Electronic Journal of Applied Statistical Analysis
EJASA (2013), Electron. J. App. Stat. Anal., Vol. 6, Issue 1, 57 – 66
e-ISSN 2070-5948, DOI 10.1285/i20705948v6n1p57

ON THE INTERPRETATION AND ESTIMATION OF THE MARKET MODEL R-SQUARE

Riccardo Bramante[*(1)], Giovanni Petrella[(2)], Diego Zappa[(1)]

[(1)]*Department of Statistical Science, Università Cattolica del Sacro Cuore, Milan, Italy*
[(2)]*Department of Management, Università Cattolica del Sacro Cuore, Milan, Italy*

Received 05 October 2012; Accepted 10 December 2012
Available online 26 April 2013

***Abstract**: The R-square of the market model is largely employed in finance and accounting studies as a measure of stock price informational efficiency. Individual firms R-squares are usually aggregated at the country-level by using the individual firm total risk over the country total risk as weighting factor. This paper shows how to interpret the country-level R-square as a Chisini mean of the firm-specific R-squares and under what conditions it may be related to the R-square of a Seemingly Unrelated Regression (SUR) model. In particular we show that for the latter a necessary constrain is that returns must be centered on zero, which appears to be in this context not only a common practice but also a methodological assumption.*

***Keywords**: R-square, market model, informational efficiency, SUR model, Chisini mean.*

1. Introduction

The R-square of the market model is largely employed in finance and accounting studies as a metric of stock price informational efficiency. The earliest studies that use the R-square as a measure of price efficiency are Morck et al.'s analysis at the country level [12] and Durnev et al.'s analysis [4] at the firm level. The authors investigate stock returns synchronicity in emerging and developed economies based on a sample of bi-weekly stock returns for 15,920 firms spanning 40 countries. [12] define the country-level R^2 as:

$$R_j^2 = \sum_i R_{i,j}^2 \cdot \frac{SST_{i,j}}{\sum_i SST_{i,j}} \tag{1}$$

[*] Email: riccardo.bramante@unicatt.it

where $SST_{i,j}$ is the sum of squared total variations for stock i in country j and $R_{i,j}^2 = SSR_{i,j}/SST_{i,j}$ measures the proportion of the variation in the bi-weekly returns of stock i in country j explained by variations in country j's market return.

Following [12], a large body of research uses the R-square as an indicator of price efficiency. A non exhaustive list of studies that employ the R-square includes [1], [2], [3], [4], [5], [8], [9], [10], [13], [14], [15], [16].

Since (1) is a weighted average of the individual stocks R-squares, it must itself be considered as an R-square. However, the weights in (1) are neither frequencies nor probabilities. Consequently, the interpretation of (1) as a coefficient of determination (i.e., as a measure of the fit of a model) must be appropriately justified and a natural question arises: does it exist a model whose R-square is equal to (1)? In this paper we address two strictly related issues: first, how to interpret the country-level R-square defined in (1) and, second, how to obtain (1) as the R-square of a regression model.

The issues we address in this paper are important both to interpret the country-level R-square in an appropriate way and to specify under what conditions (1) may be considered as an average of the R-squares of the stocks traded in a market.

2. Preliminary empirical evidence

In this Section we present some preliminary evidence about the market model R^2 estimated for the US market[1]. Specifically, our sample includes all the stocks traded on the NYSE and the NASDAQ markets. We also include dead stocks. For each common stock trading in the company's home market we collected from Thomson Reuters Datastream (*TRD*) the daily total return index (*RI* is the variable name – also known as datatype – in *TRD*), the daily adjusted price (*P*) and the market capitalization (*MV*). The sampling frequency is weekly. The data span from January 1995 through December 2011. Table 1 reports, for each year, the number of stocks and the average market capitalization for the two markets. We construct, for each year, 5 size-sorted portfolios based on the quintiles of market capitalization at the end of the previous year. Results in Table 1 are reported for both the full sample and the 5 size-sorted portfolios.

For each stock *i* and each period *t* of weekly data we run a market model regression:

$$r_{i,t} = \alpha_i + \beta_i \cdot r_{m,t} + \epsilon_{i,t} \tag{2}$$

where $r_{i,t}$ is the return for stock *i* at time *t*; $r_{m,t}$ is the return for the stock market at time *t*; $\epsilon_{i,t}$ is the residual. This procedure produces yearly estimates for market model parameters and R-square. By using (1) the stocks R-squares have been aggregated.

[1] For a more comprehensive study and details about the use of R-square as price inefficiency measure see Bramante et al. (2012)

Table 1. Sample Descriptive Statistics.

	United States - NYSE							United States - NASDAQ						
	# of	Full	Size Sorted Portfolios					# of	Full	Size Sorted Portfolios				
Year	Stocks	Sample	1	2	3	4	5	Stocks	Sample	1	2	3	4	5
1995	2,025	2,501	77	249	565	1,448	10,167	1,899	447	19	50	102	210	1,852
1996	1,928	3,210	119	366	767	1,876	12,912	2,225	532	20	56	112	239	2,239
1997	2,031	3,883	130	412	847	2,065	15,969	2,652	554	20	58	117	250	2,327
1998	2,051	4,710	98	324	740	1,935	20,461	2,821	787	15	46	96	225	3,567
1999	2,020	5,298	88	301	721	1,900	23,478	2,741	1,530	17	52	112	296	7,188
2000	1,906	5,678	82	302	807	2,172	25,040	2,838	1,185	12	37	92	265	5,515
2001	1,800	5,548	99	366	933	2,321	24,022	2,813	1,037	15	52	124	343	4,649
2002	1,814	4,462	93	331	814	2,002	19,059	2,769	795	13	45	100	255	3,569
2003	1,820	5,465	127	485	1,145	2,738	22,829	2,754	1,140	25	80	181	420	4,994
2004	1,798	6,274	160	633	1,423	3,284	25,784	2,678	1,274	31	96	214	492	5,534
2005	1,802	6,631	166	678	1,534	3,563	27,242	2,708	1,299	30	98	220	509	5,644
2006	1,765	7,632	213	823	1,810	4,024	31,300	2,710	1,388	34	111	257	576	5,960
2007	1,670	8,091	222	798	1,817	4,237	33,383	2,644	1,532	32	99	234	545	6,749
2008	1,690	4,684	95	374	947	2,270	19,786	2,604	887	13	44	121	325	3,929
2009	1,666	6,216	164	608	1,474	3,412	25,434	2,550	1,337	20	68	175	453	5,970
2010	1,649	7,104	223	777	1,836	4,095	28,575	2,517	1,579	23	84	220	572	7,001
2011	1,638	6,976	201	693	1,735	3,977	28,246	2,499	1,577	19	69	186	531	7,076

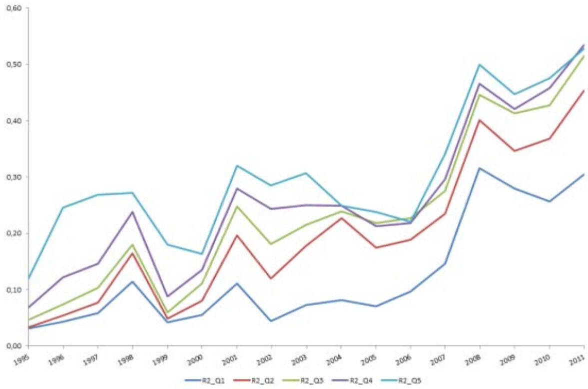
In Table 2 we report, for each market, the mean and the median of the yearly estimates of the R-square for both the full sample and the 5 size-sorted portfolios. Looking at size-sorted portfolios, both the average and the median R-squares are strictly increasing in firm size.

Table 2. R^2 by Market Capitalization in 5-size sorted portfolios.

	R^2 United States			
	NYSE		NASDAQ	
	Mean	Median	Mean	Median
Full Sample	0.222	0.203	0.143	0.105
Size Sorted Portfolios				
1 (Smallest)	0.121	0.096	0.053	0.031
2	0.194	0.183	0.077	0.051
3	0.234	0.227	0.132	0.109
4	0.260	0.254	0.184	0.166
5 (Largest)	0.301	0.293	0.255	0.239

To investigate the intertemporal behaviour of the size effect, the two graphs of Figure 1 provide, for each market, a representation of the time series dynamics of the R-squares over the entire sample period for the 5 size-sorted portfolios, while the mean and the median of the whole markets are in Table 3.

Panel A. NYSE



Panel B. NASDAQ

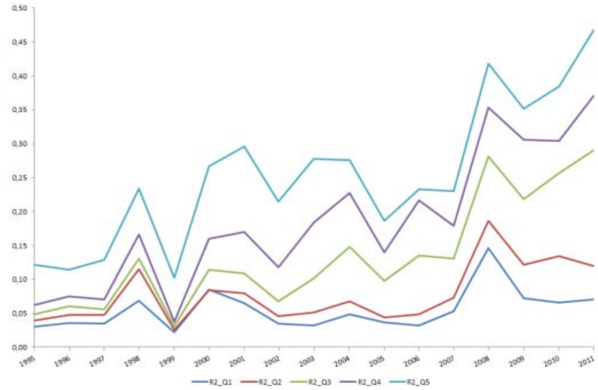


Figure 1. Time series dynamics of the R-square by quintile of market capitalization.

The R-square increases with market capitalization. The difference in R-squares across size-sorted portfolios increases over time in our sample period for the US markets. For the NYSE, in 1995, the average R-square for the 5^{th} quintile (largest stocks) was 12% and for the 1^{st} quintile (smallest) was 3%; in 2011, the average R-square for the 5^{th} quintile was 53% and for the 1^{st} quintile was 30%. Consequently, the difference in R-squares between top and bottom quintiles increased from 27 to 41 percentage points.

Table 3. R^2 by Year.

	United States								
	NYSE		NASDAQ			NYSE		NASDAQ	
Year	Mean	Median	Mean	Median	Year	Mean	Median	Mean	Median
1995	0.061	0.034	0.061	0.028	2004	0.208	0.195	0.156	0.108
1996	0.107	0.065	0.067	0.042	2005	0.181	0.158	0.103	0.066
1997	0.131	0.092	0.068	0.039	2006	0.189	0.174	0.136	0.092
1998	0.193	0.169	0.144	0.118	2007	0.258	0.249	0.136	0.104
1999	0.084	0.044	0.044	0.020	2008	0.425	0.436	0.280	0.266
2000	0.109	0.079	0.143	0.077	2009	0.381	0.388	0.218	0.185
2001	0.231	0.201	0.146	0.091	2010	0.397	0.404	0.234	0.211
2002	0.174	0.138	0.097	0.045	2011	0.467	0.491	0.270	0.249
2003	0.204	0.177	0.131	0.078					

3. The Country-Level R-Square as a Chisini mean

Since we are interested in studying the country-level R^2 and not in the comparison across countries, for the sake of simplicity, we drop the subscript j and rewrite (1) as follows:

$$R^2 = \sum_i R_i^2 \cdot \frac{\sigma_i^2}{\sum_i \sigma_i^2} = \frac{\sum_i \bar{\sigma}_i^2}{\sum_i \sigma_i^2} \tag{3}$$

where, $\bar{\sigma}_i^2 = SSR_i/T$, $\sigma_i^2 = SST_i/T$, and T is the length of the available time horizon. Since R^2 may be read as the ratio of two variables, alike most of balance or economic indexes or in physics e.g. speed, in this Section we propose an interpretation of (3) based on the Chisini approach to compute a mean (see, e.g. [6]). The intuition behind this method is explained as follows. In general, consider the variables Y, X_1, X_2,…, X_h,…, X_q and the associated sample set y_i, x_{i1}, x_{i2},…, x_{ih},…, x_{iq} for i=1, 2, …, n. Suppose that a function $f(\cdot)$ exists such that $Y = f(X_1, X_2, \dots, X_h, \dots, X_q)$. We may write:

$$\sum_i y_i = \sum_i f(x_{i1}, x_{i2}, \dots, x_{ih}, \dots, x_{iq})$$

Consider the variable X_h, for $1 \le h \le q$ and let $\bar{x}_h$ be a scalar which maps $\{x_{1h}, x_{2h}, \dots, x_{nh}\} \to \mathbb{R}^1$. For example, $\bar{x}_h$ may be the average of X_h. If:

$$\sum_i f(x_{i1}, x_{i2}, \dots, x_{ih}, \dots, x_{iq}) = \sum_i f(x_{i1}, x_{i2}, \dots, \bar{x}_h, \dots, x_{iq}) \tag{4}$$

holds, then $\bar{x}_h$ is the Chisini mean of the variable X_h. The solution $\bar{x}_h$ has the usual properties of an average operator, plus the one of keeping invariant the quantity $\sum_i y_i$. Equation (4) is indeed known as the invariance requirement. If we estimate the R-squares of individual stocks in a country and we wish to use the country-level R-square computed as in (3) as a price efficiency measure, equation (3) may be considered as the solution of a problem like (4) by supposing to keep invariant either:

a) the country systematic risk, computed as the sum of the variances explained by the market model estimated for all the stocks in the country (i.e., $\sum_i \bar{\sigma}_i^2$)),

or

b) the country total risk, computed as the sum of the overall stock return variances[2] (i.e., $\sum_i \sigma_i^2$).

If a) holds, using the Chisini approach and the fact that $\bar{\sigma}_i^2 = \sigma_i^2 \cdot R_i^2$, we may write:

$$\sum_i \bar{\sigma}_i^2 = \sum_i \sigma_i^2 R_i^2 \overset{\text{Apply Chisini}}{\Rightarrow} \sum_i \sigma_i^2 R^2 = \sum_i \sigma_i^2 R_i^2 \Rightarrow R^2 = \frac{\sum_i \sigma_i^2 \cdot R_i^2}{\sum_i \sigma_i^2} \tag{5}$$

If b) holds, using again the Chisini approach and the fact that $\sigma_i^2 = \frac{\bar{\sigma}_i^2}{R_i^2}$, we have:

$$\sum_i \sigma_i^2 = \sum_i \frac{\bar{\sigma}_i^2}{R_i^2} \overset{\text{Apply Chisini}}{\Rightarrow} \sum_i \frac{\bar{\sigma}_i^2}{R^2} = \sum_i \frac{\bar{\sigma}_i^2}{R_i^2} \Rightarrow R^2 = \frac{\sum_i \bar{\sigma}_i^2}{\sum_i \frac{\bar{\sigma}_i^2}{R_i^2}} \tag{6}$$

which may be interpreted as the harmonic mean of the R^2s with weights given by the variances explained by the market model.

Although both approaches simplify into (3), there are important differences in the interpretation of the country-level measure. According to (5) (i.e., if a) holds), we place greater weight on the R-squares that are associated with highly volatile stocks, holding constant the country systematic risk. This weighting scheme facilitates the decomposition of stock returns variation in a market-

[2] Other papers in this area also refer to this quantity as "total variance of returns" or "squared total variations" or "total sum of squares".

wide component and a firm-specific component (see, e.g., [12]). According to (6) (i.e., if b) holds), we keep invariant the country total risk and we weight individual stocks on the basis of their proportion of total risk explained by the market model.
The main difference between (5) and (6) is that in (5) the overall country systematic risk is held constant and, consequently, stock-by-stock models are implicitly assumed to be estimated. By contrast, (6) holds constant the overall country total risk and, to guarantee the equivalence in (3), assumes that a model exists such that the country systematic risk is equal to $\sum_i \bar{\sigma}_i^2$.
In the next Section we further investigate the above interpretations, looking for the conditions such that, given a model that simultaneously estimates the market model in (1) for all the stocks traded in a country, its R-square is (3).

4. The Country-Level R-Square and the corresponding Market Model

The most common model used to jointly study the multivariate responses (stock returns) as a function of one only explanatory variable (market returns) is the Seemingly Unrelated Regression (SUR) model. Focusing on that model, we need to satisfy the following two conditions: first, when a SUR have explained variance equal to $\sum_i \bar{\sigma}_i^2$ and, second when its R-square equals to (3). We emphasize that in this paper our aim is not to make inference on the model parameters or on the model itself, which are topics already well addressed in the financial econometrics literature, but only to study under which conditions a multivariate response model satisfies the previous requirements.
We will make use of the following notation:

a) $\boldsymbol{I}_k$ is the identity matrix of dimension $k\times k$ and $\mathbf{1}$ is a $T\times1$ vector of ones, otherwise specified in the text;
b) $\boldsymbol{r}_m$ is the $T\times1$ vector of market returns and $r_{m,t}$ for t=1,2,...,T is the market return at time t;
c) $\boldsymbol{R}_P$ is a $T\times k$ matrix with the i-th column, $\boldsymbol{r}_i$,representing the $T\times1$ vector of returns of the i-th stock, and $\boldsymbol{r}_{i,t}$ is the return at time t;
d) $\boldsymbol{\mu}_P$ and μ_m are, respectively, the mean vector of the stocks and the mean of the market;
e) $\boldsymbol{\beta}$ is a vector of parameters;
f) $vec(\cdot)$ and $\otimes$ are, respectively, the "*vec*" and the Kronecker matrix operators.

Conditionally to $\boldsymbol{r}_m$, the standard univariate market model in (2) has the following multivariate representation:

$$vec(\boldsymbol{R}_P) = (\boldsymbol{I}_k \otimes [\mathbf{1}\ \ \boldsymbol{r}_m])\boldsymbol{\beta} + vec(\boldsymbol{E}_P) \qquad (7)$$

Let SST_R be the sum of squared total variations for $vec(\boldsymbol{R}_P)$, i.e. $SST_R = \sum_i SST_i$. To show that (7) is the model whose R-square is (3), we start by checking whether SST_R is equal to the denominator of the country-level R-square. For a generic vector $\boldsymbol{Q}$, let $dev(\boldsymbol{Q})$= $\boldsymbol{Q}'\boldsymbol{Q}$ and let $\mathbf{1}_{Tk}$ be a vector of $T\times k$ ones. We may write:

$$SST_R = dev\big(vec(\boldsymbol{R}_P)\big) - Tk\left(\left(\mathbf{1}'_{Tk}\mathbf{1}_{Tk}\right)^{-1}\mathbf{1}'_{Tk}vec(\boldsymbol{R}_P)\right)^2 \tag{8}$$

where according to the definition of market return the last term equals to $Tk(\mu_m)^2$. The denominator of the country-level R-square is:

$$T\sum\sigma_i^2 = dev\big(vec(\boldsymbol{R}_P)\big) - T(\boldsymbol{\mu}'_P\boldsymbol{\mu}_P) \tag{9}$$

From (8) and (9) it follows that SST_R will be equal to $T\sum\sigma_i^2$ if:

$$Tk(\mu_m)^2 = T(\boldsymbol{\mu}'_P\boldsymbol{\mu}_P)$$

This condition is generally not satisfied and it may hold – for example – when $\mu_i = \mu_m\ \forall i$. To guarantee the equivalence of (8) and (9) a solution is by centring returns on zero.
It follows that this operation, commonly adopted in most of applications[3] often with no further justifications, is necessary in our context. Henceforth we use the assumption $\mu_i = \mu_m = 0\ \forall i$. From these preliminaries to show that (7) is the model whose R-square is (3), we need to verify whether the variance explained by (7) equals to $T\sum\bar{\sigma}_i^2$.
This will be true if the $\boldsymbol{i}$-th term in $\widehat{\boldsymbol{\beta}}$, i.e. the estimate of $\boldsymbol{\beta}$, corresponds to the estimate of the slope coefficient of the univariate market model (2) for the i-th stock. The parameters in (2) are usually estimated by OLS in the classical Gaussian linear model framework. This implies that the residuals are supposed to be homoschedastic and spherical. Since model (7) aggregates the k models in (2) for i=1,…,k , to be coherent with the previous assumption and introducing cross-correlation among stocks we have to assume that:

$$E[vec(\boldsymbol{E}_P)vec(\boldsymbol{E}_P)'] = \boldsymbol{V_P} = E[\boldsymbol{E}'_P\boldsymbol{E}_P] \otimes \boldsymbol{I}_k := \boldsymbol{\Sigma}_P \otimes \boldsymbol{I}_T\ . \tag{10}$$

where $\boldsymbol{\Sigma}_P$ is the variance covariance matrix of the errors[4]. Estimating by GLS, since regressors in (7) are the same for all the equations, by Kruskal's theorem we know that OLS and GLS estimators are the same:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= (\,(\,\boldsymbol{I}'_k \otimes \boldsymbol{r}'_m\,)(\,\boldsymbol{\Sigma}_P^{-1} \otimes \boldsymbol{I}_T\,)(\,\boldsymbol{I}_k \otimes \boldsymbol{r}_m\,)\,)^{-1}(\boldsymbol{I}'_k \otimes \boldsymbol{r}'_m\,)(\,\boldsymbol{\Sigma}_P^{-1} \otimes \boldsymbol{I}_T\,)\,vec(\boldsymbol{R_P}) \\
&= (\,\boldsymbol{I}'_k\,\boldsymbol{\Sigma}_P^{-1}\,\boldsymbol{I}_k \otimes \boldsymbol{r}'_m\,\boldsymbol{I}_T\,\boldsymbol{r}_m\,)^{-1}(\boldsymbol{I}'_k\,\boldsymbol{\Sigma}_P^{-1} \otimes \boldsymbol{r}'_m\boldsymbol{I}_T\,)\,vec(\boldsymbol{R_P}) \\
&= (\boldsymbol{\Sigma}_P\boldsymbol{\Sigma}_P^{-1} \otimes (\boldsymbol{r}'_m\boldsymbol{r}_m)^{-1}\boldsymbol{r}'_m\,)\,vec(\boldsymbol{R_P}) \\
&= (\boldsymbol{I}_k \otimes (\boldsymbol{r}'_m\boldsymbol{r}_m)^{-1}\boldsymbol{r}'_m\,)\,vec(\boldsymbol{R}_P) \\
&= (\,\sigma_m^2\,)^{-1}\mathrm{Cov}(\boldsymbol{R}_P, \boldsymbol{r}_m)
\end{aligned}$$

---

[3] The de-meaning procedure does not bias the estimate of R_i^2

[4] The choice to center return on zero has been shown to be relevant to guarantee equivalence of (8) and (9). By contrary it is well known that the Frish-Waugh-Lovell theorem [11] states that centring returns on zero is equivalent to the projection of the sample space onto the orthogonal complement of $(\boldsymbol{I}_k - \mathbf{1}_k(\mathbf{1}'_{\boldsymbol{k}}\mathbf{1}_k)^{-1}\mathbf{1}'_{\boldsymbol{k}})$, where $\mathbf{1}_{\boldsymbol{k}}$ is a vector of length k. Then the estimates of the betas in (10) do not change.

where Cov(·) is the column vector of covariances across the k stocks and the market returns. Hence $dev\left((\boldsymbol{I}_k \;\; \boldsymbol{r}_m)\widehat{\boldsymbol{\beta}}\right) = T\sum \frac{Cov(r_i, r_m)^2}{\sigma_m^2} = T\sum_i \bar{\sigma}_i^2$. Had we assumed more generally:

$$E[vec(\boldsymbol{E}_P)vec(\boldsymbol{E}_P)^{\mathrm{T}}] := \boldsymbol{V_P} \tag{11}$$

i.e. introducing also cross-dependencies within stocks, we could not have found the equivalence. This implies that, under assumption (11), the R^2 will not be equal to (3).
We close this section with some comments on the equivalence $SST_R = T\sum \sigma_i^2$. From that we may argue that SST_R refers to the variance of mixture of *uncorrelated* variables (i.e. returns). Since the covariances are not included, we must suppose that either the variables at the denominator of (3) are cross-sectionally uncorrelated or, less commonly, a transformation (i.e. by principal components) has been adopted to make stock returns uncorrelated. We have shown that a SUR model matches this condition centring returns on zero. Even assuming the existence of cross-correlations in (10) they vanish in the computation of R-square (but obviously (10) plays a role if we were dealing with inferential issues): on one side this is coherent with the denominator of (3) on the other it is in contradiction with the assumption (10). To include the covariances among stocks, we may set up an alternative formulation of the market model by considering stocks within a portfolio. By gathering stock returns in $\boldsymbol{r_P} = \boldsymbol{r_1} + \cdots + \boldsymbol{r_k}$ we should study the model[5]:

$$\boldsymbol{r}_P = [\boldsymbol{1} \;\; \boldsymbol{r}_m]\,\boldsymbol{\beta}_P + \boldsymbol{e}_P \quad \text{with} \quad E[\boldsymbol{e}_P\boldsymbol{e}_P^{\boldsymbol{T}}] = diag\left(\boldsymbol{\Sigma}_{\mathrm{P}_*}\right) \tag{12}$$

where, using assumptions in (10), the *ii*-th element of the diagonal matrix $diag\left(\boldsymbol{\Sigma}_{\mathrm{P}_*}\right)$ is equal to $\boldsymbol{1^T\Sigma}_{\mathrm{P}}\boldsymbol{1},$ i.e. the sum of all the elements in the matrix $\boldsymbol{\Sigma}_{\mathrm{P}}$. $\boldsymbol{\beta}_P$ is estimated by GLS. The denominator of the R-square of (12) equals to $Var(\boldsymbol{r}_P) = \sum\sigma_i^2 + 2\sum_{\boldsymbol{i<j}} Cov(\boldsymbol{r}_i, \boldsymbol{r}_j)$. Since:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_P &= \left([\boldsymbol{1} \;\; \boldsymbol{r}_m]'diag\left(\boldsymbol{\Sigma}_{\mathrm{P}_*}\right)^{-1}[\boldsymbol{1} \;\; \boldsymbol{r}_m]\right)^{-1}[\boldsymbol{1} \;\; \boldsymbol{r}_m]'diag\left(\boldsymbol{\Sigma}_{\mathrm{P}_*}\right)^{-1}\boldsymbol{r}_P \\
&= ([\boldsymbol{1} \;\; \boldsymbol{r}_m]'[\boldsymbol{1} \;\; \boldsymbol{r}_m])^{-1}[\boldsymbol{1} \;\; \boldsymbol{r}_m]'(\boldsymbol{r}_1 + \cdots + \boldsymbol{r}_k) \\
&= \widehat{\boldsymbol{\beta}}_1 + \widehat{\boldsymbol{\beta}}_2 + \cdots + \widehat{\boldsymbol{\beta}}_k
\end{aligned}$$

where $\widehat{\boldsymbol{\beta}}_i = \left[\hat{\alpha}_i \;\; \hat{\beta}_i\right]'$ is the OLS estimate of the parameters in (2), we have:

$$\begin{aligned}
Var(\hat{\boldsymbol{r}}_P) &= \widehat{\boldsymbol{\beta}}_P{}'([\boldsymbol{1} \;\; \boldsymbol{r}_m]'[\boldsymbol{1} \;\; \boldsymbol{r}_m])\widehat{\boldsymbol{\beta}}_P \\
&= \sum \widehat{\boldsymbol{\beta}}_i'([\boldsymbol{1} \;\; \boldsymbol{r}_m]'[\boldsymbol{1} \;\; \boldsymbol{r}_m])\widehat{\boldsymbol{\beta}}_i + 2\sum_{\boldsymbol{i<j}} \widehat{\boldsymbol{\beta}}_i'([\boldsymbol{1} \;\; \boldsymbol{r}_m]'[\boldsymbol{1} \;\; \boldsymbol{r}_m])\widehat{\boldsymbol{\beta}}_j \\
&= \sum \bar{\sigma}_i^2 + 2\sum_{\boldsymbol{i<j}} Cov(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j)
\end{aligned}$$

Finally the R-square of model (12) is:

[5] For the rest of the paper centering returns on zero is not a compulsory condition.

$$R^2_{[12]} = \frac{\sum \bar{\sigma}_i^2 + 2\sum_{i<j} Cov(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j)}{\sum \sigma_i^2 + 2\sum_{i<j} Cov(\boldsymbol{r}_i, \boldsymbol{r}_j)} = \frac{R^2 + 2\sum_{i<j} Cov(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j)/\sum \sigma_i^2}{1 + 2\sum_{i<j} Cov(\boldsymbol{r}_i, \boldsymbol{r}_j)/\sum \sigma_i^2}$$

which is clearly a function of (3), corrected for the cross-correlation among stocks without having constrained mean returns on zero. Considering that using (2) and the properties of OLS, stock returns can be thought of the sum of fitted values plus the residuals we have

$$\begin{aligned} Cov(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j) &= Cov([\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_i, [\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) = \\ &= E(\widehat{\boldsymbol{\beta}}_i'[\mathbf{1}\ \boldsymbol{r}_m]'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) - E(\mathbf{1}'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_i)E(\mathbf{1}'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) \end{aligned}$$

$$\begin{aligned} Cov(\boldsymbol{r}_i, \boldsymbol{r}_j) &= Cov([\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_i + \hat{\boldsymbol{e}}_i, [\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j + \hat{\boldsymbol{e}}_j) \\ &= E(\widehat{\boldsymbol{\beta}}_i'[\mathbf{1}\ \boldsymbol{r}_m]'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) + E(\widehat{\boldsymbol{\beta}}_i'[\mathbf{1}\ \boldsymbol{r}_m]'\hat{\boldsymbol{e}}_j) + E(\widehat{\boldsymbol{\beta}}_j'[\mathbf{1}\ \boldsymbol{r}_m]'\hat{\boldsymbol{e}}_i) + E(\hat{\boldsymbol{e}}_j'\hat{\boldsymbol{e}}_i) \\ &\quad - E\left(\mathbf{1}'([\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_i + \hat{\boldsymbol{e}}_i)\right)E\left(\mathbf{1}'([\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j + \hat{\boldsymbol{e}}_j)\right) = \\ &= E(\widehat{\boldsymbol{\beta}}_i'[\mathbf{1}\ \boldsymbol{r}_m]'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) + E(\hat{\boldsymbol{e}}_j'\hat{\boldsymbol{e}}_i) - E(\mathbf{1}'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_i)E(\mathbf{1}'[\mathbf{1}\ \boldsymbol{r}_m]\widehat{\boldsymbol{\beta}}_j) \end{aligned}$$

Then we expect $Cov(\hat{\boldsymbol{r}}_i, \hat{\boldsymbol{r}}_j) < Cov(\boldsymbol{r}_i, \boldsymbol{r}_j)$ and $R^2_{[12]} < R^2$.

5. Conclusion

This paper shows that the country-level R-square, computed using the individual firm total risk as weighting factor, can be interpreted as a Chisini mean of the individual firms R-squares in two ways. In the first way, the weight is the proportion of firm total risk over country total risk and the mean is computed to hold constant the country systematic risk. In the second way, the weight is the proportion of firm systematic risk over country systematic risk and the mean is computed to hold constant the country total risk. Only the second way allows the country-level R-square to be interpreted as the coefficient of determination of a unique model to be estimated jointly considering all the stocks traded in a country. Specifically, we show that the country-level R-square can be obtained, through the estimation of a Seemingly Unrelated Regressions (SUR) model even assuming correlation across stocks and by centering returns on zero.

Acknowledgement

The Authors gratefully wish to thank two referees for their valuable comments and suggestions.

References

[1]. Alves, P., K. Peasnell and P. Taylor. (2010). The Use of the R2 as a Measure of Firm-Specific Information: A Cross-Country Critique. *Journal of Business Finance & Accounting*, 37, 1–26.

[2]. Bramante R., Petrella G., Zappa D. (2012). On the use of the market model R-square as a measure of stock price efficiency. Submitted.
[3]. Chan, K., Hameed, A. (2006). Stock Price Synchronicity and Analyst Coverage in Emerging Markets. *Journal of Financial Economics*, 80, 115-147.
[4]. Durnev, A., Morck, R., Yeung, B., Zarowin, P. (2003). Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing?. *Journal of Accounting Research*, 41, 797-836.
[5]. Fernandes, N., Ferreira, M.A. (2009). Insider Trading Laws and Stock Price Informativeness. *Review of Financial Studies*, 22, 1845-1887.
[6]. Graziani, R., Veronese, P. (2009). How to Compute a Mean? The Chisini Approach and Its Application. *The American Statistician*, 63, 33-36.
[7]. Greene, W. (2003). *Econometric Analysis*. London: Prentice Hall.
[8]. Griffin, J., Kelly, P., Nardari, F. (2006). Measuring Short-Term International Stock Market Efficiency. *Working Paper*, University of Texas at Austin.
[9]. Jin, L., Myers, S. (2006). Return Synchronicity Around the World: New Theory and New Tests. *Journal of Financial Economics*, 79, 257-292.
[10]. Lai, S., Ng, L., Zhang, B. (2009). *Informed Trading Around The World*, Singapore Management University.
[11]. Lovell, M. (2008). A Simple Proof of the FWL (Frisch,Waugh,Lovell) Theorem. *Journal of economic education*, 39, 88-91.
[12]. Morck, R., Yeung, B., Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements. *Journal of Financial Economics*, 58, 215-260.
[13]. Pagano, M., Schwartz, R. (2003). A closing call's impact on market quality at Euronext Paris. *Journal of Financial Economics*, 68, 439-484.
[14]. Piotroski, J., Roulstone, D. (2004). The Influence of Analysts, Institutional Investors and Insiders on the Incorporation of Market, Industry and Firm-Specific Information into Stock Prices. *Accounting Review*, 79, 1119-1151.
[15]. Stowe, J.D. , Xing X. (2011). R2: Does it Matter for Firm Valuation?. *The Financial Review*, 46:233–250.
[16]. Wurgler, J. (2000). Financial markets and the allocation of capital. *Journal of Financial Economics*, 58, 187-214.