# Coefficients of determination *

Jean-Marie Dufour †
McGill University

First version: March 1983
Revised: February 2002, July 2011
This version: July 2011
Compiled: November 21, 2011, 11:05

---

*This work was supported by the William Dow Chair in Political Economy (McGill University), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

# Contents


**1. Coefficient of determination: $R^2$** **1**

**2. Significance tests and $R^2$** **3**

2.1. Relation of $R^2$ with a Fisher statistic . . . 3

2.2. General relation between $R^2$ and Fisher tests . . . 4

**3. Uncentered coefficient of determination: $\widetilde{R}^2$** **5**

**4. Adjusted coefficient of determination: $\overline{R}^2$** **5**

4.1. Definition and basic properties . . . 5

4.2. Criterion for $\overline{R}^2$ increase through the omission of an explanatory variable . . . 7

4.3. Generalized criterion for $\overline{R}^2$ increase through the imposition of linear constraints 8

**5. Notes on bibliography** **10**

**6. Chronological list of references** **10**

# 1. Coefficient of determination: $R^2$

Let $y = X\beta + \varepsilon$ be a model that satisfies the assumptions of the classical linear model, where $y$ and $\varepsilon$ are $T \times 1$ vectors, $X$ is a $T \times k$ matrix and $\beta$ is $k \times 1$ coefficient vector. We wish to characterize to which extent the variables included in $X$ (excluding the constant, if there is one) explain $y$.

A first method consists in computing $R^2$, the "coefficient of determination", or $R = \sqrt{R^2}$, the "coefficient of multiple correlation". Let

$$\begin{aligned}
\hat{y} &= X\hat{\beta}\ ,\ \hat{\varepsilon} = y - \hat{y}\ ,\ \bar{y} = \sum_{t=1}^{T} y_t/T = i'y/T\ , && (1.1)\\
i &= (1,1,\ldots,1)' \text{ the unit vector of dimension } T, && (1.2)\\
SST &= \sum_{t=1}^{T} (y_t - \bar{y})^2 = (y - i\bar{y})'(y - i\bar{y})\ ,\ \text{(total sum of squares)} && (1.3)\\
SSR &= \sum_{t=1}^{T} (\hat{y}_t - \bar{y})^2 = (\hat{y} - i\bar{y})'(\hat{y} - i\bar{y})\ ,\ \text{(regression sum of squares)} && (1.4)\\
SSE &= \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 = (y - \hat{y})'(y - \hat{y}) = \hat{\varepsilon}'\hat{\varepsilon}\ ,\ \text{(error sum of squares).} && (1.5)
\end{aligned}$$

We can then define "variance estimators" as follows:

$$\begin{aligned}
\hat{V}(y) &= SST/T\ , && (1.6)\\
\hat{V}(\hat{y}) &= SSR/T\ , && (1.7)\\
\hat{V}(\varepsilon) &= SSE/T\ . && (1.8)
\end{aligned}$$

**1.1 Definition** $R^2 = 1 - \left(\hat{V}(\varepsilon)/\hat{V}(y)\right) = 1 - (SSE/SST)\ .$

**1.2 Proposition** $R^2 \leq 1\ .$

PROOF This result is immediate on observing that $SSE/SST \geq 0$. □

**1.3 Lemma** $y'y = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}\ .$

PROOF We have

$$y = \hat{y} + \hat{\varepsilon} \text{ and } \hat{y}'\hat{\varepsilon} = \hat{\varepsilon}'\hat{y} = 0, \tag{1.9}$$

hence

$$y'y = (\hat{y} + \hat{\varepsilon})'(\hat{y} + \hat{\varepsilon}) = \hat{y}'\hat{y} + \hat{y}'\hat{\varepsilon} + \hat{\varepsilon}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon} = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}\ .$$

□

**1.4 Proposition** *If one of the regressors is a constant, then*

$$\begin{aligned} SST &= SSR + SSE\,, \\ \hat{V}(y) &= \hat{V}(\hat{y}) + \hat{V}(\varepsilon)\ . \end{aligned}$$

PROOF Let $A = I_T - i(i'i)^{-1} i' = I_T - \frac{1}{T} ii'$ . Then, $A'A = A$ and

$$Ay = \left[ I_T - \frac{1}{T} ii' \right] y = y - i\bar{y}.$$

If one of the regressors is a constant, we have

$$i'\hat{\varepsilon} = \sum_{t=1}^{T} \hat{\varepsilon}_t = 0$$

hence

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T} \hat{y}_t &= \frac{1}{T} i'\hat{y} = \frac{1}{T} i'(y - \hat{\varepsilon}) = \frac{1}{T} i'y = \bar{y}\,, \\ A\hat{\varepsilon} &= \hat{\varepsilon} - \frac{1}{T} ii'\hat{\varepsilon} = \hat{\varepsilon}\ , \\ A\hat{y} &= \hat{y} - \frac{1}{T} ii'\hat{y} = \hat{y} - i\bar{y}\ , \end{aligned}$$

and, using the fact that $A\hat{\varepsilon} = \hat{\varepsilon}$ and $\hat{y}'\hat{\varepsilon} = 0$,

$$\begin{aligned} SST &= (y - i\bar{y})'(y - i\bar{y}) = y'A'Ay = y'Ay \\ &= (\hat{y} + \hat{\varepsilon})' A (\hat{y} + \hat{\varepsilon}) \\ &= \hat{y}'A\hat{y} + \hat{y}'A\hat{\varepsilon} + \hat{y}'A\hat{\varepsilon} + \hat{\varepsilon}'A\hat{\varepsilon} \\ &= \hat{y}'A\hat{y} + \hat{\varepsilon}'\hat{\varepsilon} \\ &= (A\hat{y})'(A\hat{y}) + \hat{\varepsilon}'\hat{\varepsilon} = SSR + SSE\,. \end{aligned}$$

□

**1.5 Proposition** *If one of the regressors is a constant,*

$$R^2 = \frac{\hat{V}(\hat{y})}{\hat{V}(y)} = \frac{SSR}{SST} \quad \textit{and} \quad 0 \le R^2 \le 1\,.$$

PROOF By the definition of $R^2$, we have $R^2 \le 1$ and

$$R^2 = 1 - \frac{\hat{V}(\varepsilon)}{\hat{V}(y)} = \frac{\hat{V}(y) - \hat{V}(\varepsilon)}{\hat{V}(y)} = \frac{\hat{V}(\hat{y})}{\hat{V}(y)} = \frac{SSR}{SST}$$

hence $R^2 \geq 0$. $\square$

**1.6 Proposition** *If one of the regressors is a constant, the empirical correlation between $y$ and $\hat{y}$ is non-negative and equal to $\sqrt{R^2}$.*

PROOF The empirical correlation between $y$ and $\hat{y}$ is defined by

$$\hat{\rho}(y,\hat{y}) = \frac{\hat{C}(y,\hat{y})}{\left[\hat{V}(y)\hat{V}(\hat{y})\right]^{1/2}}$$

where

$$\hat{C}(y,\hat{y}) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})(\hat{y}_t - \bar{y}) = \frac{1}{T}(Ay)'(A\hat{y})$$

and $A = I_T - \frac{1}{T}ii'$. Since one of the regressors is a constant,

$$A\hat{\varepsilon} = \hat{\varepsilon}\ ,\ Ay = A\hat{y} + \hat{\varepsilon}\ ,\ \hat{\varepsilon}'(A\hat{y}) = \hat{\varepsilon}'\hat{y} = 0$$

and

$$\begin{aligned}
\hat{C}(y,\hat{y}) &= \frac{1}{T}(A\hat{y}+\hat{\varepsilon})'(A\hat{y}) = \frac{1}{T}(A\hat{y})'(A\hat{y}) = \hat{V}(\hat{y})\ , \\
\hat{\rho}(y,\hat{y}) &= \frac{\hat{V}(\hat{y})}{\left[\hat{V}(y)\hat{V}(\hat{y})\right]^{1/2}} = \left[\frac{\hat{V}(\hat{y})}{\hat{V}(y)}\right]^{1/2} = \sqrt{R^2} \geq 0\ .
\end{aligned}$$

$\square$

# 2. Significance tests and $R^2$

## 2.1. Relation of $R^2$ with a Fisher statistic

$R^2$ is descriptive statistic which measures the proportion of the "variance" of the dependent variable $y$ explained by suggested explanatory variables (excluding the constant). However, $R^2$ can be related to a significance test (under the assumptions of the Gaussian classical linear model).

Consider the model

$$y_t = \beta_1 + \beta_2 X_{t2} + \cdots + \beta_k X_{tk} + \varepsilon_t\ ,\ t = 1, \ldots, T.$$

We wish to test the hypothesis that none of these variables (excluding the constant) should appear in the equation:

$$H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0\ .$$

The Fisher statistic for $H_0$ is

$$F = \frac{(S_\omega - S_\Omega)/q}{S_\Omega/(T-k)} \sim F(q, T-k)$$

where $q = k-1$, $S_\Omega$ is the error sum of squares from the estimation of the unconstrained model

$$\Omega : y = X\beta + \varepsilon \ ,$$

where $X = [i, X_2, \ldots, X_k]$ and $S_\omega$ s the error sum of squares from the estimation of the constrained model

$$\omega : y = i\beta_1 + \varepsilon \ ,$$

where $i = (1, 1, \ldots, 1)'$. We see easily that

$$\begin{aligned} S_\Omega &= \left(y - X\hat{\beta}\right)'\left(y - X\hat{\beta}\right) = SSE \ , \\ \hat{\beta}_1 &= \left(i'i\right)^{-1} i'y = \frac{1}{T}\sum_{t=1}^{T} y_t = \bar{y} \ , \text{ (under } \omega) \\ S_\omega &= (y - i\bar{y})'(y - i\bar{y}) = SST \end{aligned}$$

and

$$\begin{aligned} F &= \frac{(SST - SSE)/(k-1)}{SSE/(T-k)} = \frac{\left[1 - \frac{SSE}{SST}\right]/(k-1)}{\frac{SSE}{SST}/(T-k)} \\ &= \frac{R^2/(k-1)}{(1-R^2)/(T-k)} \sim F(k-1, T-k) \ . \end{aligned}$$

As $R^2$ increases, $F$ increases.

## 2.2. General relation between $R^2$ and Fisher tests

Consider the general linear hypothesis

$$H_0 : C\beta = r$$

where $C : q \times k$, $\beta : k \times 1$, $r : q \times 1$ and $\text{rank}(C) = q$. The values of $R^2$ for the constrained and unconstrained models are respectively:

$$R_0^2 = 1 - \frac{S_\omega}{SST} \ , \ R_1^2 = 1 - \frac{S_\Omega}{SST},$$

hence

$$S_\omega = \left(1 - R_0^2\right) SST \ , \ S_\Omega = \left(1 - R_1^2\right) SST \ .$$

The Fisher statistic for testing $H_0$ may thus be written

$$\begin{aligned} F &= \frac{(S_\omega - S_\Omega)/q}{S_\Omega/(T-k)} = \frac{\left(R_1^2 - R_0^2\right)/q}{\left(1-R_1^2\right)/(T-k)} \\ &= \left(\frac{T-k}{q}\right)\frac{R_1^2 - R_0^2}{1-R_1^2}. \end{aligned}$$

If $R_1^2 - R_0^2$ is large, we tend to reject $H_0$. If $H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0$, then

$$q = k-1 \; , \; S_\omega = SST \; , \; R_0^2 = 0$$

and the formula for $F$ above gets reduced of the one given in section 2.1.

# 3. Uncentered coefficient of determination: $\widetilde{R}^2$

Since $R^2$ can take negative values when the model does not contain a constant, $R^2$ has little meaning in this case. In such situations, we can instead use a coefficient where the values of $y_t$ are not centered around the mean.

**3.1 Definition** $\widetilde{R}^2 = 1 - \left(\hat{\varepsilon}'\hat{\varepsilon}/y'y\right)$ .

$\tilde{R}^2$ is called the "uncentered coefficient of determination" on "uncentered $R^2$" and $\tilde{R} = \sqrt{\tilde{R}^2}$ the "uncentered coefficient of multiple correlation".

**3.2 Proposition** $0 \leq \tilde{R}^2 \leq 1$ .

PROOF This follows directly from Lemma **1.3**: $y'y = \hat{y}'\hat{y} + \hat{\varepsilon}'\hat{\varepsilon}$ . □

# 4. Adjusted coefficient of determination: $\overline{R}^2$

## 4.1. Definition and basic properties

An unattractive property of the $R^2$ coefficient comes form the fact that $R^2$ cannot decrease when explanatory variables are added to the model, even if these have no relevance. Consequently, choosing to maximize $R^2$ can be misleading. It seems desirable to penalize models that contain too many variables.

Since

$$R^2 = 1 - \frac{\hat{V}(\varepsilon)}{\hat{V}(y)} \; ,$$

where

$$\hat{V}(\varepsilon) = \frac{SSE}{T} = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 \, , \; \hat{V}(y) = \frac{SST}{T} = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^2 \; ,$$

Theil (1961, p. 213) suggested to replace $\hat{V}(\varepsilon)$ and $\hat{V}(y)$ by "unbiased estimators":

$$\begin{aligned} s^2 &= \frac{SSE}{T-k} = \frac{1}{T-k}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 \, , \\ s_y^2 &= \frac{SST}{T-1} = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \bar{y})^2 \, . \end{aligned}$$

**4.1 Definition** *$R^2$ adjusted for degrees of freedom is defined by*

$$\overline{R}^2 = 1 - \frac{s^2}{s_y^2} = 1 - \frac{T-1}{T-k}\left(\frac{SSE}{SST}\right) .$$

**4.2 Proposition** $\overline{R}^2 = 1 - \frac{T-1}{T-k}\left(1 - R^2\right) = R^2 - \frac{k-1}{T-k}\left(1 - R^2\right) \; .$

PROOF

$$\begin{aligned} \overline{R}^2 &= 1 - \frac{T-1}{T-k}\left(\frac{SSE}{SST}\right) = 1 - \frac{T-1}{T-k}\left(1 - R^2\right) \\ &= 1 - \frac{T-k+k-1}{T-k}\left(1 - R^2\right) = 1 - \left(1 + \frac{k-1}{T-k}\right)\left(1 - R^2\right) \\ &= 1 - \left(1 - R^2\right) - \frac{k-1}{T-k}\left(1 - R^2\right) = R^2 - \frac{k-1}{T-k}\left(1 - R^2\right) \; . \quad Q.E.D. \end{aligned}$$

□

**4.3 Proposition** $\overline{R}^2 \leq R^2 \leq 1 \, .$

PROOF The result follows from the fact that $1 - R^2 \geq 0$ and **(4.2)**. □

**4.4 Proposition** $\overline{R}^2 = R^2 \quad iff \quad (k = 1 \quad or \quad R^2 = 1) \; .$

**4.5 Proposition** $\overline{R}^2 \leq 0 \quad iff \; R^2 \leq \frac{k-1}{T-1} \; .$

$\overline{R}^2$ can be negative even if $R^2 \geq 0$. If the number of explanatory variables is increased, $R^2$ and $k$ both increase, so that $\overline{R}^2$ can increase or decrease.

**4.6 Remark** When several models are compared on the basis of $R^2$ or $\overline{R}^2$, it is important to have the same dependent variable. When the dependent variable $(y)$ is the same, maximizing $\overline{R}^2$ is equivalent to minimizing the standard error of the regression

$$s = \left[\frac{1}{T-k}\sum_{t=1}^{T}\hat{\varepsilon}_t^2\right]^{1/2} .$$

## 4.2. Criterion for $\overline{R}^2$ increase through the omission of an explanatory variable

Consider the two models:

$$y_t = \beta_1 X_{t1} + \cdots + \beta_{k-1} X_{t(k-1)} + \varepsilon_t \quad , \, t = 1, \ldots, T, \tag{4.1}$$

$$y_t = \beta_1 X_{t1}, + \cdots + \beta_{k-1} X_{t(k-1)} + \beta_k X_{tk} + \varepsilon_t \quad , t = 1, \ldots, T. \tag{4.2}$$

We can then show that the value of $\overline{R}^2$ associated with the restricted model (4.1) is larger than the one of model (4.2) if the $t$ statistic for testing $\beta_k = 0$ is smaller than 1 (in absolute value).

**4.7 Proposition** *If $\overline{R}_{k-1}^2$ and $\overline{R}_k^2$ are the values of $\overline{R}^2$ for models $(4.1)$ and $(4.2)$, then*

$$\overline{R}_k^2 - \overline{R}_{k-1}^2 = \frac{\left(1-\overline{R}_k^2\right)}{(T-k+1)}\left(t_k^2-1\right) \tag{4.3}$$

*where $t_k$ is the Student $t$ statistic for testing $\beta_k = 0$ in model $(4.2)$, and*

$$\overline{R}_k^2 \le \overline{R}_{k-1}^2 \quad iff \quad t_k^2 \le 1 \quad iff \quad |t_k| \le 1 \, .$$

*If furthermore $\overline{R}_k^2 < 1$, then*

$$\overline{R}_k^2 \lesseqgtr \overline{R}_{k-1}^2 \quad iff \quad |t_k| \lesseqgtr 1 \, .$$

PROOF By definition,

$$\overline{R}_k^2 = 1 - \frac{s_k^2}{s_y^2} \quad \text{and} \quad \overline{R}_{k-1}^2 = 1 - \frac{s_{k-1}^2}{s_y^2}$$

where $s_k^2 = SS_k/\,(T-k)$ and $s_{k-1}^2 = SS_{k-1}/\,(T-k+1)$. $SS_k$ and $SS_{k-1}$ are the sums of squared errors for the models with $k$ and $k-1$ explanatory variables. Since $t_k^2$ is the Fisher statistic for testing $\beta_k = 0$, we have

$$t_k^2 \quad = \quad \frac{(SS_{k-1} - SS_k)}{SS_k/\,(T-k)}$$

$$
\begin{aligned}
&= \frac{\left[(T-k+1)\,s_{k-1}^2-(T-k)\,s_k^2\right]}{s_k^2}\\
&= \frac{(T-k+1)\left(1-\overline{R}_{k-1}^2\right)-(T-k)\left(1-\overline{R}_k^2\right)}{1-\overline{R}_k^2}\\
&= (T-k+1)\left(\frac{1-\overline{R}_{k-1}^2}{1-\overline{R}_k^2}\right)-(T-k)
\end{aligned}
$$

for $s_{k-1}^2 = s_y^2\left(1-\overline{R}_{k-1}^2\right)$ and $s_k^2 = s_y^2\left(1-\overline{R}_k^2\right)$. Consequently,

$$1-\overline{R}_{k-1}^2 = \left(1-\overline{R}_k^2\right)\frac{\left[t_k^2+(T-k)\right]}{T-k+1}$$

and

$$
\begin{aligned}
\overline{R}_k^2-\overline{R}_{k-1}^2 &= \left(1-\overline{R}_{k-1}^2\right)-\left(1-\overline{R}_k^2\right)\\
&= \left(1-\overline{R}_k^2\right)\left[\frac{t_k^2+(T-k)}{T-k+1}-1\right]\\
&= \left(1-\overline{R}_k^2\right)\left[\frac{t_k^2-1}{T-k+1}\right].
\end{aligned}
$$

□

### 4.3. Generalized criterion for $\overline{R}^2$ increase through the imposition of linear constraints

We will now study when the imposition of $q$ linearly independent constraints

$$H_0 : C\beta = r$$

will raise or decrease $\overline{R}^2$, where $C : q\times k$, $r : q\times 1$ and $\text{rank}(C) = q$. Let $\overline{R}_{H_0}^2$ and $\overline{R}^2$ be the values of $\overline{R}^2$ for the constrained (by $H_0$) and unconstrained models, similarly, $s_0^2$ and $s^2$ are the values of the corresponding unbiased estimators of the error variance.

**4.8 Proposition** *Let $F$ be the Fisher statistic for testing $H_0$. Then*

$$s_0^2 - s^2 = \frac{qs^2}{T-k+q}(F-1)$$

*and*

$$s_0^2 \lesseqqgtr s^2 \quad iff \quad F \lesseqqgtr 1 \, .$$

PROOF If $SS_0$ and $SS$ are the sum of squared errors for the constrained and unconstrained models, we have:

$$s_0^2 = \frac{SS_0}{T-k+q} \quad \text{and} \quad s^2 = \frac{SS}{T-k} \, .$$

The $F$ statistic may then be written

$$\begin{aligned} F &= \frac{(SS_0 - SS)/q}{SS/(T-k)} \\ &= \frac{\left[(T-k+q)\, s_0^2 - (T-k)\, s^2\right]}{qs^2} = \frac{T-k+q}{q}\left(\frac{s_0^2}{s^2}\right) - \frac{T-k}{q} \end{aligned}$$

hence

$$\begin{aligned} s_0^2 &= s^2 \frac{[qF + (T-k)]}{(T-k)+q} \, , \\ s_0^2 - s^2 &= s^2 \frac{q\,(F-1)}{(T-k)+q} \, , \end{aligned}$$

and

$$s_0^2 \lesseqqgtr s^2 \quad \text{iff} \quad F \lesseqqgtr 1 \, .$$

□

**4.9 Proposition** *Let $F$ be the Fisher statistic for testing $H_0$. Then*

$$\bar{R}^2 - \bar{R}_{H_0}^2 = \frac{q\left(1 - \bar{R}^2\right)}{T-k+q}\,(F-1)$$

*and*

$$\bar{R}_{H_0}^2 \gtreqqless \bar{R}^2 \quad iff \quad F \lesseqqgtr 1 \, .$$

PROOF By definition,

$$\bar{R}_{H_0}^2 = 1 - \frac{s_0^2}{s_y^2} \, , \, \bar{R}^2 = 1 - \frac{s^2}{s_y^2} \, .$$

Thus,

$$\begin{aligned} \bar{R}^2 - \bar{R}_{H_0}^2 &= \frac{s^2 - s_0^2}{s_y^2} \\ &= \frac{q}{T-k+q}\left(\frac{s^2}{s_y^2}\right)(F-1) \end{aligned}$$

$$= \frac{q\left(1-\overline{R}^2\right)}{T-k+q}(F-1)$$

hence

$$\overline{R}^2_{H_0} \gtreqless \overline{R}^2 \quad \text{iff} \quad F \lesseqgtr 1 \ .$$

$\square$

On taking $q = 1$, we get property (4.3). If we test an hypothesis of the type

$$H_0 : \beta_k = \beta_{k+1} = \cdots = \beta_{k+l} = 0 \ ,$$

it is possible that $F > 1$, while all the statistics $|t_i|$ , $i = k, \ldots, k+l$ are smaller than 1. This means that $\overline{R}^2$ increases when we omit one explanatory variable at a time, but decreases when they are all excluded from the regression. Further, it is also possible that $F < 1$, but $|t_i| > 1$ for all $i$: $\overline{R}^2$ increases when all the explanatory variables are simultaneously excluded, but decreases when only one is excluded.

## 5. Notes on bibliography

The notion of $\overline{R}^2$ was proposed by Theil (1961, p. 213). Several authors have presented detailed discussions of the different concepts of multiple correlation: for example, Theil (1971, Chap. 4), Schmidt (1976) and Maddala (1977, Sections 8.1, 8.2, 8.3, 8.9). The $\overline{R}^2$ concept is criticized by Pesaran (1974). The mean and bias of $R^2$ were studied by Cramer (1987) in the Gaussian case, and by Srivastava, Srivastava and Ullah (1995) in some non-Gaussian cases.

## 6. Chronological list of references


1. Theil (1961, p. 213) _ The $\overline{R}^2$ nation was proposed in this book.
2. Theil (1971, Chap. 4) _ Detailed discussion of $R^2$, $\overline{R}^2$ and partial correlation.
3. Pesaran (1974) _ Critique of $\overline{R}^2$.
4. Schmidt (1976)
5. Maddala (1977, Sections 8.1, 8.2, 8.3, 8.9) _ Discussion of $R^2$and $\overline{R}^2$along with their relation with hypothesis tests.
6. Hendry and Marshall (1983)
7. Cramer (1987)
8. Ohtani and Hasegawa (1993)

9. Srivastava et al. (1995)

## References


Cramer, J. S. (1987), 'Mean and variance of $R^2$ in small and moderate samples', *Econometric Reviews* **35**, 253–266.

Hendry, D. F. and Marshall, R. C. (1983), 'On high and low $R^2$ contributions', *Oxford Bulletin of Economics and Statistics* **45**, 313–316.

Maddala, G. S. (1977), *Econometrics*, McGraw-Hill, New York.

Ohtani, L. and Hasegawa, H. (1993), 'On small-sample properties of $R^2$ in a linear regression model with multivariate $t$ errors and proxy variables', *Econometric Theory* **9**, 504–515.

Pesaran, M. H. (1974), 'On the general problem of model selection', *Review of Economic Studies* **41**, 153–171.

Schmidt, P. (1976), *Econometrics*, Marcel Dekker, New York.

Srivastava, A. K., Srivastava, V. K. and Ullah, A. (1995), 'The coefficient of determination and its adjusted version in linear regression models', *Econometric Reviews* **14**, 229–240.

Theil, H. (1961), *Economic Forecasts and Policy, 2nd Edition*, North-Holland, Amsterdam.

Theil, H. (1971), *Principles of Econometrics*, John Wiley & Sons, New York.