A NEW APPROACH TO ANALYZING QUASI-EXPERIMENTAL DATA: VALUE-ADDED ANALYSIS

Anthony S. Bryk, Huron Institute - Herbert I. Weisberg, Contract Research Corporation

## I. Introduction

In a true educational experiment, the experimental units (e.g., individuals, classes, schools, pre-school centers) are randomly assigned to the different treatments, or programs, under study. As the number of units per treatment group increases, the pre-experimental mean differences among the treatment groups on any background variable tend to become small. As a result of this "natural equation of groups," differences among the post-treatment outcomes can, with reasonable assurance, be attributed to treatment effects.

For a variety of political, practical, and ethical reasons, however, randomized experiments are rarely feasible in educational settings. Recently, (Cohen, 1973) even the desirability of such experiments has been called into question. These practical and theoretical considerations have led to the implementation of what Campbell and Stanley (1963) have referred to as quasi-experiments. In these designs, the experimental units are not randomly assigned to treatments. As a result, the pre-treatment equation of groups is not assured. Observed differences among the post-treatment outcomes are attributable to pre-treatment differences in addition to the effects of the treatments. We must now resort to statistical techniques to adjust away the relevant pre-experiment differences among the treatment groups. Because of the assumptions required by these techniques and certain artifacts involved in the estimation of effects, interpretation of the results of these analyses requires extreme care. Some authors (Lord, 1967; Campbell and Erlebacher, 1970) are quite pessimistic about the ability of quasi-experiments to yield any useful and valid inferences.

In this paper, we introduce the value-added strategy as an alternative approach to the analysis of data from educational quasi-experiments. This technique was developed in response to shortcomings inherent in currently existing adjustment strategies, such as analysis of covariance, matching, standardization (Wiley, 1970), gain scores, and analysis of residuals.

## II. An Alternative Strategy - Value-Added Analysis

The basic idea of the value-added analysis is to estimate for each subject in each experimental program the post-test score he would have obtained had he not been in any experimental program. Comparing this artificially-constructed post-test score with the actual pre- and post-test scores, we can estimate how much of his growth is the result of "natural" maturation, and how much is the value-added by the program in which he was enrolled. Smith (1973) used this approach to estimate overall effects of Head Start. We have refined his method and provided more theoretical underpinning.

### A. Theory

Let $Y_{ij}$ and $Y'_{ij}$ represent the observed pre- and post-test scores on some measure for individual i in treatment, or program, group j. Let $T_{ij}$ and $T'_{ij}$ be the corresponding true scores. We assume the classical test theory model that

$$Y_{ij} = T_{ij} + e_{ij} \tag{1}$$

$$Y'_{ij} = T'_{ij} + e'_{ij} \quad \text{for } i=1\ldots n_j,\ j=1\ldots J \tag{2}$$

where $E(e_{ij}|T_{ij})=E(e'_{ij}|T'_{ij})=E(e_{ij}e'_{ij})=0$ and $\mathrm{Var}\,(e_{ij}|T_{ij}) = \mathrm{Var}\,(e'_{ij}|T'_{ij}) = \sigma^2_e$.

Let us define:

$a_{ij}$=age of individual i in group j at pre-test

$a'_{ij}$=age of individual i in group j at post-test

$M_{ij}$=component of true scores, $T_{ij}$ and $T'_{ij}$, representable as a linear function of measurable covariates other than age

$U_{ij}$=nonmeasurable component of the true scores, $T_{ij}$ and $T'_{ij}$, which is independent of both age and other measurable covariates.

For any specific individual i in group j, $a_{ij}$, $a'_{ij}$, $M_{ij}$, and $U_{ij}$ are considered fixed. They do vary, however, over subjects in the population of individuals from which the sample for treatment group j was chosen. Let us define:

$$E(U_{ij}) = U_j \text{ and } \mathrm{Var}\,(U_{ij}) = \sigma^2_U \tag{3}$$

Further, since $U_{ij}$ is independent of age and the other measurable covariates, it follows that

$$\mathrm{Cov}(U_{ij},a_{ij}) = \mathrm{Cov}(U_{ij},a'_{ij}) = \mathrm{Cov}(U_{ij},M_{ij})=0 \tag{4}$$

For the simplest case, we assume that growth in the domain under study is a known linear function of age, measurable covariates, and nonmeasurable covariates. We can, then, represent the true score for subject i in group j at the pre-test as:

$$T_{ij} = \mu + \beta a_{ij} + M_{ij} + U_{ij} \tag{5}$$

Further, if there is no intervention between the pre-test and post-test - i.e., if only natural maturation is operational - then we can represent the post-test score as:

$$T'_{ij} = \mu + \beta a'_{ij} + M_{ij} + U_{ij} \tag{6}$$

Let us define a variable $\Delta_{ij}$ where there is no intervention between pre- and post-test:

$$\Delta_{ij} = T'_{ij} - T_{ij} \tag{7}$$

Clearly, $\Delta_{ij}$ represents the growth increment for subject i in group j which is attributable solely to natural maturation. Specifically, in terms of the growth model represented by equations (5) and (6),

$$\Delta_{ij} = \beta\,(a'_{ij} - a_{ij}) \tag{8}$$

In general, we could make different assumptions about the nature of growth in the domain under study than were made for our simple case. For example, one could hypothesize a growth model involving interaction terms or non-linear terms in $a_{ij}$. In this case, we would replace the mathematical model of equation (5) with a more complex model. As a result, the expression for $\Delta_{ij}$ would also become more complex. The basic derivation presented here, however, would remain unchanged.

Let us now examine the alternative situation where an intervention has occurred between the pre- and post-tests. If we assume that the effect of the intervention is to increment uniformly the growth of all children in program j, then

$$T'_{ij} = \mu + \beta a'_{ij} + M_{ij} + U_{ij} + V_j \tag{9}$$

where $V_j$ represents the incremental effect of experimental program j, or what we have termed the value-added by program j. From equations (5) and (9), we see that

$$V_j = (T'_{ij} - T_{ij}) - \Delta_{ij} \tag{10}$$

Thus, the value-added by program j, $V_j$, can be interpreted as the true difference (gain or change) score between pre- and post-test adjusted for the growth increment which is expected on the basis of natural maturation. Alternatively, $T_{ij} + \Delta_{ij}$ may be interpreted as the post-test score predicted solely on the basis of natural maturation. From this point of view, $T'_{ij} - (T_{ij} + \Delta_{ij})$ is a "residual score;" i.e., the observed post-test minus a predicted post-test.

### B. Application to the Educational Quasi-Experiment

Our goal in the value-added analysis is to estimate for each program a value-added, $\hat{V}_j$, which is a measure of the absolute program effect over and above what one might expect on the basis of natural maturation. Our model assumes that the true pre- and post-test scores can be represented as a function of age, measured covariates, and unmeasured covariates. The model also assumes that the function is known. We now consider what can be done when each of two important assumptions (true scores known; growth model known) in this idealized situation is lifted. This will lead us to a way of implementing the value-added approach in practice.

#### 1. True Scores Known: First Approach

The first problem that we encounter in applying this model is that the true scores $T_{ij}$ and $T'_{ij}$ are rarely known. The first and obvious approach is to substitute the observed pre- and post-test scores, $Y_{ij}$ and $Y'_{ij}$, in equations (5) and (9). For the pre-test, this yields

$$Y_{ij} = \mu + \beta a_{ij} + M_{ij} + V_{ij} + e_{ij} \tag{11}$$

Similarly, for the post-test,

$$Y'_{ij} = \mu + \beta a'_{ij} + M_{ij} + U_{ij} + V_j + e'_{ij} \tag{12}$$

It follows from these last two equations that

$$Y'_{ij} - (Y_{ij} + \Delta_{ij}) = V_j + e'_{ij} - e_{ij} \tag{13}$$

we may now define, $V_{ij1}$, an individual value-added for subject i in program j, where

$$V_{ij1} = V_j + e'_{ij} - e_{ij} \tag{14}$$

For any randomly selected individual i in treatment group j,

$$E\,(V_{ij1}) = V_j \text{ and } \mathrm{Var}\,(V_{ij1}) = 2\sigma^2_e \tag{15}$$

Thus, each subject provides an unbiased estimate of $V_j$ with variance $2\sigma^2_e$.

We can obtain a more efficient estimate of $V_j$ by pooling these individual estimates to obtain an average across the $n_j$ subjects in treatment group j. This approach yields

$$V_{\cdot j1} = \sum_{i=1}^{n_j} \frac{V_{ij1}}{n_j} \tag{16}$$

where, for any treatment group j,

$$E(V_{\cdot j1}) = V_j \text{ and } \mathrm{Var}\,(V_{\cdot j1}) = \frac{2\sigma^2_e}{n_j} \tag{17}$$

Like our individual subject estimate, our pooled estimate, $V_{\cdot ij1}$, is unbiased, but the variance is now reduced to $2\sigma^2_e/n_j$. Thus, the direct substitution of the observed pre- and post-test scores for the unknown true scores provides us with one approach for estimating the incremental effect of our experimental program.

#### 2. True Scores Unknown: Second Approach

While the above approach is intuitively appealing, we might ask whether we can obtain a better estimate of $V_j$ by some alternative approach. In particular, since the observed pre- and post-tests are measured with error, perhaps the substitution of an estimated or predicted true score might yield a more efficient estimate. Several approaches for estimating true scores have been reviewed by Cronbach and Furby (1970). Their "complete estimator" has intuitive appeal, and it merits further investigation. This estimator is quite complex, however, and its usefulness in the value-added approach is an unresolved question which is under further investigation.

A simple and natural spin-off of this estimated true score approach would be the use of a predicted pre-test score, $\hat{Y}_{ij}$, generated from the observable components of equation (5). Simply,

$$\hat{Y}_{ij} = \mu + \beta a_{ij} + M_{ij} \tag{18}$$

$\hat{Y}_{ij}$ may be viewed as an alternative estimate for $T_{ij}$. In terms of equation (5),

$$T_{ij} = \hat{Y}_{ij} + U_{ij} \tag{19}$$

As a second approach to dealing with the unknown true scores, we substitute the observed post-test score $Y'_{ij}$ and the predicted pre-test score $\hat{Y}_{ij*}$ into equation (10) for the unknown true scores. This yields,

$$Y'_{ij} - (\hat{Y}_{ij} + \Delta_{ij}) = V_j + U_{ij} + e'_{ij} \tag{20}$$

We now have a second individual subject estimator, $V_{ij2}$, of the value-added for subject i in program j, where

$$V_{ij2} = V_j + U_{ij} + e'_{ij} \tag{21}$$

For any randomly selected individual i in treatment group j,

$$E(V_{ij2})=V_j+U_j \text{ and } Var(V_{ij2})=\sigma^2_U + \sigma^2_e \tag{22}$$

Thus, each subject provides a biased estimate of $V_j$ (bias = $U_j$) with variance $\sigma^2_U + \sigma^2_e$.

Pooling these individual estimates across the $n_j$ subjects in treatment group j yields,

$$V_{\cdot j2} = \sum_{i=1}^{n_j} \frac{V_{ij2}}{n_j} \tag{23}$$

where, for any treatment group j,

$$E(V_{\cdot j2})=V_j+U_j \text{ and } Var(V_{\cdot j2})=\frac{\sigma^2_U + \sigma^2_e}{n_j} \tag{24}$$

Thus, the pooled estimate, $V_{\cdot j2}$, is also biased, but it has reduced variance as compared to the individual subject estimates.

### 3. Comparing the Two Estimates

If we compare equations (17) and (24), we see that neither estimate of $V_j$, $V_{\cdot j1}$, or $V_{\cdot j2}$ is clearly superior. $V_{\cdot j1}$ is unbiased, while $V_{\cdot j2}$ is in general biased, except when $U_j=0$; $j=1\ldots J$. On the other hand, $V_{\cdot j1}$ will have larger variance if $\sigma^2_e > \sigma^2_U$. The mean square error (MSE), which combines both of these criteria, is a useful measure of the accuracy of an estimator. The MSE equals the sum of the variance and the bias squared. For our two estimators,

$$MSE\ (V_{j1}) = \frac{2\sigma^2_e}{n_j} \tag{25}$$

and

$$MSE\ (V_{\cdot j2}) = \frac{\sigma^2_U + \sigma^2_e}{n_j} + (U_j)^2 \tag{26}$$

Thus, in general, $V_{\cdot j1}$ is superior if

$$\sigma^2_e < \sigma^2_U + n_j\ (U_j)^2 \tag{27}$$

As we defined $U_j$ in our simple growth model of equation (5), it can be interpreted as the expected residual pre-test score for individuals in group j over and above that which is linearly related to all measured variables. Since the model includes a constant term $\mu$, we would expect $U_j$ to differ from 0 only if group j differs from the other groups in ways unrelated to measured covariates. If the covariates are selected carefully, however, it is unlikely that any $U_j$ will differ substantially from 0. Of course, if subjects are randomly assigned to treatment groups, there should not exist a priori differences between treatment groups in the expected values on any variables. As a result, $U_j$ will equal 0.

### 4. A Combined Estimate

If $V_{\cdot j1}$ and $V_{\cdot j2}$ were independent, then an appropriate linear combination of these two estimates would have a smaller mean square error than either one individually. Although $V_{\cdot j1}$ and $V_{\cdot j2}$ are clearly not independent, we may still realize a gain in accuracy by creating such a combined estimate. This estimate would take the following general form:

$$V_{\cdot jc} = w\ V_{\cdot j1} + (1 - w)\ V_{\cdot j2} \quad 0 \le w \le 1 \tag{28}$$

It would clearly be desirable to choose w so as to minimize the mean square error of $V_{\cdot jc}$. Note that

$$E(V_{\cdot jc})=wV_j + (1-w)(V_j + U_j)=V_j + (1-w)U_j \tag{29}$$

which indicates that the combined estimate, $V_{\cdot jc}$, has bias $(1 - w)U_j$. Also,

$$Var(V_{\cdot jc})=w^2\ \frac{(2\sigma^2_e)}{n_j} + (1-w)^2\ \frac{(\sigma^2_U+\sigma^2_e)}{n_j} + 2w(1-w)\ Cov(V_{\cdot j1},V_{\cdot j2}). \tag{30}$$

For any treatment group j, if we assume that the $Cov(V_{ij1},V_{i*j2}) = 0$ for $i \neq i^*$ then

$$Cov(V_{\cdot j1},V_{\cdot j2}) = \sum_{i=1}^{n_j} Cov\left(\frac{V_{ij1}}{n_j},\frac{V_{ij2}}{n_j}\right) \tag{31}$$

It follows that

$$Var(V_{\cdot jc}) = (1+w^2)\ \frac{\sigma^2_e}{n_j} + (1-w)^2\ \frac{\sigma^2_U}{n_j} \tag{32}$$

Thus, we can now write out the mean square error for $V_{\cdot jc}$:

$$MSE(V_{\cdot jc})=(1-w^2)\ \frac{\sigma^2_e}{n_j} + (1-w)^2 \left[\frac{\sigma^2_U}{n_j} + U_j^2\right] \tag{33}$$

If we minimize this with respect to w, we find

$$w_{opt} = \frac{\sigma^2_U + n_j (U_j)^2}{\sigma^2_U + n_j (U_j)^2 + \sigma^2_e} \quad (34)$$

If, as the result of randomization or by chance $U_j = 0$, the expression for $w_{opt}$ simplifies to:

$$w_{opt} = \frac{\sigma^2_U}{\sigma^2_U + \sigma^2_e} \quad (35)$$

Examination of this expression suggests that $w_{opt}$ can be interpreted in this case as a "residual reliability" of the test after variation related to age and other measurable background variables has been removed. Thus, the weight to be placed on method 1, which uses the observed pre-test score, is simply the residual reliability of this score.

Finally, up to this point, we have assumed that the values of $\sigma^2_U$, $\sigma^2_e$, and $U_j$ are known. In most quasi-experimental settings, however, these values are unknown. Thus, $w_{opt}$ must be estimated from the data. Two alternative procedures for estimating $w_{opt}$ are presented in the Appendix.

5. Growth Model Unknown

To apply the value-added approach in the analysis of an educational quasi-experiment, we need a growth model which accurately describes the process of "natural" maturation. The second major assumption in our theoretical model which we now lift is that the natural growth model is known.

In this situation, we must develop a growth model and estimate the parameters of the model from either the quasi-experimental data or some alternative data set. This is a complex problem because the form of the growth model depends both upon the particular phenomenon under study and the conditions in the setting which bound this investigation. As a result, no single approach to the construction of such models is likely to be uniformly successful. We present here an approach developed by Smith (1973) and Weisberg (1973) which may prove useful in certain circumstances.

The key assumption in this approach to estimating the growth model is that the variation displayed in pre-test scores reflects developmental trends which can be directly related to age and other background variables. The heart of this approach involves a causal linkage of the variation in age and the background variables to the variation in pre-test scores. In particular, if the background variables are held constant, then the variation in pre-test scores may be expected to reflect only differences in the length of time exposed to the natural learning environment, or what we term age. More specifically, suppose we could look at a sub-sample of children with identical values on all measured background variables except age. Suppose then that we observe the mean score for such individuals as a function of age. The resulting curve based on this cross-sectional data is an approximation to the longitudinal growth curve that these children would actually display as they grow older.

This approach of attempting longitudinal inference from cross-sectional data is a general and well-known strategy. Kodlin and Thompson (1958) have considered the limitations of this approach in some detail. Under the conditions of a "stable universe" - i.e., a stable growth process - and a stable population across age levels, the cross-sectional approach can be used in place of the longitudinal for the estimation of mean growth. Thus, in settings where: 1) there are no significant external influences - other than the experimental intervention - to disturb the natural growth process; and 2) the pre-test sample is selected effectively at random with respect to age and the background variables; the cross-sectional approach should provide an excellent approximation to the natural growth curves.

## VI. Summary and Conclusions

The basic idea of the value-added analysis is quite intuitive. We develop a growth model to predict normal growth in the absence of an experimental treatment. By combining this expected outcome with the observed outcome, we estimate the program effect - the growth increment over and above natural maturation.

As one approach to estimating the natural growth model, we can assume a causal linkage between the variation in age and background variables, and the variation in pre-test scores. If the research setting is stable, variation as a function of age will be attributable solely to growth, and not to other differences among age cohorts. In such a situation, we can estimate the effect of natural maturation for a treatment group, and the program effect over and above this natural growth.

A particularly perplexing problem for traditional techniques, such as ANCOVA, is the biased and inconsistent estimation of program effects resulting from measurement errors in the covariates (e.g., pre-test scores). Under the value-added model, however, measurement error in the pre-test will not bias estimates of the program effects.

Further, we believe that the concept of the value-added effect is more meaningful than the adjusted treatment mean differences in ANCOVA. The value-added is an absolute measure of program effect. Although control group data is useful as a check on the fitted mathematical model, it is not mandatory. Under the strong assumption that the growth model is correct, the value-added technique generates a statistical control group.

In a pre-school setting such as Head Start, for example, child's age from birth seems suitable to use in the growth model. In certain situations, our model may be applicable with other measures of "age." For example, suppose we were comparing a traditional and an experimental high school foreign language program. A logical choice here for "age at pre-test" might be the length of time studying the language via the traditional approach.

Although we feel that the value-added approach is quite general, two factors cause difficulties in applying it in a school context. First, much testing in schools uses a standardized metric. Standardized tests do not allow measurement of absolute growth, only "relative standing." Our model in its present form would not be suitable for such a metric. Second, most schools have summer vacations, so that growth is not a strictly monotone function of chronological age or length of exposure to schooling. We need to investigate thoroughly the effects of this summer discontinuity. The development of complex growth models incorporating such summer effects would make the value-added approach feasible.

Although we hope that the methods described in this paper will provide analysts with a useful alternative to traditional adjustment strategies, we view this work as only a beginning. There are many possible extensions and refinements. For example, we need to develop realistic models to represent growth in educational settings, and practical ways of estimating model parameters. A better understanding of the sampling theory associated with our methods is needed, so that significance tests and confidence intervals can be obtained. Also, the development of quasi-experimental designs which facilitate this type of analysis should be pursued. Lastly, a critical problem with traditional adjustment strategies is that they implicitly embody a static model which conceptualizes a program effect as a constant increment to a static base. Educational programs, on the other hand, are usually attempts to alter some developmental growth process. An intervention is typically a dynamic change in an on-going process. We view the value-added approach as a tentative effort to operationalize this conceptual approach in the analysis of quasi-experiments.

---

*The use of a predicted post-test score might also be considered here. This is problematic because treatment effects must be included in the predicted post-test. In reality, there are numerous possible predicted or estimated true scores, such as the Cronbach and Furby estimates, which could be utilized in the value-added setting. Because of the introductory nature of this paper, we consider here only two simple approaches.

---

REFERENCES:

Campbell, D.T. and Erlebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), Compensatory Education: A National Debate. Vol. III. Disadvantaged Child. New York: Brunner-Mazel, 1970.

Campbell, D.T. and Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Cochran, W.G. Errors of measurement is statistics. Technometrics. 1968, 10, 637-660.

Cohen, D. Social experiments with schools: what has been learned? to be published in Planned Variation Experiments: Should We Give Up or Try Harder?, A Rivlin (Ed.), Brookings Institute.

Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968-70, 426-443.

Kodlin, D. and Thompson, D.J. An appraisal of the longitudinal approach to studies of growth and development. Monographs of the Society for Research in Child Development, 1958, 23 (No. 67).

Lord, F.M. A paradox in the interpretation of group comparisons. Psychological Bulletin, 1967-68, 304-405.

Lord, F.M. and Novick, M.R. Statistical Theories of Mental Test Scores. Reading, Massachusetts: Addison-Wesley, 1968.

Smith, M.S. and Bissell, J.S. Report analysis: the impact of Head Start. Harvard Educational Review, 1970, 40, 51-104.

Weisberg, H.I. Short term cognitive effects of Head Start programs: a report of the third year of planned variation - 1971-1972. Cambridge, Massachusetts: Huron Institute, 1973.

Wiley, D.E. Approximations to ceteris paribus: Data adjustment in educational research. Paper presented at International Research Seminar on Learning and the Educational Process, Munich, Germany, 1971.

---

## APPENDIX

In this Appendix, we present two methods of estimating the optimal weight w in our combined estimator $V_{.jc}$. Both of these methods assume that $U_j = 0$ for all j, so that

$$w_{opt} = \frac{\sigma^2_U}{\sigma^2_U + \sigma^2_e}.$$

It is this quantity we wish to estimate. If $U_j$ is substantially different from 0 for some j's, $w_{opt}$ may differ substantially from the true optimum given by equation (34). If measured covariates are selected judiciously, this is unlikely to occur.

### Estimator #1

Suppose the reliability ρ of our outcome measure Y (and Y') is known. Then,

$$\rho = \frac{V(T_{ij})}{V(Y_{ij})} = \frac{V(\alpha + \beta a_{ij} + M_{ij}) + \alpha^2_U}{V(Y_{ij})}$$

Let

$$\frac{V(\alpha + \beta a_{ij} + M_{ij})}{V(Y_{ij})} \equiv \eta^2$$

Then

$$\rho = \eta^2 + \frac{\sigma^2_U}{V(Y_{ij})}$$

$$\frac{\rho - \eta^2}{1 - \eta^2} = \frac{\sigma^2_U}{V(Y_{ij})(1-\eta^2)} = \frac{\sigma^2_U}{\sigma^2_U + \sigma^2_e} = w_{opt}$$

From the regression analyses used to produce our model, we obtain an estimate $R^2$ of $\eta^2$. A natural estimator of $w_{opt}$ is thus given by

$$w_{opt} = \frac{\rho - R^2}{1 - R^2}$$

Estimator #2

Our second estimator does not require independent information about $\rho$. From equations (29) and (32), we see that for any individual i in group j, assuming $U_j = 0$:

$$E(V_{ijc}) = V_j$$

$$Var(V_{ijc}) = (1+w^2)\,\sigma^2_e + (1-w^2)\,\sigma^2_U$$

Ignoring sampling variation in the regression coefficients of our predictor equation, the $V_{ijc}$'s (for any value of w) are independent random variables. Thus, if we perform a one-way analysis of variance, using these as outcomes, the mean square error term provides an unbiased estimate, $\hat{Var}(V_{ijc})$ of $Var(V_{ijc})$. Since

$$Var(V_{\cdot ij}) = \frac{Var(V_{ijc})}{n_j}$$

$w_{opt}$ is the value of w which minimizes $Var(V_{ijc})$. It seems reasonable, then, to estimate $w_{opt}$ by the value of w which minimizes $\hat{Var}(V_{ijc})$. Let

$$SS_{11} = \sum_{j}\sum_{i}(V_{ij1} - V_{\cdot j1})^2$$

$$SS_{22} = \sum_{j}\sum_{i}(V_{ij2} - V_{\cdot j2})^2$$

$$SS_{12} = \sum_{j}\sum_{i}(V_{ij1} - V_{\cdot j1})(V_{ij2} - V_{\cdot j2})$$

Then the mean square error for the one-way ANCOVA is easily shown to be

$$MSE = w^2 SS_{11} + (1-w)^2\, SS_{22} + 2w(1-w)\, SS_{12}$$

Minimizing this with respect to w, we find

$$w_{opt} = \frac{SS_{22} - SS_{12}}{SS_{22} + SS_{11} - 2SS_{12}} .$$