CENTRAL BANK OF CYPRUS
EUROSYSTEM

WORKING PAPER SERIES

Back to Basics: Is Statistical Significance All That Matters?

Nektarios A. Michail

Constantinos I. Massouras

September 2014

Working Paper 2014-3

Address
80 Kennedy Avenue
CY-1076 Nicosia, Cyprus

Postal Address
P. O. Box 25529
CY-1395 Nicosia, Cyprus

E-mail
publications@centralbank.gov.cy

Website
http://www.centralbank.gov.cy

Fax
+357 22 378153

Papers in the Working Paper Series may be downloaded from:
http://www.centralbank.gov.cy/nqcontent.cfm?a_id=5755

Back to basics: is statistical significance all that matters?

Nektarios A. Michail* and Constantinos I. Massouras**

September 2014

Abstract

We examine whether statistical significance can convey all the information necessary for the econometrician to judge the performance of his models. We find that in some cases, the t-statistic and the R-squared can be biased due to high correlation between the variables and variables can appear significant even though their correlation with the dependent variable is very low. The solution we propose is a back-to-basics approach: examine correlations between variables and variable variances in order for a correct interpretation of the regression outcome.

Keywords: correlation, statistical significance, hypothesis testing, t test, F test

JEL Classification: C12, C21, C31

---

* Economic Research Department, Central Bank of Cyprus and Department of Commerce, Finance and Shipping, Cyprus University of Technology ** Economic Research Department, Central Bank of Cyprus and Department of Economics, University of Cyprus. We thank Angelos Roussos and Maria Papageorghiou for their comments and suggestions

Corresponding author: Nektarios Michail, Economic Research Department, Central Bank of Cyprus. Email: nektarios.michail@centralbank.gov.cy

Back to basics: is statistical significance all that matters?

1. Introduction

The aim of this paper is to examine whether statistical significance can provide assurance to the econometrician on the relationship between the variables under examination. Even though regression analysis is the most widely used tool in econometrics, our understanding of both statistical and economic significance resembles something of a black box. How do the relationships between independent variables or between the dependent and independent variables affect the regression? Is it possible for variables which are insignificant to the regression to appear statistically significant?

Statistical significance is often used by economists (or other scientists) to substantiate their claims of their instruments' validity; that is, if the variables on the right hand-side of the regression equation (i.e. the X's) are good in explaining or causing the dependant variable (i.e. the Y). In other words, as Ziliak and McCloskey (1996) state "...they are equating statistical significance with economic importance" where economic importance[1] (or significance) is defined as "the size and sign of beta hat" (Wooldridge, 2004). Another widely accepted principle is that "correlation does not mean causation". However, when economists are using the t-statistic or more generally statistical significance, this is actually disguised correlation. According to Black's (1982) criticism, "Econometric models are supposed to tell us about causation. Most often, they actually tell us about correlation".

Motivation for the creation of this paper can be found in the debate concerning the importance of statistical significance versus economic significance, which originated in McCloskey (1985) and continued with McCloskey and Ziliak (1996), yet built on predecessors such as Hendry (1980) and Keynes (1939). The discussion on the subject was further pursued and several arguments which continue until today were made on both sides [Ziliak and McCloskey (2004 a, b) Elliot and Granger (2004),

[1] A variable is considered as economically significant when it's inclusion in a model is justified by economic theories and thinking.

Leamer (2004), Zellner (2004), Mayer (2012), Ziliak and McCloskey (2013) to mention only a few]. On the two sides of this debate we have, Ziliak and McCloskey (1996) who support economic significance and argue that statistical significance is often misused and misinterpreted, and Spanos (2008) who severely criticises this argument and comments that the authors "...delude themselves far more than those economists at whom they wag their fingers...", and advocates that their solution is not credible.

In pursuing an answer to this debate, we de-construct the most widely used means of statistical testing (t-test, F-test and the goodness-of-fit) to get to the essence of statistical significance. In addition, we devote some space to discuss the effects of coefficient size in order to examine the relationship between the coefficients and the correlation between the variables. As our results show, statistical significance on its own cannot determine whether a variable can be considered as economically significant or whether it is important in a model specification. Significance is affected by other factors such as a variable's standard deviation and the correlation between the independent variables, which can, at times, drive the t-statistic so that a variable appears significant when in fact its correlation with the endogenous variable is negligible.

This does not mean that the t-statistic is unimportant when determining a model structure, yet knowledge of the above can assist in better deductions. Thus, we believe the econometrician should turn to the basics, i.e. the correlation between variables and sample variances or standard deviations, in order to be better equipped to interpret their output. Knowledge of the above can be of great assistance in any econometric model but is of particular importance in large-scale econometric models where the amount of data can easily mask relationships between variables. This is very evident in the case of policy-oriented models, such as the ones usually employed in central banking, where the introduction of variables with no correlation with the endogenous (even if variable interrelations make them appear statistically significant) can seriously affect the results and, more importantly, forecasts which can greatly influence policy advice.

2. De-construction of statistical testing

2.1. Simple linear regression

According to standard econometric textbooks (e.g. Stock and Watson (2010), Wooldridge (2012)) in the simplest case of the linear regression model the t-statistic is defined, in the case of $H_{null} = 0$, as

$$t_{stat} = \frac{\hat{\beta}-0}{s.e(\hat{\beta})} \qquad (1)$$

If we replace the estimates of $\hat{\beta}$ in the estimation for the standard error of $\hat{\beta}$ in (1) above, we get[2,3]

$$s.e(\hat{\beta}) = \sqrt{Var(Y) - 2\hat{\beta}\sum(y-\bar{y})(x-\bar{x}) + \hat{\beta}^2 Var(X)}$$

Using the relationship of correlation and covariance we get

$$s.e(\hat{\beta}) = \sqrt{Var(Y) - r_{xy}\frac{Sy}{Sx}r_{xy}\frac{Sy}{Sx}Var(X)}$$

which can be shown to equal the following after substituting in equation (1):

$$t_{stat} = \frac{r_{xy}}{\sqrt{(1-r^2{}_{xy})}} \qquad (2)$$

What equation (2) tells us is that both the numerator and the denominator depend on the correlation between the dependent and independent variable, which is the same regardless of the order as $r_{xy} = r_{yx}$. Hence, the relationship between the variables is the main driver of the statistic. This shows that the result of the t-statistic is the same regardless of the order of the variables (i.e. if we set the endogenous as the exogenous and vice versa) or, simply put, that significance does not mean causation.

For example, if we use weather as the independent variable and agricultural production as the dependent variable, then it is no surprise that there is a high value

[2] A detailed exposition of the steps for each equation can be found in the Appendix.

[3] For simplicity, we do not account for the sample size effect here since linear regression uses a balanced data pool. It is known that sample size affects the standard deviation estimates (McCloskey, 1985, and Black, 1982 for comments) but as the dataset is balanced, these cancel out in the derivations which follow.

of the t-statistic, indicating that weather is a good variable for explaining agricultural production. However, if we reverse them, the same value for the t-statistic will hold, indicating that agricultural production is good in explaining the weather. Even though this might hold, it is definitely not a good indicator of causality[4]. In the case of the simple linear regression, statistical significance can indicate a high degree of correlation between the variables.

2.2. Linear regression with two independent variables

A more interesting case occurs when we consider the case of a regression with two independent variables. Using the same procedure as in 2.1 we find that in this case, the t-statistic is influenced by the existence of the second variable[5], as presented below.

$$\text{t-stat}(\widehat{\beta 1}) = \frac{(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2})\sqrt{(1-r_{1,2}^2)}}{\sqrt{A}},$$

where A= $\left(1 - {r^2}_{x_1,x_2}\right)^2 + \left(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2}\right)^2$

$$+\left(r_{x_2,y} - r_{x_1,y}\, r_{x_1,x_2}\right)^2 - 2\left(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2}\right) r_{1,y}\left(1 - {r^2}_{x_1,x_2}\right)$$

$$-2\left(r_{x_2,y} - r_{x_1,y}\, r_{x_1,x_2}\right) r_{x_2,y}\left(1 - {r^2}_{x_1,x_2}\right)$$

$$+2\left(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2}\right)\left(r_{x_2,y} - r_{x_1,y}\, r_{x_1,x_2}\right) r_{x_1 x_2} \qquad (3)$$

Under this specification, we examine several scenarios to see whether statistical significance can be relied upon to provide coherent estimates of the relationship between the variables.

Numerical Test 1: In this specification we assume that the correlation between the two explanatory variables is zero (or approximately zero), i.e. that the two variables are independent; a requirement econometricians usually seek from their explanatory variables. Under this assumption, (2) simplifies to

---

[4] The usual econometrics technique employed for causality is what is commonly known as Granger-causality, i.e. testing whether a time-series is significant is predicting another time series. The interested reader may refer to Granger (1969), Granger (1980) for a theoretical background and Sims (1972) for an early application of the method.

[5] As in Section 2.1 detailed derivations of this result can be found in the Appendix

$t_{stat}(\widehat{\beta_1}) = \frac{r_{x_1,y}}{\sqrt{A^*}}$ where $A^*=1 - r^2_{x_1,y} - r^2_{x_2,y}$

We can observe that the probability of $\widehat{\beta_1}$ being significant relies on the relationship of the two variables with Y. Nevertheless, this has some implications: For example, if $r_{x_1,y} = 0.2$ while $r_{x_1,x_2} \approx 0$ then, after solving for the values of X_2 which make the statistic higher than 1.96 (the critical value at the 95% significance level), we find that X_1 will be statistically significant if $r_{x_2,y} > 0.97$ while X_2 will at the same time also be statistically significant. In essence, a high correlation between one variable and Y can mask the low correlation between the other variable and Y. At the same time

$\hat{\beta}_1 = \frac{Sy(r_{x_1,y})}{Sx_1}$ and $\hat{\beta}_2 = \frac{Sy(r_{x_2,y})}{Sx_2}$

indicating that the higher the correlation between the explanatory and dependent variable will result, *ceteris paribus*, in higher beta values. In the case where the relation between the independent variables is very low, the beta value can assist in assuring statistical significance, depending on the relative values of the variable variance. If the researcher deals with data in logarithm or standardised form, where variable variances do not differ much, the beta value can provide a useful proxy of the significance.

Numerical Test 2: What is even more interesting is the case where X_1 and X_2 are correlated. In order to set a cut off point to prevent multicollinearity between our variables we set the correlation between X_1 and X_2 at 0.8944 which gives us a variance inflation factor (VIF) statistic of exactly 5 (a common threshold for rejecting multicollinearity). Under this assumption, if $r_{x_1,y} = 0.1$ then, solving for the values which make the t-statistic higher than 1.96, it appears that for $\mathrm{r}_{\mathrm{x}_2,\mathrm{y}} > 0.499$ or $\mathrm{r}_{\mathrm{x}_2,\mathrm{y}} < -0.311$ both X_1 and X_2 will be statistically significant, even though they have a low or modest correlation to Y. Again, as beta values are influenced by the standard deviations of Y and X's interpretation of variable importance is very vague. Yet beta size can be a good rule of thumb if data are in either standardised or logarithmic form.

Numerical Test 3: Similarly to the above, we also estimate what the value of $r_{x_2,y}$ should be in order to prove that both X_1 and X_2 are statistically significant for $r_{x_1,y} = 0.1$ and $r_{x_1,x_2} = 0.8$ (i.e. VIF = 2.77). Our estimations show that both variables appear to be statistically significant if the correlation between X_2 and Y is $r_{x_2,y} > 0.6386$ or $r_{x_2,y} < -0.4662$ which, as expected, is greater (in absolute values) than when $r_{x_1,x_2} = 0.8944$. Nevertheless, both a variable with very low correlation with Y and a variable with moderate correlation with Y appear to be statistically significant because of the correlation between them. In addition, the same situation concerning beta values as in Numerical Test 2 also holds here.

2.3. Goodness of fit

Turning to the goodness-of-fit statistic, the R-squared is defined as:

$R^2 = r_{x_1,y}^2 + r_{x_2,y}^2$ for $r_{x_1,x_2} = 0$

$R^2 = \beta_1 r_{x_1,y} + \beta_2 r_{x_2,y}$ for $r_{x_1,x_2} \neq 0$

In the first case where $r_{x_1,x_2} = 0$ we can see that the high correlation of X_2 with the dependent variable (0.97) will force the R^2 to reach quite high values even though the correlation of X_1 with Y is extremely low. In this case, the R^2 and subsequently the F-statistic (defined as $\frac{R^2/k}{(1-R^2)/(n-k-1)}$) is not very useful in determining whether the model specification is correct as it can be biased upwards due to the high correlation of one variable with the endogenous variable. In the second case where $r_{x_1,x_2} \neq 0$, the standard deviation of the dependent variable and of the independent variables plays a major role. This happens as S_y, S_{x1} and S_{x2} enter the equation of R^2 through the beta estimation.

$\widehat{\beta_1} = \frac{Sy(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2})}{Sx_1(1 - r^2{}_{x_1,x_2})}$ and $\widehat{\beta_2} = \frac{Sy(r_{x_2,y} - r_{x_1,y}\, r_{x_1,x_2})}{Sx_2(1 - r^2{}_{x_1,x_2})}$

Thus, just like in Tests 2 and 3, the ratio between $\frac{S_y}{S_{x_1}}, \frac{S_y}{S_{x_2}}$ will contribute to the determination of the R^2 value. When these ratios are over 1, then the R^2 will be higher than when the ratios are less than 1. In the case where these ratios are equal to 1 then the R^2 will depend on the correlations, and most importantly to r_{x_1,x_2}. If

$\frac{s_y}{s_{x_1}} > 1$ and $\frac{s_y}{s_{x_2}} < 1$ and since in our case $r_{x_2,y} > r_{x_1,y} = 0.1$ then the R^2 is expected to be lower than in the case where $\frac{s_y}{s_{x_1}} < 1$ and $\frac{s_y}{s_{x_2}} > 1$ where the correlation of X_2 with Y will drive the R^2. This shows that a high correlation of an independent variable with the endogenous can severely affect the results and produce a high value for R^2 value (and subsequently F-test value), even though the other variable is in fact insignificant. As a general statement we note that in order to be able to interpret the R^2 we should also know standard deviation of our variables, in addition to their correlations.

3. Conclusions

Does statistical significance matter? We cannot reject this statement as the t-statistic (or the R-squared and F-statistic for that matter) will, most of the times, correctly reject a variable which is not correlated to the endogenous.

Nevertheless, as we have shown in this paper, there can be instances where statistical significance fails to reject the null hypothesis of non-significance due to either high correlation between the variables or high correlation of one variable with the endogenous. In these cases, the size of the beta coefficient could be employed as a rule of thumb to gauge the relationship between the variables and the endogenous especially if the data are in logarithm or standardised form.

Under these specifications, a higher beta value reflects a higher correlation between the endogenous and the exogenous variables. Yet, if the variance of each variable differs then this can easily mask its relationship with the dependent variable. Hence, our results would lead us to recommend a back-to-basics approach: testing for correlation and estimating the variance of the exogenous variables (while at the same time accounting for the effect of sample size) before running regression analysis can assist in avoiding such issues.

As for causality, given that the econometric arsenal can provide only Granger-causality tests to test for this relation, we can follow the advice of Black (1982) where “... all we can do is to keep making our theories more plausible and to keep testing the theories against measured correlations. Constructing an econometric model is just a way of testing a theory against measured correlations".

Appendix

Derivation of t-statistic in the Simple Linear Regression Framework:

$$t_{stat} = \frac{\hat{\beta} - 0}{s.e(\hat{\beta})} = \frac{r_{xy}\frac{Sy}{Sx}}{\frac{\sqrt{\sum Ui^2}}{\sqrt{Var(X)}}} = \frac{r_{xy}\frac{Sy}{Sx}}{\frac{\sqrt{\sum Ui^2}}{Sx}} = r_{xy}\frac{Sy}{\sqrt{\sum Ui^2}}$$

where $\sqrt{\sum Ui^2} = \sqrt{\sum (y - \hat{y})^2}$

$$= \sqrt{\sum (y - \hat{a} - \hat{\beta}x)^2}$$

Then substituting for $\hat{a} = \bar{y} - \hat{\beta}\bar{\chi}$

$$= \sqrt{\sum (y - \bar{y} + \hat{\beta}\bar{x} - \hat{\beta}\chi)^2}$$

$$= \sqrt{\sum [(y - \bar{y}) - \hat{\beta}(x - \bar{x})]^2}$$

$$= \sqrt{\sum \{(y - \bar{y})^2 - 2(y - \bar{y})[\hat{\beta}(x - \bar{x})] + \hat{\beta}^2(x - \bar{x})^2\}}$$

$$= \sqrt{Var(Y) - 2\hat{\beta}\sum (y - \bar{y})(x - \bar{x}) + \hat{\beta}^2 Var(X)}$$

$$= \sqrt{Var(Y) - 2\hat{\beta}Cov(X,Y) + \hat{\beta}^2 Var(X)}$$

Using $\hat{\beta} = \frac{Cov(X,Y)}{Var(X)}$ we get

$$= \sqrt{Var(Y) - 2\frac{Cov(X,Y)}{Var(X)}Cov(X,Y) + \frac{Cov^2(X,Y)}{Var^2(X)}Var(X)}$$

$$= \sqrt{Var(Y) - 2\frac{Cov(X,Y)}{Var(X)}Cov(X,Y) + \frac{Cov^2(X,Y)}{Var(X)}}$$

$$= \sqrt{Var(Y) - 2\frac{Cov^2(X,Y)}{Var(X)} + \frac{Cov^2(X,Y)}{Var(X)}}$$

$$= \sqrt{Var(Y) - \frac{Cov^2(X,Y)}{Var(X)}} = \sqrt{Var(Y) - \frac{Cov(X,Y)}{Var(X)}Cov(X,Y)}$$

Substituting $Cov(X,Y) = r_{xy}\frac{Sy}{Sx}Var(X)$ yields

$$= \sqrt{Var(Y) - r_{xy}\frac{Sy}{Sx}r_{xy}\frac{Sy}{Sx}Var(X)} = \sqrt{Var(Y) - r^2{}_{xy}Var(Y)}$$

$$= \sqrt{Var(Y)(1 - r^2{}_{xy})} = Sy\sqrt{(1 - r^2{}_{xy})}$$

Thus, $t_{stat} = \frac{\hat{\beta}-0}{s.e(\hat{\beta})} = \frac{r_{xy}\frac{Sy}{Sx}}{\frac{\sqrt{\sum Ui^2}}{\sqrt{Var(X)}}} = \frac{r_{xy}\frac{Sy}{Sx}}{\frac{\sqrt{\sum Ui^2}}{Sx}} = r_{xy}\frac{Sy}{\sqrt{\sum Ui^2}} = r_{xy}\frac{Sy}{Sy\sqrt{(1-r^2{}_{xy})}}$

and $t_{stat} = \frac{r_{xy}}{\sqrt{(1-r^2{}_{xy})}}$

<u>Derivation of the t-statistic in the linear regression with two independent variables framework:</u>

$$\min \mathrm{A} = \sum Ui^2 => \min A = \sum (Yi - \hat{a} - \hat{\beta}Xi)^2$$

$$\frac{dA}{da} = -2\sum\left(Yi - \hat{a} - \hat{b}X\right) = 0$$

$$\frac{dA}{db} = -2\sum\left(Yi - \hat{a} - \hat{\beta}X\right)X = 0$$

$$\sum Y - \sum \hat{\beta}X = \sum \hat{a} = n\hat{a}$$

$=> \hat{a} = \frac{1}{n}\sum Y - \hat{\beta}\sum X = \bar{y} - \hat{\beta}\bar{x}$

To solve for $\hat{\beta}$ we substitute for $\hat{a}$, which means that $X = \sum(Xi - \bar{X})$.

<u>Solution:</u>

$B = (X'X)^{-1}X'y = \frac{X'y}{X'X} = \frac{\begin{bmatrix} X_1 & X_2 \\ X_2 & X_1 \end{bmatrix}y}{\begin{bmatrix} X_1 & X_2 \\ X_2 & X_1 \end{bmatrix}\begin{bmatrix} X_1 & X_2 \\ X_2 & X_1 \end{bmatrix}} = \begin{bmatrix} \sum {X_1}^2 & \sum X_1 \quad X_2 \\ \sum X_2 \quad X_1 & \sum {X_2}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum X_1 y \\ \sum X_2 y \end{bmatrix}$

$$= \frac{\begin{bmatrix} \sum X_2{}^2 & -\sum X_2 \;\; X_1 \\ -\sum X_1 \;\; X_2 & \sum X_1{}^2 \end{bmatrix} \begin{bmatrix} \sum X_1 y \\ \sum X_2 y \end{bmatrix}}{\sum X_1^2 \sum X_2^2 - (\sum X_1 x_2)^2} = \frac{\begin{bmatrix} \sum X_2{}^2 \sum X_1 y & -\sum X_2 \;\; X_1 \sum X_2 y \\ -\sum X_1 \;\; X_2 \sum X_1 y & \sum X_1{}^2 \sum X_2 y \end{bmatrix}}{\sum X_1^2 \sum X_2^2 - (\sum X_1 x_2)^2}$$

$$=> B = \frac{\begin{bmatrix} Var(X_2)\, Cov(X_1,y) & -Cov(X_2,y) Cov(X_1,X_2) \\ -Cov(X_1,y) Cov(X_1,X_2) & Var(X_1) Cov(X_2,y) \end{bmatrix}}{Var(X_1) Var(X_2) - Cov^2(X_1,X_2)}$$

using $Cov(X,Y) = r_{xy} \frac{Sy}{Sx} Var(X)$ yields

$$\widehat{\beta_1} = \frac{Sx_2^2 r_{x_1,y} Sx_1 Sy - \; r_{x_2,y} Sx_2 Sy \; r_{x_1,x_2} Sx_2 Sx_1}{Sx_1^2 Sx_2^2 - \; (r_{x_1,x_2} Sx_1 Sx_2)^2}$$

$$\widehat{\beta_1} = \frac{Sx_2^2 Sx_1 Sy (r_{x_1,y} - \; r_{x_2,y} \; r_{x_1,x_2})}{Sx_1^2 Sx_2^2 (1 - \; r^2{}_{x_1,x_2})} = \frac{Sy(r_{x_1,y} - \; r_{x_2,y} \; r_{x_1,x_2})}{Sx_1 (1 - \; r^2{}_{x_1,x_2})}$$

Similarly for $\widehat{\beta_2} = \frac{Sy(r_{x_2,y} - r_{x_1,y} \, r_{x_1,x_2})}{Sx_2 (1 - r^2{}_{x_1,x_2})}$

$t_{stat}\, (\widehat{\beta_1}) = \frac{\widehat{\beta_1}}{s.e(\widehat{\beta}1)}$, *where* $s.e(\hat{\beta}_1) = \sqrt{\frac{\sum U_1^2}{Var(\beta_1)(1 - r_{1,2}^2)}}$ and

$$\hat{\beta}_1 = \frac{Sy(r_{x_1,y} - \; r_{x_2,y} \; r_{x_1,x_2})}{Sx_1 (1 - \; r^2{}_{x_1,x_2})}$$

$$=> t_{stat}\, (\widehat{\beta_1}) = \frac{\frac{Sy(r_{x_1,y} - r_{x_2,y} \, r_{x_1,x_2})}{Sx_1 (1 - r^2{}_{x_1,x_2})}}{\sqrt{\frac{S{y_{1,2}}^2}{\sum X_1^2 (1 - r_{1,2}^2)}}}$$

$$S{y_{1,2}}^2 = \sum (y - x_1 \; \beta_1 \;\; - \chi_2 \; \beta_2 \;)^2$$

$$= \sum y^2 - 2 \sum y x_1 \beta_1 - 2 \sum y x_2 \beta_2 + 2 \sum \chi_1 \chi_2 \beta_1 \beta_2 + \beta_1^2 \sum \chi_1^2 + \beta_2^2 \sum \chi_2^2$$

$$= Var(y) + \beta_1^2 Var\,(x_1) + \beta_2^2 Var\,(x_2) - \; 2\beta_1 Cov(y, x_1) - \; 2\beta_2 Cov(y, x_2) + 2\beta_1 \beta_2 Cov(\chi_1, x_2)$$

$$= S_y^2 + \; \beta_1^2 S_1^2 + \beta_2^2 S_2^2 - 2\beta_1 r_{1,y} \, S_y S_1 - 2\beta_2 r_{2,y} \, S_y S_2 + 2\beta_1 \beta_2 r_{1,2} \, S_1 S_2$$

Replacing β_1 and β_2

$$=> S{y_{1,2}}^2 = S_y^2 + \frac{S_y^2 (r_{x_1,y} - r_{x_2,y} \, r_{x_1,x_2})^2 S_{x_1}^2}{S_{x_1}^2 (1 - r^2{}_{x_1,x_2})^2} + \frac{S_y^2 (r_{x_2,y} - r_{x_1,y} \, r_{x_1,x_2})^2 S_{x_2}^2}{S_{x_2}^2 (1 - r^2{}_{x_1,x_2})^2}$$

$$-2\frac{S_y\,(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})S_{x_1}\,S_y r_{1,y}}{S_{x_1}\,(1-r^2{}_{x_1,x_2})}-2\frac{S_y\,(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})S_{x_2}\,S_y r_{2,y}}{S_{x_2}\,(1-r^2{}_{x_1,x_2})}$$

$$+2\frac{S_y\,(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})S_{x_1}}{S_{x_1}\,(1-r^2{}_{x_1,x_2})}\frac{Sy(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})S_{x_2}r_{1,2}}{Sx_2(1-r^2{}_{x_1,x_2})}$$

$$=S_y^2+\frac{S_y^2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})^2}{(1-r^2{}_{x_1,x_2})^2}+\frac{S_y^2(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})^2}{(1-r^2{}_{x_1,x_2})^2}-2\frac{S_y^2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})r_{1,y}}{(1-r^2{}_{x_1,x_2})}$$

$$-2\frac{S_y^2(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})r_{2,y}}{(1-r^2{}_{x_1,x_2})}+2\frac{S_y^2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})(r_{2,y}-r_{x_1,y}\,r_{x_1,x_2})r_{x_1,x_2}}{(1-r^2{}_{x_1,x_2})^2}$$

Taking $\frac{S_y^2}{(1-r^2{}_{x_1,x_2})^2}$ *as a common factor*

$$=>S_{y_{1,2}}^{\ 2}=\frac{S_y^2}{(1-r^2{}_{x_1,x_2})^2}[(1-r^2{}_{x_1,x_2})^2+(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})^2$$

$$+(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})^2-2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})r_{1,y}(1-r^2{}_{x_1,x_2})$$

$$-2(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})r_{2,y}(1-r^2{}_{x_1,x_2})$$

$$+2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})(r_{2,y}-r_{x_1,y}\,r_{x_1,x_2})r_{1,2}]$$

We denote **A** to be $A=[(1-r^2{}_{x_1,x_2})^2+(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})^2$

$$+(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})^2-2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})r_{1,y}(1-r^2{}_{x_1,x_2})$$

$$-2(r_{x_2,y}-r_{x_1,y}\,r_{x_1,x_2})r_{2,y}(1-r^2{}_{x_1,x_2})$$

$$+2(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})(r_{2,y}-r_{x_1,y}\,r_{x_1,x_2})\,r_{x_1,x_2}]$$

$$=>t_{stat}\,(\widehat{\beta_1})=\frac{\frac{Sy(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})}{Sx_1(1-r^2{}_{x_1,x_2})}}{\frac{\sqrt{S{y_{1,2}^2}}}{\sqrt{\sum X_1^2(1-r_{1,2}^2)}}}=\frac{Sy(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})\sqrt{\sum X_1^2(1-r_{1,2}^2)}}{Sx_1(1-r^2{}_{x_1,x_2})\sqrt{S{y_{1,2}^2}}}$$

$$=\frac{Sy(r_{x_1,y}-r_{x_2,y}\,r_{x_1,x_2})\sqrt{\sum X_1^2(1-r_{1,2}^2)}}{Sx_1(1-r^2{}_{x_1,x_2})\frac{\sqrt{S_y^2A}}{\sqrt{(1-r^2{}_{x_1,x_2})^2}}}$$

$$= \frac{Sy(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2})\sqrt{\sum X_1^2(1 - r_{1,2}^2)}\,(1 - r^2{}_{x_1,x_2})}{Sx_1(1 - r^2{}_{x_1,x_2})\sqrt{S_y^2 A}}$$

$$= \frac{Sy\big(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2}\big)\sqrt{\sum X_1^2(1 - r^2{}_{x_1,x_2})}}{Sx_1 S_y\sqrt{A}}$$

and simplifying yields $= \frac{(r_{x_1,y} - r_{x_2,y}\, r_{x_1,x_2})\sqrt{(1 - r^2{}_{x_1,x_2})}}{\sqrt{A}}$

References

Black, F. (1982). The Trouble with Econometric Models. *Financial Analysts Journal* , 29-37.

Elliot, G., & Granger, C. W. (2004). Evaluating Significance: Comments on "Size Matters". *Journal of Socio-Economics 33(5)* , 547-550.

Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica 37(3)* , 424-438.

Granger, C. (1988). Some recent development in a concept of causality. *Journal of Econometrics 39(1-2)* , 199-211.

Hendry, D. F. (1980). Econometrics - Alchemy or Science? *Economics, New Series, Vol.47, No.188* , 387-406.

Keynes, J. M. (1939). Professor Tinbergen's method. *Economic Journal 49* , 558-568.

Leamer, E. E. (2004). Are the Roads Red? Comments on "Size Matters". *Journal of Socio-Economics 33(5)* , 555-557.

Mayer, T. (2012). Ziliak and McCloskey's Criticism of Significance Tests: An Assessment. *Econ JournalWatch 9(3)* , 256-297.

McCloskey, D. N. (1985a). The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests. *American Economic Review Supplement 75 (2)* , 201-205.

McCloskey, D. N., & Ziliak, S. T. (1996). The Standard Error of Regressions. *Journal of Economic Literature Vol. XXXIV* , 97-114.

Sims, C. (1972). Money, Income and Causality. *American Economic Review 62(4)* , 540-552.

Spanos, A. (2008). Review of Stephen T. Ziliak and Deirdre N.mcCloskey's The cult of statistical significance: how the standard error costs us jobs, justice, and lives. *Erasmus Journal for Philosophy and Economics Volume 1, Issue 1* , 154-164.

Stock, J. H., & Watson, M. W. (2010). *Introduction to Econometrics.* Addison-Wesley.

Wooldridge, J. (2012). *Intorductory Econometrics: A Modern Approach (5th Edition).* South-Western College Publishers.

Wooldridge, J. M. (2004). Statistical Significance is Okay, Too: Comment on "Size Matters". *Journal of Socio-Economics 33(5)* , 577-579.

Zellner, A. (2004). To Test or Not to Test and if So, How? Comments on " Size Matters". *Journal of Socio-Economics 33(5)* , 581-586.

Ziliak, S. T., & McCloskey, D. N. (2004b). Significance Redux. *Journal of Socio-Economics 33(5)* , 665-675.

Ziliak, S. T., & McCloskey, D. N. (2004). Size matters: the standard error of regressions in the American Economic Review. *The Journal of Socio-Economics 33* , 527-546.

Ziliak, S. T., & McCloskey, D. N. (2004a). Size Matters: The Standard Error of Regressions in the American Economic Review. *Journal of Socio-Economics 33(5)* , 527-546.

Ziliak, S. T., & McCloskey, D. N. (2008). *The cult of statistical significance: how the standard error costs us jobs, justice, and lives,.* The University of Michigan Press.

Ziliak, S. T., & McCloskey, D. N. (2013). We Agree That Statistical Significance Proves Essientially Nothing: A Rejoinder to Thomas Mayer. *Econ Journal Watch 10(1)* , 97-107.