DF-Analyses of Heritability
with Double-Entry Twin Data:
Asymptotic Standard Errors and
Efficient Estimation

Hans-Peter Kohler (kohler@demogr.mpg.de)
Joseph Lee Rodgers (jrodgers@ou.edu)

DF-Analyses of Heritability with Double-Entry Twin Data: Asymptotic Standard Errors and Efficient Estimation

Hans-Peter Kohler[*]   Joseph Lee Rodgers[†]

May 10, 2000

Abstract

In this paper we establish the asymptotic distribution of DeFries–Fulker (1985) regression estimates for heritability and shared environmental influences with double-entry twin data. A simple formula to estimate the covariance matrix of the coefficients in DF-regressions is provided, and applications to simulated data and Danish twin data show that this method can substantially increase the statistical power of the analyses. We also provide an 'efficient DF-analysis' that yields more precise estimates when additional covariates are included among the explanatory variables.

Keywords: DF-analysis, twin studies

(this version: twdfefficient01a.tex)

1   Introduction

DeFries and Fulker (1985) propose a method of estimating heritability (h^2) and common environmental influences (c^2) with twin data by a simple linear regression of a twin's trait on her co-twin's trait and the degree of genetic relatedness. The initial development of the

---

[*]Head of Research Group on Social Dynamics and Fertility, Max Planck Institute for Demographic Research, Doberaner Str. 114, 18057 Rostock, Germany. *Tel:* +49-381-2081-123, *Fax:* +49-381-2081-423, *Email:* kohler@demogr.mpg.de, *www:* http://user.demogr.mpg.de/kohler.

[†]Professor of Psychology, Department of Psychology, University of Oklahoma, Norman, Ok 73019. *Email:* jrodgers@ou.edu

DeFries and Fulker (DF) method focused on selected twin data in which one member of each pair is a proband. The method has subsequently been generalized to include random samples of twins (Cherny, Cardon, Fulker, and DeFries 1992), and several authors have shown that the regression coefficients in the DF-analysis provide unbiased estimates for heritability and shared environmental influences (LaBuda, DeFries, and Fulker 1986; Rodgers and McGue 1994). In addition, several extensions have been proposed that allow the consideration of genetic non-additivity (Waller 1994), observed differences in non-shared environment (Rodgers, Rowe, and Li 1994), and binary or censored observations (Kohler and Rodgers 1999). DF-regressions in these different variants have been widely used with selected and unselected twin or kinship data because of their simple and flexible implementation (Cherny et al. 1992; Cyphers et al. 1990; Fulker et al. 1991; LaBuda and DeFries 1990; Rodgers and Rowe 1987; Rodgers et al. 1994; Rodgers et al. 1998; Zieleniewski et al. 1987).

The ambiguity in unselected samples as to which twin's trait should be used as the dependent, and which as the independent variable, is frequently resolved by using double-entry (Haggard 1958). Each twin pair is entered twice in the data, and each member of a twin pair provides once the dependent and once the explanatory variable.[1] While the consistency of the regression estimates for heritability and environmental influences is not affected by double-entry, the standard errors of the coefficients are obviously biased. It has been common practice to adjust these standard errors for the correct degrees of freedom. In large samples this adjustment is achieved by multiplying the standard errors obtained from the double-entry regression with an adjustment factor $\sqrt{2}$, or alternatively, by multiplying the covariance matrix obtained in the initial regression by the factor 2.

This approach is problematic for at least two reasons. First, the DF-regression is heteroscedastic—even with single entry data—because a co-twin's trait is a better predictor of the proband's trait in MZ twin pairs as compared to DZ twin pairs. This implies that the residual variance in the DF-regression is larger for DZ twins than for MZ twins. This heteroscedasticity potentially distorts the estimated covariance matrix of the parameters. The above adjustment for the degrees of freedom in double-entry data may therefore be based

---

[1]Cherny et al. (1992) point out two alternative possibilities. They note that each of the 2^N possible combinations of twin assignment as proband could be considered and the average estimate obtained. Alternatively the average of the estimate obtained from a smaller number of analyses based on random assingment could be used.

on a distorted and inappropriate starting point, which subsequently may lead to erroneous significance tests even after the adjustment. Second, the degrees-of-freedom adjustment achieved by multiplying covariance matrix with a factor 2 is conservative. It assumes that double-entering a twin pair does not increase the information—or loosely speaking, the overall degrees of freedom—in the sample. However, this is not the case and double-entering each twin pair not only resolves the ambiguity in the assignment of the proband, but also increases the information available for the inference of heritability and shared environmental effects.

In this paper we establish the asymptotically correct covariance matrix of the coefficients in DF-regressions with double-entry data and we provide a simple estimator for this covariance matrix.[2] The analysis is based on the interpretation of DF-regressions as a generalized method of moment (GMM) estimator. This interpretation also suggests that DF-analysis is inefficient when it is combined with additional covariates, and we propose an 'efficient DF-estimation' for these cases. This efficient method yields more accurate estimates than the linear regression used in standard DF-analysis. STATA programs to estimate both the correct covariance matrix and the efficient DF-estimator are available from the first author of this paper.

The remainder of the paper is structured as follows. Section 2 investigates DF-analysis with double-entry data within the perspective of standard linear regression. Section 3 shows how DF-regressions can be interpreted within the framework of generalized methods of moments. This Section also establishes the asymptotic distribution of the estimated coefficients and derives an efficient DF-estimation that is asymptotically preferable when DF-regressions contains additional covariates. Section 4 provides simulations that compare the different estimators in this paper with results obtained from bootstrapping the data. Section 5 applies the methods to body height and body mass index in a sample of Danish twins, and Section 6 concludes.

---

[2]With a few straight-forward modifications, this method resolves the heteroscedasticity problem of DF regressions also in selected twin samples with single-entry data.

2 DF-Analysis from a Regression Perspective

In order to motivate our GMM approach to the estimation of heritability and shared environmental effects, it is useful to first revisit DF-analyses with double-entry data in the framework of linear regression.[3]

The 'augmented DF-analysis' estimates the regression $w_{1j} = \beta_0 + \beta_1 w_{2j} + \beta_2 R(z_j) + \beta_3 R(z_j) w_{2j}$, where w_{ij} is the trait value of twin $i = 1, 2$ in pair j, z_j is the zygosity, and $R(z_j)$ is the degree of genetic relatedness of the twin pair. In the development of this method, DeFries and Fulker (1985) assume an additive genetic model with no assortative mating and with equal environmental influences across kinship categories. The phenotype w_{ij} of twin i within twin-pair j can therefore be linearly decomposed as $w_{ij} = c_j + a_{ij} + e_{ij}$, where c_j represents the shared environmental influences experienced by the twins in pair j, a_{ij} is the additive genetic value of twin i in pair j, and e_{ij} denotes influences on the trait w_{ij} that are individual-specific to twin i in pair j. This relation can also be written in terms of deviations from the mean as

$$w_{ij} = \bar{w} + \tilde{c}_j + \tilde{a}_{ij} + \tilde{e}_{ij}, \tag{1}$$

where $\bar{w} = \mathrm{E}\, w_{ij}$ is the overall expectation of the trait, and $\tilde{c}_j = c_j - \bar{c}$, $\tilde{a}_{ij} = a_{ij} - \bar{a}$, $e_{ij} = \tilde{e}_{ij} - \bar{e}$ are the deviations from the expected values $\bar{c} = \mathrm{E}\, c_j$, $\bar{a} = \mathrm{E}\, a_{ij}$, and $\bar{e} = \mathrm{E}\, e_{ij}$ of the underlying genetic and environmental determinants of the trait. The additive genetic model then implies that the observations $(w_{1j}, w_{2j})'$ are independently distributed according to a bivariate distribution with a mean vector $(\bar{w}, \bar{w})'$ and a covariance matrix $\Omega(z_j)$ that depends on the zygosity z_j of a twin pair. In particular, the standard genetic model described above implies that

$$\Omega(z_j) = \sigma_w^2 \begin{pmatrix} 1 & \rho(z_j) \\ \rho(z_j) & 1 \end{pmatrix}, \tag{2}$$

where σ_w^2 is the variance of the trait w_{ij}, and $\rho(z_j) = c^2 + h^2 R(z_j)$ is the within-twin pair correlation of the trait as a function of the zygosity z_j. Some weak regularity conditions on this distribution are necessary to ensure asymptotic normality of regression and related estimators, and we assume throughout this paper that all fourth moments of $(w_{1j}, w_{2j})'$ are finite.

[3]For related analyses see also LaBuda, DeFries, and Fulker (1986), Rodgers and McGue (1994), Waller (1994) and Rodgers, Rowe, and Li (1994).

The expected value of the first twin's trait w_{1j}, conditional on the second twin's trait w_{2j}, in the above model equals $\mathrm{E}[w_{1j}|w_{2j}] = \bar{w} + c^2(w_{2j} - \bar{w}) + h^2 R(z_j)(w_{2j} - \bar{w}) = (1 - c^2)\bar{w} + c^2 w_{2j} - h^2 \bar{w} R(z_j) + h^2 R(z_j) w_{2j}$, where $h^2 = \mathrm{var}(a_{ij})/\mathrm{var}(w_{ij})$ is the coefficient of heritability of the trait, and $c^2 = \mathrm{var}(c_{ij})/\mathrm{var}(w_{ij})$ is the coefficient of shared environmental influences associated with the trait. This conditional expectation provides a direct intuition for DF-analysis. Since the expectation $\mathrm{E}[w_{1j}|w_{2j}]$ is a linear function of w_{2j}, $R(z_j)$ and $R(z_j)w_{2j}$, it can be estimated using standard regression techniques. Moreover, because two of the coefficients in this linear relation correspond to c^2 and h^2, the regression coefficients provide direct estimates for the heritability and shared environmental influence.

In the following we introduce some additional notation in order to investigate the properties of this DF-regression further. First we define the coefficient vector $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)'$ which we intend to estimate using DF-regression analysis. The regression coefficients correspond to the above genetic model as $\beta_0 = (1 - c^2)\bar{w}$, $\beta_1 = c^2$, $\beta_2 = -h^2 \bar{w}$, and $\beta_3 = h^2$. Second, we adopt the standard notion 'y' and 'x' to denote left- and right-hand-side variables in regression analysis. In particular, we denote as $y_{1j} = w_{1j}$ and $x_{1j} = (1, w_{2j}, R(z_j), R(z_j)w_{2j})$ the left- and right-hand-side variables in the DF-regression for twin 1 in pair j. The corresponding vectors y_{2j} and x_{2j} are also defined for the second twin with the proband assignment reversed. We then stack the observations for both twins in a pair and write $y_j = \binom{y_{1j}}{y_{2j}}$ and $x_j = \binom{x_{1j}}{x_{2j}}$.

When only the subscript j is present, the vectors/matrices 'y' and 'x' refer to the data in a twin pair, while the presence of the additional subscript '1' and '2' specifies the data for either the first or second twin within a pair. We also define the stacked matrixes $Y = (y_1', \dots, y_n')'$ and $X = (x_1', \dots, x_n')'$ which contain the complete double-entry data for all twin pairs $j = 1, \dots, n$.

The expectation of $y_j = (w_{1j}, w_{2j})'$, conditional on the respective co-twin's realization of the traits $(w_{2j}, w_{1j})'$, can now be written as $\binom{\mathrm{E}[y_{1j}|x_{1j}]}{\mathrm{E}[y_{2j}|x_{2j}]} = x_j\beta$, which directly leads to the DF-regression specification $y_j = x_j\beta + \varepsilon_j$. The term $\varepsilon_j = (\varepsilon_{1j}, \varepsilon_{2j})'$ is the residual in the DF regression of pair j. It is important to observe that this DF-regression satisfies for each twin i within pair j the condition $\mathrm{E}[\varepsilon_{ij}|x_{ij}] = \mathrm{E}[y_{ij} - x_{ij}\beta|x_{ij}] = 0$. This condition is essential for OLS in order to yield unbiased and consistent estimates for the parameters of interest β since DF-analysis is sometimes erroneously criticized for an apparent correlation

between the variables in x_{ij} and the corresponding residual term ε_{ij}.

The absence of any correlation between x_{ij} and ε_{ij} is seen by investigating the relation of the residual term ε_{ij} to the various components of the underlying genetic model. In order to establish this relation, we first define a matrix $Q(z_j) = \begin{pmatrix} 1 & -\rho(z_j) \\ -\rho(z_j) & 1 \end{pmatrix}$. We then stack relation (1) for both twins in a pair, and we premultiply this stacked relation with the matrix $Q(z_j)$ to obtain

$$\underbrace{Q(z_j)\begin{pmatrix} w_{1j} - \bar{w} \\ w_{2j} - \bar{w} \end{pmatrix}}_{=\varepsilon_j} = Q(z_j)\begin{pmatrix} \tilde{c}_j + \tilde{a}_{1j} + \tilde{e}_{1j} \\ \tilde{c}_j + \tilde{a}_{2j} + \tilde{e}_{2j} \end{pmatrix}. \tag{3}$$

After carrying out the matrix multiplication on the left side we find that the terms on the left side of (3) equal $y_j - x_j\beta = \varepsilon_j$. The right hand side of (3) is hence a representation of the residuals ε_j in terms of the genetic and environmental determinants of the trait w_{ij}.

Since we have defined $y_j = (w_{1j}, w_{2j})'$, the expectation $\mathrm{E}[y_j\varepsilon_j']$ reveals the correlation between each trait value and the residual in the DF-regression

$$\mathrm{E}[y_j\varepsilon_j'] = \mathrm{E}\left[\begin{pmatrix} \bar{w} + \tilde{c}_j + \tilde{a}_{1j} + \tilde{e}_{1j} \\ \bar{w} + \tilde{c}_j + \tilde{a}_{2j} + \tilde{e}_{2j} \end{pmatrix}\begin{pmatrix} \tilde{c}_j + \tilde{a}_{1j} + \tilde{e}_{1j} \\ \tilde{c}_j + \tilde{a}_{2j} + \tilde{e}_{2j} \end{pmatrix}' Q(z_j)'\right] = \sigma_w^2(1 - \rho(z_j))^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4}$$

The first equality in (4) follows from (1) and (3). The second equality follows from the observation that the matrix $Q(z_j)$, multiplied by $[\sigma_w^2(1 - \rho(z_j))^2]^{-1}$, is inverse of the covariance matrix $\Omega(z_j)$ in (2). Since the expectation $\mathrm{E}[y_j\varepsilon_j']$ is a diagonal matrix, it follows that the residual term ε_{1j} is uncorrelated with w_{2j}, and similarly, that ε_{2j} is uncorrelated with w_{1j}. Since the matrix x_{1j} (or x_{2j}) consists only of constant terms and w_{2j} (or w_{1j}), the residual term ε_{ij} is uncorrelated with the respective right-hand-side variables contained in x_{ij}.

At the same time, the residuals ε_{1j} and ε_{2j} are neither independent of each other nor are they perfectly dependent. In particular, the covariance matrix of the error term ε_j follows from (3) and (2) as

$$\mathrm{E}\,\varepsilon_j\varepsilon_j' = Q(z_j)\Omega(z_j)Q'(z_j) = \sigma_w^2(1 - \rho(z_j))^2\begin{pmatrix} 1 & -\rho(z_j) \\ -\rho(z_j) & 1 \end{pmatrix}. \tag{5}$$

This covariance matrix is useful because it reveals two important aspects in which double-entry DF-analysis deviates from the standard OLS model. First, the residuals across individuals in the sample are not independent but correlated within twin pairs. This correlation will typically be negative and equal to $-\rho(z_j)$, i.e., it equals the negative correlation of the

trait within twin pairs. Second, since $\rho(z_j)$ is typically larger for MZ twins than for DZ twins, the residuals are heteroscedastic across monozygotic and dizygotic twins whenever the trait w_{ij} has a genetic influence, i.e., when $h^2 > 0$.

In addition to the correlation and heteroscedasticity of the residuals, the x_{ij}'s within a twin pair are not independent and the usual interpretation of the estimated coefficient as 'conditional on the data X' is not appropriate because the x_j's consist of realizations that are used both in the dependent variable as well as in the explanatory variables. Neither of these complications affect the consistency and unbiasedness of the estimated coefficients $\hat{\beta}_{\mathrm{DF}}$ in a regression-based DF-analysis. However, they do affect the standard errors associated with this estimate. In addition, a mere adjustment for the degrees of freedom is not sufficient because the correlation and heteroscedasticity in the residuals may distort the OLS results in a more complicated fashion that cannot be adjusted for by a proportional degrees-of-freedom adjustment of the standard errors obtained in DF-analysis.

The complications in the application of DF-analysis with double-entry data, however, do not imply that we cannot determine the correct distribution of the estimated coefficient β and of the estimated heritability h^2 and shared environmental influence c^2. In order to establish this distribution, however, it is necessary to adopt a more general framework than that of standard linear regression.

3 A Generalized Method of Moment Approach to DF-Estimation

3.1 Specification of a GMM Estimator

The important characteristic of double-entry twin data is that one can no longer treat each individual as an independent observation since the realizations of the trait w_{ij} within a twin pair are drawn from a bivariate distribution with mean $(\bar{w}, \bar{w})'$ and the covariance matrix in (2). However, it is still appropriate to consider the observations $(w_{1j}, w_{2j})'$ across twin pairs as independent realizations.

Generalized method of moments (GMM) estimation is based on the idea that the parameters of an empirical model can be identified and estimated by comparing the theoretical moments of a statistical model with their empirical counterparts in the sample (Hansen

1982; Newey and McFadden 1994). One advantage of this approach, as compared to maximum likelihood estimation, is that the inferences are based only on the moments of a statistical model, without the requirement to define the respective distributions exactly. For instance, standard OLS regression emerges within the GMM framework from the moment condition that the residuals of the regression are orthogonal (i.e., uncorrelated) with the right-hand-side variables in the regression. While in this regression example the number of moment conditions equals the number of parameters to be estimated (i.e., each right-hand-side variable in the a regression model yields one moment condition to identify the respective parameter), GMM estimation easily accommodates overidentified models where the number of moment conditions exceeds the number of parameters. This feature of GMM estimation makes it particularly useful for the analysis of twin data, which are characterized by the fact that in each twin pair we obtain two observations of a trait.

3.2 Asymptotic Variance in Double-Entry DF-Analysis

In the following we continue to use the notation introduced in Section 2 above. The only modification to our notation is that from now on we allow for additional covariates to be included among the right-hand-side variables x_{ij}. Useful covariates are, for instance, observed differences in the non-shared environment of twins that can contribute to the explanations of within-twin pair differences of a trait (Rodgers, Rowe, and Li 1994). Moreover, the co-twin's trait value can be eliminated from the vectors x_{ij} in order to estimate DF-models that contain only additive-genetic and non-shared environmental effects (AE models) instead of the standard ACE models that include additive genetic, shared and non-shared environmental effects.

We denote the number of right-hand-side variables contained in x_{ij} as K. The vector x_{ij} thus has $K+1$ columns, x_j is a $2 \times (K+1)$ matrix containing the RHS variables of each twin pair, and β is a vector with $K+1$ elements (each vector/matrix includes a constant term in addition to the K right-hand-side variables).

One possibility to define relevant moment conditions for the method-of-moment estimation of β is to require that within each twin pair $\mathrm{E}[x'_{1j}\varepsilon_{1j}] = 0$ and $\mathrm{E}[x'_{2j}\varepsilon_{2j}] = 0$. However, these conditions yield a total of $2 \times (K+1)$ moment restrictions that need to be satisfied by estimating $K+1$ parameters. Due to randomness in the data, it is usually impossible

to find an estimate $\hat{\beta}$ to solve this overidentified problem.

The symmetry of double-entry data suggests that we choose the moment restriction

$$E[x_j'\varepsilon_j] = 0, \tag{6}$$

which implies that the sum $x_{1j}'\varepsilon_{1j} + x_{2j}'\varepsilon_{2j}$ has zero expectation. The empirical counterpart to this moment condition is

$$\frac{1}{n}\sum_{j=1}^{n} x_j'\varepsilon_j = \sum_{j=1}^{n} x_j'(y_j - x_j\beta) = \frac{1}{n}X'(Y - X\beta) = 0, \tag{7}$$

which yields $K+1$ equations for the unknown $K+1$ parameters contained in β. The solution to (7) is $\hat{\beta}_{\mathrm{DF}} = (X'X)^{-1}X'Y$, which is identical to the OLS coefficient obtained from regressing Y on X. The consistency of this estimator follows directly from the properties of OLS regression, and the asymptotic variance of this estimator is immediately obtained from the large-sample properties of GMM estimators (e.g., Hansen 1982; Newey and McFadden 1994).

Result 1 *(asymptotic variance of DF-estimator) The estimator $\hat{\beta}_{DF} = (X'X)^{-1}X'Y$ with double-entry twin data is asymptotically distributed as*

$$\hat{\beta}_{DF} \xrightarrow{a} N(\beta, \frac{1}{n}\{E[x_j'x_j]\}^{-1}\Sigma\{E[x_j'x_j]\}^{-1}), \tag{8}$$

where $\Sigma = E[x_j'\varepsilon_j\varepsilon_j'x_j]$ is the covariance matrix of the moment restrictions. A consistent estimator for the asymptotic variance of $\hat{\beta}_{DF}$ in (8) is given by $(X'X)^{-1}\hat{S}(X'X)^{-1}$ with $\hat{S} = [\sum_{j=1}^{n} x_j'\hat{\varepsilon}_j\hat{\varepsilon}_j'x_j] = [\sum_{j=1}^{n}(x_{1j}'\hat{\varepsilon}_{1j} + x_{2j}'\hat{\varepsilon}_{2j})(\hat{\varepsilon}_{1j}'x_{1j} + \hat{\varepsilon}_{2j}'x_{2j})]$, where $\hat{\varepsilon}_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - x_{ij}\hat{\beta}$ is the residual of the DF-regression.[4]

The above estimator for the variance has the advantage that it is easily computed, and we will show in Section 4 that it performs well when compared to the standard errors derived from bootstrapping the double-entry twin data. Moreover, since the matrices x_j can contain covariates in addition to the co-twin's trait and the degree of genetic relatedness, Result 1 also applies to the various extensions of the DF-model that are based on adding more

---

[4]The small sample properties of the variance estimator $(X'X)^{-1}\hat{S}(X'X)^{-1}$ in Result 1 may be improved by multiplying it with $n/(n - K - 1)$, where n is the number of twin pairs in the data and K is the number of right-hand-side variables in the DF-regression. See MacKinnon and White (1985).

variables to the right-hand-side of the regression (see for instance Rodgers, Rowe, and Li (1994) for a discussion of such extensions).

The asymptotic distribution in (8) also reveals why a mere degrees-of-freedom adjustment after DF-analyses with double-entry data may be misleading. The covariance matrix that is reported with standard regression software after a DF-analysis is $\hat{\sigma}_e^2 (X'X)^{-1}$, where $\hat{\sigma}_e^2$ is the residual variance of the regression. A degrees-of-freedom adjustment of this covariance matrix approximately increases all elements of this matrix by a factor 2. Such a proportional adjustment of the covariance matrix can only yield correct standard errors for $\hat{\beta}_{\mathrm{DF}}$ if the variance of the moment restrictions $\Sigma = \mathrm{E}[x_j' \varepsilon_j \varepsilon_j' x_j]$ can be factored into a constant term and $(X'X)$. Since the DF-regression is heteroscedastic, this is not possible and any proportional 'degrees of freedom adjustment' after DF-regressions with double-entry data may lead to inappropriate standard errors. Our simulations below reveal that the correct estimation based on Result 1 can deviate quite substantially from that of a 'degrees of freedom adjustment'.

While we established the asymptotic distribution of $\hat{\beta}$ using a GMM framework, alternative derivations are also possible. In particular, the variance estimator in Result 1 can also be interpreted as a modification of the heteroscedasticity-consistent or 'robust' covariance estimators that are commonly used when OLS is applied to data with unknown heteroscedasticity (Newey and West 1987; White 1980). This interpretation is useful because the variance estimator in Result 1 not only provides means to account for the particularities of double-entry twin data, but it also accommodates more general forms of unknown heteroscedasticity. This feature is a welcome side-effect of Result 1 when DF-analyses are applied to traits whose distribution is less 'well-behaved' as the additive genetic model suggests. In Section 5 we investigate body mass index as one example of a heteroscedastic trait. Alternative examples are binary or censored observation where the distribution of the trait usually suggests the application of alternative methods (Kohler and Rodgers 1999).

3.3   Efficient Estimation with Double-Entry Data

In Section 2 we have pointed out the heteroscedasticity of DF-regressions that is due to the fact that a co-twin's trait is a better predictor of the proband's trait in MZ twin pairs as compared to DZ twin pairs whenever $h^2 > 0$. The same heteroscedasticity persists in the

GMM framework above, where the variance of the moment restriction $\mathrm{E}[x_j'\varepsilon_j\varepsilon_j'x_j]$ is different between MZ and DZ twins.[5] This heteroscedasticity does not affect the consistency of β_{DF}, and in Result 1 we provide an estimator for the covariance matrix of β_{DF} that accounts for this heteroscedasticity. Nevertheless, an analogy with regression analysis suggests that we may be able to improve the efficiency of DF-analysis with alternative estimator, quite similar to the well-known fact that a weighted regression—with weights equal the inverse of the residual variance—is more efficient in the presence of heteroscedasticity than standard OLS.

In order to derive an 'efficient DF-analysis', we denote as $\Sigma(z_j) = \mathrm{E}[x_j'\varepsilon_j\varepsilon_j'x_j|z_j]$ the covariance matrix of the moment restrictions (6) conditional on the zygosity of a twin pair. Because these covariance matrices $\Sigma(z_j)$ are not of full rank in the subset of MZ or DZ twins, we denote in this Section as $\Sigma(z_j)^-$ the Moore-Penrose inverse, or pseudo-inverse, of $\Sigma(z_j)$.

Following the 'general efficiency framework' in Newey and McFadden (1994), a more efficient estimator can be derived by specifying the moment condition

$$\mathrm{E}[\mathrm{E}(x_j'x_j|z_j)\Sigma(z_j)^-x_j'\varepsilon_j] = 0. \tag{9}$$

This moment condition is a weighted version of our earlier condition in equation (6), and the weights depend on the zygosity-specific covariance matrix $\Sigma(z_j)$ of this earlier moment condition. Since the estimation of β based on (9) involves the unknown covariance matrices $\Sigma(z_j)$, this efficient estimator needs to be implemented via a two-step procedure. In the first step a consistent—but inefficient—estimator for β is used, such as for instance the standard double-entry DF-regression analyzed in the previous Section. The residuals of this initial estimation can then be used to estimate $\Sigma(z_j)$. The second step then reestimates β based on the weighted moment condition (9) using the estimate of $\Sigma(z_j)$ obtained in the first step. The asymptotic equivalence of such a two-step approach to the actual knowledge of $\Sigma(z_j)$ is one central insight of GMM and related methods (e.g., see Newey and McFadden 1994 for a formal discussion).

---

[5]In the case of standard DF analysis without additional covariates it is possible to solve for $\mathrm{E}[x_j'\varepsilon_j\varepsilon_j'x_j]$ explicitly in terms of σ_w^2, $\rho(z_j)$, and other higher-order moments of the joint distribution of $(w_{1j}, w_{2j})'$. However, this solution is cumbersome and not particularly informative, and it is therefore not reported here.

The empirical counterpart of the moment restriction in (9) is given by

$$\frac{1}{n}[X'_{\mathrm{MZ}}X_{\mathrm{MZ}}\hat{S}^-_{\mathrm{MZ}}X'_{\mathrm{MZ}}(Y_{\mathrm{MZ}} - X_{\mathrm{MZ}}\beta) + X'_{\mathrm{DZ}}X_{\mathrm{DZ}}\hat{S}^-_{\mathrm{DZ}}X'_{\mathrm{DZ}}(Y_{\mathrm{DZ}} - X_{\mathrm{DZ}}\beta)] = 0,$$

where $\hat{S}^-_{\mathrm{MZ}}$ and $\hat{S}^-_{\mathrm{DZ}}$ are the Moore-Penrose inverses of $\hat{S}_{\mathrm{DZ}} = [\sum_{\text{all DZ twin pairs}} x'_j \hat{e}_j \hat{e}'_j x_j]$ and $\hat{S}_{\mathrm{MZ}} = [\sum_{\text{all MZ twin pairs}} x'_j \hat{e}_j \hat{e}'_j x_j]$, $\hat{e}_j$ are the residuals obtained from a consistent first-step estimation of β, and the subscript 'MZ' or 'DZ' on Y and X denotes the fact that these matrices contain only the monozygotic or dizygotic twin pairs in the sample. The estimator for β that solves this empirical moment condition is

$$\begin{aligned} \hat{\beta}_{\mathrm{GMM}} &= H^{-1}(X'_{\mathrm{MZ}}X_{\mathrm{MZ}}\hat{S}^-_{\mathrm{MZ}}X'_{\mathrm{MZ}}Y_{\mathrm{MZ}} + X'_{\mathrm{DZ}}X_{\mathrm{DZ}}\hat{S}^-_{\mathrm{DZ}}X'_{\mathrm{DZ}}Y_{\mathrm{DZ}}), \qquad \text{where} \quad (10)\\ H &= (X'_{\mathrm{MZ}}X_{\mathrm{MZ}}\hat{S}^-_{\mathrm{MZ}}X'_{\mathrm{MZ}}X_{\mathrm{MZ}} + X'_{\mathrm{DZ}}X_{\mathrm{DZ}}\hat{S}^-_{\mathrm{DZ}}X'_{\mathrm{DZ}}X_{\mathrm{DZ}}). \qquad\qquad (11) \end{aligned}$$

The asymptotic distribution of $\hat{\beta}_{\mathrm{GMM}}$ follows directly from the standard GMM theory:

Result 2 (efficient DF estimation) *The estimator $\hat{\beta}_{GMM}$ defined in (10) is asymptotically distributed as $\hat{\beta}_{GMM} \xrightarrow{a} N(\beta, H^{-1})$, where H is defined in (11) above. This estimator is efficient in the class of minimum distance estimators (which includes all weighted regression analyses).*

The GMM estimator $\hat{\beta}_{\mathrm{GMM}}$ is more efficient than the standard DF–estimator $\hat{\beta}_{\mathrm{DF}}$ in the sense that the difference in the covariance matrices of these two estimators, asy.var$(\hat{\beta}_{\mathrm{DF}})$−asy.var$(\hat{\beta}_{\mathrm{GMM}})$, is positive semi-definite. In large samples $\hat{\beta}_{\mathrm{GMM}}$ is therefore equivalent or preferable to the standard DF-regression. The efficient estimation in Result 2 is closely related to similar efficient GMM estimators proposed by Im (1998) and Cragg (1983) for heteroscedastic data. However, our approach additionally accounts for the twin-structure in which only observations on twin pairs, and not on individuals, can be treated as independent observations.

It turns out, however, that $\hat{\beta}_{\mathrm{GMM}}$ and $\hat{\beta}_{\mathrm{DF}}$ have the same asymptotic distribution in the 'augmented DF-model'. This frequently used version of the DF-regression is without additional covariates, and the right-hand-side variables only consist of a constant, the co-twin's value to the trait, the degree of genetic relatedness and its interaction between the co-twin's trait value (see Section 2). The asymptotic distribution of the coefficients $\hat{\beta}_{\mathrm{DF}}$ is given in Result 1, and its asymptotic variance is not improved by using the GMM estimator in Result 2.

This finding can be seen by choosing a reparameterization of the DF-model that is statistically equivalent to the augmented DF-model. Denote as $1_{\mathrm{MZ}}(z_j)$ the indicator variable that equals one if a twin pair is monozygotic and let $1_{\mathrm{DZ}}(z_j)$ be the corresponding indicator for dizygotic twins. The augmented DF-model is statistically equivalent to regressing (in double-entry data) the first twin's trait w_{1j} on $\tilde{x}_{1j} = (1_{\mathrm{MZ}}(z_j), 1_{\mathrm{MZ}}(z_j)w_{2j}, 1_{\mathrm{DZ}}(z_j), 1_{\mathrm{DZ}}(z_j)w_{2j})$. The parameters in this modified model therefore estimate the mean and the within twin-pair correlation of the trait w_{ij} separately for MZ and DZ. The parameters therefore no longer have the interpretation of c^2 and h^2 as in the DF-model, but statistically this second model is equivalent to the augmented DF-model. In this modified regression it is clear that there is no efficiency gain in estimating the model with data that combines both MZ and DZ twin pairs—the first two parameter could equally well be estimated using only the MZ twin pairs, and the latter using only the DZ twin pairs. However, the problem of heteroscedasticity is absent within the subset of only MZ or DZ twins. Since the GMM estimator in Result 2 is equivalent to OLS in the absence of heteroscedasticity, it is clear that the augmented DF-model is not improved by adopting the GMM estimator $\hat{\beta}_{\mathrm{GMM}}$.

In applications of DF-analysis where additional covariates are included among the explanatory variables—such as observed environmental differences between the twins or sex in tests for gender differences—the two estimators $\hat{\beta}_{\mathrm{GMM}}$ and $\hat{\beta}_{\mathrm{DF}}$ will usually not be equivalent and the efficiency of the estimates can be improved by using our GMM based 'efficient DF-estimator'. Our simulations in the next Section will reveal the extent to which such efficiency gains are to be expected with data. However, since the computation of $\hat{\beta}_{\mathrm{GMM}}$ is not much more complicated that than that of $\hat{\beta}_{\mathrm{DF}}$, the former is also easily implemented and directly compared with the latter.

4   Simulation

In this Section we apply the augmented DF-model to simulated data and compare the standard errors given in Result 1 with a degrees of freedom adjustment and a bootstrap estimate of these standard errors. We choose three different levels of c^2 and h^2 in these simulated data in order to assess the importance of a correct variance estimation in a context

Table 1: Summary statistics for the estimated c^2 and h^2 in simulations

	Estimated value of		Estimated std. errors of	
	c^2	h^2	c^2	h^2
Simulation 1:	$\mathrm{E}[c^2] = .15$	$\mathrm{E}[h^2] = .15$		
mean	0.1385	0.1584	0.1260	0.1672
1st quartile	0.0503	0.0337	0.1223	0.1613
median	0.1459	0.1423	0.1257	0.1663
3rd quartile	0.2316	0.2784	0.1297	0.1730
Simulation 2:	$\mathrm{E}[c^2] = .15$	$\mathrm{E}[h^2] = .35$		
mean	0.1388	0.3579	0.1155	0.1469
1st quartile	0.0560	0.2509	0.1114	0.1412
median	0.1451	0.3396	0.1153	0.1469
3rd quartile	0.2206	0.4609	0.1191	0.1522
Simulation 3:	$\mathrm{E}[c^2] = .25$	$\mathrm{E}[h^2] = .5$		
mean	0.2392	0.5084	0.0919	0.1063
1st quartile	0.1745	0.4294	0.0873	0.1014
median	0.2474	0.4918	0.0918	0.1063
3rd quartile	0.3063	0.5835	0.0965	0.1106

with (a) relatively low heritability and shared environmental influences, and (b) a modest and high importance of genetic and shared environmental influences on a trait.

The simulated data consist of 300 DZ and 200 MZ twin pairs. We first generate two i.i.d. $N(0, 1)$ distributed random variables for each twin pair and then calculate the realizations of the trait w_{1j} and w_{2j} as a linear combination of these random variables. We choose this linear combination so that the trait w_{ij} has an expected value of zero, a variance of one, and a correlation within twin pairs of $\rho(z_j) = c^2 + h^2 R(z_j)$, where $R(z_j)$ equals one for MZ and 0.5 for DZ twins. We create 100 realizations of this data-set and then apply the augmented DF–regression with different estimators of the standard errors to these data. We focus our discussion on the two primary parameters of interest in these DF-regressions, namely the heritability h^2 and the shared environment influence c^2.

The expected values of c^2 and h^2 in the three different simulation specifications are (a) $\mathrm{E}(c^2) = .15$, $\mathrm{E}(h^2) = .15$ (low influence of genetic and shared environmental factors), (b) $\mathrm{E}(c^2) = .15$, $\mathrm{E}(h^2) = .35$ (moderately strong genetic influence on trait), and (c) $\mathrm{E}(c^2) = .25$, $\mathrm{E}(h^2) = .5$ (relatively high influence of genetic and shared environmental factors). Table 1 reports summary statistics of the estimated values of c^2 and h^2 that were obtained from these simulated data. On average, the estimates of c^2 and h^2 are quite close to the underlying

expected values, but at the same time there are substantial fluctuations in these estimates caused by random variations in the simulated data sets. The above Table also summarizes the standard errors associated with the estimates of c^2 and h^2 (the standard errors are based on Result 1).

The comparison of these standard errors with alternative estimates based on bootstraps of the simulated data[6] and on a degrees-of-freedom adjustment of the double entry results[7] is given in Figure 1. The boxplots on the left-hand side of this Figure show the ratio of the bootstrapped standard errors to the asymptotic standard errors obtained from Result 1. The graphs reveal a very high agreement of the asymptotic standard errors with the respective bootstrap estimates. On average the two estimates agree almost completely, and in most simulations the deviation is less than 5%. Such deviations are to be expected since the bootstrap is in itself a random process that is subject to sampling variation. These bootstrap estimates therefore support the argument that the 'large-sample theory', which underlies the asymptotic results in Section 3, provides a good approximation for the properties of the estimator already in samples with about 500 twin pairs.

More relevant are the boxplots on the right-hand side of Figure 1, which depict the ratio of the degrees-of-freedom adjusted standard errors to the asymptotic standard errors. In Simulation 1, where both genetic and shared environmental factors have only a small influence on the trait, the two estimates differ only to a modest extent. The degrees-of-freedom adjustment yields generally larger estimates for the standard errors, but the overestimate is quite modest and usually less than 10%.

The extent to which the degrees-of-freedom adjustment overestimates the asymptotic standard errors increases in Simulation 2, which is based on a modest heritability of $\mathrm{E}\, h^2 =$ .35 and a relatively small relevance of shared environmental factors of $\mathrm{E}\, c^2 = .15$. The boxplot confirms our earlier arguments that a proportional increase of standard errors is not optimal in DF-analyses since the regression is heteroscedastic. In Simulation 2, the degrees-of-freedom adjustment overestimates the standard errors of h^2 substantially more than those of c^2: the overestimate for h^2 is mostly between 10–18 percent, while that for c^2

---

[6]The bootstrap estimates are based on 150 random drawings (with replacement) of twin pairs in each of the simulated data sets.

[7]The variance-covariance matrix obtained after the regression with double-entry data was multiplied by a factor of 2, or equivalently, the standard errors were increased by a factor of $\sqrt{2}$.
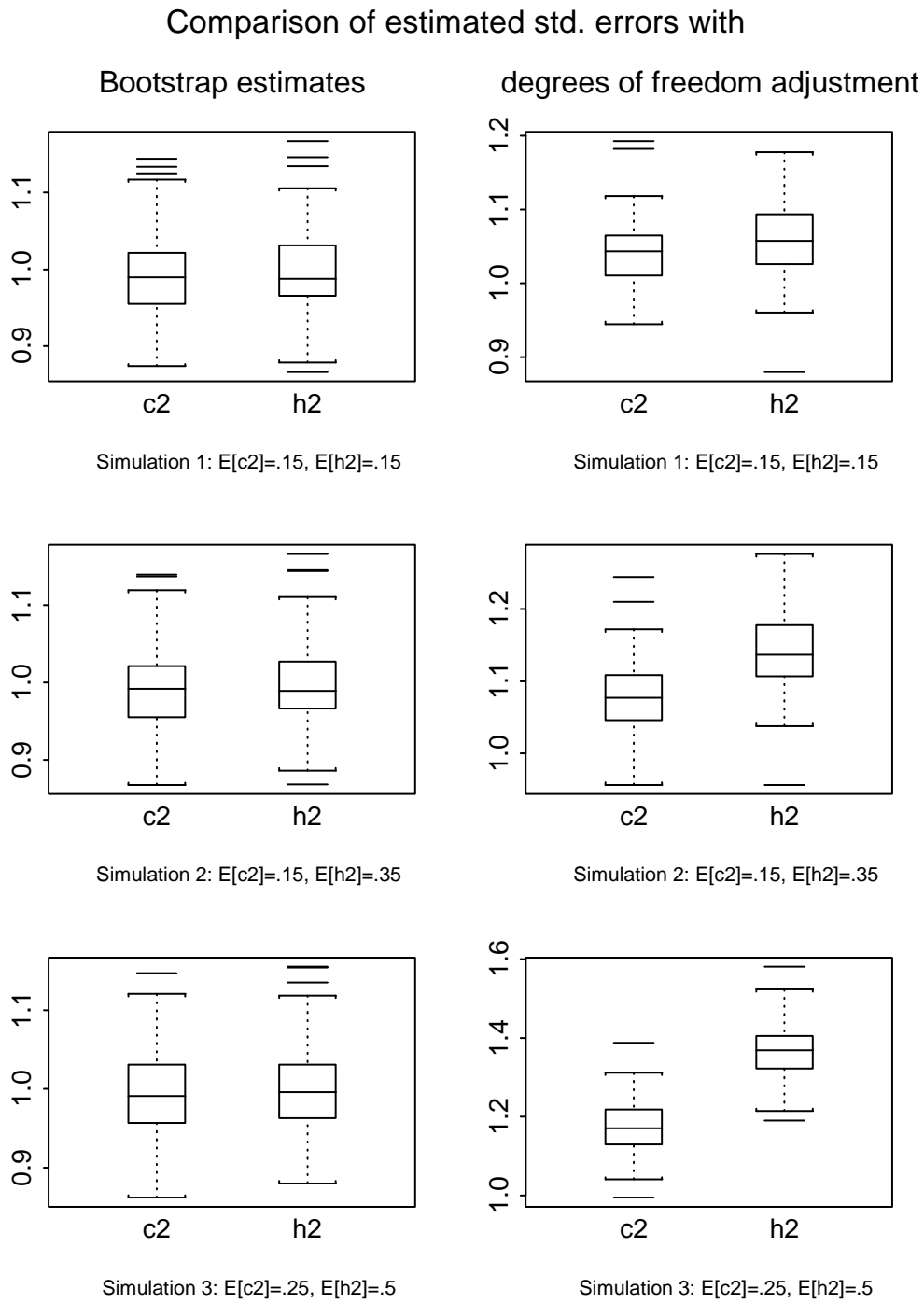
Figure 1: Comparison of standard errors on the basis of Result 1 with estimates based on bootstraps and a degrees of freedom adjustment. The boxplots show the ratio of the bootstrapped and degrees of freedom adjusted standard errors of c^2 and h^2 to the asymptotic standard errors established in Result 1

is mostly between 5–10 percent.

This pattern is even more pronounced in Simulation 3, which represents situations with modest shared environmental influences and strong genetic influences on a trait ($\mathrm{E}\,c^2 = .25$, $\mathrm{E}\,h^2 = .5$). The degrees-of-freedom adjustment overestimates again the standard errors associated with c^2 and h^2, and the extent of this overestimate has further increased. Moreover, the divergence between the coefficients c^2 and h^2 has become stronger. The standard error of h^2 tends to be overestimated by 32–40%, while those of c^2 are "only" overestimated by 13–20%.

Although the estimates for c^2 and h^2 are of primary interest in a DF-analysis, it is worth mentioning that the degrees-of-freedom adjustment also overestimates the standard errors of the remaining coefficients β_0 (constant term) and β_2 (coefficient on $R(z_j)$) in the augmented DF-analysis. These overestimates can be quite substantial and range from an average 65% in Simulation 1 to an average of 100% in Simulation 3. At the same time, the asymptotic standard errors for β_0 and β_2 conform very well with the respective estimates obtained from bootstrapping the data.

The comparison between the degrees-of-freedom adjustment and the asymptotically correct estimation of the standard errors can be interpreted in two directions. The first interpretation emphasizes that the degrees-of-freedom adjustment tends to be conservative and provides standard errors that exceed the correct values. A researcher therefore would not erroneously interpret an estimated genetic or shared environmental influence as statistically significant, and she would report p-values that tend to understate the true significance.

The alternative interpretation emphasizes that the overestimates by a degrees-of-freedom adjustment can be of a magnitude that substantially reduces the power to detect significant genetic or shared environmental factors. In many cases it may be that the exaggeration of the standard errors renders an estimate of c^2 or h^2 statistically insignificant, and the researcher may fail to infer relevant genetic or environmental influences. Adopting an asymptotically correct estimate of the standard errors is therefore not only more satisfactory from a statistical standpoint, but it is also facilitates the statistical identification of relevant genetic and environmental influences. The simulation results suggest that the gains from adopting the asymptotic estimates substantially outweigh the computational costs and the reliance on large-sample theory in the derivation of the asymptotic distribution.

In Table 2 we provide a comparison of the 'efficient DF-analysis' developed in Section 3.3 with the standard DF-analysis using the correct standard errors in Result 1. The analysis is based on our earlier 'Simulation 3', augmented by an observed non-shared environmental effect ΔE. We generate this effect via a twin specific $N(0,1)$ random variable with is added to the trait w_{ij} via a coefficient $\beta_{\Delta E} = 0.2$. The difference ΔE in the observed environment is then calculated as the difference of these two random variables. We reduce the sample size to 180 DZ and 120 MZ twins in this simulation because the coefficient h^2 in the case of relatively high heritability is very accurately estimated already in smaller samples.

The top panel of Table 2 provides summary statistics of the estimated coefficients and their standard errors when the augmented DF model is applied to these data. If the difference in the environment is accounted for, the estimated parameters—on average—agree again with their expected values. If the environmental difference had not been accounted for, the estimate c^2 would have been reduced by about 20% on average.

The bottom panel of Table 2 reports the respective coefficient and standard errors obtained from the efficient DF-analysis based on GMM. It is quite apparent that the efficient analysis has almost no effect on the estimated coefficients itself—not surprising given that both analysis provide consistent estimates of the respective coefficients—and it is therefore difficult to rank the two methods in terms of their estimated coefficients. As expected, the GMM estimation leads to smaller standard errors of the coefficients due to a gain in efficiency, but the reductions in the estimated standard errors of the coefficients c^2 and h^2 are very modest.

In summary, the simulations in this Section suggest that the correct estimation of standard errors in DF-analysis is of considerable importance. Depending on the strength of genetic and shared environmental influences on the trait value, the degrees-of-freedom adjustment can lead to overestimates of the standard errors exceeding 30% for the main parameters of interest. Although these overestimates will generally not lead to an erroneous inference of genetic or shared environmental influences on a trait, the power of the DF-analysis to detect heritable traits is substantially reduced by this overestimation of the standard errors. The asymptotic standard errors established in Section 3.2 avoid this 'loss of statistical power'. Moreover, these asymptotic results compare very well with estimates of the standard errors that are derived through bootstrapping of the data. The 'large-sample

Table 2: Comparison of efficient DF-estimation with standard DF-regression (Simulation is based on the settings of 'Simulation 3' with an additionalal observed nonshared environmental influences)

	Estimated value of			Estimated std. errors of		
	c^2	h^2	$\beta_{\Delta E}$	c^2	h^2	$\beta_{\Delta E}$
DF-analysis	$E[c^2] = .15$	$E[h^2] = .15$	$E[\beta_{\Delta E}] = .2$			
mean	0.2589	0.4921	0.2012	0.1145	0.1328	0.0291
1st quartile	0.1873	0.3964	0.1785	0.1083	0.1251	0.0278
median	0.2573	0.5028	0.1981	0.1140	0.1326	0.0291
3rd quartile	0.3293	0.5722	0.2200	0.1209	0.1410	0.0303
efficient DF-analysis						
mean	0.2577	0.4973	0.2043	0.1143	0.1323	0.0283
1st quartile	0.1891	0.3964	0.1806	0.1083	0.1245	0.0271
median	0.2563	0.5140	0.2023	0.1139	0.1308	0.0284
3rd quartile	0.3304	0.5802	0.2230	0.1207	0.1409	0.0296

theory' used to establish these asymptotic standard errors therefore seems to be a good approximation already in twin samples of a modest size.

The efficient DF-analysis of Section 3.3 provides an improvement of the usual regression approach when DF-analysis is combined with additional covariates. At the same time, the gains from implementing this efficient analysis seem considerably less relevant than the correct estimation of the standard errors in DF-regressions. Nevertheless, small improvements in the standard errors of our estimates did occur after the implementation of efficient GMM-based DF-analysis, and in some studies these small gains in efficiency may improve the analysis.

5   An Application to Danish Twin Data

In order to support the conclusion obtained from simulated data, we also apply our modifications of the DF-analysis to age-standardized body height and body mass index (BMI) of female Danish twins. The data consist of 696 female twin pairs who were born during 1890–1920 and interviewed in 1966. These data have earlier been used by (Herskind et al. 1996) in order to assess the age- and sex-specific heritability of BMI. This earlier study also contains a description of the Danish twin register and the survey in which the data were collected.

The respondents have an average body height of 160.3 cm, with a standard deviation of 6.04, and they have an average body mass index of 25.3 with a standard deviation of 4.5. Since the distribution of BMI is skewed to the right, we follow Herskind et al. and apply the DF-regression to the natural logarithm of BMI. Moreover, we also standardize height and BMI by regressing the respective traits on a third-order polynomial in age and removing any age-differences in the average value of the traits.[8]

Height and body mass index constitute 'classic' traits for behavioral genetic studies with high heritability and low shared environmental influences. Their analysis with DF-analysis is interesting because of the comparability to several other studies, and because the distribution of these traits in our data exhibit some subtle differences that affect the performance of traditional DF-analyses with degrees-of-freedom adjusted standard errors. In particular, our simulations in Section 4 have suggested that the degrees-of-freedom adjustment in double-entry DF-analysis is conservative, i.e., that it tends to overestimate the correct standard errors. While this pattern holds when DF-analysis is applied to body height, it is reversed when DF-analysis is applied to body mass index: the degrees-of-freedom adjustment in the BMI application *underestimates* the standard errors due to the presence of additional heteroscedasticity. The asymptotic standard errors in Result 1, however, are unaffected by this additional heteroscedasticity, and they continue to yield correct estimates of the variability of the inferred parameters. The comparison of DF-analysis for height and BMI therefore emphasizes the need to estimate standard errors that are correct in large samples and also robust against deviations from the assumptions of the underlying genetic model.

In Table 3 we first report the results obtained from the DF-analysis of body height. Model 1 uses a degrees-of-freedom adjustment to establish the standard errors of the estimates, while Model 2 calculates the asymptotically correct standard errors on the basis of Result 1. Since the difference between these two methods is only in the calculation of the covariance matrix, both models yield identical coefficients. The DF-estimates for heritability agree with earlier studies of body height (e.g., Carmichael and McGue 1995; Plomin 1990), and our results reveal a relatively high heritability ($h^2 = 0.64$) with very modest

---

[8]These age-standardized traits have a mean of 160.3 cm (height) and 3.26 (log of BMI), and a standard deviation of 6.01 (height) and 0.164 (log of BMI).

Table 3: Application of DF-analysis to the age-standardized body weight and fertility of Danish twins born 1953–63

Method Std. Errors	DF degrees of freedom adjustment	DF asymptotic std. errors (Result 1)	DF degrees of freedom adjustment	DF asymptotic std. errors (Result 1)	efficient DF (Result 2)
Body height	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
c^2	0.133	0.133	0.137	0.137	0.137
	(0.092)	(0.088)	(0.092)	(0.088)	(0.087)
degree of genetic	-103.7	-103.7	-102.9	-102.9	-103.3
relatedness	(20.20)**	(16.48)**	(20.19)**	(16.45)**	(16.33)**
h^2	0.644	0.644	0.640	0.640	0.642
	(0.126)**	(0.103)**	(0.126)**	(0.103)**	(0.102)**
Δ ever smoked[a]	—	—	0.470	0.470	0.461
			(0.419)	(0.375)	(0.340)
Δ number of	—	—	0.005	0.005	-0.0002
cigarettes[b]			(0.035)	(0.035)	(0.030)
constant	139.3	139.3	138.6	138.6	138.7
	(14.72)**	(14.17)**	(14.72)**	(14.14)**	(14.02)**
N (twin pairs)	696	696	696	696	696
R^2	0.356	0.356	0.359	0.359	—
log of BMI	*Model 6*	*Model 7*	*Model 8*	*Model 9*	*Model 10*
c^2	0.021	0.021	0.036	0.036	0.016
	(0.099)	(0.109)	(0.097)	(0.108)	(0.106)
degree of genetic	-2.143	-2.143	-2.133	-2.133	-2.313
relatedness	(0.427)**	(0.459)**	(0.422)**	(0.444)**	(0.418)**
h^2	0.666	0.666	0.663	0.663	0.719
	(0.133)**	(0.143)**	(0.131)**	(0.138)**	(0.130)**
Δ ever smoked[a]	—	—	-0.0249	-0.0249	-0.0261
			(0.0121)*	(0.0098)*	(0.0098)**
Δ number of	—	—	-0.0018	-0.0018	-0.0017
cigarettes[b]			(0.0010)+	(0.0009)*	(0.0008)*
constant	3.149	3.149	3.101	3.101	3.164
	(0.317)**	(0.351)**	(0.314)**	(0.349)**	(0.342)**
N (twin pairs)	696	696	696	696	696
R^2	0.265	0.265	0.284	0.284	—

Notes: Standard errors in parentheses. *p-values:* $+\ p < 0.10$; $*\ p < 0.05$; $**\ p < 0.01$. All analyses are based on double-entry data. (*a*) Ever smoked more than 100 cigarettes, 50 cigars, or 5 packs of tobacco. (*b*) Cigars and tobacco were converted into 'cigarette-equivalents' by adjusting the respective distributions of cigar and tobacco consumption to the same mean and variance as the distribution of cigarette consumption.

and statistically insignificant shared environmental influences. Moreover, this heritability increases to $h^2 = 0.82$ if the insignificant shared-environment effect is dropped from the analysis and the estimate for heritability is obtained from an AE model only. Since we are mainly interested in the methodological aspects of the analysis, however, we do not report these AE estimates and continue our discussion of the ACE model shown in Table 3.

The difference between Model 1 and 2 occurs with respect to the precision that is associated with these estimates. Most relevantly, the standard error of h^2 is about 20% lower and that of c^2 is about 5% lower, if they are based on the correct calculation in Result 1 instead of the degrees-of-freedom adjustment. This pattern agrees with our earlier simulation results in which the degrees-of-freedom adjustment usually provided a conservative estimate of the standard errors.

One important environmental difference that can affect body height and BMI is smoking. Models 3–5 therefore include the difference in the twins' smoking behavior in order to test whether the estimates of heritability for height are affected by potentially different smoking behavior of the twins. The results in all models suggest, however, that body height is not affected by smoking. Moreover, including smoking in the analysis leaves the estimates for c^2 and h^2 almost unaffected.

As in our earlier analyses, the degrees-of-freedom adjustment overstates the variability of the coefficients. The extent of this overestimation differs quite considerable between the coefficients, ranging from virtually zero to more than 25%. The efficient DF analysis in Model 5 provides an additional small reduction of the standard errors in the DF-analysis, and it results in small changes in the point estimates for the coefficients.

The bottom part of Table 3 reports the corresponding analyses for the logarithm of body mass index. Model 6 and 7 present the results of the augmented DF-analysis with degrees-of-freedom adjusted and asymptotic standard errors. The coefficients suggest that the heritability of BMI is similar to that of body height. The estimated h^2 equals 0.67 and there is a virtual absence of shared environmental influences (the estimate for heritability increases to $h^2 = 0.69$ in an AE model). This pattern agrees with the earlier analysis of the same data with structural equation models (Herskind et al. 1996), although this study has found a somewhat higher heritability based on the estimation of an AE model without shared environmental influences. Moreover, our results agree with analyses of body

weight that have also found heritabilities of approximately 0.70 (Grilo and Progue-Geile 1991; Rodgers et al. 1994).

The striking aspect in the comparison of Models 1 and 2 is that the degrees-of-freedom adjustment yields standard errors that are approximately 10% below the asymptotic standard errors. The latter, however, seem to be more appropriate because they correspond closely with estimates of the standard errors obtained by bootstrapping the data.[9] The asymptotic standard errors therefore not only provide a correct assessment of the variability of the estimated coefficients, but they also reveal that the degrees-of-freedom adjustment would have *underestimated* this variability.[10]

The reason for this pattern is due to the fact that the logarithm of BMI (and also BMI itself) is heteroscedastic beyond heteroscedasticity that is due to the twin design. In Section 2 we have shown that the variance of the residual ε_{ij} in DF-analysis depends on zygosity (see equation 5), but not on the value of the co-twin's trait. We investigate this second condition using a Cook-Weisberg test for heteroscedasticity. This test reveals whether the variance of the residual depends on the value of co-twin's trait. The null-hypotheses of homoscedastic residuals with respect to the co-twin's trait value is rejected in the BMI analysis (p-value < 0.0001), while this null-hypotheses is not rejected in the analysis of body height (p-value $= .24$). Although height and body mass index appear like very similar traits, the above tests reveal that they differ in our data in an important aspect: the residual variance in DF-regressions does not depend on the co-twin's trait value in the analysis of height, but it does in the analysis of BMI (taking the logarithm reduces this heteroscedasticity somewhat, but does not eliminate it).

It is well-known that such heteroscedasticity biases the estimates for the standard errors. The degrees-of-freedom adjustment, which proportionally inflates the standard errors obtained after a DF-regression, is therefore based on an inappropriate starting point. In this specific example this implies that the degrees-of-freedom adjustment is no longer conservative but actually underestimates the variability of the estimated coefficients.

The presence of this additional heteroscedasticity, however, does not affect the asymp-

---

[9]The bootstrap was based an 500 random samples from the data. The bootstrapped standard errors exceed the asymptotic standard errors by 1–1.5% for all four coefficients.

[10]Although we do not pursue this question further, the distortion of standard errors due to such additional heteroscedasticity also exists in structural equation models and is not restricted to DF-analyses.

totic standard errors in Result 1. In particular, the derivation of this Result within the framework of GMM estimation does not require that the residuals ε_{ij} are homoscedastic. Some relatively weak regularity conditions are sufficient in order for Result 1 to yield asymptotically correct standard errors. As indicated earlier, this robustness of the estimate in Result 1 is a welcome side-effect of our method. The estimated standard errors remain asymptotically correct even if the residual variance in DF-analysis is heteroscedastic and depends on the co-twin's trait value (recall that DF-analysis is always heteroscedastic with respect to zygosity whenever $h^2 > 0$).

Models 8–10 in Table 3 additionally include differences in smoking behavior in the DF-regressions. The inclusion of these non-shared environments does not substantially affect the estimates. The heritability remains at $h^2 = .66$, and shared environmental influences remain absent. Smoking itself, however, is clearly negatively associated with BMI.

Most relevant in the present context is the comparison of the standard errors in these three models. The comparison reveals that the asymptotic standard errors tend to be larger than those of the degrees-of-freedom adjustment for the coefficients c^2 and h^2—similar to our earlier discussion of Models 6 and 7—while the pattern is reversed for the covariates pertaining to smoking behavior.

The efficient estimation in Model 10 yields smaller standard errors than Model 9, and it yields slightly different estimates for the coefficients. In particular, the shared environmental influence is somewhat diminished and heritability is slightly increased. The coefficients of the smoking variables remain almost unchanged. Although the efficient estimation yields smaller standard errors, it is generally not possible to rank the point estimates of Model 9 and 10 in terms of their accuracy. The latter are more efficient, but both estimates are consistent and thus 'close' to the true coefficients in large samples.

The application of our methods to the Danish data supports the findings of our simulation study. The degrees-of-freedom adjustment of the standard errors after a DF-regression with double entry data provides an overestimate of the correct standard errors when the data conform with the theoretical model. When the data are heteroscedastic like BMI, however, this may no longer be the case and the degrees-of-freedom adjustment can yield standard errors that are too small. The calculation of the asymptotically correct standard errors is therefore advisable, both because it increases the power of the statistical tests

and because these standard errors are also 'robust' with respect to more general forms of heteroscedasticity which are not directly related to the twin-design.

The efficient DF-analysis seems to provide coefficient estimates that are very close to those of the standard DF-analysis and the 'efficiency gains' in terms of smaller standard errors on these coefficients is relatively modest. At the same time, the computation of this GMM estimator is relatively simple and even modest efficiency gains may contribute to the statistical properties of an analysis.

6   Conclusions

DeFries and Fulker's (1985) regression procedure for estimating heritability (h^2) and common environmental influences (c^2) has been frequently applied to twin data and other selected and unselected samples with multiple kinship levels. The ambiguity in unselected samples as to which twin should be used as the proband is frequently resolved by double-entry: Each twin pair is entered twice, and each member of a twin pair provides once the dependent and once the explanatory variable. While the consistency of the regression estimates for heritability and shared environmental influences is not affected by this procedure, the standard errors obtained after regressions with double-entry data need to be adjusted. It has been common practice to adopt a degrees-of-freedom adjustment in which the standard errors are based on the number of twin pairs, instead of the number of observations in the double-entry data. This procedure, however, is problematic due to the correlation and heteroscedasticity of the residuals. The degrees-of-freedom adjusted standard errors may therefore deviate, possibly substantially, from the correct distribution of the estimated coefficients.

In this paper we establish the correct asymptotic distribution of the coefficients of DF-analyses with double entry data and we provide a simple formula to estimate the covariance matrix of the coefficients in DF-regressions. Applications of this method to simulated data sets show that the degrees-of-freedom adjustment overestimates the correct standard errors of the coefficients. The extent of this overestimation exceeds 20% for traits with moderate heritability, and it tends to become larger for traits with higher heritability. Moreover, the overestimation of the standard errors is not uniform across coefficients and it can vary

considerably between c^2 and h^2.

Although the degrees-of-freedom adjustment seems to provide conservative estimates of the standard errors, the extent of this overestimation substantially reduces the statistical power of DF analyses to detect relevant genetic and shared-environment influences on a trait. The estimation of the asymptotically correct standard errors proposed in this paper avoids this loss of statistical power, and the empirical validity of these asymptotic standard errors is confirmed by bootstrap estimates.

Our analyses also show that DF-regressions are inefficient when additional covariates, such as observed differences in non-shared environments, are included in the analysis. We provide an 'efficient DF analysis' that is preferable in such situations because it provides more precise estimates in large samples. Our simulations show that such efficiency gains are possible. At the same time, these gains are only of modest importance, especially when compared to the effect of correctly estimating the asymptotic standard errors.

The applications of our methods to the body height and body mass index of female Danish twins supports the conclusions of our simulation study: the estimated asymptotic standard errors tend to be below those obtained from a degrees-of-freedom adjustment, and the statistical power to detect genetic effects by DF analysis is substantially improved by adopting the methods proposed in this paper. Moreover, our method is relatively robust to additional sources of heteroscedasticity which are not due to the twin design. In the presence of such heteroscedasticity, as for instance in the analysis of body mass index, the degrees-of-freedom adjustment can yield standard errors that too low, while our modified estimation continues to yield an appropriate assessment of the variability of the estimates.

References

Carmichael, C. M. and M. McGue (1995). A cross-sectional examination of height, weight and body-mass index in adult twins. *Journals of Gerontology Series A – Biological Sciences and Medical Sciences 50*(4), B237–244.

Cherny, S. S., L. Cardon, D. Fulker, and J. DeFries (1992). Differential heriability across levels of cognitive ability. *Behavior Genetics 22*(2), 153–162.

Cragg, J. G. (1983). More efficient estimation in the presence of heteroscedasticity of unknown form. *Econometrica 51*(3), 751–763.

Cyphers, L. H., K. Philips, D. W. Fulker, and D. A. Mrazek (1990). Twin temperament during the transition from infancy to early childhood. *Journal of Personality and Social Psychology 51*, 339–346.

DeFries, J. and D. Fulker (1985). Multiple regression analysis of twin data. *Behavior Genetics 15*(5), 467–73.

Fulker, D. W., L. R. Cardon, J. C. DeFries, W. J. Kimberling, B. F. Pennington, and S. D. Smith (1991). Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. *Reading and Writing: An Interdisciplinary Journal 3*, 299–313.

Grilo, C. M. and M. R. Progue-Geile (1991). The nature of environmental influences on weight and obesity: A behavior genetic analysis. *Psychological Bulletin 110*, 520–537.

Haggard, E. A. (1958). *Intraclass Correlations and the Analysis of Variance*. New York: Dryden Press.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*, 1029–1054.

Herskind, A. M., M. McGue, T. I. A. Sorensen, and B. Harvald (1996). Sex and age specific assessment of genetic and environmental influences on body mass index in twins. *International Journal of Obesity 20*(2), 106–113.

Im, K. S. (1998). Efficient estimation with grouped data. *Economic Letters 59*(2), 169–174.

Kohler, H.-P. and J. L. Rodgers (1999). DF-like analyses of binary, ordered and censored variables using Probit and Tobit approaches. *Behavior Genetics 29*(4), 221–232.

LaBuda, M. C. and J. DeFries (1990). Genetic etiology of reading disability: evidence from a twin study. In G. T. Paylidis (Ed.), *Perspectives on Dyslexia Vol. 1*. New York: Wiley & Sons.

LaBuda, M. C., J. C. DeFries, and D. W. Fulker (1986). Multiple regression analysis of twin data obtained from selected samples. *Genetic Epidemiology 3*(6), 425–433.

MacKinnon, J. G. and H. White (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics 29*, 305–325.

Newey, W. and K. West (1987). A simple positive semi-definite, heteroscadisticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Newey, W. K. and D. L. McFadden (1994). Large sample estimation and hypothesis testing. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics, Volume 4*, pp. 2111–2241.

Plomin, R. (1990). *Nature and Nurture: An Introduction to Human Behavioral Genetics*. Pacific Grove, CA: Brooks/Cole.

Rodgers, J. L., H.-P. Kohler, K. Kyvik, and K. Christensen (1998). Genes affect human fertility via fertility motivations: Findings from a contemporary Danish twin study. Max Planck Institute for Demographic Research, Rostock, Germany: Mimeo.

Rodgers, J. L. and M. McGue (1994). A simple algebraic demonstration of the validity of DeFries - Fulker analysis in unselected samples with multiple kinship levels. *Behavior Genetics 24*(3), 259–62.

Rodgers, J. L., D. Rowe, and C. Li (1994). Beyond natures versus nuture: Df analysis of nonshared influences on problem behavior. *Developmental Psychology 30*(3), 374–84.

Rodgers, J. L. and D. C. Rowe (1987). IQ similarity it twins, siblings, half siblings, cousins and random pairs. *Intelligence 11*, 199–206.

Waller, N. G. (1994). A DeFries and Fulker regression model for genetic nonadditivity. *Behavior Genetics 24*(2), 149–153.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica 48*, 817–838.

Zieleniewski, A. M., D. W. Fulker, J. C. DeFries, and M. C. LaBuda (1987). Multiple regression analysis of twin and sibling data. *Personality and Individual Differences 8*, 787–791.