Perfect Multicollinearity and Dummy Variable Trap: Explaining the Unexplained

Pillai N., Vijayamohanan and A., Rju Mohan

Gulati Institute of Finance and Taxation, Trivandrum, Kerala, India,
Gulati Institute of Finance and Taxation, Trivandrum, Kerala, India

March 2024

Perfect Multicollinearity and Dummy Variable Trap:

Explaining the Unexplained

Vijayamohanan Pillai N.

Rju Mohan A.

Gulati Institute of Finance and Taxation, Trivandrum, Kerala, India

March 2024

Perfect Multicollinearity and Dummy Variable Trap: Explaining the Unexplained

Vijayamohanan Pillai N.

Rju Mohan A.

Abstract

Recently we have come across some confused references to 'dummy variable trap' (DVT) during an Econometrics workshop organized at a University in Kerala, India. A google search has generated a large number of so-called 'machine learning'-based tutorials of the very same content. In addition to these internet sources of such insufficient/incorrect information, a number of (new generation) econometrics text books also have unfortunately been found to cater to such confusions. The confusion arises from the inadequate care in discussion by some textbook authors that spreads through the mass of new generation half-wit tutorial bloggers and other media, who further venture to *simplify* it, and finally grips the careless novices, who get lured by the 'simple logic' of it. Unfortunately, they choose to ignore the authoritative text books as well as the need for an approach of mathematical logic. Note that these text books are also often insufficient to bring to light the concepts clearly. Hence this paper seeks to explain this issue in the framework of its mathematical logic.

Perfect Multicollinearity and Dummy Variable Trap:

Explaining the Unexplained

Vijayamohanan Pillai N.

Rju Mohan A.

1. Introduction

Recently we have come across some confused references to 'dummy variable trap' (DVT) during an Econometrics workshop organized at a University in Kerala, India; some self-styled pundits of little wit tried to explain DVT with reference to the dummy variables only, disregarding the intercept. According to them, for example in the case of a categorical variable like 'gender', "Including both the dummy variable[s] can cause redundancy because if a person is not male in such case that person is a female, hence, we don't need to use both the variables in regression models."[1] "In this case ['male' and 'female'] are perfectly correlated and have a correlation coefficient of -1."[2] "These two dummy variables are multi-collinear".[3] A google search has generated a large number of so-called 'machine learning'-based tutorials of the very same content. In addition to these internet sources of such insufficient/incorrect information, a number of new generation econometrics text books also have unfortunately been found to cater to such confusions, for example, Panchanan Das (2019); A. H. Studenmund (2017):

> "If we incorporate three dummies for three categories, we cannot estimate the model uniquely because the set of three dummy variables is perfectly collinear [*in the presence of the intercept*; the author should have explicitly included this also]. This is known as dummy variable trap." (Das 2019: 159). Again, "If we incorporate *p* number of dummies for *p* categories, we cannot estimate the model uniquely because the set of three [not *three*, but *p*] dummy variables is perfectly collinear [*in the presence of the intercept*; the author should have explicitly included this also]. This is known as dummy variable trap." (Das 2019: 166).

> "Note that in this example only one dummy variable is used even though there were two conditions. This is because one fewer dummy variable is constructed than conditions. The event not explicitly represented by a dummy variable, the omitted condition, forms the basis against which the included conditions are compared [*the*

[1] https://www.geeksforgeeks.org/ml-dummy-variable-trap-in-regression-models/.

[2] https://www.statology.org/dummy-variable-trap/.

[3] https://www.learndatasci.com/glossary/dummy-variable-trap/.

> *omitted condition in turn being represented by the intercept*; the author should have explicitly included this also]. Thus, for dual situations only one dummy variable is entered as an independent variable; the coefficient is interpreted as the [*marginal*] effect of the included condition relative to the omitted condition. Be careful never to use two dummy variables to describe the two conditions. If you were to make this mistake, sometimes called a dummy variable trap, you'd have perfect multicollinearity." (Studenmund 2017: 80-81).

Strangely, insufficient/incorrect information marks the initial discussion by even such an econometrician as Wooldridge in his famous textbook: "Using two dummy variables would introduce perfect collinearity because *female* + *male* = 1, which means that *male* is a perfect linear function of *female*. Including dummy variables for both genders is the simplest example of the so-called **dummy variable trap**, which arises when too many dummy variables describe a given number of groups" [in the presence of the intercept; the author should have explicitly included this also]. (Jeffrey M. Wooldridge 2020: 222-223; emphasis as in the original); however, later on, he points out the significance of the presence of intercept (on the same page 223): "Some researchers prefer to drop the overall intercept in the model and to include dummy variables for each group. …. There is no dummy variable trap in this case because we do not have an overall intercept."

The confusion arises from the inadequate care in discussion by some textbook authors that spreads through the mass of new generation half-wit tutorial bloggers and other media, who further venture to *simplify* it, and finally grips the careless novices, who get lured by the 'simple logic' of it. Unfortunately, they choose to ignore the authoritative text books as well as the need for an approach of mathematical logic. Note that most of the text books are also often insufficient to bring to light the concepts clearly. Hence this paper seeks to explain this issue in the framework of its mathematical logic. In the next section we explain dummy variable trap in terms of its mathematical logic, both empirical and theoretical and Section 3 concludes the study.

2. Dummy Variable Trap Explained: Theoretical and Practical Substantiation

In regression analysis, a dummy variable or an indicator variable (or simply a dummy) is one with a binary value (one or zero) to indicate the presence or absence of a category that we may expect to have some effect on the dependent variable. For example, in gender discrimination models with, say, agricultural wages as dependent variable, we use a dummy variable to represent the categorical (nominal) variable of gender, with the dummy variable taking a value of one (1) for female workers and zero (0) for male workers (or vice versa). Note that since the nominal variable gender has two categories here (female and male), we can also have two dummy variables, one for each category, where one dummy takes a value of one (1) for female workers and zero (0) other wise and the other dummy takes a value of one (1) for male workers and zero (0) otherwise. Thus we have two cases: for a single categorical variable with two categories, we can have a model with (i) one dummy variable or (ii) two dummy variables.

Dummy variable trap is basically a case of perfect multicollinearity. So, what is perfect (exact) multicollinearity, and what is its consequence for estimation?

2.1 Perfect Multicollinearity

One of the basic assumptions of the ordinary least squares (general linear) model is that the data matrix X, which is of order n × k, has full column rank k, where k denotes the number of parameters to be estimated (number of columns in the data matrix), and n is the number of observations (number of rows of the data matrix), that is, that there is no linear dependence among the explanatory variables. In this case, $\mathrm{X'X}$, the matrix of cross-products of the data matrix X, is of dimension k × k, and is square, symmetric, and of full rank k. The reason for this assumption is that the least-squares estimator $\hat{\beta} = (\mathrm{X'X})^{-1}\mathrm{X'Y}$ requires the inversion of $(\mathrm{X'X})$, which is impossible if the rank of X and hence the rank of $\mathrm{X'X}$, is less than k. In a more general context, if any two columns of X are linearly dependent, then rank $(\mathrm{X'X}) < k$, and $\hat{\beta} = (\mathrm{X'X})^{-1}\mathrm{X'Y}$ becomes incalculable. This is because $(\mathrm{X'X})^{-1}$ involves division by the determinant of $\mathrm{X'X}$, that is, $(\mathrm{X'X})^{-1} = \frac{\text{Adjoint (X'X)}}{|\mathrm{X'X}|}$, where $|\mathrm{X'X}|$ refers to the determinant. If the determinant of $\mathrm{X'X}$ happens to be zero, that is, if the matrix $\mathrm{X'X}$ is singular, then inversion of $(\mathrm{X'X})$ becomes impossible; and this can happen when any two or more columns of X are linearly dependent, that is, rank $(\mathrm{X'X}) < k$. This situation is called perfect multicollinearity, which occurs when some or all of the explanatory variables are perfectly collinear. Thus, essentially, perfect multicollinearity may be evidenced from a singular $\mathrm{X'X}$ matrix, or a singular data matrix, X, itself, if it is a square matrix.

Perfect multicollinearity in the data matrix X appears in the form of linear dependency among the variables (columns), which in turn occurs in linear combination of the variables. Linear combination can appear in different ways.

Example 1: For a simple empirical example, let us consider a 2x2 data matrix $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$; the number of columns in the data matrix (the number of parameters to be estimated) k = 2, and let the first column (C_1) represent variable X_1 and the second column (C_2), variable X_2. The determinant |X| = (1·4 – 2·3) = –2; the data matrix is non-singular, it has a rank of 2 = k; and there is no perfect multicollinearity. Alternatively, here,

$$\mathrm{X'X} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix},$$

with a determinant of 4 (= 10·20 – 14·14); the $\mathrm{X'X}$ matrix is non-singular, it has a rank of 2 = k; there is no linear dependence, no perfect multicollinearity, and the parameters are estimable.

Example 2: Now consider another 2x2 data matrix $X = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$; the determinant |X| = (1·4 – 2·2) = 0; the data matrix is singular, it has a rank of 1 < k; the problem here is that the second column (row) is two-times the first column (row): $C_2 = 2C_1$ or $R_2 = 2R_1$ (that is, $X_2 = 2X_1$), and this is perfect multicollinearity. Alternatively, here,

$$X'X = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix},$$

with a zero determinant; the $X'X$ matrix is singular, it has a rank of 1 < k; and there is perfect multicollinearity and the estimation breaks down.

In the above example (2), also note that there is perfect positive correlation between the two columns (variables): $r(X_1, X_2) = 1$; *but,* so is the case with the first example also! With $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, we have $C_2 = 1 + C_1$ or $R_2 = 2 + R_1$; (that is, $X_2 = 1 + X_1$), and there is perfect positive correlation between the two columns (variables): $r(X_1, X_2) = 1$! Yet, there is no perfect multicollinearity in this case! That is, $X'X$ is invertible and the estimates are obtainable here! It simply means that perfect correlation between two variables need not result in perfect multicollinearity between them.

Mathematically, the first case ($C_2 = 2C_1$) relates to a linear function, such as $cX_2 = bX_1$, where b, c ≠ 0, (in our example, b = 2, and c = 1), and the second case ($C_2 = 1 + C_1$) to an affine function, a linear-plus-constant function, such as $cX_2 = a + bX_1$, where a, b, c ≠ 0, (in our example, a = b = c = 1). Now the question arises: Why in the second (affine function) case, there is no perfect multicollinearity here?

Note that usually a regression equation represents an affine function, that is with a slope and a constant (intercept). In this specification, the X data matrix now includes a column of ones (units) to account for the intercept. That is, in our first example, the data matrix now appears as $X = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \end{bmatrix}$, where the first column represents the intercept. Now,

$$X'X = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 & 2 \\ 1 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 10 & 14 \\ 6 & 14 & 20 \end{bmatrix},$$

and the determinant [(2·10·20) + (4·14·6) + (4·14·6) – (6·10·6) – (4·4·20) – (14·14·2)] of this matrix is zero, because there is perfect multicollinearity (linear dependency among the three columns) in the data matrix X (as well as in $X'X$): Intercept = $X_2 - X_1$, or X_2 = Intercept + X_1. Since $X'X$ is singular, least squares estimation fails. But note that in the absence of the intercept in the data matrix, there was no problem at all, despite perfect positive correlation between the two variables! However, it does not suggest that the presence of the intercept always presents this problem. For example, consider the data matrix

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \\ 1 & 5 & 6 \end{bmatrix},$$

a non-singular matrix with a determinant [(1·3·6) + (1·4·1) + (1·5·1) – (1·3·1) – (1·1·6) – (4·5·1)] of –2. The corresponding cross-product matrix is

$$X'X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 5 \\ 1 & 4 & 6 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 4 \\ 1 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 11 \\ 9 & 35 & 43 \\ 11 & 43 & 53 \end{bmatrix},$$

and the determinant of this matrix is 4; the matrix is non-singular, and the parameters are estimable, despite close to unit correlation between the second and third columns of X (that is, between X_1 and X_2 variables: $r_{12} = 0.9934$).

Thus, the upshot of this empirical exercise is very simple: when X_1 and X_2 variables are in linear function (such as $cX_2 = bX_1$, c, b ≠ 0), perfect multicollinearity appears, irrespective of the presence/absence of an intercept. On the other hand, when X_1 and X_2 variables are in affine function (such as $cX_2 = a + bX_1$, where a, b, c ≠ 0), perfect multicollinearity appears only in the presence of an intercept, with $a = cX_2 – bX_1$ accounting for the intercept column, in the data matrix.

That is the significance of the intercept column in the data matrix X. This is because the affine function denotes a non-zero (non-homogeneous) linear combination of the two variables; that is $X_2 = a + bX_1$ implies $cX_2 + bX_1 = a$. And if in the data matrix there is a column of values equivalent to this constant a, then linear dependency occurs. However, a linear function (such as $cX_2 = bX_1$, c, b ≠ 0) is a homogeneous linear combination: here $cX_2 – bX_1 = 0$, and the above condition is irrelevant here.

Note that a (as well as b) in the affine function can take any value (other than zero), even though the intercept column in the data matrix is in ones, as the following illustrates:

Example 3: Consider $X_2 = 3 + 2X_1$ in the X-data matrix

$$\mathrm{X} = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 7 \\ 1 & 3 & 9 \end{bmatrix},$$

where the first column represents intercept and the second and third columns are X_1 and X_2 (where $X_2 = 3 + 2X_1$) respectively. This is a linear combination: $X_2 – 2X_1 = 3$, a column of constant; and the intercept also is a column of constant; hence X is a singular matrix (with zero determinant). The corresponding

$$\mathrm{X}'\mathrm{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 5 & 7 & 9 \end{bmatrix}\begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 7 \\ 1 & 3 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 21 \\ 6 & 14 & 46 \\ 21 & 46 & 155 \end{bmatrix},$$

also is a singular matrix, and the parameters are inestimable. However, if we consider the data matrix without the unit (intercept) column, that is,

$$X = \begin{bmatrix} 1 & 5 \\ 2 & 7 \\ 3 & 9 \end{bmatrix}, \text{ the corresponding}$$

$$\mathrm{X}'\mathrm{X} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 9 \end{bmatrix}\begin{bmatrix} 1 & 5 \\ 2 & 7 \\ 3 & 9 \end{bmatrix} = \begin{bmatrix} 14 & 46 \\ 46 & 155 \end{bmatrix}$$

is non-singular, with a determinant of 54, even though X_1 and X_2 are perfectly positively correlated (as $X_2 = 3 + 2X_1$)! Thus, perfect multicollinearity appears only in the presence of an intercept column in the data matrix when X_1 and X_2 variables are in affine function (such as $cX_2 = a + bX_1$, where a, b, c ≠ 0), involving a non-homogeneous linear combination.

But note that linear combination need not involve only the intercept term. It can also appear among the explanatory variables:

Example 4: Consider three explanatory variables in the data matrix X: X_1, X_2, and X_3, where $X_1 + X_2 = X_3$, a linear combination, where a = b = c = 1:

$$\mathrm{X} = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 1 & 3 \\ 3 & 3 & 6 \end{bmatrix}.$$

The determinant here is zero. The corresponding

$$\mathrm{X'X} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 3 \\ 5 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 & 5 \\ 2 & 1 & 3 \\ 3 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 14 & 15 & 29 \\ 15 & 26 & 41 \\ 29 & 41 & 70 \end{bmatrix},$$

also is a singular matrix, and the parameters are inestimable. In the presence of an intercept also, we will have the same result:

Given $\mathrm{X} = \begin{bmatrix} 1 & 1 & 4 & 5 \\ 1 & 2 & 1 & 3 \\ 1 & 3 & 3 & 6 \end{bmatrix}$, where the first column represents intercept,

$$\mathrm{X'X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 4 & 1 & 3 \\ 5 & 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 & 4 & 5 \\ 1 & 2 & 1 & 3 \\ 1 & 3 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 8 & 14 \\ 6 & 14 & 15 & 29 \\ 8 & 15 & 26 & 41 \\ 14 & 29 & 41 & 70 \end{bmatrix},$$

with a determinant of zero. Note that in the X′X also we have the same linear dependnecy in both the cases.

Example 5: Linear combination among the explanatory variables in the data matrix X can also appear as: $X_1 + X_2 = X_3 + X_4$:

In $\mathrm{X} = \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 2 & 1 \\ 3 & 3 & 4 & 2 \end{bmatrix}$, we have a linear dependency, sum of the first two columns (variables) = sum of the last two columns.

$$\mathrm{X'X} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 3 \\ 3 & 2 & 4 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 4 & 3 & 2 \\ 2 & 1 & 2 & 1 \\ 3 & 3 & 4 & 2 \end{bmatrix} = \begin{bmatrix} 14 & 15 & 19 & 10 \\ 15 & 26 & 26 & 15 \\ 19 & 26 & 29 & 16 \\ 10 & 15 & 16 & 9 \end{bmatrix}.$$

The determinant here also is zero. The same linear dependency is there in the X′X.

Generalization

We can generalise the above using a data matrix with a unit column for intercept and two variables, X_1 and X_2 over n observations:

$$X = \begin{bmatrix} 1 & X_{11} & X_{21} \\ .. & .. & .. \\ 1 & X_{1n} & X_{2n} \end{bmatrix} \text{ such that}$$

$$\mathrm{X'X} = \begin{bmatrix} 1 & .. & 1 \\ X_{11} & .. & X_{1n} \\ X_{21} & .. & X_{2n} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{21} \\ .. & .. & .. \\ 1 & X_{1n} & X_{2n} \end{bmatrix} = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i} X_{2i} \\ \sum X_{2i} & \sum X_{1i} X_{2i} & \sum X_{2i}^2 \end{bmatrix}.$$

As long as there is no perfect multicollinearity in the data matrix, X′X is non-singular, hence invertible and the three estimates are determinate. Now consider three cases: (i) a linear function: $X_2 = bX_1$, $b \neq 0$; (ii) an affine function: $X_2 = a + bX_1$, $a, b \neq 0$ in the presence of the intercept column and (iii) an affine function: $X_2 = a + bX_1$, $a, b \neq 0$ without the intercept column.

(i) A linear function: $X_2 = bX_1$, $b \neq 0$. The cross-product matrix now becomes:

$$\mathrm{X'X} = \begin{bmatrix} n & \sum X_{1i} & b\sum X_{1i} \\ \sum X_{1i} & \sum X_{1i}^2 & b\sum X_{1i}^2 \\ b\sum X_{1i} & b\sum X_{1i}^2 & b^2\sum X_{1i}^2 \end{bmatrix}.$$

The determinant of this matrix is:

$$|\mathrm{X'X}| = \mathrm{n}\sum X_{1i}^2\,(b^2\sum X_{1i}^2)$$

$$+\sum X_{1i}\,(b\sum X_{1i}^2)(b\sum X_{1i})$$

$$+\sum X_{1i}\,(b\sum X_{1i}^2)(b\sum X_{1i})$$

$$-\ (b\sum X_{1i})(b\sum X_{1i})\ \sum X_{1i}^2$$

$$-\ \sum X_{1i}\,(b^2\sum X_{1i}^2)\sum X_{1i}$$

$$-\ n(b\sum X_{1i}^2)(b\sum X_{1i}^2)\ =\ 0.$$

(ii) An affine function: $X_2 = a + bX_1$, $a, b \neq 0$ in the presence of the intercept column. The cross-product matrix now becomes:

$$\mathrm{X'X} = \begin{bmatrix} n & \sum X_{1i} & (na + b\sum X_{1i}) \\ \sum X_{1i} & \sum X_{1i}^2 & (a\sum X_{1i} + b\sum X_{1i}^2) \\ (na + b\sum X_{1i}) & (a\sum X_{1i} + b\sum X_{1i}^2) & (na^2 + b^2\sum X_{1i}^2 + 2ab\sum X_{1i}) \end{bmatrix}.$$

The determinant of this matrix is:

$$|\mathrm{X'X}| = \mathrm{n}\sum X_{1i}^2\,(na^2 + b^2\sum X_{1i}^2 + 2ab\sum X_{1i})$$

$$+\sum X_{1i}\,(a\sum X_{1i} + b\sum X_{1i}^2)(na + b\sum X_{1i})$$

$$+\sum X_{1i}\,(a\sum X_{1i} + b\sum X_{1i}^2)(na + b\sum X_{1i})$$

$$-\ (na + b\sum X_{1i})(na + b\sum X_{1i})\ \sum X_{1i}^2$$

$$-\ \sum X_{1i}\,(na^2 + b^2\sum X_{1i}^2 + 2ab\sum X_{1i})\sum X_{1i}$$

$$-\ n(a\sum X_{1i} + b\sum X_{1i}^2)(a\sum X_{1i} + b\sum X_{1i}^2)\ =\ 0.$$

(iii) An affine function: $X_2 = a + bX_1$, a, b ≠ 0 without the intercept column. The data matrix has now only two columns for the two variables X_1 and X_2 over n observations, without the unit column for intercept: $X = \begin{bmatrix} X_{11} & X_{21} \\ .. & .. \\ X_{1n} & X_{2n} \end{bmatrix}$ such that the cross-product matrix is:

$$X'X = \begin{bmatrix} X_{11} & .. & X_{1n} \\ X_{21} & .. & X_{2n} \end{bmatrix} \begin{bmatrix} X_{11} & X_{21} \\ .. & .. \\ X_{1n} & X_{2n} \end{bmatrix} = \begin{bmatrix} \Sigma X_{1i}^2 & \Sigma X_{1i} X_{2i} \\ \Sigma X_{1i} X_{2i} & \Sigma X_{2i}^2 \end{bmatrix}.$$

With $X_2 = a + bX_1$, a, b ≠ 0, this matrix becomes

$$X'X = \begin{bmatrix} \Sigma X_{1i}^2 & (a\Sigma X_{1i} + b\Sigma X_{1i}^2) \\ (a\Sigma X_{1i} + b\Sigma X_{1i}^2) & (na^2 + b^2\Sigma X_{1i}^2 + 2ab\Sigma X_{1i}) \end{bmatrix}.$$

And the determinant of this matrix is:

$$|X'X| = \sum X_{1i}^2 \,(na^2 + b^2 \sum X_{1i}^2 + 2ab \sum X_{1i})$$
$$- \,(a \sum X_{1i} + b \sum X_{1i}^2)(a \sum X_{1i} + b \sum X_{1i}^2) \; = a^2[n \sum X_{1i}^2 \; - \; (\sum X_{1i})^2] \neq 0.$$

Note that when a = 0 here, the affine function reduces to linear function, and the determinant will be zero, rendering OLS estimators indeterminate.

Thus, when X_1 and X_2 variables are in linear function (such as $X_2 = bX_1$, b ≠ 0), perfect multicollinearity appears, irrespective of the presence/absence of an intercept column in the data matrix. On the other hand, when X_1 and X_2 variables are in affine function (such as $X_2 = a + bX_1$, where a, b ≠ 0), perfect multicollinearity appears only in the presence of an intercept column in the data matrix.

Now let us also consider the correlation coefficient between the two variables X_1 and X_2, denoted as r_{12}.

$$r_{12} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\cdot\text{Var}(X_2)}} = \frac{\Sigma(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)/n}{\sqrt{\text{Var}(X_1)\cdot Var(X_2)}}.$$

Given the general affine function between the two variables ($X_{2i} = a + bX_{1i}$, where a, b ≠ 0),

$$\bar{X}_2 = a + b\bar{X}_1, \text{ and } (X_{2i} - \bar{X}_2) = b(X_{1i} - \bar{X}_1).$$

Similarly, variance of X_2 = Var(a + bX_1) = b^2Var(X_1). Substituting these in the expression for the correlation coefficient and noting that $Var(X_1) = \sum(X_{1i} - \bar{X}_1)^2/n$,

$$r_{12} = \frac{\Sigma(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)/n}{\sqrt{\text{Var}(X_1)\cdot\text{Var}(X_2)}} = \frac{\Sigma(X_{1i} - \bar{X}_1)b(X_{1i} - \bar{X}_1)/n}{\sqrt{\text{Var}(X_1)\cdot b^2\text{Var}(X_1)}} = \frac{\text{bVar}(X_1)}{\text{bVar}(X_1)} = 1.$$

Now consider the third case above of the affine function: $X_2 = a + bX_1$, a, b ≠ 0 without the intercept column. In this case, there is perfect correlation between the two variables, but that *does not necessitate* perfect multicollinearity, as we have seen.

However, surprisingly, many Econometrics textbooks spread insufficient information in this important issue. For example, Jan Kmenta (1986: 433) writes: “For the multiple regression model with two explanatory variables, perfect multicollinearity means that we can write …. $X_{i2} = a + bX_{i3}$, where a and b are some fixed numbers and $b \neq 0$. In this case there is perfect correlation between the two explanatory variables in the sample.” And “the first [least squares] normal equation [for the model with two explanatory variables] is exactly equal to the second normal equation multiplied by b. Therefore, the two equations are not independent, and the solution for [the parameters] β_2 and β_3 is indeterminate.”

Anna Koutsoyiannis (1977: 233) says: “If the explanatory variables are perfectly linearly correlated, that is, if the correlation coefficient for these variables is equal to unity, the parameters become indeterminate: it is impossible to obtain numerical values for each parameter separately and the method of least squares breaks down.”

Yongmiao Hong (2020: 72) just writes: “… we have exact or perfect multicollinearity if the correlation between two explanatory variables is equal to 1 or −1.”

A. H. Studenmund (2017: 222) states: “Perfect multicollinearity violates Classical Assumption VI, which specifies that no explanatory variable is a perfect linear function of any other explanatory variable. The word *perfect* in this context implies that the variation in one explanatory variable can be *completely* explained by movements in another explanatory variable. Such a perfect linear function between two independent variables would be: $X_{1i} = \alpha_0 + \alpha_1 X_{2i}$.”

2.2 Dummy Variable Trap

Let us recap our findings on perfect multicollinearity:

The problem occurs with affine functions such as

1. $aX_1 = bX_2$, $a, b \neq 0$,
2. $aX_1 = bX_2 + cX_3$, $a, b, c \neq 0$,
3. $aX_1 + bX_2 = cX_3 + dX_4$, $a, b, c, d \neq 0$.

Perfect multicollinearity appears also with linear functions; for example, in the second of the above functions, ($aX_1 = bX_2 + cX_3$, $a, b, c \neq 0$), if the first variable (X_1) is a column of constant (say, $X_1 = 1$), we have a linear function: $a = bX_2 + cX_3$ or $bX_2 = a – cX_3$. As the intercept represents a column of units in the regression data matrix X, we have $bX_2 = 1 – cX_3$. If X_2 and X_3 are two dummy variables for a two-category qualitative variable (say, gender) and $b = c = 1$, the data matrix X suffers from dummy variable trap, as there is perfect multicollinearity among the three columns (including the intercept), as we have seen above. But if there is no intercept column, there would not be any linear dependency, despite a linear relationship between the two dummy variables: $X_2 = 1 – X_3$, again demonstrated above. This we further explain below.

First, we consider the case of a single qualitative (categorical) variable and then extend it to cases of two and more qualitative variables.

2.2.1 Case of One Qualitative (Categorical) Variable

Case 1: Let us consider a very simple model (as in Adrian C. Darnell (1994, *A Dictionary of Econometrics*, cited above) of gender discrimination in wage rate, with wage rate (Yi) as the dependent variable and gender of the individual worker as the independent variable. We know gender is a qualitative (categorical, nominal) variable and a dummy variable (Gi) is used to represent it, where $G_i = 1$ if worker i is male and $G_i = 0$ if female. The equation of relationship is then $Y_i = \alpha + \beta G_i + \varepsilon_i$, which implies that the expected wage rate of a male $[E(Y_i| Gi = 1)]$ is given by $\alpha + \beta$, and that for a female $[E(Y_i| Gi = 0)]$ is given by α. The parameter β thus represents the difference between the expected wage rates of male and female workers.

Example 1: Suppose the data matrix in this case has six workers, the first three being male and the last three, female. The matrix with an intercept column of ones for α and the gender dummy variable is given below. The second column shows that the first three workers are male and the remining are female.

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \text{ and the } X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 3 & 3 \end{bmatrix}.$$

The X′X is non-singular with a determinant value of nine, and the parameters are estimable.

Note that in this example there are two distinct categories, namely, male and female, but only one dummy variable is used.

Case 2: Suppose we use two dummy variables: M_i takes the value one if the worker is male and zero otherwise, while F_i takes the value one if the worker is female and zero otherwise; then the equation becomes $Y_i = \alpha + \delta Mi + \gamma F_i + \varepsilon_i$.

Example 2: Let us continue with the above data matrix of six workers, the first three being male and the last three, female. The matrix with an intercept column of ones for α and the two (male and female) dummy variables is given below. The second and third columns are according to the definitions of M_i and F_i as already given above.

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}; \quad X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

The X′X in this case is singular with a determinant value of zero, and the parameters are inestimable, because the model suffers perfect multicollinearity; with the two dummy variables for male and female and the intercept column of ones, there is an exact linear relationship among the three columns: if α is taken to represent the intercept column of the data matrix, α = M$_i$ + F$_i$, that is, the first column equals the sum of second and third columns. This case of perfect multicollinearity is the dummy variable trap and occurs when, with q dichotomous qualitative categories, the regression equation is specified to include q dummy variables (for the q categories) and also a constant term. Note that there is an affine function between the two dummy variables, M$_i$ and F$_i$, such that M$_i$ = a + bF$_i$, where a = 1 (= α) and b = –1, and thus perfect negative correlation also. This we show below:

$$r_{MF} = \frac{\mathrm{Cov}(M,F)}{\sqrt{\mathrm{Var}(M)\cdot\mathrm{Var}(F)}} = \frac{\sum(M_i - \bar{M})(F_i - \bar{F})/n}{\sqrt{\mathrm{Var}(M)\cdot\mathrm{Var}(F)}}.$$

and the general affine function between the two variables (M$_i$ = 1 – F$_i$),

$$\bar{M} = (1 - \bar{F}), \text{ and } (M_i - \bar{M}) = -(F_i - \bar{F}).$$

Similarly, variance of M = Var(1 – F) = Var(F). Substituting these in the expression for the correlation coefficient and noting that $\mathrm{Var}(F) = \sum(F_i - \bar{F})^2/n$,

$$r_{MF} = \frac{\mathrm{Cov}(M,F)}{\sqrt{\mathrm{Var}(M)\cdot\mathrm{Var}(F)}} = \frac{\sum(M_i - \bar{M})(F_i - \bar{F})/n}{\sqrt{\mathrm{Var}(M)\cdot\mathrm{Var}(F)}} = \frac{\sum(-(F_i - \bar{F})(F_i - \bar{F})/n}{\sqrt{\mathrm{Var}(F)\cdot\mathrm{Var}(F)}} = \frac{(-)\mathrm{Var}(F)}{\mathrm{Var}(F)} = -1.$$

Now, given the definitions of the two dummy variables, that is, M$_i$ takes the value one if the worker is male and zero otherwise, while F$_i$ takes the value one if the worker is female and zero otherwise, and the equation of relationship Y$_i$ = α + δM$_i$ + γF$_i$ + ε$_i$, let us find out the expected wage rates of the two categories:

The expected wage rate of a male [E(Y$_i$| M$_i$ = 1)] is given by α + δ, and that for a female [E(Y$_i$| F$_i$ = 1)] is given by α + γ (note that when M$_i$ = 1, F$_i$ = 0 by definition, and vice versa). The parameter (δ – γ) thus represents the difference between the expected wage rates of male and female workers, rendering the parameter α an irrelevant one, (its presence signifying perfect multicollinearity), as we will see below.

Case 3: Now let us dispense with the intercept column in the above model and in the data matrix. The model is now Y$_i$ = δMi + γF$_i$ + ε$_i$, with the usual definitions of the two dummy variables.

Example 3: The data matrices are now:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}; \quad X'X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

The X′X in this case is non-singular with a determinant value of nine, and the parameters are estimable, because there is no perfect multicollinearity here, despite a perfect negative correlation between the two columns of the data matrix.

The expected wage rate of a male in this case [$E(Y_i| M_i = 1)$] is given by δ, and that for a female [$E(Y_i| F_i = 1)$] is given by γ. The parameter ($\delta - \gamma$) thus represents the difference between the expected wage rates of male and female workers. And this is the same as the above case including the intercept α: the expected wage rate of a male [$E(Y_i| M_i = 1)$] = $\alpha + \delta$, and that of a female [$E(Y_i| F_i = 1)$] = $\alpha + \gamma$. The difference between the two is ($\delta - \gamma$). This shows that α in Example 2 above is irrelevant in the equation used there and that it is its presence that has activated perfect multicollinearity.

Case 4: Now, let us keep the intercept in the equation and remove one of the dummy variables, say F_i. The equation of relationship is now: $Y_i = \alpha + \delta M_i + \varepsilon_i$, which is the same as the one that we used earlier (Case 1): $Y_i = \alpha + \beta G_i + \varepsilon_i$. This is because, from the first one, the expected wage rate of a male [$E(Y_i| M_i = 1)$] is given by $\alpha + \delta$, and that for a female [$E(Y_i| M_i = 0)$] is given by α. The parameter δ thus represents the difference between the expected wage rates of male and female workers. From the second equation, at the same time, we have the expected wage rate of a male [$E(Y_i| Gi = 1)$] given by $\alpha + \beta$, and that for a female [$E(Y_i| Gi = 0)$] given by α. The parameter β thus represents the difference between the expected wage rates of male and female workers. This means $\delta = \beta$, and the two equations are the same.

These examples illustrate the remedies for the dummy variable trap. Case one is to use q – 1 dummy variables for q categories along with an intercept term in the model; the omitted category is set as the ‘base’ here, such that the associated parameters of the q – 1 dummy variables are then interpreted as differences with respect to the base. In our first example, $Y_i = \alpha + \beta G_i + \varepsilon_i$. female ($G_i = 0$) is taken as the base such that the expected female wage rate is represented by the intercept term (α), and β represents the difference between expected wages of male and female workers.

Case two is to use q dummy variables for q categories without an intercept term in the model. This is illustrated in our later example, $Y_i = \delta M_i + \gamma F_i + \varepsilon_i$, where we get the expected wage rate of male and that of female directly from the parameters associated with the corresponding dummy variables.

However, if our aim is to test for differences (say gender discrimination in wage rates), then case one (q – 1 dummy variables for q categories along with an intercept) allows a direct examination of hypotheses of differences, by way of simple t-tests of the coefficients on the dummy variables.

Generalization

As earlier, we can now generalise the above

(i) using a data matrix with a unit column for intercept and two dummy variables, D$_1$ and D$_2$ for two categories (say, male and female workers) over n observations:

$$X = \begin{bmatrix} 1 & D_{11} & D_{21} \\ .. & .. & .. \\ 1 & D_{1n} & D_{2n} \end{bmatrix} \text{ such that}$$

$$\mathrm{X'X} = \begin{bmatrix} 1 & .. & 1 \\ D_{11} & .. & D_{1n} \\ D_{21} & .. & D_{2n} \end{bmatrix} \begin{bmatrix} 1 & D_{11} & D_{21} \\ .. & .. & .. \\ 1 & D_{1n} & D_{2n} \end{bmatrix} = \begin{bmatrix} n & \sum D_{1i} & \sum D_{2i} \\ \sum D_{1i} & \sum D_{1i}^2 & \sum D_{1i} D_{2i} \\ \sum D_{2i} & \sum D_{1i} D_{2i} & \sum D_{2i}^2 \end{bmatrix}.$$

Suppose D$_1$ takes value one if the worker is male and zero otherwise, while D$_2$ takes value one if the worker is female and zero otherwise, such that D$_{1i}$ + D$_{2i}$ = 1, or D$_{2i}$ = 1 – D$_{1i}$ (an affine function, such as X$_2$ = a + bX$_1$, with a = 1 and b = –1). We prove that this data matrix involves dummy variable trap. Substituting for D$_{2i}$ = 1 – D$_{1i}$ in the above cross-product matrix,

$$\mathrm{X'X} = \begin{bmatrix} n & \sum D_{1i} & \sum (1 - D_{1i}) \\ \sum D_{1i} & \sum D_{1i}^2 & \sum D_{1i}(1 - D_{1i}) \\ \sum (1 - D_{1i}) & \sum D_{1i}\,(1 - D_{1i}) & \sum (1 - {D_{1i}}^2 \end{bmatrix}. \text{ Or}$$

$$\mathrm{X'X} = \begin{bmatrix} n & \sum D_{1i} & n - \sum D_{1i} \\ \sum D_{1i} & \sum D_{1i}^2 & \sum D_{1i} - \sum D_{1i}^2 \\ n - \sum D_{1i} & \sum D_{1i} - \sum D_{1i}^2 & n + \sum D_{1i}^2 - 2\sum D_{1i} \end{bmatrix}.$$

The determinant of this matrix is:

$$|\mathrm{X'X}| = \mathrm{n} \sum D_{1i}^2 \,(n + \sum D_{1i}^2 - 2\sum D_{1i})$$

$$+ \sum D_{1i}\,(\sum D_{1i} - \sum D_{1i}^2)(n - \sum D_{1i})$$

$$+ \sum D_{1i}\,(\sum D_{1i} - \sum D_{1i}^2)(n - \sum D_{1i})$$

$$-\ (n - \sum D_{1i})(n - \sum D_{1i})\ \sum D_{1i}^2$$

$$-\ \sum D_{1i}\,(n + \sum D_{1i}^2 - 2\sum D_{1i}) \sum D_{1i}$$

$$-\ n(\sum D_{1i} - \sum D_{1i}^2)(\sum D_{1i} - \sum D_{1i}^2)\ =\ 0.$$

Thus, this X′X is singular with a zero determinant value, and the parameters are inestimable, because there is perfect multicollinearity here: the third column is the difference between the first two columns, that is, C3 = C1 – C2, or C1 = C2 + C3, the affine function involving the intercept column of ones and the two gender dummy variables of the X-data matrix.

(ii) Now if we omit the second dummy variable, the X-data matrix has then only two columns, one for the intercept (the column of ones) and the other for the dummy variable D$_1$ (say for male) over n observations:

$$\mathrm{X} = \begin{bmatrix} 1 & D_{11} \\ .. & .. \\ 1 & \mathrm{D}_{1n} \end{bmatrix}$$ such that the cross-product matrix is:

$$\mathrm{X'X} = \begin{bmatrix} 1 & .. & 1 \\ D_{11} & .. & \mathrm{D}_{1n} \end{bmatrix} \begin{bmatrix} 1 & D_{11} \\ .. & .. \\ 1 & \mathrm{D}_{1n} \end{bmatrix} = \begin{bmatrix} \mathrm{n} & \Sigma D_{1i} \\ \Sigma D_{1i} & \Sigma D_{1i}^2 \end{bmatrix}.$$

The determinant $[n\Sigma D_{1i}^2 - (\Sigma D_{1i})^2]$ is not zero, and the solution is determinate.

(iii) Finally, if we omit the intercept, the X-data matrix is left with only two columns for the two dummy variables D_1 and D_2 over n observations:

$$\mathrm{X} = \begin{bmatrix} \mathrm{D}_{11} & D_{21} \\ .. & .. \\ \mathrm{D}_{1n} & \mathrm{D}_{2n} \end{bmatrix}$$ such that the cross-product matrix is:

$$\mathrm{X'X} = \begin{bmatrix} \mathrm{D}_{11} & .. & \mathrm{D}_{1n} \\ D_{21} & .. & \mathrm{D}_{2n} \end{bmatrix} \begin{bmatrix} \mathrm{D}_{11} & D_{21} \\ .. & .. \\ \mathrm{D}_{1n} & \mathrm{D}_{2n} \end{bmatrix} = \begin{bmatrix} \Sigma D_{1i}^2 & \Sigma D_{1i} D_{2i} \\ \Sigma D_{1i} D_{2i} & \Sigma D_{2i}^2 \end{bmatrix}.$$

With $D_{2i} = 1 – D_{1i}$, this matrix becomes

$$\mathrm{X'X} = \begin{bmatrix} \sum D_{1i}^2 & \sum D_{1i} - \sum D_{1i}^2 \\ \sum D_{1i} - \sum D_{1i}^2 & n + \sum D_{1i}^2 - 2\sum D_{1i} \end{bmatrix}.$$

And the determinant of this matrix is:

$$|\mathrm{X'X}| = \sum \mathrm{D}_{1i}^2\,(\mathrm{n} + \sum \mathrm{D}_{1i}^2 - 2\sum \mathrm{D}_{1i})$$
$$– (\sum \mathrm{D}_{1i} - \sum \mathrm{D}_{1i}^2)(\sum \mathrm{D}_{1i} - \sum \mathrm{D}_{1i}^2) \;= [n\sum \mathrm{D}_{1i}^2 \;-\; (\sum \mathrm{D}_{1i})^2] \neq 0.$$

Thus, the cross-product matrix is non-singular and the parameters are estimable.

The last two cases are the possible solutions for dummy variable trap that occurs in the first case, as perfect multicollinearity appears there in the presence of an intercept column in the X-data matrix.

Thus, when X_1 and X_2 variables are in linear function (such as $X_2 = bX_1$, $b \neq 0$), perfect multicollinearity appears, irrespective of the presence/absence of an intercept column in the data matrix. On the other hand, when X_1 and X_2 variables are in affine function (such as $X_2 = a + bX_1$, where $a, b \neq 0$),

Now let us also consider the correlation coefficient between the two dummy variables D_1 and D_2, denoted as r_{12}.

$$r_{12} = \frac{\mathrm{Cov}(D_1, D_2)}{\sqrt{\mathrm{Var}(D_1)\cdot\mathrm{Var}(D_2)}} = \frac{\sum(D_{1i} – \bar{D}_1)(D_{2i} – \bar{D}_2)/n}{\sqrt{\mathrm{Var}(D_1)\cdot Var(D_2)}}.$$

Given the relationship between the two dummy variables ($D_{2i} = 1 – D_{1i}$),

$$\bar{D}_2 = 1 - \bar{D}_1, \text{ and } (D_{2i} – \bar{D}_2) = -(D_{1i} – \bar{D}_1).$$

Similarly, variance of D_2 = Var($1 – D_1$) = Var(D_1). Substituting these in the expression for the correlation coefficient and noting that $Var(D_1) = \sum(D_{1i} – \bar{D}_1)^2/n$,

$$r_{12} = \frac{\sum(D_{1i} - \bar{D}_1)(D_{2i} - \bar{D}_2)/n}{\sqrt{\mathrm{Var}(D_1)\cdot \mathrm{Var}(D_2)}} = \frac{\sum(D_{1i} - \bar{D}_1)[-(D_{1i} - \bar{D}_1)]/n}{\sqrt{\mathrm{Var}(D_1)\cdot Var(D_1)}} = \frac{-\mathrm{Var}(D_1)}{\mathrm{Var}(D_1)} = -1.$$

Now consider the third case above with the two dummy variables only. In this case, there is perfect (negative) correlation between the two variables, but that *does not necessitate* perfect multicollinearity, as there is no intercept column to materialize the perfect multicollinearity in the affine function ($D_{2i} = 1 - D_{1i}$).

2.2.2 Case of Two Categorical Variables

See Jack Johnston and John DiNardo (1997: 137); Badi H. Baltagi (2021: 98); and William Greene (2018: 162) for brief discussion on this topic.

Let us now extend our earlier example of wage discrimination model to include one more qualitative (categorical) variable, say, race, in addition to the nominal variable gender.

Case 1: Here also we use a dummy variable to represent the nominal variable of race that takes a value of one (1) for 'white' workers and zero (0) for 'black' workers (or vice versa). Now, given the definitions of the two dummy variables, that is, M_i takes the value one if the worker is male and zero otherwise, while W_i takes the value one if the worker is 'white' and zero otherwise, and the equation of relationship $Y_i = \alpha + \delta M_i + \beta W_i + \varepsilon_i$, let us find out the expected wage rates of the two categories.

The expected wage rate of a white male worker $[E(Y_i| M_i = 1, W_i = 1)]$ is given by $\alpha + \delta + \beta$, and that for a white female worker $[E(Y_i| M_i = 0, W_i = 1)]$ is given by $\alpha + \beta$. The parameter δ thus represents the difference between the expected wage rates of white male and white female workers. Similarly, the expected wage rate of a black male worker $[E(Y_i| M_i = 1, W_i = 0)]$ is given by $\alpha + \delta$, and that for a black female worker $[E(Y_i| M_i = 0, W_i = 0)]$ is given by α. The parameter δ thus represents the difference between the expected wage rates of black male and black female workers. Thus, we have the following events:

(i) difference between the expected wage rates of white male and white female workers $= (\alpha + \delta + \beta) - (\alpha + \beta) = \delta$: gender discrimination.

(ii) difference between the expected wage rates of black male and black female workers $= (\alpha + \delta) - (\alpha) = \delta$: gender discrimination.

(iii) difference between the expected wage rates of white male and black male workers $= (\alpha + \delta + \beta) - (\alpha + \delta) = \beta$: racial discrimination.

(iv) difference between the expected wage rates of white female and black female workers $= (\alpha + \beta) - (\alpha) = \beta$: racial discrimination.

(v) difference between the expected wage rates of white male and black female workers $= (\alpha + \delta + \beta) - (\alpha) = (\delta + \beta)$: compounded gender-race discrimination.

(vi) difference between the expected wage rates of white female and black male workers $= (\alpha + \beta) - (\alpha + \delta) = (\beta - \delta)$: compounded racial supremacy, if $\beta > \delta$.

If the parameter estimates are significant, the corresponding sample data exemplify a case of both gender and racial discrimination in agricultural wage rates.

Example 1: Let us continue with the above data matrix of six workers, the first three being male and the last three, female, now modified by including one more dummy for race, every second worker being black. The matrix with an intercept column of ones for α (first column) and the two dummy variables for gender (second column, M_i = 1 for male and zero for female) and race (third column, W_i = 1 for white and zero for black) is given below.

$$X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}; \quad X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 \\ 3 & 3 & 2 \\ 3 & 2 & 3 \end{bmatrix}$$

The X′X in this case is non-singular with a determinant value of 12, and the parameters are estimable, as the X-data matrix harbours no perfect multicollinearity.

Case 2: Now, what will happen if we include two more dummy variables to represent female workers and black workers? In this case, the equation of relationship becomes $Y_i = \alpha + \delta M_i + \gamma F_i + \beta W_i + \phi B_i + \varepsilon_i$, where F_i takes value one if the worker is female and zero otherwise, while B_i takes value one if the worker is 'black' and zero otherwise. It is evident that in this case we have perfect multicollinearity with $M_i + F_i = W_i + B_i = \alpha$, as the example below shows.

Example 2: In this case, we include two more dummy variables, one for female (third column, F_i = 1 for female and zero otherwise) and the other for black (fifth column, B_i = 1 for black and zero otherwise).

$$X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix};$$

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 & 3 & 3 \\ 3 & 3 & 0 & 2 & 1 \\ 3 & 0 & 3 & 1 & 2 \\ 3 & 2 & 1 & 3 & 0 \\ 3 & 1 & 2 & 0 & 3 \end{bmatrix}$$

The X′X in this case is singular with a determinant value of zero, leaving the parameters inestimable, as in this case we have perfect multicollinearity with $\alpha = M_i + F_i = W_i + B_i$, evident from the columns of the X-data matrix and the X′X matrix: $C_1 = C_2 + C_3 = C_4 + C_5$.

Case 3: Now let us dispense with the intercept column in the above model and in the data matrix, as we have done earlier in the case of a single categorical variable as regressor to avoid perfect multicollinearity. The model is now $Y_i = \delta M_i + \gamma F_i + \beta W_i + \phi B_i + \varepsilon_i$,, with the usual definitions of the four dummy variables. With this, let us consider the following events:

(i) difference between the expected wage rates of white male and white female workers = $(\delta + \beta) - (\gamma + \beta) = (\delta - \gamma)$: gender discrimination.

(ii) difference between the expected wage rates of black male and black female workers = $(\delta + \phi) - (\gamma + \phi) = (\delta - \gamma)$: gender discrimination.

(iii) difference between the expected wage rates of white male and black male workers = $(\delta + \beta) - (\delta + \phi) = (\beta - \phi)$: racial discrimination.

(iv) difference between the expected wage rates of white female and black female workers = $(\gamma + \beta) - (\gamma + \phi) = (\beta - \phi)$: racial discrimination.

(v) difference between the expected wage rates of white male and black female workers = $(\delta + \beta) - (\gamma + \phi)$: compounded gender-race discrimination.

(vi) difference between the expected wage rates of white female and black male workers = $(\gamma + \beta) - (\delta + \phi)$: compounded racial supremacy if $(\gamma + \beta) > (\delta + \phi)$.

Note that the results from the events (i) and (ii) represent gender discrimination, as in Case 1 above, and from (iii) and (iv) represent racial discrimination. However, this model also suffers from perfect multicollinearity: $M_i + F_i = W_i + B_i$. Thus, the parameters remain indeterminate, as the following example shows.

Example 3: The data matrix is now:

$$X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix};$$

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 2 & 1 \\ 0 & 3 & 1 & 2 \\ 2 & 1 & 3 & 0 \\ 1 & 2 & 0 & 3 \end{bmatrix}$$

The X′X here also is singular with a determinant value of zero, leaving the parameters inestimable, with perfect multicollinearity given by $M_i + F_i = W_i + B_i$, evident from the columns of the X-data matrix and the X′X matrix: $C_1 + C_2 = C_3 + C_4$. This suggests that we need to omit one dummy variable here.

Case 4 (a): Let us now omit one dummy variable in the above model and in the data matrix, say, B_i. The model is now $Y_i = \delta M_i + \gamma F_i + \beta W_i + \varepsilon_i$,, with the usual definitions of the three dummy variables. With this, let us consider the following events:

(i) difference between the expected wage rates of white male and white female workers $= (\delta + \beta) - (\gamma + \beta) = (\delta - \gamma)$: gender discrimination.

(ii) difference between the expected wage rates of black male and black female workers $= (\delta - \gamma)$: gender discrimination.

(iii) difference between the expected wage rates of white male and black male workers $= (\delta + \beta) - (\delta) = (\beta)$: racial discrimination.

(iv) difference between the expected wage rates of white female and black female workers $= (\gamma + \beta) - (\gamma) = (\beta)$: racial discrimination.

(v) difference between the expected wage rates of white male and black female workers $= (\delta + \beta) - (\gamma)$: compounded gender-race discrimination.

(vi) difference between the expected wage rates of white female and black male workers $= (\gamma + \beta) - (\delta)$: compounded racial supremacy, if $(\gamma + \beta) > (\delta)$.

The results from the events (i) and (ii) represent gender discrimination, same as in Case 3 above, and from (iii) and (iv) represent racial discrimination, same as in Case 1. Note that this model does not suffer from perfect multicollinearity.

Example 4(a): The data matrix is now:

$$X = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \quad X'X = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 2 \\ 0 & 3 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

The X′X now is non-singular with a determinant value of 12, capable of estimating the parameters, as there is no perfect multicollinearity now.

Case 4 (b): The same model we can re-specify by omitting one of the two dummy variables of 'gender', say, F_i, and including both the dummies of 'race'. The model is now $Y_i = \delta M_i + \beta W_i + \phi B_i + \varepsilon_i$,, with the usual definitions of the three dummy variables. Then we have the following events:

(i) difference between the expected wage rates of white male and white female workers $= (\delta + \beta) - (\beta) = (\delta)$: gender discrimination.

(ii) difference between the expected wage rates of black male and black female workers $= (\delta + \phi) - (\phi) = (\delta)$: gender discrimination.

(iii) difference between the expected wage rates of white male and black male workers $= (\delta + \beta) - (\delta + \phi) = (\beta - \phi)$: racial discrimination.

(iv) difference between the expected wage rates of white female and black female workers = (β – ϕ): racial discrimination.

(v) difference between the expected wage rates of white male and black female workers = (δ + β) – (ϕ): compounded gender-race discrimination.

(i) difference between the expected wage rates of white female and black male workers = (β) – (δ + ϕ): compounded racial supremacy, if β > (δ + ϕ).

The results from the events (i) and (ii) here represent gender discrimination, same as in Case 1 above, and from (iii) and (iv) represent racial discrimination, same as in Case 3. This model also does not have any perfect multicollinearity.

Example 4(b): The data matrix is now:

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}; \quad \mathrm{X'X} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

The X′X here also is non-singular with the same determinant value of 12; there is no perfect multicollinearity, and the parameter estimates are possible.

2.2.3 Case of Three Categorical Variables

Unfortunately, no available textbooks deal with this case. So, let us now extend the above discrimination model (with gender and race variables) to include one more qualitative (categorical) variable, say, skill, with two categories, skilled and unskilled.

Case 1: When we use only one (q – 1) dummy variable for two (q) categories of every qualitative variable along with an intercept, no dummy variable trap appears. Thus, the equation of relationship $Y_i = \alpha + \delta M_i + \beta W_i + \lambda S_i + \varepsilon_i$, where $S_i = 1$ if the worker is skilled and zero otherwise, is valid. Some of the possible events are:

(ii) difference between the expected wage rates of white skilled male and white skilled female workers = (α + δ + β + λ) – (α + β + λ) = (δ): gender discrimination.

(iii) difference between the expected wage rates of black skilled male and black skilled female workers = (α + δ + λ) – (α + λ) = (δ): gender discrimination.

(iv) difference between the expected wage rates of white skilled male and black skilled male workers = (α + δ + β + λ) – (α + δ + λ) = (β): racial discrimination.

(v) difference between the expected wage rates of white skilled female and black skilled female workers = (α + β + λ) – (α + λ) = (β): racial discrimination.

(vi) difference between the expected wage rates of white skilled male and black skilled female workers = $(\alpha + \delta + \beta + \lambda) - (\alpha + \lambda) = (\delta + \beta)$: compounded gender and racial discrimination.

(vii) difference between the expected wage rates of white skilled female and black skilled male workers = $(\alpha + \beta + \lambda) - (\alpha + \delta + \lambda) = (\beta - \delta)$: compounded racial supremacy, if $\beta > \delta$.

Among the unskilled workers also, we have the same results, with $\lambda = 0$. Considering both the skilled and unskilled workers,

(viii) difference between the expected wage rates of white skilled male and white unskilled skilled female workers = $(\alpha + \delta + \beta + \lambda) - (\alpha + \beta) = (\delta + \lambda)$: compounded gender-skill discrimination.

(ix) difference between the expected wage rates of black skilled male and black unskilled female workers = $(\alpha + \delta + \lambda) - (\alpha) = (\delta + \lambda)$: compounded gender-skill discrimination.

(x) difference between the expected wage rates of white skilled male and black unskilled male workers = $(\alpha + \delta + \beta + \lambda) - (\alpha + \delta) = (\beta + \lambda)$: compounded race-skill discrimination.

(xi) difference between the expected wage rates of white skilled female and black unskilled female workers = $(\alpha + \beta + \lambda) - (\alpha) = (\beta + \lambda)$: compounded race-skill discrimination.

(xii) difference between the expected wage rates of white skilled male and black unskilled female workers = $(\alpha + \delta + \beta + \lambda) - (\alpha) = (\delta + \beta + \lambda)$: compounded gender-race-skill discrimination.

(xiii) difference between the expected wage rates of white skilled female and black unskilled male workers = $(\alpha + \beta + \lambda) - (\alpha + \delta) = (\beta + \lambda - \delta)$: compounded race-skill supremacy, if $(\beta + \lambda) > \delta$.

Example 1: Suppose in the X-data matrix with an intercept and two dummy variables for gender (column 2) and race (column 3), we have one more column, representing skilled workers; assume that the first two of the three male (and female) workers are skilled, such that the X-data matrix is now

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix};$$

$$\mathrm{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 & 4 \\ 3 & 3 & 2 & 2 \\ 3 & 2 & 3 & 2 \\ 4 & 2 & 2 & 4 \end{bmatrix}$$

The X′X is non-singular with a determinant value of 16; there is no perfect multicollinearity, and the parameter estimates are possible.

Case 2: However, let us see what happens if we extend the Case 4(a) or 4(b) above to include the two categories of the new variable skill (without an intercept). In this case, the model becomes, say, $Y_i = \delta M_i + \beta W_i + \phi B_i + \lambda S_i + \rho U_i + \varepsilon_i$, where $U_i = 1$ if the worker is unskilled and zero otherwise. That is, we use one dummy for gender and two dummies each for race and skill variables. Evidently, we can find perfect multicollinearity here: $W_i + B_i = S_i + U_i$.

Example 2: The X-data and the X′X matrices in this case are:

$$X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix};$$

$$\mathrm{X'X} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 & 2 & 1 \\ 2 & 3 & 0 & 2 & 1 \\ 1 & 0 & 3 & 2 & 1 \\ 2 & 2 & 2 & 4 & 0 \\ 1 & 1 & 1 & 0 & 2 \end{bmatrix}$$

Perfect multicollinearity in terms of $W_i + B_i = S_i + U_i$ appears here in the columns of both the X-data and X′X matrices ($C_2 + C_3 = C_4 + C_5$), rendering the latter a singular matrix with zero determinant.

Case 3: The above means we have to omit any one dummy variable out of these four. This yields the new model $Y_i = \delta M_i + \beta W_i + \phi B_i + \lambda S_i + \varepsilon_i$, or $Y_i = \delta M_i + \beta W_i + \lambda S_i + \rho U_i + \varepsilon_i$, or $Y_i = \delta M_i + \gamma F_i + \beta W_i + \lambda S_i + \varepsilon_i$, without any perfect multicollinearity.

Example 2: The X-data and the X′X matrices for the last model ($Y_i = \delta M_i + \gamma F_i + \beta W_i + \lambda S_i + \varepsilon_i$,) are:

$$X = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}; \ \mathrm{X'X} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 3 & 1 & 2 \\ 2 & 1 & 3 & 2 \\ 2 & 2 & 2 & 4 \end{bmatrix}$$

The X′X is non-singular with a determinant value of 16; there is no perfect multicollinearity, and the parameter estimates are determinate.

In short, we have the following points to avoid dummy variable trap in the context of two or more qualitative (categorical) variables:

(i) If the intercept (constant) is retained, use q – 1 dummy variables for q categories of *every* qualitative variable;
(ii) If the intercept is dropped, use q – 1 dummy variables for q categories of p – 1 of p qualitative variables; that is, we can use q dummy variables for all q categories of only one qualitative variable; with all other qualitative variables, use only q – 1 dummy variables for q categories.

3. Conclusion

It is good to remember the summary assessment given by Adrian Pagan in 1989 of the work on Granger causality:

> "There was a lot of high-powered analysis of this topic, but I came away from a reading of it with the feeling that it was one of the most unfortunate turnings for econometrics in the last two decades, and it has probably generated more nonsense results than anything else during that time." (Pagan 1989: 325).

Strangely, it is the negatives of the information revolution that seem to have captured the fancy of some of the new generation self-styled pundits of econometrics; the proliferation of (mostly pirated) econometric software on the one hand and the 'impressively simplifying' online tutorials on the other hand has "generated more nonsense results than anything else" from these package-driven self-styled econometricians at the cost of theoretical understanding and assessment. Unfortunately, most of the econometrics text books are also often insufficient to bring to light the concepts clearly. Hence this paper has sought to explain the issue of perfect multicollinearity and dummy variable trap in the framework of mathematical logic.

Appendix
Dummy Variable Trap Explained: Some Authoritative Sources

(a) Case of One Categorical Variable

(b) We start with Adrian C. Darnell (1994, *A Dictionary of Econometrics*). He considers in the section on 'Dummy variables', "a very simple model" with hours of labour supplied (Yi) as the dependent variable (DV, regressand) and wage rate (Xi) and sex of the individual worker (Gi) as the independent variables (IVs, regressors). A dummy (Di) is used to represent the sex "where $D_i = 1$ if individual i is male and $D_i = 0$ if female". The equation of relationship is then "$Y_i = \alpha + \beta X_i + \gamma D_i + \varepsilon_i$, which implies that the expected hours of labour supplied by a male, facing a wage rate of X_0, is given by $\alpha + \beta X_0 + \gamma$; that for a female facing the same wage rate is given by $\alpha + \beta X_0$. The parameter γ thus represents the difference between the expected labour supplies of men and women who face the same wage rate. It is to be noted immediately that in this case there are two categories to be distinguished, namely, male and female, but only one dummy is used. Suppose two dummies were used: $D^m{}_i$ takes the value one if the individual is male and zero otherwise while $D^f{}_i$ takes the value one if the individual is female and zero otherwise; then $Y_i = \alpha + \beta X_i + \gamma D^m{}_i + \xi D^f{}_i + \varepsilon_i$. This model cannot be estimated, for it suffers perfect multicollinearity. This may be noted from the fact that the parameter α may be seen as the coefficient on a dummy variable, ι, everyone of whose elements is unity; hence, in the model with the two dummies there is an exact linear relationship between the regressors ι, D^m and D^f: $\iota = D^m + D^f$. This situation is referred to as the dummy variable trap and occurs when, with p dichotomous qualitative categories, the regression equation is specified to include p linearly independent 0-1 dummies and also a constant term. With p categories, if a constant term is to be included, one category should be set as the 'base' and only p – 1 dummies are included so that their associated parameters are then interpreted as differences with respect to the base. Hence in the model with only one dummy for the individual's sex, female labour supply is the base and γ represents the difference between men and women. One could, of course, use two dummies in such a situation but one would then omit the constant term: $Y_i = \beta X_i + \gamma_m D^m{}_i + \gamma_f D^f{}_i + \varepsilon_i$. This model, without a constant term, may be estimated and here $\gamma_m + \beta X_0$ is the expected labour supply of a male facing a wage rate of X_0, while a female is expected to supply $\gamma_f + \beta X_0$. In the model with only one dummy, the corresponding supplies are $\alpha + \beta X_0 + \gamma$ and $\alpha + \beta X_0$; hence $\alpha + \gamma = \gamma_m$ and $\alpha = \gamma_f$, demonstrating that the use of the one dummy sets the female as the base. Including a constant term, then, in a model which incorporates dichotomous dummy variables demands that with p categories only p – 1 dummies are used, and their associated parameters represent differences with respect to a chosen base; the choice of base is without loss of generality. The alternative is to exclude the constant explicitly and use all p dummies; however, one use of dummies is to test for differences, and in the former approach allows a direct examination of hypotheses of differences, via simple t-tests of the coefficients on the dummy variables." P. 108-109.

(c) Next comes J. Johnston's *Econometric Methods* (1972, 2nd edition); he considers a consumption function with two dummy variables, one for wartime and the other for peacetime; that is, "X_1 = 1 in each wartime year and zero in each peacetime year; X_2 = 1 in each wartime year and zero in each peacetime year.… At this stage, we must warn the readers of the dummy variable trap. If explanatory variables such as [these dummy variables] are used in conjunction with a regression program that automatically produces an intercept term, then the estimating procedure breaks down, for this is equivalent to using an expanded data matrix … in which the first is a column of units and the remaining" two columns are those of the two dummy variables. "A linear dependence then exists between the columns, that is, col(1) – col(2) – col(3) … = **0** [that is, a null column vector] so that (**X′X**) is singular." (p. 178-179).

(d) William H. Greene in his *Econometric Analysis* (2020, 8th edition) explains: "Note that only three of the four quarterly dummy variables are included in the model. If the fourth were included, then the four dummy variables would sum to one at every observation, which would replicate the constant term—a case of perfect multicollinearity. This is known as the **dummy variable trap**. To avoid the dummy variable trap, we drop the dummy variable for the fourth quarter. (Depending on the application, it might be preferable to have four separate dummy variables and drop the overall constant. See Suits (1984) and Greene and Seaks (1991)." (p. 197; emphasis as in the original; also see pp. 199, 202).

(e) According to James H. Stock and Mark W. Watson (2020. *Introduction to Econometrics*, 4th ed.), "[One] possible source of perfect multicollinearity arises when multiple binary, or dummy, variables are used as regressors. For example, suppose you have partitioned the school districts into three categories: rural, suburban, and urban. Each district falls into one (and only one) category. Let these binary variables be $Rural_i$, which equals 1 for a rural district and equals 0 otherwise; $Suburban_i$; and $Urban_i$. If you include all three binary variables in the regression along with a constant, the regressors will be perfectly multicollinear: Because each district belongs to one and only one category, $Rural_i + Suburban_i + Urban_i = 1 = X0_i$, where $X0_i$ denotes the constant regressor introduced in Equation (6.6). Thus, to estimate the regression, you must exclude one of these four variables, either one of the binary indicators or the constant term. By convention, the constant term is typically retained, in which case one of the binary indicators is excluded. For example, if $Rural_i$ were excluded, then the coefficient on $Suburban_i$ would be the average difference between test scores in suburban and rural districts, holding constant the other variables in the regression.

"In general, if there are *G* binary variables, if each observation falls into one and only one category, if there is an intercept in the regression, and if all *G* binary variables are included as regressors, then the regression will fail because of perfect multicollinearity. This situation is called the **dummy variable trap**. The usual way to avoid the dummy variable trap is to exclude one of the binary variables from the multiple regression, so only G - 1 of the G binary variables are included as regressors. In this case, the coefficients on the included binary variables represent the incremental effect of being in that category, relative to the base case of the omitted category, holding constant the other regressors. Alternatively, all G binary regressors can be included if the intercept is omitted from the regression." (pp. 229-230, emphasis as in the original; also see p. 777)

(f) Christopher Dougherty (2011, *Introduction to Econometrics*. 4th ed.) writes: "What would happen if you included a dummy variable for the reference category [in addition to the constant]? There would be two consequences.

"First, were it possible to compute regression coefficients, you would not be able to give them an interpretation. The coefficient b_1 is a basic estimate of the intercept, and the coefficients of the dummies are the estimates of the increase in the intercept from this basic level, but now there is no definition of what is basic, so the interpretation collapses.

"The other consequence is that the numerical procedure for calculating the regression coefficients will break down and the computer will simply send you an error message (or possibly, in sophisticated applications, drop one of the dummies for you). Suppose that there are m dummy categories and you define dummy variables D_1,…., Dm. Then, in observation i, ΣD_{ji} = 1 because one of the dummy variables will be equal to 1 and all the others will be equal to 0. But the intercept β_1 is really the product of the parameter β_1 and a special variable whose value is 1 in all observations …. Hence, for all observations, the sum of the dummy variables is equal to this special variable, and one has an exact linear relationship among the variables in the regression model. This is known as the dummy variable trap. As a consequence, the model is subject to a special case of exact multicollinearity, making it impossible to compute regression coefficients.

"An alternative procedure for avoiding the dummy variable trap is to drop the intercept from the model. The special unit variable is thereby dropped and there is no longer an exact linear relationship among the variables." (p. 235-237).

(g) Chris Brooks in his *Introductory econometrics for finance* (2019, 4th edition), says: "We could either include all three dummy variables together (and not include an intercept in the regression equation) or only include two of the dummy variables and retain the intercept. If we include all three dummy variables and the intercept at the same time, the regression model could not be estimated and this is known as the *dummy variable trap*." *(*p. 224; emphasis as in the original). In Chapter 10, he explains in detail: "The sum of the four dummies would be 1 in every time period. Unfortunately, this sum is of course identical to the variable that is implicitly attached to the intercept coefficient. Thus, if the four dummy variables and the intercept were both included in the same regression, the problem would be one of perfect multicollinearity so that (X′X)−1 would not exist and none of the coefficients could be estimated. This problem is known as the *dummy variable trap*. The solution would be either to just use three dummy variables plus the intercept, or to use the four dummy variables with no intercept." (P. 578; emphasis as in the original).

(h) In Badi H. Baltagi's *Econometrics* (2021, 6th Edition), we find "dummy variable trap" as follows: "Briefly stated, there will be *perfect multicollinearity* between MALE, FEMALE, and the constant. In fact, MALE + FEMALE = 1. Some researchers may choose to include the intercept and exclude one of the gender dummy variables, say MALE." (P. 97; emphasis as in the original).

(i) Damodar N. Gujarati and Dawn C. Porter (2009. *Basic econometrics*. 5th ed.) give a detailed discussion: "Although they are easy to incorporate in the regression models, one must use the dummy variables carefully. In particular, consider the following aspects:

"1. In Example 9.1, to distinguish the three regions, we used only two dummy variables, D_2 and D_3. Why did we not use three dummies to distinguish the three regions? Suppose we do that and write the model (9.2.1) as:

$$Y_i = \alpha + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i \qquad \textbf{(9.2.6)}$$

"where D_{1i} takes a value of 1 for states in the West and 0 otherwise. Thus, we now have a dummy variable for each of the three geographical regions. Using the data in Table 9.1, if you were to run the regression (9.2.6), the computer would "refuse" to run the regression (try it).6 (6Actually you will get a message saying that the data matrix is singular.) Why? The reason is that in the setup of Eq. (9.2.6) where you have a dummy variable for each category or group and also an intercept, you have a case of **perfect collinearity,** that is, exact linear relationships among the variables. Why? Refer to Table 9.1. Imagine that now we add the D_1 column, taking the value of 1 whenever a state is in the West and 0 otherwise.

"Now if you add the three *D* columns horizontally, you will obtain a column that has 51 ones in it. But since the value of the intercept α is (implicitly) 1 for each observation, you will have a column that also contains 51 ones. In other words, the sum of the three *D* columns will simply reproduce the intercept column, thus leading to perfect collinearity. In this case, estimation of the model (9.2.6) is impossible.

"The message here is: **If a qualitative variable has *m* categories, introduce only (*m* − 1) dummy variables.** In our example, since the qualitative variable "region" has three categories, we introduced only two dummies. If you do not follow this rule, you will fall into what is called the **dummy variable trap,** that is, the situation of perfect collinearity or perfect multicollinearity, if there is more than one exact relationship among the variables. This rule also applies if we have more than one qualitative variable in the model, an example of which is presented later. Thus we should restate the preceding rule as: **For each qualitative regressor, the number of dummy variables introduced must be one less than the categories of that variable.** Thus, if in Example 9.1 we had information about the gender of the teacher, we would use an additional dummy variable (but not two) taking a value of 1 for female and 0 for male or vice versa.

"2. The category for which no dummy variable is assigned is known as the **base, benchmark, control, comparison, reference,** or **omitted category.** And all comparisons are made in relation to the benchmark category…….

"6. We warned above about the dummy variable trap. There is a way to circumvent this trap by introducing as many dummy variables as the number of categories of that variable, *provided we do not introduce the intercept in such a model.* Thus, if we drop the intercept term from Eq. (9.2.6), and consider the following model,

$$Y_i = \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i \qquad \textbf{(9.2.7)}$$

we do not fall into the dummy variable trap, as there is no longer perfect collinearity. *But make sure that when you run this regression, you use the no-intercept option in your regression package………*

"*…. with the intercept suppressed, and allowing a dummy variable for each category, we obtain directly the mean values of the various categories.*" (pp. 281-282, emphasis as in the original).

(j) According to Dimitrios Asteriou and Stephen G. Hall (2021, *Applied Econometrics* 4th Ed.), "This concept [perfect multicollinearity] can be understood better by using the dummy variable trap. Take, for example, X_1 to be the intercept (so $X_1 = 1$), and X_2, X_3, X_4 and X_5 to be seasonal dummies for quarterly time series data (that is, X_2 takes the value of 1 for the first quarter, 0 otherwise; X_3 takes the value of 1 for the second quarter, 0 otherwise and so on). Then in this case $X_2 + X_3 + X_4 + X_5 = 1$; and because $X_1 = 1$ then $X_1 = X_2 + X_3 + X_4 + X_5$. So, …. this set of variables is linearly dependent." They note that "linear dependence means that one variable can be expressed as a linear combination of one or more, or even all, of the other variables." (Pp. 104-105; also see p. 223, 226).

From Some Internet Sources:

(a) "In the above model, the sum of all category dummy variable for each row is equal to the intercept value of that row - in other words there is perfect multi-collinearity (one value can be predicted from the other values). Intuitively, there is a duplicate category: if we dropped the male category it is inherently defined in the female category (zero female value indicate male, and vice-versa). The solution to the dummy variable trap is to drop one of the categorical variables (or alternatively, drop the intercept constant) - if there are m number of categories, use m – 1 in the model, the value left out can be thought of as the reference value and the fit values of the remaining categories represent the change from this reference."[4]

(b) "If dummy variables for all categories were included, their sum would equal 1 for all observations, which is identical to and hence perfectly correlated with the vector-of-ones variable whose coefficient is the constant term; if the vector-of-ones variable were also present, this would result in perfect multicollinearity, so that the matrix inversion in the estimation algorithm would be impossible. This is referred to as the dummy variable trap."[5]

(c) "One common error is the *dummy variable trap,* in which a complete set of dummy variables and an intercept, or more than one complete set of dummy variables, are included in a regression. For example, including a variable for female gender (coded 1/0), a variable for male gender, and an intercept would cause the regression to fail."[6]

(b) Case of More Than One Categorical Variable

(a) Jack Johnston and John DiNardo (1997, Econometric Methods 4th ed.) explains: "If there are *two* sets of dummy variables in a relationship [two separate categorical variables, say, gender and race] and the constant is suppressed, the estimation procedure still breaks down, because the included dummy variables are linearly dependent (the sum of the first set minus the sum of the second set gives the zero vector). If a constant is retained, one dummy variable must be dropped from each set; and this rule obviously extends to three or more sets." (p 137).

(b) Badi H. Baltagi (2021 Econometrics 6th ed.) continues with his example of dummy variable trap using the categorical variable gender. "What happens when another qualitative variable is included, to depict another classification of the individuals in the sample, say for example, race? If there are three race groups in the sample, WHITE, BLACK, and HISPANIC, one could create a dummy variable for each of these

[4] https://www.algosome.com/articles/dummy-variable-trap-regression.html.

[5] https://en.wikipedia.org/wiki/Dummy_variable_(statistics)

[6] https://www.encyclopedia.com/social-sciences-and-law/sociology-and-social-reform/sociology-general-terms-and-concepts/multicollinearity.

classifications. For example, WHITE will take the value 1 when the individual is White and 0 when the individual is non-White. Note that the dummy variable trap does not allow the inclusion of all three categories as they sum up to 1. Also, even if the intercept is dropped, once MALE and FEMALE are included, perfect multicollinearity is still present because MALE + FEMALE = WHITE + BLACK + HISPANIC. Therefore, one category from race should be dropped. Suits (1984) argues that the researcher should use the dummy variable category omission to his or her advantage, in interpreting the results, keeping in mind the purpose of the study. For example, if one is interested in comparing earnings across gender holding race constant, the omission of MALE is natural, whereas if one is interested in the race differential in earnings holding gender constant, one of the race variables should be omitted. Whichever variable is omitted, this becomes the base category for which the other earnings are compared. Most researchers prefer to keep an intercept, although regression packages allow for a no intercept option. In this case one should omit one category from each of the race and gender classifications." (p. 98)

(c) William Greene (2018 Econometric Analysis, 8th ed) discusses the case of sets of categories: "The case in which several sets of dummy variables are needed is much the same as those we have already considered, with one important exception. Consider a model of state-wide per capita expenditure on education, *y*, as a function of state-wide per capita income, *x*. Suppose that we have observations on all n = 50 states for T = 10 years. A regression model that allows the expected expenditure to change over time as well as across states would be

yit = a + bxit + di + ut + eit.

As before, it is necessary to drop one of the variables in each set of dummy variables to avoid the dummy variable trap. For our example, if a total of 50 state dummies and 10 time dummies is retained, a problem of perfect multicollinearity remains; the sums of the 50 state dummies and the 10 time dummies are the same, that is, 1. One of the variables in each of the sets (or the overall constant term and one of the variables in one of the sets) must be omitted." (p.162)

REFERENCES

Asteriou, Dimitrios and Stephen G. Hall 2021, *Applied Econometrics* 4th Edition, Macmillan Education Limited, London:

Baltagi, Badi H. 2021 *Econometrics* Sixth Edition. Springer.

Brooks, Chris 2019 *Introductory econometrics for finance*, Fourth edition, Cambridge University Press

Darnell, Adrian C. 1994. *A Dictionary of Econometrics*. Edward Elgar, Cheltenham, UK.

Das, Panchanan 2019. *Econometrics in Theory and Practice: Analysis of Cross Section, Time Series and Panel Data with Stata 15.1*. Springer.

Dougherty, Christopher 2011. *Introduction to Econometrics*. 4th ed. Oxford University Press Inc., New York

Greene, W., and T. Seaks. "The Restricted Least Squares Estimator: A Pedagogical Note." *Review of Economics and Statistics*, 73, 1991, pp. 563–567.

Greene, William H. 2020. *Econometric Analysis* 8th ed., Pearson.

Gujarati, Damodar N. and Dawn C. Porter. 2009. *Basic econometrics*. 5th ed. McGraw-Hill

Hong, Yongmiao 2020. *Foundations of Modern Econometrics: A Unified Approach*.

Johnston, J 1972 *Econometric methods* 2nd ed. McGraw-Hill. p. 178.

Johnston, J. and John DiNardo 1997, *Econometric methods* 4th ed. McGraw-Hill.

Koutsoyiannis, A. 1977. *Theory of Econometrics: An Introductory Exposition of Econometric Methods*. 2nd ed. MacMillan Press Ltd., London.

Pagan, A.R. (1989), "20 Years after: Econometrics 1966-1986," in B. Cornet and H. Tulkens (eds.), *Contributions to Operations Research and Econometrics,* The MIT Press, Cambridge, MA.

Stock, James H. and Mark W. Watson, 2020. *Introduction to Econometrics,* 4th ed. Pearson Education Limited 2020.

Studenmund, A. H. 2017. *Using econometrics: a practical guide*. 7th Edition. Pearson Education, Inc.

Suits, D. "Dummy Variables: Mechanics vs. Interpretation." *Review of Economics and Statistics,* 66, 1984, pp. 177–180.

Wooldridge, Jeffrey M. 2020 *Introductory Econometrics: A Modern Approach,* Seventh Edition, Cengage Learning, Inc, World Scientific Publishing Co. Pte. Ltd.