

Article

Not peer-reviewed version

A Statistical Model for Identifying Collusive Bidding Based on Historical Quotations

Jieni Li , Fei Xie , [Hongchao Guo](#) *

Posted Date: 8 January 2024

doi: 10.20944/preprints202401.0487.v1

Keywords: collusive bidding; historical quotations; coefficient of variation; deviation rate; statistical model



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

A Statistical Model for Identifying Collusive Bidding Based on Historical Quotations

Jieni Li ¹, Fei Xie ² and Hongchao Guo ^{1,*}

¹ Material and Equipment Center of Yanshan Petrochemical, Sinopec Group, Beijing 102400, China;

² China Ship Development and Design Center, Wuhan 430064, China; acute.xjtu@foxmail.com

* Correspondence: jiaoxiner@163.com

Abstract: Open bidding is widely used in Sinopec's material procurement due to its advantages of openness, transparency, and cost savings. However, there have been many complaints about collusive bidding in recent years. Collusive bidding will pose significant challenges to fair competition and increase project costs for material procurement. Therefore, there is an urgent need for an analysis model to identify and prevent the collusive bidding. In this paper, we propose a statistical model for identifying collusive bidding using the coefficient of variation c_v , the ratio of high to low average quotation r_p and the deviation rate of quotation δ as indicators, based on the statistics of the historical quotations for material A. Through a case study of the framework agreement bidding for material B, we identified the collusion behavior of bidders effectively, thereby verifying the applicability of this statistical model.

Keywords: collusive bidding; historical quotations; coefficient of variation; deviation rate; statistical model

1. Introduction

The open bidding system can not only standardize the market order, improve the efficiency of procurement fund utilization, and reasonably allocate resources, but also enhance the competitive and innovative awareness of social enterprises. This will be conducive to select supplier enterprises that are scientific and international for bidding.

Competitive bidding is one of the most common methods for tenders to select suppliers [1,2]. Open bidding is an important measure to standardize procurement behavior and improve the efficiency of capital use [3]. Since 2015, open bidding has made great progress and achieved an outstanding effect in material supply of Sinopec. But there have been a lot of complaints about collusive bidding in recent years. Among Sinopec's complaint cases in the past three years, 48.6% were complaints about collusive bidding.

Collusive bidding will pose a serious challenge to the fair competition with suppliers [4,5]. Contracts awarded on the basis of collusive bidding by bidders may result in increased costs for material procurement projects [6,7]. The essence of collusive bidding is that bidders quote higher or lower than the reasonable price of the project [8,9], aiming to gain an advantage over other competitors and achieve higher profits [10,11]. As far as the tender is concerned, the collusive bidding will damage the interests of the tender [12]. Through the formal bidding process, the tender could have obtained high-quality services or products at a lower purchase price, but the collusive bidding behavior makes the high-quality bidders unable to win the bid. The result is that the tender has to pay a higher price for a relatively inferior service or product [13]. For the society, collusive bidding not only disrupts the fair competition in the market, but also interferes with the effective allocation of social and economic resources, thus reducing the economic benefits and production efficiency of the whole society.

The phenomenon of collusion bidding has existed for a considerable period of time, but its timely detection and identification is a pervasive challenge [14,15]. Therefore, a scientific and effective analysis model is needed to identify the collusive bidding behavior of bidders. Taking the framework agreement bidding of material A of Sinopec as an example, this paper establishes a statistical model

based on historical quotations, which can be widely used in the bidding and procurement of the framework agreement of materials to prevent the collusive bidding behavior of bidders.

2. Materials and Methods

2.1. The procurement process of material A

At present, the number of professional manufacturers producing the material A is nearly 100, and there are differences in product quality and price among various manufacturers. In general, suppliers can be divided into three levels. The first level of manufacturers is less than 10%, and its production scale is large, technology research and product series have reached the international advanced level. The second level of manufacturers accounted for about 30%, with considerable production capacity and research and development capabilities, but in the product series and qualification certification is slightly inferior. Other manufacturers are the third level, and their product series have a higher degree of general research and development capabilities, and relatively low product prices.

Due to the large demand for material A in petrochemical enterprises, the procurement methods are mainly based on framework agreements. First of all, through the centralized open bidding of the headquarters to determine the shortlisted suppliers, and sign a unified framework agreement with them. Subsequently, each enterprise signed specific orders with suppliers according to the framework agreement of the headquarters.

In the bidding process, business rating is usually calculated by the benchmark price method. Among them, the arithmetic average price of each bidder is taken as the price benchmark, and the business score of each bidder is calculated according to the difference between the price and the benchmark price [16]. This method can avoid the risk of winning bids at low prices and defective products, and is the most commonly used business score calculation method for large-scale or important material procurement [17]. However, the large interest drives some bidders to form a community to raise the benchmark price and win the bid at a high price for illegal profit, thus increasing the risk of collusion bidding.

2.2. Establishment of a statistical model for identifying collusive bidding

In order to prevent the occurrence of collusion bidding by bidders, this paper aims to determine the price rule based on the statistical analysis of the historical quotations for material A, and establish a statistical model that can identify the collusion bidding, to be applicable to the bidding and procurement under various material frameworks agreements.

2.2.1. Analysis of historical quotations for material A

The historical quotations database is the most commonly used tool for purchasing personnel [18]. Using the principle of statistics to analyze the historical quotations, we can determine the market price rule of this kind of material. The historical quotations of 10 commonly used types of material A are shown in Table 1 and Figure 1.

Table 1. Historical quotations for Material A.

Material A	Historical quotations									
Type I	512	495	440	480	470	507	420	488	398	485
Type II	526	518	470	515	500	522	420	522	439	498
Type III	524	523	478	515	500	522	420	525	448	498
Type IV	640	638	568	660	620	631	530	645	530	625
Type V	602	590	546	615	530	629	500	580	506	585
Type VI	770	770	840	805	820	770	690	790	665	775
Type VII	801	788	800	780	832	812	720	808	730	785
Type VIII	792	780	875	815	830	791	700	805	698	785

Type IX	920	950	940	980	960	925	785	960	785	940
Type X	875	813	900	850	830	868	745	810	758	820

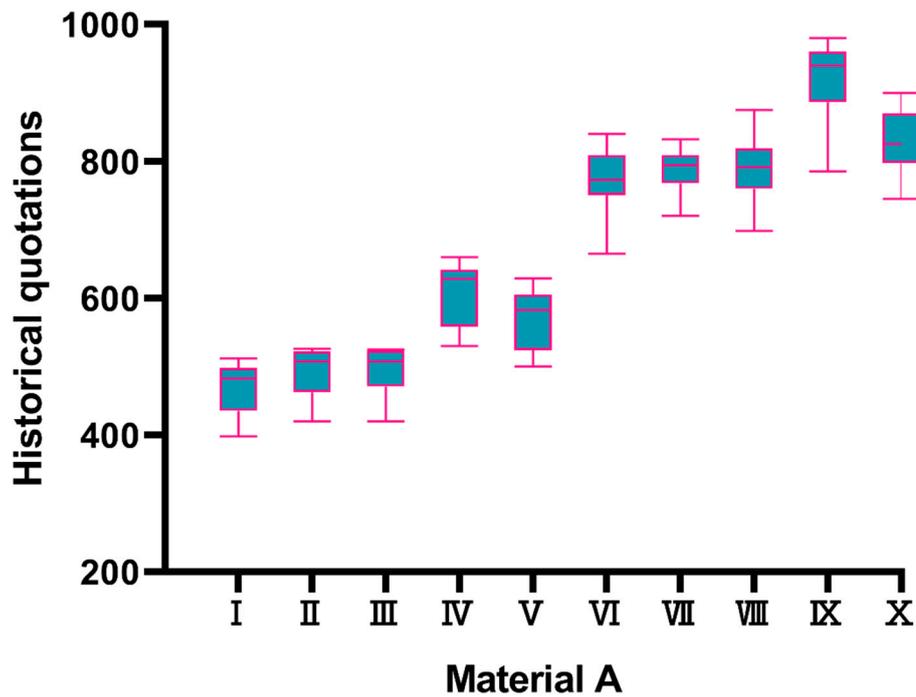


Figure 1. Distribution of historical quotations for Material A.

2.2.2. Statistics of historical quotations for material A

The average quotations \bar{x} , standard deviation S , coefficient of variation c_v and deviation rate δ between the maximum and minimum value of the material A are calculated by historical quotations, where

$$\delta = \frac{P_{\max} - P_{\min}}{P_{\min}} \quad (1)$$

The results are shown in Table 2. Excluding the factors of inflation and rising cost, the price rule of material A is found through these statistics as follows:

Table 2. Statistics of historical quotations for Material A.

Material A	Mean (\bar{x})	SD (S)	CV (c_v)	DR (δ)
Type I	469.50	36.04	8%	29%
Type II	493.00	35.68	7%	25%
Type III	495.30	34.37	7%	25%
Type IV	608.70	45.52	7%	25%
Type V	568.30	42.89	8%	26%
Type VI	769.50	51.45	7%	26%
Type VII	785.60	33.51	4%	16%
Type VIII	787.10	51.27	7%	25%
Type IX	914.50	66.84	7%	25%
Type X	826.90	46.77	6%	21%

Abbreviations: SD, standard deviation; CV, coefficient of variation; DR, deviation rate.

a) The historical quotations are relatively stable, and the coefficient of variation c_v is no more than 10%;

b) The difference between bid quotations is small, and the deviation ratio δ between the maximum and the minimum value does not exceed 30%.

2.2.3. Estimate of the market price for material A

Then the statistics of historical quotations for the material A are used as the sample to estimate the range of the market price of this set of materials.

The market price distribution of the material A conforms to the normal distribution,

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \sim N(0,1) \quad (2)$$

Where, the equation for calculating the average quotations \bar{x} and standard deviation S based on historical quotations for material A is

$$\bar{x} = \frac{\sum x}{n} \quad (3)$$

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} \quad (4)$$

According to Equation 2, the population mean of the market price of material A in the confidence interval $1 - \alpha$ is

$$\mu = \bar{x} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}} \quad (5)$$

Considering the increase in raw material and labor costs during the framework agreement period, the market price is multiplied by the increase correction coefficient K , and the K of the material A is taken as 1.05 per year. The average market price of the material A with 95% confidence interval can be calculated, as shown in Table 3.

Table 3. Limit of market price for Material A.

Material A	Upper limit	Lower limit
Type I	491.02	456.48
Type II	510.29	477.55
Type III	511.53	479.97
Type IV	632.85	588.31
Type V	594.29	552.87
Type VI	796.63	749.20
Type VII	803.93	772.90
Type VIII	812.20	765.46
Type IX	953.19	890.14
Type X	854.63	806.87

2.2.4. Evaluation indicators of statistical model for identifying collusive bidding

(A) *Quotation coefficient of variation cv*: The coefficient of variation is a characteristic quantity representing the degree of data dispersion, that is, the ratio of the standard deviation of the data to its corresponding average, which is mainly used to compare the statistical dispersion of different samples [19]. Here, the quotation coefficient of variation is used to test the deviation level of each supplier. Due to the similar raw materials and no significant differences in production processes for the material A, there is not a significant price difference. Based on the analysis of historical quotations as a sample in the previous section, it can be seen that the coefficient of variation of the material A is within 10%, and the maximum value of other categories of materials does not exceed 15%. Therefore, we presumed that a quotation coefficient of variation cv exceeding 15% is considered a possibility of collusive bidding.

(B) *Ratio of high to low average quotation r_P* : Collusive bidding will increase the benchmark price, resulting in the supplier winning the bid at a high price. How many collusion bidders will result in the benchmark price being raised?

Based on the deviation rate between the maximum and minimum historical quotations mentioned above, it was found that collusion suppliers would increase their prices by at least 20% of profit, resulting in a deviation in the benchmark price. Excluding the annual increase in labor costs, we presume that there is collusion behavior when the benchmark price is raised by 7%.

The calculation equation for the number of collusion bidders is as follows:

$$(N - n)P + 1.2nP = 1.07NP \quad (6)$$

where N represents the total number of bidding suppliers, n represents collusive bidding suppliers, and P represents the average quotation of materials.

The result calculated by Equation 6 is

$$n = 0.35N$$

We found that when about 1/3 of suppliers engage in collusion bidding, it will cause a deviation in the benchmark price.

Therefore, after ranking the quotations of bidders from high to low, we set the ratio of high to low average prices as

$$r_P = \frac{P_{1/3}}{P_{2/3}} \quad (7)$$

where $P_{1/3}$ is the average high price quoted by the top 1/3 bidders, and $P_{2/3}$ is the average low price quoted by the bottom 2/3 bidders.

When the r_P exceeds 20%, it indicates a deviation in the benchmark price quoted by the bidder, and we speculate that there is a possibility of collusion bidding.

(C) *Deviation rate of quotation δ* : One of the important characteristics of collusive bidding is that the bidding quotations are similar. According to the analysis results above, 1/3 of the bidders colluding may lead to an increase in the benchmark price.

The deviation rate that characterizes the proximity of higher bidders' quotations is:

$$\delta = \frac{P_{\max} - P_{1/3}}{P_{1/3}} \quad (8)$$

where P_{\max} is the highest price, and $P_{1/3}$ is the average high price quoted by the top 1/3 bidders.

When δ is less than 5%, it indicates that bidders with high quotations have similar prices and there is a possibility of collusion bidding.

2.2.5. Analysis results of statistical model for identifying collusive bidding

By combining the three important parameters of c_v , r_P and δ , the statistical model can be established to identify collusive bidding. The analysis results are shown in Table 4.

Table 4. List of bid analysis based on c_v , r_P and δ .

CV (c_v)	RP (r_P)	DR (δ)	Result
$c_v \cong 15\%$			Collusion
$10\% \cong c_v < 15\%$	$r_P \cong 20\%$	$\delta < 5\%$	Collusion
	$r_P \cong 20\%$	$\delta \cong 5\%$	Suspicious
	$r_P < 20\%$		Cost difference
$c_v < 10\%$	$r_P \cong 20\%$	$\delta \cong 5\%$	Suspicious
Other cases			No collusion

Case 1: When $c_v \geq 15\%$, the coefficient of variation has exceeded the statistical dispersion of the market price. The benchmark price can no longer represent the market price and it is considered a collusive bidding.

Case 2: When $10\% \leq c_v < 15\%$, the dispersion of bidding quotations is relatively high, which may be due to collusive bidding or cost differences caused by factors such as materials and raw materials. This case requires further analysis of other indicators.

When $r_p \geq 20\%$ and $\delta < 5\%$, it indicates that the quotations are similar and can be considered as collusive bidding. When $r_p \geq 20\%$ and $\delta \geq 5\%$, it indicates that the benchmark price quoted by the bidder is relatively high and can be considered as suspicious.

When $r_p < 20\%$, it indicates that there is no deviation in the benchmark price although the quotations are discrete and may be caused by cost differences.

Case 3: When $c_v < 10\%$, the bidding quotations are stable and concentrated, indicating a low possibility of collusion bidding.

However, when $r_p \geq 20\%$ and $\delta \geq 5\%$, there is a possibility that the statistical dispersion does not deviate significantly due to a small number of collusive suppliers, which requires special attention.

Other cases: Except for the three cases mentioned above, the bidding quotations conform to the market price rule and there is no collusion.

2.3. Statistical analyses

All data were analyzed by GraphPad Prism 8.0 software (GraphPad Software Inc., USA).

3. Results and discussion

3.1. Statistics of quotations for material B

It is necessary to verify the validity of the statistical model for identifying collusive bidding. We take the framework agreement bidding for material B as a case study, and the quotations of 12 suppliers are shown in Table 5.

Table 5. Bid quotations for Material B.

Material B	Bid quotations											
Type I	600	525	509	466	590	509	498	585	445	580	422	514
Type II	580	520	515	490	530	520	590	610	430	620	460	528
Type III	590	520	515	490	580	520	560	620	430	625	460	528
Type IV	760	650	615	602	640	635	740	755	562	750	562	663
Type V	750	650	615	602	640	635	755	740	562	760	562	663
Type VI	900	800	825	810	848	820	880	890	700	910	700	810
Type VII	900	750	720	790	770	800	900	880	610	900	700	810
Type VIII	950	825	830	840	830	750	950	970	636	960	700	820
Type IX	960	890	850	845	950	990	850	890	730	965	830	880
Type X	990	862	848	954	954	901	950	970	684	970	803	869

According to Figure 2 and Table 5, the characteristics of the quotations can be analyzed as follows:

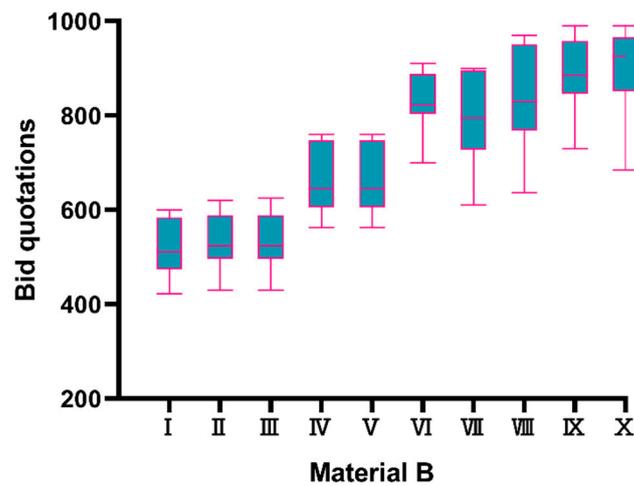


Figure 2. Distribution of bid quotations for Material B.

a) The standard deviation of bid quotations was large, and $10\% \leq c_v < 15\%$ indicates that these quotations are discrete.

b) There was a price gap in bid quotations. For example, bidders 1, 5, 8, and 10 quoted 40% higher for type VIII than other suppliers, and $r_P \geq 20\%$.

c) Some bidders have quoted high and similar prices. For example, the prices quoted by bidders 1, 5, 8, and 10 are significantly higher than the average price, and $\delta < 5\%$.

3.2. Analysis results of statistical model in material B

The statistical model for identifying collusive bidding was used to analyze the quotations of material B. The results in Table 6 indicate that there is a possibility of collusion bidding for types I, V, VII, VIII, and IX of material B. Further analysis revealed that the abnormal bidding quotations were from bidders 1, 5, 8, and 10. After investigation and evidence collection, it was found that there were indeed four suppliers colluding in this bidding. Therefore, the statistical model for identifying collusive bidding has been verified in practice.

Table 6. Analysis results of bid quotations for Material B based on statistical model.

Material B	CV (c_v)	RP (r_P)	DR (δ)	Result
Type I	11%	21.14%	3.45%	Collusion
Type II	11%	20.21%	6.90%	Suspicious
Type III	11%	20.97%	7.76%	Suspicious
Type IV	11%	21.96%	2.70%	Suspicious
Type V	11%	21.96%	2.70%	Collusion
Type VI	8%	13.42%	3.41%	No collusion
Type VII	12%	20.34%	2.27%	Collusion
Type VIII	13%	22.93%	2.11%	Collusion
Type IX	24%	54.49%	15.38%	Collusion
Type X	10%	13.05%	3.77%	No collusion

4. Conclusions

Based on the statistical analysis of the historical quotations for the material A, we proposed a statistical model for identifying collusive bidding using the coefficient of variation c_v , the ratio of high to low average quotation r_P and deviation rate of quotation δ as indicators. The validity and

applicability of this statistical model were further verified through a case study of the framework agreement bidding for material B.

This statistical model could be used in multiple large-scale framework agreement bidding by Sinopec in the future. This will help the tender quickly and effectively identify the possibility of collusive bidding without conducting market analysis or investigating historical prices, thereby improving the quality of bidding. Furthermore, this is also beneficial for preventing the collusive bidding.

Author Contributions: Conceptualization, Jieni Li and Hongchao Guo; Data curation, Fei Xie; Formal analysis, Jieni Li; Investigation, Jieni Li; Methodology, Jieni Li, Fei Xie and Hongchao Guo; Software, Jieni Li, Fei Xie and Hongchao Guo; Supervision, Hongchao Guo; Validation, Fei Xie and Hongchao Guo; Writing – original draft, Jieni Li; Writing – review & editing, Fei Xie and Hongchao Guo.

Funding: This research was funded by the Technology Transformation and Development Project of Sinopec Yanshan Petrochemical (grant nos. YSAC190181).

Data Availability Statement: All other data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that there are no conflicts of interest.

References

1. Takano, Y., Ishii, N., Muraki, M. Determining bid markup and resources allocated to cost estimation in competitive bidding. *Automation in Construction* **2018** *85*, 358-368. doi:10.1016/j.autcon.2017.06.007.
2. Cattell, D. W., Bowen, P. A., Kaka, A. P. Review of unbalanced bidding models in construction. *Journal of Construction Engineering and Management-Asce* **2007** *133*, 562-57. doi:10.1061/(asce)0733-9364(2007)133:8(562).
3. Li, H., Su, L., Zuo, J., An, X., Dong, G., Wang, L., Zhang, C. The framework of data-driven and multi-criteria decision-making for detecting unbalanced bidding. *Engineering, Construction and Architectural Management* **2021** *30*, 598-622. doi:10.1108/ecam-08-2020-0603.
4. Hyari, K. H. A Procedure for Rebalancing Unbalanced Bidding in Unit Price Contracts. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction* **2023** *15*, doi:10.1061/jladah.Ladr-939.
5. Choi, S., Choi, S., Kim, J., Lee, E.-B. AI and Text-Mining Applications for Analyzing Contractor's Risk in Invitation to Bid (ITB) and Contracts for Engineering Procurement and Construction (EPC) Projects. *Energies* **2021** *14*, doi:10.3390/en14154632.
6. Alhyari, O., Hyari, K. H. Handling Unbalanced Pricing in Bidding Regulations for Public Construction Projects. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction* **2022** *14*, doi:10.1061/(asce)la.1943-4170.0000547.
7. Hu, H., Deng, X., Mahmoudi, A. A cognitive model for understanding fraudulent behavior in construction industry. *Engineering, Construction and Architectural Management* **2021** *30*, 1423-1443, doi:10.1108/ecam-08-2021-0703.
8. Lee, J.-S. Simulating Competitive Bidding in Construction Collusive Bidding Cases. *Journal of Management in Engineering* **2022** *38*, doi:10.1061/(asce)me.1943-5479.0001081.
9. Alhyari, O., Hyari, K. H. An International Overview of Unbalanced Pricing under Bidding Regulations for Public Construction Projects. *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction* **2023** *15*, doi:10.1061/jladah.Ladr-852.
10. An, X., Li, H., Zuo, J., Ojuri, O., Wang, Z., Ding, J. Identification and Prevention of Unbalanced Bids Using the Unascertained Model. *Journal of Construction Engineering and Management* **2018** *144*, doi:10.1061/(asce)co.1943-7862.0001563.
11. Padhi, S. S., Mohapatra, P. K. J. Detection of collusion in government procurement auctions. *Journal of Purchasing and Supply Management* **2011** *17*, 207-221, doi:10.1016/j.pursup.2011.03.001.
12. Kwas, M., Rubaszek, M. Forecasting Commodity Prices: Looking for a Benchmark. *Forecasting* **2021** *3*, 447-459, doi:10.3390/forecast3020027.
13. Bergeron, C. Benchmark, relative return, and asset pricing. *Applied Economics Letters* **2021** *29*, 1498-1503, doi:10.1080/13504851.2021.1940080.

14. Abrantes-Metz, RM ., Bajari, P. Screens for conspiracies and their multiple applications. *Competition Policy Int* **2010** *6*, 129–144.
15. Signor, R; Love, PED; Oliveira, PS. Public Infrastructure Procurement: Detecting Collusion in Capped First-Priced Auctions. *Journal of Infrastructure Systems* **2020** *26*, 1-12, doi:10.1061/(ASCE)IS.1943-555X.0000543
16. Lee, J.-S. Simulating Competitive Bidding in Construction Collusive Bidding Cases. *Journal of Management in Engineering* **2022** *38*, doi:10.1061/(asce)me.1943-5479.0001081.
17. Mehrbod, A., Zutshi, A., Grilo, A.Jardim-Gonsalves, R. Application of a semantic product matching mechanism in open tendering e-marketplaces. *Journal of Public Procurement* **2018** *18*, 14-30, doi:10.1108/jopp-03-2018-002.
18. Gunadi, M. P.Evangelidis, I. The Impact of Historical Price Information on Purchase Deferral. *Journal of Marketing Research* **2022** *59*, 623-640, doi:10.1177/00222437211060359.
19. Kvalseth, T. O. Coefficient of variation: the second-order alternative. *Journal of Applied Statistics* **2016** *44*, 402-415, doi:10.1080/02664763.2016.1174195

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.