

Swarthmore College

## Works

---

Economics Faculty Works

Economics

---

2005

# Deriving The GLS Transformation Parameter In Elementary Panel Data Models

Philip N. Jefferson

*Swarthmore College*, [pjeffer1@swarthmore.edu](mailto:pjeffer1@swarthmore.edu)

Follow this and additional works at: <https://works.swarthmore.edu/fac-economics>



Part of the [Economics Commons](#)

[Let us know how access to these works benefits you](#)

---

### Recommended Citation

Philip N. Jefferson. (2005). "Deriving The GLS Transformation Parameter In Elementary Panel Data Models". *American Economist*. Volume 49, Issue 1. 45-48. DOI: 10.1177/056943450504900103  
<https://works.swarthmore.edu/fac-economics/11>

This work is brought to you for free and open access by . It has been accepted for inclusion in Economics Faculty Works by an authorized administrator of Works. For more information, please contact [myworks@swarthmore.edu](mailto:myworks@swarthmore.edu).

---

Deriving the GLS Transformation Parameter in Elementary Panel Data Models

Author(s): Philip N. Jefferson

Source: *The American Economist*, Vol. 49, No. 1 (Spring, 2005), pp. 45-48

Published by: [American Economist](#)

Stable URL: <http://www.jstor.org/stable/25604312>

Accessed: 17/07/2014 14:15

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*American Economist* and *Omicron Delta Epsilon* are collaborating with JSTOR to digitize, preserve and extend access to *The American Economist*.

<http://www.jstor.org>

# DERIVING THE GLS TRANSFORMATION PARAMETER IN ELEMENTARY PANEL DATA MODELS

by Philip N. Jefferson\*

## Abstract

The Generalized Least Squares (GLS) transformation that eliminates serial correlation in the error terms is central to a complete understanding of the relationship between the pooled OLS, random effects, and fixed effects estimators. A significant hurdle to attainment of that understanding is the calculation of the parameter that delivers the desired transformation. This paper derives this critical parameter in the benchmark case typically used to introduce these estimators using nothing more than elementary statistics (mean, variance, and covariance) and the quadratic formula.

*“Deriving the GLS transformation that eliminates serial correlation in the error terms requires sophisticated matrix algebra.”*

—Jeffrey M. Wooldridge (2003)

## I. Introduction

This paper reconsiders the derivation of the Generalized Least Squares (GLS) transformation that eliminates serial correlation in the error terms in elementary panel data models. Such a reconsideration is warranted because traditional derivations are presented at a relatively high level of analytical sophistication. This creates a hurdle to a complete understanding of the pooled OLS, random effects, and fixed effects estimators and the connection between them. This is unfortunate because these estimators have risen in importance as the availability and use of panel data sets for testing economic hypotheses and policy analysis have dramatically expanded in recent years.

The approach taken to deriving the GLS transformation in this paper is unabashedly elementary. A benchmark case is used to illustrate how the GLS transformation parameter can be explicitly derived using a restriction on the error covariance. Although not perfectly general, this benchmark case is the one typically used to introduce the pooled OLS, random effects, and fixed effects estimators. A

value of examining this case is that the derivation requires the use of only elementary statistics (mean, variance, and covariance) and the quadratic formula.

Table 1 indicates how this transformation is treated in selected econometrics textbooks. As suggested by the comment by Wooldridge above, the hurdle is high indeed. For researchers making a first approach to these estimators, it appears that one’s ability to fully appreciate them is bounded by one’s facility with sophisticated matrix algebra. This paper seeks to lower this hurdle. Sections II and III present the benchmark model and the problem that causes OLS estimation to be inefficient, respectively. Section IV presents a simple scalar-based covariance restriction method for deriving the GLS transformation parameter in the benchmark case. Section V concludes.

## II. The Benchmark Model

The elementary error components model has the following structure:

$$y_{it} = \alpha_i + \beta x_{it} + w_{it} \quad (1)$$

for  $i = 2, \dots, N$  and  $t = 2, \dots, T$ . In this notation,  $i$  is an index for cross section units and  $t$  is an index for time periods. We assume that  $E(w_{it}) = 0$ ,  $E(w_{it}^2) = \sigma_w^2$ , and  $E(w_{it}w_{is}) = 0$  for  $t \neq s$ . In equation (1),  $\alpha_i$

\* Economics Department, Swarthmore College, Swarthmore, PA 19081, e-mail: pjeffer1@swarthmore.edu, voice: (610) 690-6856, fax: (610) 328-7352

I thank Paul Ruud and an anonymous referee for very helpful comments. This research is supported by a Swarthmore College Faculty Research Grant.

TABLE 1  
Textbook Treatment of GLS Transformation

Author(s)	Level	Treatment	Framework	Derivation
Ashenfelter et al. (2003)	U	Yes, p. 272	Scalar	No
Griffiths et al. (1993)	U/G	Yes, p. 577	Matrix	No
Gujarati (2003)	U/G	Yes, p. 649	Scalar	No
Pindyck & Rubinfeld (1998)	U/G	Yes, p. 254	Scalar	No
Ramanathan (2002)	U	Yes, p. 481	Scalar	No
Wooldridge (2003)	U/G	Yes, p. 470	Scalar	No
Green (2003)	G	Yes, p. 295	Matrix	Yes
Johnston & Dinardo (1997)	G	Yes, p. 392	Matrix	Yes
Judge et al. (1985)	G	Yes, p. 524	Matrix	Yes
Kmenta (1986)	G/U	Yes, p. 627	Matrix	Yes
Ruud (2000)	G	Yes, p. 638	Matrix	Yes
Wooldridge (2002)	G	Yes, p. 286	Matrix	Yes

Notes: U = Undergraduate, G = Graduate

permits differences across cross section units. These differences may have a random component thus

$$\alpha_i = \alpha + u_i \quad (2)$$

where  $E(u_i) = 0$ ,  $E(u_i^2) = \sigma_u^2$ , and  $E(u_i u_j) = 0$  for  $i \neq j$ . Substituting (2) into (1) yields

$$y_{it} = \alpha + \beta x_{it} + u_i + w_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \quad (3)$$

where the composite error is  $\varepsilon_{it} = u_i + w_{it}$  and by assumption  $E(u_i w_{it}) = 0$ ,  $E(u_i x_{it}) = 0$ , and  $E(w_{it} x_{it}) = 0$ , for all  $i, t$ .

The interpretation of the error components is that  $u_i$  represents an individual effect and that  $w_{it}$  represents unsystematic variation across time and cross section unit. Many important issues surrounding parameter estimation and interpretation using panel data can be considered in the benchmark case with  $N$  of arbitrary size and  $T = 2$ . In fact, it is not uncommon for the benchmark case to serve as the gateway to analysis of the pooled OLS, random effects, and fixed effects estimators.<sup>1</sup> Considerable emphasis is placed on the benchmark case below.

### III. The Problem

It is well-known that OLS estimation of the parameters in equation (3) is inefficient. The source of the problem is the individual effect that induces correlation in the error terms within each cross section unit. In the benchmark case, this is easy to see. For individual  $i$  the errors are  $\varepsilon_{i1} = u_i + w_{i1}$  and  $\varepsilon_{i2} = u_i + w_{i2}$ . Thus, the covariance of the error terms is

$$Cov(\varepsilon_{i1}, \varepsilon_{i2}) = \sigma_u^2 \quad (4)$$

An analogous result holds for each cross section unit. Thus, a generalized least squares (GLS) procedure that explicitly takes this covariance structure into account is needed.

### IV. Deriving Two Solutions

A more transparent derivation of the GLS transformation parameter in elementary panel data models requires some form of differencing for solutions to the problem. Intuition for this conclusion may be drawn from a more familiar but different problem: first order autocorrelated error terms. A common prescription for dealing with first order autocorrelation is generalized differencing. That prescription is based on knowledge of the first order autocorrelation parameter,  $\rho$ , that allows the researcher to uncover the underlying *iid* disturbance. The transformed model then paves the way for the assertion that the conditions of the Gauss-Markov theorem hold.<sup>2</sup>

Analogous to the case of first order autocorrelation, a non-zero error covariance in equation (4) is problematic for OLS estimation in the panel data case. Therefore, it seems intuitive to appeal to some form of differencing for possible solutions. Consider the following transformation of the errors for cross section unit  $i$ ,

$$\begin{aligned} \varepsilon_{i1}^* &= \varepsilon_{i1} - \theta z \\ \varepsilon_{i2}^* &= \varepsilon_{i2} - \theta z \end{aligned} \quad (5)$$

where  $\theta$  is a constant (the GLS transformation parameter) to be determined and  $z$  is a random variable to be determined.<sup>3</sup> If we could extract the individual effect from the original composite error terms, then it might be possible to eliminate the within unit autocorrelation. Equation (3) suggests two possible choices for  $z$ .

The first choice for  $z$  is taken directly from equation (3). Since  $u_i$  is the source of the problem, it might seem reasonable to set  $z = u_i$  in equations (5). Given that  $\theta$  is unknown, however, it is not immediately clear how this selection solves the problem. Fortunately,  $\theta$  is a free parameter. Thus, we may choose it in an appropriate way. Consider the restriction:

$$\text{Cov}(\epsilon_{i1}^*, \epsilon_{i2}^*) = E[(\epsilon_{i1} - \theta z)(\epsilon_{i2} - \theta z)] = 0 \quad (6)$$

where  $z = u_i$ . In the appendix, it is shown that the covariance restriction in equation (6) holds if

$$\theta^2 - 2\theta + 1 = 0 \quad (7)$$

An appropriate choice of  $\theta$  is given by the quadratic formula:  $\theta = 1$ . Since equation (3) holds for all  $t$ , plugging  $\theta = 1$  into equations (5) with  $z = u_i$  suggests that first differencing is one solution to the problem. The covariance restriction method easily reproduces the standard solution to the within autocorrelation problem in the benchmark case.

The second choice for  $z$  is also taken from equation (3). An alternative way of isolating the impact of within the unit factors is by averaging the error term over time. Thus, set  $z = u_i + w_i$  where  $w_i = \sum_{t=1}^T w_{it} / T$  and  $u_i = \sum_{t=1}^T u_{it} / T$ . Notice that  $\text{Var}(u_i + w_i) = \sigma_u^2 + (\sigma_w^2 / T)$ . Now, reconsider the covariance restriction in equation (6). Again,  $\theta$  is a free parameter. In this alternative case, it is shown in the appendix that the covariance restriction in equation (6) holds if

$$\left(\frac{\sigma_w^2}{T}\right)\theta^2 - 2\left(\frac{\sigma_w^2}{T}\right)\theta + \sigma_u^2 = 0 \quad (8)$$

where  $\sigma^2 = T\sigma_u^2 + \sigma_w^2$ . Of course, the quadratic formula yields two solutions for  $\theta$  in equation (8). They are

$$\theta_{\pm} = 1 \pm \frac{\sigma_w}{\sqrt{T\sigma_u^2 + \sigma_w^2}} \quad (9)$$

In applied work, the root  $\theta_-$  is preferred as its range is  $[0,1]$ . This property of  $\theta_-$  gives it a natural interpretation as a weighting parameter. Plugging  $\theta_-$  into

equations (5) with  $z = u_i + w_i$  suggests that generalized differencing is an alternative solution to the problem. In equations (5),  $\theta_-$  answers the question, How much weight should be placed on  $z$  in the generalized differencing procedure? Equation (9) indicates that the answer to this question depends on the variability of the individual effect,  $\sigma_u^2$  relative to the variability of unsystematic error,  $\sigma_w^2$ . The parameter  $\theta_-$  in equation (9) is the GLS transformation parameter that eliminates serial correlation in the error terms in panel data models.

The parameter  $\theta_-$  is central to understanding the relationship between the pooled OLS, random effects, and fixed effects estimators. As several of the authors cited in table 1 note, either estimator obtains depending on the value  $\theta_-$  takes on in its range. There are three cases:  $\theta_- = 0$ ,  $\theta_- = 1$ , and  $0 < \theta_- < 1$ . First,  $\theta_- = 0$  can occur only if  $\sigma_u^2 = 0$ . In equations (5),  $\theta_- = 0$  implies that no weight is placed on  $z$ . Thus, the transformed errors are the same as the composite errors. Since  $\text{Cov}(\epsilon_{i1}, \epsilon_{i2}) = \sigma_u^2 = 0$ , there is no autocorrelation and application of the pooled OLS estimator is appropriate. Second,  $\theta_- = 1$  can occur only if  $\sigma_w^2 = 0$ . In equations (5),  $\theta_- = 1$  implies that full weight is placed on  $z$ . Thus, the individual (fixed) effect, embedded in the composite errors and presumed to be correlated with the regressors, is totally removed from the transformed errors. This is what the fixed effect estimator does. Finally,  $0 < \theta_- < 1$  can occur when both  $\sigma_u^2 \neq 0$  and  $\sigma_w^2 \neq 0$ . In equations (5),  $0 < \theta_- < 1$  implies that partial weight is placed on  $z$ . This partial weighting combined with the additional assumption that the individual effect is uncorrelated with the regressors yields the random effects estimator.<sup>4</sup>

## V. Conclusion

The scalar-based covariance restriction method of deriving the GLS transformation parameter in the benchmark case is computationally direct and intuitive. It requires nothing more than knowledge of elementary statistics (mean, variance, and covariance) and the quadratic formula. Drawing on the intuition from a more familiar case, it relies on the principle of differencing the data in the search for an appropriate GLS estimator. Most importantly, however, it opens up the possibility that the pooled OLS, random effects, and fixed effects estimators can be understood at a deeper level by

researchers making their first approach to this important class of estimators.

## Appendix

**Derivation of Equation (7).** With  $z = u_i$ , the covariance of the transformed errors is

$$\begin{aligned} \text{Cov}(\varepsilon_{i1}^*, \varepsilon_{i2}^*) &= E[(\varepsilon_{i1} - \theta u_i)(\varepsilon_{i2} - \theta u_i)] \\ &= E[(u_i + w_{i1} - \theta u_i)(u_i + w_{i2} - \theta u_i)] \\ &= E[u_i^2 + u_i w_{i1} + u_i w_{i2} + \\ &\quad w_{i1} w_{i2} - \theta(2u_i^2 + u_i w_{i1} + u_i w_{i2}) + \\ &\quad \theta^2 u_i^2] \\ &= \sigma_u^2 - 2\sigma_u^2 \theta + \sigma_u^2 \theta^2 \end{aligned}$$

Applying the covariance restriction in equation (6) yields equation (7).

**Derivation of Equation (8).** Let  $z = u_i + w_i$ ,  $T = 2$ , and  $\sigma^2/T = \sigma_u^2 + (\sigma_w^2/T)$ .

The covariance of the transformed errors is

$$\begin{aligned} \text{Cov}(\varepsilon_{i1}^*, \varepsilon_{i2}^*) &= E[(\varepsilon_{i1} - \theta(u_i + w_i))(\varepsilon_{i2} - \theta(u_i + w_i))] \\ &= E[(u_i + w_{i1}) - \theta(u_i + w_i)][(u_i + w_{i2}) \\ &\quad - \theta(u_i + w_i)] \\ &= E[u_i^2 + u_i w_{i1} + u_i w_{i2} + w_{i1} w_{i2} - \\ &\quad \theta(2u_i^2 + 2w_i u_i + u_i w_{i1} + w_i w_{i1} + u_i w_{i2} \\ &\quad + w_i w_{i2}) + \theta^2(u_i^2 + 2w_i u_i + w_i^2)] \\ &= \sigma_u^2 - 2(\sigma_u^2 + \sigma_w^2/T)\theta + (\sigma_u^2 + \sigma_w^2/T)\theta^2 \\ &= \sigma_u^2 - 2\left(\frac{\sigma^2}{T}\right)\theta + \left(\frac{\sigma^2}{T}\right)\theta^2 \end{aligned}$$

Applying the covariance restriction in equation (6) yields equation (8).

## Notes

1. Alternative assumptions about the correlation between the errors and the regressors determine which estimator is under consideration.
2. Of course, special treatment of the first observation is required.
3. The ordering of the observations within the cross section unit is immaterial in the panel data setting. Thus, all of the observations should be given the same treatment, unlike the first order autocorrelation case, by virtue of symmetry.
4. In practice,  $\theta_1$  will typically lie in the interior of its range and the estimator employed will be determined by the explicit assumptions about the correlation between the errors and the regressors made by the researcher.

## References

- Ashenfelter, Orley, Phillip B. Levine, and David Zimmerman (2003): *Statistics and Econometrics: Methods and Applications*, John Wiley & Sons, Inc. New York.
- Greene, William H. (2003): *Econometric Analysis*, 5th edition, Prentice Hall, New Jersey.
- Griffiths, William E., R. Carter Hill, and George G. Judge (1993): *Learning and Practicing Econometrics*, John Wiley & Sons, Inc. New York.
- Gujarati, Damodar N. (2003): *Basic Econometrics*, 4th edition, McGraw-Hill Irwin, New York.
- Johnson, Jack and John Dinardo (1997): *Econometric Methods*, 4th edition, McGraw-Hill, New York.
- Judge, George G., Griffiths, W. E., R. Carter Hill, Helmut Lutkepohl, Tsoung-Chao Lee (1985): *The Theory and Practice of Econometrics*, 2nd edition, John Wiley & Sons, Inc. New York.
- Kmenta, Jan (1986): *Elements of Econometrics*, 2nd edition, Macmillan Publishing Company, New York.
- Pindyck, Robert S. and Daniel L. Rubinfeld (1998): *Econometric Models and Economic Forecasts*, 4th edition, McGraw-Hill Irwin, New York.
- Ramanathan, Ramu (2002): *Introductory Econometrics with Applications*, 5th edition, Harcourt College Publishers, Fort Worth.
- Ruud, Paul A. (2000): *An Introduction to Classical Econometric Theory*, Oxford University Press, New York.
- Wooldridge, Jeffrey M. (2003): *Introductory Econometrics: A Modern Approach*, 2nd edition, South-Western, Ohio.
- Wooldridge, Jeffrey M. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.