

Published in final edited form as:

*Biometrics*. 2009 June ; 65(2): 584–589. doi:10.1111/j.1541-0420.2008.01098.x.

## Analysis of Twin Data Using SAS

Rui Feng<sup>1</sup>, Gongfu Zhou<sup>2</sup>, Meizhuo Zhang<sup>2</sup>, and Heping Zhang<sup>2,\*</sup>

<sup>1</sup> Department of Biostatistics, the University of Alabama at Birmingham, Birmingham, AL 35294, U.S.A

<sup>2</sup> Department of Epidemiology and Public Health, Yale University, New Haven, CT 06520, U.S.A

### SUMMARY

Twin studies are essential for assessing disease inheritance. Data generated from twin studies are traditionally analyzed using specialized computational programs. For many researchers, especially those who are new to twin studies, understanding and using those specialized computational programs can be a daunting task. Given that SAS is the most popular software for statistical analysis, we suggest the use of SAS procedures for twin data may be a helpful alternative and demonstrate that we can obtain similar results from SAS to those produced by specialized computational programs. This numerical validation is practically useful, because a natural concern with general statistical software is whether it can deal with data that are generated from special study designs such as twin studies and whether it can test a particular hypothesis. We conclude through our extensive simulation that SAS procedures can be used easily as a very convenient alternative to specialized programs for twin data analysis.

### Keywords

Twin study; Variance components method; Heritability; Generalized linear mixed model; SAS PROC MIXED; SAS PROC NL MIXED

## 1. Introduction

Twin studies are often used to evaluate the inheritance of a trait by dissecting the genetic and environmental contributions to the trait (e.g., Allison et al. 1995; Martin, Boomsma, and Machin, 1997; Mulder et al., 2003; Evans et al., 2003; Peltonen and GenomEUtwin, 2003; Franks et al., 2005; Bhandari et al., 2006; Bizzarro et al., 2006). Monozygotic (MZ) twins, also called identical twins, share the same genetic materials, whereas dizygotic (DZ) twins or fraternal twins share an average 50% of their genes like the usual siblings. Thus, for an inheritable trait, we expect a higher concordance or correlation in MZ twins than in DZ twins, and we can use the excess in concordance among MZ twins to determine the level of genetic contribution, provided that twins of the same parents are raised in similar environments such as the same families.

Statistical models have been well established for analyzing twin data (Kempthorne and Osborne, 1961; Haseman and Elston, 1970). The most commonly used models are variance components methods, also called structure equation models in social sciences, which decompose the observed phenotypic variance into additive genetic, dominant genetic, common environmental components, and random noises including measurement errors (Neale et al., 1989). Heritability measures the proportion of the total phenotypic variation that is attributable to genes. Several software programs are available for twin analyses, such as LISREL (Jöreskog

---

\*email: E-mail: Heping.Zhang@yale.edu.

and Sörbom, 1986), Mx (Neale et al., 1999), and Mplus (Muthén and Muthén, 1998; Prescott, 2004). All of these programs require thorough understanding and sophisticated use of matrix operations. Among them, Mx is a popular choice, and is free. In recent years, SAS, Statistical Analysis Software, has enhanced its procedures for fitting variance component models (McArdle, 2006). Franks et al. (2005) used SAS PROC MIXED to assess the heritability of physical activity in children. In addition, SASPairs (<http://psych.colorado.edu/~carey/SASPairs/>) provides a comprehensive series of macros and IML (Interactive Matrix Language) modules written in SAS. With basic functions similar to Mx, it does everything that Mx can do for sibpairs within SAS. In other words, SASPairs can be viewed as a re-implementation of Mx in SAS.

All software packages are based on similar statistical models, but the computational algorithms may differ and can lead to different results, even for simple models. Thus, it is useful to evaluate both similarities and differences resulting from uses of different software. Such comparisons have been reported in analyses of correlated data. Ferrer, Hamagami, and McArdle (2004) showed nearly identical results from different software programs including LISREL, Mx, Mplus, AMOS (Arbuckle, 1999), and SAS when fitting latent growth curve models to longitudinal data. A comparison of Mplus and LISREL by Asparouhov and Muthén (2006) demonstrated that the Mplus algorithm produced more accurate parameter estimates in mixed effect modeling than LISREL. Only until recently did researchers begin to compare the performance of the various programs in analyzing twin data. McArdle (2006) examined Mx, Mplus and SAS by fitting mixed effect models to a simulated longitudinal twin dataset and found identical maximum likelihood parameter estimates and close standard errors from those programs. However, more simulations and thorough evaluations are warranted to understand the effect on the heritability estimate as a result of different software packages.

In this paper, we compare the performances of SAS and Mx for twin data based on both simulations and a published data set in a neonatal twin study. Because SAS is the most widely used statistical software package, it is practically very useful to be able to perform analyses of twin data using standard SAS procedures and syntax, specifically, SAS PROC MIXED and PROC NL MIXED. The advantage is that the users will not need to install any additional packages, nor are we required to learn new syntax.

## 2. Methods

### 2.1. Genetic Model for Quantitative Traits

Suppose there are  $n_1$  MZ twin pairs and  $n_2$  DZ twin pairs. A quantitative response  $Y$  and  $p$  covariates, denoted by  $X$ , are available for each twin. The covariates could be any observed demographic and environmental variables. A genetic model that includes covariate effects, an additive genetic effect, a dominant genetic effect, a common environmental effect shared by a twin pair (no matter which zygosity it has), and a residual environmental effect can be written as:

$$y_{ij} = \mathbf{X}_{ij}\beta + a_{ij} + d_{ij} + c_i + \varepsilon_{ij}, \quad (1)$$

where  $i$  is the index for each of the  $n_1 + n_2$  twin pairs and  $j$  (1 or 2) is the index for one of the twins in a pair. Also,  $\beta$  measures the covariate effects (including an intercept),  $a_{ij}$ ,  $d_{ij}$ ,  $c_i$ , and  $\varepsilon_{ij}$  are, respectively, the additive genetic, dominance genetic, common environmental and residual environmental random effects on the  $i$ -th twin pair. Because MZ twins are genetically identical, for them  $a_{i1} = a_{i2}$  and  $d_{i1} = d_{i2}$ . We assume that  $a$ ,  $d$ ,  $c$ , and  $\Sigma$  (the subscripts are sometimes suppressed for convenience) are independently normally distributed with mean 0 and variance  $\sigma_a^2$ ,  $\sigma_d^2$ ,  $\sigma_c^2$ , and  $\sigma_e^2$ , respectively. If the  $i$ -th twin pair is MZ, the covariances of  $a$

and  $d$  between the two twins are  $\text{cov}(a_{i1}, a_{i2}) = \sigma_a^2$  and  $\text{cov}(d_{i1}, d_{i2}) = \sigma_d^2$ , respectively; If the  $i$ -th twin pair is DZ, the covariances of  $a$  and  $d$  between the two twins are  $\text{cov}(a_{i1}, a_{i2}) = \frac{\sigma_a^2}{2}$  and  $\text{cov}(d_{i1}, d_{i2}) = \frac{\sigma_d^2}{4}$ , respectively (Falconer and Mackay, 1996).

The genetic heritability then can be estimated using the ratio of estimated genetic variance and the total variance of the quantitative trait, i.e.,  $\frac{\text{Var}(A) + \text{Var}(D)}{\text{Var}(A) + \text{Var}(D) + \text{Var}(C) + \text{Var}(E)}$ .

Due to the identifiability problem, we may not be able to dissect additive genetic effects, dominance genetic effects and shared environmental effects. In that case, we may have to consider simpler models such as ACE model (model (1) without  $d_{ij}$ ) or ADE model (model (1) without  $c_i$ ).

## 2.2. Genetic Model for Binary Traits

Suppose instead the response  $Y$  is binary, for example, disease status (affected or normal). Then we can use probit model to fit twin data:

$$\text{probit}(\Pr(y_{ij}=1)) = \mathbf{X}_{ij}\beta + a_{ij} + d_{ij} + e_i,$$

where  $\text{probit}(\cdot)$  is the inverse function of the cumulative standard normal distribution. Other notations are similar to those above. The difference is that we envision a liability underlying the binary disease and that  $a_{ij}$ ,  $d_{ij}$ , and  $c_i$  are the additive genetic, dominance genetic, common environmental effects on the liability for the  $i$ -th twin pair. This probit model is commonly referred to as the threshold model (Gianola and Foulley, 1983; Falconer and Mackay, 1996).

## 2.3. Software – SAS and Mx

Mx is a flexible statistical modeling package, initially designed for structural equation modeling. It is freely available. SAS is a commercial software package, but widely available. SAS procedures MIXED and NLMIXED are well developed to fit linear and nonlinear mixed models that facilitate fixed covariate effects, random genetic and environmental effects. PROC MIXED and NLMIXED fit mixed models by maximizing an approximation to the likelihood integrated over the random effects (Latour, Latour, and Wolfinger, 1994; Wolfinger, 1999). PROC NLMIXED can specify any conditional distribution for the response variable given the random effects, including normal, binomial, and Poisson. To fit the ACE or ADE model for twin data, we specify the variance-covariance structures of  $a_{ij}$ ,  $d_{ij}$ ,  $c_i$ , and  $\varepsilon_{ij}$  defined in model (1) through the “type” option in the random effect statement. For people familiar with SAS, the syntax of using PROC MIXED and NLMIXED is very simple.

## 3. Simulations

### 3.1. Quantitative trait

We simulated 200 MZ and 200 DZ twin pairs in each data set. A covariate,  $x$ , was generated from  $\text{uniform}(0,1)$ . For the  $i$ -th twin pair, a common familial risk score, denoted by  $c_i$ , was generated from  $\text{Normal}(0, \sigma_c^2)$ . We considered both the ACE and ADE models. For each model, we simulated 100 datasets.

For the ACE models, the additive genetic effect,  $a_{ij}$  ( $j=1, 2$ ), was generated with mean 0 and covariance matrix  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \sigma_a^2$  for MZ twins or  $\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \sigma_a^2$  for DZ twins. The response  $Y_{ij}$  equals  $\beta_0 + \beta_1 x_i + a_{ij} + c_i + e_{ij}$ , for  $j = 1, 2$ , where  $e_{ij}$  is a normal noise with mean 0 and variance  $\sigma_e^2$ . For the ADE models, we generated the response  $Y_{ij}$  from  $\beta_0 + \beta_1 x_i + a_{ij} + d_{ij} + e_{ij}$ , where  $d_{ij}$  was the dominant genetic effect and generated for MD twins with covariance matrix  $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \sigma_d^2$  and for DZ twins with  $\begin{pmatrix} 1 & 1/4 \\ 1/4 & 1 \end{pmatrix} \sigma_d^2$ .

Without loss of generality, we fixed variance  $\sigma_e^2$  at 1.0. For convenience, we set  $\sigma_c^2 = 0.5\sigma_a^2$  in ACE model simulation and  $\sigma_d^2 = \sigma_a^2$  in ADE model simulation (other specification give similar results for the comparison). We varied each variance components to set the heritability ranging from 0.2 to 0.6 with the step size of 0.1. We compared the estimated heritability  $\hat{h}^2$  and fixed effects  $\hat{\beta}_1$  from SAS and Mx.

Figure 1 shows the box-plots of parameter estimates when the fitted model (ACE or ADE) is the same as the data generating model (ACE or ADE). Figure 1a is for heritability estimates and Figure 1b is for the estimates of the fixed effects.

To assess the robustness of the estimates when the fitted model is different from the data generating model, we examined the parameter estimates from the ACE and ADE models when the data were simulated from an ACDE model. Without loss of generality, we again set variance  $\sigma_e^2$  at 1 and set  $\sigma_a^2, \sigma_d^2$ , and  $\sigma_c^2$  to be equal (other specification give similar results for the comparison). We varied each variance components to allow the heritability ranging from 0.2 to 0.6 at an interval of 0.1. The sample sizes were the same as before. Again SAS and Mx yielded similar results. The average and sample standard deviation of the estimated heritability are presented in Figure 2. Both SAS and Mx overestimated the heritability, although the ACE models tend to yield less biased heritability estimates than the ADE models.

In summary, whether the fitted models are same as the true models or not, SAS and Mx produce similar results. Both SAS and Mx have good estimates of the fixed effect ( $\beta_1 = 1$ ) in all data sets, although the estimates of the fixed effect  $\beta_1$  become less accurate with increased inheritability. For the heritability estimates, the standard deviations from the ADE model are smaller than those of the ACE model because the estimates of  $\sigma_a^2$  and  $\sigma_d^2$  are negatively correlated (Williams, 1993).

### 3.2 Binary Trait

We also examined the performance of SAS and Mx for qualitative or binary traits. Following the same procedure as in Section 3.1, we first simulated a quantitative trait  $Z$  as a liability variable. We then defined a binary trait  $Y$  taking value of 1 or 0 according to whether  $Z > 2$  (2 was arbitrarily chosen) or not. We varied the variance components, which in turn controls the heritability of the liability variable  $Z$ . The box plots of parameter estimates are shown in Figure 3. When the fitted model is same as the data generating model, the parameter estimates from SAS and Mx are similar and the estimated fixed effects from both ACE and ADE models are close to the true values.

As in Section 3.1, we also examined the parameter estimates when the true liability was simulated from an ACDE model, whereas we fitted either ACE or ADE model to the data. The results from SAS and Mx were similar, as shown in Figure 4. Again, from both SAS and Mx, the ACE models tend to produce less biased heritability estimates than the ADE models.

#### 4. Application to a Study on Premature Twins

Preterm birth is defined as birth less than 37 weeks of gestation. About half millions of infants in the U.S. are born prematurely, and preterm infants have increased morbidity and mortality rate compared with full-term infants (Martin, Boomsma, and Machin, 1997).

Bronchopulmonary dysplasia (BPD), a chronic lung injury, is a prevalent cause of morbidity in preterm infants. The etiology of BPD is multifactorial and genes play very important parts (Bokodi et al., 2007). We intended to compare the heritability estimates of BPD from SAS and Mx using a dataset from a premature twin study (Bizzarro et al., 2006).

All twins were born between January 1, 1994 and December 31, 2004 at three medical centers, Karolinska Institute, the University of Connecticut, and Yale University and the data we used include those who survived beyond 36 weeks postmenstrual age (PMA). The zygosity of each twin pair was determined by histopathologic examination of the placenta with additional confirmation using gender concordance or discordance. 63 monozygotic (MZ) and 137 dizygotic (DZ) twin pairs were identified. Birth weight (BW), gestation age (GA), sex (1: male, 0: female), presence of respiratory distress syndrome (RDS), duration of ventilation and supplemental oxygen use, and length of stay were also available for each infant. See Bhandari et al. (2006) and Bizzarro et al. (2006) for more details. We fitted the ACE, ADE probit models using SAS and Mx for BPD. Bhandari et al. (2006) also documented a study of BPD using a different dataset in which they used the logistic model and included only those variables with significant fixed effects. Thus, the parameter estimates we reported are different from those reported in Bizzarro et al. (2006). But the inheritability estimates are quite close even they used a smaller dataset and fitted a different model. We included all important covariates to compare the fixed effect estimates between SAS and Mx. Table 1 displays the estimated covariates effects and heritability estimates for BPD. Mx and SAS give very similar results even though the dataset might not even come from an ACDE model. The estimated heritability is 56.97% (SAS) and 56.77% (Mx) using the ACE model and 78.18% (SAS) or 78.94% (Mx) using the ADE model.

#### 5. Conclusion and Discussion

This work has three useful implications in practice. First, although PROC MIXED has been used for analysis of twin data in which the trait has a normal distribution, we now demonstrated how to use PROC NLMIXED when the trait is binary. In other words, we presented how to estimate disease heritability from twin studies whether the trait is quantitative or binary. Second, the usage of PROC MIXED and PROC NLMIXED for twin data analysis is straightforward. Lastly, we examined whether use of SAS and Mx resulted in different estimates and inference when we assess the inheritance of quantitative or binary trait. We found the results similar. This is assuring, because a major disadvantage of a general software package such as SAS is that it is not necessarily designed to address a specific question and take into account a particular design. Our evaluation relieves us from this concern.

As statistical software adopt advanced statistical methods and computational algorithms, it will become more convenient and simple to conduct genetic analyses under the standard syntax of those software. Similar to our objective, Rabe-Hesketh et al. (2008) explained in detail how to use the *glamm* function in STATA and *lme* function in R or S-Plus to analyze data from twin studies.

## 5. Supplementary Materials

Web-based supplementary materials, including the SAS code and data sets, will be distributed through the Biometrics website <http://www.tibs.org/biometrics> as well as the authors' website: <http://c2s2.yale.edu/software/twin>.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

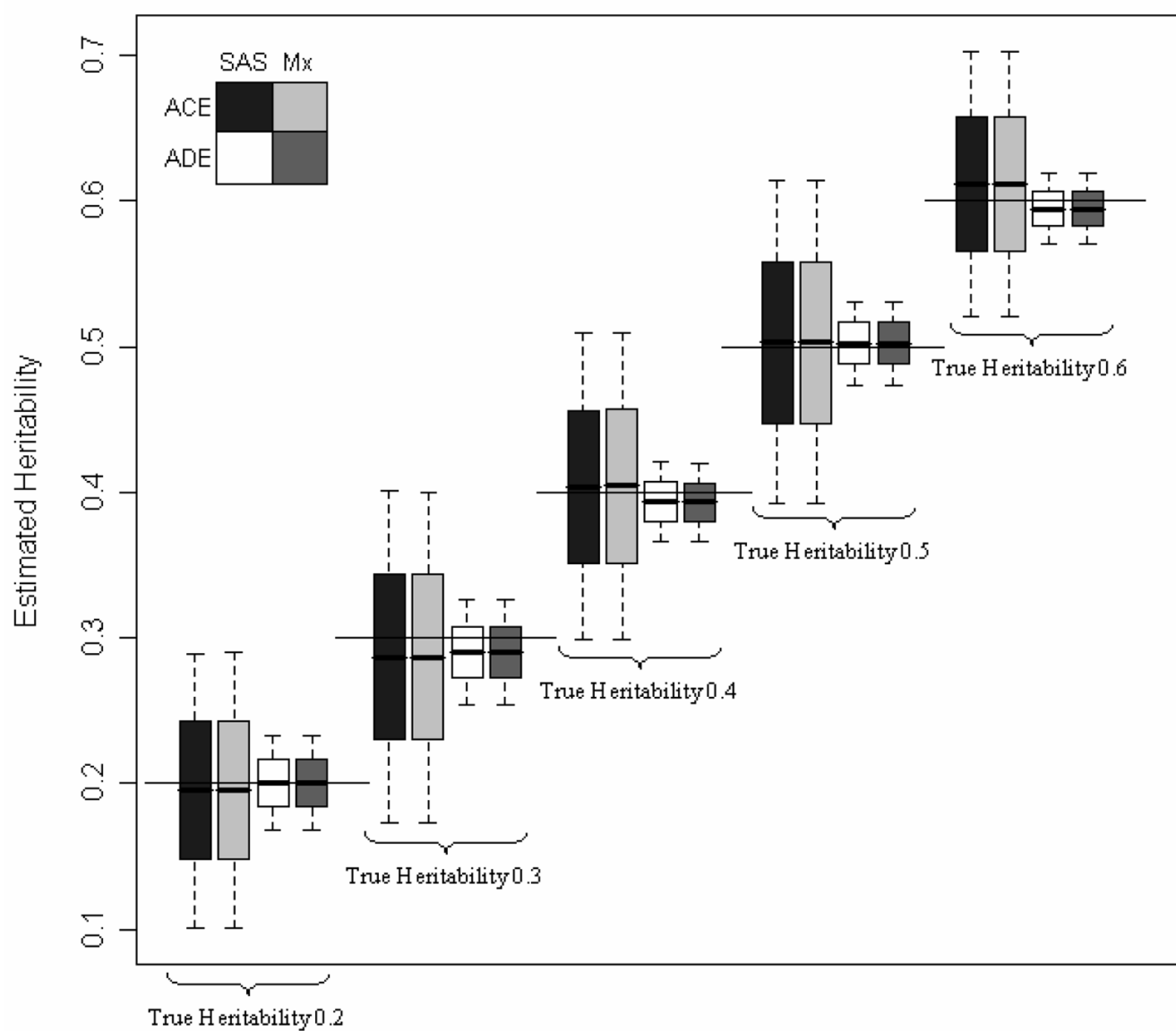
The authors thank Dr. David Allison for his suggestions and Dr. Michael Neal for his help with Mx programming. This research is supported in part by grants K02DA017713 and R01DA016750 from the National Institutes on Drug Abuse.

## References

- Allison DB, Heshka S, Neale MC, Tishler PV, Heymsfield SB. Genetic, environmental, and phenotypic links between body mass index and blood pressure among women. *American journal of medical genetics* 1995;55:335–341. [PubMed: 7726233]
- Arbuckle, JL. AMOS 4.0. Chicago, IL: Smallwaters; 1999.
- Asparouhov T, Muthen B. Comparison of estimation methods for complex survey data analysis. Mplus Web Notes. 2006
- Bhandari V, Bizzarro MJ, Shetty A, Zhong X, Page GP, Zhang H, Ment LR, Gruen JR. Familial and genetic susceptibility to major neonatal morbidities in preterm twins. *Pediatrics* 2006;117:1901–1906. [PubMed: 16740829]
- Bizzarro MJ, Hussain N, Jonsson B, Feng R, Ment LR, Gruen JR, Zhang H, Bhandari V. Genetic susceptibility to retinopathy of prematurity. *Pediatrics* 2006;118:1858–1863. [PubMed: 17079555]
- Bokodi G, Treszl A, Kovacs L, Tulassay T, Vasarhelyi B. Dysplasia: a review. *Pediatric pulmonology* 2007;42:952–961. [PubMed: 17726703]
- Evans A, Van Baal GC, McCarron P, DeLange M, Soerensen TI, De Geus EJ, Kyvik K, Pedersen NL, Spector TD, Andrew T, Patterson C, Whitfield JB, Zhu G, Martin NG, Kaprio J, Boomsma DI. The genetics of coronary heart disease: the contribution of twin studies. *Twin Res* 2003;6:432–441. [PubMed: 14624727]
- Falconer, DS.; Mackay, TFC. Introduction to quantitative genetics. Harlow, England; New York, Prentice Hall: 1996.
- Ferrer E, Hamagami F, McArdle JJ. Modeling latent growth curves with incomplete data using different types of structural equation modeling and multilevel software. *Structural Equation Modeling-a Multidisciplinary Journal* 2004;11:452–483.
- Franks PW, Ravussin E, Hanson RL, Harper IT, Allison DB, Knowler WC, Tataranni PA, Salbe AD. Habitual physical activity in children: the role of genes and the environment. *The American journal of clinical nutrition* 2005;82:901–908. [PubMed: 16210723]
- Gianola D, Foulley JL. Sire evaluation for ordered categorical data with a threshold model. *Genetics, Selection, Evolution* 1983;15:201–224.
- Hao K, Wang X, Niu T, Xu X, Li A, Chang W, Wang L, Li G, Laird N, Xu X. A candidate gene association study on preterm delivery: application of high-throughput genotyping technology and advanced statistical methods. *Human molecular genetics* 2004;13:683–691. [PubMed: 14976157]
- Haseman JK, Elston RC. The estimation of genetic variance from twin data. *Behavior genetics* 1970;1:11–19. [PubMed: 5535855]
- Jöreskog, KG.; Sörbom, D. Lisrel VI. Mooresville, Indiana: Scientific Software; 1986.
- Kemphorne O, Osborne RH. The interpretation of twin data. *Am J Hum Genet* 1961;13:320–339. [PubMed: 13752449]

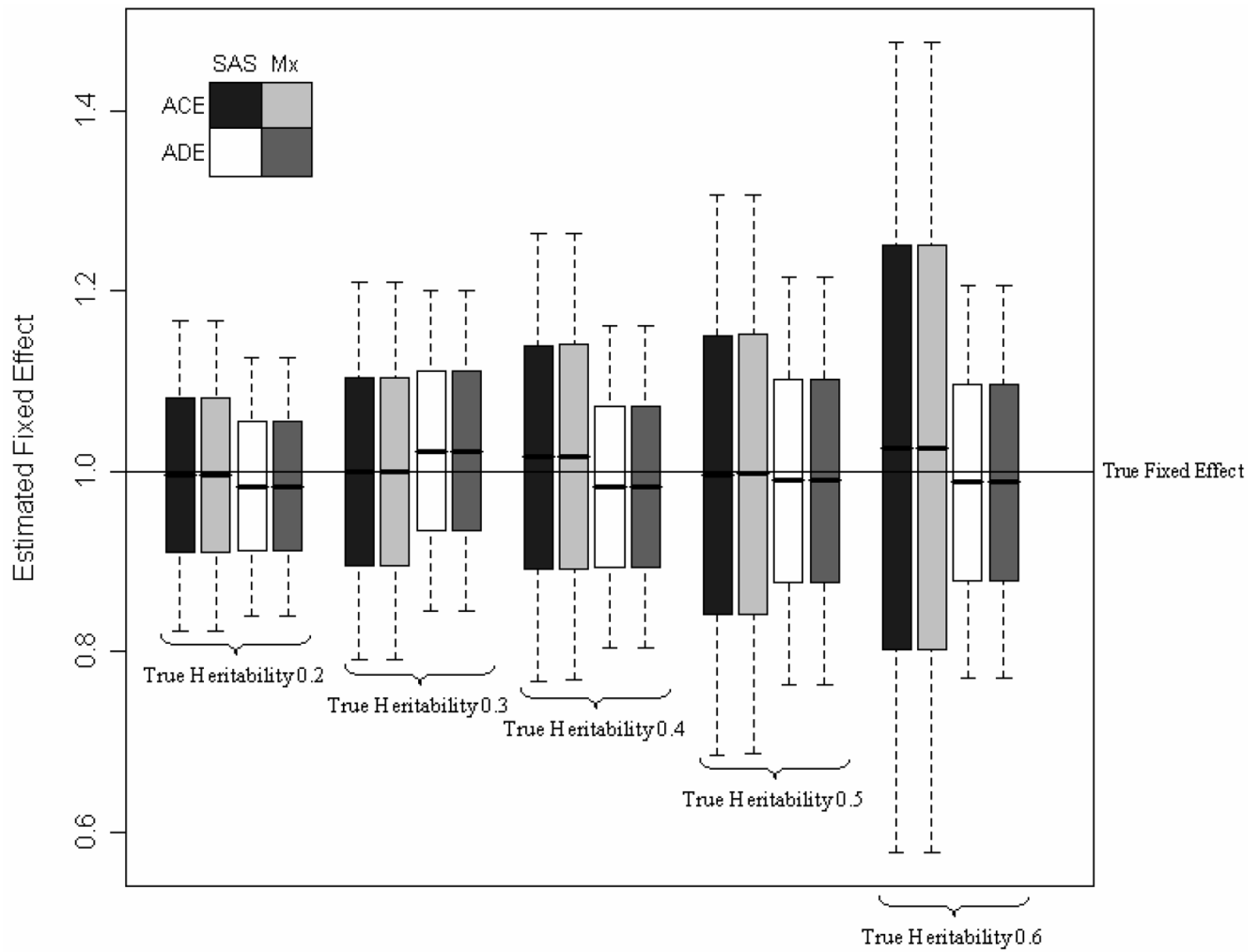


- Latour, D.; Latour, K.; Wolfinger, RD. Getting Started with PROC MIXED. Software Sales and Marketing Department, SAS Institute Inc.; Cary, NC: 1994.
- Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Munson ML. Births: final data for 2003. *Natl Vital Stat Rep* 2005;54:1–116.
- Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet* 1997;17:387–392. [PubMed: 9398838]
- McArdle JJ. Latent curve analyses of longitudinal twin data using a mixed-effects biometric approach. *Twin Res Hum Genet* 2006;9:343–359. [PubMed: 16790145]
- Mulder EJ, Van Baal C, Gaist D, Kallela M, Kaprio J, Svensson DA, Nyholt DR, Martin NG, MacGregor AJ, Cherkas LF, Boomsma DI, Palotie A. Genetic and environmental influences on migraine: a twin study across six countries. *Twin Res* 2003;6:422–431. [PubMed: 14624726]
- Muthén, LK.; Muthén, BO. Mplus User's Guide. Los Angeles, CA: Muthén & Muthén; 1998.
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical Modeling. Richmond: Dept of Psychiatry, Medical College of Virginia of Virginia Commonwealth University; 1999.
- Neale, MC.; Cardon, LR. Methodology for Genetic Studies of Twins and Families. Dordrecht, the Netherlands; Kluwer Academic: 1992.
- Neale MC, Heath AC, Hewitt JK, Eaves LJ, Fulker DW. Fitting genetic models with LISREL: hypothesis testing. *Behavior genetics* 1989;19:37–49. [PubMed: 2712812]
- Peltonen L. GenomEUtwin: a strategy to identify genetic influences on health and disease. *Twin Res* 2003;6:354–360. [PubMed: 14624719]
- Prescott CA. Using the Mplus computer program to estimate models for continuous and categorical data from twins. *Behavior genetics* 2004;34:17–40. [PubMed: 14739694]
- Rabe-Hesketh S, Skrondal A, Gjessing HK. Biometrical modeling of twin and family data using standard mixed model software. *Biometrics* 2008;64:280–288. [PubMed: 17484777]
- Williams CJ, Christian JC, Norton JA Jr. TWINAN90: a FORTRAN program for conducting ANOVA-based and likelihood-based analyses of twin data. *Comput Methods Programs Biomed* 1992;38:167–176. [PubMed: 1458866]
- Wolfinger R. Fitting Nonlinear Mixed Models with the New NLMIXED Procedure. Proceedings of the 24th Annual SAS Users Group International Conference (SUGI 24) 1999:287–224.
- Zhang H, Bracken MB. Tree-based risk factor analysis of preterm delivery and small-for-gestational-age birth. *American journal of epidemiology* 1995;141:70–78. [PubMed: 7801968]



**Figure 1 (a)**



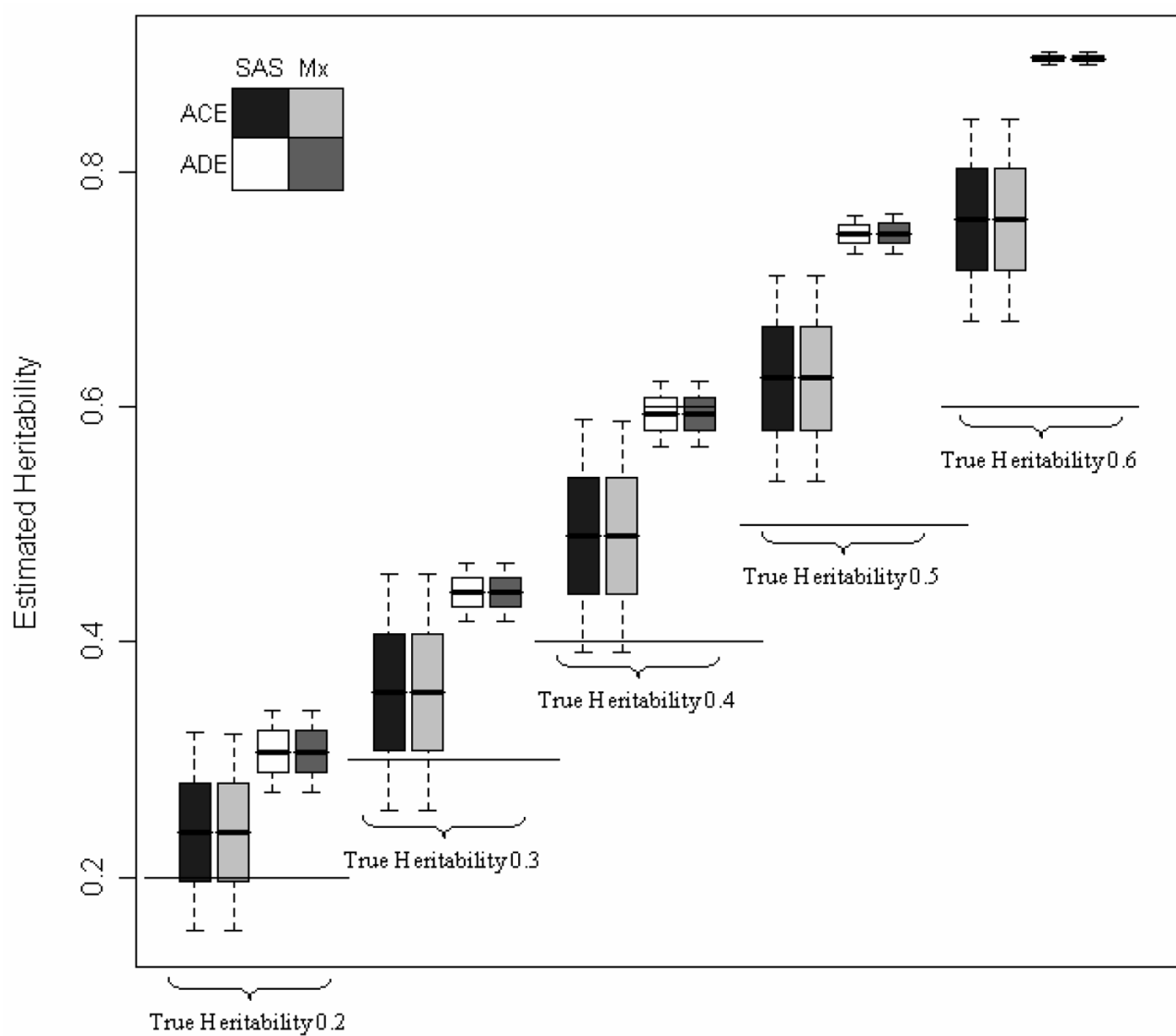


**Figure 1 (b)**

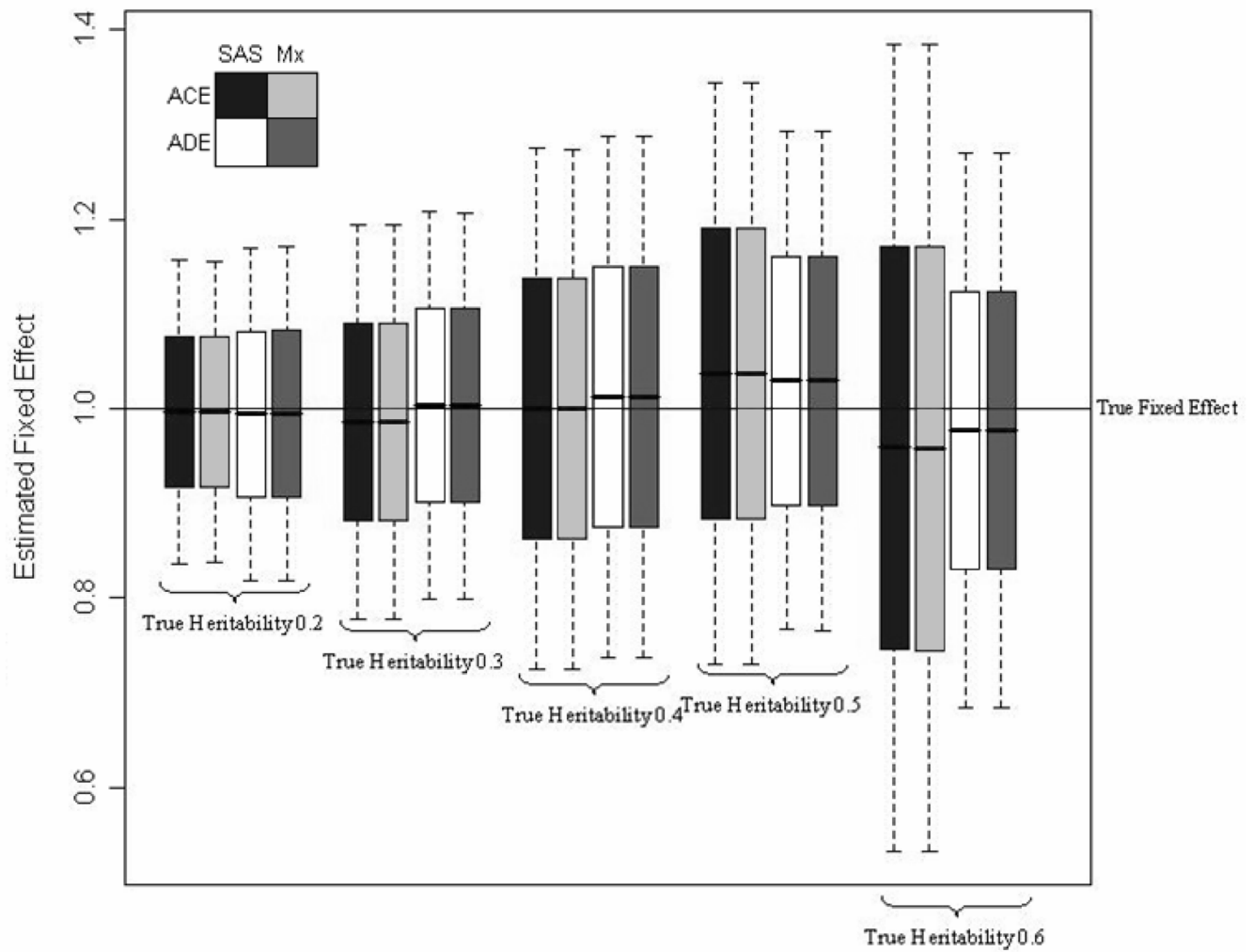
**Figure 1.**

Heritability and Fixed Effect Estimated from SAS PROC NLMIXED & Mx Based on 100 datasets, each dataset contains 200 MZ Twin Pairs and 200 DZ Twin Pairs for a Quantitative Trait under Correct Genetic Models

- Estimated heritability when the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively
- Estimated fixed effect when the true mixed effect is 1.0 and the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively.



**Figure 2(a)**

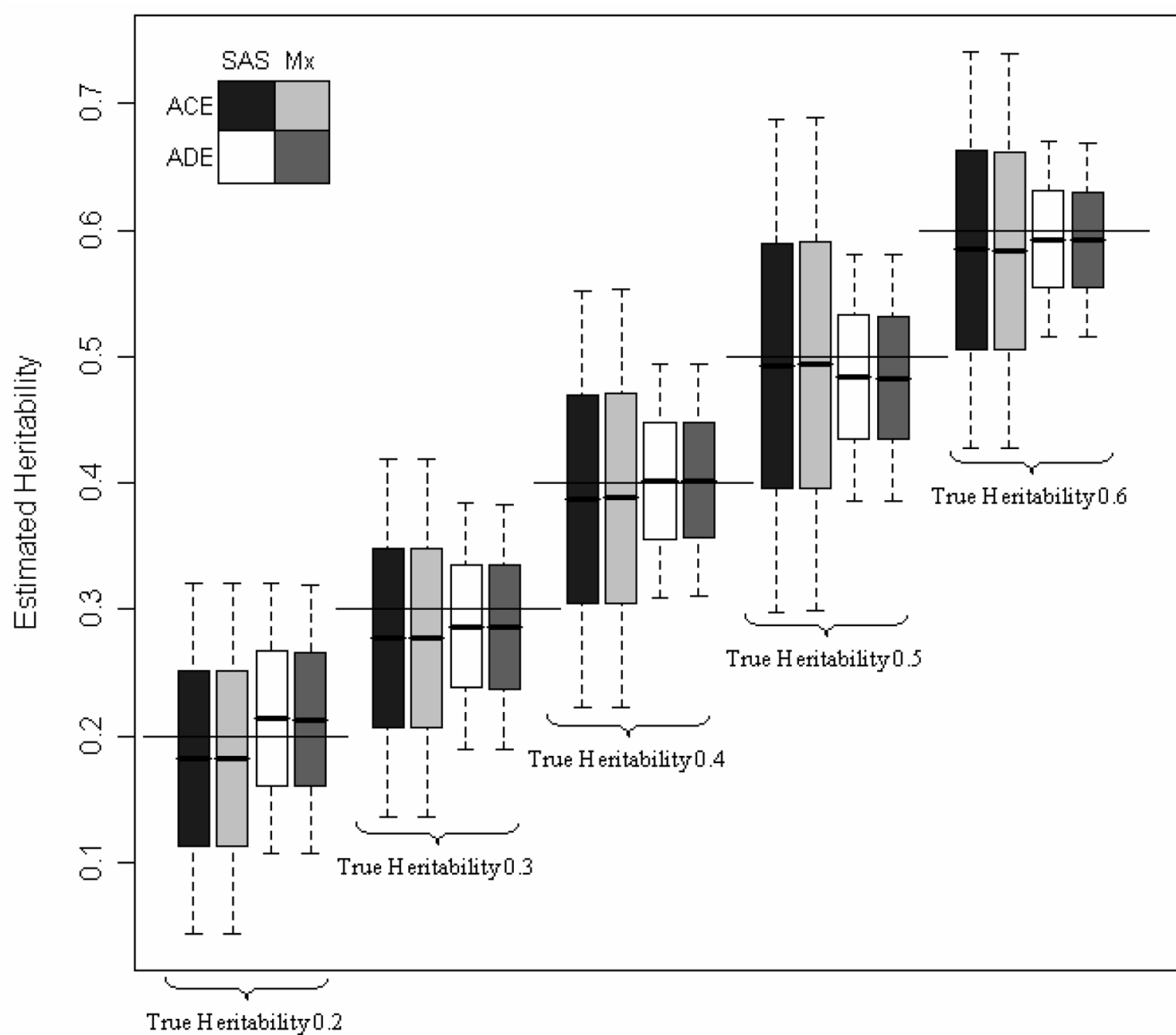


**Figure 2 (b)**

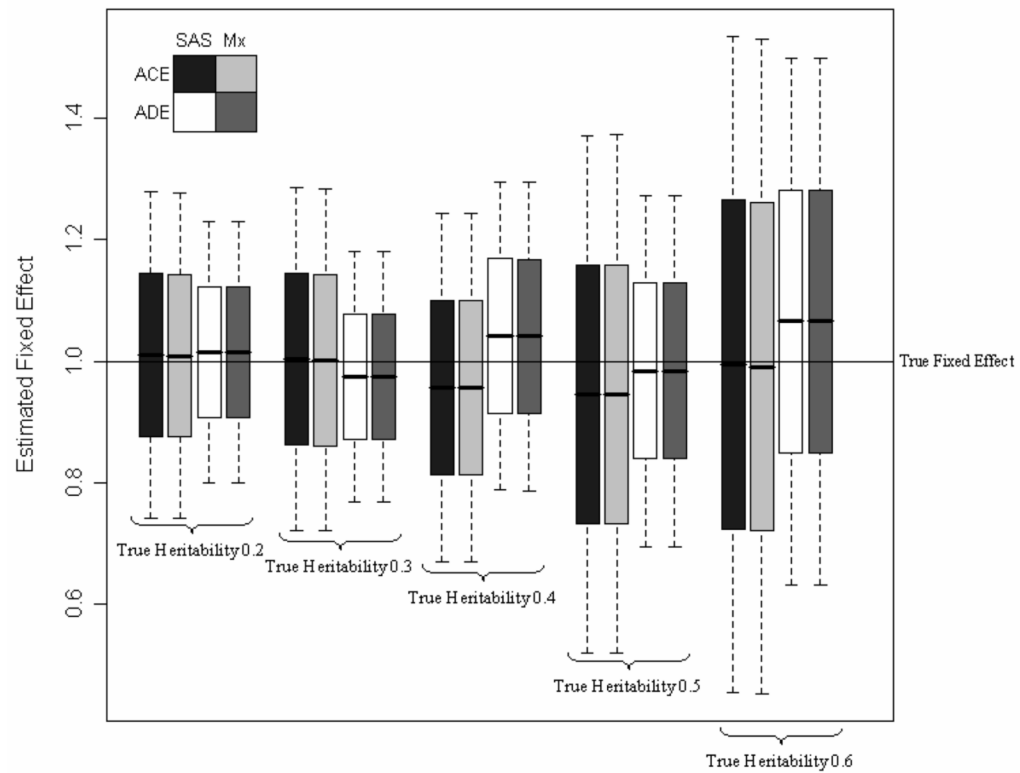
**Figure 2.**

Heritability and Fixed Effect Estimated from SAS PROC NLMIXED & Mx Based on 100 datasets, each dataset contains 200 MZ Twin Pairs and 200 DZ Twin Pairs for a Quantitative Trait under Mis-specified Models

- Estimated heritability when the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively
- Estimated fixed effect when the true mixed effect is 1.0 and the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively.



**Figure 3 (a)**



**Figure 3 (b)**

**Figure 3.**

Heritability and Fixed Effect Estimated from SAS PROC NLMIXED & Mx Based on 100 datasets, each dataset contains 200 MZ Twin Pairs and 200 DZ Twin Pairs for a Binary Trait under Correct Genetic Models

- Estimated heritability when the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively
- Estimated fixed effect when the true mixed effect is 1.0 and the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively.

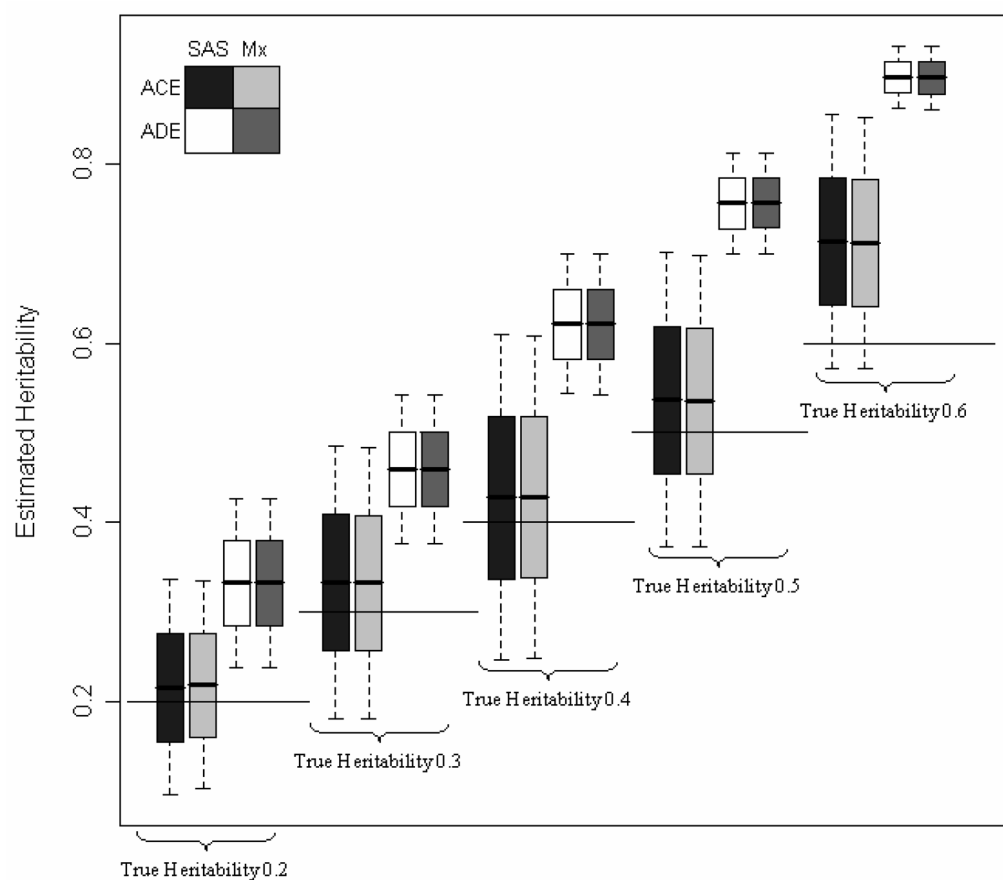
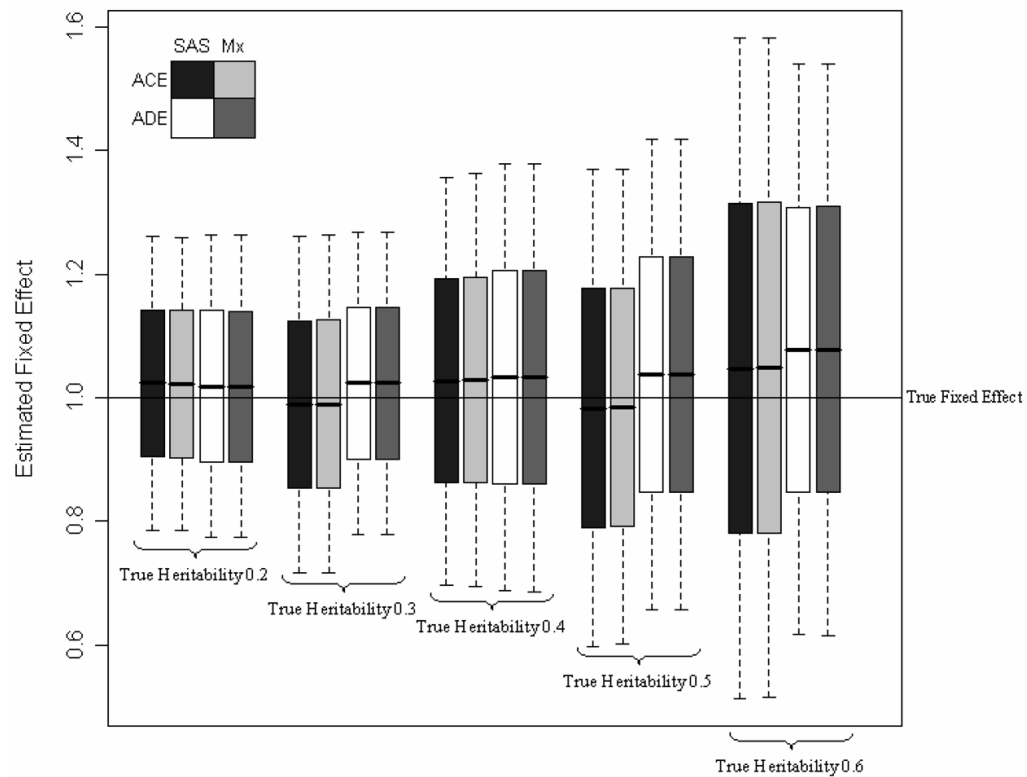


Figure 4 (a)



**Figure 4 (b)**

**Figure 4.**

Heritability and Fixed Effect Estimated from SAS PROC NLMIXED & Mx Based on 100 datasets, each dataset contains 200 MZ Twin Pairs and 200 DZ Twin Pairs for a Binary Trait under Mis-specified Models

- Estimated heritability when the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively
- Estimated fixed effect when the true mixed effect is 1.0 and the true heritability is 0.2, 0.3, 0.4, 0.5 and 0.6, respectively.



**Table 1**  
The estimated covariate effects and heritability for BPD.

Variables	ADE			ACE		
	Estimate	P-Value	Mx	Estimate	P-Value	Mx
Sex	-0.3477	0.3673	-0.3527	-0.3571	0.3386	-0.3572
Ga	0.1859	0.2418	0.1832	0.1602	0.3183	0.1601
BW	0.0024	0.0279	0.0025	0.0025	0.0250	0.0025
RDS	-2.4320	<0.01	-2.4585	-2.4064	<0.01	-2.4071
INST1	1.8818	<0.01	1.9149	1.8498	<0.01	1.8504
INST2	0.4633	0.5448	0.4753	0.4553	0.5551	0.4554
$h^2$	0.7818	<0.01	0.7894	0.5697	0.1911	0.5677