Università di Pavia

Specificazione funzione di investimento

Eduardo Rossi

	Variabili
I_t	investimento reale
i_t	tasso d'interesse nominale
$\Delta \log p_t$	tasso d'inflazione
Y_t	prodotto reale
t	trend

Dati: 1950.1 – 2000.4 204 osservazioni.

$$\log I_t = \beta_1 + \beta_2 i_t + \beta_3 \Delta p_t + \beta_4 \log Y_t + \beta_5 t + \varepsilon_t \quad t = 1, \ldots, N$$

il modello dice che gli investitori sono sensibili al tasso d'interesse nominale, i_t, al tasso d'inflazione, Δp_t, all'output reale ($\log Y_t$), e ad altri fattori che hanno un trend crescente.

Stime OLS con 203 osservazioni 1950 : 2 – 2000 : 4
Variabile Dipendente: $\log I_t$

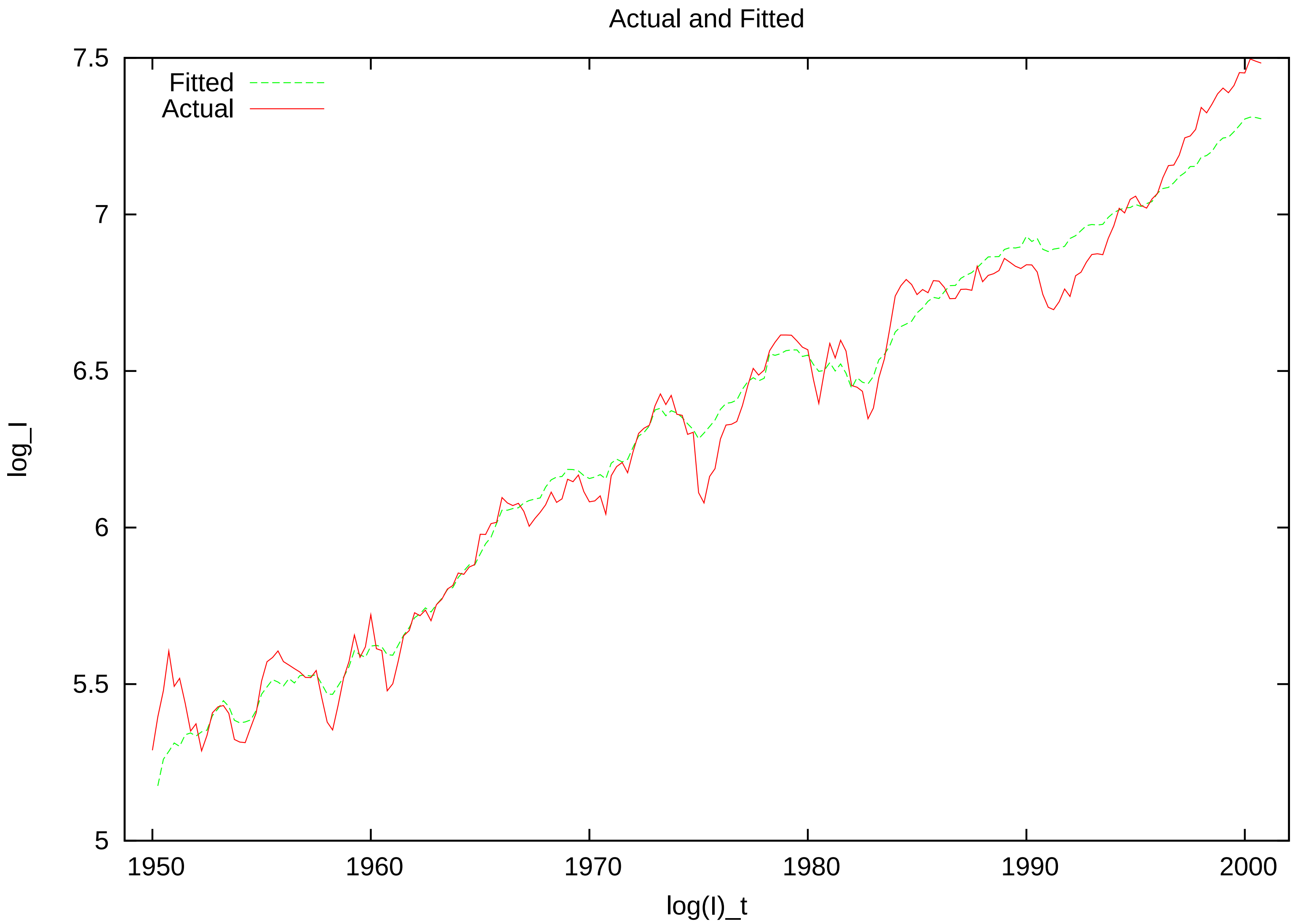
Variabile	Stima	Std. Error	t	p-value
const	-9,1284	1,3650	-6,6876	0,0000
i_t	-0,0086	0,0032	-2,6906	0,0077
Δp_t	0,0033	0,0023	1,415	0,1587
$\log Y_t$	1,930	0,183	10,532	0,000
t	-0,0057	0,0015	-3,8030	0,0002

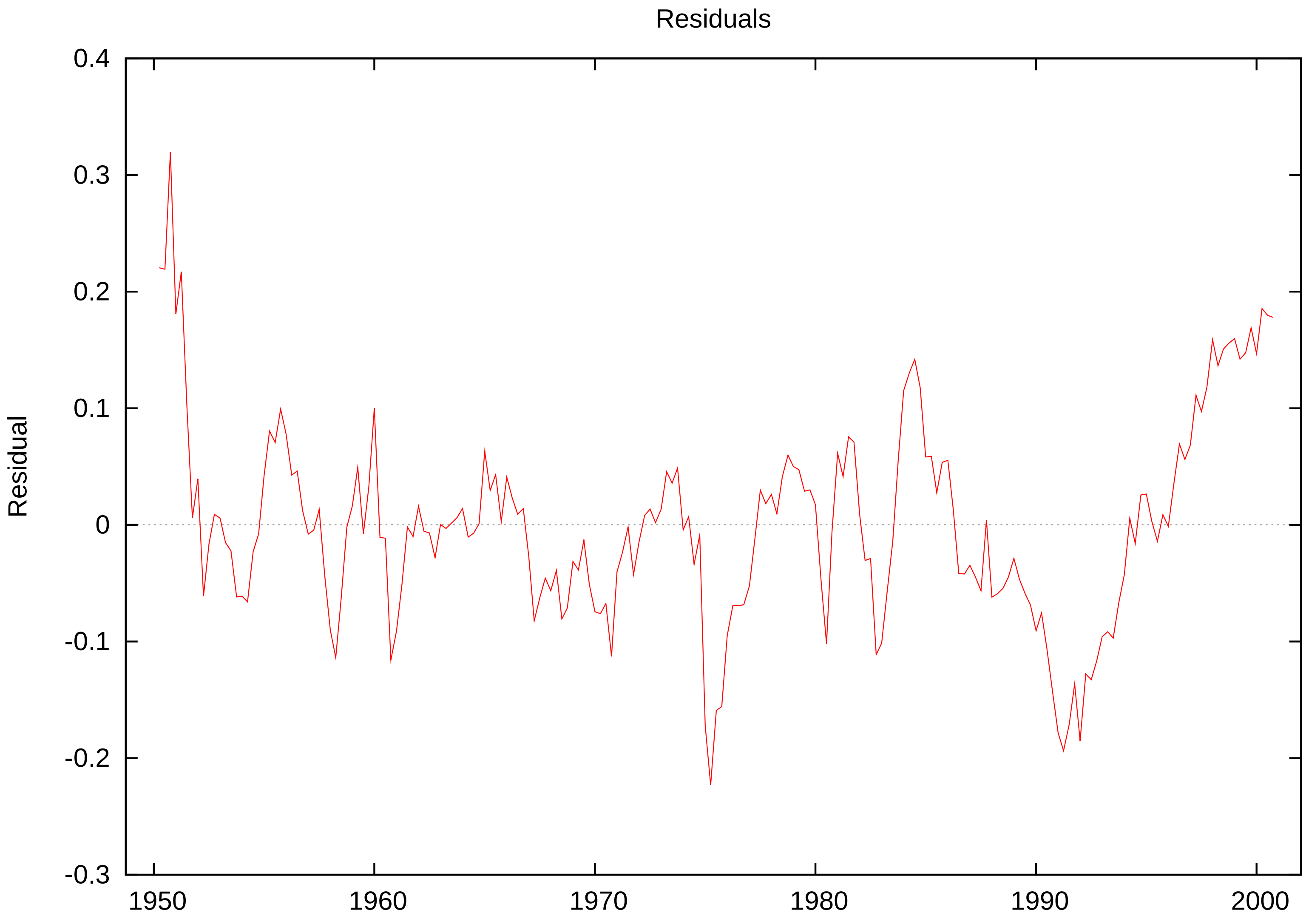
Media campionaria della y	6,30947
Deviazione standard della y	0,599625
SSR	1,4706
Standard error dei residui (s)	0,0861806
R^2	0,979752
$\bar{R}^2$	0,979343
$F(4.198)$	2395,23
Statistica Durbin-Watson	0,213960

t(198, .025) = 1,972

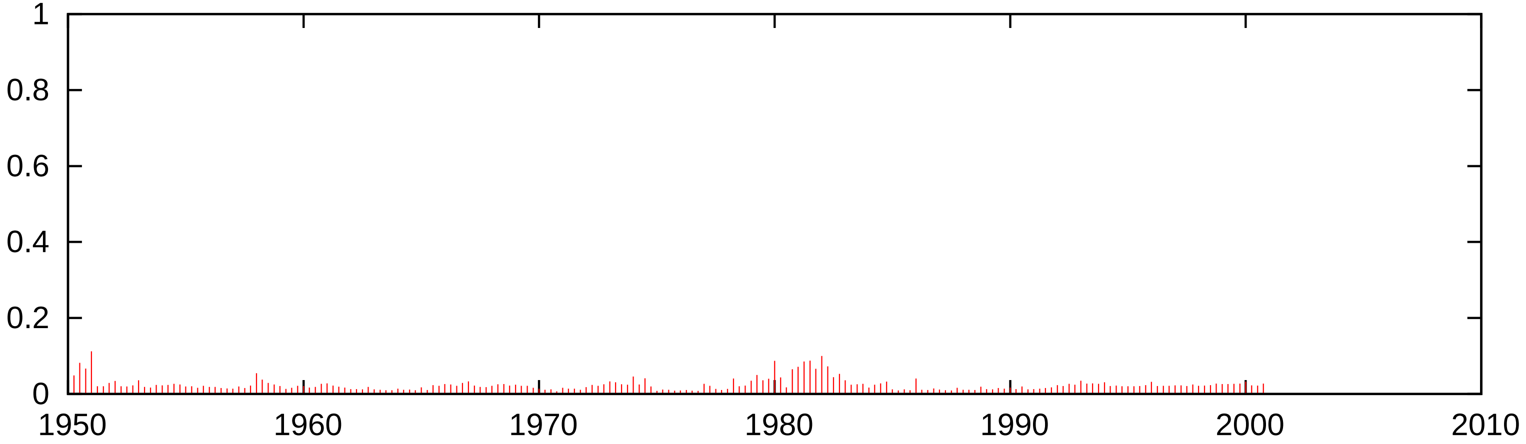
Variabile	Stima	Intervallo di confidenza 95%
const	-9,12843	(-11,8202;-6,4367)
i_t	-0,00859785	(-0,0149;-0,0022963)
$\log Y_t$	1,93016	(1,5687; 2,2916)
Δp_t	0,00330	(-0,5200;3,1659)
t	-0,00565889	(-0,008593;-0,002725)

Modello investimentto

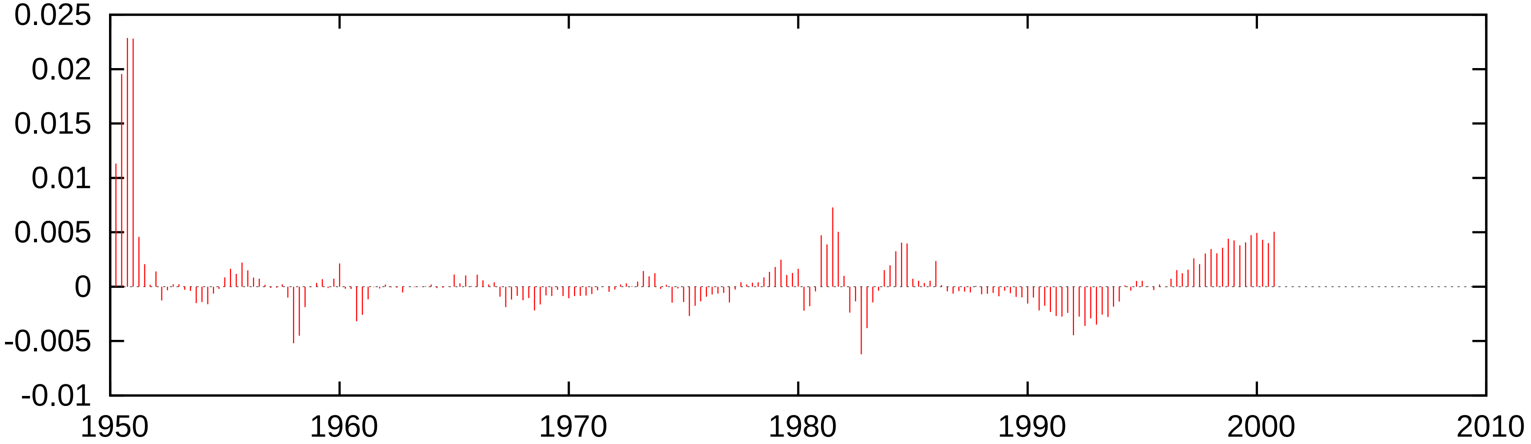


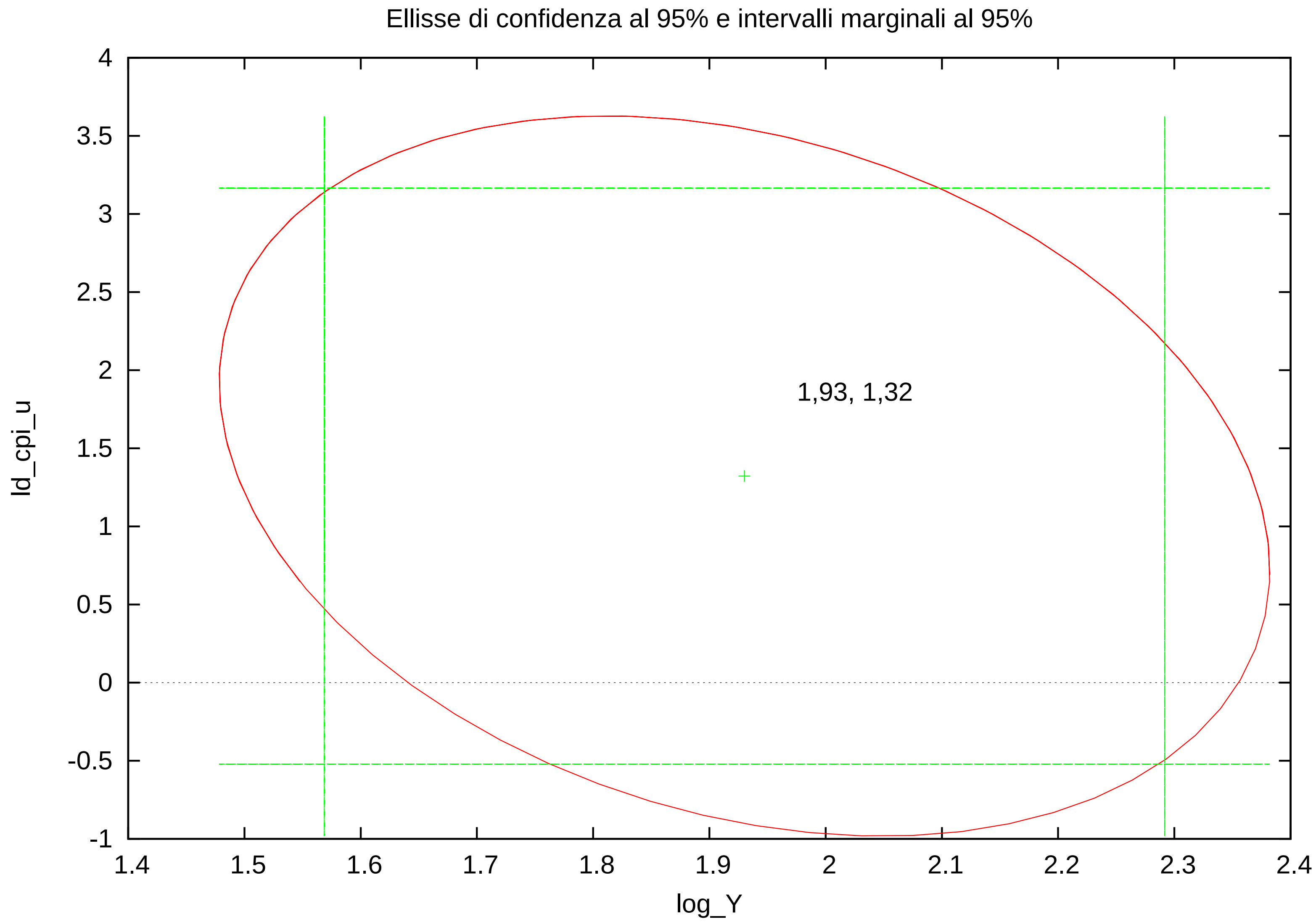

Residuals
Residual
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
1950
1960
1970
1980
1990
2000

Leverage



Influenza




Ellisse di confidenza al 95% e intervalli marginali al 95%
1,93, 1,32
ld_cpi_u
log_Y
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
1.4
1.5
1.6
1.7
1.8
1.9
2
2.1
2.2
2.3
2.4

Se gli investitori rispondono solo a variazioni dei tassi d'interesse reali, nell'equazione

$$\log I_t = \beta_1 + \beta_2 i_t + \beta_3 \Delta p_t + \beta_4 \log Y_t + \beta_5 t + \varepsilon_t \quad t = 1, \ldots, N$$

ciò implica che

$$\beta_2 + \beta_3 = 0.$$

Dobbiamo verificare

$$H_0 : \beta_2 + \beta_3 = 0$$

$$H_1 : \beta_2 + \beta_3 \neq 0$$

$$\mathbf{R}\widehat{\boldsymbol{\beta}} = [0,1,1,0,0]\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_5 \end{bmatrix}$$

$$\begin{aligned} \widehat{F} &= \frac{(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{c})'[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\widehat{\boldsymbol{\beta}} - \mathbf{c})}{s^2} \\ &= \frac{(\widehat{\beta}_2 + \widehat{\beta}_3)[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\widehat{\beta}_2 + \widehat{\beta}_3)}{s^2} \end{aligned}$$

$$\widehat{Var(\mathbf{R}\widehat{\boldsymbol{\beta}}|\mathbf{X})} = s^2[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']$$

$$\text{s.e.}(\widehat{\beta}_2 + \widehat{\beta}_3) = \left(\widehat{Var(\mathbf{R}\widehat{\boldsymbol{\beta}}|\mathbf{X})}\right)^{1/2} = \left(s^2[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']\right)^{1/2}$$

$$\text{s.e.}(\widehat{\beta}_2 + \widehat{\beta}_3) = \left(\widehat{Var[\widehat{\beta}_2]} + \widehat{Var[\widehat{\beta}_3]} + 2\widehat{Cov[\widehat{\beta}_2, \widehat{\beta}_2]}\right)^{1/2}$$

$$\widehat{F} = \frac{(\widehat{\beta}_2 + \widehat{\beta}_3)^2}{\left(\widehat{Var}[\widehat{\beta}_2] + \widehat{Var}[\widehat{\beta}_3] + 2\widehat{Cov}[\widehat{\beta}_2, \widehat{\beta}_2]\right)} \sim F_{(1,198)}$$

$$\widehat{t} = \widehat{F}^{1/2} = \frac{(\widehat{\beta}_2 + \widehat{\beta}_3)}{\sqrt{\widehat{Var}[\widehat{\beta}_2] + \widehat{Var}[\widehat{\beta}_3] + 2\widehat{Cov}[\widehat{\beta}_2, \widehat{\beta}_2]}} \sim t_{(198)}$$

with $\widehat{Cov}[\widehat{\beta}_2, \widehat{\beta}_2] = -3.718 \times 10^{-6}$

$$\begin{aligned} \widehat{s.e.(\widehat{\beta}_2, \widehat{\beta}_3)} &= (0,00319^2 + 0,00234^2 + 2 \times (-3,718 \times 10^{-6}))^{1/2} \\ &= 0,002866 \end{aligned}$$

$$\widehat{t} = \frac{(-0,00860 + 0,00331)}{0,002866} = -1,845$$

Stime ristrette

Variabile	Stima	Std. Error	t	p-value	
const	$-7,90275$	$1,19931$	$-6,589$	$< 0,00001$	***
i_t	$-0,00442654$	$0,00227018$	$-1,950$	$0,05260$	*
Δp_t	$0,00442654$	$0,00227018$	$1,950$	$0,05260$	*
$\log Y_t$	$1,76406$	$0,160561$	$10,987$	$< 0,00001$	***
t	$-0,00440260$	$0,00133078$	$-3,308$	$0,00111$	***

Standard error dei residui $= 0,0866985$.
$\widetilde{\boldsymbol{\epsilon}}'\widetilde{\boldsymbol{\epsilon}} = (0,0866985)^2 \times (203-4)$, $\widehat{F} = 3,39913$, p-value $= 0,066725$
$F(1,198)$.

Riparametrizzazione del modello in modo tale che la statistica t prodotta dal software sia usata per verificare un'ipotesi congiunta. In questo caso,

$$\log I_t = \beta_1 + \beta_2(i_t - \Delta p_t) + \beta_3 \Delta p_t + \beta_4 \log Y_t + \beta_5 t + \varepsilon_t \quad t = 1, \ldots, N$$

in questa regressione

$$H_0 : \beta_3 = 0$$

corrisponde alla nulla nel modello precedente

$$\beta_2 + \beta_3 = 0$$

Variabile	Stime	Std.error	t	p-value	
const	-9,12885	1,36496	-6,688	$< 0,00001$	***
$\log Y_t$	1,93021	0,18327	10,532	$< 0,00001$	***
Δp_t	-0,00529	0,00287	-1,844	0,06663	*
t	-0,00566	0,001488	-3,803	0,00019	***
$i_t - \Delta \log p_t$	-0,00859	0,0031954	-2,692	0,00772	***

$\widehat{t}(\widehat{\beta_3}) = -1,844$

Infine, consideriamo le ipotesi congiunte:

$$\begin{aligned} \beta_2 + \beta_3 &= 0 \\ \beta_4 &= 1 \\ \beta_5 &= 0 \end{aligned}$$

Variabile	Stime	Std.error	t	p-value	
const	-2,01646	0,0107646	-187,32	$< 0,00001$	***
i_t	0,00689615	0,00339908	2,029	0,04379	**
$\Delta \log p_t$	-0,00689615	0,00339908	-2,029	0,04379	**
$\log Y_t$	1	0			
t	0	0			

$F(3, 198) = 109,841,$ con $p\text{-}value = 6,59341e - 042.$

I ricercatori usano le procedure di test per sepcificare gli elementi della matrice $\mathbf{X}$ basandosi su criteri statistici.
Le specificazioni del modello condotte in questo modo non sono giustificate.
Applicazione di test statistici in modo informale
Procedura *pretest*

1. Scegliamo un livello di significatività α
2. Calcoliamo $\widehat{\boldsymbol{\beta}}$
3. Calcoliamo il rapporto $\widehat{F}$ per $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{c}$
4. Stimiamo $\boldsymbol{\beta}$ con

$$\widehat{\boldsymbol{\beta}}_\alpha = \begin{cases} \widehat{\boldsymbol{\beta}} & \widehat{F} > F^*_{(r,N-K),\alpha} \\ \widetilde{\boldsymbol{\beta}} & \widehat{F} \leq F^*_{(r,N-K),\alpha} \end{cases} \qquad \widehat{\boldsymbol{\beta}}_\alpha = \widetilde{\boldsymbol{\beta}} + \mathbf{1}\{\widehat{F} > F^*_{(r,N-K),\alpha}\}(\widehat{\boldsymbol{\beta}} - \widetilde{\boldsymbol{\beta}})$$

Lo stimatore *pretest* non è uno stimatore lineare di $\mathbf{y}$. La teoria della distribuzione è analiticamente difficile. La stima pretest è distorta, la sua matrice di varianza è stimata in modo scorretto e la distribuzione non è normale.

1. Calcoliamo stime iniziali OLS ed esaminiamo le varie statistiche (R^2, statistiche t, segni e dimensioni dei coefficienti stimati).

2. Scartiamo quelle stime con R^2 troppo bassi, statistiche t troppo piccole, o con segni e dimensioni dei coefficienti stimati implausibili.

3. Si provano nuove stime, con nuove variabili esplicative ottenute come trasformazioni non lineari

regressioni *stepwise*: sequenza di stimatori *pretest* ottenuti aggiungendo (o rimuovendo) le variabili a seconda della loro statistica t quando le variabili sono incluse.

Il pericolo di tali procedure sequenziali è che non permettano di arrivare ad un modello *generale* bensì che finiscano per adattare il modello alle caratteristiche del campione in esame che non si verificherebbere in media se il campionamento fosse ripetuto.