

PhD thesis title: **Variance screen as a tool for detecting price collusion. Fundamentals and meaning of the imputation for missing data**
Author: Adam Korczyński

The PhD thesis titled “Variance screen as a tool for detecting price collusion. Fundamentals and meaning of the imputation for missing data” is a part of the research on the economic tools for detecting cartels. The study is focused on missing data as one of the practical aspects of data analysis in variance screen for collusion. The thesis consists of introduction, five chapters and conclusions. The theoretical part outlines the economics of cartels and contains an overview of the findings on the variance of prices under price collusion. The second part considers the methods and models used for missing data analysis, which were described and assessed based on the properties of the estimates provided by these methods. The empirical part presents the use of imputation techniques in the variance screen for collusion conducted for a group of petrol stations. The empirical study contains a comparison of the estimates provided by two independent imputation techniques: multiple imputation and the author’s original version of the expectation-maximization (EM) algorithm supplemented by Newton-Raphson algorithm.

Research hypothesis:

According to the research questions and goals stipulated in the thesis, the following hypotheses were tested:

1. Variance screen makes it possible to identify price collusion.
2. Data imputation techniques reduce the bias of the estimates in variance screen if data are missing at random.
3. Author’s original version of the expectation-maximization algorithm (EM) is a more computationally efficient tool reducing bias in variance screen than multiple imputation.

Research methods:

The research was based on three sources of data. The first hypothesis has been verified based on the time series of prices in industries where price collusions were known to have taken place. The second and third hypothesis have been verified using simulation study, where a random sample of prices have been generated out of a statistical model describing changes of prices over time, as well as time series data on prices recorded for a group of petrol stations.

Conclusions:

The results of the empirical studies and the analysis of cartel of Polish cement producers provide evidence for periods of smaller variance of prices under collusive regimes. However small variance of prices is observed only during certain period of the collusion. Low variance is expected to characterize the second phase of collusion starting when the price stabilizes at high level after a phase of price increase. We may also observe higher variance as a result of behaviour of individual companies, which might want to increase their short-term profits by enlarging the market share. The first hypothesis, stating that price collusions can be detected by using variance screen, was proven to be partially true.

Both multiple imputation and EM algorithm reduce bias caused by missing data. This observation proved the second hypothesis to be true. The empirical analysis has shown the differences between the estimates calculated using complete cases and using the imputation techniques. If data are missing at random the complete case analysis will provide biased estimates and will lead to false conclusions.

The EM algorithm is more efficient than multiple imputation in terms of the computational requirements. Both methods provide consistent estimates. The advantage of the EM algorithm over multiple imputation lies in its deterministic character. It does not require multiple draws from posterior distributions, which are built-in the multiple imputation technique. Multiple imputation adjusts the standard errors and provides confidence intervals corrected for the uncertainty due to missing data. In this respect multiple imputation surpasses the EM algorithm. The third hypothesis was found to be partially true.

Variance screen as a tool for detecting collusion requires further research and verification. One of the aspects to be verified is the identification of the collusion phases when prices are stabilized, as opposed to the phase of price increase. Multiple imputation, as well as the EM algorithm, allow to restore the information lost due to missing data and in that sense the imputation techniques make the variance screen practically useful.

Both multiple imputation and the EM algorithm can be applied to wide range of estimation problems where samples are incomplete. One of the possible further developments of the EM algorithm supplemented by Newton-Raphson procedure would be to expand the method in a way which would enable its use in multivariate models. The complexity of economic phenomena requires incorporating more elaborate relationships observed in the time series of prices. The prices might be affected by some parallel process, which would need to be incorporated in the imputation models. The version of the EM algorithm formulated in the thesis can be developed to allow for correction of the standard errors. In terms of multiple imputation technique, the method requires further research on the stability of the estimates provided by Gibbs sampling used for imputing purpose.

The variance screen would need to be studied further to identify the optimal windows for observing the prices. The theory and empirical studies have proven cartels to be inherently unstable. The identification of collusion would require one to consider the phases of collusive regimes. Further research is needed to define the transformations of the original time series of prices to be applied in variance screen. Lastly, the variance screen itself requires further verification with more empirical data. This will require more observational studies and analyses to be conducted by the antitrust agencies. The results of the controls carried out by the competition protection institutions are basis for identification of industries and periods of collusion which can be then compared against the external sources of data provided for example by the statistical offices.