Testing for model selection in predicting aggregate variables*

Giacomo Sbrana†

Abstract

This paper focuses on the choice between aggregate and disaggregate models, consisting in both univariate and multivariate specifications, in predicting aggregate variables. Here, we suggest a formal hypothesis testing procedure for in-sample model selection. The empirical size and power of the test are investigated via the use of Monte Carlo simulations. Empirical results show that the test has good performance not only when the competitive models are non-nested specifications, but also when considering nested competitors.

Keywords: Aggregation, Model evaluation and selection, Hypothesis testing, Monte Carlo simulation.

JEL Classification: C43, C52, C12, C15

---

*This paper has been written during my visiting research period at the Department of Economics, University of Southampton. I am particularly grateful to John Aldrich, Grayham Mizon and Jan Podivinsky for their helpful suggestions and comments. I also wish to thank Alberto Baffigi, Massimiliano Marcellino, Hashem Pesaran and the participants of the 58th European Meeting of the Econometric Society (Stockholm, 2003). Financial support of the Royal Economic Society, Dipartimento di Economia, Universita degli Studi "Roma Tre" and the Marie Curie fellowship (HPMT-CT-2001-00353) are gratefully acknowledged.

†United Nations, Department of Economics and Social Affairs, Population Division, 2 UN Plaza, DC2-1940, New York, NY 10017, USA. E-mail address: sbrana@un.org

1 Introduction

The aggregation problem over individuals has been extensively investigated in the econometric literature[1]. In particular, it has been analysed by assuming different hypotheses and assumptions, such that it is necessary to divide it into three different main theoretical approaches and one empirical approach in order to clarify the particular area of our analysis. According to Pesaran (2003), we can identify the following different theoretical approaches to the aggregation problem: 1) The deterministic approach, 2) The statistical approach and 3) The forecasting approach. The deterministic approach was originally explored by Gorman (1953), Theil (1954) and Fisher (1969). It is mainly concerned with obtaining conditions under which it is legitimate to set up an aggregate model containing parameters that are direct measures of the underlying micro parameters. In other words, this area of the econometric literature has been primarily interested in whether an aggregate function exists which is identical to the individual functions. An alternative approach is the so-called statistical approach which has been advanced by Kelejian (1980), Stoker (1984) and Lewbel (1994). This approach is less stringent than the previous one. In fact it mainly focuses on the relationship between the aggregate and the disaggregate specification, where the aggregate function can be considered as a joint probability distribution of the micro variables and parameters. The last and more recent methodology is the "forecasting approach" proposed by Pesaran (2003)[2]. This kind of approach focuses the analysis on the optimal prediction of the aggregate variables conditional on the available aggregate information. In this framework, the optimal prediction can be seen with respect to the minimum mean squared error as a particular objective loss function.

Another important area of the aggregation literature is concerned with the empirical approach, namely the "model selection problem"[3], which is the main topic of our analysis. In that context, the major object of interest is to choose whether to predict aggregated (or macro) variables by using aggregate or disaggregate specifications [4]. This kind of approach was laid out for linear models by the pioneering paper of Grunfeld and Griliches (1960). Here

---

[1] A complete survey of the problem can be found in Stoker (1993)

[2] For an application see also Garderen, Lee and Pesaran (2000)

[3] A good introduction to the empirical approach can be found in Maddala (1977)

[4] For an interesting overview see also: Barker and Pesaran (1990)

it has been argued that if micro-data are subject to large errors (compared with macro-data) and if the micro relations are not correctly specified, there would be a gain in using the aggregate specification rather then the disaggregate specification even if there is an aggregation error arising from the use of a common specification. However, if the micro relations are correctly specified, the use of an aggregate specification would generally led to worse predictions than those of the disaggregate model. In order to choose the model to be used, Grunfeld and Griliches proposed an "in-sample goodness of fit" criterion to discriminate between the aggregate and the disaggregate model. Such a criterion, generally called the "GG criterion", is based on the comparison of the R^2 of the aggregate model with that of the disaggregate model. A few years later Aigner and Goldfeld (1973, 1974) suggested a formal analysis of the case where aggregates are measured more correctly than the disaggregates. In particular, they argued in favor of the macro-equation specification to be considered superior to the micro (or individual) equations. A completely different opinion can be found in Orcutt, Watts, Edwards (1968), who show that, by using some simulation experiments, the predictions from the disaggregate specification are far better than that of the aggregate. Other important contributions to the model selection problem are those of Pesaran, Pierce and Kumar (1989), Pesaran, Pierce and Lee (1994), Garderen, Lee and Pesaran (2000). The GG criterion has been extensively employed by the empirical literature. In particular, most of the contributions on predicting Euro-area macroeconomic variables comparing country specific versus area-wide models such as Wesche (1997), Den Butten and Van Djiken (1997), Clausen (1998), Fagan and Henry (1998), Dedola, Gaiotti and Silipo (2001), Golinelli and Pastorelo (2002), Baffigi, Parigi and Golinelli (2002, 2004), Marcellino , Stock and Watson (2003), Hubrich (2004), made use of the GG criterion to investigate on the use of aggregate specifications as opposed to the disaggregate national models. Furthermore, Hsiao, Shen and Fujiki (2005) also employed the GG criterion in estimating money demand function in Japan.

The aim of this paper is to introduce a statistical procedure for model selection in predicting aggregate variables. Here, we employ the Vuong testing procedure in order to test whether the two rival models can be considered observationally equivalent. The properties of the empirical size and power of the likelihood-ratio test are tested by Monte Carlo simulation. In particular, we aim to expand the model selection problem to take into account for a more formal procedure to discriminate between competitive models where the aggregate and disaggregate specifications consist in both single and multiple equations models. It is important to notice that most of the literature on the model selection problem has been primarily concerned with the single equation case. However, we consider it important to extend the econometric framework of the aggregation problem to the multivariate case. This is because prediction of aggregated variables is often carried out in a context of multiple equations models (e.g. in modelling macroeconomic time series). The paper is organized as follows: in Section 2 the disaggregated and the aggregated specifications are introduced in order to set up the basic assumptions. In Section 3 we introduce the classical model selection criterion and we generalize the GG criterion in the multivariate case. Section 4 focuses on the Vuong testing procedure in the context of prediction of aggregate variables. Section 5 contains a Monte Carlo simulation for analyzing the small sample properties of the Likelihood-ratio test statistic. Some concluding remarks are provided in Section 6.

2 The econometric framework

As a starting point we focus on the disaggregate model. Suppose that there are m micro-units (such as individuals, firms, regions, countries) whose decision variables are contained in the following vector, $\boldsymbol{y}_{it}$, observed over T time periods. We can define the multivariate model specification for the i-th unit as follows:

$$\boldsymbol{y}_{it} = \boldsymbol{B}_i \boldsymbol{x}_{it} + \boldsymbol{u}_{it} \tag{1}$$

with $i = 1, ..., m$; and $t = 1, ..., T$;. Here $\boldsymbol{y}_{it} = (y_{1t}, ..., y_{Kt})'$ are $(K \times 1)$ vectors of endogenous modelled variables, $\boldsymbol{x}_{it}$ are $(p_i \times 1)$ vectors of explanatory variables related to the specific micro-unit, $\boldsymbol{B}_i$ are $(K \times p_i)$ coefficients matrices and $\boldsymbol{u}_{it}$ are error terms. In addition, the variables contained in $\boldsymbol{x}_{it}$ can be either lagged endogenous variables, exogenous variables or stationary linear combinations ($i.e.$ $\boldsymbol{y}_{it-1}, ..., \boldsymbol{y}_{it-p_i}, \boldsymbol{z}_{it}, ..., \boldsymbol{z}_{it-g_i}, \boldsymbol{\beta}'\boldsymbol{x}_{it-1}$).[5]

---

[5] Note that $\boldsymbol{\beta}'\boldsymbol{x}_{it-1}$ stands for the so called cointergating relationships

We now turn to consider the aggregate specification (i.e. the macro specification). In this particular model the vector of decision variables can be expressed as a sum of the endogenous variables over the micro units. Hence, we can write the aggregate specification as follows:

$$\boldsymbol{y}_{at} = \boldsymbol{B}_a \boldsymbol{x}_{at} + \boldsymbol{u}_{at} \tag{2}$$

In the previous ad-hoc specification the vector of aggregate variables $\boldsymbol{y}_{at} = \sum_{i=1}^{m} \boldsymbol{y}_{it}$ is function of some explanatory variables. These will include aggregate variables contained in $\boldsymbol{x}_{at}$[6], a coefficients matrix $\boldsymbol{B}_a$, plus a vector of disturbances $\boldsymbol{u}_{at}$. In addition, we consider the following assumption:

Assumption 2.1 The disturbance terms of the micro and macro specifications, $\boldsymbol{u}_{it}$ and $\boldsymbol{u}_{at}$, are Normally distributed vectors.

In particular, we have:

$$E(\boldsymbol{u}_{it}) = \boldsymbol{0}, \qquad E(\boldsymbol{u}_{at}) = \boldsymbol{0}$$

and

$$E(\boldsymbol{u}_{it}\boldsymbol{u}'_{i\tau}) = \begin{cases} \boldsymbol{\Omega}_i & \text{for } t=\tau \\ \boldsymbol{0} & \text{otherwise} \end{cases} \quad E(\boldsymbol{u}_{it}\boldsymbol{u}'_{j\tau}) = \begin{cases} \boldsymbol{\Omega}_{ij} & \text{for } t=\tau \\ \boldsymbol{0} & \text{otherwise} \end{cases}$$

$$E\left[\left(\sum_{i=1}^{m} \boldsymbol{u}_{it}\right)\left(\sum_{i=1}^{m} \boldsymbol{u}_{i\tau}\right)'\right] = \begin{cases} \boldsymbol{\Omega}_d & \text{for } t=\tau \\ \boldsymbol{0} & \text{otherwise} \end{cases}$$

$$E(\boldsymbol{u}_{at}\boldsymbol{u}'_{a\varsigma}) = \begin{cases} \boldsymbol{\Omega}_a & \text{for } t=\varsigma \\ \boldsymbol{0} & \text{otherwise} \end{cases}$$

Here we assume that $\boldsymbol{\Omega}_d$ and $\boldsymbol{\Omega}_a$ are both positive definite.

In this paper we focus the analysis on forecasting the vector of aggregated variables $\boldsymbol{y}_{at}$ where m is finite[7]. Due to the availability of the micro-units data, we face the choice of whether to employ the disaggregate specification $\widehat{\boldsymbol{y}}_{dt} = \sum_{i=1}^{m} \widehat{\boldsymbol{y}}_{it}$, by forecasting the micro units first and pooling the forecasts *ex-post*, or the aggregated one $\widehat{\boldsymbol{y}}_{at} = \widehat{\boldsymbol{B}}_a \boldsymbol{x}_{at}$ by aggregating the endogenous variables *ex-ante* and estimating an ad-hoc aggregate model[8].

---

[6] Note that $\boldsymbol{x}_{at}$ is a $(p_a \times 1)$ vector of exogenous aggregate variables. Therefore we allow for heterogeneous specifications across both micro and macro models.

[7] Notice that the vector y_{at} can also be defined as a weighted average across the endogenous variables of the micro-units (i.e. $y_{at} = \sum_{i=1}^{m} \omega_i y_{it}$ with $\sum_{i=1}^{m} \omega_i = 1$).

[8] Here it is relevant to mention that our analysis do not allow for including micro exogenous variables within the aggregate model as suggested by Hendry and Hubrich (2005). Therefore disaggregate information is not contained in the ad-hoc aggregate (or macro) specification. The extension to this particular case, recently suggested, is feasible. However, at this stage we prefer to refer to the classical model selection approach.

3 Model Selection: a review and an extension to the multivariate case

In this sections we review the non-formal GG criterion. Moreover, we also extend the GG criterion to the multivariate case when we aim to predict more than one endogenous variables of interest. Recall the micro-unit specification (1). We can rewrite the model as follows:

$$\boldsymbol{Y}_i = \boldsymbol{B}_i \boldsymbol{X}_i + \boldsymbol{U}_i \tag{3}$$

where

$$\underset{(K\times T)}{\boldsymbol{Y}_i} = (\boldsymbol{y}_{1t}, ..., \boldsymbol{y}_{Kt})' \tag{4}$$

$$\underset{(K\times p_i)}{\boldsymbol{B}_i} = (\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_{p_i}) \tag{5}$$

$$\underset{(p_i\times T)}{\boldsymbol{X}_{it}} = (\boldsymbol{x}_{1t}, ..., \boldsymbol{x}_{p_i t})' \tag{6}$$

$$\underset{(K\times T)}{\boldsymbol{U}_i} = (\boldsymbol{u}_{1t}, ..., \boldsymbol{u}_{Kt})' \tag{7}$$

such that the OLS estimator is[9]:

$$\widehat{\boldsymbol{B}}_i = \boldsymbol{Y}_i \boldsymbol{X}_i' (\boldsymbol{X}_i \boldsymbol{X}_i')^{-1} \tag{8}$$

Furthermore, we assume the aggregate model having the following specification:

$$\boldsymbol{Y}_a = \boldsymbol{B}_a \boldsymbol{X}_a + \boldsymbol{U}_a \tag{9}$$

where

$$\underset{(K\times T)}{\boldsymbol{Y}_a} = \sum_{i=1}^{m} \boldsymbol{Y}_i \tag{10}$$

$$\underset{(p_a\times T)}{\boldsymbol{X}_a} = \sum_{i=1}^{m} \boldsymbol{X}_i \tag{11}$$

Therefore we have:

$$\widehat{\boldsymbol{B}}_a = \boldsymbol{Y}_a \boldsymbol{X}_a' (\boldsymbol{X}_a \boldsymbol{X}_a')^{-1} \tag{12}$$

---

[9] An extension of the GG criterion to the instrumental variables estimation in the case of simultaneous equations models can be found in Pesaran, Pierce and Lee (1994).

Following the past literature on the "in-sample goodness of fit" criteria, we now focus on the estimated residuals of the aggregate and disaggregate specifications. From the OLS estimator we have:

$$\widehat{\boldsymbol{U}}_d = \sum_{i=1}^{m} \boldsymbol{Y}_i \boldsymbol{M}_i \qquad \widehat{\boldsymbol{U}}_a = \boldsymbol{Y}_a \boldsymbol{M}_a \tag{13}$$

with

$$\boldsymbol{M}_i = (\boldsymbol{I}_T - \boldsymbol{X}_i'(\boldsymbol{X}_i\boldsymbol{X}_i')^{-1}\boldsymbol{X}_i) \qquad \boldsymbol{M}_a = (\boldsymbol{I}_T - \boldsymbol{X}_a'(\boldsymbol{X}_a\boldsymbol{X}_a')^{-1}\boldsymbol{X}_a) \tag{14}$$

Thus, from the aggregate and disaggregate specifications we have:

$$\widehat{\boldsymbol{\Omega}}_a = T^{-1}\widehat{\boldsymbol{U}}_a\widehat{\boldsymbol{U}}_a' = T^{-1}(\boldsymbol{Y}_a\boldsymbol{M}_a\boldsymbol{Y}_a') \tag{15}$$

$$\widehat{\boldsymbol{\Omega}}_d = T^{-1}\widehat{\boldsymbol{U}}_d\widehat{\boldsymbol{U}}_d' = T^{-1}\left(\sum_{i=1}^{m}\boldsymbol{Y}_i\boldsymbol{M}_i\right)\left(\sum_{i=1}^{m}\boldsymbol{Y}_i\boldsymbol{M}_i\right)' \tag{16}$$

with

$$\widehat{\boldsymbol{\Omega}}_d = T^{-1}\begin{pmatrix} \hat{\boldsymbol{u}}_{d(1)}\hat{\boldsymbol{u}}_{d(1)}' & \dots & \hat{\boldsymbol{u}}_{d(1)}\hat{\boldsymbol{u}}_{d(K)}' \\ \vdots & \ddots & \vdots \\ \hat{\boldsymbol{u}}_{d(K)}\hat{\boldsymbol{u}}_{d(1)}' & \dots & \hat{\boldsymbol{u}}_{d(K)}\hat{\boldsymbol{u}}_{d(K)}' \end{pmatrix} = \widehat{\boldsymbol{\Omega}}_{d_{ii}} + \widehat{\boldsymbol{\Omega}}_{d_{ij}}$$

and

$$\widehat{\boldsymbol{\Omega}}_{d_{ii}} = \begin{pmatrix} \sum_{i=1}^{m}\hat{\sigma}_{1(i)}^2 & \sum_{i=1}^{m}\hat{\sigma}_{1,2(i)} & \dots & \sum_{i=1}^{m}\hat{\sigma}_{1,K(i)} \\ \sum_{i=1}^{m}\hat{\sigma}_{2,1(i)} & \sum_{i=1}^{m}\hat{\sigma}_{2(i)}^2 & \dots & \sum_{i=1}^{m}\hat{\sigma}_{2,K(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{m}\hat{\sigma}_{K,1(i)} & \dots & \dots & \sum_{i=1}^{m}\hat{\sigma}_{K(i)}^2 \end{pmatrix} = \sum_{i=1}^{m}\widehat{\boldsymbol{\Omega}}_i$$

and

$$\widehat{\boldsymbol{\Omega}}_{d_{ij}} = \begin{pmatrix} \sum_{i\neq j}^{m}\hat{\sigma}_{1(i,j)} & \sum_{i\neq j}^{m}\hat{\sigma}_{1,2(i,j)} & \dots & \sum_{i\neq j}^{m}\hat{\sigma}_{1,K(i,j)} \\ \sum_{i\neq j}^{m}\hat{\sigma}_{2,1(i,j)} & \sum_{i\neq j}^{m}\hat{\sigma}_{2(i,j)} & \dots & \sum_{i\neq j}^{m}\hat{\sigma}_{2,K(i,j)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i\neq j}^{m}\hat{\sigma}_{K,1(i,j)} & \dots & \dots & \sum_{i\neq j}^{m}\hat{\sigma}_{K(i,j)} \end{pmatrix} = \sum_{i\neq j}^{m}\widehat{\boldsymbol{\Omega}}_{ij}$$

Let us now assume that K=1 (single equation). According to the "within-sample goodness of fit" criterion suggested by Grunfeld and Griliches, we can discriminate between the aggregate and the disaggregate specifications by comparing the variances of the prediction errors as follows:

Choose the disaggregate specification whenever:

$$\hat{\sigma}_d^2 \leq \hat{\sigma}_a^2 \tag{17}$$

Otherwise choose the aggregate model.

Here the variance of the aggregate model is the simple sum of squared of the aggregate residuals[10].

As we mentioned previously, most of the literature on model selection criteria has focused the analysis on the single equation specification. In particular, the GG criterion and the other generalizations of the same criterion (Pesaran, Pierce and Lee (1994), Garderen, Lee and Pesaran (2000)) are based on the comparison of single scalars (variances of errors). On the other hand, suppose that we allow for $K > 1$. In this case we deal with covariance matrices of the prediction errors of multivariate specifications. Here the comparison of the determinants allow us to reproduce the same theoretical approach used by Grunfeld and Griliches in a context where we aim to select between rival models employing more than one endogenous variable over time.

Therefore we can discriminate between the aggregate and the disaggregate specifications as follows:

Choose the disaggregate model whenever:

$$|\widehat{\mathbf{\Omega}}_d| \leq |\widehat{\mathbf{\Omega}}_a| \tag{19}$$

Otherwise choose the aggregate one.

where $|\widehat{\mathbf{\Omega}}_d|$ and $|\widehat{\mathbf{\Omega}}_a|$ are the "generalized variances" of the two different specifications. In fact, they represent a measure of the dispersion of the error terms. As already pointed out by the previous literature, it is worth noticing

---

[10] Similarly, when the vector y_{at} has been defined as a weighted average across the endogenous variables of the micro-units (i.e. $y_{at} = \sum_{i=1}^{m} \omega_i y_{it}$ with $\sum_{i=1}^{m} \omega_i = 1$) than:

$$\hat{\sigma}_d^2 = \sum_{i=1}^{m} \omega_i^2 \hat{\sigma}_i^2 + 2 \sum_{i>j}^{m} \omega_i \hat{\sigma}_{i,j} \omega_j \tag{18}$$

that the disaggregate model will fit better than the aggregate one if the micro-models are correctly specified. In particular, as shown by Lutkepohl(1987), under the hypothesis of correct specification of the disaggregate model the following property holds:

$$E(\boldsymbol{U}_d \boldsymbol{U}_d') \leq E(\boldsymbol{U}_a \boldsymbol{U}_a') \tag{20}$$

Nevertheless, in empirical research we are not aware about the Data Generation Processes of each single micro models, such that misspecification problems do usually arise. Therefore, the model selection criteria introduced in this section can be helpful tools in discriminating which is the model that better forecast the behavior of our aggregate variables.

4 A likelihood ratio test for model selection

In the previous sections we focused the attention on model selection criteria. Here it is worth noticing that model selection is a specific procedure that ends in a definite outcome. With this respect, the model selection criteria suggested in the previous aggregation literature allow the researcher to choose the best model in predicting aggregate variables. We believe that this model evaluation process is an important but restricted analysis. The use of an appropriate hypothesis test would allow investigating whether there is any statistically significant evidence of departure from a Null hypothesis focusing on the equal performances of two competitive models. The importance of testing relies on the fact that usually the researcher not only targets the best model predicting aggregate variables but also he aims to focus whether the two competitive models could be considered statistically equivalent. In addition, as already noted in Pesaran and Weeks (2001) on comparing model selection and hypothesis testing: *"only in the unlikely event that the true model is known or knowable will the selected model be universally applicable. In the real world where the truth is elusive and unknowable both approaches to model evaluation are worth pursuing"*. Before introducing the statistical test it is important to focus on the fact that the rival models under consideration are usually non-nested specifications. Hence, we need to employ a statistical procedure which take into account for this particular environment. According to the previous literature, we suggest using the test introduced by Vuong

(1989). The Vuong procedure is a useful tool to test the hypothesis that the models under consideration are equally close to the true model[11]. In general, disaggregate models are expected to outperform the aggregate models in absence of misspecification of the micro models. However, this outcome should not necessarily lead the researcher to discard the competitor aggregate model. Suppose for example that the disaggregate model outperform the aggregate one but the difference between the two is not statistically significant. Than the researcher could be induce to employ the aggregate model because of the more parsimonious information required. Therefore the use of a hypothesis testing procedure consists in expanding the empirical analysis on assessing predicting performances.

It is interesting to notice that a different hypothesis testing procedure has been employed for model selection in forecasting univariate models out of sample. In particular, Granger and Newbold (1986) suggest using the usual test for zero correlation, based on the sample correlation coefficient. Here, they assumed that the individual forecasts are unbiased and the forecast errors are not correlated. In addition, Lutkepohl (1987) considers the same procedure to compare the MSE of the aggregate and the disaggregate models when forecasting VARMA processes. However, we consider the zero correlation assumption almost unrealistic when analyzing the aggregation problem. In fact, the outcome of the aggregate model and that of the disaggregate one are likely to be strongly correlated. Thus we turned to consider a different hypothesis testing procedure[12].

We now concentrate the attention on the main features of the Vuong test statistic in order to show how it is linked with our aggregation context. First of all, Here it is assumed that a data generation process for the aggregate data is given, but it is unknown. Therefore, we look for the "best" model (between the aggregate and the disaggregate one) which is closer to the true process. Hence, Vuong (1989) suggested a likelihood-ratio test procedure.

---

[11] Note that the distance between the true process and a specified model is based on the minimum KLIC (Kullback-Leibler (1951) Information Criterion)

[12] Here it is worth emphasizing that our analysis focus on the in-sample predictive ability of the rival models. Thus, the Vuong procedure we suggest to use is not implemented for ranking models out of sample. Although we believe that the extension to the out of sample context is feasible, we decide to concentrate the attention only on the within sample goodness of fit analysis.

The LR test focuses on the following Null hypothesis:

$$H_0 : E\left(\log \frac{L_d(\boldsymbol{Y}_{at}, \boldsymbol{\theta})}{L_a(\boldsymbol{Y}_{at}, \boldsymbol{\gamma})}\right) = 0 \tag{21}$$

In other words, the H_0 hypothesis tests that the rival models[13] are *equivalent*, against

$$H_d : E\left(\log \frac{L_d(\boldsymbol{Y}_{at}, \boldsymbol{\theta})}{L_a(\boldsymbol{Y}_{at}, \boldsymbol{\gamma})}\right) > 0 \tag{22}$$

meaning that L_d is *better* than L_a, or

$$H_a : E\left(\log \frac{L_d(\boldsymbol{Y}_{at}, \boldsymbol{\theta})}{L_a(\boldsymbol{Y}_{at}, \boldsymbol{\gamma})}\right) < 0 \tag{23}$$

meaning, viceversa, that L_a is *better* than L_d. In particular the statistical test procedure can be expressed as follows:

$$v = T^{-\frac{1}{2}} \cdot \frac{LR(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\gamma}})}{\widehat{\omega}_T} \tag{24}$$

where

$$LR(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\gamma}}) = \sum_{t=1}^{T} \log \frac{L_d(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\theta}})}{L_a(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\gamma}})}) \tag{25}$$

and

$$\hat{\omega}_T^2 = \frac{1}{T}\sum_{t=1}^{T}\left(\log \frac{L_d(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\theta}})}{L_a(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\gamma}})}\right)^2 - \left(\frac{1}{T}\sum_{t=1}^{T}\log \frac{L_d(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\theta}})}{L_a(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\gamma}})}\right)^2 \tag{26}$$

As shown in Vuong (1989), under the Null-hypothesis, the previous statistic has the following distribution:

$$v \underset{D.}{\rightarrow} N(0,1) \tag{27}$$

We now focus on the aggregation context where the two competing models are the aggregate and the disaggregate specifications. More specifically, consider the Likelihood function of the aggregate model that can be written as follows:

$$L_a(\boldsymbol{\gamma}) = \left((2\Pi)^H |\boldsymbol{\Omega}_a|\right)^{-\frac{T}{2}} \cdot exp\left(-\frac{1}{2}\sum_{t=1}^{T} \boldsymbol{u}'_{at}\boldsymbol{\Omega}_a^{-1}\boldsymbol{u}_{at}\right) \tag{28}$$

Taking the logarithm of the previous formula we have:

---

[13] Here d and a stand for disaggregate and aggregate specifications.

$$l_a(\boldsymbol{\gamma}) = -\frac{TH}{2}\log(2\Pi) - \frac{T}{2}\log|\boldsymbol{\Omega}_a| - \frac{1}{2}\left(\sum_{t=1}^{T}\boldsymbol{u}'_{at}\boldsymbol{\Omega}_a^{-1}\boldsymbol{u}_{at}\right) \tag{29}$$

If we concentrate with respect to the maximum likelihood estimator of the parameters of interest we have the following concentrated log-Likelihood function:

$$l_a(\widehat{\gamma}) = -\frac{TH}{2}\log(2\Pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Omega}}_a| - \frac{TH}{2} \tag{30}$$

On the other hand, the "composite" likelihood function of the disaggregate model is:

$$L_d(\boldsymbol{\theta}) = \left((2\Pi)^H|\boldsymbol{\Omega}_d|\right)^{-\frac{T}{2}} \cdot exp\left(-\frac{1}{2}\sum_{t=1}^{T}\boldsymbol{u}'_{dt}\boldsymbol{\Omega}_d^{-1}\boldsymbol{u}_{dt}\right) \tag{31}$$

Taking the logarithm we have:

$$l_d(\boldsymbol{\theta}) = -\frac{TH}{2}\log(2\Pi) - \frac{T}{2}\log|\boldsymbol{\Omega}_d| - \frac{1}{2}\left(\sum_{t=1}^{T}\boldsymbol{u}'_{d,t}\boldsymbol{\Omega}_d^{-1}\boldsymbol{u}_{d,t}\right) \tag{32}$$

After concentrating with respect to the maximum likelihood estimator of the parameters of interest we have the following concentrated log-Likelihood function:

$$l_d(\widehat{\boldsymbol{\theta}}) = -\frac{TH}{2}\log(2\Pi) - \frac{T}{2}\log|\widehat{\boldsymbol{\Omega}}_d| - \frac{TH}{2} \tag{33}$$

Moreover, since the following property holds:

$$\frac{1}{T}LR(\widehat{\boldsymbol{\theta}}, \widehat{\gamma}) \underset{a.s.}{\rightarrow} E\left(\log\frac{L_d(\boldsymbol{Y}_{at}, \boldsymbol{\theta})}{L_a(\boldsymbol{Y}_{at}, \gamma)}\right) \tag{34}$$

and

$$\frac{1}{T}\sum_{t=1}^{T}\log\frac{L_d(\boldsymbol{Y}_{at}, \widehat{\boldsymbol{\theta}})}{L_a(\boldsymbol{Y}_{at}, \widehat{\gamma})} = \frac{1}{2}(\log|\widehat{\boldsymbol{\Omega}}_a| - \log|\widehat{\boldsymbol{\Omega}}_d|) \tag{35}$$

Therefore, the Null hypothesis of the Vuong test can be written as:

$$H_0 : \frac{1}{2}(\log|\boldsymbol{\Omega}_a| - \log|\boldsymbol{\Omega}_d|) = 0 \tag{36}$$

Or, more explicitly as:

$$H_0: \quad |\boldsymbol{\Omega}_a| = |\boldsymbol{\Omega}_d| \tag{37}$$

meaning that the two rival models are *equivalent*, against:

$$H_d: \quad |\boldsymbol{\Omega}_a| > |\boldsymbol{\Omega}_d| \tag{38}$$

that states that the disaggregate model is closer to the true process, or:

$$H_a: \quad |\boldsymbol{\Omega}_a| < |\boldsymbol{\Omega}_d| \tag{39}$$

that states that the aggregate specification is closer to the true process. In addition, since the log-likelihood functions of the aggregate and the disaggregate models, relative to the t-th observation, are respectively:

$$l_{at} = -\frac{H}{2}\log(2\Pi) - \frac{1}{2}\log|\boldsymbol{\Omega}_a| - \frac{1}{2}\boldsymbol{u}'_{at}\boldsymbol{\Omega}_a^{-1}\boldsymbol{u}_{at} \tag{40}$$

and

$$l_{dt} = -\frac{H}{2}\log(2\Pi) - \frac{1}{2}\log|\boldsymbol{\Omega}_d| - \frac{1}{2}\boldsymbol{u}'_{dt}\boldsymbol{\Omega}_d^{-1}\boldsymbol{u}_{dt} \tag{41}$$

Hence:

$$\log\frac{L_d(\boldsymbol{Y}_{at},\widehat{\boldsymbol{\theta}})}{L_a(\boldsymbol{Y}_{at},\widehat{\boldsymbol{\gamma}})} = \frac{1}{2}(\log|\widehat{\boldsymbol{\Omega}}_a| - \log|\widehat{\boldsymbol{\Omega}}_d|) - \frac{1}{2}\widehat{\boldsymbol{u}}'_{dt}\widehat{\boldsymbol{\Omega}}_d^{-1}\widehat{\boldsymbol{u}}_{dt} + \frac{1}{2}\widehat{\boldsymbol{u}}'_{at}\widehat{\boldsymbol{\Omega}}_a^{-1}\widehat{\boldsymbol{u}}_{at} \tag{42}$$

Such that the Vuong test statistic for the aggregation context can be written as follows:

$$v_a = T^{-\frac{1}{2}} \cdot \frac{\frac{T}{2}(\log|\widehat{\boldsymbol{\Omega}}_a| - \log|\widehat{\boldsymbol{\Omega}}_d|)}{[\frac{1}{T}\sum_{t=1}^{T}(l_{dt}(\widehat{\boldsymbol{\theta}}) - l_{at}(\widehat{\boldsymbol{\gamma}}))^2 - (\frac{1}{2}(\log|\widehat{\boldsymbol{\Omega}}_a| - \log|\widehat{\boldsymbol{\Omega}}_d|))^2]^{\frac{1}{2}}}$$

where:

$$l_{dt}(\widehat{\boldsymbol{\theta}}) - l_{at}(\widehat{\boldsymbol{\gamma}}) = \frac{1}{2}(\log|\widehat{\boldsymbol{\Omega}}_a| - \log|\widehat{\boldsymbol{\Omega}}_d|) - \frac{1}{2}\widehat{\boldsymbol{u}}'_{dt}\widehat{\boldsymbol{\Omega}}_d^{-1}\widehat{\boldsymbol{u}}_{dt} + \frac{1}{2}\widehat{\boldsymbol{u}}'_{at}\widehat{\boldsymbol{\Omega}}_a^{-1}\widehat{\boldsymbol{u}}_{at}$$

and:

$$v_a \underset{D.}{\rightarrow} N(0,1) \tag{43}$$

Note that the statistic v_a is easy to compute. It is in fact sufficient to compute the covariance matrices of the two specifications and the second moment of the likelihood ratio. The latter can be easily calculated by computing first the matrices:

$$\mathbf{G}_a = \underset{(T\times K)}{\widehat{\boldsymbol{U}}_a}{}' \underset{(K\times K)}{\widehat{\boldsymbol{\Omega}}_a^{-1}} \underset{(K\times T)}{\widehat{\boldsymbol{U}}_a} \quad and \quad \mathbf{G}_d = \underset{(T\times K)}{\widehat{\boldsymbol{U}}_d}{}' \underset{(K\times K)}{\widehat{\boldsymbol{\Omega}}_d^{-1}} \underset{(K\times T)}{\widehat{\boldsymbol{U}}_d}$$

Secondly we can extract the diagonal elements of the two matrices and set up the following vector:

$$\boldsymbol{r} = \boldsymbol{f} - \boldsymbol{g}_d + \boldsymbol{g}_a$$

where $\boldsymbol{f}$ is a $(T \times 1)$ vector of $\frac{1}{2}\log|\widehat{\boldsymbol{\Omega}}_d^{-1}\widehat{\boldsymbol{\Omega}}_a|$ terms and $\boldsymbol{g}_d$, $\boldsymbol{g}_a$ are $(T \times 1)$ vectors of diagonal elements of respectively the matrices $\boldsymbol{G}_d$ and $\boldsymbol{G}_a$. Therefore, the second moment can be estimated by computing $T^{-1}\boldsymbol{r}'\boldsymbol{r}$.

In the next section we evaluate the properties of the likelihood-ratio test by running a Monte Carlo simulation. The Vuong test statistic will be computed in order to analyse how the test behaves considering the aggregation of linear prediction models.

5 A Monte Carlo simulation

In this section we attempt to focus on how the Vuong test reacts whenever the number of observations is not considerably large. This is relevant in order to assess the validity of the test since empirical analyses are usually carried out using finite samples. Hence, a Monte Carlo simulation will try to shed light on the small sample properties of both the size and the power of our test statistic[14]. More specifically, the main object of the first part of our simulation is to compute and compare the empirical rejection frequencies of the test with those of the standard normal distribution under our particular assumptions. The second part of the simulation is devoted on investigating the power of the test when the disaggregate model perform better than the aggregate one. In the next two subsections we introduce the structures of the simulation taken into account.

---

[14] The Monte Carlo simulation has been run by using Ox version 3.00 (J.A. Doornik, 1994-2001)

5.1 Simulation analysis using non-nested models

In this subsection we introduce, step by step, the procedure used for running the Monte Carlo simulation when the aggregate and the disaggregate models are non-nested. The main feature of our analysis consists in the fact that both estimated models are not far from the Data Generation Process but none of them contains the true and hence, is not correctly specified. More specifically, both the micro and the macro estimated models are misspecified since they all miss some information (exogenous variables). We first describe the procedure to evaluate the empirical size of the test by calculating the rejection frequencies when the Null-hypothesis is true.

1^{st} *stage*:

We first generate the micro models alternatively as follows:

$$\boldsymbol{y}_{it} = \mathbf{c}_i + \boldsymbol{\beta}_{i1}\boldsymbol{x}_{it} + \boldsymbol{\beta}_{i2}\boldsymbol{z}_{it} + \boldsymbol{u}_{it};$$

$$\boldsymbol{y}_{jt} = \mathbf{c}_j + \boldsymbol{\theta}_{j1}\boldsymbol{x}_{jt} + \boldsymbol{\theta}_{j2}\boldsymbol{z}_{jt} + \boldsymbol{u}_{jt};$$

Here t=1,2,...,T with T=70, 150, 200, 300, 500, 1000. Moreover $\boldsymbol{x}$ and $\boldsymbol{z}$ are Uniformly distributed U(0,1) and $\boldsymbol{u} \sim N(0,1)$; $\mathbf{c} = \mathbf{2}$ and hence both $\boldsymbol{y}_{it} \sim I(0)$ and $\boldsymbol{y}_{jt} \sim I(0)$ are generated as stationary $(K \times 1)$ processes. In addition we let both the number of micro units and the number of endogenous variables vary such as $m = i+j$ =2, 3, 4, 5, 10, 20 and K=1, 2, 3, 4, 5. Note that $\boldsymbol{\beta}_i$ and $\boldsymbol{\theta}_i$ are $(K \times K)$ matrices of parameters set up alternatively as follows:

For K=1:

$$\beta_{i1} = 1,\ \beta_{i2} = 1\ , \theta_{i1} = .9,\ \theta_{j2} = .9$$

For K=2:

$$\boldsymbol{\beta}_{i1} = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix},\ \boldsymbol{\beta}_{i2} = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix},\ \boldsymbol{\theta}_{j1} = \begin{bmatrix} .9 & .4 \\ .4 & .9 \end{bmatrix},\ \boldsymbol{\theta}_{j2} = \begin{bmatrix} .9 & .4 \\ .4 & .9 \end{bmatrix},\ \ldots$$

For K=3:

$$\boldsymbol{\beta}_{i1} = \begin{bmatrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{bmatrix},\ \boldsymbol{\beta}_{i2} = \begin{bmatrix} 1 & .5 & .5 \\ .5 & 1 & .5 \\ .5 & .5 & 1 \end{bmatrix},$$

$$\boldsymbol{\theta}_{j1} = \begin{bmatrix} .9 & .4 & .4 \\ .4 & .9 & .4 \\ .4 & .4 & .9 \end{bmatrix},\ \boldsymbol{\theta}_{j2} = \begin{bmatrix} .9 & .4 & .4 \\ .4 & .9 & .4 \\ .4 & .4 & .9 \end{bmatrix},\ \ldots$$

We then estimate our micro models as follows:

$$\boldsymbol{y}_{it} = \widehat{\mathbf{c}}_i + \widehat{\boldsymbol{\beta}}_{i1}\boldsymbol{x}_{it} + \widehat{\boldsymbol{u}}_{it}$$

$$\boldsymbol{y}_{jt} = \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\theta}}_{j2}\boldsymbol{z}_{jt} + \widehat{\boldsymbol{u}}_{jt}$$

Note that each micro model is misspecified due to the omission of exogenous variables (alternatively $\boldsymbol{z}$ and $\boldsymbol{x}$). We than collect the micro residuals obtaining:

$$\widehat{\boldsymbol{u}}_{dt} = \widehat{\boldsymbol{u}}_{it} + \widehat{\boldsymbol{u}}_{jt}$$

2^{nd} *stage*:
Now we aggregate the endogenous variables of the macro-models such as:

$$\boldsymbol{y}_{at} = \boldsymbol{y}_{it} + \boldsymbol{y}_{jt};$$

We then estimate an aggregate specification as follows:

$$\boldsymbol{y}_{at} = \widehat{\mathbf{c}}_a + \widehat{\boldsymbol{\beta}}_a\boldsymbol{x}_{at} + \widehat{\boldsymbol{u}}_{at}$$

with

$$\boldsymbol{x}_{at} = \boldsymbol{x}_{it} + \boldsymbol{x}_{jt};$$

Note that this procedure allow us to evaluate the empirical size of the test when the Null-hypothesis is true[15].

We now turn to consider the power test analysis in order to observe the reaction of the test statistic whenever the Null hypothesis is false. We first generate the micro models as before. However, when estimating the micro model we allow the first micro regression to be correctly specified.

hence:

$$\boldsymbol{y}_{1t} = \widehat{\mathbf{c}}_1 + \widehat{\boldsymbol{\beta}}_{11}\boldsymbol{x}_{1t} + \widehat{\boldsymbol{\beta}}_{12}\boldsymbol{z}_{1t} + \widehat{\boldsymbol{u}}_{1t}$$

on the other hand, the remaining micro models are misspecified as before:

---

[15] In fact, let us assume the case when aggregating 2 micro units with K=1. In this specific case the disaggregate and the aggregate models estimate our vector of aggregate endogenous variables alternatively as follows:
$\widehat{\boldsymbol{y}}_{at} = \widehat{\mathbf{c}}_i + \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\beta}}_{11}\boldsymbol{x}_{1t} + \widehat{\boldsymbol{\theta}}_{22}\boldsymbol{z}_{2t}$ and $\widehat{\boldsymbol{y}}_{at} = \widehat{\mathbf{c}}_a + \widehat{\boldsymbol{\beta}}_a(\boldsymbol{x}_{1t} + \boldsymbol{x}_{2t})$. Both specifications make use of the constant and two exogenous variables which estimated coefficients are biased due to the omission of relevant variables. Hence, the distance (in the Kullback-Leibler sense) between the true aggregate models and the two competitive specification is approximatively similar and we can hardly claim that one model outperform the rival. Indeed, it is possible to see that for T=10000000 the following equality holds $|\boldsymbol{\Omega}_a| = |\boldsymbol{\Omega}_d|$ and hence H_0 is true.

$$\boldsymbol{y}_{it} = \widehat{\mathbf{c}}_i + \widehat{\boldsymbol{\beta}}_{i1}\boldsymbol{x}_{it} + \widehat{\boldsymbol{u}}_{it}$$

$$\boldsymbol{y}_{jt} = \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\theta}}_{j2}\boldsymbol{z}_{jt} + \widehat{\boldsymbol{u}}_{jt}$$

With $i \neq 1$. Furthermore, the aggregate endogenous variables are set up and estimated as before[16]. Here, we aim to consider the behavior of the test when the H_1 hypothesis is true, more specifically when $|\boldsymbol{\Omega}_a| > |\boldsymbol{\Omega}_d|$.

5.2 Simulation analysis using nested models

As already mentioned in the text, the use of the Vuong procedure is recommended whenever the two competing models are non-nested specifications. In general, it is likely that the aggregate and the micro specifications do not contain the same regressors. Nevertheless, it is interesting to investigate how the test react when the two competing models are nested. Therefore, in this subsection we describe a Monte Carlo exercise considering the two specifications using the same regressors.

First we generate the micro models alternatively as before. As a second step we estimate our micro models as follows:

$$\boldsymbol{y}_{it} = \widehat{\mathbf{c}}_i + \widehat{\boldsymbol{\beta}}_{i1}\boldsymbol{x}_{it} + \widehat{\boldsymbol{u}}_{it}$$

$$\boldsymbol{y}_{jt} = \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\theta}}_{j1}\boldsymbol{x}_{jt} + \widehat{\boldsymbol{u}}_{jt}$$

Note that each micro model is misspecified due to omission of the exogenous variables $\boldsymbol{z}$. We than collect the micro residuals obtaining:

$$\widehat{\boldsymbol{u}}_{dt} = \widehat{\boldsymbol{u}}_{it} + \widehat{\boldsymbol{u}}_{jt}$$

Regarding the aggregate model we used the same procedure adopted in the previous subsection. This exercise allow us to investigate on the empirical size of the test when the Null-hypothesis is true[17]

Regarding the power test analysis, we used the same framework as before. In particular, the first micro equation is correctly estimated while both the

---

[16] using only $\boldsymbol{x}_{at} = \boldsymbol{x}_{it} + \boldsymbol{x}_{jt}$ as regressors.

[17] Here it is more evident that $|\boldsymbol{\Omega}_a| = |\boldsymbol{\Omega}_d|$. In fact, assuming that m=2 and K=1 than the disaggregate and the aggregate models estimate our vector of aggregate endogenous variables alternatively as follows:
$\widehat{\boldsymbol{y}}_{at} = \widehat{\mathbf{c}}_i + \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\beta}}_{11}\boldsymbol{x}_{1t} + \widehat{\boldsymbol{\beta}}_{12}\boldsymbol{x}_{2t}$ and $\widehat{\boldsymbol{y}}_{at} = \widehat{\mathbf{c}}_a + \widehat{\boldsymbol{\beta}}_a(\boldsymbol{x}_{1t} + \boldsymbol{x}_{2t})$. Both specifications use the same regressors with the aggregate model imposing the same coefficient to both $\boldsymbol{x}_{1t}$ and $\boldsymbol{x}_{2t}$.

remaining micro equations and the aggregate model are misspecified. Such that:

$$\boldsymbol{y}_{1t} = \widehat{\mathbf{c}}_1 + \widehat{\boldsymbol{\beta}}_{11}\boldsymbol{x}_{1t} + \widehat{\boldsymbol{\beta}}_{12}\boldsymbol{z}_{1t} + \widehat{\boldsymbol{u}}_{1t}$$

and:

$$\boldsymbol{y}_{it} = \widehat{\mathbf{c}}_i + \widehat{\boldsymbol{\beta}}_{i1}\boldsymbol{x}_{it} + \widehat{\boldsymbol{u}}_{it}$$

$$\boldsymbol{y}_{jt} = \widehat{\mathbf{c}}_j + \widehat{\boldsymbol{\theta}}_{j2}\boldsymbol{x}_{jt} + \widehat{\boldsymbol{u}}_{jt}$$

while:

$$\boldsymbol{y}_{at} = \widehat{\mathbf{c}}_a + \widehat{\boldsymbol{\beta}}_a\boldsymbol{x}_{at} + \widehat{\boldsymbol{u}}_{at}$$

With $i \neq 1$. Also in the nested case we investigated the power of the test when the non-null hypothesis $|\boldsymbol{\Omega}_a| > |\boldsymbol{\Omega}_d|$ is true.

5.3 Results

5.3.1 Size of the test

In this subsection we focus on Tables 1 and 2 showing the empirical rejection frequencies of the computed Likelihood-ratio test when the Null-hypothesis is true. The reported values have to be compared with those of the Standard normal distribution with nominal rejection frequencies $\alpha = 0.01, \alpha = 0.05$. Furthermore, each cell of the tables reports paired results relative to both the non-nested and the nested case. Despite the huge amount of results, it is possible to shed light on the performance of the test statistic. First of all it can be easily seen that the empirical rejection frequencies of Vuong's testing procedure are particularly similar to the nominal ones when the rival models are non-nested specifications. This is an important result showing the good performance of the test when the number of observations is not particularly high. On the other hand, it is possible to note that the Vuong test tends to under-reject the null-hypothesis whenever we employ nested models. This is specially true when the number of observations is below T=500. Most likely, the reason relies on the fact that, as already observed before, in the nested case the aggregate and the disaggregate models employ the same exogenous variables and therefore the test has more difficulties to reject the Null-hypothesis. However, for $T > 500$ the rejection frequencies are not far from the nominal values and, for T=1000, they also tend to over-reject the Null-hypothesis.

5.3.2 Power test analysis

We now concentrate the attention on the power test analysis. Note that for Table 3 the minimum power is equal to 0.01[18] while for Table 4 the minimum power is equal to 0.05. We focused on the power of the test when $H_1 : |\mathbf{\Omega}_a| > |\mathbf{\Omega}_d|$ is true, in other words the disaggregate is closer to the Data Generation Process compared with the aggregate specification. As before results relative to the non-nested and the nested cases are reported. Here, it is possible to observe that the empirical power of the test increases whenever m decreases (going from 20 to 2). This is an expected results, since the less the micro units, the less the exogenous variables, the more relevant the role of the first correctly specified micro regression in explaining the aggregate endogenous vector. Moreover, the power increases together with the K. This is also reasonable since the higher K, the higher the number of micro exogenous variables in the micro models, the more relevant the explanatory role of the first correctly specified model. Here, it is particularly evident the rise of the power whenever we increase K from 1 to 2. In addition, the empirical power increase together with the sample size. It is also important to observe that the power of the test is higher for the nested case. In particular, it is possible to note that in same cases the empirical power of the nested case is double or more than double the power of the non-nested case. This result is not unexpected. In fact, the correct specification of the first micro regression has a stronger impact when the aggregate and the disaggregate models use the same explanatory variables. More specifically, the test tends to detect more easily the presence of a correctly specified model, compared with the non-nested case, and hence the better performance of the disaggregate specification. Thus, we can conclude that the Vuong test seems to have very good performance both for the size and the power test analysis. This is true not only for the non-nested but also for the nested case. We believe that this result is relevant since it legitimates the use of the Vuong test in the aggregation context regardless the type of specification of the two rival models.

---

[18] Meaning that, when the Null-hypothesis is true, the minimum power is 0.01

6 Conclusions

In this paper we attempted to expand the "model selection" analysis in predicting aggregate variables, when the micro units consist in both univariate and multivariate specifications. More specifically, we both reviewed the classical Grunfeld and Griliches selection criterion and extended the same criterion to the case when the number of endogenous variables is greater than 1 by comparing the generalized variances of the rival models. In addition, we introduced a statistical procedure for model selection. The likelihood ratio test considered is that suggested by Vuong (1989). A Monte Carlo simulation has shown good properties of the test for different sample sizes when the aggregate and the disaggregate are both non-nested and nested specifications. This is an important result legitimating the use of the Vuong test in the aggregation context regardless the specification of the rival models. The paper focused on the in-sample prediction analysis of aggregate variables. Further research is needed to extend the analysis when predicting aggregate variables out-of sample.

References

[1] Aigner D. J. and Goldfeld S. G. (1973), "Simulation and Aggregation: A reconsideration", *Review of Economics and Statistics*, 5, 3, 114-118

[2] Aigner D. J., and Goldfeld S. G. (1974), "Estimation and prediction from aggregate data when aggregates are measured more accurately than their components", *Econometrica*, 42, 1, 113-133

[3] Baffigi A., Golinelli R. and Parigi G. (2002), "Real-time GDP forecasting in the Euro area", Banca d'Italia, Temi di Discussione, 456

[4] Baffigi A., Golinelli R. and Parigi G. (2004), "Bridge models to forecast the euro area GDP", *International Journal of Forecasting*

[5] Barker T. and Pesaran M. H., (1990), *Disaggregation in Econometric Modelling*, Routledge, London

[6] Den Butten F. A. G., and Van Djiken S. (1997), "The information contents of aggregate money demand in the EMU", *Open economies Review*, 8(3), 233-244

[7] Fagan G. and Henry J. (1998), Long run money demand in euro area, *Empirical Economics*,Vol. 23, 483-506

[8] Fisher W. D. (1969), *Clustering and Aggregation in Economics*, Baltimore: Johns Hopkins University Press.

[9] Dedola L., Gaiotti E. and Silipo L. (2001), Money Demand in the Euro Area: Do National Differences Matter?, Banca d'Italia, Temi di Discussione, 405

[10] Garderen K. J. V., Lee K. and Pesaran M. H. (2000), "Cross-sectional aggregation of non-linear models", *Journal of Econometrics*,95, 285-331

[11] Golinelli R. and Pastorello S. (2002), Modelling the demand for M3 in the Euro Area, *The European Journal of Finance*, Vol. 8, N. 4, 371-401

[12] Gorman, W.M. (1953), "Community preference fields", *Econometrica*, 21, 63-80

[13] Granger C. and Newbold P. (1986), *Forecasting Economic Time Series*, 2nd Edn., New York: Academic Press.

[14] Grunfeld Y. and Griliches Z. (1960), "Is aggregation necessarily bad?", *Review of Economics and Statistics*, 42, 1-13

[15] Hendry D. F, and Hubrich K. (2005), "Forecasting Aggregates by Disaggregates", *Computing in Economics and Finance 2005*, 270, Society for Computational Economics

[16] Hsiao C., Shen Y. and Fujiki H. (2005), "Aggregate vs. disaggregate data analysis-a paradox in the estimation of a money demand function of Japan under the low interest rate policy", *Journal of Applied Econometrics*, 20,579-601

[17] Hubrich K. (2004),"Forecasting euro area inflation: Does aggregating forecasts by HICP components improve forecat accuracy?", *Computing in Economics and Finance 2004*, 230, Society for Computational Economics

[18] Kelejian, H.H., (1980), "Aggregation and disaggregation of non-linear equations", in Kmenta, J. and Ramsay, J.B. (Eds.), Evaluation of econometric models, Academic Press, New York

[19] Lutkepohl, H, (1987), *Forecasting aggregated vector ARMA process*, Berlin: Springer-Verlag.

[20] Maddala G. S. (1977), *Econometrics*, McGraw-Hill, 207-217

[21] Marcellino M., Stock J. H. and Watson M. W. (2003), "Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information", *European Economic Review*

[22] Orcutt G. H., Watts H.W. and Edwards J.B. (1968), "Data aggregation and information loss", *American Economic Review*, 773-787

[23] Pesaran M. H., Pierse R. G., Kumar M. S. (1989), "Econometric analysis of aggregation in the context of linear prediction models", *Econometrica*, 57, 4, 861-888

[24] Pesaran M. H., Pierse R. G. and Lee K. (1994), "Choice between disaggregate and aggregate specifications estimated by instrumental variables methods", *Journal of Business and Economic Statistics*, 12, 1, 11-21

[25] Pesaran M. H. and Weeks M. (2001), "Non-Nested Hyothesis Testing: An Overview", In Badi H Baltagi (Eds.), Companion to Theoretical Econometrics, Basil Blackwell, Oxford

[26] Pesaran M. H. (2003), "Aggregation of Linear Dynamic Models: An Application to the Life-Cycle Consumption Models Under Habit Formation", *Economic Modelling*, 20, 2, 227-435

[27] Stoker M.T. (1984), "Completeness, distribution restrictions, and the form of aggregate functions", *Econometrica*, 52, 887-907

[28] Stoker M.T. (1993), "Empirical Approaches to the Problem of Aggregation Over Individuals", *Journal of Economic Literature*, 31, 1827-1874

[29] Theil H. (1954), *Linear Aggregation of Economic Relations*, North Holland, Amsterdam

[30] Vuong (1989), "Likelihood ratio tests for model selection and non-nested hypothesis", *Econometrica*, 57, 2, 307-333

[31] Wesche K. (1997), The stability of European money demand: an investigation of M3H, *Open Economies Review*, 8, 371-391

Appendix

Table 1. Rejection frequencies (with $\alpha = 0.01$)

Non-Nested; Nested

		K=1	K=2	K=3	K=4	K=5
T=70	m=2	*.008 ; .001*	*.008 ; .000*	*.010 ; .002*	*.008 ; .004*	*.010 ; .008*
	m=3	*.006 ; .004*	*.009 ; .006*	*.008 ; .004*	*.006 ; .002*	*.011 ; .002*
	m=4	*.011 ; .004*	*.013 ; .005*	*.007 ; .004*	*.013 ; .003*	*.012 ; .003*
	m=5	*.009 ; .005*	*.011 ; .008*	*.009 ; .006*	*.010 ; .007*	*.008 ; .004*
	m=10	*.011 ; .007*	*.010 ; .007*	*.013 ; .006*	*.010 ; .004*	*.009 ; .006*
	m=20	*.010 ; .014*	*.008 ; .009*	*.012 ; .008*	*.005 ; .005*	*.011 ; .007*
T=150	m=2	*.008 ; .002*	*.010 ; .000*	*.007 ; .000*	*.011 ; .000*	*.010 ; .008*
	m=3	*.008 ; .002*	*.01 ; .002*	*.009 ; .004*	*.014 ; .004*	*.011 ; .002*
	m=4	*.012 ; .008*	*.012 ; .004*	*.010 ; .006*	*.010 ; .006*	*.012 ; .004*
	m=5	*.007 ; .004*	*.013 ; .006*	*.012 ; .009*	*.013 ; .006*	*.009 ; .007*
	m=10	*.010 ; .008*	*.011 ; .013*	*.011 ; .008*	*.007 ; .008*	*.013 ; .007*
	m=20	*.011 ; .010*	*.012 ; .012*	*.011 ; .011*	*.009 ; .010*	*.007 ; .008*
T=200	m=2	*.008 ; .002*	*.007 ; .002*	*.011 ; .002*	*.009 ; .003*	*.014 ; .007*
	m=3	*.008 ; .003*	*.006 ; .004*	*.009 ; .004*	*.012 ; .003*	*.012 ; .002*
	m=4	*.014 ; .005*	*.010 ; .004*	*.009 ; .004*	*.012 ; .004*	*.015 ; .006*
	m=5	*.009 ; .008*	*.013 ; .006*	*.011 ; .006*	*.011 ; .003*	*.010 ; .006*
	m=10	*.012 ; .010*	*.017 ; .011*	*.009 ; .006*	*.006 ; .006*	*.010 ; .006*
	m=20	*.012 ; .011*	*.011 ; .014*	*.010 ; .008*	*.010 ; .012*	*.011 ; .007*
T=300	m=2	*.007 ; .002*	*.012 ; .001*	*.011 ; .004*	*.011 ; .006*	*.012 ; .011*
	m=3	*.011 ; .003*	*.008 ; .002*	*.008 ; .005*	*.012 ; .002*	*.011 ; .003*
	m=4	*.016 ; .007*	*.008 ; .002*	*.015 ; .007*	*.011 ; .004*	*.011 ; .007*
	m=5	*.013 ; .006*	*.013 ; .006*	*.011 ; .004*	*.011 ; .007*	*.014 ; .006*
	m=10	*.011 ; .008*	*.009 ; .011*	*.008 ; .006*	*.013 ; .011*	*.013 ; .007*
	m=20	*.010 ; .008*	*.014 ; .012*	*.010 ; .014*	*.012 ; .010*	*.014 ; .011*
T=500	m=2	*.010 ; .002*	*.009 ; .004*	*.010 ; .003*	*.012 ; .004*	*.013 ; .013*
	m=3	*.009 ; .002*	*.012 ; .007*	*.009 ; .006*	*.010 ; .008*	*.013 ; .007*
	m=4	*.010 ; .008*	*.013 ; .008*	*.008 ; .006*	*.010 ; .011*	*.011 ; .007*
	m=5	*.010 ; .006*	*.010 ; .012*	*.009 ; .006*	*.012 ; .010*	*.010 ; .009*
	m=10	*.012 ; .007*	*.010 ; .012*	*.011 ; .011*	*.010 ; .012*	*.013 ; .012*
	m=20	*.015 ; .013*	*.012 ; .011*	*.010 ; .011*	*.014 ; .015*	*.011 ; .013*
T=1000	m=2	*.012 ; .002*	*.011 ; .008*	*.012 ; .010*	*.009 ; .011*	*.011 ; .017*
	m=3	*.011 ; .002*	*.015 ; .010*	*.012 ; .018*	*.011 ; .014*	*.011 ; .015*
	m=4	*.012 ; .007*	*.013 ; .011*	*.013 ; .016*	*.014 ; .019*	*.011 ; .022*
	m=5	*.012 ; .010*	*.012 ; .011*	*.009 ; .018*	*.012 ; .014*	*.012 ; .021*
	m=10	*.012 ; .010*	*.015 ; .015*	*.011 ; .007*	*.012 ; .014*	*.008 ; .023*
	m=20	*.012 ; .010*	*.009 ; .013*	*.013 ; .015*	*.015 ; .011*	*.007 ; .017*

Table 2. Rejection frequencies with ($\alpha = 0.05$)
Non-Nested; Nested

		K=1	K=2	K=3	K=4	K=5
T=70	m=2	*.042 ; .018*	*.042 ; .022*	*.054 ; .023*	*.054 ; .024*	*.062 ; .017*
	m=3	*.048 ; .030*	*.054 ; .040*	*.061 ; .030*	*.058 ; .031*	*.065 ; .019*
	m=4	*.053 ; .038*	*.055 ; .035*	*.050 ; .035*	*.061 ; .030*	*.060 ; .025*
	m=5	*.043 ; .040*	*.062 ; .041*	*.047 ; .035*	*.048 ; .038*	*.052 ; .028*
	m=10	*.057 ; .055*	*.063 ; .041*	*.060 ; .041*	*.055 ; .038*	*.058 ; .039*
	m=20	*.057 ; .063*	*.060 ; .050*	*.060 ; .052*	*.049 ; .045*	*.062 ; .048*
T=150	m=2	*.053 ; .015*	*.056 ; .017*	*.052 ; .015*	*.064 ; .022*	*.060 ; .019*
	m=3	*.039 ; .023*	*.053 ; .021*	*.054 ; .028*	*.060 ; .025*	*.058 ; .024*
	m=4	*.054 ; .042*	*.051 ; .034*	*.050 ; .033*	*.059 ; .038*	*.064 ; .030*
	m=5	*.045 ; .036*	*.050 ; .039*	*.054 ; .040*	*.055 ; .028*	*.052 ; .036*
	m=10	*.056 ; .043*	*.048 ; .046*	*.053 ; .045*	*.047 ; .043*	*.060 ; .039*
	m=20	*.059 ; .054*	*.046 ; .049*	*.056 ; .050*	*.052 ; .050*	*.048 ; .057*
T=200	m=2	*.050 ; .013*	*.052 ; .024*	*.047 ; .024*	*.050 ; .017*	*.053 ; .022*
	m=3	*.043 ; .024*	*.056 ; .029*	*.044 ; .034*	*.051 ; .031*	*.062 ; .026*
	m=4	*.056 ; .033*	*.059 ; .038*	*.054 ; .029*	*.057 ; .036*	*.059 ; .033*
	m=5	*.054 ; .035*	*.062 ; .043*	*.055 ; .040*	*.052 ; .032*	*.049 ; .035*
	m=10	*.056 ; .049*	*.059 ; .049*	*.050 ; .043*	*.041 ; .043*	*.060 ; .044*
	m=20	*.054 ; .051*	*.058 ; .059*	*.051 ; .044*	*.047 ; .049*	*.046 ; .047*
T=300	m=2	*.048 ; .016*	*.053 ; .020*	*.055 ; .030*	*.058 ; .032*	*.055 ; .030*
	m=3	*.052 ; .026*	*.042 ; .034*	*.056 ; .035*	*.051 ; .031*	*.053 ; .036*
	m=4	*.056 ; .039*	*.051 ; .036*	*.057 ; .040*	*.055 ; .039*	*.054 ; .033*
	m=5	*.051 ; .047*	*.052 ; .043*	*.057 ; .039*	*.055 ; .042*	*.061 ; .034*
	m=10	*.055 ; .044*	*.047 ; .051*	*.041 ; .051*	*.057 ; .048*	*.053 ; .043*
	m=20	*.054 ; .043*	*.057 ; .061*	*.050 ; .059*	*.057 ; .060*	*.059 ; .058*
T=500	m=2	*.046 ; .016*	*.052 ; .034*	*.056 ; .038*	*.058 ; .036*	*.062 ; .039*
	m=3	*.047 ; .027*	*.058 ; .041*	*.048 ; .031*	*.053 ; .042*	*.060 ; .045*
	m=4	*.055 ; .037*	*.054 ; .050*	*.046 ; .041*	*.050 ; .062*	*.056 ; .043*
	m=5	*.058 ; .054*	*.062 ; .051*	*.051 ; .054*	*.065 ; .057*	*.058 ; .055*
	m=10	*.056 ; .043*	*.058 ; .062*	*.050 ; .064*	*.050 ; .058*	*.057 ; .051*
	m=20	*.058 ; .060*	*.050 ; .056*	*.054 ; .054*	*.059 ; .066*	*.064 ; .048*
T=1000	m=2	*.055 ; .020*	*.062 ; .050*	*.059 ; .067*	*.058 ; .079*	*.063 ; .095*
	m=3	*.056 ; .032*	*.064 ; .059*	*.054 ; .073*	*.050 ; .075*	*.058 ; .080*
	m=4	*.061 ; .040*	*.046 ; .067*	*.060 ; .076*	*.065 ; .093*	*.059 ; .102*
	m=5	*.052 ; .043*	*.053 ; .063*	*.057 ; .080*	*.050 ; .078*	*.056 ; .096*
	m=10	*.057 ; .055*	*.063 ; .072*	*.055 ; .080*	*.060 ; .089*	*.053 ; .103*
	m=20	*.053 ; .049*	*.055 ; .062*	*.055 ; .060*	*.055 ; .071*	*.055 ; .080*

Table 3. Power of the test when H_1: $|\boldsymbol{\Omega}_a| > |\boldsymbol{\Omega}_d|$ is true and $\alpha = 0.01$
Non-Nested; Nested

		K=1	K=2	K=3	K=4	K=5
T=70	m=2	*.024 ; .030*	*.881 ; .976*	*.941 ; .995*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.025 ; .022*	*.377 ; .713*	*.661 ; .952*	*.860 ; .986*	*.952 ; .999*
	m=4	*.023 ; .017*	*.224 ; .494*	*.328 ; .780*	*.644 ; .915*	*.758 ; .978*
	m=5	*.017 ; .015*	*.113 ; .339*	*.143 ; .476*	*.334 ; .768*	*.431 ; .860*
	m=10	*.018 ; .018*	*.037 ; .093*	*.047 ; .140*	*.076 ; .250*	*.107 ; .302*
	m=20	*.016 ; .017*	*.019 ; .046*	*.026 ; .041*	*.031 ; .068*	*.033 ; .077*
T=150	m=2	*.049 ; .069*	*.968 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.034 ; .058*	*.751 ; .979*	*.969 ; 1.00*	*.998 ; 1.00*	*1.00 ; 1.00*
	m=4	*.028 ; .044*	*.547 ; .915*	*.722 ; .994*	*.962 ; 1.00*	*.993 ; 1.00*
	m=5	*.024 ; .036*	*.231 ; .749*	*.362 ; .928*	*.712 ; .996*	*.822 ; 1.00*
	m=10	*.013 ; .019*	*.063 ; .313*	*.089 ; .461*	*.172 ; .748*	*.236 ; .835*
	m=20	*.017 ; .015*	*.030 ; .116*	*.031 ; .137*	*.047 ; .285*	*.055 ; .334*
T=200	m=2	*.062 ; .121*	*.996 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.049 ; .082*	*.879 ; .998*	*.995 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.039 ; .066*	*.704 ; .976*	*.970 ; .999*	*.993 ; 1.00*	*.998 ; 1.00*
	m=5	*.029 ; .048*	*.349 ; .892*	*.488 ; .982*	*.860 ; 1.00*	*.941 ; 1.00*
	m=10	*.018 ; .034*	*.077 ; .481*	*.113 ; .658*	*.248 ; .917*	*.340 ; .956*
	m=20	*.019 ; .019*	*.032 ; .179*	*.039 ; .231*	*.062 ; .419*	*.078 ; .526*
T=300	m=2	*.096 ; .208*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.059 ; .129*	*.976 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.044 ; .085*	*.880 ; .997*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.035 ; .062*	*.506 ; .981*	*.703 ; .999*	*.967 ; 1.00*	*.988 ; 1.00*
	m=10	*.020 ; .026*	*.117 ; .705*	*.163 ; .885*	*.381 ; .993*	*.494 ; .999*
	m=20	*.010 ; .018*	*.044 ; .304*	*.046 ; .457*	*.086 ; .721*	*.131 ; .829*
T=500	m=2	*.164 ; .436*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.095 ; .257*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.059 ; .182*	*.978 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.048 ; .139*	*.777 ; 1.00*	*.915 ; 1.00*	*.999 ; 1.00*	*1.00 ; 1.00*
	m=10	*.027 ; .061*	*.198 ; .929*	*.310 ; .992*	*.629 ; 1.00*	*.766 ; 1.00*
	m=20	*.017 ; .042*	*.066 ; .578*	*.071 ; .765*	*.142 ; .961*	*.206 ; .989*
T=1000	m=2	*.376 ; .854*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.211 ; .584*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.119 ; .450*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.094 ; .343*	*.977 ; 1.00*	*.999 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=10	*.033 ; .139*	*.419 ; 1.00*	*.605 ; 1.00*	*.927 ; 1.00*	*.979 ; 1.00*
	m=20	*.018 ; .057*	*.116 ; .930*	*.145 ; .988*	*.308 ; 1.00*	*.433 ; 1.00*

Table 4. Power of the test when H_1: $|\boldsymbol{\Omega}_a| > |\boldsymbol{\Omega}_d|$ is true and $\alpha = 0.05$

Non-Nested; Nested

		K=1	K=2	K=3	K=4	K=5
T=70	m=2	*.116 ; .127*	*.975 ; .998*	*.999 ; .999*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.107 ; .113*	*.659 ; .903*	*.882 ; .993*	*.966 ; .999*	*.988 ; 1.00*
	m=4	*.086 ; .106*	*.508 ; .763*	*.632 ; .951*	*.856 ; .985*	*.930 ; .997*
	m=5	*.080 ; .078*	*.313 ; .618*	*.373 ; .761*	*.606 ; .932*	*.728 ; .970*
	m=10	*.072 ; .067*	*.149 ; .281*	*.173 ; .369*	*.257 ; .520*	*.304 ; .610*
	m=20	*.069 ; .074*	*.098 ; .156*	*.109 ; .172*	*.124 ; .232*	*.143 ; .256*
T=150	m=2	*.183 ; .268*	*.999 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.137 ; .188*	*.921 ; .996*	*.993 ; 1.00*	*.999 ; 1.00*	*1.00 ; 1.00*
	m=4	*.110 ; .154*	*.795 ; .982*	*.911 ; .999*	*.994 ; 1.00*	*.999 ; 1.00*
	m=5	*.097 ; .142*	*.504 ; .919*	*.643 ; .985*	*.896 ; .999*	*.948 ; 1.00*
	m=10	*.067 ; .096*	*.203 ; .588*	*.246 ; .739*	*.434 ; .919*	*.504 ; .955*
	m=20	*.072 ; .094*	*.116 ; .304*	*.117 ; .383*	*.160 ; .557*	*.182 ; .615*
T=200	m=2	*.220 ; .364*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.170 ; .251*	*.967 ; .999*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.150 ; .206*	*.885 ; .996*	*.965 ; 1.00*	*.998 ; 1.00*	*1.00 ; 1.00*
	m=5	*.125 ; .171*	*.619 ; .973*	*.748 ; .997*	*.967 ; 1.00*	*.989 ; 1.00*
	m=10	*.074 ; .112*	*.224 ; .747*	*.307 ; .868*	*.505 ; .982*	*.623 ; .994*
	m=20	*.064 ; .084*	*.122 ; .402*	*.144 ; .484*	*.195 ; .706*	*.242 ; .789*
T=300	m=2	*.288 ; .522*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.191 ; .362*	*.994 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.146 ; .278*	*.973 ; 1.00*	*.997 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.123 ; .211*	*.779 ; .996*	*.889 ; 1.00*	*.996 ; 1.00*	*.999 ; 1.00*
	m=10	*.091 ; .127*	*.293 ; .895*	*.411 ; .973*	*.649 ; .998*	*.753 ; 1.00*
	m=20	*.064 ; .101*	*.151 ; .572*	*.158 ; .730*	*.249 ; .899*	*.325 ; .958*
T=500	m=2	*.402 ; .756*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.265 ; .543*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.188 ; .420*	*.999 ; 1.00*	*.999 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.155 ; .350*	*.922 ; 1.00*	*.981 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=10	*.093 ; .212*	*.444 ; .986*	*.563 ; .999*	*.844 ; 1.00*	*.927 ; 1.00*
	m=20	*.073 ; .137*	*.191 ; .814*	*.224 ; .924*	*.362 ; .991*	*.442 ; .997*
T=1000	m=2	*.647 ; .970*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=3	*.470 ; .841*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=4	*.313 ; .719*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=5	*.252 ; .629*	*.995 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*	*1.00 ; 1.00*
	m=10	*.124 ; .354*	*.682 ; .999*	*.823 ; 1.00*	*.982 ; 1.00*	*.995 ; 1.00*
	m=20	*.094 ; .191*	*.312 ; .983*	*.370 ; .998*	*.587 ; 1.00*	*.691 ; 1.00*