariate context of SEM is increasingly compelling.

Our data application of SEM was based on a causal structure elicited on the basis of epidemiological knowledge from field experts and was further supported by a comparative fit assessment of alternative causal structures (Cha et al. 2017). While an approach based on expert knowledge can be useful, its focus on a pre-selected subset of causal networks arguably impairs exploration of the structure space and potentially hinders discovery. Instead, researchers may be interested in data-driven selection of causal structures, such as in observational studies that use farm recorded data (Rosa and Valente 2013), similar to the feedlot data discussed here. In this context, the search algorithm known as inductive causation (Verma and Pearl 1991) may be relevant, particularly given that is has been recently extended to enable accounting for

hierarchical data structure in the causal search (Valente et al. 2011). The effect of mispecified data architecture on structure learning needs further study, as it is likely to compound the results described here for parameter estimation and inference.

A concern specific to the beef cattle feedlot data used here is the skewed-to-the-right nature of the response variable BRD$ (Table 2.1), which clearly departs from the SEM modeling assumption of normality. Initially, we considered applying a transformation to BRD$; however, a SEM fitted to responses expressed on a transformed scale poses challenges for expression of indirect and total effects, as well as for interpretation of structural coefficients as linear rates of change of one outcome as a function of another. Additional difficulties include the mediating status of BRD$ in the network under study (Figure 2.2), whereby BRD$ receives the effect of AW and has an effect on ADG. Further work is needed to assess how best to handle this type of responses in a SEM context.

Finally, we also acknowledge that SEM-based causal inference from observational data can be challenging, partly due to the need to make additional non-trivial assumptions that are qualitative in nature and not amenable to be tested on data (Pearl 2009). Specifically, causal homogeneity, Markov condition, faithfulness, and, particularly, causal sufficiency are critical causal assumptions required any time that causal inference is intended from observational data (Pearl 2009), such as the type of operational feedlot data used here.

2.5. CONCLUSIONS

Inferential implications of mispecified data architecture and ensuing covariance structure in the context of multilevel structural equation modeling were examined by simulation and by application to an operational feedlot dataset in beef cattle. As expected, when data architecture was properly modeled, global SEM fit was enhanced and parameter estimation was accurate; on

the other hand, SEM inference was impaired when data architecture was ignored or mispecified. Arguably, standing statistical recommendations that call for proper model specification capturing relevant hierarchical levels in data structure extend to the multivariate context of structural equation modeling. Ensuring that structural equation modelling yields sound estimation and inference requires careful consideration of the data generation process and ensuing multilevel correlation patterns in data.

ACKNOWLEDGEMENTS

This project was partially funded by the United States Department of Agriculture National Institute of Food and Agriculture Award # 2015-67015-23079. Computing for this project was partially performed on the Beocat Research Cluster at Kansas State University, which is funded in part by NSF Grants CNS-1006860, EPS-1006860 and EPS-0919443.

REFERENCES

Babot, D., Noguera, J. L., Alfonso, L., and Estany, J. 2003. "Fixed or Random Contemporary Groups in Genetic Evaluation for Litter Size in Pigs Using a Single Trait Repeatability Animal Model," *Journal of Animal Breeding and Genetics*, 120(1), 12-22.

Bae, H., Monti, S., Montano, M., Steinberg, M. H., Perls, T. T., and Sebastiani, P. 2016. "Learning Bayesian Networks from Correlated Data," *Scientific Reports*, 6, 25156, 1-14.

Bello, N. M., Steibel, J. P., and Tempelman, R. J. 2010. "Hierarchical Bayesian Modeling of Random and Residual Variance–covariance Matrices in Bivariate Mixed Effects Models," *Biometrical Journal*, 52(3), 297-313.

Cernicchiaro, N., White, B. J., Renter, D. G., and Babcock, A. H. 2013. "Evaluation of Economic and Performance Outcomes Associated with the Number of Treatments after an Initial Diagnosis of Bovine Respiratory Disease in Commercial Feeder Cattle," *American Journal of Veterinary Research*, 74(2), 300-309.

Cha, E., Sanderson, M., Renter, D., Jager, A.,Cernicchiaro, N., and Bello, N. M. 2017. "Implementing Structural Equation Models to Observational Data from Feedlot Production Systems," *Preventive Veterinary Medicine*, 147, 163-171.

de los Campos, G., Gianola, D., Boettcher, P., and Moroni, P. 2006. "A Structural Equation Model for Describing Relationships between Somatic Cell Score and Milk Yield in Dairy Goats," *Journal of Animal Science*, 84(11), 2934-2941.

Dohoo, I., Martin, W., and Stryhn, H. 2014. *Veterinary Epidemiologic Research* (2nd ed.), Canada: VER Inc.

Dohoo, I. R. 2008. "Quantitaive epidemiology: Progress and challenges," *Preventive Veterinary Medicine*, 86(3), 260-269.

Duncan, O. D. 1966. "Path Analysis: Sociological Examples," *American Journal of Sociology*, 72 (1), 1-16.

Gbur, E. E., Stroup, W., McCarter, W., Kevin, S., Durham, S., Young, L. J., Christman, M., West, M., and Kramer, M. 2012. *Analysis of Generalized Linear Mixed Models in the Agricultural and Natural Resources Sciences*, Madison, WI, USA: American Society of Agronomy, Soil Science Society of America, Crop Science Society of America, Inc.

Gelman, A. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis*, 1(3), 515-533.

Gianola, D. and Sorensen, D. 2004. "Quantitative Genetic Models for Describing Simultaneous and Recursive Relationships between Phenotypes," *Genetics*, 167(3), 1407-1424.

Haavelmo, T. 1943. "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11 (1), 1-12.

Hay, K. E., Barnes, T. S., Morton, J. M., Clements, A. C. A., and Mahony, T. J. 2014. "Risk Factors for Bovine Respiratory Disease in Australian Feedlot Cattle: Use of a Causal

Diagram-informed Approach to Estimate Effects of Animal Mixing and Movements before Feedlot Entry," *Preventive Veterinary Medicine*, 117(1), 160-169.

Inoue, K., Valente, B. D., Shoji, N., Honda, T., Oyama, K., and Rosa, G. J. 2016. "Inferring Phenotypic Causal Structures among Meat Quality Traits and the Application of a Structural Equation Model in Japanese Black Cattle," *Journal of Animal Science*, 94(10), 4133-4142.

Johnson, Richard A., and Dean W. Wichern. (2007). *Applied Multivariate Statistical Analysis* (6th ed), Upper Saddle River, New Jersey: Pearson Prentice Hall.

Joreskog, K.G. 1973. *A General Method for Estimating a Linear Structural Equation System*, Edited by A. S. Goldberger and O. D. Duncan, *Equation Models in the Social Sciences*, New York: Senimar Press.

Konig, S., Wu, X. L., Gianola, D., Heringstad, B., and Simianer, H. 2008. "Exploration of Relationships between Claw Disorders and Milk Yield in Holstein Cows via Recursive Linear and Threshold Models," *Journal of Dairy Science*, 91(1), 395-406.

Lauritzen, S. L. (1996). *Graphical models*. Oxford, UK: Oxford University Press.

Littell, R. C., Milliken G. A. , Stroup W., Russell, D. W., and Schabenberger, O. 2006. *SAS for Mixed Models* (2nd ed.), Cary, NC: SAS Institute Inc.

Lopez de Maturana, E. , Wu, X. L., Gianola, D., Weigel, K. A. and Rosa, G. J. 2009. "Exploring biological relationships between calving traits in primiparous cattle with a Bayesian recursive model," *Genetics*,181 (1), 277-87.

Milliken, G. A. and Johnson, D. E. 2009. *Analysis of Messy Data - Volume 1: Designed Experiments* (2nd ed.), Boca Raton, Florida, USA: Chapman and Hall/CRC Press.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inference* (2nd ed.), Cambridge University Press.

Peñagarican, F.,Valente, B. D., Steibel, J. P., Bates, R. O., Ernst, C. W., Khatib, H., and Rosa, G. J. 2015. "Searching for Causal Networks Involving Latent Variables in Complex Traits: Application to Growth, Carcass, and Meat Quality Traits in Pig," *Journal of Animal Science*, 93(10), 4617-4623.

Plummer, M., Best, N., Cowles, K., and Vines, K. 2006. "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News*, 6, 7-11.

Raftery, A. and Lewis, S. 1992. "How many iterations in the Gibbs sampler," *In Bayesian Statistics 4*, 763-773, Oxford University Press.

Robinson, G. K. 1991. "That BLUP is a Good Thing: The Estimation of Random Effects," *Statistics Science*, 6(1), 15-32.

Rosa, G. J. M. and Valente, B. D. (2013). "BREEDING AND GENETICS SYMPOSIUM: Inferring causal effects from observational data in livestock." *Journal of Animal Science*, 91(2), 553-564.

Rosa, G. J., Valente, B. D., de los Campos, G., Wu, X. L., Gianola, D., and Silva, M. A. 2011. "Inferring Causal Phenotype Networks Using Structural Equation Models," *Genetics Selection Evolution,* 43, 6-18.

R Development Core Team. 2017. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria.

Sanderson, M., Dargatz, D. A., and Wagner, B. A. 2008. "Risk Factors for Initial Respiratory Disease in United States' Feedlots based on Producer-collected Daily Morbidity Counts," *Canadian Veterianary Journal*, 49(4), 373-378.

Shipley, B. 2002. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*, Cambridge University Press.

Sorensen, D., Andersen, S., Gianola, D., and Korsgaard., I. 1995. "Bayesian-inference in Threshold Models Using Gibbs Sampling, *Genetics Selection Evolution,* 27(3), 229-249.

Sorensen, D. and Gianola, D. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*, New York, Springer-Verlag.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. 2002. "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.

Stroup, W. 2013. *Generalized Linear Mixed Models*, Boca Raton, Florida, CRC Press Taylor & Francis Group.

Tempelman, R. J. 2009. "Invited Review: Assessing Experimental Designs for Research Conducted on Commercial Dairies," *Journal of Dairy Science*, 92(1), 1-15.

Valente, B. D., Morota, G., Penagaricano, F., Gianola, D., Weigel, K., and Rosa, G. J. 2015. "The Causal Meaning of Genomic Predictors and How It Affects Construction and Comparison of Genome-Enabled Selection Models," *Genetics,* 200(2), 483-494

Valente, B. D., Rosa, G. J., de los Campos, G., Gianola, D., and Silva, M. A. 2010. "Searching for Recursive Causal Structures in Multivariate Quantitative Genetics Mixed Models," *Genetics*, 185(2), 633-644.

Valente, B. D., Rosa G. J., Silva, M. A., Teixeira, R. B., and Torres, R. A. 2011. "Searching for Phenotypic Causal Networks Involving Complex Traits: an Application to European Quail," *Genetics Selection Evolution,* 43,37-48.

Varona, L. and Sorensen, D. 2014. "Joint Analysis of Binomial and Continuous Traits with a Recursive Model: A Case Study Using Mortality and Litter Size of Pigs," *Genetics*, 196 (3), 643-651.

Verma, Thomas, and Judea Pearl. 1991. "A Theory of Inferred Causation," In Allen, J. A., Fike, R. snd Sandwall, E. (editors), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*, 441-452, Morgan Kaufmann, San Mateo.

Visscher, P. M. and Goddard, M. E. 1993. "Fixed and Random Contemporary Groups," *Journal of Dairy Science*, 76 (5), 1444-1454.

Wright, S. 1921. "Correlation and Causation.," *Journal of Agricultural Research,* 20, 557-585.

Wright, S. 1934. "The Method of Path Coefficients," *The Annals of Mathematical Statistics,* 5 (3), 161-215.

Wu, X. L., Heringstad, B., Chang, Y. M., de Los Campos, G., and Gianola, D. 2007. "Inferring Relationships between Somatic Cell Score and Milk Yield Using Simultaneous and Recursive models," *Journal of Dairy Science*, 90 (7), 3508-3521.

Yates, F. (1940). "The Recovery of Inter-block Information in Balanced Incomplete Block Designs," *Annals of Eugenics*, 10(1), 317-325.

FIGURES AND TABLES

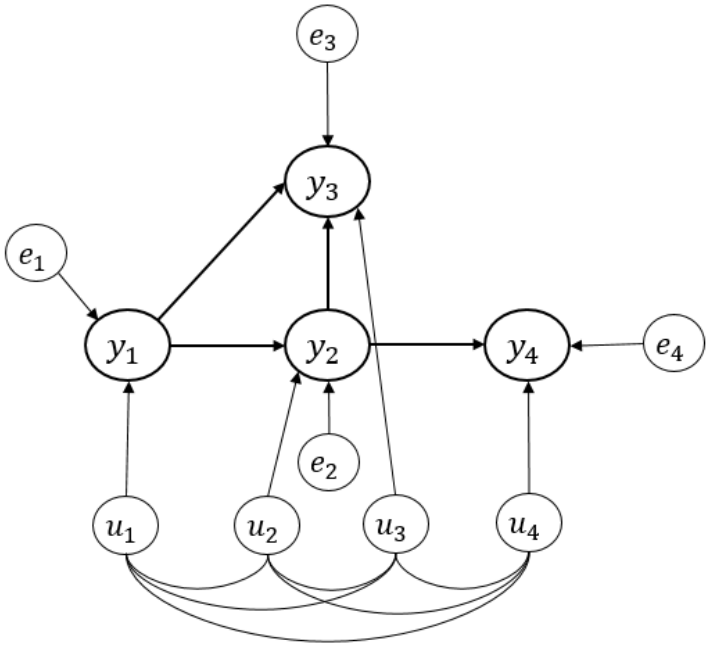


Figure 2.1 Directed acyclic graph depicting functional links between response variables y_1, y_2, y_3 and y_4, following the structural equation model in Equation (1). Directed edges, or arrows, indicate that y_1 has a direct effect on y_2 (i.e. $y_1 \rightarrow y_2$) and on y_3 (i.e. $y_1 \rightarrow y_3$). Also, y_2 directly affects y_3 (i.e. $y_2 \rightarrow y_3$) and y_4 (i.e. $y_2 \rightarrow y_4$). Random cluster effects u_j (j = 1, 2, 3, 4) are defined on each response, and can potentially be mutually correlated, as indicated by undirected edges connecting them. Each response y_j is also directly affected by a corresponding residual e_j, which are assumed mutually independent of each other.

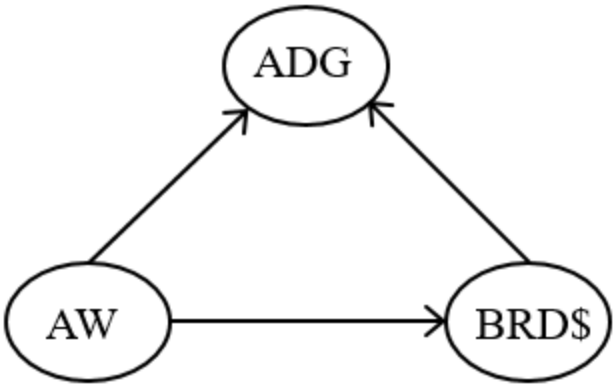


Figure 2.2 Directed acyclic graph of expert-elicited functional relationships between arrival weight (AW), Bovine Respiratory Disease related treatment costs (BRD$) and average daily gain (ADG) in feedlot cattle. Contributing factors are described in the text.

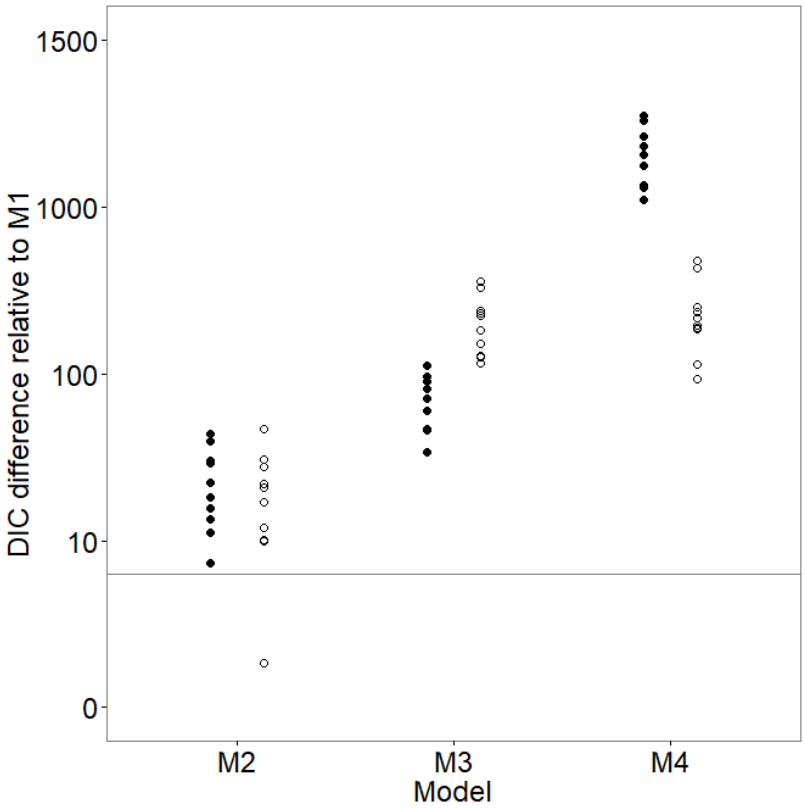


Figure 2.3 Differences in Deviance Information Criteria (DIC), expressed relative to Model 1 (M1; true data generating model), for alternative SEM specifications for which data architecture was mispecified (M2, M3) or ignored (M4), under simulated conditions of balanced cluster size. Each dot represents a simulated dataset. White and black dots represent indicate simulation scenarios of moderate (~10%) or large (~30-50%) magnitude of cluster-level variance relative to the total variance of each response, respectively. A reference horizontal line indicates the threshold value indicative of a difference in fit relative to M1.

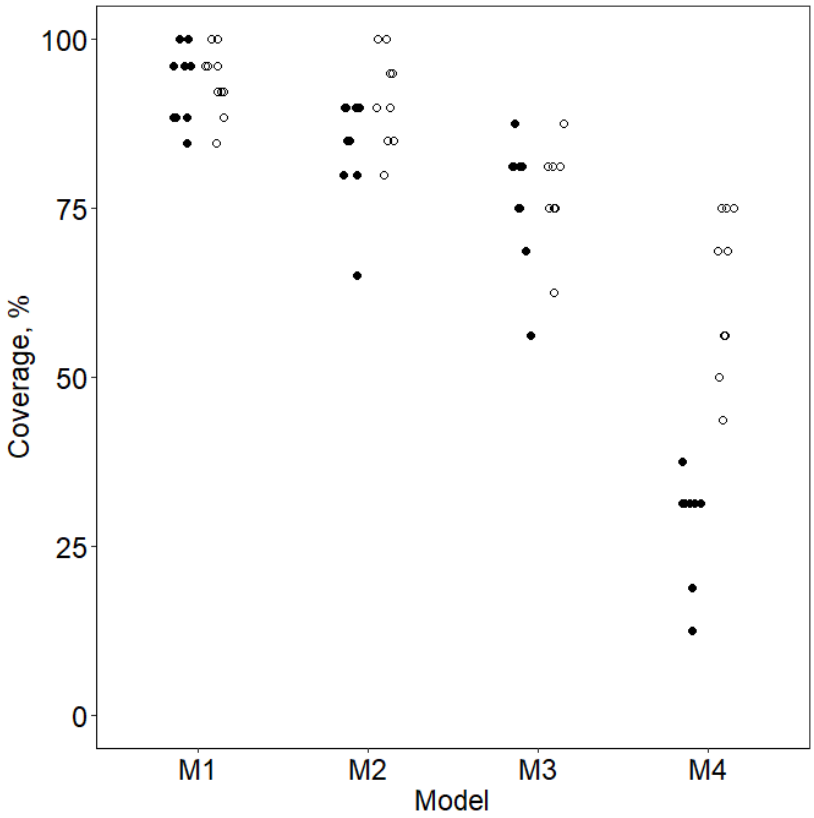


Figure 2.4 Overall accuracy of parameter estimation, expressed as proportion of coverage (in %) based on 95% HPD intervals, for alternative SEM specifications for which data architecture was correctly specified (M1), mispecified (M2, M3) or ignored (M4) under simulated conditions of balanced cluster sizes. White and black dots indicate simulation scenarios of moderate (~10%) or large (~30-50%) magnitude of cluster-level variance relative to the total variance of each response, respectively.

Table 2.1 Descriptive statistics for response variables collected on 1,430 lots of cattle representing a total of 178,983 animals from 9 feedlots located across the US Great Plains during 2014 and 2015.

Response variables	Mean	Min	25th Percentile	Median	75th Percentile	Max
Arrival weight (AW; kg per head)	320.9	144.3	290.3	332.7	363.6	397.6
Bovine Respiratory Disease-related treatment costs (BRD$; $ per head)	7.4	0.0	0.0	0.5	4.4	91.7
Average daily gain (ADG; g per head per day)	1.5	0.05	1.4	1.5	1.7	2.3

Table 2.2 Coverage (expressed as %) of selected parameter types for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of balanced cluster size.

		Moderate cluster-level variance				Large cluster-level variance			
Parameters	No. of simulated events	M1	M2	M3	M4	M1	M2	M3	M4
$\lambda_{jj'}$	40	90.0	80.0↓	90.0	47.5↓	87.5	62.5↓	87.5	20.0↓
$\beta_{1,j}$	40	95.0	90.0	27.5↓	70.0↓	95.0	85.0	30.0↓	42.5↓
$\beta_{2,j}$	40	95.0	95.0	97.5	87.5	97.5	92.5	100.0	52.5↑
$\sigma^2_{e_j}$	40	92.5	95.0	92.5	45.0↑	92.5	90.0	90.0	0↑
$\sigma^2_{u_j}$	40	90.0	100.0	-	-	95.0	92.5	-	-
$\sigma_{u_{jj'}}$	60	98.3	-	-	-	95.0	-	-	-

For j = 1, 2, 3, and 4; $j' < j$

Upward and downward arrows indicate whether a given parameter type was consistently overestimated or underestimated, respectively.

M1: True data-generating model.

M2: Model ignores cluster-level correlations between responses.

M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).

M4: Model ignores data architecture all together.

Table 2.3 Precision of parameter estimation, expressed as width of the 95% HPD intervals (mean and SD) across 10 simulated datasets, for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of moderate magnitude of cluster-level variability and balanced cluster size.

	Alternative SEM specifications of data architecture							
	M1		M2		M3		M4	
Parameters	Mean	SD	Mean	SD	Mean	SD	Mean	SD
λ_{21}	0.0261	0.0004	0.0253	0.0005	0.0274	0.0003	0.0247	0.0005
λ_{31}	0.1059	0.0039	0.1003	0.0033	0.1085	0.0041	0.0992	0.0035
λ_{32}	0.1981	0.0081	0.1890	0.0069	0.2073	0.0078	0.1904	0.0060
λ_{42}	0.0814	0.0025	0.0788	0.0026	0.0860	0.0035	0.0791	0.0024
$\beta_{1,1}$	6.29	0.28	6.06	0.28	3.67	0.09	4.85	0.10
$\beta_{1,2}$	2.90	0.07	2.80	0.08	2.56	0.05	2.66	0.07
$\beta_{1,3}$	7.36	0.32	7.10	0.29	6.79	0.26	7.25	0.29
$\beta_{1,4}$	6.86	0.25	6.60	0.24	6.24	0.29	6.53	0.23
$\beta_{2,1}$	7.89	0.19	7.89	0.16	9.19	0.23	8.13	0.19
$\beta_{2,2}$	1.76	0.05	1.76	0.05	2.12	0.05	1.79	0.04
$\beta_{2,3}$	3.34	0.13	3.29	0.11	3.88	0.17	3.39	0.13
$\beta_{2,4}$	3.37	0.15	3.33	0.13	3.96	0.18	3.38	0.12
$\sigma^2_{e_1}$	161.7	7.7	164.4	7.5	163.1	7.4	171.9	6.0
$\sigma^2_{e_2}$	6.6	0.3	6.7	0.3	6.6	0.3	6.9	0.3
$\sigma^2_{e_3}$	15.0	0.7	15.2	0.7	15.2	0.7	15.9	0.8
$\sigma^2_{e_4}$	18.6	1.0	18.7	1.1	18.7	1.0	19.5	0.8
$\sigma^2_{u_1}$	120.7	15.0	108.7	12.9	-	-	-	-
$\sigma^2_{u_2}$	4.9	0.5	4.3	0.4	-	-	-	-
$\sigma^2_{u_3}$	11.8	1.9	10.1	1.6	-	-	-	-
$\sigma^2_{u_4}$	13.5	1.6	12.1	1.5	-	-	-	-
$\sigma_{u_{21}}$	19.0	1.8	-	-	-	-	-	-
$\sigma_{u_{31}}$	28.5	2.1	-	-	-	-	-	-
$\sigma_{u_{32}}$	5.7	0.6	-	-	-	-	-	-
$\sigma_{u_{41}}$	30.1	3.2	-	-	-	-	-	-
$\sigma_{u_{42}}$	5.6	0.4	-	-	-	-	-	-
$\sigma_{u_{43}}$	9.0	0.8	-	-	-	-	-	-

"-" indicates that estimates are not applicable due to SEM specification.
M1: True data-generating model.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together.

Table 2.4 Comparison of global model fit based on Deviance Information Criteria (DIC) of alternative SEM specifications for data architecture fitted to feedlot data.

Alternative SEM specifications	p_D	DIC	DIC difference relative to M1
M1	262.7	24,323.8	
M2	259.1	24,319.9	- 3.9
M3	339.6	24,399.8	76.0
M4	11.0	26,178.2	1,854.4

M1: Model accounts for cluster-level correlation between responses and within-cluster correlation between observations for a given response.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together.
p_D: Effective number of parameters, indicative of model complexity (Spiegelhalter et al. 2002).

Table 2.5 Posterior summaries of parameters of interest obtained from alternative SEM specifications of data architecture M1 – M4 fitted to data from commercial feedlot operations. Posterior means and 95% highest posterior density intervals (in brackets) are shown.

Parameter	Alternative SEM specifications of data architecture			
	M1	M2	M3	M4
$\lambda_{BRD\$,AW}$	-0.0785 [-0.0904, -0.0667]	-0.0819 [-0.0932, -0.0704]	-0.0790 [-0.0904, -0.0667]	-0.1280 [-0.1395, -0.1163]
$\lambda_{ADG,AW}$	0.0020 [0.0017, 0.0022]	0.0020 [0.0017, 0.0022]	0.0020 [0.0017, 0.0022]	0.0020 [0.0018, 0.0022]
$\lambda_{ADG,BRD\$}$	-0.0039 [-0.0050, -0.0027]	-0.0036 [-0.0046, -0.0027]	-0.0039 [-0.0050, -0.0027]	-0.0033 [-0.0042, -0.0025]
$\beta_{1,AW}$	301.9 [291.8, 312.2]	301.8 [291.9, 311.4]	324.3 [313.0, 335.5]	314.6 [310.9, 318.1]
$\beta_{1,BRD\$}$	33.5 [29.3, 37.6]	34.5 [30.4, 38.4]	32.6 [28.3, 36.8]	47.4 [43.8, 51.3]
$\beta_{1,ADG}$	0.886 [0.800, 0.972]	0.885 [0.808, 0.963]	0.939 [0.841, 1.019]	0.873 [0.799, 0.945]
$\beta_{2,AW}$	18.2 [13.9, 22.8]	18.3 [13.8, 22.7]	18.3 [13.2, 23.3]	16.7 [10.8, 22.5]
$\beta_{2,BRD\$}$	2.2 [1.2, 3.2]	2.3 [1.3, 3.3]	2.3 [1.1, 3.4]	2.7 [1.4, 4.0]
$\beta_{2,ADG}$	0.1346 [0.1130, 0.1563]	0.1341 [0. 1130, 0.1555]	0.1383 [0.1137, 0.1629]	0.1232 [0.1018, 0.1448]
$\sigma^2_{e_{AW}}$	1540.0 [1422.7, 1657.5]	1538.0 [1422.5 1658.3]	1534.0 [1420.5, 1655.3]	2981.0 [2759.1, 3197.8]
$\sigma^2_{e_{BRD\$}}$	76.1 [70.4, 82.1]	76.1 [70.20, 81.8]	76.1 [70.4, 82.0]	152.3 [141.5, 163.8]
$\sigma^2_{e_{ADG}}$	0.035 [0.032, 0.0378]	0.035 [0.032, 0.038]	0.035 [0.032, 0.038]	0.040 [0.036, 0.042]
$\sigma^2_{u_{AW}}$	2543.0 [1763.4, 3371.3]	2425.0 [1717.1, 3212.3]	-	-
$\sigma^2_{u_{BRD\$}}$	103.3 [73.5, 136.8]	97.7 [68.8, 127.9]	-	-
$\sigma^2_{u_{ADG}}$	0.0055 [0.0031, 0.0082]	0.0049 [0.0027, 0.0074]	-	-
$\sigma_{u_{BRD\$,AW}}$	-119.2 [-246.0, 1.0]	-	-	-
$\sigma_{u_{ADG,AW}}$	-0.22 [-1.56, 1.17]	-	-	-
$\sigma_{u_{ADG,BRD\$}}$	0.07 [-0.18, 0.32]	-	-	-

M1: Model accounts for cluster-level correlation between responses and within-cluster correlation between observations for a given response.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together

Table 2.6 Width of 95% HPD interval for parameters fitted to feedlot data with alternative SEM specifications of data architecture.

	Alternative SEM specifications of data architecture			
Parameter	M1	M2	M3	M4
$\lambda_{BRD\$,AW}$	0.0236	0.0228	0.0237	0.0232
$\lambda_{ADG,AW}$	0.0005	0.0005	0.0005	0.0004
$\lambda_{ADG,BRD\$}$	0.0023	0.0019	0.0023	0.0017
$\beta_{1,AW}$	20.46	19.46	22.45	7.20
$\beta_{1,BRD\$}$	8.29	8.02	8.56	7.48
$\beta_{1,ADG}$	0.17	0.16	0.18	0.15
$\beta_{2,AW}$	8.96	8.94	10.12	11.66
$\beta_{2,BRD\$}$	2.03	2.02	2.28	2.67
$\beta_{2,ADG}$	0.0433	0.0424	0.0491	0.0430
$\sigma^2_{e_{AW}}$	234.89	235.73	234.75	438.69
$\sigma^2_{e_{BRD\$}}$	11.65	11.64	11.59	22.30
$\sigma^2_{e_{ADG}}$	0.0053	0.0052	0.0054	0.0057
$\sigma^2_{u_{AW}}$	1607.88	1495.21	-	-
$\sigma^2_{u_{BRD\$}}$	2.73	59.14	-	-
$\sigma^2_{u_{ADG}}$	0.0052	0.0047	-	-
$\sigma_{u_{BRD\$,AW}}$	244.97	-	-	-
$\sigma_{u_{ADG,AW}}$	63.28	-	-	-
$\sigma_{u_{ADG,BRD\$}}$	0.4975	-	-	-

M1: Model accounts for cluster-level correlation between responses and within-cluster correlation between observations for a given response.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together.

Chapter 3 - Investigating Causal Biological Relationships between Reproductive Performance Traits in High-performing Gilts and Sows

ABSTRACT

Efficient management of swine production systems requires understanding of complex reproductive physiological mechanisms. Our objective in this study was to investigate potential causal biological relationships between reproductive performance traits in high-producing gilts and sows. Data originated from a nutrition experiment and consisted of 200 sows and 440 gilts arranged in body weight blocks and randomly assigned to dietary treatments during late gestation at a commercial swine farm. Reproductive performance traits consisted of weight gain during late gestation, total number born (TB) and number born alive (BA) in a litter, born alive average birth weight (BABW), wean-to-estrous interval, and total litter size born in the subsequent farrowing. Structural equation models combined with the inductive causation algorithm, both adapted to a hierarchical Bayesian framework, were employed to search for, estimate and infer upon causal links between the traits within each parity group. Results indicated potentially distinct reproductive networks for gilts and for sows. Sows showed sparse connectivity between reproductive traits whereas the network learned for gilts was densely interconnected, suggesting closely linked physiological mechanisms in younger females, with a potential for ripple effects throughout their productive lifecycle in response to early implementation of tailored managerial interventions. Crossvalidation analyses indicated substantial network stability both for the general structure and for individual links, though results about directionality of such links were unstable in this study and will need further investigation. An assessment of relative statistical

power in sows and gilts indicated that the observed network discrepancies may be partially explained on a biological basis. In summary, our results suggest distinctly heterogeneous mechanistic networks of reproductive physiology for gilts and sows, consistent with physiological differences between the groups. These findings have potential practical implications for integrated understanding and differential management of gilts and sows to enhance efficiency of swine production systems.

Key words: hierarchical Bayesian models, structural equation model, structure learning, swine reproductive physiology.

3.1 INTRODUCTION

Female reproductive physiology involves complex mechanisms that need to be integrated and managed efficiently in swine production systems. Understanding the functional links that underlie the relationship between multiple reproductive performance traits can provide valuable mechanistic insight, and thus enhance efficiency of the decision making process in swine production systems. Classical multiple-trait models (MTM) can be used to simultaneously consider multiple traits and study their probabilistic relationships, expressed as correlations or covariances (Henderson and Quaas, 1976; Van Vleck and Edlin, 1984). However, classical MTM are limited in that they cannot assess directionality of such relationships (Valente and Rosa 2013), thus hindering their use for exploring causal relationships within a network, as might be of interest in complex systems with multiple interrelated traits.

Structural Equation Models (SEM) (Haavelmo, 1943) are a special type of MTM that, under some assumptions, can accommodate directionality and thus potentially causal relationships in the links that define a functional network (Pearl, 2009). Thus, an SEM approach can facilitate a more comprehensive understanding of physiological mechanisms as an interconnected system, as opposed to an assortment of individual outcome traits, the relationship among which is evaluated only anecdotally or, at best, one at a time. Originally evolved from path analysis (Wright, 1934), SEM were recently adapted to a mixed models framework (Gianola and Sorensen, 2004), thereby enabling specification of data architecture, as is typically the case with designed experiments. Search algorithms, such as inductive causation (IC; Verma and Pearl, 1991), are also available to help investigate the network space while accounting for multiple levels of data structure (Valente et al., 2010). Given these developments, SEM are being increasingly used in the animal sciences, particularly in quantitative genetics. Examples include

beef cattle (Inoue et al., 2016), dairy cattle (de Maturana et al., 2009), dairy goats (de los Campos et al., 2006), swine (Varona and Sorensen, 2014), and quail (Valente et al., 2011), among others.

Here, our interest in causal network-type relationships is motivated by a recent swine experiment conducted to evaluate the effect of nutritional management during late gestation on multiple reproductive performance traits in high performing gilts and sows (Gonçalves et al., 2016). In that experimental study, each trait was analyzed separately using a generalized linear mixed model approach, whereby each trait represented related, though arguably different, aspects of swine reproductive physiology. The randomization process in that experiment granted causal inference of nutritional dietary treatment on each individual trait. In the study herein, we pursue a more comprehensive understanding of the reproductive physiological system as a whole, recognizing that, in addition to treatment effects, reproductive performance traits may also influence each other within the system.

The objective of this study was to investigate potential causal biological relationships between reproductive performance traits in high-producing gilts and sows in the context of a designed nutrition experiment. Using hierarchical SEM combined with IC, we search for, estimate and infer upon potential causal network-type interrelationships between reproductive traits in sows and gilts. We further evaluate the stability of the learned network structures and assess potential differences in statistical power between the parity groups.

3.2 MATERIALS AND METHODS

3.2.1 Data

Data were obtained from an experimental study on swine nutrition (Gonçalves et al., 2016) conducted at a commercial swine farm in northern Ohio under an experimental protocol approved by the Kansas State University Institutional Animal Care and Use Committee, as described by Gonçalves et al. (2016). The complete experimental data consisted of observations from 1,102 females, including 361 sows and 741 gilts. For each parity group (i.e. sows and gilts), the experimental design used for data collection consisted of a randomized complete block design with blocks of 4 females defined along a body weight gradient. Within each block, four dietary treatments consisting of combinations of energy intake and dietary amino acids were randomly allocated. The complete description of the data is available in Gonçalves et al. (2016). The reproductive performance traits of interest, labeled $j = 1, \ldots, J,$ for $J = 6$, included female weight gain during late (d 90 to d 111) gestation (GAIN, in kg; $j = 1$), total number of piglets born (TB; $j = 2$) and number born alive (BA; $j = 3$) in a litter, and born alive average birth weight (BABW, in g; $j = 4$) , wean-to-estrous interval (WEI, in days; $j = 5$), and total litter size born in the subsequent farrowing (SuTB; $j = 6$). Observations on GAIN were obtained as the difference in female body weight observed at d 90 and at d 111 of gestation. Observations on TB, BA and BABW were collected simultaneously at the time of farrowing and preceded those observations on WEI, which in turn preceded data collection on SuTB. Both TB and SuTB refer to total litter size and included piglets born alive, stillborns and mummies. Only females with complete records on all six reproductive performance traits (i.e. no missing data) were considered for analyses in this study. Following data editing for removal of incomplete records, the final dataset used for analysis consisted of 200 sows and 440 gilts arranged in 97 and 222 body weight blocks,

respectively. As a result, the dataset analyzed consisted of incomplete blocks that ranged in size from 1 to 4 females.

3.2.2 The Structural Equation Model

For each parity group, namely sows and gilts, we write a hierarchical SEM following Gianola and Sorensen (2004), as follows:

$$\boldsymbol{y}_i = \boldsymbol{\Lambda}\boldsymbol{y}_i + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b} + \boldsymbol{e}_i \tag{3.1}$$

where $\boldsymbol{y}_i' = [y_{i1}\ y_{i2}\ y_{i3}\ y_{i4}\ y_{i5}\ y_{i6}]'$ is a vector of $J = 6$ reproductive performance traits observed on the i^{th} female ($i = 1, 2, \ldots, n$; corresponding to $n = 200$ sows or $n = 440$ gilts, respectively). Also, $\boldsymbol{\Lambda}$ is a $J \times J$ matrix composed by zeroes along the main diagonal and on the upper triangle, and unknown structural coefficient parameters $\lambda_{jj'}$ on the lower triangle (under a fully recursive specification) representing functional links between responses in the network. The structural coefficient $\lambda_{jj'}$ $(j' < j)$ describes the direct effect of trait $y_{j'}$ on trait y_j. For example, λ_{21} represents the expected change in y_2 per unit increase in y_1. Further, $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1'\ \boldsymbol{\beta}_2'\ \boldsymbol{\beta}_3'\ \boldsymbol{\beta}_4'\ \boldsymbol{\beta}_5'\ \boldsymbol{\beta}_6']$ is a vector of unknown fixed-effect location parameters associated with treatment factors through the design matrix $\boldsymbol{X}_i = diag(\boldsymbol{x}_{i1}'\ \boldsymbol{x}_{i2}'\ \boldsymbol{x}_{i3}'\ \boldsymbol{x}_{i4}'\ \boldsymbol{x}_{i5}'\ \boldsymbol{x}_{i6}')$ unique to each subject i. Next, $\boldsymbol{b}' = [\boldsymbol{b}_1'\ \boldsymbol{b}_2'\ \boldsymbol{b}_3'\ \boldsymbol{b}_4'\ \boldsymbol{b}_5'\ \boldsymbol{b}_6']$ is a vector of unknown random effects associated with body weight blocks expressed in the design matrix $\boldsymbol{Z}_i = diag(\boldsymbol{z}_{i1}'\ \boldsymbol{z}_{i2}'\ \boldsymbol{z}_{i3}'\ \boldsymbol{z}_{i4}'\ \boldsymbol{z}_{i5}'\ \boldsymbol{z}_{i6}')$. Random effects $\boldsymbol{b}$ are assumed multivariate normally distributed with null mean vector and covariance matrix $\boldsymbol{B}$, such that:

$$\boldsymbol{b} \sim MVN\left(\boldsymbol{0},\ \boldsymbol{B} \otimes \boldsymbol{I}_q = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_{12}} & \sigma_{b_{13}} & \sigma_{b_{14}} & \sigma_{b_{15}} & \sigma_{b_{16}} \\ \sigma_{b_{12}} & \sigma_{b_2}^2 & \sigma_{b_{23}} & \sigma_{b_{24}} & \sigma_{b_{25}} & \sigma_{b_{26}} \\ \sigma_{b_{13}} & \sigma_{b_{23}} & \sigma_{b_3}^2 & \sigma_{b_{34}} & \sigma_{b_{35}} & \sigma_{b_{36}} \\ \sigma_{b_{14}} & \sigma_{b_{24}} & \sigma_{b_{34}} & \sigma_{b_4}^2 & \sigma_{b_{45}} & \sigma_{b_{46}} \\ \sigma_{b_{15}} & \sigma_{b_{25}} & \sigma_{b_{35}} & \sigma_{b_{45}} & \sigma_{b_5}^2 & \sigma_{b_{56}} \\ \sigma_{b_{16}} & \sigma_{b_{26}} & \sigma_{b_{36}} & \sigma_{b_{46}} & \sigma_{b_{56}} & \sigma_{b_6}^2 \end{bmatrix} \otimes \boldsymbol{I}_q \right) \tag{3.2}$$

where q is the number of blocks within each parity group. Finally, $\boldsymbol{e_i'} = [e_{i1}\ \ e_{i2}\ \ e_{i3}\ \ e_{i4}\ e_{i5}\ e_{i6}]$ is the corresponding set of residuals for the i^{th} female, assumed multivariate normally distributed with null mean vector and diagonal (co)variance matrix $\boldsymbol{R}$, such that:

$$\boldsymbol{e}_i \sim MVN\left(\mathbf{0}, \boldsymbol{R} = \begin{bmatrix} \sigma^2_{e_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_{e_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_{e_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_{e_4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_{e_5} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_{e_6} \end{bmatrix}\right) \tag{3.3}$$

The assumption of a diagonal residual (co)variance matrix $\boldsymbol{R}$ is standard in the context of SEM to ensure parameter identifiability in an acyclic causal framework (Gianola and Sorensen, 2004). Any additional source of covariability (besides random effects) is attributed to structural coefficients in $\boldsymbol{\Lambda}$ and refers to functional relationships between traits (Gianola and Sorensen, 2004; Wu et al., 2007)

3.2.3 The Inductive Causation Algorithm

We implemented the IC algorithm as adapted to mixed models by Valente et al. (2010) to search the space of network structures separately for each parity group. The search was conducted at the residual level of a MTM, that is, after conditioning on the random effects that characterize the experimental design. This approach prevents the structure search from being confounded by correlations between traits that might be induced by the design structure of the data; this is analogous to the approach proposed by Valente et al. (2010) in the context of quantitative genetics. Briefly, we first fit a hierarchical Bayesian SEM with a fully recursive specification equivalent to a standard MTM to the data. Posterior samples of the SEM parameters in matrices $\boldsymbol{R}$ and $\boldsymbol{\Lambda}$ were used to compute the posterior densities for parameters of $\boldsymbol{R}^*$ defined as the residual (co)variance matrix under the equivalent MTM, whereby $\boldsymbol{R}^* =$

$(\boldsymbol{I}-\boldsymbol{\Lambda})^{-1}\boldsymbol{R}(\boldsymbol{I}-\boldsymbol{\Lambda})^{-1'}$**.** Next, we used posterior samples of $\boldsymbol{R}^{*}$ to compute all partial residual correlations for each pair of reproductive performance traits, given all possible conditioning sets consisting of combinations of the remaining traits. Finally, we implemented the IC algorithm on such partial residual correlations between traits to make a series of statistical decisions intended to yield a learned network structure (Valente et al., 2010), as follows:

Step 1: For each pair of traits y_j and $y_{j'}$, evaluate all partial residual correlations conditional on every possible conditioning set of traits besides the *(j, j')* pair. if all such partial correlations differ from zero, connect the traits by an undirected link (e.g. y_j ⎯⎯ $y_{j'}$). Otherwise, no connecting link is placed between the traits in that pair. After considering all pairs of traits, this step yields an undirected (skeleton) graph.

Step 2: From the skeleton obtained in Step 1, consider every pair of disconnected traits that share a common adjacent trait (e.g. y_j and $y_{j''}$ in y_j ⎯⎯ $y_{j'}$ ⎯⎯ $y_{j''}$) and evaluate all partial residual correlations between the pair that include the common adjacent trait ($y_{j'}$) in the conditioning set. If all such conditional partial correlations differ from zero, direct the edges towards the common adjacent trait (i.e. $y_j \rightarrow y_{j'} \leftarrow y_{j''}$); this identifies a network structure known as an unshielded collider. This step yields a partially directed graph (Pearl, 2009).

Step 3: Based on the partially directed graph obtained from Step 2, orient as many undirected edges as possible without creating any new unshielded colliders nor any cycles. Graphs obtained in Step 3 represent a class of equivalent causal structures (i.e. structures that would result in the same space of joint probability distributions) (Pearl, 2009). In a hierarchical modeling context, the IC algorithm is implemented on the residual (co)variance matrix $\boldsymbol{R}^{*}$ (Valente et al., 2010), such that equivalence classes are defined in the space of residual joint probability distributions.

The inferential basis for each of the statistical decisions made in Steps 1 and 2 of the IC algorithm was the highest posterior density (HPD) interval of the partial residual correlations between traits and whether such intervals contained the value zero. We applied different probability contents of the HPD interval, namely 70, 75, 80, 85, 90, and 95%, to learn the causal structures. Appendix B - Figures B.1 (B.2 and B.3) show the network structures recovered for sows (gilts) across the range of evaluated probability contents of HPD intervals. For sows, when 85, 90 and 95% HPD intervals were applied, no links between any of the reproductive traits were recovered; an 80% HPD interval was the highest probability content that recovered links in the network structure. For gilts, IC steps 1 and 2 using 95, 90 and 85% HPD intervals detected links between some of the reproductive traits (Appendix B - Figure B.2), though GAIN remained disconnected in all cases. An 80% HPD interval was the largest probability content that connected all traits in the gilt dataset (Appendix B - Figure B.3). For consistency of reporting, we chose to proceed with 80% HPD intervals for both sows and gilts in an attempt to balance the probability of missing true connections (Type II error) with the probability of including false ones (Type I error). This is consistent with previous applications (Inoue et al., 2016).

After Step 3 of the IC algorithm, we used additional information to direct any left-over undirected edges, including temporal arrangement of traits and model fit comparisons based on Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). Models that better fit the data are characterized by smaller DIC values, and generally differences of 7 or greater are considered indicative of improved model fit (Spiegelhalter et al., 2002).

3.2.4 Hierarchical Bayesian Implementation and Posterior Inference

The selected SEM specification was fit in a hierarchical Bayesian framework implemented with Markov Chain Monte Carlo (MCMC). The joint posterior distribution of all unknown parameters in the model was:

$$p(\boldsymbol{\beta}, \boldsymbol{b}, \boldsymbol{B}, \boldsymbol{R}, \boldsymbol{\Lambda}, |\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{b}, \boldsymbol{R}, \boldsymbol{\Lambda}) p(\boldsymbol{\beta}) p(\boldsymbol{b}|\boldsymbol{B}) p(\boldsymbol{B}) p(\boldsymbol{R}) p(\boldsymbol{\Lambda}) \quad (3.4)$$

We used conditionally conjugate prior distributions to facilitate Gibbs sampling. Prior distributions for fixed effects location parameters $\boldsymbol{\beta}_j$ and each of the non-zero structural coefficients $\lambda_{jj'}$ in matrix $\boldsymbol{\Lambda}$ were assumed flat such that $p(\boldsymbol{\beta}_j) \propto$ constant and $p(\lambda_{jj'}) \propto$ constant for all j and j'. For each residual variance $\sigma^2_{e_j}$ along the main diagonal of $\boldsymbol{R}$, we generated a proxy for a flat, improper prior using as an instrument the density of a scale-inverse chi-square distribution and setting degrees of freedom v_{e_j} = -1 and scale parameter $s^2_{e_j}$ = 0. This proxy is consistent with the prior $\sqrt{\sigma^2_{e_j}} \sim U(0, A)$, for any finite but sufficiently large value of A, such that the resulting distribution is vague, as recommended for variance components by Gelman (2006). For the block effects $\boldsymbol{b}$, we specified a structural prior such that $p(\boldsymbol{b}|\boldsymbol{B}) \sim N(\boldsymbol{0}, \boldsymbol{B})$ to allow for borrowing of information across traits within each block (i.e. block level covariance parameters), and also across body weight blocks for each outcome (i.e. block level variance parameters) (Robinson, 1991), as is common for random effects in a multivariate mixed models framework. For the (co)variance matrix **B**, we implemented a proxy for a vague prior using as an instrument the density of an inverse-Wishart (IW) distribution, setting degrees of freedom v_B = -(J+1) and specifying a $J \times J$ scale matrix of zeroes; this can be interpreted as a multivariate extension of the proxy used for prior specification on the residual variances.

All MCMC implementations were programmed in R software (R Development Core Team, 2016). For each parity group, the fully recursive SEM of Equation (3.1) was fit using a single MCMC chain run for 100,000 iterations after a burn-in period of 10,000 cycles; one of every 3 samples were saved and fed into the IC algorithm for structure search, as explained in the previous section. For final inference, the selected SEM with a specification of causal structure determined by the IC algorithm was fit using a single MCMC chain run for 300,000 iterations for a burn-in period of 10,000 cycles; one of every 3 samples were saved for inference. In all cases, convergence diagnostics were performed using the R package CODA (Cowles and Carlin, 1996). Specifically, we monitored convergence for all hyperparameters using trace plots and the diagnostic testing approach proposed by Raftery and Lewis (1992). Furthermore, effective sample size (ESS) was estimated to evaluate the number of effectively independent samples amongst the autocorrelated MCMC samples for each hyperparameter (Sorensen et al., 1995). Length of MCMC chains was adjusted to ensure that ESS was greater than 800 for all hyperparameters. For each parameter of interest, we summarized posterior inference using posterior means and 95% HPD intervals.

3.2.5 Assessment of Network Stability

For each parity group, we evaluated the stability of the learned networks using a Jackknife resampling approach adapted from Peñagaricano et al. (2015) for implementation to a network learning context. More specifically, we modified the resampling strategy to leave-one-*block*-out of the dataset at a time (as opposed to leave-one-*observation*-out). In each resampling iteration, the network structure was inferred from a new dataset by implementation of the IC algorithm on a fully recursive SEM specification based on 80% HPD intervals, as described in the previous section. We evaluated the stability of each individual link by assessing the

frequency of presence or absence, as well as direction, in the networks learned from the new datasets obtained from resampling.

3.2.6 Power Assessment in Network Learning

Recall that the size and structure of the datasets available for gilts (i.e. 440 females; 222 blocks) was approximately twice of that available for sows (i.e. 200 females; 97 blocks). Thus, it was of interest to assess whether any discrepancies in network learning might be due, at least partially, to differential power in the statistical decisions made to infer network structure. To evaluate this, we constructed five subsets of the gilt dataset by sampling it at random and without replacement; each gilt data subset mimicked the structure of the sow dataset in number of observations and number of blocks. For each gilt data subset, we learned network structure by implementing the mixed-model-adapted IC algorithm (Valente et al., 2011) as described in previous sections. We then compared the networks learnt from the gilt data subsets with those learnt from the complete datasets for gilts and for sows.

3.3 RESULTS

3.3.1 Descriptive Data Analyses

Descriptive statistics of the reproductive performance traits are presented separately for sows and gilts in Table 3.1. Also, for each parity group, the empirical distribution of each trait is presented on the main diagonal of Figure 3.1, along with pairwise scatterplots and empirical marginal Pearson correlations in the lower and upper triangles of the figure, respectively. Except for WEI, reproductive performance traits showed approximately symmetrical empirical distributions, thereby supporting a normal approximation for modeling. By contrast, the empirical distribution of WEI was skewed to the right for both parity groups.

Pairwise scatterplots and marginal correlations described general similarities for both gilts and sows in the pairwise relationships between traits (Figure 3.1), except for a few distinct patterns. For example, GAIN was positively correlated to TB, BA and BABW in both parity groups but the numerical magnitude of the correlation coefficient appeared to be at least doubled in sows relative to gilts. Furthermore, the marginal Pearson correlation coefficient between WEI and SuTB was significantly different from zero for both groups, though of positive sign for gilts, and negative sign for sows. Overall, preliminary descriptive analyses suggested potential differences between gilts and sows in the network of interconnected reproductive traits. This finding motivated separate network analyses for each parity group.

3.3.2 Learned Network Structures Using the IC Algorithm

For each parity group, network structure was learned separately using the IC algorithm. Based on statistical decisions made on 80% HPD intervals, the network recovered for sows included links only among TB, BA and BABW, but GAIN, WEI and SuTB remained unconnected. Moreover, the IC algorithm did not resolve directionality for any of the links connecting TB, BA and BABW for sows, thus yielding an undirected graph (Figure 3.2A). It was not possible to further orient the links between BABW, BA and TB based on temporal information because, in this study, these traits were realized simultaneously at farrowing. Therefore, for sows, we considered alternative SEMs connecting BABW, BA and TB (Figure 3.2B) that did not introduce any new unshielded colliders (which would presumably have been detected in Step 2 of the IC algorithm) and compare them for fit to data using DIC. Figure 3.2B shows competing SEM alternatives for sows, of which network (3) (i.e. BABW → BA → TB) was selected for further inference based on a DIC value smaller by 20 and 38 points from that of directed networks (1) or (2) respectively.

For gilts, Steps 1 and 2 of the IC algorithm yielded a partially oriented graph depicting an interconnected network that linked all six reproductive performance traits (Figure 3.3A). Of special interest was the link between GAIN and BABW, the detection of which was pivotal to assess directionality in the gilt network because this link formed an unshielded collider with BA (i.e. GAIN → BABW ← BA). So directed, the link GAIN → BABW supported directionality for other links by negating unshielded colliders with TB (i.e. GAIN → BABW → TB) as well as with traits indicative of longer-term reproductive performance, namely WEI (i.e. GAIN → BABW → WEI) and SuTB (i.e. GAIN → BABW → SuTB). These results further oriented the remaining link connecting BA and TB as BA → TB to prevent formation of a cycle, as dictated by the IC algorithm. Overall, the IC algorithm applied to the gilt dataset concluded on an effect of GAIN on BABW (i.e. GAIN → BABW), which in turn affected TB (i.e. BABW → TB). Also detected were effects of BA on TB (i.e. BA → TB) and on BABW (i.e. BA → BABW). Notably, the link connecting BA and BABW (i.e. BA → BABW) prevented TB to function as an unshielded collider between BABW and BA. Further, effects of BABW were detected on subsequent reproductive performance of gilts, as characterized by WEI and SuTB (i.e. BABW → WEI and BABW → SuTB, respectively). Finally, a link between WEI and SuTB was detected in gilts, though the IC algorithm was inconclusive about directionality. We used temporal information to tentatively orient the link as WEI → SuTB and relied on DIC to further evaluate this model choice. The difference in DIC between a learned network presuming the selected direction (i.e. WEI → SuTB) versus the same network with a reversed link (i.e. WEI ← SuTB) was smaller than 1 point (i.e. DIC values were 11974.86 and 11974.09, respectively). Such small DIC difference suggested that there was no information available in the data to select one

network over the other. As a result, we relied solely on qualitative information (e.g. temporal arrangement of the traits) to inform directionality of this link (Figure 3.3B).

3.3.3 SEM-based Inference

Figure 3.4 shows the causal network structures selected for final inference on the potentially causal relationships between reproductive performance traits of (A) sows and (B) gilts. Table 3.2 shows posterior summaries of the corresponding structural coefficients representing direct effects between traits.

For sows, BABW was found to have a direct effect on BA (i.e. BABW → BA) by which every 100 g increase in average BABW resulted in an estimated decrease in BA of about (posterior mean [95% HPD]) 0.58 [0.41, 0.77] piglets. Subsequently, every additional piglet BA increased the TB in a sow litter (i.e. BA → TB) by approximately 0.96 [0.89, 1.02] piglets.

For gilts, BA had a direct effect on both TB and BABW. Specifically, every unit increase in BA resulted in an increase in TB of 0.89 [0.85, 0.94] piglets and in an estimated decrease of about 360 g [240, 480] in BABW. Moreover, in gilts, BABW also had a direct effect on TB and on SuTB. In particular, a l00 g increase in average BABW resulted in an estimated decrease of 0.11 [0.05, 0.17] in TB and of 0.42 [0.20, 0.65] in SuTB. Finally, WEI had a direct effect on SuTB, whereby a 1-day increase in WEI resulted in an estimated increase in SuTB of 0.16 [0.09, 0.24] piglets per litter. Additional direct effects yielded by the network learning process are presented in Table 3.2 but are not discussed further, as the 95% HPD interval for the corresponding structural coefficients overlapped with the null value zero.

For completeness of network interpretation, we also report indirect effects, which under linearity assumptions can be computed as the product of the corresponding structural coefficients, as described by Shipley (2002). For example, for sows, the indirect, BA-mediated,

effect of BABW on TB (i.e. BABW → BA → TB) was estimated as the product of the corresponding direct effects, namely $\lambda_{TB,BA} \times \lambda_{BA,BABW}$, the posterior density of which had a mean of -0.56 TB per 100 g increase in BABW with 95% HPD = [-0.74, -0.39]. Indirect effects for gilts are presented in Table 3.3 and can be interpreted in a similar manner. For gilts, most indirect effects had 95% HPD intervals that overlapped with zero, except for the BABW-mediated effect of BA on TB (i.e. BA → BABW → TB) and of BA on SuTB (i.e. BA → BABW → SuTB) (Table 3.3).

Finally, Table 3.3 also shows total effects, reported as the sum of the direct effect and all indirect effects connecting two traits, as described by Shipley (2002). For example, the total effect of BA on TB is obtained as $(\lambda_{TB,BABW} \times \lambda_{BABW,BA}) + \lambda_{TB,BA}$ and had a posterior mean of 0.94 TB per additional BA with 95% HPD = [0.89, 0.98].

3.3.4 Assessment of Network Stability

The stability of the learned network was assessed using a Jackknife resampling approach modified to leave-one-block-out at each resampling iteration. Figure 3.5 shows results of network stability for (A) sows and (B) gilts at the level of individual links and expressed as the the percent of resampled datasets (i.e. 97 for sows, 222 for gilts) for which each link was present in the learned network structure.

For sows, the link between TB and BA was recovered in 100% of the resampled datasets, whereas the link between BA and BABW appeared in 92% of the cases. The reproductive traits GAIN, WEI and SuTB remained unconnected to any other nodes in all cases. Notably, no link directionality was recovered from any of the resampled datasets. Overall, the network structure learned for sows was considered stable, though not so the directionality of the links.

For gilts, most links connecting reproductive traits showed substantial stability as their presence ranged from 99% to 100% of the resampled datasets (Figure 3.5B). A notable exception was the link connecting GAIN and BABW, which was present in only 54% of the cases. This link showed the least stability in the gilt network, as removal of a single block negated its presence in almost half (101 out of 222) of the resampled datasets. Recall that the link between GAIN and BABW was deemed pivotal for assessing directionality of effects within the gilt network; specifically, this link informed an unshielded collider with BA (i.e. GAIN → BABW ← BA), while negating any unshielded colliders between GAIN and other non-adjacent traits, namely TB, WEI and SuTB. Similar to sows, the general network structure of reproductive traits in gilts was deemed stable, though the evidence for directionality of effects was considered weak.

3.3.5 Relative Power Differences in Network Learning

Figure 3.6 shows the network structures learned from each of the five gilt data subsets that were sampled to mimic the sow dataset in size and structure. First, we note that no unshielded colliders were detected in any of the gilt data subsets; as a result, it was not possible to learn any directionality of individual links from any dataset smaller than the complete gilt dataset. This suggests a relative sample-size-based power differential between parity groups for learning directionality within the network; admittedly, thought, results on directionality were weak for both gilts and sows. Second, the network structure learned from each of the five gilt data subsets showed a number of links ranging from three to five (Figure 3.6), in contrast to only two links learned from the complete sow data and seven links learned from the complete gilt dataset. The intermediate number of links detected from the gilt data subsets suggests non-power-based differences in network structure, presumably of a biological basis, to partially

explain the parity difference in links recovered. In particular, all five gilt data subsets consistently identified a specific link that was not apparent from the sow dataset, namely the link between WEI and SuTB.

3.4 DISCUSSION

In this study, we investigated potentially causal biological relationships among reproductive performance traits in experimental data from high-performing gilts and sows. To accomplish this, we used a network approach implemented with structural equation models and the IC algorithm to search the structure space of interrelationships between reproductive traits. we followed recent methodological developments that extended networks to a mixed model framework (Gianola and Sorensen, 2004; Valente et al., 2010) in order to accommodate the inherently hierarchical nature of experimental data. The analytic approach of hierarchical network modeling formalizes integration from a systems perspective, thus enabling more comprehensive insight into the complex mechanisms that underlie animal production.

Consistent with the distinct reproductive physiology of sows and gilts (Da Silva et al., 2016; Da Silva et al., 2017), our results indicate disparate network connectivity across parity groups. Gilts showed a densely interconnected network between all six reproductive performance traits considered in this study. Notably, direct effects connected reproductive events starting at first gestation (i.e. GAIN) through first farrowing (i.e. BABW, TB and BA) and into the subsequent reproductive cycle (i.e. WEI, SuTB). By contrast, sows showed a sparse network depicting only connections within the immediate farrowing event (i.e. BABW, BA and TB), but without any temporal connections with earlier or later stages of the reproductive lifecycle (i.e. GAIN, WEI and SuTB were completely disconnected from the rest of the traits). This differential network connectivity between parity groups may be partially explained by the fact that sows are

mature adults, whereas gilts are immature animals still growing after the first farrowing event and into the first lactation (Kraeling and Webel, 2015). As such, gilts are subjected to competing requirements for growth and reproduction, which can be expected to impact subsequent reproductive performance. Indeed, the high level of network connectivity observed in gilts suggests potential propagating ripple effects across the physiological network and productive lifecycle. This is consistent with the current state of the literature, which indicates that performance of parity one sows (which are referred to as gilts in this study) is a critical contributing factor to subsequent reproductive performance and ultimately, sow longevity in the herd (Mabry et al., 1996; Rozeboom et al., 1996). Indeed, our results show direct effects and total effects connecting stages of the gilt lifecycle (e.g. BABW and SuTB), suggesting longer-term impact of early reproductive events.

In contrast, the sparser network recovered in sows suggests a mitigated potential for long-term propagation of intervention effects once females have reached maturity and adult size. One may interpret this result as an opportunity for compounding long-term beneficial effects in response to early managerial interventions. As such, special attention and tailored management of gilts during the first farrowing and lactation seems warranted. Undoubtedly, modern swine females across the parity spectrum require continuous updating of nutrient requirements to maximize productive efficiency and are highly responsive to management practices and technologies (Boyd et al., 2002; Kraeling and Webel, 2015). Specially so for young females from modern maternal lines, as they seem to be particularly sensitive to finely-tuned managerial practices because of a lower appetite and exceptional lean growth potential despite large competing nutritional demands from gestation (Bortolozzo et al., 2009). This is consistent with current recommendations for tailored management of dietary nutrient density and feeding intake

of gilts during first lactation, as this may enhance their long-term reproductive performance and longevity in the production system (Kraeling and Webel, 2015). For example, tailored feeding strategist such as segregated phase feeding is often recommended (Kraeling and Webel 2015) to ensure feed intake is matched with nutrient requirements of younger gilts still undergoing growth.

In addition, the influence of season and environmental conditions on reproductive performance of swine females is well documented (Kraeling and Webel 2015). Gilts seem to be most vulnerable to seasonal infertility in response to extreme temperature and disruptions in photoperiod, with reported consequences of delayed estrus after weaning and disrupted estrus behavior, amongst others (Britt et al., 1983; Tummaruk et al., 2004; Auvigne et al., 2010) , as consistent with results from this study. This is relevant to our discussion because data collection for this study was conducted during summer months. Due to enhanced sensitivity, gilts might also be expected to respond to tailored modulations of the ambient environment (e.g. cooling devices, group size, equipment and facilities) and to dietary management of nutrient density (McGlone et al., 1988; McGlone et al., 2004; Kraeling and Webel, 2015).

As we continue to ponder connectivity of the reproductive network in sows and gilts, it is worth mentioning that all links recovered from the sow dataset and most of those from the gilt dataset were validated as inferentially stable using Jacknife resampling. Moreover, a power assessment acknowledged a potential biological basis for differences in network connectivity. Admittedly, differences in sample size between the parity groups may underlie and partially explain some of the observed differences, as the gilt dataset was almost twice as large as the sow dataset. Yet, the power assessment recovered more links (i.e. ranging from 3 to 5; Figure 3.6) in every one of the five gilt data subsets compared to the 2 links recovered from the complete sow

data, suggesting additional non-power related (presumably biological) reasons for the observed network discrepancies between parity groups. Worth highlighting is the link connecting WEI and SuTB, both indicators of reproductive performance in the subsequent gestation. This link was consistently recovered from all 5 gilt data subsets as well as from the complete gilt dataset but was not apparent in sows. This may be partially explained by the fact that, in the post-weaning period, gilts are typically less effective in the recovery of mobilized tissue and body condition score compared to sows and tend to redirect energy and nutrients towards ensuring mature size growth over reproductive functions (Rempel et al., 2015), thus delaying estrous and rebreeding. It is then not unexpected that subsequent reproductive performance of gilts is more sensitive to WEI than that of sows. For this reason, maximizing nutrient intake during lactation seems to be especially important in gilts.

The network-type SEM approach implemented in this study raises the issue of statistical power as an important point in need of serious consideration for proper implementation of these methods in the context of animal agriculture. Recall that the swine datasets used in this study might be considered substantial in size for a designed experiment (i.e. > 3,800 observations from 640 females). Yet, this experiment was arguably not designed for network analysis and the dataset may still not be big enough to have adequate power for learning network structure or assessing network differences. In particular, inference on directionality of links seems to be specially sensitive to power. For illustration, consider the sow dataset (the smallest of the two datasets), consisting of a total of 1,200 observations collected from 200 females (i.e. 6 traits × 200 sows), for which the IC algorithm failed to assign direction to any of the links. Further, Jacknife resampling failed to identify direction for any link in sows, such that link orientation for final inference in this parity group was based on relative model fit to the observed data. Even

with the complete gilt dataset, which consisted of a total of 2,640 observations collected from 440 females (i.e. 6 traits × 440 gilts), directionality of links turned out to be the weakest of our results. Specifically, recall that directionality of the gilt network hinged on the link recovered between GAIN and BABW, which in turn oriented most other links in the network by detecting (or negating) the role of BABW as an unshielded collider with other nodes. Yet, the link between GAIN and BABW was not inferentially stable, as it was detected in only 54% of the Jacknife resampling folds. Overall, it appears that even larger datasets with informative structures will be needed to ensure sufficient statistical power for network learning in order to reliably explore how traits affect each other in an integrated system.

In pork production systems, an important aspect of profitability is the number of piglets produced per sow per year (Kraeling and Webel, 2015). To this end, swine breeding programs have emphasized selection for increased litter sizes (Johnson et al., 1999) over the past few decades. Larger litter sizes have been associated with lower birth weight (Town et al., 2004), which in turn has implications for pre-weaning mortality rate and further piglet development (Milligan et al., 2002). This is likely due to compromised placental development in the crowded uterus (Père et al., 1997) and associated competition for nutrients and space throughout gestation (Geisert and Schmitt, 2002). Indeed, piglets of lower average birth weight are weaker and less likely to survive, thus posing economic and welfare concerns. Our findings are conceptually consistent with this understanding, particularly regarding the interrelationships between piglet birth weight and survivability. Specifically, BABW as a node played a key role in both sow and gilt networks; further, the structural coefficients connecting BABW and BA showed negative signs and thus supported inverse relationships between the traits in both parity groups.

Also, both parity groups showed a connection between litter size and piglet survivability, as depicted by the recovery of a link connecting BA and TB. Our next discussion emphasizes presence, rather than direction, of this link, as results on directionality were considered weak in this study, as discussed earlier. Piglet mortality at birth seems to be a multifactorial problem, with contributing factors including litter size and parity, amongst others (Vanderhaeghe et al., 2013). While not a formal comparison, we observed seemingly disparate magnitudes of the association between BA and TB in gilts and in sows. Specifically, in sows, every unit increase in BA resulted in a larger litter size by approximately one unit, as the 95% HPD of $\lambda_{TB,BA}$ = [0.89, 1.02] contained the value one. For gilts, a piglet increase in BA translated into a proportionally smaller increase in litter size, as the corresponding 95% HPD (i.e. [0.85, 0.94]) was below the value one. This could be partially explained by physiological differences in uterine capacity between the parity groups (Ford et al., 2002), and the resulting timing of embryonic mortality due to uterine overcrowding (Foxcroft et al., 2009). Uterine capacity is critical for the development and survivability of the pig conceptus (Chen and Dziuk, 1993) due to its influence on placental growth and thus, supply of nutrients to the conceptus (Ford et al., 2002). Swine females in general, but particularly gilts, typically have limited uterine capacity relative to their ovulation rate (Da Silva et al., 2016). In gilts, this precipitates uterine crowding and conceptus losses in earlier stages of gestation (Da Silva et al., 2017). Embryos that die before d 30-35 of gestation are often reabsorbed, but any fetuses that are lost after d 30-35 are accounted for at farrowing either as mummies or stillborn piglets (Foxcroft et al., 2009). Therefore, our results can be explained by gilts having higher embryonic loss in the early gestation period (Da Silva et al., 2017), thus unobservable at birth; whereas sows with higher uterine capacity may be able to

maintain fetal development longer into the pregnancy, thereby resulting in proportionally higher losses later state of gestation, which are then observable at birth.

Identification of parity differences in the magnitude of effects along the reproductive network is arguably of interest to inform tailored management and efficient decision making of each category of animals. In this study, each parity groups was analyzed separately such that it is not possible to conduct formal tests for parity differences on the structural coefficients connecting, say, BA and TB. Yet, separate analyses were inevitable in this case because of the standard assumption of SEMs that structural coefficients be homogeneous across the population (Gianola and Sorensen, 2004), thereby curtailing formal testing, even if a joint analysis of sows and gilts were conducted. Methodological developments are warranted to explicitly allow specification of sources of heterogeneity on structural coefficients in SEM. Additional methodological extensions relevant to SEMs in animal agriculture include heterogeneity of variance parameters, as a way to stabilize the variance of traits with skewed distributions, such as WEI in this study. Here, we considered a variance-stabilizing transformation on WEI, though this was not pursued because it is unclear what the implications of doing so might be for simultaneous modeling of WEI as a response and as a predictor, as well as for interpretation of indirect and total effects given linearity assumptions of SEM. The inferential implications of SEM assumptions of normality and homogeneous variances for all traits requires further study.

Specification of the network structure is critical to SEM. The IC algorithm is one of many strategies available to conduct a data-driven search for network structure and directionality compatible with the joint probability distribution of the data (Spirtes et al., 1993). In a rather informal follow-up step to the IC algorithm, a researcher can incorporate additional information to assist with remaining undirected links and thus differentiate between models within an

equivalence class. Admittedly, such incorporation of information is subject-matter-based and rather *ad-hoc* in nature, and thus, does not guarantee causality. For example, in this study, the direction of the link connecting WEI and SuTB in gilts was not resolved by the IC algorithm. To fully orient the gilt network, we incorporated temporal information and set WEI → SuTB, as later events cannot have an effect on earlier events. Another type of information that can be used to assist in orienting left-over undirected links in the context of structured data is an indicator of model fit. Here, we used DIC (Spiegelhalter et al., 2002) to select between SEM specifications compatible with the undirected graph yielded by the IC algorithm for the sow dataset, as shown in Figure 3.2B. These three SEM specifications correspond to the same equivalence class defined at the residual level by the mixed-models adapted IC algorithm (Valente et al., 2010). In this adaptation of the IC algorithm, the residual-level equivalence class is defined after fitting random effects to account for correlation patterns in the data due to experimental design or other structural components of the data, such as genetic relationships (Inoue et al., 2016). By contrast, recall that the DIC is an indicator of model fit computed based on the deviance statistic and as such, uses the joint likelihood of the observed data (as opposed to that of residuals). That is, the mixed-models-adapted IC algorithm searches for network structure on the residual (co)variance matrix $\boldsymbol{R}^*$ (rather than on the (co)variance matrix of the observed data $var(\boldsymbol{y}) = \boldsymbol{Z}\boldsymbol{B}^*\boldsymbol{Z} + \boldsymbol{R}^*$, for $\boldsymbol{B}^* = (\boldsymbol{I} - \boldsymbol{\Lambda})^{-\mathbf{1}}\boldsymbol{B}(\boldsymbol{I} - \boldsymbol{\Lambda})^{-\mathbf{1}'}$), thus yielding a class of equivalent models defined at the residual level (Valente et al., 2010). Differences in goodness of data fit between otherwise residual-level equivalent models should then not be surprising, provided there is structure of distribution of random effects; in fact, DIC differences are to be anticipated in such case. Indeed, discrepancies in the selection of SEM structure from DIC relative to the mixed-models adapted IC algorithm were also reported by Inoue et al. (2016). Specifically for this study, data

architecture proved useful in the implementation of DIC to differentiate between competing sow networks (Figure 3.2A), as indicated by DIC differences of 20 points or more between competing SEM specifications.

In using the IC algorithm, it is important to assess stability of the learned network structures before proceeding further with inference. Following Peñagaricano et al. (2015), we adapted Jacknife resampling to network learning and implemented the sampling strategy so as to leave-one-*block*-out at a time instead of the more traditional leave-one-*observation*-out. This was intended as a trade-off for computational efficiency in the assessment of network stability. This approach is consistent with Dórea et al. (2018), who used a machine-learning strategy to leave-one-*trial*-out of the dataset at a time when validating feed intake predictions in lactating dairy cows. So adapted, Jacknife resampling might be considered a validation strategy in a broader scope of inference (Bello and Renter, 2018) across the population for which the body weight blocks used in this study (or trials used by Dórea et al. (2018)) might be considered a representative, if not random, sample.

Admittedly, in and of themselves, the statistical methodology that underlies SEM does not guarantee causal claims. As an example, we refer the reader to the inferential weaknesses described for this study in the learning of directionality of direct effects. In designed experiments, the conclusion of a causal effect of a treatment of interest is supported by randomization, which is considered the gold standard for causal inference (Bello et al., 2018). It is often overlooked, though, that such randomization-based causal claims are restricted to the effect of treatment on a response variable of interest; yet, randomization does not support causality of one response variable on another, as is of interest for the type of causal phenotypic networks that motivated this study. That is, even in designed experiments, causal relationships

between response variables of interest may be confounded by other variables (either observed or unobserved) or even by correlations induced by the design structure (i.e. within-block correlations). It is for this reason that causal identification requires that non-trivial causal assumptions be made (Pearl, 2009), even in the context of designed experiments. Briefly, these assumptions comprise the Markov condition, the assumption of faithfulness or stability, and causal sufficiency. The latter assumption, causal sufficiency, is probably the most delicate one because it implies that any confounders in the system are known and either have been measured or can be controlled for by other measured variables. Importantly, these causal assumptions are not directly testable from data and yet are fundamental for further inference, as explained by Bello et al. (2018). Despite these limitations, we strongly believe that the framework of causal inference offers promising practical advantages for scientific progress (Bello et al., 2018). For example, insight gained from causal inference can help subject-matter scientists identify and refine research hypotheses that, ideally, could then be tested in tailored randomized experiments conducted in-vivo or in-vitro. From a study design standpoint, causal inference can further help elucidate what set of variables should be measured and which others might be redundant, thereby enhancing efficiency in the allocation and use of research funding. More broadly, causal inference undeniably contributes to the growing body of scientific evidence on which science advances.

3.5 CONCLUSION

This study provides insight into potentially distinct causal networks of interconnected reproductive performance traits in high-producing sows and gilts, based on structural equation models and the IC algorithm adapted to mixed models. Evidence suggests distinct networks for these parity groups, consistent with differences in their reproductive physiology, thereby

substantiating tailored reproductive management specific for each category of animals. Further investigation of network connectivity, directionality and sources of heterogeneity is warranted to improve integrated understanding and efficient management of swine production systems.

ACKNOWLEDGEMENTS

Computing for this project was partially performed on the Beocat Research Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860 and EPS-0919443. This research was also performed using the computing resources and assistance of the UW-Madison Center for High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison, the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

REFERENCES

Auvigne, V., P. Leneveu, C. Jehannin, O. Peltoniemi, and E. Salle. 2010. Seasonal infertility in sows: a five year field study to analyze the relative roles of heat stress and photoperiod. Theriogenology. 74: 60-66. doi:10.1016/j.theriogenology.2009.12.019

Bello, N. M., V. C. Ferreira, D. Gianola, and G. J. M. Rosa. 2018. Conceptual framework for investigating causal effects from observational data in livestock. J. Anim. Sci. 96:4045-4062. doi:10.1093/jas/sky277

Bello, N. M., and D. G. Renter. 2018. Invited review: Reproducible research from noisy data: Revisiting key statistical principles for the animal sciences. J. Dairy Sci. 101:5679-5701. doi:10.3168/jds.2017-13978

Bortolozzo, F. P., M. L. Bernardi, R. Kummer, and I. Wentz. 2009. Growth, body btate and breeding performance in gilts and primiparous sows. Soc. Reprod. Fertil. Suppl. 66: 281-291.

Boyd, D. R., G. C. Castro and R. A. Cabrera. 2002. Nutrition and management of the sow to maximize lifetime productivity. Adv. Pork Prod. 13: 47-59.

Britt, J. H., V. E. Szarek, and D. G. Levis. 1983. Characterization of summer infertility of sows in large confinement units. Theriogenology 20: 133-140.

Chen, Z. Y., and P. J. Dziuk. 1993. Influence of initial length of uterus per embryo and gestation stage on prenatal survival, development, and sex ratio in the pig. J. Anim. Sci. 71: 1895-1901. doi:10.2527/1993.7171895x

Cowles, M. K., and B. P. Carlin. 1996. Markov chain Monte carlo convergence diagnostics: A comparative review. J. Am. Stat. Assoc. 91:883-904. doi:10.1080/01621459.1996.10476956

Da Silva, C. L. A., B. F. A. Laurenssen, E. F. Knol, B. Kemp, and N. M. Soede. 2017. Validation of transrectal ultrasonography for assessment of corpora lutea characteristics in pregnant sows and its relationship with litter characteristics at birth. Transl. Anim. Sci. 1: 507-517. doi:10.2527/tas2017.0057

Da Silva, C. L. A., H. van den Brand, B. F. A. Laurenssen, M. L. W. L. Broekhuijse, E. F. Knol, B. Kemp, and N. M. Soede. 2016. Relationships between ovulation rate and embryonic and placental characteristics in multiparous sows at 35 days of pregnancy. Animal. 10:1192-1199. doi:10.1017/S175173111600015X

de los Campos, G., D. Gianola, P. Boettcher, and P. Moroni. 2006. A structural equation model for describing relationships between somatic cell score and milk yield in dairy goats. J. Anim. Sci. 84:2934-2941. doi:10.2527/jas.2006-016

de Maturana, E. L., X. L. Wu, D. Gianola, K. A. Weigel, and G. J. M. Rosa. 2009. Exploring biological relationships between calving traits in primiparous cattle with a Bayesian recursive model. Genetics 181: 277-287. doi:10.1534/genetics.108.094888

Dórea, J. R. R., G. J. M. Rosa, K. A. Weld, and L. E. Armentano. 2018. Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows. J. Dairy Sci. 101:5878-5889. doi:10.3168/jds.2017-13997

Ford, S. P., K. A. Vonnahme, and M. E. Wilson.. 2002. Uterine capacity in the pig reflects a combination of uterine environment and conceptus genotype effects. J. Anim. Sci. 80(E. Suppl. 1): E66-E73. doi:10.2527/animalsci2002.0021881200800ES10010x

Foxcroft, G. R., W. T. Dixon, M. K. Dyck, S. Novak, J. C. Harding, and F. C. Almeida. 2009. Prenatal programming of postnatal development in the pig. Soc. Reprod. Fertil Suppl. 66:213-231.

Geisert, R. D., and R. A. M. Schmitt. 2002. Early embryonic survival in the pig: Can it be improved? J. Anim. Sci. 80: E54-E65. doi:10.2527/animalsci2002.0021881200800ES10009x

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 1:515-533. doi:10.1214/06-BA117A

Gianola, D., and D. Sorensen. 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. Genetics. 167:1407-1424. doi: 10.1534/genetics.103.025734

Gonçalves, M. A. D., K. M. Gourley, S. S. Dritz, M. D. Tokach, N. M. Bello, J. M. DeRouchey, J. C. Woodworth, and R. D. Goodband. 2016. Effects of amino acids and energy intake during late gestation of high-performing gilts and sows on litter and reproductive performance under commercial conditions. J. Anim. Sci. 94:1993-2003. doi: 10.2527/jas.2015-0087

Haavelmo, T. 1943. The Statistical Implications of a System of Simultaneous Equations. Econometrica 11:1-12. doi:10.2307/1905714

Henderson, C. R., and R. L. Quaas. 1976. Multiple Trait Evaluation Using Relatives Records. J. Anim. Sci. 43:1188-1197. doi:10.2527/jas1976.4361188x

Inoue, K., B. D. Valente, N. Shoji, T. Honda, K. Oyama, and G. J. M. Rosa. 2016. Inferring phenotypic causal structures among meat quality traits and the application of a structural equation model in Japanese Black cattle1. J. Anim. Sci. 94:4133-4142. doi: 10.2527/jas.2016-0554

Johnson, R. K., M. K. Nielsen, and D. S. Casey. 1999. Responses in ovulation rate, embryonal survival, and litter traits in swine to 14 generations of selection to increase litter size. J Anim Sci 77: 541-557.

Kraeling, R. R., and S. K. Webel. 2015. Current strategies for reproductive management of gilts and sows in North America. J. Anim. Sci. Biotechnol. 6:3. doi:10.1186/2049-1891-6-3

Mabry, J., M. S. Culbertson, and D. Reeves. 1996. Effects of lactation length on weaning-to-first-service interval, first-service farrowing rate, and subsequent litter size. Swine Health Prod. 4: 185-188

McGlone, J. J., W. F. Stansbury, and L. F. Tribble. 1988. Management of lactating sows during heat stress: effects of water drip, snout coolers, floor type and a high energy-density diet. J. Anim. Sci. 66: 885-891.

McGlone, J. J., E. H. von Borell, J. Deen, A. K. Johnson, D. G. Levis, M. Meunier-Salaon, J. Morrow, D. Reeves, J. L. Salak-Johnson, and P. L. Sundberg. 2004. Reviews: Compilation of the scientific literature comparing housing systems for gestating sows and gilts using measures of physiology, behavior, performance, and health. Prof. Anim. Scient. 20:105-117. doi:10.15232/S1080-7446(15)31285-7

Milligan, B. N., D. Fraser, and D. L. Kramer. 2002. Within-litter birth weight variation in the domestic pig and its relation to pre-weaning survival, weight gain, and variation in weaning weights. Livest. Prod. Sci. 76: 181-191. doi:10.1016/S0301-6226(02)00012-X

Pearl, J. 2009. Causality: Models, Reasoning, and Inference. 2nd ed. Cambridge University Press, New York.

Peñagaricano, F., B. D. Valente, J. P. Steibel, R. O. Bates, C. W. Ernst, H. Khatib, and G. J. M. Rosa. 2015. Exploring causal networks underlying fat deposition and muscularity in pigs through the integration of phenotypic, genotypic and transcriptomic data. BMC Syst. Biol. 9:58. doi:10.1186/s12918-015-0207-6

Père, M.-C., J.-Y. Dourmad, and M. Etienne. 1997. Effect of number of pig embryos in the uterus on their survival and development and on maternal metabolism. J. Anim. Sci. 75: 1337-1342.

Raftery, A., and S. Lewis. 1992. How many iterations in the Gibbs sampler?. Bayesian Stat. 4:763-773. doi:10.1.1.41.6474

Robinson, G. K. 1991. That BLUP is a Good Thing: The Estimation of Random Effects. Stat. Sci. 6:15-32. doi:10.1214/ss/1177011926

Rozeboom, D. W., J. E. Pettigrew, R. L. Moser, S. G. Cornelius, and S. M. El Kandelgy. 1996. Influence of gilt age and body composition at first breeding on sow reproductive performance and longevity. J. Anim. Sci. 74: 138-150.

R Development Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Shipley, B. 2002. Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference.1st ed. Cambridge University Press, New York.

Sorensen, D. A., S. Andersen, D. Gianola, and I. Korsgaard. 1995. Bayesian Infernce in Threshold Model using Gibbs Sampling. Genet. Sel. Evol. 27:229-249. doi:10.1051/gse:19950303

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde. 2002. Bayesian measures of model complexity and fit. J. Royal Stat. Soc. 64(Series B): 583-639. doi:10.1111/1467-9868.00353

Spirtes, P., Glymour, C., and Scheines, R. 1993. Causation, Prediction, and Search. Springer-Verlag, New York, NY.

Town, S. C., Putman, C. T., Turchinsky, N. J., Dixon, W. T., and Foxcroft, G. R. 2004. Number of conceptuses in utero affects porcine fetal muscle development. Reproduction. 128: 443-454.

Tummaruk, P., W. Tantasuparuk, M. Techakumphu, and A. Kunavongkrit, 2004. Effect of season and outdoor climate on litter size at birth in purebred landrace and yorkshire sows in Thailand. J. Vet. Med. Sci. 66: 477-482.

Valente, B. D., G. J. M. Rosa, G, M. A. Silva, R. B. Teixeira, and R. A. Torres. 2011. Searching for phenotypic causal networks involving complex traits: an application to European quail. Genet. Sel. Evol. 43: 37. doi:10.1186/1297-9686-43-37

Valente, B. D., G. J. M. Rosa, G. de Los Campos, D. Gianola, and M. A. Silva. 2010. Searching for recursive causal structures in multivariate quantitative genetics mixed models. Genetics. 185:633-644. doi:10.1534/genetics.109.112979

Van Vleck, L. D., and K. M. Edlin. 1984. Multiple Trait Evaluation of Bulls for Calving Ease. J. Dairy Sci. 67:3025-3033. doi: 10.3168/jds.S0022-0302(84)81668-9

Varona, L., and D. Sorensen. 2014. Joint Analysis of Binomial and Continuous Traits with a Recursive Model: A Case Study Using Mortality and Litter Size of Pigs. Genetics. 196:643-651. doi:10.1534/genetics.113.159475

Verma, T., and J. Pearl. 1991. Equivalence and synthesis of causal models Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence. Elsevier Science Inc, New York. p. 255-270.

Wettemann, R. P., and Bazer, F. W. 1985. Influence of environmental temperature on prolificacy of pigs. J. Reprod. Fertil. Suppl. 33: 199-208.

Wright, S. 1934. An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs. Genetics 19:506-536.

Wu, X. L., B. Heringstad, Y. M. Chang, G. de Los Campos, and D. Gianola. 2007. Inferring relationships between somatic cell score and milk yield using simultaneous and recursive models. J. Dairy Sci. 90:3508-3521. doi:10.3168/jds.2006-762

FIGURES AND TABLES

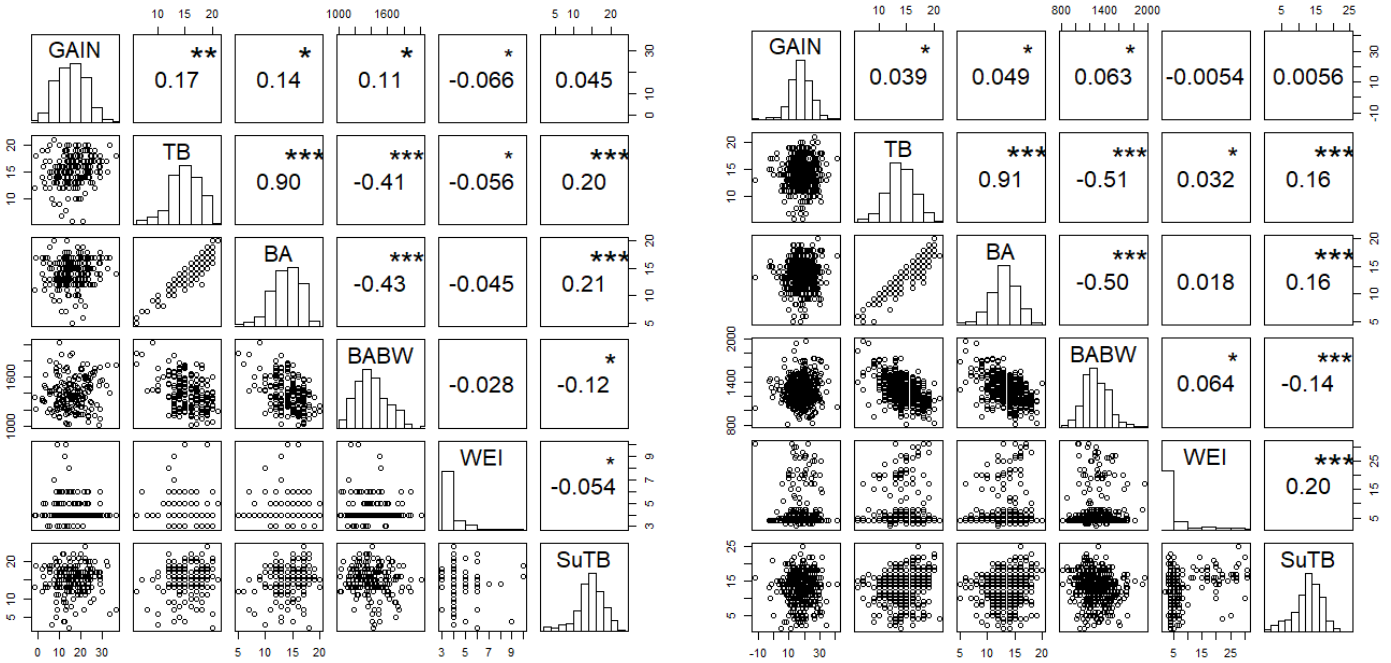


Figure 3.1 Scatterplot matrices depicting the empirical distributions and marginal associations between reproductive performance traits in sows (left panel) and in gilts (right panel). For each panel, histograms are shown along the main diagonal whereas bivariate scatterplots are shown on the lower triangle and estimated Pearson correlation coefficients are presented on the upper triangle. Asterisks indicate that the coefficient is significantly different from zero (*** *P*-value < 0.0001, ** *P*-value < 0.01, * *P*-value < 0.05). GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

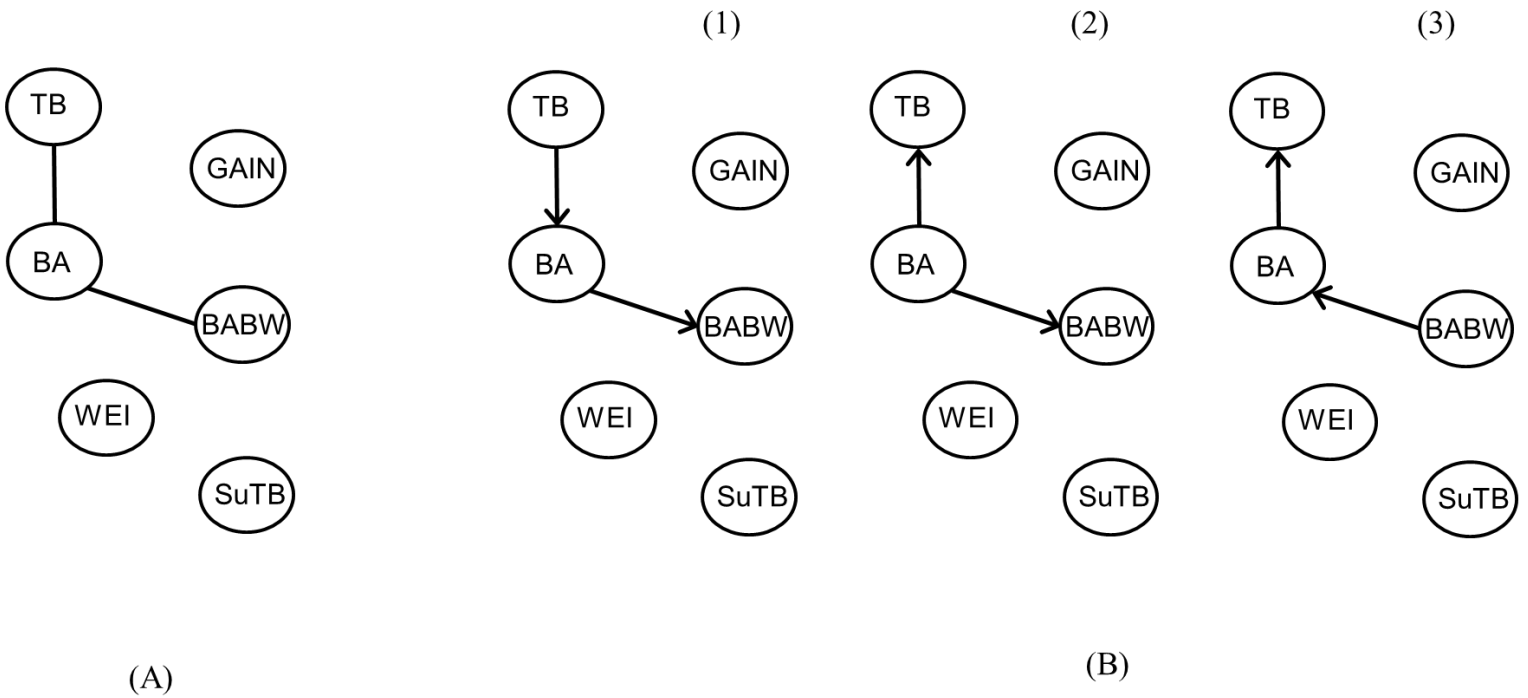


Figure 3.2 (A) Undirected graph of reproductive performance traits in sows detected by the inductive causation algorithm implemented with 80% highest posterior density intervals. (B) Plausible structures within the equivalence class defined by traits connected in Panel (A). Links without arrowheads represent associations between traits; links with arrowheads represent causal effects from the trait on the arrowtail to the trait on the arrowhead. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

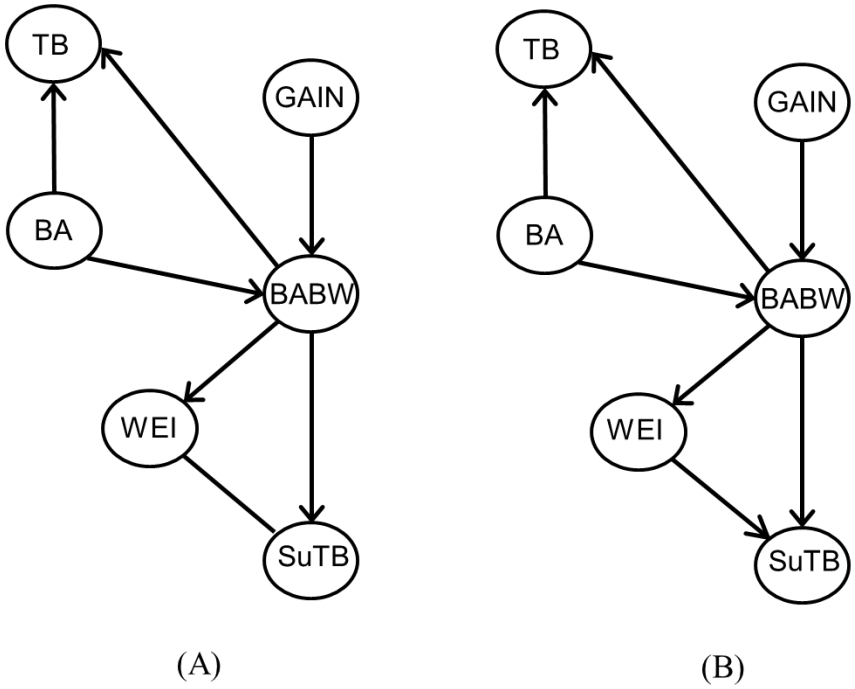


Figure 3.3 (A) Partially oriented graph of reproductive performance traits in gilts detected by the inductive causation algorithm implemented with 80% highest posterior density intervals. (B) Fully oriented graph obtained after incorporating additional temporal information to (A). Links without arrowheads represent associations between traits; links with arrowheads represent causal effects from the trait on the arrowtail to the trait on the arrowhead. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

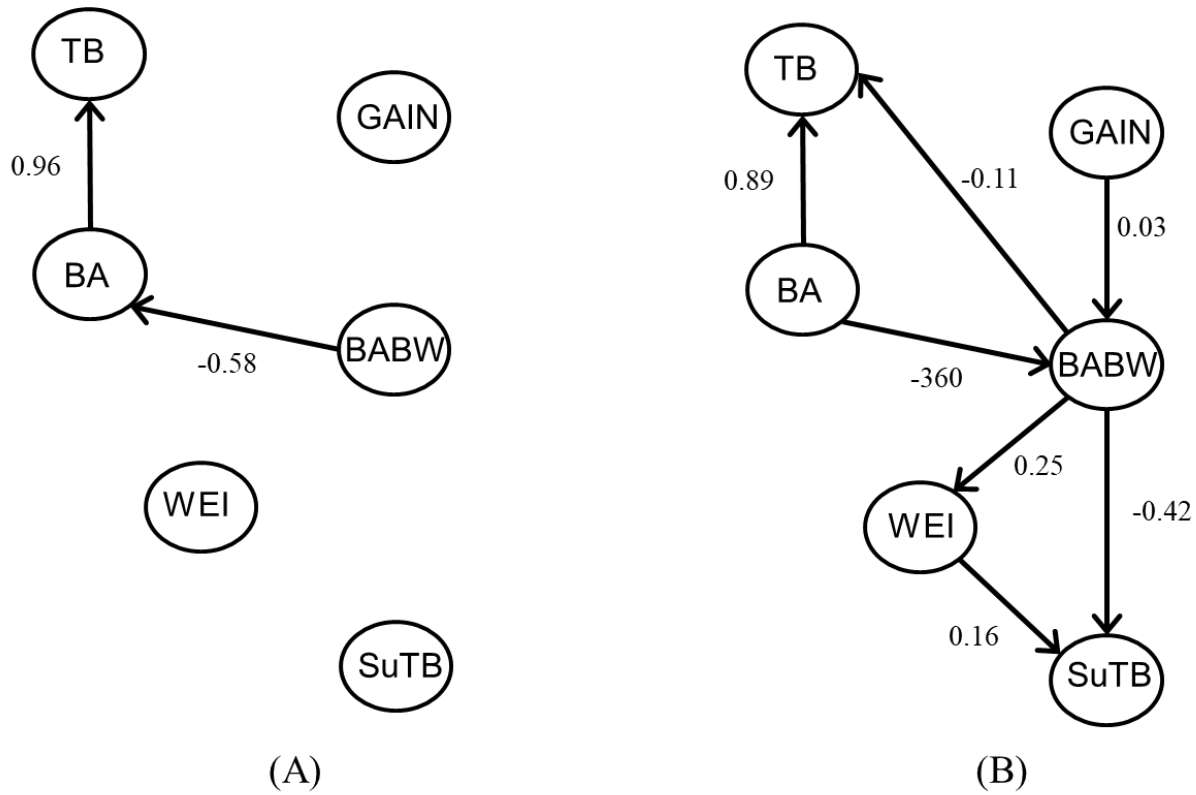


Figure 3.4 Links and posterior means of structural coefficients between reproductive performance traits in sows (A) and in gilts (B) learned using a mixed-models adapted inductive causation algorithm implemented with 80% highest posterior density intervals. Refer to Table 2 for further details. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

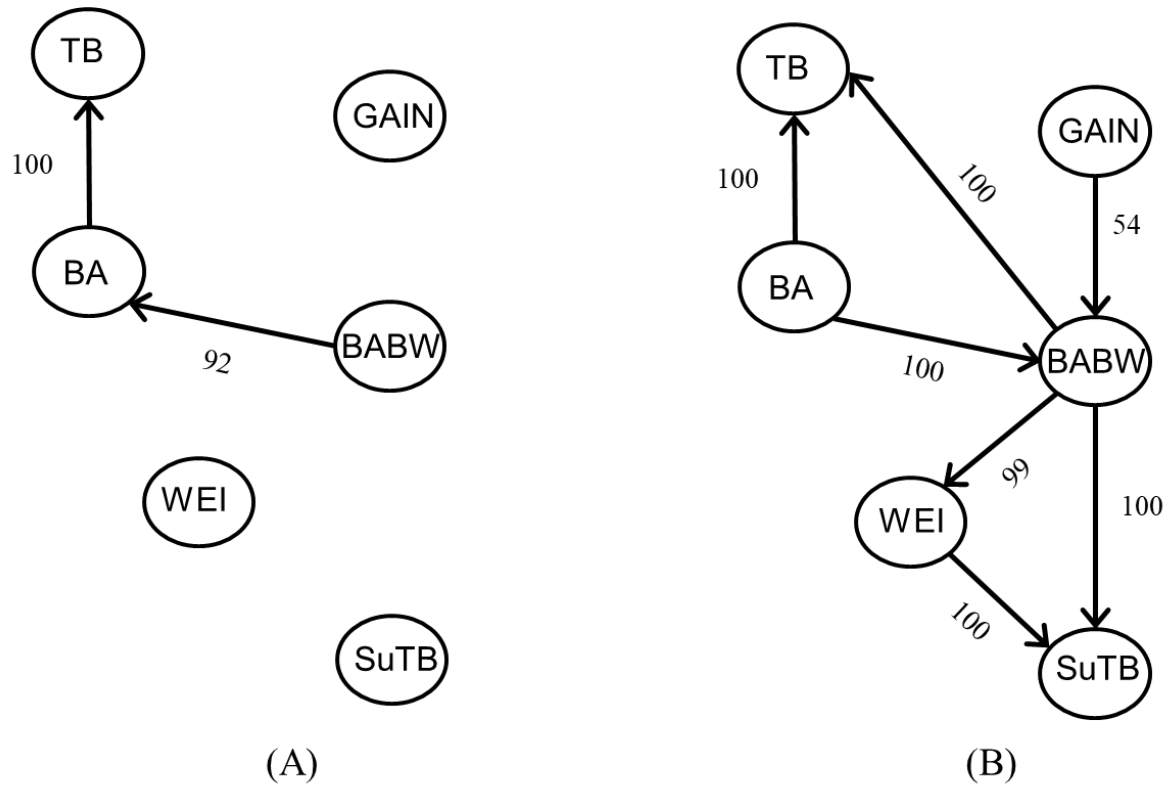


Figure 3.5 Stability analysis of the learned network of reproductive performance traits for sows (A) and gilts (B) using a leave-one-block-out Jackknife resampling approach. Values (%) indicate the percent of resampled datasets for which each link was present in the learned network structure. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

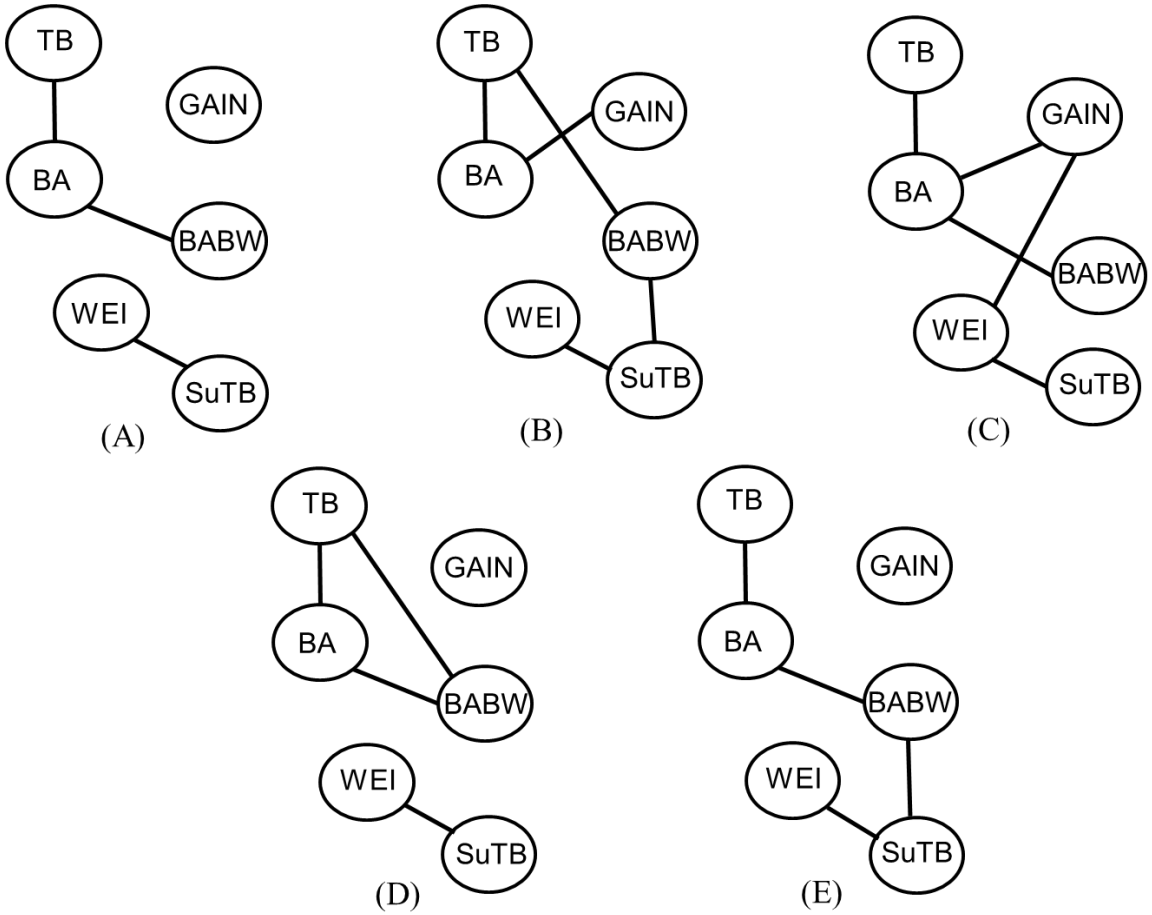


Figure 3.6 Network structures learned from five subsets created from the gilt dataset by random sampling without replacement to mimic the sow dataset in size and structure. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

Table 3.1 Descriptive statistics of reproductive performance traits in sows and gilts.

	Sows (n = 200)						Gilts (n = 440)					
Trait[1]	Mean	Min	25th Percentile	Median	75th Percentile	Max	Mean	Min	25th Percentile	Median	75th Percentile	Max
GAIN, kg	16.0	-1.8	11.2	15.4	20.9	36.3	17.9	-12.3	14.1	17.7	22.0	41.3
TB	15.4	6.0	14.0	15.0	17.0	21.0	14.3	6.0	13.0	14.0	16.0	21.0
BA	14.3	5.0	13.0	14.5	16.0	20.0	13.5	5.0	12.0	14.0	15.0	20.0
BABW, g	1404	1018	1251	1520	1520	2018	1279	817	1163	1267	1392	1960
WEI, day	4.3	3.0	4.0	4.6	5.0	10.0	6.2	2.0	4.0	4.0	5.0	31.0
SuTB	14.9	2.0	13.0	15.0	17.0	24.0	13.2	1.0	11.0	14.0	16.0	25.0

[1] GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

Table 3.2 Posterior means and 95% highest posterior density (HPD) intervals of structural coefficients from the final structural equation model selected for inference separately for sows and for gilts based on the inductive causation algorithm.

			Sows		Gilts	
Structural coefficient	From[1]	To[1]	Posterior mean	95% HPD interval	Posterior mean	95% HPD interval
$\lambda_{BA,BABW}$, unit per 100g	BABW	BA	-0.58	[-0.77, - 0.41]	-	-
$\lambda_{BABW,BA}$, g per unit	BA	BABW	-	-	- 360	[- 480, - 240]
$\lambda_{BABW,GAIN}$, g per kg	GAIN	BABW	-	-	0.03	[- 0.02, 0.09]
$\lambda_{TB,BA}$, unit per unit	BA	TB	0.96	[0.89, 1.02]	0.89	[0.85, 0.94]
$\lambda_{TB,BABW}$, unit per 100g	BABW	TB	-	-	- 0.11	[- 0.17, - 0.05]
$\lambda_{WEI,BABW}$, day per 100g	BABW	WEI	-	-	0.25	[- 0.04, 0.55]
$\lambda_{SuTB,BABW}$, unit per 100g	BABW	SuTB	-	-	- 0.42	[- 0.65, - 0.20]
$\lambda_{SuTB,WEI}$, unit per day	WEI	SuTB	-	-	0.16	[0.09, 0.24]

[1]GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

Table 3.3 Posterior means and 95% highest posterior density (HPD) intervals of indirect and total causal effects from the final structural equation models selected for inference separately for gilts based on the inductive causation algorithm

		Gilts	
Causal effect	Formulaic expression	Posterior mean	95% HPD interval
Indirect effects[1]			
GAIN → BABW → TB[3], unit per kg	$\lambda_{TB,BABW} \times \lambda_{BABW,GAIN}$	-0.004	[-0.011, 0.003]
GAIN → BABW → WEI, day per kg	$\lambda_{WEI,BABW} \times \lambda_{BABW,GAIN}$	0.007	[-0.008, 0.029]
GAIN → BABW → SuTB, unit per kg	$\lambda_{SuTB,BABW} \times \lambda_{BABW,GAIN}$	-0.01	[-0.04, 0.01]
GAIN → BABW → WEI → SuTB, unit per kg	$\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW} \times \lambda_{BABW,GAIN}$	0.001	[-0.001, 0.005]
BA → BABW → TB, unit per unit	$\lambda_{TB,BABW} \times \lambda_{BABW,BA}$	0.04	[0.02, 0.07]
BA → BABW → SuTB, unit per unit	$\lambda_{SuTB,BABW} \times \lambda_{BABW,BA}$	0.16	[0.06, 0.25]
BA → BABW →WEI→ TB, unit per unit	$\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW} \times \lambda_{BABW,BA}$	-0.015	[-0.036, 0.004]
BABW → WEI → SuTB, unit per 100g	$\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW}$	0.41	[-0.10, 0.95]
Total effects[2]			
GAIN → SuTB, unit per kg	$(\lambda_{SuTB,BABW} \times \lambda_{BABW,GAIN}) + (\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW} \times \lambda_{BABW,GAIN})$	-0.01	[-0.04, 0.01]
BA → TB, unit per unit	$(\lambda_{TB,BABW} \times \lambda_{BABW,BA}) + \lambda_{TB,BA}$	0.94	[0.89, 0.98]
BA → SuTB, unit per unit	$(\lambda_{SuTB,BABW} \times \lambda_{BABW,BA}) + (\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW} \times \lambda_{BABW,BA})$	0.14	[0.05, 0.24]
BABW → SuTB, unit per 100g	$(\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW}) + \lambda_{SuTB,WEI}$	0.57	[0.05, 1.15]

[1] Indirect effects are computed as the product of the corresponding direct effects represented by structural coefficients.

[2]Total effects are computed as the sum of the direct effect and all indirect effects.

[3]GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

Chapter 4 - A General Approach for Hierarchical Modeling of Heterogeneous Structural Coefficients in Structural Equation Models

ABSTRACT

Understanding the interconnections between performance outcomes in a system is increasingly important for integrated management. Structural equation models (SEM) are a type of multiple-variable modeling strategy that allows investigation of directionality in the association between outcome variables, thereby providing insight into their interconnections as putative causal links defining a functional network. A key assumption underlying SEM is that of a homogeneous network, where the structural coefficients defining functional links are assumed homogeneous and impervious to environmental conditions or management factors. This assumption seems questionable as systems are regularly subjected to explicit interventions to optimize the necessary trade-offs between outcomes. Using a Bayesian approach, we propose methodological extensions to hierarchical SEM that explicitly specify structural coefficients as functions of systematic and non-systematic sources of variation, thus allowing for hierarchical heterogeneity in the network links and recognizing design structure in the data. We validate our proposed method using a simulation study and show that hierarchical sources of heterogeneity on structural coefficients can be estimated and inferred upon accurately. Further, we show that networks can be consistently identified as homogeneous or heterogeneous based on model fit statistics that compare competing SEMs with flexible specifications of structural coefficients. We apply the proposed methodological extensions to a dataset from a designed experiment in swine production consisting of six interrelated reproductive performance outcomes to explore

physiological links that differed by parity. Overall, our results indicate that explicit hierarchical SEM-based modeling of heterogeneous functional networks can be used to advance understanding of complex networks of performance outcomes in an animal production system.
Key words: Animal production systems, Hierarchical Bayesian, Heterogeneous structural coefficients, and Structural equation models

4.1. INTRODUCTION

Efficient agricultural production systems play a key role in ensuring a safe and secure food supply for the rapidly growing global population (Godfray et al. 2010). To succeed at this challenge, food production systems will require truly integrated strategies that cut across scientific disciplines and provide multifaceted solutions. Pivotal to such integrated management is an in-depth mechanistic understanding of the functional connections that underlie the relationships between system outcomes. Classical multiple-trait models (MTM) are often considered to investigate associations between outcomes (Johnson and Wichern 2007). However, classical MTM cannot evaluate directionality of such associations (Rosa et al. 2011), thereby impairing their use for inferring potential causal links between outcome variables in a complex system.

In contrast to classical MTM, structural equation models (SEM) (Haavelmo 1943) define a special type of MTM that can, under certain conditions, be used to evaluate the directionality of interconnections between outcomes. So-called direct effects can then be used to inform putative causal links that describe functional networks in a system. Methodologically, SEM (Haavelmo 1943) evolved from path analysis (Wright 1921, Duncan 1966) and were later extended to multivariate linear models (Joreskog 1973), mixed models (Gianola and Sorensen 2004) and graphical models (Lauritzen 1996), thereby integrating SEM with the causal inference framework (Pearl 2009).

A key assumption underlying SEM is that of homogeneity of the structural coefficients (Gianola and Sorensen 2004) defining functional links in a network, so that such links are considered impervious to environmental conditions or management factors. This assumption seems questionable as functionality in a system may differ between subpopulations and is often

subjected to explicit interventions to optimize the necessary trade-offs between competing outcomes. For example, early work by Wu et al. (2007) showed that direct effects obtained from a simultaneous and recursive modeling framework to describe the associations between milk yield and udder health in dairy cows were yield-dependent and changed during a lactation period. Also, a recent study from our group (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation) observed distinct causal networks for two parity-defined subpopulations of female pigs based on separate SEM fitted to reproductive traits in a swine production system. Taken together, empirical evidence warrants methodological extensions to SEM that explicitly accommodate heterogeneity of structural coefficients into a single analysis to better capture heterogeneous functional networks in complex systems.

In a SEM, structural coefficients may be interpreted as substitutes of residual covariances of the classical MTM (Gianola and Sorensen 2004, Wu et al. 2010) in that any source of covariance between observations beyond those explicitly specified by the model are considered functional relationships between variables (Gianola and Sorensen 2004, Wu et al. 2010). Several methodological developments have been proposed to specify sources of variability in MTM (co)variances, and functions thereof (Bello et al. 2010, Yang and Tempelman 2012); these methods have seen growing interest in agricultural applications (Bello et al. 2013, Tempelman et al. 2015, Ou et al. 2016). Furthermore, Wu et al. (2007) proposed an extension to simultaneous and recursive models that recognized structural coefficients as specific to pre-defined subpopulations. Here, we extend Wu et al. (2007) and propose a general approach to hierarchical SEM that enables explicit specification of systematic and non-systematic sources of variation on the structural coefficients determining functional links in a network. The hierarchical framework

of our proposed approach can easily accommodate data architecture and is thus well suited to structured data from designed experiments or observational studies in agriculture.

The objectives of this Chapter are 1) to propose a general approach for explicit specification of heterogeneity of SEM structural coefficients as functions of systematic and non-systematic sources of variability, 2) to validate the properties of the proposed methodological extensions using a simulation study, and 3) to apply the proposed method to an agricultural dataset corresponding to a designed experiment in swine reproduction, for which preliminary analyses suggested parity-dependent networks. In section 4.2, we briefly introduce SEM, delineate the proposed methodological extension and describe the hierarchical Bayesian framework we used for implementation. This is followed by a description of the simulation study and the data application in swine reproduction. Results are presented in section 4.3 and discussed in section 4.4. Finally, section 4.5 provides concluding remarks.

4.2. METHOD

4.2.1. Structural Equation Model specification

Consider the directed acyclic graph in Figure 4.1 representing a hypothetical causal network between outcome variables y_j (j = 1,2, …, J; J = 4). Arrows in Figure 4.1 indicate that y_1 has a direct effect on y_2 and on y_3, whereas, y_2 and y_3 directly affect y_4. Figure 4.1 also portrays mutually independent residuals $\boldsymbol{e}_j$ and mutually correlated random effects $\boldsymbol{b}_j$, both of which directly affect the corresponding j^{th} outcome variable.

For each i^{th} subject (i = 1, 2, …, n), a hierarchical SEM can be expressed following Gianola and Sorensen (2004) as:

$$\boldsymbol{y}_i = \boldsymbol{\Lambda}\boldsymbol{y}_i + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b} + \boldsymbol{e}_i \quad (4.1)$$

where $\boldsymbol{y}_i' = [y_{i1}\ \ y_{i2}\ \ y_{i3}\ \ y_{i4}\ ]$ is a vector of $J = 4$ observed outcomes, and $\boldsymbol{\Lambda}$ is a $J \times J$ square matrix of zeroes, except that some elements of the lower triangle are replaced by unknown structural coefficients $\lambda_{jj'}$ $(j > j')$. The structural coefficient $\lambda_{jj'}$ represents a functional link whereby outcome j' has a direct effect on outcome j, as indicated by the arrows in Figure 4.1; otherwise, $\lambda_{jj'}$ is set to zero. As such, the matrix $\boldsymbol{\Lambda}$ corresponding to Figure 4.1 can be expressed as:

$$\boldsymbol{\Lambda} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 \\ 0 & \lambda_{42} & \lambda_{43} & 0 \end{bmatrix} \tag{4.2}$$

For instance, λ_{21} represents the magnitude of the direct effect of $y_1$on y_2 and represents the expected change in y_2 per unit increase in y_1. Additional parameters in equation (4.1) consist of $\boldsymbol{\beta}' = [\boldsymbol{\beta}_1'\ \ \boldsymbol{\beta}_2'\ \ \boldsymbol{\beta}_3'\ \ \boldsymbol{\beta}_4']$**,** a vector of unknown fixed-effect location parameters associated with factors and covariates (e.g. treatments, demographics) through the corresponding incidence matrix $\boldsymbol{X}_i = diag[\boldsymbol{x}_{i1}'\ \boldsymbol{x}_{i2}'\ \boldsymbol{x}_{i3}'\ \boldsymbol{x}_{i4}']$ unique to the i^{th} subject. Also, $\boldsymbol{b}' = [\boldsymbol{b}_1'\ \ \boldsymbol{b}_2'\ \ \boldsymbol{b}_3'\ \ \boldsymbol{b}_4']$ is a vector of unknown random effects associated with blocking factors or other components of the data architecture through the design matrix $\boldsymbol{Z}_i = diag[\boldsymbol{z}_{i1}'\ \boldsymbol{z}_{i2}'\ \boldsymbol{z}_{i3}'\ \boldsymbol{z}_{i4}']$. Random effects $\boldsymbol{b}$ are assumed multivariate normally distributed such that:

$$\boldsymbol{b} \sim MVN\big(\boldsymbol{0},\ \boldsymbol{B} \otimes\ \boldsymbol{I}_q\big) \tag{4.3}$$

where $\boldsymbol{B} = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_{12}} & \sigma_{b_{13}} & \sigma_{b_{14}} \\ \sigma_{b_{21}} & \sigma_{b_2}^2 & \sigma_{b_{23}} & \sigma_{b_{24}} \\ \sigma_{b_{31}} & \sigma_{b_{32}} & \sigma_{b_3}^2 & \sigma_{b_{34}} \\ \sigma_{b_{41}} & \sigma_{b_{42}} & \sigma_{b_{43}} & \sigma_{b_4}^2 \end{bmatrix}$, and q is the number of levels of the random effect factor.

Finally, $\boldsymbol{e}_i' = [e_{i1}\ \ e_{i2}\ \ e_{i3}\ \ e_{i4}]$ is the corresponding set of residuals for subject *i* and is assumed multivariate normal, as $\boldsymbol{e}_i \sim MVN(\boldsymbol{0}, \boldsymbol{R})$, where

$$\boldsymbol{R} = \begin{bmatrix} \sigma^2_{e_1} & 0 & 0 & 0 \\ 0 & \sigma^2_{e_2} & 0 & 0 \\ 0 & 0 & \sigma^2_{e_3} & 0 \\ 0 & 0 & 0 & \sigma^2_{e_4} \end{bmatrix} \tag{4.4}$$

The underlying assumption of a diagonal residual covariance matrix $\boldsymbol{R}$ (i.e. $\sigma_{e_{jj'}} = 0$) is standard in the context of recursive SEM to ensure parameter identifiability (Wu et al. 2010, Gianola and Sorensen 2004). This implies that any correlations between outcomes above and beyond those due to random effects in the design structure, are attributed to the causal effects between outcomes which are represented by structural coefficients $\lambda_{jj'}$ (Gianola and Sorensen 2004).

4.2.2. Heterogeneous structural coefficients

The standard specification of SEM assumes causal homogeneity (Shipley 2002); that is, structural coefficients $\lambda_{jj'}$ $(j' < j)$ are considered of the same magnitude for all subjects in the population. Next, we propose methodological extensions that relax this assumption and, instead, specify each structural coefficient $\lambda_{jj'}$ as a function of fixed and random effects, thereby reflecting potential systematic and non-systematic sources of variability in the direct effects connecting outcomes in a network. Specifically, for each i^{th} subject, we specify:

$$\lambda_{jj',i} = \boldsymbol{x}'_{jj',i}\boldsymbol{\delta}_{jj'} + \boldsymbol{z}'_{jj',i}\boldsymbol{v}_{jj'} \tag{4.5}$$

where $\boldsymbol{\delta}_{jj'}$ is a vector of unknown fixed-effect location parameters (e.g. parity, demographics, *etc*.), $\boldsymbol{v}_{jj'}$ is a vector of unknown random effects associated with blocking factors or other design components of the data architecture such that $\boldsymbol{v}_{jj'} \sim NIID(\mathbf{0}, \sigma^2_{v_{jj'}}\mathbf{I}_q)$, and $\boldsymbol{x}'_{jj',i}$ and $\boldsymbol{z}'_{jj',i}$ are known corresponding row incidence vectors unique to each i^{th} subject and specific to the jj'^{th} structural coefficient. Further, $\boldsymbol{x}'_{jj',i}$ and $\boldsymbol{z}'_{jj',i}$ need not be the same for all jj' structural coefficients, neither as the rows of matrices $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$, respectively, as defined in Equation (4.1).

4.2.3. Data Simulation

A simulation study was used to validate the proposed methodological extensions for specifying heterogeneous structural coefficients on a SEM and to study inferential properties for the corresponding parameters $\boldsymbol{\delta}_{jj'}$, $\boldsymbol{v}_{jj'}$, and $\sigma^2_{\boldsymbol{v}_{jj'}}$. We designed two simulation scenarios that followed the data-generating process reflected by the network in Figure 4.1 and in Equation (4.1). Simulation scenario A comprised all non-zero structural coefficients being homogeneous across subjects. Simulation scenario B was similar to scenario A except that structural coefficients $\lambda_{21,i}$ and $\lambda_{43,i}$ were specified as heterogeneous following Equation (4.5). For each simulation scenario, we generated 10 datasets, with each dataset consisting of observations for *J* = 4 outcome variables from 2,000 subjects arranged in 100 clusters. In both scenarios, the choice of parameter values for data simulation reflected the data application (refer to section 4.2.4)).

For simulation scenario A, structural coefficients were specified as $\lambda_{21} = 0.25$, $\lambda_{31} = -0.6$, $\lambda_{42} = 0.4$, and $\lambda_{43} = 0.8$ for all subjects. Scenario B was similar to scenario A except that structural coefficients $\lambda_{21,i}$ and $\lambda_{43,i}$ were deemed heterogeneous and generated as in Equation (4.5), with $\boldsymbol{\delta}'_{21} = [\,0.25 \;\; -0.25]$, and $\boldsymbol{\delta}'_{43} = [\,0.5 \;\; 0.3]$ for $\lambda_{21,i}$ and $\lambda_{43,i}$, respectively. The elements of $\boldsymbol{\delta}_{21}$ and $\boldsymbol{\delta}_{43}$ identified the effects of two levels of a single fixed effect factor measured on the i^{th} subject and expressed in the design vector $\boldsymbol{x}'_{jj',i}$ using a set-to-zero parameterization (Milliken and Johnson 2009). Also in simulation scenario B, the cluster-level variances for the random effects $\boldsymbol{v}_{21}$ and $\boldsymbol{v}_{43}$ on structural coefficients $\lambda_{21,i}$ and $\lambda_{43,i}$ were specified as $\sigma^2_{v_{21}} = 0.1$ and $\sigma^2_{v_{43}} = 0.2$, respectively.

In all cases, we simulated location parameters $\boldsymbol{\beta}$ analogous to a fixed effect factor with two levels using a set-to-zero parameterization (Miliken and Johnson 2009) such that $\boldsymbol{\beta}'_1 = [25 \;\; -12]$, $\boldsymbol{\beta}'_2 = [18 \;\; -6]$, $\boldsymbol{\beta}'_3 = [5 \;\; -3]$, $\boldsymbol{\beta}'_4 = [15 \;\; -5]$. The diagonal elements of the

residual (co)variance matrix $\boldsymbol{R}$ were specified as $\sigma^2_{e_1} = 9$, $\sigma^2_{e_2} = 1$, $\sigma^2_{e_3} = 3$, and $\sigma^2_{e_4} = 6$, whereas the diagonal elements of the random effects (co)variance matrix $\boldsymbol{B}$ were specified as $\sigma^2_{b_1} = 9$, $\sigma^2_{b_2} = 2$, $\sigma^2_{b_3} = 1$, and $\sigma^2_{b_4} = 3$. Finally, the random-level correlations between outcomes were as follows:

$$\boldsymbol{\rho_b} = \begin{bmatrix} 1 & -0.7 & 0.4 & 0.1 \\ -0.7 & 1 & -0.5 & -0.2 \\ 0.4 & -0.5 & 1 & 0 \\ 0.1 & -0.2 & 0 & 1 \end{bmatrix}$$

For each simulation scenario, the full (co)variance matrix $\boldsymbol{B}$ was obtained by pre- and post-multiplying matrix $\boldsymbol{\rho_b}$ by a diagonal matrix containing the square root of the corresponding variance parameters.

4.2.4. Application to swine data

Data were obtained from an experimental study on swine reproduction conducted at a commercial swine farm in northern Ohio (Gonçalves et al. 2016). The complete description of the data is available in Gonçalves et al. (2016). Briefly, the outcomes variables consisted of descriptors of reproductive performance, namely female weight gain during late (d 90 to 111) gestation (GAIN, in kg; j = 1), number born alive in a litter (BA; j = 2), born alive average birth weight (BABW, in g; j = 3), total number of piglets born (TB; j = 4), wean-to-estrous interval (WEI, in days; j = 5), and total litter size born in the subsequent gestation (SuTB; j = 6). Both TB and SuTB are defined as total litter size and consisted of the summation or piglets born alive, stillborns and mummies. In this analysis, we only considered females with complete records on all six outcomes (i.e. no missing data). After editing, the final dataset used for analysis consisted of 691 females, 200 of which were multiparous adults (i.e. sows) and 491 were primiparous youngsters (i.e. gilts). Within each parity group, females were arranged in 97 and 222 body weight blocks for sows and gilts, respectively. Dietary treatments consisting of four

combinations of energy intake and dietary amino acids were randomly assigned to individual females within each body weight block.

For each j^{th} outcome, the $\boldsymbol{x}_{ij}$ vector corresponding to Equation (4.1) included an intercept and indicator variables for parity groups (sows or gilts), dietary treatments and their combination, such that the corresponding $\boldsymbol{\beta}_j$ was expressed in set-to-zero full-rank parameterization (Milliken and Johnson 2009). The $\boldsymbol{z}_i$ vector corresponding to Equation (4.1) consisted of indicator variables that identified the body weight block within parity group for each i^{th} observation.

Specification of the network structure was based on preliminary analyses (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation) conducted separately for each parity group using a SEM that assumed homogeneous structural coefficients within groups. Briefly, for each parity group, the network structure and thus, the specification of matrix $\boldsymbol{\Lambda}$, was learned from the data using the Inductive Causation algorithm (Verma and Pearl 1991) adapted to a hierarchical Bayesian framework (Valente et al. 2010). Results from such preliminary analyses are depicted in Figure 4.2A and 4.2B for sows and for gilts, respectively, and suggest distinct causal networks for each parity group. Specifically, sows showed a sparsely connected network with few direct effects and thus, few non-zero structural coefficients (i.e. BABW → BA and BA → TB; remaining outcomes were unconnected) (Figure 4.2A). By contrast, gilts showed a more densely interconnected network between reproductive outcomes (Figure 4.2B) (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation). These preliminary results prompted the methodological SEM extension proposed herein to incorporate heterogeneity due to both treatment structure and design structure onto the modeling of structural coefficients. When structural coefficients were specified as heterogeneous, as in Equation (4.5), the

corresponding $\boldsymbol{x}'_{jj',i}$ vector included indicator variables for parity group and $\boldsymbol{z}'_{jj',i}$ consisted of indicator variables for body weight blocks within each parity group.

Preliminary analysis showed a sow network structure that was nested within that of gilts, though the direction of the link connecting BA and BABW was reversed (i.e. BA ← BABW in sows vs. BA → BABW in gilts; Figure 4.2). Therefore, for the joint analysis presented herein, we implemented the most general network structure (i.e. that obtained from the preliminary analyses of gilt data) and further considered two causal structures that differed in the direction of the link connecting BA and BABW (Figure 4.3). We refer to the causal orientation with direct effect BA ← BABW as option A, and to that with direct effect BA → BABW as option B (Figure 4.3). The matrices $\boldsymbol{\Lambda}$ of structural coefficients corresponding to options A and B can be expressed as:

$$\boldsymbol{\Lambda}_{\mathrm{A}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{BA,BABW} & 0 & 0 & 0 \\ \lambda_{BABW,GAIN} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{TB,BA} & \lambda_{TB,BABW} & 0 & 0 & 0 \\ 0 & 0 & \lambda_{WEI,BABW} & 0 & 0 & 0 \\ 0 & 0 & \lambda_{SuTB,BABW} & 0 & \lambda_{SuTB,WEI} & 0 \end{bmatrix} \tag{4.7}$$

$$\boldsymbol{\Lambda}_{\mathrm{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_{BABW,GAIN} & \lambda_{BABW,BA} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{TB,BA} & \lambda_{TB,BABW} & 0 & 0 & 0 \\ 0 & 0 & \lambda_{WEI,BABW} & 0 & 0 & 0 \\ 0 & 0 & \lambda_{SuTB,BABW} & 0 & \lambda_{SuTB,WEI} & 0 \end{bmatrix} \tag{4.8}$$

For both $\boldsymbol{\Lambda}_{\mathrm{A}}$ and $\boldsymbol{\Lambda}_{\mathrm{B}}$, the outcomes are ordered as described earlier in Section (4.2.4). However, we recognized that parameter $\lambda_{BA,BABW}$ in $\boldsymbol{\Lambda}_{\mathrm{A}}$ was not included in the lower triangle. Therefore, to abide by standard SEM notation, we re-ordered the outcomes as $j_1^* =$ GAIN, $j_2^* =$ BABW, $j_3^* =$ BA, $j_4^* =$ TB, $j_5^* =$ WEI, and $j_6^* =$ SuTB and re-wrote $\boldsymbol{\Lambda}_{\mathrm{A}}$ as $\boldsymbol{\Lambda}_{\mathrm{A}^*}$,

$$\mathbf{\Lambda}_{A^*} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \lambda_{BABW,GAIN} & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{BA,BABW} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{TB,BABW} & \lambda_{TB,BA} & 0 & 0 & 0 \\ 0 & \lambda_{WEI,BABW} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{SuTB,BABW} & 0 & 0 & \lambda_{SuTB,BABW} & 0 \end{bmatrix} \quad (4.9)$$

such that all structural coefficients could be located in the lower triangle of $\mathbf{\Lambda}_{A^*}$.

4.2.5. Alternative SEM specifications

We fitted alternative SEM specifications to the simulated data in order to validate inferential properties of the proposed methodological extensions. Specifically, Model 0 (M0) was a fully-recursive SEM with a complete lower triangular matrix $\mathbf{\Lambda}$ such that all the $\lambda_{jj'}$ indicative of recursive effects were estimated. The fully recursive SEM specification in M0 can be shown to be statistically equivalent to a classical MTM (Gianola and Sorensen 2004). Model 1 (M1) was devised as a SEM with selected $\lambda_{jj'}$ set to zero to reflect the causal structure in Figure 4.1 and Equation (4.1). It is noted that M1 represents the true data-generation process for simulation scenario A. Both M0 and M1 specify homogeneity of any structural coefficient $\lambda_{jj'}$ not set to zero but rather estimated from data. Model 2 (M2) expands M1 with a heterogeneous specification of structural coefficients such that all non-zero $\lambda_{jj'}$ are expressed as functions of fixed (e.g. subpopulations of interest) and random (e.g. design components of data architecture) effects, as in Equation (4.5). Model 2* (M2*) expands M1 by enabling a heterogeneous specification of structural coefficients $\lambda_{21,i}$ and $\lambda_{43,i}$ only, thereby reflecting the true data-generating process for simulation scenario B. In the simulation study, the alternative SEM specifications M0, M1, M2 and M2* were fitted to each of the 20 simulated datasets.

For the swine data application, we implemented SEM specifications M0, M1 and M2 as described for the simulation study. For models M1 and M2, we fitted the two competing causal

structures represented by $\mathbf{\Lambda}_{A^*}$ and $\mathbf{\Lambda}_{B}$ in Equations (4.1) and (4.5), and in Figure 4.3. Preliminary analyses further indicated substantial evidence for heterogeneity of residual variances as a function of body weight blocks on GAIN, BA, TB, WEI, and SuTB (results not shown). Therefore, we devised Model 3 (M3), which expanded M2 to accommodate heterogeneous residual variances for the corresponding outcomes, following Kizilkaya and Tempelman (2005), such that:

$$\sigma^2_{e_{j,i}} = \sigma^2_{e_{ref,j}} w_{e_{j,i}} \tag{4.9}$$

where $\sigma^2_{e_{ref,j}}$ represents an unknown reference residual variance for the j^{th} outcome (i.e. GAIN, BA, TB, WEI, and SuTB), akin to the intercept in location parameters, and $w_{e_{j,i}}$ represents the random effect of the body-weight block for the j^{th} outcome that the i^{th} female corresponds to.

4.2.6. Prior specifications

The proposed methodological extension to SEM was implemented in a hierarchical Bayesian framework, where all unknown parameters are considered to be random effects. In previous sections, we have referred to so-called fixed effect factors, which may be defined from a Bayesian perspective as those with effects specified by a vague prior (Gianola and Sorensen 2004). In turn, our reference to random effect factors reflects notation often used in mixed models and may be represented in the Bayesian framework as those with effects specified with a structural prior dependent on unknown parameters estimated from data (Sorensen and Gianola Sorensen 2002).

The elements of location parameters $\boldsymbol{\beta}_j$ were specified to have vague prior densities; i.e. $p(\boldsymbol{\beta}_j) \propto$ constant for all j (Sorensen and Gianola 2002). For each residual variance $\sigma^2_{e_j}$, we generated a proxy for a vague prior using as an instrument the density of a truncated scaled -inverse chi-square distribution and setting degrees of freedom υ_{e_j}= -1 and scale parameter $s^2_{e_j}$=

0. This proxy is consistent with the prior $\sqrt{\sigma^2_{e_j}} \sim U(0, A)$, for any finite but appropriately large value of *A* such that the resulting distribution is vague, as recommended for variance components by Gelman (2006). On the random effects $\boldsymbol{b}$, we specified a structural prior such that $p(\boldsymbol{b}|\boldsymbol{B}) = N(\mathbf{0}, \boldsymbol{B} \otimes \boldsymbol{I}_q)$ to allow for borrowing of information across levels of the random effect factor in a j^{th} outcome, as is common in a mixed models framework (Robinson 1991). For the random-level (co)variance matrix $\boldsymbol{B}$, we implemented a proxy for a vague prior using as an instrument the density of an inverse-Wishart distribution and setting degrees of freedom υ_B = -(*J*+1) and specifying a *J* × *J* scale matrix of zeroes; this can be interpreted as a multivariate extension of the proxy used for the prior specified above on the residual variances.

For SEM with homogeneous structural coefficients, the prior distribution for each non-zero $\lambda_{jj'}$ was specified to be vague such that $p(\lambda_{jj'}) \propto$ constant for all jj'. For SEM that allowed heterogeneity of structural coefficients, priors were specified on $\boldsymbol{\delta}_{jj'}$, $\boldsymbol{v}_{jj'}$ and $\sigma^2_{v_{jj'}}$. More specifically, in turn, random effects $\boldsymbol{v}_{jj'}$ were assigned a structural prior such that $p\left(\boldsymbol{v}_{jj'}|\sigma^2_{v_{jj'}}\right) \sim N(\mathbf{0}, \sigma^2_{v_{jj'}}\mathbf{I}_q)$. For the random-level variance $\sigma^2_{v_{jj'}}$ of structural coefficients, we implement a proxy for a vague prior as described above for other variance components.

For additional Model 3 (M3), we followed Kizilkaya and Tempelman (2005) to specify heterogeneity on the diagonal elements $\sigma^2_{e_j}$ of matrix $\boldsymbol{R}$ corresponding to outcomes GAIN, BA, TB, WEI and SuTB. For each reference variance parameters $\sigma^2_{e_{ref,j}}$, we generated a proxy for a vague prior using as an instrument the density of a scaled-inverse chi-square distribution and setting degrees of freedom υ_{e_j}= -1 and scale parameter $s^2_{e_j}$= 0. Further, each element of the outcome-specific set of random effects $\boldsymbol{w}_{e_j} = \left\{w_{e_{j,l}}\right\}_{l=1}^{q}$ was assigned an independent structural

prior $IG(\alpha_{e_j}, \alpha_{e_j} - 1)$ such that $E(\boldsymbol{w}_{e_j} | \alpha_{e_j}) = 1$ and $var(\boldsymbol{w}_{e_j} | \alpha_{e_j}) = (\alpha_{e_j} - 2)^{-1}$ (Kizilkaya and Tempelman 2005) for $\alpha_{e_j} > 2$. Finally, we assigned a vaguely informative yet proper prior density to each α_{e_j}, whereby $\alpha_{e_j} \sim p\left(\alpha_{e_j}\right) \propto (1 + \alpha_{e_j})^{-2}$, which is commonly used for strictly positive parameters (Kizilkaya and Tempelman 2005).

Following from the conditionally conjugate prior specifications described above, full conditional densities can be easily recognized for all unknown parameters in M0, M1 and M2, thus facilitating a Gibbs sampling implementation of the Markov Chain Monte Carlo (MCMC). For M3, the full conditional density of α_{e_j} is not recognizable and thus calls for sampling of α_{e_j} from a Metropolis-Hastings algorithm, as described in Kizilkaya and Tempelman (2005).

4.2.7. Bayesian implementation

Implementation of the proposed methodological extensions used a MCMC programmed in R software (R Development Core Team, 2016). In the simulation study, alternative SEM were fitted to each simulated dataset using a single MCMC chain run for 260,000 iterations after a burn-in period of 60,000 cycles; one of every 2 MCMC samples was saved and used for posterior inference. For the data application, each alternative SEM was fitted using 10 MCMC chains ran for 100,000 iterations after a burn-in period of 80,000 cycles; one of every 2 samples were saved for posterior inference. In all cases, convergence was monitored using trace plots on all higher-order hyperparameters. Diagnostic tests for convergence were also conducted, specifically the single-chain approach proposed by Raftery and Lewis (1992) in the simulation study and the multiple-chain testing approach proposed by Gelman and Rubin (1992) in the data application. Effective sample size (ESS) was estimated to evaluate the number of effectively independent samples among the autocorrelated MCMC samples for each hyperparameter (Sorensen et al. 1995). Length of MCMC chains was adjusted to ensure that ESS was greater

than 400 for all hyperparameters. In all cases, convergence diagnostics were performed using the R package CODA (Plummer et al. 2005). For each parameter of interest, we summarized posterior inference using posterior means and 95% highest posterior density (HPD) intervals, as well as selected posterior probabilities. For assessing heterogeneity of structural coefficients between parity groups, we computed the 95% HPD interval of a parity difference expressed as $\lambda_{jj'}^{\text{Gilt}} - \lambda_{jj'}^{\text{Sow}}$ and the posterior probability of such difference as $max\left(p\left(\lambda_{jj'}^{\text{Gilt}} - \lambda_{jj'}^{\text{Sow}}|y\right) > 0,\right.$ $\left.p\left(\lambda_{jj'}^{\text{Gilt}} - \lambda_{jj'}^{\text{Sow}}|y\right) < 0\right)$ and 95% HPD interval of the parity differences between gilts and sows.

For completeness, we include results on heteroskedasticity across body-weight blocks for GAIN, BA, TB, WEI and SuTB by defining the coefficient of variation $CV_j = 1/\sqrt{\alpha_{e_j} - 2}$ (Kizilkaya and Tempelman 2005) such that a greater value of CV is indicative of the degree of residual heteroskedasticity across blocks.

4.3. RESULTS

4.3.1. Simulation study

Figure 4.4 shows comparative model fit of alternative SEM specifications under simulation scenarios A (panel A) and B (panel B), expressed as DIC differences relative to the corresponding true data-generation model in each case. For networks generated under conditions of homogeneous structural coefficients (i.e. simulation scenario A), any of the alternative SEM specifications M0, M2 and M2* showed DIC differences indicative of impaired global fit relative to the true-data generation model (i.e. M1). The difference in fit can be explained by the loss in parsimony of M0, M2 and M2* relative to M1. Indeed, the effective number of parameters pD (Spiegelhalter et al. 2002) was estimated to be greater by [minimum, maximum]

[368.6, 380.4], [390.9, 412.8], and [377.5, 391.0] for M0, M2 and M2*, respectively, relative to the true data-generation model M1 across simulated datasets under scenario A.

Impaired global fit of alternative SEM misspecifications was even more dramatic under simulation scenario B, as indicated by the logarithmic scale of the y-axis in Figure 4.4B. Recall that simulation scenario B was characterized by heterogeneity of selected structural coefficients in the network, namely $\lambda_{21,i}$ and $\lambda_{43,i}$, such that M2* depicted the data-generation process. Model specifications that incorrectly assumed homogeneity of all structural coefficients (i.e. M0 and M1) showed drastically impaired fit, with DIC differences ranging from 3713 to 4344 points relative to the true model M2* (Figure 4.4B). In fact, the observed differences in global fit for M0 and M1 relative to M2* were explained by differences in model adequacy, as quantified by the deviance term of DIC (Spiegelhalter et al. 2002). The value of the deviance term was estimated to be greater by [minimum, maximum] [3832.2, 4479.2] and [3826.4, 4472.45] for alternative SEM M0 and M1, respectively, relative to the true data-generation model M2*. In turn, SEM M2, which allowed all structural coefficients to be heterogeneous, also showed impaired fit relative to M2*, though the magnitude of DIC differences was considerably mitigated, ranging from 59 to 113 points across simulated datasets.

Figures 4.5A and 4.5B illustrate overall accuracy of parameter estimation, expressed as percent coverage of 95% HPD intervals, for alternative SEM specifications M0, M1, M2 and M2* fitted to simulation scenarios A and B, respectively. For networks generated under conditions of homogeneous structural coefficients (i.e. M1 in simulation scenario A), estimation accuracy was within probabilistic expectation for all model parameters regardless of SEM specification. On average, the coverage was 93.1, 95.0, 97.5, 96.8, 94.4, and 99.2 % for parameter sets $\lambda_{jj'}, \beta_{1,j}, \ \beta_{2,j}, \ \sigma^2_{e_j}, \ \sigma^2_{u_j}$ and $\sigma_{u_{jj'}}$, respectively. This was to be expected as all the

fitted SEM specifications were at least as general and flexible as the one used for data generation (i.e. M1).

Under simulation scenario B with heterogeneous $\lambda_{21,i}$ and $\lambda_{43,i}$, estimation accuracy for both SEM M2 and M2* was on target for all parameters (Figure 4.5B), with average coverage of ~100.0 ~100.0, ~98.8, ~100.0, ~91.3, ~98.3, ~85.0, ~100.0 and ~100.0%, for $\lambda_{jj'}, \beta_{1,j}, \beta_{2,j}, \sigma^2_{e_j}$, $\sigma^2_{u_j}, \sigma_{u_{jj'}}, \delta_{1,jj'}, \delta_{2,jj'}$ and $\sigma^2_{v_{jj'}}$, respectively. In contrast, coverage of the simpler SEM specifications M0 and M1 were decreased, thus indicating impaired estimation accuracy (Figure 4.5B). In these cases, decreased coverage was explained primarily by underestimation of structural coefficients $\lambda_{jj'}$ (i.e. average coverage of approximately ~75%), and by biased estimation of location parameters $\beta_{1,j}$ and $\beta_{2,j}$ (i.e. average coverage of ~68.8% and ~47.5%, respectively). Furthermore, (co)variance components $\sigma^2_{e_j}, \sigma^2_{u_j}, \sigma_{u_{jj'}}$ were mostly overestimated (average coverage of ~47.5, ~43.8 and ~71.3%, respectively) when M0 and M1 were fitted to datasets under simulation scenario B, thus failing to recognize heterogeneous $\lambda_{21,i}$ and $\lambda_{43,i}$.

4.3.2. Application to swine data

Table 4.1 shows DIC-based global fit assessment of alternative SEM specifications M0, M1, M2 and M3 fitted to the swine data with causal structures A (i.e. BA ← BABW) and B (i.e. BA → BABW) (Figure 4.3). Overall, M3 yielded the smallest DIC values of all SEM specifications, thus indicating evidence for heterogeneity of both structural coefficients (based on DIC differences of M3 relative to M0 and M1) and residual variances (based on the DIC difference of M3 relative to M2). Furthermore, DIC comparisons within M3 seemed to slightly favor a causal structure with a direct effect from BA to BABW (i.e. BA → BABW), consistent with Figure 4.3B option, though the DIC difference between causal structures was small in magnitude (~4.4 points; Table 4.1) and potentially inconclusive. Based on these results, we

selected as best fitting SEM M3 with a casual structure that included a direct effect BA → BABW, and used it for final inference on a heterogeneous network of reproductive performance outcomes in swine females.

4.3.2.1. Heterogeneous direct effects

Evidence for heterogeneity of direct effects was substantial, as a function of both the systematic effect of parity group and the non-systematic effects of body weight block, as dictated by the design structure. Table 4.2 shows the posterior summary of the structural coefficients for sows and gilts, as well as their parity difference, based on the final model selected for inference. Systematic heterogeneity of structural coefficients was mostly apparent for the direct effects of GAIN on BABW (i.e. GAIN → BABW) and of WEI on SuTB (i.e. WEI → SuTB), with posterior probabilities of 0.93 and 0.99 for parity differences, respectively. Specifically, every kg increase in sow GAIN resulted in an estimated increase of BABW of (posterior mean [95% HPD]) 0.07 g [0.02, 1.11]. By contrast, the magnitude of the corresponding values for gilts seemed to be halved, with a 95% HPD interval that slightly overlapped zero (Table 4.2). Furthermore, the direct effect of WEI on SuTB (i.e. WEI → SuTB) was of opposite signs in each parity group; every additional WEI day decreased SuTB in sows by approximately 0.26 [0.07, 0.48] piglets but increased SuTB in gilts by an estimated 0.16 [0.09, 0.23] piglets (Table 4.2).

Other direct effects showed a moderate posterior probability of heterogeneity between sows and gilts. In particular, our results indicate an 88% posterior probability of a greater direct effect of BA on TB (i.e. BA → TB) in sows compared to gilts. For sows, every additional piglet BA increased TB by 0.91 [0.84,0.99] piglets, whereas for gilts, the increase was smaller at 0.86 [0.80, 0.92]. We further detected a 89% posterior probability of the direct effect of BABW on TB (i.e. BABW → TB) being more pronounced in gilts that in sows. For gilts, every additional

100 g BABW decreased TB in gilts by 0.14 [0.05, 0.22] piglets, whereas the corresponding 95% HPD interval for sows was broader and overlapped with zero. A similar pattern was observed on the long-term reproductive performance outcome SuTB, whereby the direct effect of BABW on SuTB was also more pronounced in gilts than in sows, with 81% posterior probability (Table 4.2). In contrast, the direct effects of BA on BABW and of BABW on WEI seemed to be more homogeneous across parity groups, as indicated by posterior probabilities of a parity difference being closer to 0.50 for both effects (Table 4.2).

Table 4.3 shows a posterior summary of the between-block (i.e. random-level) variance $\sigma^2_{v_{jj'}}$ on the corresponding structural coefficient $\lambda_{jj'}$; this variance component describes non-systematic sources of variability on the corresponding direct effects across the observed body weight gradient. Assuming the $\boldsymbol{v}_{jj'}$ to be multivariate normal for each jj'^{th} direct effect and using the posterior mean of the between-block variance $\sigma^2_{v_{jj'}}$ as a point estimate (Table 4.3) , one may anticipate direct effects to range by approximately $\pm 1.96\sqrt{\sigma^2_{v_{jj'}}}$ between the most extreme body-weight blocks, based on the Empirical Rule (Moore et al. 2018) (Table 4.3). For example, given posterior inference on $\sigma^2_{v_{BABW,GAIN}}$, one might expect a range of ±0.02 g BABW per kg GAIN between direct effects from the most extreme body-weight blocks. Overlaying this range on the posterior mean of $\lambda_{BABW,GAIN}$, one might expect the direct effect GAIN → BABW to range from 0.05 to 0.09 g BABW per kg GAIN in-sows and from 0.01 to 0.05 g BABW per kg GAIN in gilts. Furthermore, given posterior inferences on $\sigma^2_{v_{TB,BABW}}$, $\sigma^2_{v_{WEI,BABW}}$, and $\sigma^2_{v_{SuTB,BABW}}$, one might anticipate a range of ±0.81 piglet SuTB per 100g BABW, of ±1.02 day WEI per 100g BABW, and of ±1.87 piglet per 100g BABW between direct effects from the most extreme body-weight blocks, respectively (Table 4.3).

Finally, given posterior inference on $\sigma^2_{v_{SuTB,WEI}}$, one might expect a range of ±0.13 piglet per 100 g BABW between direct effects from the most extreme body-weight blocks (Table 3). Centered on overall posterior mean of $\lambda_{SuTB,WEI}$, one might expect the direct effect WEI → SuTB ranges from -0.39 to -0.13 piglet in sow litter, and in turn from 0.03 to 0.29 piglet in gilt litter for each additional 100 g in BABW.

Remaining random-level variances characterizing non-systematic variability across body-weight blocks on structural coefficients were negligible in magnitude (Table 4.3) and are not discussed further.

4.3.2.2. Indirect and total effects

In addition to direct effects, SEM allow assessment of indirect effects and total effects between outcomes. Under linearity assumptions, the indirect effect of one outcome to another can be computed as the product of the structural coefficients on a directed path connecting those outcomes, and the total effect can be expressed as the sum of the direct effect and all indirect effects between them, as described by Wright (1934) and Shipley (2002).

The total effect of BA on TB, consisting of the corresponding direct effect (reported in the previous section) and the indirect effect mediated by BABW (i.e. BA → BABW → TB), can be expressed as $\lambda_{TB,BA}. + (\lambda_{TB,BABW} \times \lambda_{BABW,BA})$. Posterior inference shows total effect of BA on TB of 0.94 [0.88, 1.00] and of 0.91 [0.87, 0.95] piglets per additional piglet BA in sow and gilt litters, respectively, with a moderate posterior probability of 0.75 for difference between parity groups.

For long-term reproductive performance, we consider the total effect of BA on SuTB, which consisted of the sum of the indirect effects mediated by BABW (i.e. BA → BABW → SuTB) and by BABW and WEI (i.e. BA → BABW → WEI → SuTB). As such, this total effect

can be expressed as $(\lambda_{SuTB,BABW} \times \lambda_{BABW,BA}) + (\lambda_{SuTB,WEI} \times \lambda_{WEI,BABW} \times \lambda_{BABW,BA})$ and was estimated at 0.10 piglets [0.01, 0.21] and 0.05 [0.04, 0.23] SuTB per additional BA for sows and for gilts, respectively. We further detected an 81% posterior probability of the BABW-mediated effect of BA on TB (i.e. BABW → TB) being more pronounced in sows that in gilts. Remaining total effects were negligible in magnitude (results not shown) and are not discussed further.

4.3.2.3. Heteroskedasticity

Heterogeneity of residual variances across body-weight blocks was substantial for reproductive performance outcomes GAIN and WEI, moderate for BA, and negligible for TB and SuTB, as indicated by posterior summaries on the corresponding α_{e_j} (Table 4.4). We also interpret the coefficient of variation CV for residual variances between body-weight blocks, which is defined as $CV_j = 1/\sqrt{\alpha_{e_j} - 2}$. Hence, the magnitude of the heterogeneous residual variances across levels of body-weight blocks diminishes (i.e. $\sigma^2_{w_{e_j}} \to 0$) with larger values of α_{e_j} (i.e. $\alpha_{e_j} \to \infty$) (Kizilkaya and Tempelman 2005).

The magnitude of the heterogeneity of residual variances across body weight blocks was largest for GAIN and WEI, as indicated by upper bounds of the 95% HPD of the $\alpha_{v_{GAIN}}$ and $\alpha_{v_{WEI}}$ below 4 and 2, respectively. Indeed, the CV of residual variances across blocks was concentrated above 0.77 and 1.12, respectively. In fact, the largest and smallest posterior means on elements of the body-weight block effects $\boldsymbol{w}_{e_{GAIN}}$ and $\boldsymbol{w}_{e_{WEI}}$ were (85.84 and 0.52), and (433.95 and 0.69), for estimated fold changes of 165 and 628, respectively (Table 4.4).

For the rest of the reproductive outcomes, the magnitude of residual heteroskedasticity across body weight blocks was moderate (i.e. BA) to low (i.e. TB, SuTB). For BA, the posterior mean of $\alpha_{e_{BA}}$ was 13.5 with a posterior mean CV of 0.39. In turn, for TB and SuTB, the

posterior means of the corresponding $\alpha_{v_{TB}}$ and $\alpha_{v_{SuTB}}$ were in the hundreds (Table 4.4), indicating mild, if any, heterogeneity of residual variances across body weight blocks.

4.4. DISCUSSION

In this study, we proposed a general methodological approach that extends structural equation models to explicitly specify heterogeneity of structural coefficients in the context of network modeling. More specifically, structural coefficients were parameterized as a function of systematic and non-systematic sources of variability in a mixed-models framework. Our general approach relaxes the assumption of causal homogeneity, thereby enabling us to capture heterogeneous structural coefficients when present. As the current hierarchical SEM framework (Gianola and Sorensen 2004) makes the rather strong assumption of causal homogeneity, thereby forcing our preliminary analyses of the swine data application to be based on separate models fitted to each parity group (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation).The proposed methodological extension was implemented using a hierarchical Bayesian approach and this facilitated integration of hierarchical data architecture due to treatment structure and design structure, as is commonly the case for designed experiment. We used a simulation study to validate the inferential properties of the proposed methodological development and applied it to a dataset from a designed experiment in swine reproduction. Explicitly accommodating these sources of heterogeneity in structural coefficients is likely to often additional benefits as it can enhance more insightful information, thus allowing for effective management and decision strategies in specific subpopulations. Further, this data application illustrated additional subject-matter insights that can be gained from incorporating heterogeneity into network modeling, thereby enhancing understanding of the underlying mechanisms of a complex animal production system.

Previous work has proposed tailored approaches to incorporate heterogeneity into the general structural equation framework. For instance, Wu et al. (2007) extended a simultaneous and recursive (SIR) model to specify the whole matrix of structural coefficients $\mathbf{\Lambda}$ to be unique to a limited number of pre-defined subpopulations. Keeping with a recursive SEM, we generalize the proposal of Wu et al. (2007) and allow for a more flexible specification of structural coefficients that can accommodate multifactorial treatment and design structures with crossed or nested combinations. Moreover, our approach is computationally simpler as it relies on a Gibbs sampler with fully recognizable full conditional densities, as opposed to the more intensive implementation in Wu et al. (2007) that relied on a series of Metropolis-Hastings algorithms.

Within our proposed implementation, it also seemed feasible to diagnose cases that might warrant specifying an SEM with heterogeneous structural coefficients based on model fit comparisons. Indeed, our simulation study showed that when data were generated under conditions of network heterogeneity (i.e. simulation scenario B), DIC-based differences in global fit between SEM that assumed structural coefficients as homogeneous (i.e. M1) or heterogeneous (i.e. M2* or M2) were easily detectable, with DIC consistently selecting the appropriate model. In fact, DIC seemed to be quite sensitive a diagnostic tool as model fit indicated preference for a SEM with heterogeneous structural coefficients even when only a few of the coefficients were actually heterogeneous; this was supported by the comparison of M2 vs. M1 in simulation scenario B. In turn, an underspecified SEM (i.e. M0 and M1 vs. M2* in scenario B) showed clear evidence for impaired fit and led to decreased estimation accuracy not only for structural coefficients but also for mean-level location parameters, thus emphasizing the need for a more general SEM specification. In contrast, the DIC was also able to identify cases in which a SEM with the standard assumption of structural homogeneity sufficed (i.e. simulation scenario A),

thus preventing unnecessary model complexity. These results suggest that DIC-based model fit may be used to identify departures from the assumption of structural homogeneity in SEM. This desirable diagnostic behavior of DIC is consistent with that already observed in SEM applications (Valente et al. 2010, Inoue et al. 2016, Chitakasempornkul et al. 2018b) as well as in other hierarchical multivariate models that specify multi-level heterogeneity of (co)variance parameters (Bello et al. 2010) or functions thereof (Bello et al. 2013).

Despite the encouraging results for using DIC as a diagnostic tool in SEM, challenges remain, particularly for assessing directionality of network links. In our data application, power for assessing directionality of the link connecting BA and BABW was low. The small magnitude of model fit differences between alternative causal structures suggested that DIC was inconclusive to support selection of one causal structure as better fitting over the other. This is also consistent with our previous work (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation), for which the data were admittedly least informative to infer upon directionality of links.

Previous work from our group illustrated the importance of correctly recognizing hierarchical data architecture in the context of SEM (Chitakasempornkul et al. 2018b) (Chapter 2 of this dissertation). Specifically, the fit of SEM that ignore or mis-specify data architecture is likely to suffer, thus leading to impaired inference due to bias and loss of precision in the estimation process. The practical implications of correctly specifying data architecture are particularly relevant for agricultural applications in which designed experiments with data characterized by treatment structure and a design structure are common. The current study further expands this recommendation for SEM to account for data architecture not only on mean-level location parameters but also on structural coefficients. Our proposed methodological

extension contributes to the current SEM literature by explicitly specifying a hierarchical model on the non-zero structural coefficients that defines functional links in a multivariate network. In so doing, we argue for extending integrity of the inferential scope beyond mean-level location parameters (Bello and Renter 2018).

Specific to the data application, our results indicated evidence for heterogeneity of structural coefficients, some of which differed by parity groups and across body weight blocks. Most notable, the direct effect of WEI → SuTB differed between sows and gilts, and this was consistent with preliminary analyses that fitted separate SEM to each parity (Chitakasempornkul et al. 2018a) (Chapter 3 of this dissertation). From a biological standpoint, this may be explained by the fact that gilts are young females with competing nutritional requirements for growth and reproductive performance (Kraeling and Webel 2015), as opposed to adult size sows. Notably, our proposed methodological extension enabled us to formalize direct inference on the heterogeneity of this coefficient based on a flexible and more general SEM specification that incorporates all data (as opposed to separate analyses of data subsets). Further, accounting for a design structure was important because BW blocks identified the unit of experimental replication for parity, which is the factor that defines heterogeneity in this case. Indeed, our methodological extension to SEM can certainly recognize different sizes of experimental unit or a case where the units of observation and experimental replication are different. Therefore, this will ensure powerful inference on SEM structural coefficients that is also properly calibrated to the true amount of replication available, thus enhancing the inferential implications for functional links in a network

Additionally, heterogeneity of residual variances seems to be a common phenomenon in in animal agriculture in general (Cernicchiaro et al. 2013), and in swine production (Gonçalves et

al. 2016) in particular. If ignored, heteroskedasticity can lead to inefficient and possibly biased inference on mean-level parameters (Wiggans and VanRaden 1991). In our swine data application, evidence of heteroscedasticity was substantial, particularly for WEI. A potential approach to deal with heteroskedasticity might be to use variance stabilization transformation (Milliken and Johnson 2009). However, this approach would arguably complicate the interpretation of causal effects onto and from non-terminal nodes in the causal structure, as well as indirect effects involving such nodes. In addition, linearity of effects and normality of distributions would need to be assumed for transformed variables. Here, we expanded the proposed SEM with heterogeneous structural coefficients to also accommodate explicit modeling of heteroscedasticity, following the framework proposed by Kizilkaya and Tempelman (2005). Though details are not shown here, we did notice differences between M2 and M3 in the posterior summaries of the structural coefficients involving WEI. Therefore, it is recommended that sources of heterogeneity of residual variances are considered when investigating causal structure in data application.

Outcome variables of discrete nature are common in agricultural production systems. In our data application, BA, TB and SuTB can be formally defined considered as to be discrete counts. Normal approximations to discrete variables is often implemented to expedite model fitting and facilitate the estimation process (Stroup 2015); this was indeed our approach in this study. However, we recognize that inappropriately forcing a normal distribution approximation for data analysis can mislead the estimation process and hinder interpretation (Larrabee et al. 2014). Instead, proper recognition of the discrete nature of an outcome variables is increasingly promoted to ensure sound inference (Stroup 2013, 2015). In the context of SEM, one might consider incorporating threshold models (Sorensen et al. 1998) to the hierarchical Bayesian

framework, thus enabling the joint network analysis of continuous and discrete outcomes. Recent studies by López de Maturana et al. (2007), Konig et al. (2008) and Wu et al. (2008) jointly analyzed normal and non-normal outcomes using a threshold modeling extension to hierarchical SEM in genetic applications. Extensions that incorporate heterogeneous structural coefficients to hierarchical threshold SEM for non-normal responses is warranted, though challenges with interpretation are to be expected due to back-transformation from the link scale (Gianola 1982).

Lastly, we revisit the causal interpretation that can be attached to structural coefficients in an SEM and recognize the need for non-trivial causal assumptions that are not testable from data. Specifically, we refer to the Markov condition, faithfulness or stability and causal sufficiency (Pearl 2009). The latter, causal sufficiency, is probably the most important and the most controversial one, as it implies that all relevant players in the network are identified, measured and/or controlled for, thus circumventing any confounding effects. In the context of a designed experiment, the randomization only substantiates causal effects of treatments on each individual outcome variable, but not on casual claims from one outcome to another. Hence, causal assumptions are still needed to interpret structural coefficients in an SEM as functional network links, even in a designed experiment.

4.5. CONCLUSIONS

In this study, we propose a general approach for explicit specification of heterogeneous structural coefficients in network-type structural equation models. The proposed methodological extension is implemented in a hierarchical Bayesian framework and can accommodate structural coefficients as a function of systematic and non-systematic sources of variability, as would be expected from a designed experiment. Our proposed extended SEM can reliably infer upon heterogeneity of network links. Relevant applications include complex systems, including

agricultural production, for which an in-depth understanding of the mechanism underlying the system are critical to efficient management and optimal decision-making.

ACKNOWLEDGEMENTS

This research was performed using the Beocat Research Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860 and EPS-0919443. This research was also performed using the computing resources and assistance from the UW-Madison Center for High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison, the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

REFERENCES

Bello, N. M., Steibel, J. P., Erskine, R. J., and Tempelman, R. J. 2013. "Cows and Herds Constitute Distinct Hierarchical Levels of Heterogeneity in the Variability of and Association between Milk Yield and Pregnancy Outcome in Dairy Cows," *Journal of Dairy Science*, 96(4), 2314-2326.

Bello, N. M., Ferreira, V. C., Gianola, D., and Rosa, G. J. M. 2018. "Conceptual Framework for Investigating Causal Effects from Observational Data in Livestock," *Journal of Animal Science*, 96(10), 4045-4062.

Bello, N. M., Steibel, J. P., and Tempelman, R. J. 2010. "Hierarchical Bayesian Modeling of Random and Residual Variance–covariance Matrices in Bivariate Mixed Effects Models," *Biometrical Journal*, 52(3), 297-313.

Cernicchiaro, N., Renter, D. G., Xiang, S., White, B. J., and Bello, N. M. 2013. "Hierarchical Bayesian Modeling of Heterogeneous Variances in Average Daily Weight Gain of Commercial Feedlot Cattle," *Journal of Animal Science*, 91(6), 2910-2919.

Chitakasempornkul, K., Menegat, M. B., Rosa, G. J. M., Lopes, F. B., Jager,A., Gonçalves, M. A. D., Dritz, S. S., Tokach,M. D., Goodband, R. D., and Bello, N. M. 2018a. "Exploring Causal Biological Relationships between Reproductive Performance Traits in High-Performing Gilts and Sows," Conference on *American Dairy Science Association and American Society of Animal Science*, Omaha, NE, USA.

Chitakasempornkul, K, Sanderson, M. W., Cha, E., Renter, D. G., Jager, A., and Bello, N. M. 2018b. "Accounting for Data Architecture on Structural Equation Modeling of Feedlot Cattle Performance," *Journal of Agricultural, Biological and Environmental Statistics*, 23(4), 529-549.

Duncan, O. D. 1966. "Path Analysis: Sociological Examples," *American Journal of Sociology*, 72(1), 1-16.

Gelman, A. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models," *Bayesian Analysis*, 1(3), 515-533.

Gelman, A., and Rubin, D. B. 1992. "Inference from Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7(4), 457-472.

Gianola, D., and Sorensen, D. 2004. "Quantitative Genetic Models for Describing Simultaneous and Recursive Relationships between Phenotypes," *Genetics*, 167(3), 1407-1424.

Gianola, D. 1982. "Theory and Analysis of Threshold Characters," *Journal of Animal Science*, 54(5), 1079-1096.

Gianola, D., and Foulley, J. L. 1983. "Sire Evaluation for Order Categorical Data with a Threshold Model," *Genetics Selection Evolution*, 15(2), 201-224.

Godfray, H. C. J., Beddington, J. R., Crute, L. H., Lawrence, D., James, F. M., Pretty, J., Robinson, S., Thomas, S. M., and Toulmin, C. 2010. "Food Security: The Challenge of Feeding 9 Billion People," *Science*, 327(5967),812-818.

Gonçalves, M. A. D., Gourley, K. M., Dritz, S. S., Tokach, M. D., Bello, N. M., DeRouchey, J. M., Woodworth, J. C., and Goodband, R. D. 2016. "Effects of Amino Acids and Energy Intake during Late Gestation of High-performing Gilts and Sows on Litter and Reproductive Performance under Commercial Conditions," *Journal of Animal Science*, 94(5), 1993-2003.

Haavelmo, T. 1943. "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11(1), 1-12.

Inoue, K., Valente, B. D., Shoji, N., Honda, T., Oyama, K., and Rosa, G. J. M. 2016. "Inferring Phenotypic Causal Structures among Meat Quality Traits and the Application of a Structural Equation Model in Japanese Black cattle," *Journal of Animal Science*, 94(10), 4133-4142.

Johnson, R. A., and Wichern, D. W. 2007. *Applied Multivariate Statistical Analysis* (6th ed.), Upper Saddle River, New Jersey: Pearson Prentice Hall.

Joreskog, K.G. 1973. *A General Method for Estimating a Linear Structural Equation System*, . Edited by A. S. Goldberger and O. D. Duncan, *Equation Models in the Social Sciences*, New York: Senimar Press.

Kizilkaya, K., and Tempelman, R. J. 2005. "A General Approach to Mixed Effects Modeling of Residual Variances in Generalized Linear Mixed Models." *Genetics Selection Evolution*, 37(1), 31-56.

Konig, S., Wu, X. L., Gianola, D., Heringstad, B., and Simianer, H. 2008. "Exploration of Relationships between Claw Disorders and Milk Yield in Holstein Cows via Recursive Linear and Threshold Models," *Journal of Dairy Science*, 91(1), 395-406.

Kraeling, R. R., and Webel, S. K. 2015. "Current Strategies for Reproductive Management of Gilts and Sows in North America," *Journal of Animal Science and Biotechnology*, 6(1), 3.

Larrabee, B., Scott, M. H., and Bello.N. B. (2014). "Ordinary Least Squares Regression of Ordered Categorical Data: Inferential Implications for Practice," *Journal of Agricultural, Biological, and Environmental Statistics*, 19(3), 373-386.

Lauritzen, S. L. 1996. *Graphical models*. Oxford, UK: Oxford University Press.

Limpert, E., and Stahel, W. A. 2011. "Problems with Using the Normal Distribution – and Ways to Improve Quality and Efficiency of Data Analysis," *PLoS One*, 6(7), e21403.

López de Maturana, E. L., Legarra, A., Varona, L., and Ugarte, E. 2007. "Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data," *Journal of Dairy Science*, 90(4), 2012-2024.

Milliken, G. A. and Johnson, D. E. 2009. *Analysis of Messy Data - Volume 1: Designed Experiments* (2nd ed.), Boca Raton, Florida, USA: Chapman and Hall/CRC Press.

Moore, D. S., Notz, W. I., and Fligner, M. A. 2018. *The Basic Practice of Statistics* (8th ed), New York, NY: W. H. Freeman Macmillan Learning.

Ou, Z., Tempelman, R. J., Steibel, J. P., Ernst, C. W., Bates, R. O., and Bello, N. M. 2016. "Genomic Prediction Accounting for Residual Heteroskedasticity,"*Genes Genomes Genetics*, 6(1), 1-13.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inference* (2nd ed.), Cambridge University Press.

Plummer, M., Best, N., Cowles, K., and Vines, K. 2006. "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News*, 6, 7-11.

Raftery, A. and Lewis, S. 1992. "How many iterations in the Gibbs sampler," *Bayesian Statistics*, 4.

Robinson, G. K. 1991. "That BLUP is a Good Thing: The Estimation of Random Effects," *Statistics Science*, 6(1), 15-32.

Rosa, G. J. M., and Valente, B. D. 2013. "Breeding and Genetics Symposium: Inferring Causal Effects from Observational Data in Livestock," *Journal of Animal Science*, 91(2), 553-564.

Rosa, G. J. M., Valente, B.D., de los Campos, G., Wu, X. L., Gianola, D., and Silva, M. A. 2011. "Inferring Causal Phenotype Networks Using Structural Equation Models," *Genetics Selection Evolution*, 43, 6-18.

Shipley, B. 2002. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*, Cambridge University Press.

Sorensen, D., Andersen, S., Gianola, D., and Korsgaard., I. 1995. "Bayesian-inference in Threshold Models Using Gibbs Sampling, *Genetics Selection Evolution*, 27(3), 229-249.

Sorensen, D., Gianola, D., and Korsgaard, I. R. 1998. "Bayesian mixed-effects Model Analysis of a Censored Normal Distribution with Animal Breeding Applications," *Agriculturae Scandinavica Section a-Animal Science*, 48(4), 222-229.

Sorensen, D. and Gianola, D. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*, New York, Springer-Verlag.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. 2002. "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.

Stroup, W. 2013. *Generalized Linear Mixed Models*, Boca Raton, Florida, CRC Press Taylor & Francis Group.

Stroup, W. W. 2015. "Rethinking the Analysis of Non-Normal Data in Plant and Soil Science." *Agronomy Journal*, 107(2), 811-827.

R Development Core Team. 2016. R: A Language and Environment for Statistical Computing, Version 3.3.1 In *R, Foundation for Statistical Computing*.

Tempelman, R. J., Spurlock, D. M., Coffey, M., Veerkamp, R. F., Armentano, L. E., Weigel, K. A., de Haas, Y., Staples, C. R., Connor, E. E., Lu, Y., and VandeHaar, M. J. 2015. "Heterogeneity in Genetic and Nongenetic Variation and Energy Sink Relationships for Residual Feed Intake across Research Stations and Countries," *Journal of Dairy Science*, 98(3):2013-2026.

Valente, B. D., Rosa, G. J. M., Silva, M. A., Teixeira, R. B., and Torres, R. A. 2011. "Searching for Phenotypic Causal Networks Involving Complex Traits: an Application to European quail," *Genetics Selection Evolution*, 43(1), 37.

Valente, B. D., Rosa, G. J. M., de Los Campos, G., Gianola, D., and Silva, M. A. 2010. "Searching for Recursive Causal Structures in Multivariate Quantitative Genetics Mixed Models," *Genetics*, 185(2), 633-644.

Verma, T., and Pearl, J. 1991. "Equivalence and synthesis of causal models," Proceedings of the Sixth Annual Conference on *Uncertainty in Artificial Intelligence*, New York, Elsevier Science Inc, p. 255-270.

Wiggans, G. R., and VanRaden, P. M. 1991. "Method and Effect of Adjustment for Heterogeneous Variance," *Journal of Dairy Science*, 74(12), 4350-4357.

Wright, S. 1934a. "An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs," *Genetics*, 19(6), 506-536.

Wright, S. 1921. "The Method of Path Coefficients," *Annals of. Mathtematics and Statistics*, 5(3), 161-215.

Wu, X. L., Heringstad, B., Chang, Y. M., de Los Campos, G., and Gianola, D. 2007. "Inferring Relationships between Somatic Cell Score and Milk Yield Using Simultaneous and Recursive Models," *Journal of Dairy Science*, 90(7), 3508-3521.

Wu, X. L., Heringstad, B., and Gianola, D. 2008. "Exploration of Lagged Relationships between Mastitis and Milk Yield in Dairy Cows Using a Bayesian Structural Equation Gaussian-threshold Model," *Genetics Selection Evolution*, 40(4), 333-357.

Wu, X. L., Heringstad, B., and Gianola, D. 2010. "Bayesian Structural Equation Models for Inferring Relationships between Phenotypes: a Review of Methodology, Identifiability, and Applications," *Journal of Animal Breeding and Genetics*, 127(1), 3-15.

Yang, W., and Tempelman, R. J. 2012. "A Bayesian Antedependence Model for Whole Genome Prediction," *Genetics*, 190(4), 1491-1501.

FIGURES AND TABLES

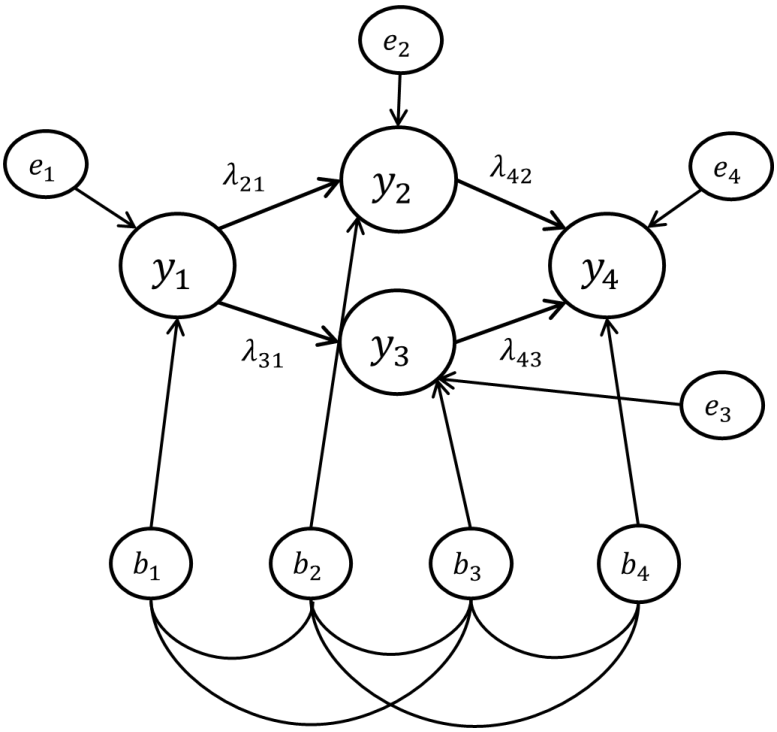


Figure 4.1 Directed acyclic graph depicting putative functional links between outcome variables y_1, y_2, y_3 and y_4. Directed links with arrowheads indicate that y_1 has a direct effect on y_2 (i.e. $y_1 \rightarrow y_2$) and on y3 (i.e. $y_1 \rightarrow y_3$). Both y_2 and y_3 have a direct effect on y_4 (i.e. $y_2 \rightarrow y_4$ and $y_3 \rightarrow y_4$, respectively). Each outcome variable is directly affected by a corresponding residual, e_1, e_2, e_3 and e_4 and by mutually correlated random effects b_1, b_2, b_3 and b_4. Correlations between random effects are indicated by undirected arcs.

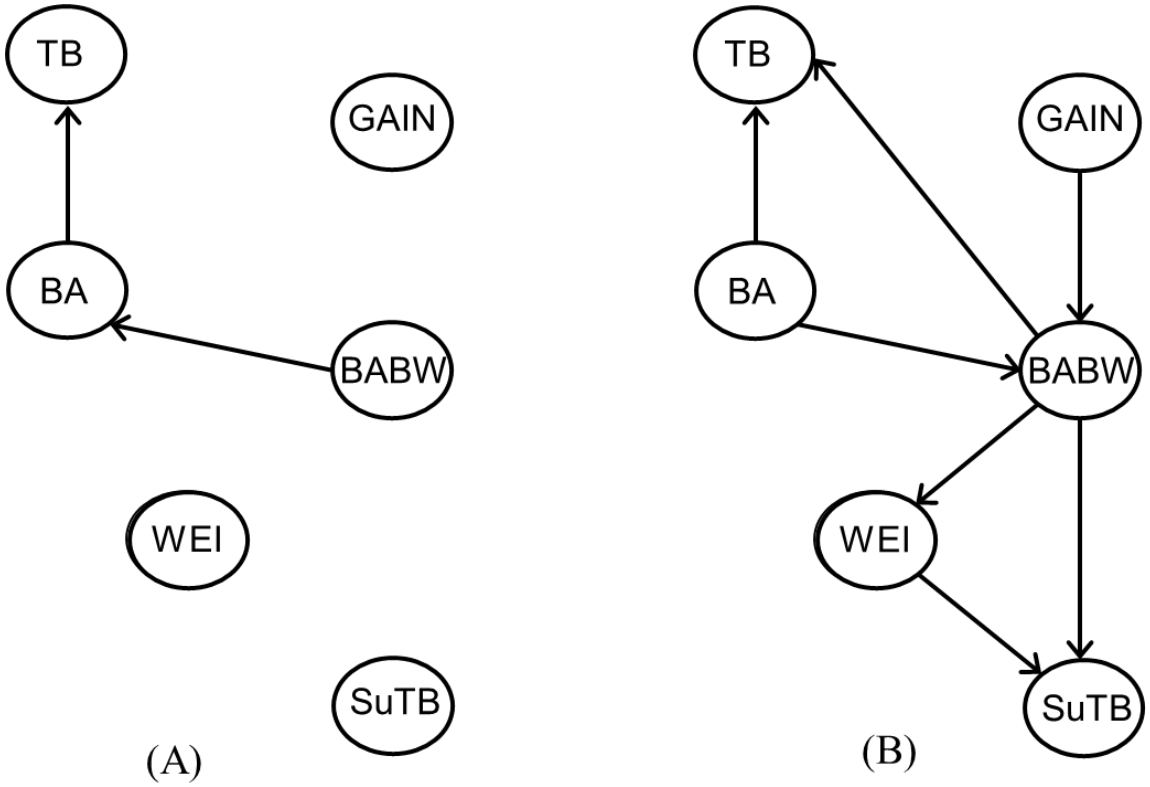
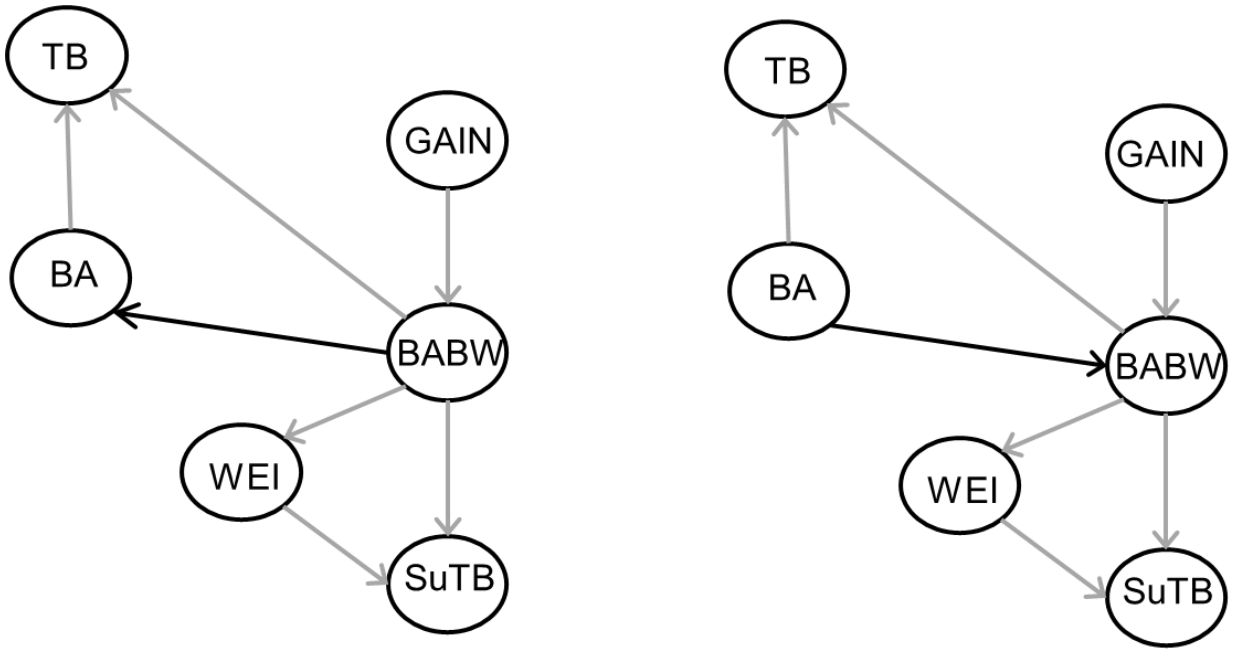


Figure 4.2 Results of preliminary analyses (Chitakasempornkul et al. 2018a) showing functional links between reproductive performance outcomes in sows (A) and in gilts (B) obtained using the mixed-model adapted inductive causation algorithm (Valente et al. 2010) implemented with 80% highest posterior density intervals. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.



(A) (B)

Figure 4.3 Alternative network structures connecting reproductive performance outcomes in the swine data. The black arrow indicates alternative directions of the link between BABW and BA. (A) Causal effect of BABW on BA (i.e. BA ← BABW). (B) Causal effect of BA on BABW (i.e. BA → BABW). GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

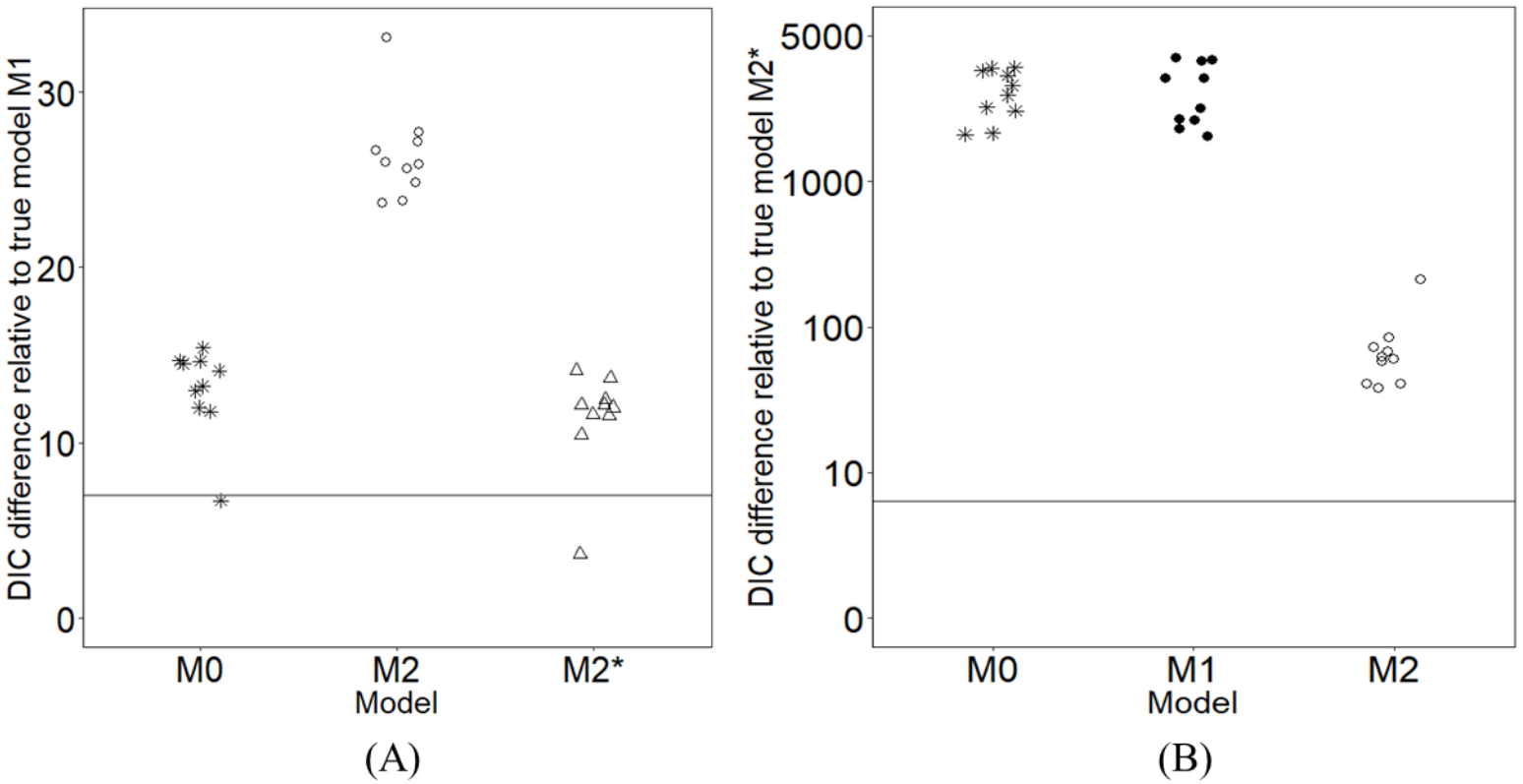


Figure 4.4 Model fit comparison, expressed as differences in Deviance Information Criteria (DIC) relative to the true data-generation model under simulation scenario A (i.e., Panel A: homogeneous structural coefficients, true model = M1; Black dots) and simulation scenario B (i.e. Panel B: heterogeneous $\boldsymbol{\lambda_{21,i}}$ and $\boldsymbol{\lambda_{43,i}}$, true model = M2*; Triangles). M0 = Fully recursive SEM, equivalent to a traditional multiple-trait model (Stars). M1 = SEM with homogeneous specification of structural coefficients defining a pre-selected causal structure; M2 = SEM with heterogeneous specification of all $\boldsymbol{\lambda_{jj',i}}$ for $\boldsymbol{j' < j}$, j = 1, 2, …, J and i = 1, 2, …, n defining a pre-selected causal structure (White dots)*;* M2* = SEM with heterogeneous specification of structural coefficients $\boldsymbol{\lambda_{21,i}}$ and $\boldsymbol{\lambda_{43,i}}$ only defining a pre-selected causal structure. Each point represents a dataset generated under the corresponding scenario and fitted with one of SEM M0, M1, M2 or M2*. A reference horizontal line represents a threshold value of 7 indicative of a substantial difference in model fit relative to the true data-generation model in each case.

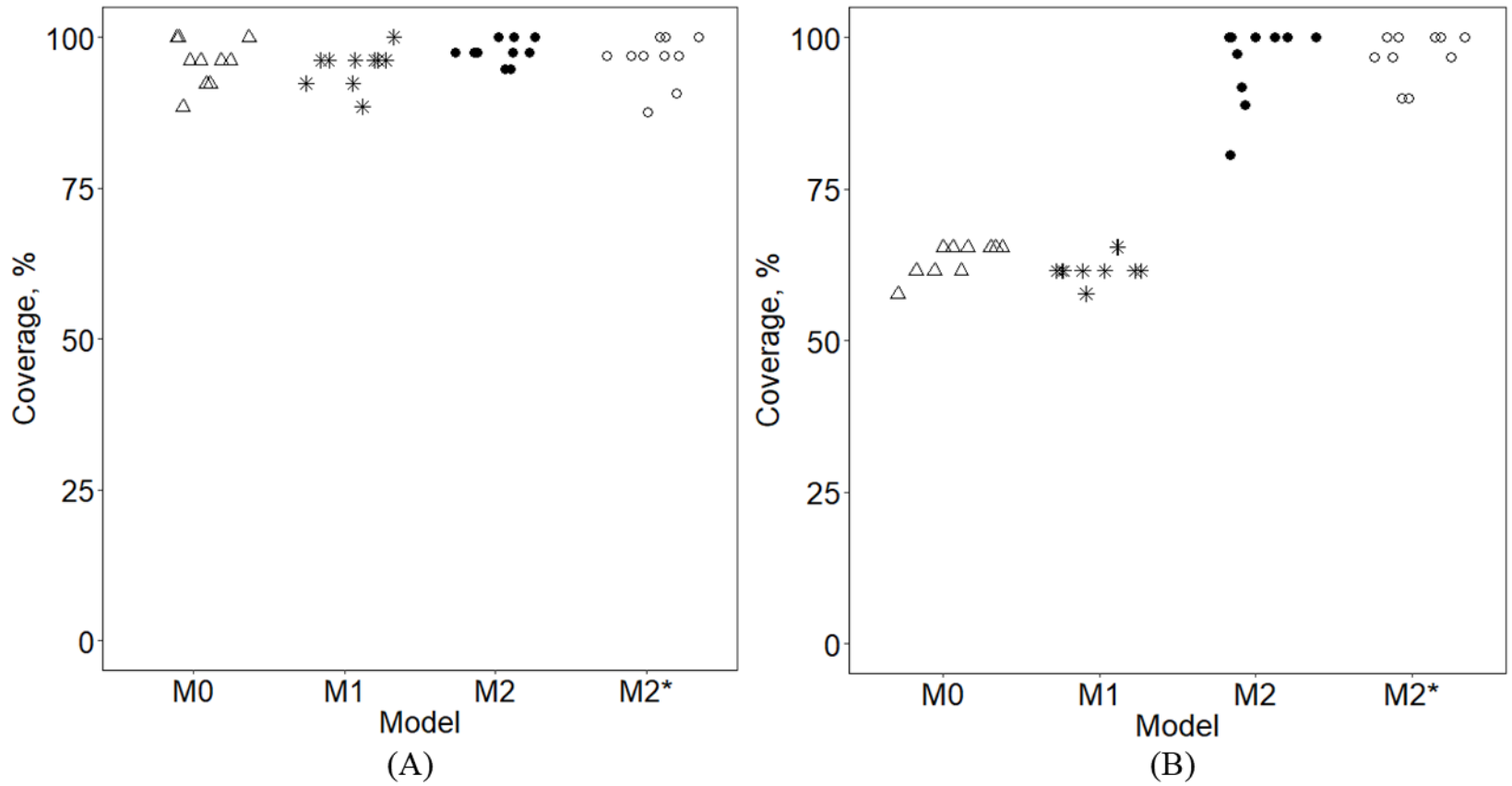


Figure 4.5 Overall accuracy of parameter estimation, defined as percent coverage (%) of 95% highest posterior density intervals under simulation scenario A (i.e., Panel A: homogeneous structural coefficients, true model = M1) and B (i.e. Panel B: heterogeneous $\boldsymbol{\lambda_{21,i}}$ and $\boldsymbol{\lambda_{43,i}}$, true model = M2*). M1 = SEM with homogeneous specification of structural coefficients defining a pre-selected causal structure; M2 = SEM with heterogeneous specification of all $\boldsymbol{\lambda_{jj',i}}$ for $\boldsymbol{j' < j}$, j = 1, 2, …, J and i = 1, 2, …, n defining a pre-selected causal structure; M2* = SEM with heterogeneous specification of structural coefficients $\boldsymbol{\lambda_{21,i}}$ and $\boldsymbol{\lambda_{43,i}}$ only defining a pre-selected causal structure. Each dot represents all parameters from a simulated dataset.

Table 4.1 Model fit of alternative SEM specifications and causal structures to swine data, expressed as deviance information criteria (DIC) and effective number of parameters (pD) (Spiegelhalter et al. 2002).

Alternative SEM specification	Alternative causal link	pD	DIC
M0	(A) BA ← BABW	310.4	17075.1
	(B) BA → BABW	310.2	17073.9
M1	(A) BA ← BABW	298.3	17044.2
	(B) BA → BABW	297.1	17043.7
M2	(A) BA ← BABW	389.2	17036.4
	(B) BA → BABW	387.3	17027.4
M3	(A) BA ← BABW	393.7	17019.1
	(B) BA → BABW	392.5	17014.7

For $j' < j,\ j = 1, 2, \ldots, J$ and $i = 1, 2, \ldots, n$.
M0: Fully recursive SEM, equivalent to a traditional multiple-trait model.
M1: SEM with homogeneous specification of structural coefficients defining a pre-selected causal structure.
M2: SEM with heterogeneous specification of all structural coefficients defining a pre-selected causal structure.
M3: SEM with heterogeneous specification of all structural coefficients defining a pre-selected causal structure and heterogeneous residual variances $\sigma^2_{e_j}$.
BA: Number born alive in a litter; BABW: Born alive average body weight.

Table 4.2 Posterior summary of structural coefficients for sows and gilts of typical body-weight block based on the best-fitting SEM selected for final inference.

			Sows	Gilts	Difference Gilts - Sows	
Structural coefficient	From	To	Posterior mean [95% HPD]	Posterior mean [95% HPD]	Posterior mean [95% HPD]	Posterior probability*
$\lambda_{BABW,GAIN}$, g per kg	GAIN	BABW	0.07 [0.02, 1.11]	0.03 [-0.02, 0.57]	0.04 [-0.09, 0.14]	0.93
$\lambda_{BABW,BA}$, g per unit	BA	BABW	-356.3 [-447.5, -267.7]	-358.1 [-421.7, -291.8]	-1.80 [-2.02, 1.03]	0.51
$\lambda_{TB,BA}$, unit per unit	BA	TB	0.91 [0.84, 0.99]	0.86 [0.80,0.92]	-0.05 [-0.13, 0.04]	0.88
$\lambda_{TB,BABW}$, unit per 100g	BABW	TB	-0.64 [-1.64, 0.31]	-0.14 [-0.22, -0.05]	-0.5 [-0.80, 0.04]	0.89
$\lambda_{WEI,BABW}$, day per 100g	BABW	WEI	0.01 [-0.05, 0.07]	0.29 [-0.31, 0.83]	-0.17 [-0.61, 0.94]	0.66
$\lambda_{SuTB,BABW}$, unit per 100g	BABW	SuTB	-0.26 [-0.48, -0.06]	-0.39 [-0.64, -0.15]	-0.13 [-0.43, 0.18]	0.81
$\lambda_{SuTB,WEI}$, unit per day	WEI	SuTB	-0.26 [-0.48, -0.07]	0.16 [0.09, 0.23]	0.10 [0.08, 0.50]	0.99

"*" defined as $max\ \left(p\left(\lambda_{jj'}^{\text{Gilt}} - \lambda_{jj'}^{\text{Sow}} | y\right) > 0,\ p\left(\lambda_{jj'}^{\text{Gilt}} - \lambda_{jj'}^{\text{Sow}} | y\right) < 0\right)$.

GAIN: Female weight gain during late gestation.
BA: Number born alive in a litter.
BABW: Born alive average body weight.
TB: Total number born in a litter.
WEI: Wean-to-estrous interval.
SuTB: Total number born in the subsequent gestation.

Table 4.3 Posterior summary of random-level (i.e. between-block) variance $\sigma^2_{v_{jj'}}$ of structural coefficients linking reproductive performance outcomes in swine data, and their interpretation in terms of expected range of the corresponding $\lambda_{jj'}$ based on empirical rule.

Direct effect			Posterior mean [95% HPD]		Expected range of $\lambda_{jj'}$ based on empirical rule = $\left(\pm 1.96\sqrt{\sigma^2_{v_{jj'}}}\right)$
From	To	Between-block variance component			
GAIN	BABW	$\sigma^2_{v_{BABW,GAIN}}$	1.54×10^{-4} $[2.23\times10^{-10}, 5.38\times10^{-4}]$	$(\text{g per kg})^2$	±0.02 g per kg
BA	BABW	$\sigma^2_{v_{BABW,BA}}$	1.05×10^{-3} $[6.22\times10^{-12}, 3.91\times10^{-4}]$	$(\text{g per unit})^2$	±0.06 g per unit
BA	TB	$\sigma^2_{v_{TB,BA}}$	6.98×10^{-5} $[2.18\times10^{-11}, 2.68\times10^{-4}]$	$(\text{unit per unit})^2$	±0.02 unit per unit
BABW	TB	$\sigma^2_{v_{TB,BABW}}$	0.17 $[7.05\times10^{-8}, 0.58]$	$(\text{unit per 100g})^2$	±0.81 unit per 100g
BABW	WEI	$\sigma^2_{v_{WEI,BABW}}$	0.27 $[8.24\times10^{-6}, 0.98]$	$(\text{day per 100g})^2$	±1.02 day per 100g
BABW	SuTB	$\sigma^2_{v_{SuTB,BABW}}$	0.91 $[7.21\times10^{-7}, 3.44]$	$(\text{unit per 100g})^2$	±1.87 unit per 100g
WEI	SuTB	$\sigma^2_{v_{SuTB,WEI}}$	4.64×10^{-3} $[6.96\times10^{-8}, 1.65\times10^{-2}]$	$(\text{unit per day})^2$	±0.13 unit per day

GAIN: Female weight gain during late gestation.
BA: Number born alive in a litter.
BABW: Born alive average body weight.
TB: Total number born in a litter.
WEI: Wean-to-estrous interval.
SuTB: Total number born in the subsequent gestation.

Table 4.4 Posterior summary of heterogeneity of residual-level variances σ_e^2 for reproductive performance outcomes GAIN, BA, TB, WEI and SuTB, obtained from the structural equation model selected for final inference based on best fit.

Performance outcome	Posterior mean of $\sigma^2_{e_{ref}}$ [95% HPD]	Posterior mean of α_e [95% HPD]	Posterior mean of CV [95% HPD]
GAIN	22.10	2.79	0.77
	[17.96, 26.65]	[1.83, 3.90]	[0.55, 1.01]
BA	5.49	13.49	0.39
	[4.72, 6.31]	[2.44, 37.19]	[0.13, 0.63]
TB	1.02	184.90	0.12
	[0.87, 1.16]	[8.67, 660.93]	[0.02, 0.22]
WEI	0.65	1.80	1.12
	[0.51, 0.79]	[1.68, 1.92]	[1.04, 1.21]
SuTB	16.28	408.60	0.10
	[13.44, 19.44]	[7.25, 1466.79]	[0.01, 0.21]

Coefficient of variation (CV) = $1/\sqrt{\alpha_e - 2}$ for reproductive performance outcomes GAIN, BA, TB, WEI and SuTB.
GAIN: Female weight gain during late gestation.
BA: Number born alive in a litter.
TB: Total number born in a litter.
WEI: Wean-to-estrous interval.
SuTB: Total number born in the subsequent gestation.

Chapter 5 - Conclusion

Understanding interrelationships between multiple outcomes of interest is increasingly important in agricultural production systems. To do that, a system approach toward animal production systems based on hierarchical SEM provides an in-depth insight into the mechanisms of animal agricultural production systems and how they interact altogether. In this dissertation, I pursue to characterize and address challenges associated with hierarchical SEM modeling in animal agriculture. In Chapter 2, I evaluated the inferential implications of properly accounting for (or ignoring) existent correlation structure due to data architecture in a SEM approach using a simulation study and for a proof-of-concept evaluation, applied to beef cattle data application. In Chapter 3, I investigated the space of causal structure and estimated and inferred upon potential functional links that are connected between reproductive outcomes in sows and gilts in the context of a designed nutrition experiment. Also, I evaluated the stability of the learned network structures and showed how to investigate potential differences in statistical power between the parity group. In Chapter 4, I developed hierarchical Bayesian extensions to SEM to accommodate explicit specification of heterogeneous structural coefficients as a function of systematic and non-systematic sources of variation as results from Chapter 3 suggested heterogeneous interconnections in the swine system. Further, I validated the properties of our proposed methodological extensions using a simulation study and applied our proposed extension to data application in a swine production system.

5.1. Address of specific objectives

1) To evaluate the inferential implications of accounting for (or ignoring) existent correlation structure due to data architecture in a SEM approach.

The inferential implications presented in Chapter 2 of this dissertation address this goal in two respects. First, the contribution shows the potential consequences of incorrectly specifying data architecture in a SEM framework. Second, particularly in the context of observational study, it represents a novel application in animal production (i.e. feedlot data) that is of great importance in the beef cattle industry. Further, the study in Chapter 2 represents recent recognition of data architecture in the SEM literature and provides statistical recommendations that require proper model specification accounting for relevant hierarchical levels in data architecture extend to the context of SEM-based modeling. However, data architecture is limited to clusters which are analogous to blocks in experimental design setting; more complex components of data architecture included nested effects, split plots, and subsampling are represented in Chapter 4, as described later. Also, the causal structure of SEM in feedlot application is elicited on the epidemiological knowledge and supported by Cha et al. (2017). Hence, the search algorithm (Valente et al. 2010) with accounting for data architecture in designed swine nutrition experiment is considered in Chapter 3.

2) Investigating causal biological relationships between reproductive performance outcomes in high-performing gilts and sows.

This specific objective was accomplished by using the IC search algorithm adapted to a mixed-model SEM framework (Valente et al. 2010) to investigate potential biological interconnections between performance outcomes. Indeed, results from Chapter 2 provide insight into potentially distinct causal networks of interconnected reproductive performance outcomes in sows and gilts. However, the current hierarchical SEM framework (Gianola and Sorensen 2004) makes the assumption of causal homogeneity. This strong assumption inevitably forced the preliminary analysis of investigating causal interconnections between outcomes to be based on

separated models fitted to each subpopulation (i.e. sows and gilts). Hence, the heterogeneous nature of the networks could be lost to inference; this constraint is circumvented in Chapter 4. It should be noted that the search algorithm that accounts for data architecture provides a great deal of learned network structures between sows and gilts, thus leading to a new methodological extension to accommodate the sources of heterogeneity in hierarchical SEM.

3) A general approach for hierarchical modeling of heterogeneous structural coefficients in structural equation models.

The development of a methodological extension of SEM to accommodate heterogeneous structural coefficients in a network successfully addresses a questionable concern on the strong assumption of causal homogeneity. In Chapter 3, I proposed the methodological extension to hierarchical SEM in a more complex designed experiment and implemented the most general network structure that obtained from Chapter 3. The need to incorporate the nature of heterogeneous interconnections between outcomes is driven by the search results as described in Chapter 2. In particular, the motivating question of this dissertation calls for the assumption of causal homogeneity, a causal network can be differed across subjects, and it is of importance in investigating interconnections of animal production performance in a way that is meaningful to farm management.

5.2. Implications associated with SEM-based modeling in animal agriculture

Results from Chapter 2, 3 and 4 in this dissertation address challenges associated with hierarchical SEM. Results from this dissertation indicate three important aspects. First, data architecture is not uncommon in animal agriculture data for both an observational study and a designed experiment. Specific objectives (1) in this dissertation (Chitakasempornkul et al. 2018) suggests that when data architecture is recognized and properly modeled, hierarchical SEM

inference is enhanced; if not, impaired. Gbur (2012) and Stroup (2013) described that animal agriculture data usually show a well-defined hierarchical structure, by which observations at animal level generally share a common background and are thus not mutually independent and correlated within pens or cohorts; yet, pens or cohorts within a farm or operation. This novel is immensely important to capture applicable hierarchical levels in data architecture in a mixed-model framework (Dohoo et al. 2014). Yet, the results from Chapter 2 of this dissertation (Chitakasempornkul et al. 2018) certainly indicate a similar recommendation that requires careful consideration of the data generation process and verification of correlation structures in animal agricultural data.

Results from Chapter 3 in this dissertation indicate that the proposed methodological extension of hierarchical SEM can easily obtain direct inference on the heterogeneity of structural coefficients if apparent. Further, the proposed model provides a more flexible and general SEM specification that accommodates both observational and designed experimental data. In particular, based on Chapter 2, the model can readily recognize the different size of experimental unit or in the case where experimental unit and observational unit are different in a design experimental setting. Hence, this gains in more powerful SEM inference (Bello et al. 2016) and enhances the inferential implication of functional links in a causal network. Indeed, if sources of heterogeneity are existing, incorporating the heterogenous components into functional links ultimately enhance a more comprehensive, though in-depth insight, understanding of interrelations between outcomes in subpopulations in a complex system. In doing so, I believe this proposed approach attempts to provide realistic insight from which to design management recommendations targeted to optimize overall performance of animal agricultural production.

Further statistical methodological extensions will undeniably be used to improve and clarify what management recommendations may be in specific circumstances in animal agriculture.

5.3. Opportunities for future research

The research presented in this dissertation provides a groundwork of SEM-based modeling for future works in the domain of statistical methodology as well as that of the recommended utilization in animal agriculture. Further, these works could also be implemented and applied to other subject-matter applications across a wide range of sciences. In animal agriculture research, many outcome variables of importance are of a discrete nature (Stroup 2015). For example, total litter size and number of born alive in pigs, calving difficulty in cattle, presence or absence of a disease, survival, or death. Generally, they are often modeled using normal approximation to facilitate the estimation exercise. Notably, Wu et al. (2008) proposed a Bayesian structural equation Gaussian-threshold model that assumed causal homogeneity to infer potential causal relationships between binary and continuous outcome variables. The hierarchical SEM-based modeling that relaxed assumption of causal homogeneity presented in Chapter 4 provides a general framework that can be easily established to accommodate non-normal outcome variables. The methodological extension is to combine the concept of threshold model (Wright 1934) with hierarchical SEM that accommodate the heterogeneity of structural coefficients in a system.

The concept of threshold models was first introduced in the context of quantitative genetics by Wright (1934) for the analysis of the number of toes in Guinea pigs. Threshold models assume a hypothetical liability expressed on a continuous scale that underlies the observed categorical outcome (Sorensen and Gianola 2002). The value of the liability relative to thresholds determine the realization of a categorical outcome. This liability or latent variable can

be shown to fully specify the likelihood of the categorical outcomes yet can be modeled as a Gaussian distribution. Just for analogy, a threshold model can be thought of as a generalized linear mixed model to a non-Gaussian variable, namely binary or ordered categorical that uses a probit link function to connect the linear predictor to the expected value of the categorical outcome (Sorensen et al. 1998). Moreover, working with liabilities facilitates sampling from FCD that now have a closed form and can thus be easily sampled using a Gibbs sampling algorithm (Sorensen and Gianola 2002, Korsgaard et al. 2003).

Another extension is that of accounting for missing data in hierarchical SEM modeling. Recent SEM-based modeling has limited to using complete records (i.e. no missing data for any outcomes of interest). This can lead to loss of valuable information in data, thereby providing losing in statistical power, especially for network learning. Missing data is a common problem in multivariate setting in learning networks. Potential way to incorporate missing value in outcome variables to the methodological extension of hierarchical SEM is using latent variables to augment information from data proposed by Albert and Chib (1993) and the concept of Bayesian latent variable model by Dunson (2000).

Further, integrating additional DNA polymorphism information and knowledge about genes is a great methodological extension to hierarchical SEM that is plausible to be of main interest to animal and plant breeders in their investigation to the heritable component of genetic architecture for correlated phenotypes and potential sources of heterogeneity. Potential approaches to integrate a genome-to-phenome causal network to the propose hierarchical SEM of this dissertation include an interesting proposal by Liu et al. (2008), the joint analysis of causal network and genetic architecture (Chaibub Neto et al. 2010), and incorporating design of

microarray experiments (Rosa and Vazquez 2009). Therefore, these methodological extensions to hierarchical SEM might usefully enhance more understanding of a complex system.

5.4. Conclusion

Overall, understanding the functional causality that underlie interconnections between outcomes can provide valuable mechanistic information and predictive ability to system changes, thereby facilitating efficient decision making in animal production system. A system approach based on a hierarchical SEM provides a powerful and flexible strategies for studying causality in the context of data from designed experiments and observational studies. Perhaps, the hierarchical SEM framework could be the best arena to answer research questions pertaining to how outcome variables are connected, structured and organized in a system, thus bringing a more comprehensive understanding of biological mechanisms.

REFERENCES

Bello, N. M., Kramer, M., Tempelman, R. J., Stroup, W. W., St-Pierre, N. R., Craig, B. A., Young, L. J., and Gbur, E. E. 2016. "Short Communication: On Recognizing the Proper Experimental Unit in Animal Studies in the Dairy Sciences," *Journal of Dairy Science*, 99(11), 8871-8879.

Cha, E., Sanderson, M. W., Renter, D., Jager, A., Cernicchiaro, N., and Bello, N. M. 2017. "Implementing Structural Equation Models to Observational Data from Feedlot Production Systems," *Preventive Veterinary Medicine*, 147,163-171.

Chaibub, N., E., Keller, M. P., Attie, A. D., and Yandell, B. S. 2010. "Causal Graphical Models in Systems Genetics: A Unified Framework for Joint Inference of Causal Network and Genetic Architecture for Correlated Phenotypes," *Annals of Applied Statistics*, 4, 320-339.

Albert, J. H., and Chib, S. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88(422), 669-679.

Chitakasempornkul, K., Sanderson, M. W., Cha, E., Renter, D. G., Jager, A., and Bello, N. M. 2018. "Accounting for Data Architecture on Structural Equation Modeling of Feedlot Cattle Performance," *journal of Agricultural, Biological and Environmental Statistics*, 23(4), 529-549.

Dohoo, I., Martin, W., and Stryhn, H. 2014. *Veterinary Epidemiologic Research* (2nd ed.), Canada: VER Inc.

Dunson, D. B. 2000. "Bayesian Latent Variable Models for Clustered Mixed Outcomes," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 62(2), 355-366.

Gbur, E. E., Stroup, W. W., McCarter, W., Kevin ,S., Durham, S., Young, L. J., Christman, M., West, M., Kramer, M. 2012. *Analysis of Generalized Linear Mixed Models in the Agricultural and Natural Resources Sciences*, Madison, WI, USA: American Society of Agronomy, Soil Science Society of America, Crop Science Society of America, Inc.

Gianola, D., and Sorensen, D. 2004. "Quantitative Genetic Models for Describing Simultaneous and Recursive Relationships between Phenotypes," *Genetics*, 167(3), 1407-1424.

Korsgaard, I. R., Lund, M. S., Sorensen, D., Gianola, D., Madsen, P., and Jensen, J. 2003. "Multivariate Bayesian Analysis of Gaussian, Right Censored Gaussian, Ordered Categorical and Binary Traits Using Gibbs sampling," *Genetics Selection Evolution*, 35(2), 159-183.

Liu, B., Fuente, A., and Hoeschele, I. 2008. "Gene Network Inference via Structural Equation Modeling in Genetical Genomics Experiments," *Genetics*, 178, 1763-1776.

Rosa, G. J. M., and Vazquez, A. I. 2009. "Integrating Biological Information into the Statistical Analysis and Design of Microarray Experiments," *Animal*, 4(2), 165-172.

Sorensen, D. A., Gianola, D., and Korsgaard, R. I. 1998. "Bayesian Mixed-effects Model Analysis of a Censored Normal Distribution with Animal Breeding Applications," *Acta Agriculturae Scandinavica, Section A-Animal Science*, 48(4), 222-229.

Sorensen, D., and Gianola, D. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*, New Your: Springer-Verlag.

Stroup, W. W. 2013. *Generalized Linear Mixed Models*, Boca Raton, FL: CRC Press Taylor & Francis Group.

Stroup, W. W. 2015. "Rethinking the Analysis of Non-Normal Data in Plant and Soil Science," *Agronomy Journal*, 107(2), 811-827.

Valente, B. D., Rosa, G. J. M., de Los Campos, G., Gianola, D., and Silva, M. A. 2010. "Searching for Recursive Causal Structures in Multivariate Quantitative Genetics Mixed Models," *Genetics*, 185(2), 633-644.

Wright, S. 1934. "An Analysis of Variability in Number of Digits in an Inbred Strain of Guinea Pigs," *Genetics*, 19 (6),506-536.

Wu, X. L., Heringstad, B., and Gianola, D. 2008. "Exploration of Lagged Relationships between Mastitis and Milk Yield in Dairy Cows Using a Bayesian Structural Equation Gaussian-threshold Model," *Genetics Selection Evolution*, 40(4), 333-357.

Appendix A - Supplementary Materials for Chapter 2

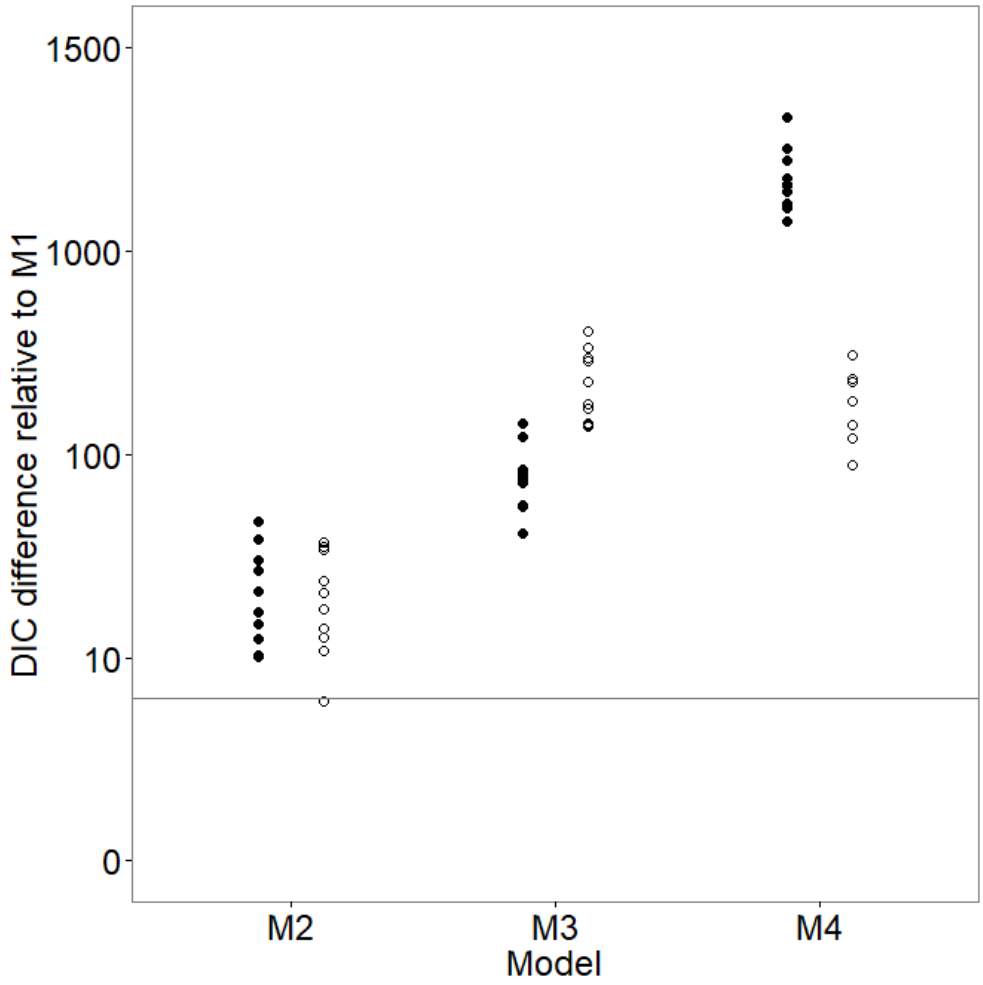


Figure A.1 Differences in Deviance Information Criteria (DIC), expressed relative to Model 1 (M1; true data generating model), for alternative SEM specifications for which data architecture was mispecified (M2, M3) or ignored (M4), under simulated conditions of unbalanced cluster size. White and black dots indicate simulation scenarios of moderate (~10%) large (~30-50%) magnitude of cluster-level variance relative to the total variance of each response, respectively. A reference horizontal line indicates the threshold value indicative of a decisive fit difference relative to M1.

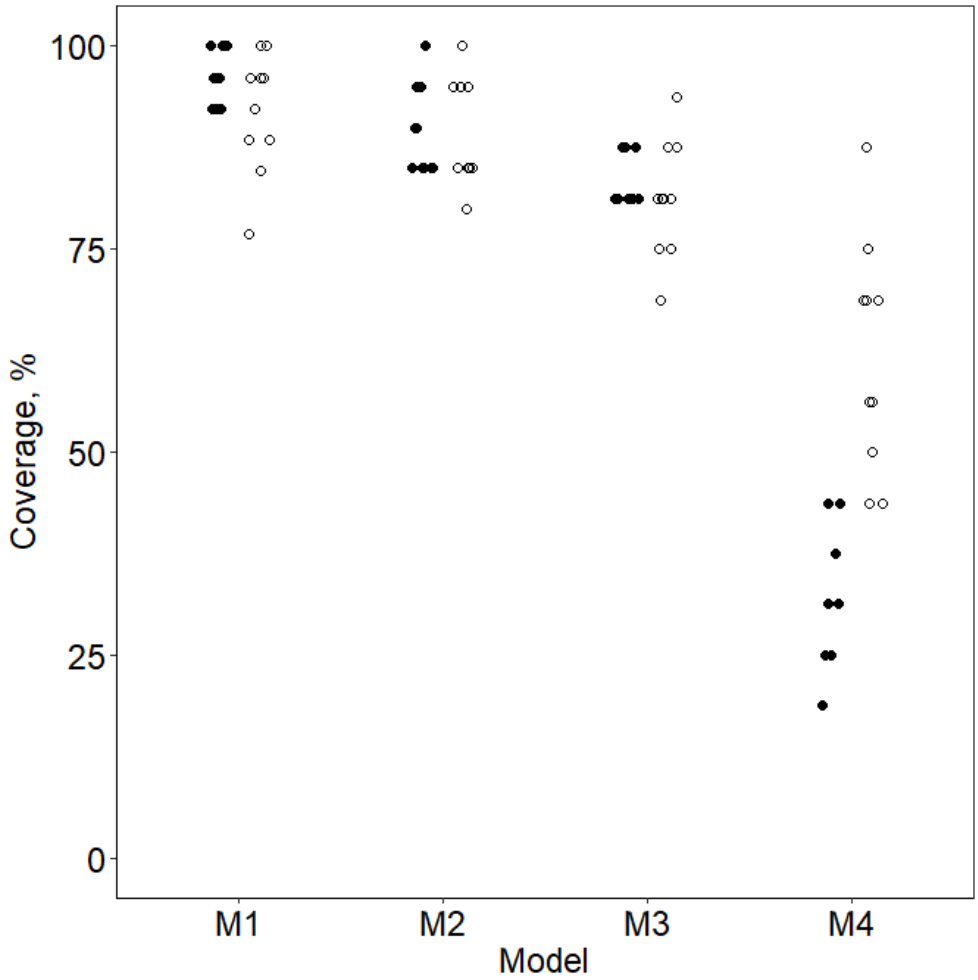


Figure A.2 Overall accuracy of parameter estimation, expressed as proportion of coverage (in %) based on 95% HPD intervals, for alternative SEM specifications for which data architecture was correctly specified (M1), mispecified (M2, M3) or ignored (M4) under simulated conditions of balanced cluster sizes. White and black dots indicate simulation scenarios of moderate (~10%) or large (~30-50%) magnitude of cluster-level variance relative to the total variance of each response, respectively.

Table A.1 Coverage (expressed in %) of selected parameter types for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of unbalanced cluster size.

		Moderate cluster-level variance				Large cluster-level variance			
Parameters	No. of simulated data	M1	M2	M3	M4	M1	M2	M3	M4
$\lambda_{jj'}$	40	92.5	82.5↓	92.5	57.5↓	97.5	70.0↓	100.0	22.5↓
$\beta_{1,j}$	40	87.5	87.5	42.5↓	72.5↓	97.5	95.0	42.5↓	52.5↓
$\beta_{2,j}$	40	87.5	90.0	95.0	90.0	95.0	95.0	97.5	57.5↑
$\sigma^2_{e_j}$	40	97.5	95.0	95.0	27.5↑	95.0	92.5	95.0	0↑
$\sigma^2_{u_j}$	40	90.0	90.0	-	-	87.5	95.0	-	-
$\sigma_{u_{jj'}}$	60	95.0	-	-	-	100.0	-	-	-

For j = 1, 2, 3, and 4 responses and $j' < j$

Upward and downward arrows indicate whether a given parameter type was consistently overestimated or underestimated, respectively.

M1: True data-generating model.

M2: Model ignores cluster-level correlations between responses.

M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).

M4: Model ignores data architecture all together.

Table A.2 Precision of parameter estimation, expressed as width of the 95% highest posterior density intervals for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of large magnitude of cluster-level variability and balanced cluster size. Means and standard deviation of precision of parameter estimation are shown.

	Alternative SEM specifications for data architecture							
	M1		M2		M3		M4	
Parameters	Mean	SD	Mean	SD	Mean	SD	Mean	SD
λ_{21}	0.0262	0.0007	0.0259	0.0008	0.0272	0.0009	0.0304	0.0013
λ_{31}	0.1042	0.0043	0.0996	0.0034	0.1067	0.0050	0.0837	0.0041
λ_{32}	0.1959	0.0068	0.1885	0.0063	0.2042	0.0090	0.1722	0.0069
λ_{42}	0.0819	0.0042	0.0812	0.0038	0.0856	0.0048	0.0987	0.0052
$\beta_{1,1}$	9.49	0.62	9.14	0.59	3.74	0.08	5.54	0.17
$\beta_{1,2}$	4.01	0.24	3.85	0.24	2.52	0.10	3.29	0.14
$\beta_{1,3}$	8.19	0.15	7.89	0.25	6.70	0.27	7.95	0.25
$\beta_{1,4}$	7.81	0.36	7.73	0.32	6.21	0.36	8.05	0.48
$\beta_{2,1}$	8.18	0.24	8.16	0.23	9.36	0.24	9.31	0.28
$\beta_{2,2}$	1.81	0.05	1.80	0.06	2.11	0.05	2.45	0.12
$\beta_{2,3}$	3.35	0.08	3.34	0.09	3.83	0.109	4.02	0.14
$\beta_{2,4}$	3.40	0.15	3.38	0.14	3.93	0.18	4.34	0.23
$\sigma^2_{e_1}$	169.3	9.0	168.8	9.2	168.0	9.5	226.1	13.2
$\sigma^2_{e_2}$	6.8	0.4	6.7	0.4	6.7	0.4	13.6	1.2
$\sigma^2_{e_3}$	15.0	0.7	15.1	0.7	15.1	0.6	26.1	1.9
$\sigma^2_{e_4}$	18.8	1.4	18.7	1.4	18.8	1.4	33.8	3.0
$\sigma^2_{u_1}$	305.9	44.8	271.1	40.6	-	-	-	-
$\sigma^2_{u_2}$	34.9	6.1	30.0	5.2	-	-	-	-
$\sigma^2_{u_3}$	60.3	7.6	49.4	6.7	-	-	-	-
$\sigma^2_{u_4}$	67.7	11.0	60.9	10.0	-	-	-	-
$\sigma_{u_{21}}$	86.5	15.0	-	-	-	-	-	-
$\sigma_{u_{31}}$	105.3	14.7	-	-	-	-	-	-
$\sigma_{u_{32}}$	37.4	5.3	-	-	-	-	-	-
$\sigma_{u_{41}}$	105.2	7.1	-	-	-	-	-	-
$\sigma_{u_{42}}$	35.5	4.3	-	-	-	-	-	-
$\sigma_{u_{43}}$	45.6	4.1	-	-	-	-	-	-

"-" indicates that estimates are not applicable due to SEM specifications

M1: True data-generating model.

M2: Model ignores cluster-level correlations between responses.

M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).

M4: Model ignores data architecture all together.

Table A.3 Precision of parameter estimation, expressed as width of the 95% highest posterior density intervals for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of moderate magnitude of cluster-level variability and unbalanced cluster size**.** Means and standard deviation of precision of parameter estimation are shown.

	Alternative SEM specifications for data architecture							
	M1		M2		M3		M4	
Parameters	Mean	SD	Mean	SD	Mean	SD	Mean	SD
λ_{21}	0.0252	0.0008	0.0244	0.0010	0.0262	0.0009	0.0240	0.0007
λ_{31}	0.0996	0.0056	0.0951	0.0047	0.1014	0.0057	0.0932	0.0050
λ_{32}	0.1875	0.0090	0.1806	0.0087	0.1949	0.0098	0.1773	0.0070
λ_{42}	0.0775	0.0026	0.0754	0.0025	0.0831	0.0031	0.0763	0.0025
$\beta_{1,1}$	6.61	0.32	6.36	0.32	3.61	0.17	4.78	0.13
$\beta_{1,2}$	2.89	0.12	2.77	0.13	2.47	0.09	2.62	0.11
$\beta_{1,3}$	7.00	0.20	6.80	0.27	6.39	0.26	6.77	0.18
$\beta_{1,4}$	6.62	0.21	6.41	0.21	6.07	0.21	6.39	0.19
$\beta_{2,1}$	7.78	0.32	7.78	0.35	9.15	0.44	8.11	0.28
$\beta_{2,2}$	1.73	0.07	1.74	0.07	2.08	0.09	1.78	0.07
$\beta_{2,3}$	3.17	0.12	3.14	0.13	3.66	0.16	3.19	0.10
$\beta_{2,4}$	3.26	0.11	3.22	0.10	3.85	0.13	3.33	0.10
$\sigma^2_{e_1}$	161.5	9.8	161.9	10.6	164.0	11.8	173.3	7.5
$\sigma^2_{e_2}$	6.5	0.4	6.5	0.4	6.5	0.4	7.0	0.4
$\sigma^2_{e_3}$	14.2	0.8	14.3	0.7	14.2	0.9	15.1	0.8
$\sigma^2_{e_4}$	18.0	1.1	17.9	1.0	18.1	1.1	19.1	0.9
$\sigma^2_{u_1}$	131.6	20.3	115.9	18.1	-	-	-	-
$\sigma^2_{u_2}$	5.7	0.8	4.8	0.6	-	-	-	-
$\sigma^2_{u_3}$	11.1	1.3	9.3	1.2	-	-	-	-
$\sigma^2_{u_4}$	15.0	3.7	13.2	3.4	-	-	-	-
$\sigma_{u_{21}}$	21.5	2.7	-	-	-	-	-	-
$\sigma_{u_{31}}$	29.7	3.8	-	-	-	-	-	-
$\sigma_{u_{32}}$	6.2	0.7	-	-	-	-	-	-
$\sigma_{u_{41}}$	33.0	3.8	-	-	-	-	-	-
$\sigma_{u_{42}}$	6.7	0.7	-	-	-	-	-	-
$\sigma_{u_{43}}$	9.3	1.0	-	-	-	-	-	-

"-" indicates that estimates are not applicable due to SEM specifications.
M1: True data-generating model.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together.

Table A.4 Precision of parameter estimation expressed as width of the 95% highest posterior density intervals for alternative SEM specifications of data architecture M1 – M4 under simulated conditions of large magnitude of cluster-level variability and unbalanced cluster size. Means and standard deviation of precision of parameter estimation are shown.

	Alternative SEM specifications for data architecture							
	M1		M2		M3		M4	
Parameters	Mean	SD	Mean	SD	Mean	SD	Mean	SD
λ_{21}	0.0255	0.0013	0.0256	0.0016	0.0267	0.0013	0.0296	0.0018
λ_{31}	0.1026	0.0048	0.0972	0.0039	0.1045	0.0052	0.0830	0.0051
λ_{32}	0.1909	0.0098	0.1833	0.0089	0.1984	0.0108	0.1721	0.0109
λ_{42}	0.0800	0.0031	0.0780	0.0032	0.0840	0.0037	0.0968	0.0036
$\beta_{1,1}$	9.89	0.89	9.55	0.93	3.65	0.15	5.51	0.35
$\beta_{1,2}$	4.05	0.17	3.91	0.19	2.48	0.14	3.22	0.19
$\beta_{1,3}$	7.98	0.51	7.74	0.44	6.52	0.44	7.90	0.54
$\beta_{1,4}$	7.72	0.20	7.47	0.20	6.08	0.28	8.04	0.30
$\beta_{2,1}$	7.86	0.33	7.89	0.34	9.12	0.38	9.20	0.63
$\beta_{2,2}$	1.77	0.08	1.78	0.07	2.09	0.09	2.39	0.13
$\beta_{2,3}$	3.23	0.14	3.21	0.12	3.72	0.15	3.99	0.26
$\beta_{2,4}$	3.32	0.11	3.30	0.12	3.85	0.14	4.31	0.17
$\sigma^2_{e_1}$	161.6	9.3	162.0	8.6	163.2	8.4	228.0	21.9
$\sigma^2_{e_2}$	6.5	0.42	6.5	0.4	6.5	0.4	13.3	1.2
$\sigma^2_{e_3}$	14.3	0.9	14.4	1.0	14.3	0.9	26.0	3.0
$\sigma^2_{e_4}$	18.2	0.8	18.2	0.9	18.2	0.8	33.8	2.9
$\sigma^2_{u_1}$	348.1	70.5	309.1	62.7	-	-	-	-
$\sigma^2_{u_2}$	36.8	3.3	31.9	3.0	-	-	-	-
$\sigma^2_{u_3}$	65.7	11.8	55.6	10.5	-	-	-	-
$\sigma^2_{u_4}$	70.7	10.0	62.5	8.7	-	-	-	-
$\sigma_{u_{21}}$	94.9	11.8	-	-	-	-	-	-
$\sigma_{u_{31}}$	119.4	20.3	-	-	-	-	-	-
$\sigma_{u_{32}}$	39.4	4.1	-	-	-	-	-	-
$\sigma_{u_{41}}$	114.2	14.4	-	-	-	-	-	-
$\sigma_{u_{42}}$	37.0	2.8	-	-	-	-	-	-
$\sigma_{u_{43}}$	48.8	5.9	-	-	-	-	-	-

"-" indicates that estimates are not applicable due to SEM specifications.
M1: True data-generating model.
M2: Model ignores cluster-level correlations between responses.
M3: Model further ignores within-cluster correlation between observations for a given response (i.e. clusters as fixed effects). Also ignores cluster-level correlations between responses (similar to M2).
M4: Model ignores data architecture all together.

Appendix B - Supplementary Materials for Chapter 3

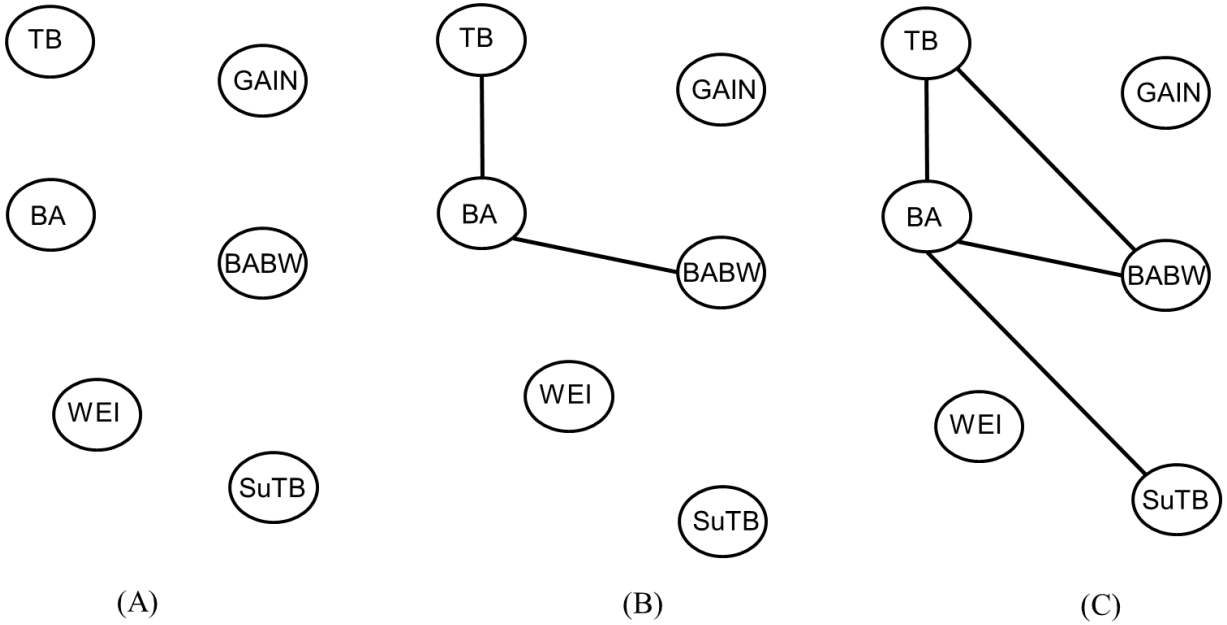


Figure B.1 Undirected graph of reproductive performance traits in sows detected by the inductive causation algorithm implemented with (A) 85, 90, and 95%, (B) 80% and (C) 75 and 70% highest posterior density intervals. Links represent associated between traits. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

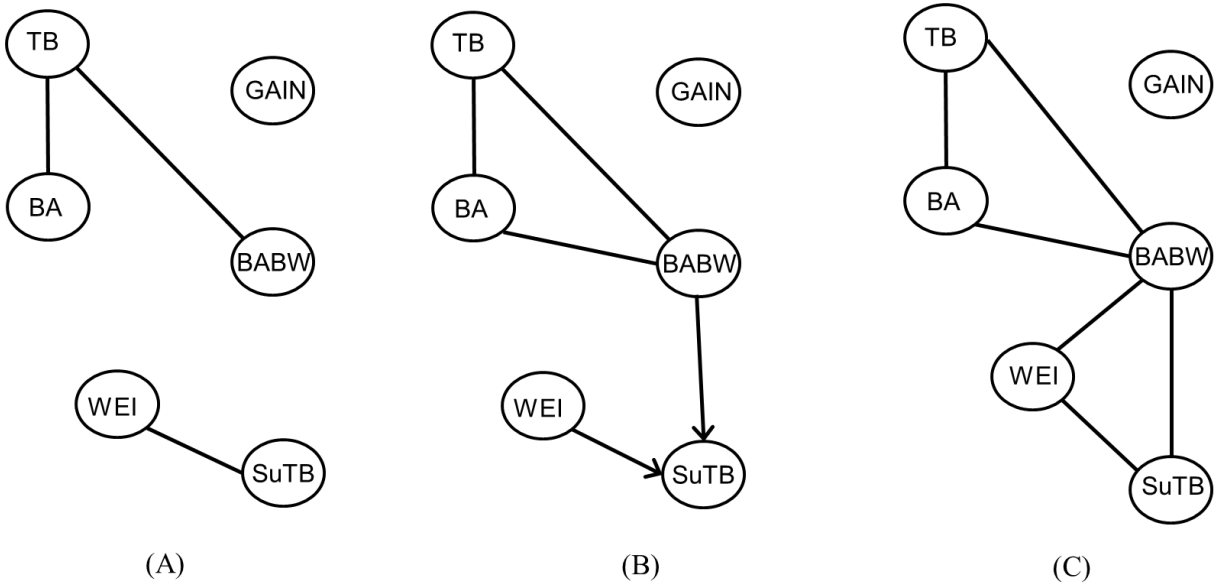


Figure B.2 Partially oriented graph of reproductive performance traits in gilts detected by the inductive causation algorithm implemented with (A)95%, (B) 90% and (C) 85% highest posterior density intervals. Links without arrowheads represent connections between traits; links with arrowheads represent causal effects from the trait on the arrowtail to the trait on the arrowhead GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.

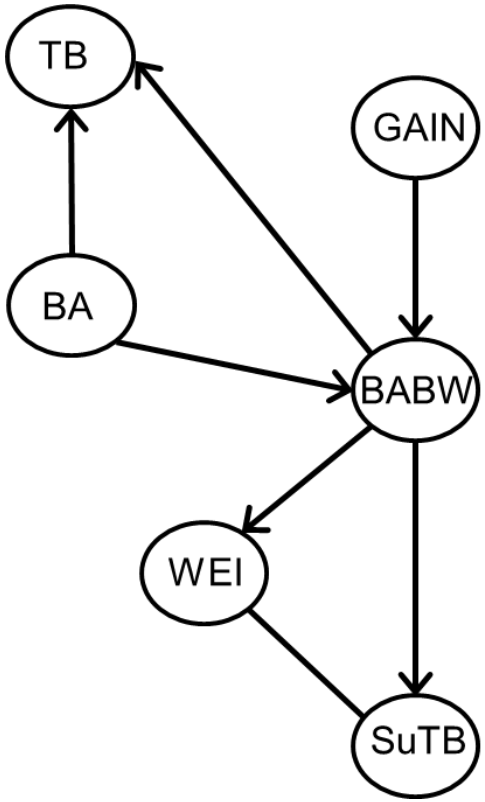


Figure B.3 Partially oriented graph of reproductive performance traits in gilts detected by the inductive causation algorithm implemented with 80, 75 and 70% highest posterior density intervals. Links without arrowheads represent connections between traits; links with arrowheads represent causal effects from the trait on the arrowtail to the trait on the arrowhead. GAIN = female weight gain during late gestation; TB = total number born in a litter; BA = number born alive in a litter; BABW = born alive average body weight; WEI = wean-to-estrous interval; SuTB = total number born in the subsequent gestation.