SA-TIED

Southern Africa – Towards Inclusive Economic Development

SA-TIED Technical Note 5

# COVID-19 emergency procurement trends, deviations, expansions, and excessive pricing

Daniel Page, David McClelland, Adeola Oyenubi, Grace Bridgman, and Uma Kollamparambil

December 2021

UNITED NATIONS UNIVERSITY
UNU-WIDER

national treasury
Department: National Treasury
REPUBLIC OF SOUTH AFRICA

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
IFPRI

planning, monitoring & evaluation
Department: Planning, Monitoring and Evaluation
REPUBLIC OF SOUTH AFRICA

the dtic
Department: Trade, Industry and Competition
REPUBLIC OF SOUTH AFRICA

SARS

TIPS

## About the project

### Southern Africa –Towards Inclusive Economic Development (SA-TIED)

SA-TIED is a unique collaboration between local and international research institutes and the government of South Africa. Its primary goal is to improve the interface between research and policy by producing cutting-edge research for inclusive growth and economic transformation in the southern African region. It is hoped that the SA-TIED programme will lead to greater institutional and individual capacities, improve database management and data analysis, and provide research outputs that assist in the formulation of evidence-based economic policy.

The collaboration is between the United Nations University World Institute for Development Economics Research (UNU-WIDER), the National Treasury of South Africa, the International Food Policy Research Institute (IFPRI), the Department of Monitoring, Planning, and Evaluation, the Department of Trade and Industry, South African Revenue Services, Trade and Industrial Policy Strategies, and other universities and institutes. It is funded by the National Treasury of South Africa, the Department of Trade and Industry of South Africa, the Delegation of the European Union to South Africa, IFPRI, and UNU-WIDER through the Institute's contributions from Finland, Sweden, and the United Kingdom to its research programme.



Corresponding author: Uma.Kollamparambil@wits.ac.za

# COVID-19 emergency procurement trends, deviations, expansions, and excessive pricing

Daniel Page,[1] David McClelland,[1] Adeola Oyenubi,[1]
Grace Bridgman,[2] and Uma Kollamparambil[1]

December 2021

**Abstract:** The arrival of the COVID-19 pandemic and the response by the South African government was consistent with global trends and prompted unprecedented public procurement. Government responses included the application of a staged lockdown as well as large-scale mobilization of procurement measures for emergency and medical services, business support as well as other social incentives. Given the necessary swift action by provincial and national government, an unfortunate consequence of the pandemic-related procurement was the heightened possibility of illicit and corrupt procurement activities. To date, there have been numerous instances of corrupt COVID-19-related procurement activities being exposed in local media, however the extent of COVID-19-specific corruption has yet to be identified or quantified. The purpose of this study is to analyse procurement data pre-COVID-19 against the available COVID-19 emergency procurement data to identify possible instances of corruption.

**Key words:** COVID-19, pandemic, government procurement, South Africa

**JEL classification:** C53, C55, C61, G28

[1] School of Economics and Finance, University of the Witwatersrand, Johannesburg, South Africa, corresponding author: Uma.Kollamparambil@wits.ac.za; [2] UNU-WIDER (at the time of writing)

# 1 Introduction and framework

The arrival of the COVID-19 pandemic and the response by the South African government was consistent with global trends and prompted unprecedented public procurement. Government responses included the application of a staged lockdown as well as large-scale mobilization of procurement measures for emergency and medical services, business support as well as other social incentives. Given the necessary swift action by provincial and national government, an unfortunate consequence of the pandemic-related procurement was the heightened possibility of illicit and corrupt procurement activities. To date, there have been numerous instances of corrupt COVID-19-related procurement activities being exposed in local media, however, the extent of COVID-19-specific corruption has yet to be identified or quantified. The purpose of this study is to analyse procurement data pre-COVID-19 against the available COVID-19 emergency procurement data to identify possible instances of corruption.

This note begins by analysing a pilot dataset on the procurement of goods and services for 2017–20, made available by the Office of the Chief Procurement Officer, National Treasury of South Africa, in conjunction with UNU-WIDER. Initial analysis is conducted using Benford's law[1] to determine whether COVID-19-specific procurement data, specifically *spend amount* displays distributional idiosyncrasies which may indicate the presence of corrupt procurement activity. The note proceeds to develop and discuss a multivariate, multi-test empirical framework that relies on an ensemble of tests for the statistical identification of probable cases of illicit spending practices.

# 2 Data and assumptions

The dataset applied in the analysis was the combined join of the LOGIS, BAS, and CSD datasets and contained a cross-section of 2.942 million payments from Q2/2017 until Q4/2020. For the analysis, the subset of key variables extracted is described below with the variable type in parenthesis.

- *document date* (Date)
- *government entity* (Categorical)
- *industry classification code* (Categorical)
- *payment unit price* (Rand value)
- *order quoted price* (Rand value)
- *payment quantity* (Unit)
- *icn description* (Categorical)
- *unit of issue* (Unit)
- *fund description* (Categorical)
- *turnover* (Categorical)
- *Black Economic Empowerment ('BEE') level* (Categorical)
- *women-owned* (Categorical)

[1] Benford's Law (Benford 1939) is applied in determining the normal level of number duplicates in a dataset which allows for the identification of outliers.

- *spend amount* (Rand value)
- *spend per quotation amount* (Rand value)
- *spend per contract amount* (Rand value)
- *disbursed amount* (Rand value)
- *awaiting disbursed amount* (Rand value)
- *payment unit price* (Rand value)

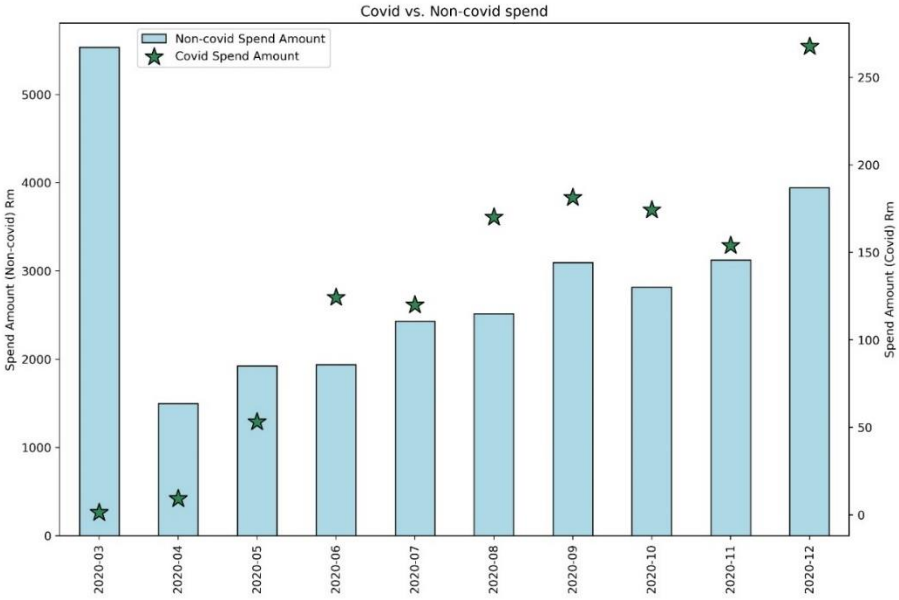
To ensure consistency of analysis on a like-for-like basis through time, all Rand value variables were deflated using a CPI adjustment based to 2016. *Document_date* was applied as the date key as it was found to be the most consistent transaction and contract date variable. In terms of the categorical variables considered, *government_entity* represents the government procuring party for each contract, *industry_classification_code* is the industry short-hand code of the supplier or service[2] and *icn_description* represents the type of good or service contracted. Lastly, *fund_description* was central to the analysis as it describes the government source of funds utilized for the respective procurement spend. *Fund_description* was therefore used to define a COVID-19 binary variable and was labelled true in instances where the fund description contained the word COVID.

## 3 Preliminary analysis of COVID-19 spending

The section that follows provides a summarized overview of COVID-19-related spending over the latter 10 months of 2020. As mentioned, all analysis was conducted on the joined LOGIS, BAS and CSD database where COVID-related procurement was isolated via the *fund_description* which typically specified the purpose of the procurement spend. Any label that included the term 'COVID' was assigned a dummy variable of one, allowing for the clear separation of general procurement over the period from COVID-specific procurement. Lastly, to minimize any impact of inflation against prior years of the database, all monetary variables were deflated to 2016 values.

[2] See Appendix A2 for *industry_classification_code* and *industry_classification_name*

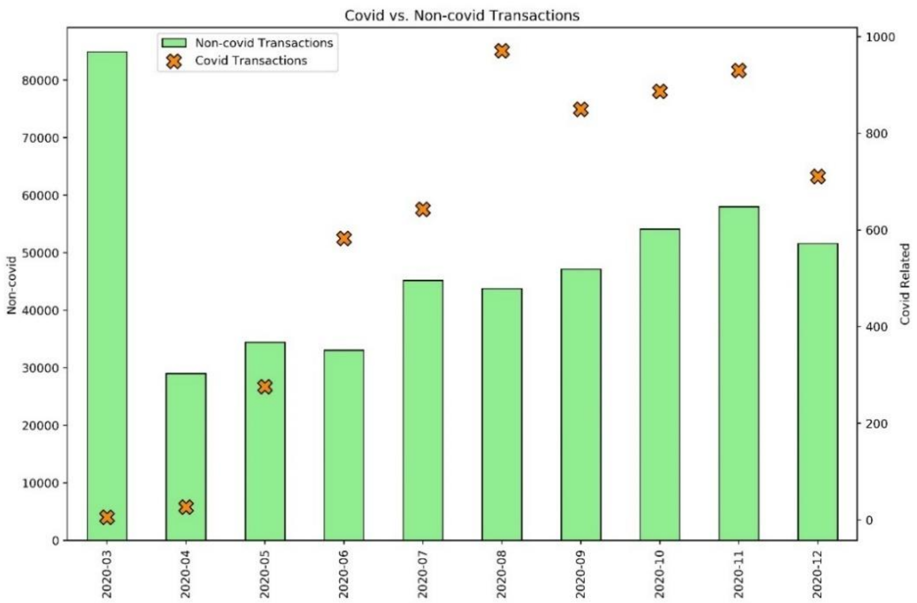
Figure 1: COVID vs. non-COVID expenditure




Source: authors' calculations based on data presented in Section 2.

Figure 1 describes total non-COVID and COVID-related procurement per month (based on real *spend_amount*) over the period of investigation. The results show that COVID expenditure initially diverged from non-COVID spending but accelerated and converged on towards the end of 2020. The analysis further indicated that non-COVID-related expenditure (in 2016 terms) amounted to R28.82bn compared to the COVID spend of ZAR1.26bn.[3]

Figure 2: COVID vs. non-COVID transactions



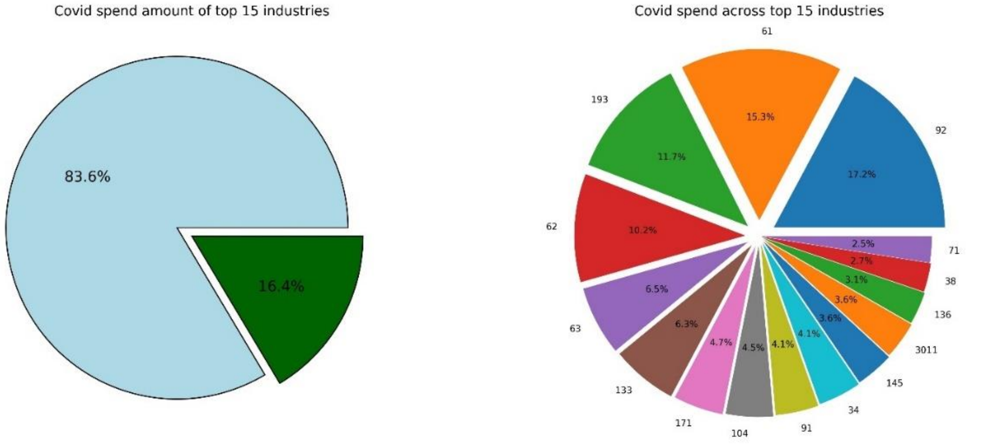

Source: authors' calculations based on data presented in Section 2.

Figure 2 describes the number of COVID versus non-COVID procurement transactions per the joined database of the period of study. The results show higher levels of divergence specifically

[3] In the rest of the note, the currency code 'ZAR' is shortened to 'R'.

around August and December 2020, implying that the actual expended amount per contract increased dramatically in December. Numerically, the total number of COVID specific transaction recorded over the sample period totalled 5,884 compared to 481,385 non-COVID-related transactions.

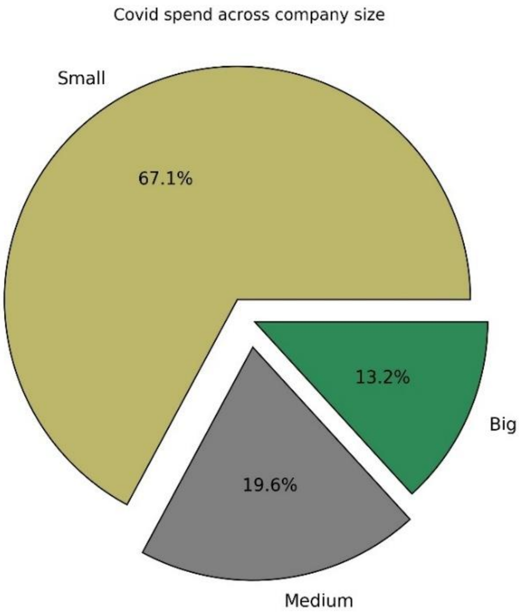
Figure 3: COVID spend across supplied goods/services industry




Source: authors' calculations based on data presented in Section 2.

Focusing on COVID spending, Figure 3 describes COVID-19 procurement split according to industry. For brevity, numerical codes per the combined database have been used and are described in Appendix A2. Out of 106 possible industries, the top 15 supplier-/contract-related industries covered 83.6 per cent of all COVID-related expenditure in 2020 while the five largest industries per the right-hand figure were food and beverage service activities (92–17.2 per cent), construction of buildings (61–15.3 per cent), other personal service activities (193–11.7 per cent), civil engineering (62–10.2 per cent) and specialized construction activities (62–10.2 per cent).

Figure 4: COVID spend across supplier size

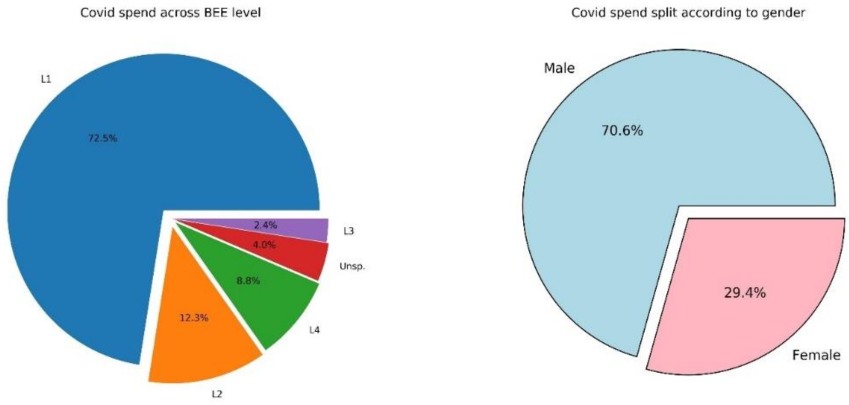



Source: authors' calculations based on data presented in Section 2.

Figure 4 describes the split of COVID-related expenditure across supplier size, proxied by sales turnover where small entails turnover less than R10m per annum, medium between R10m and

R50m and big being suppliers with turnover greater than R50m. The analysis indicates that the 67 per cent of suppliers responsible for COVID-related goods and services were small while only 13 per cent of services and goods were supplied by large ('big') suppliers.

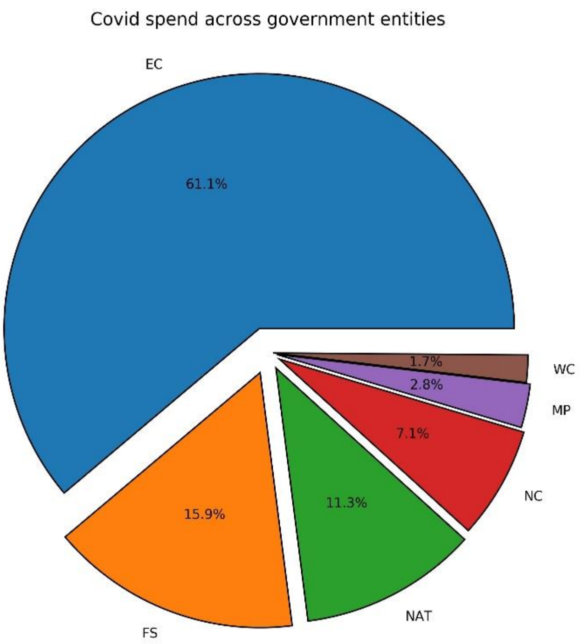
Figure 5: COVID spend across BEE level and gender




Source: authors' calculations based on data presented in Section 2.

Figure 5 describes COVID spend across BEE level and clearly indicates that most goods and services (72.5 per cent) were sourced through level 1 suppliers. Suppliers with levels 2, 3, and 4 provided 12.3, 2.4, and 8.8 per cent respectively. Considering the gender split based on supplier shareholding and management, the 70.6 per cent of COVID-specific suppliers were male/predominantly male with the remainder being female.

Figure 6: COVID spend per government entity




Source: authors' calculations based on data presented in Section 2.

Figure 6 describes the split of COVID spending across government entity where government entity represents the procuring arm of government (provincial or national). The figure shows that the Eastern Cape was the biggest procurer of COVID services, being responsible for 61.1 per cent (R767.7m in 2016 terms) of the total spend. The next biggest procuring government entity was the

Free State, constituting 15.9 per cent of the overall COVID spend (R199.2m) while national government was responsible for 11.3 per cent of the total spend, equating to R141.4m of the total COVID spend over the period.

## 4 A framework for identifying potential illicit spending

In order to identify potential corruption and illicit spending practices associated with COVID-19, an empirical testing framework was developed that considers three core components:

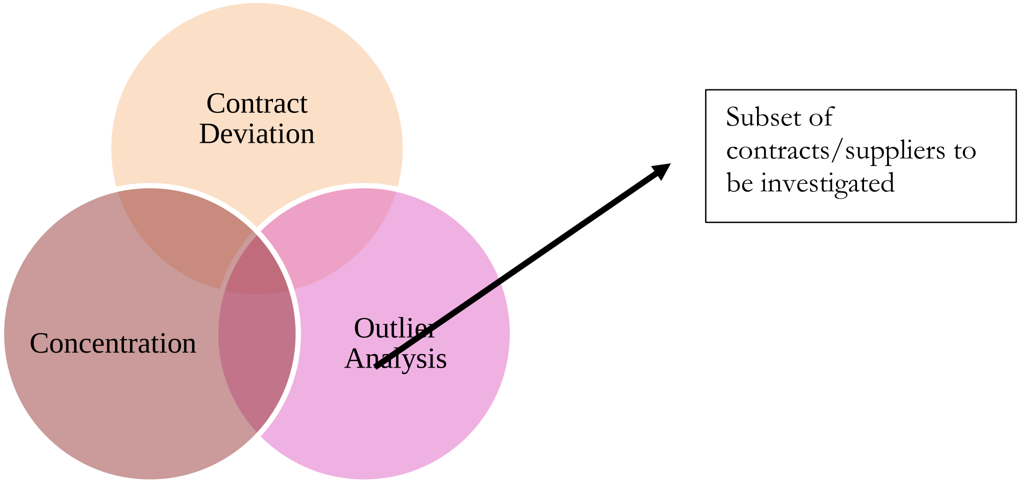
- **Contract deviation analysis**: Evaluation of differences between actual spending versus contracted procurement spend across COVID-19 and non-COVID-19-related expenditure. The focus of the analysis therefore attempts to identify instances of contract spend deviations specific to COVID-19 procurement and provide a direct comparison to deviation levels across non-COVID-19 procurement.

- **Supplier concentration analysis:** Estimation of the Herfindahl-Hirschman Index (HHI) provides insight to the extent of supplier concentration. Increases in concentration can be considered evidence of unfair or corrupt procurement activity as procurement within a pre-defined 'economy' is not spread equitably across suppliers. The higher the HHI index, the closer an 'economy' is to a monopoly and the higher the possible propensity for illicit or corrupt procurement activity. Importantly, for the purpose of this study, the change in HHI (as opposed to static HHI) between non-COVID-19 and COVID-19-specific expenditure is explored.

- **Price and contract deviation outlier analysis:** The analysis and detection of outliers is commonplace across data science and econometric applications, as outliers can result in incorrect inferences or relationships being identified. For the purpose of this study, the identification and classification of outliers provides an additional layer of evaluation in attempt to detect corrupt and illicit procurement instances. To detect possible procurement datapoint outliers, the Mahalanobis distance (MD) developed by Mahalanobis (1936) is applied to determine whether the Euclidian distance between two points is significant in a multivariate space. The basis of the MD measure is to determine whether a dependent variable can be considered an outlier based on the distributional properties of its independent explanatory variables. The test allows for the evaluation COVID-19 procurement in conjunction with like-for-like pre-COVID-19 procurement data to determine whether the former can be considered an outlier based on the distribution of the latter. If COVID-19-specific prices or contract payment differences is found to be a statistically significant outlier, the result may point to illicit or corrupt procurement practices.

A benefit of the described methods is that each provides unique information regarding the detection of illicit procurement activities. A further benefit of the independence of each test, beyond the informational analysis, is the ability to create a combined or ensemble test for corrupt/illicit procurement identification.

Figure 7: Ensemble/combined test Venn diagram



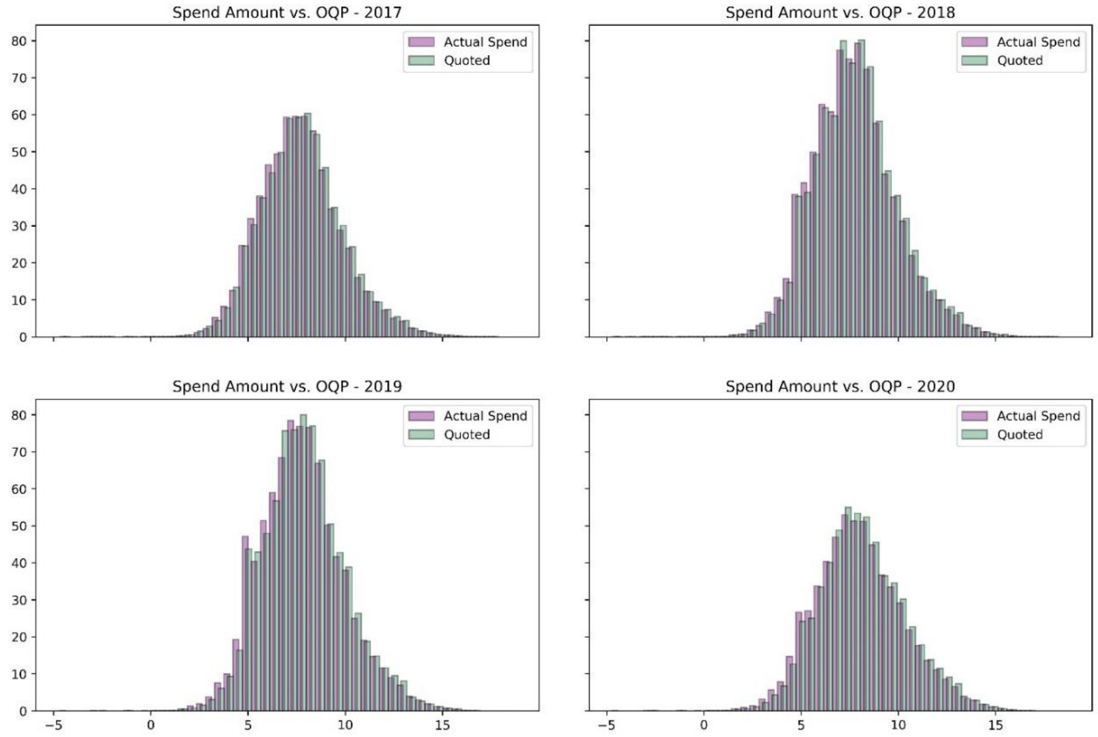

Source: authors' illustration.

Figure 7 describes a Venn diagram that considers the outcome of each test run on the combined LOGIS, BAS and CSD dataset. The outcome of each test can be converted to a binary outcome labelled 'high concern' and 'low concern' across the procurement data. The combination of all three tests allows for the stratification of the procurement data in subsets, where the highest concern stratum considers contract/supplier instances that achieved three 'flags', represented by the union of the three test circles above. Unfortunately, without *a priori* knowledge of specific contracts that have subsequently found to be corrupt/illicit, determination of accuracy is untestable but is an avenue of further research.

## 5 Contract deviation analysis

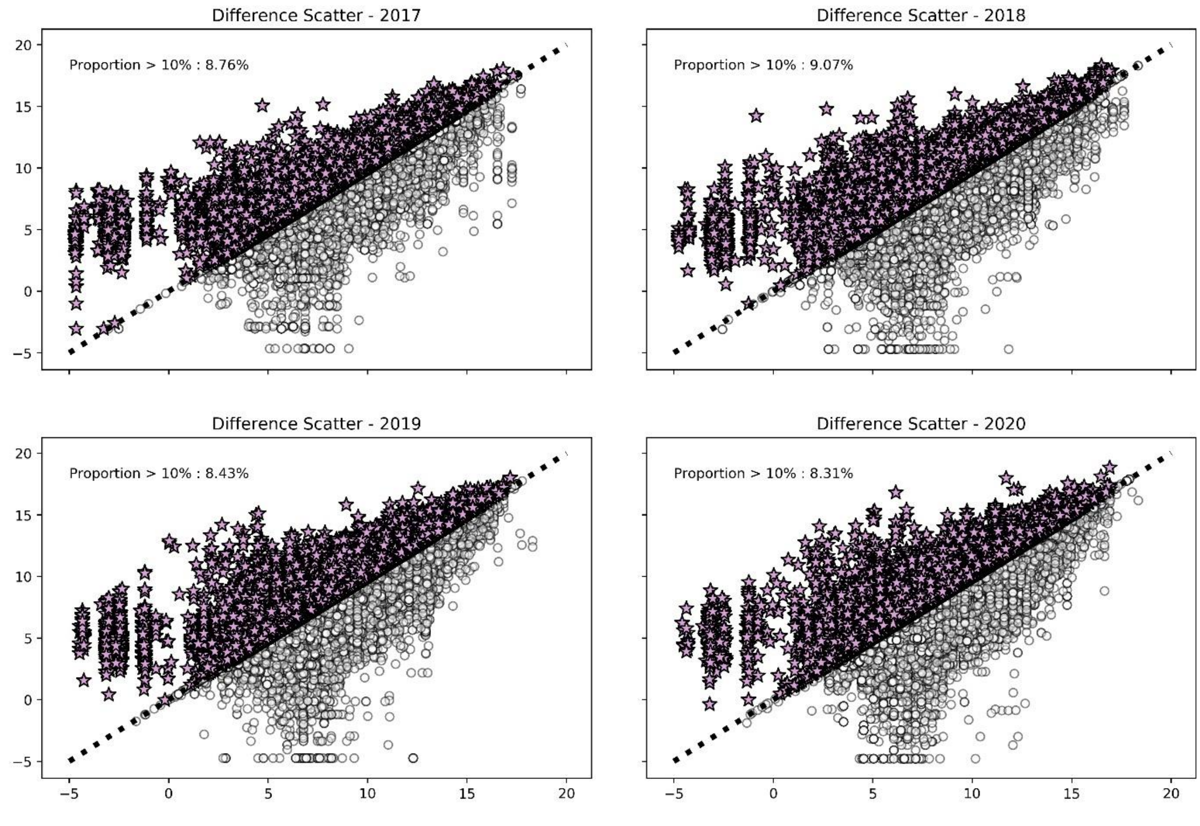
Contract deviation analysis focused on three key variables namely, *spend amount*, *order quoted price* and *payment_quantity* where the *spend_amount* represents the actual amount expended ('actual spend') while the product of the latter two represents the contracted spend amount ('contracted spend'). An additional variable was created which calculated the log difference (percentage change) between actual and contracted spend was labelled *difference*. Figure 8 that follows plots the distributions of actual versus contracted spend over each year of the time period analysed. The results indicate that in general, actual spend largely overlaps with the contracted amount. The figure further indicates that the distributional plots do not show any distinct change in pattern across the years, barring 2020 displaying lower levels of kurtosis. The outcome is expected as 2020 was unique in terms of conventional government spending being replaced by emergency COVID-19-related procurement.

Figure 8: Quoted vs. actual spend distribution




Source: authors' calculations based on data presented in Section 2.

Figure 9: Quoted vs. actual spend scatter

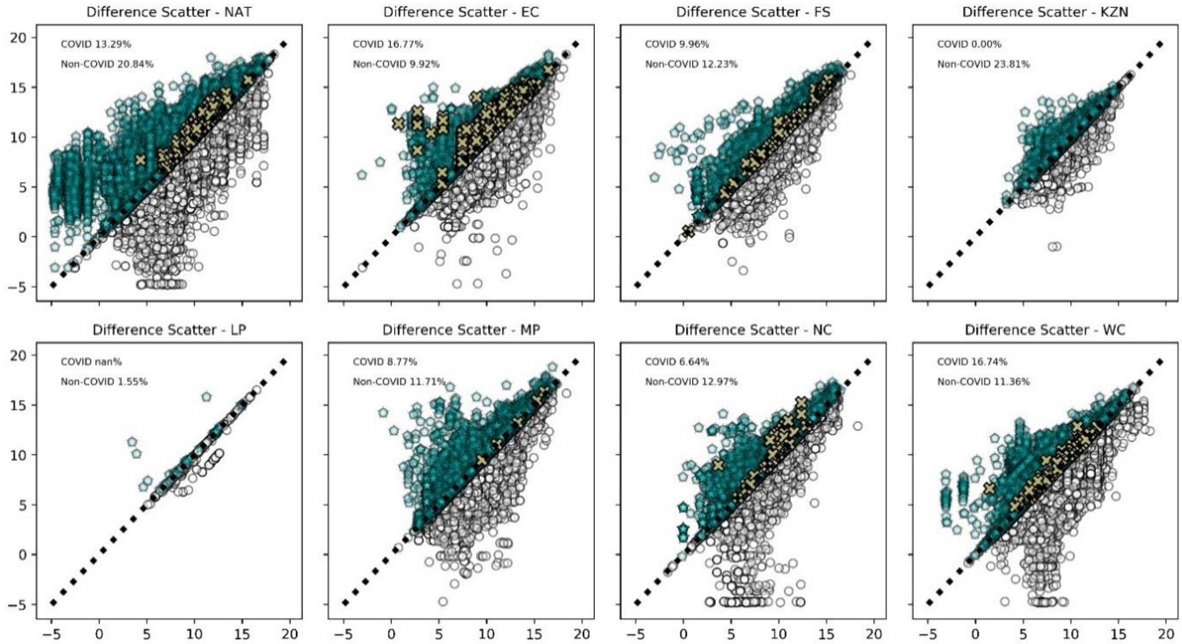



Source: authors' calculations based on data presented in Section 2.

Figure 9 describes scatter plots per year of actual spend (y-axis) plotted against contracted spend (x-axis). The white dots represent contract *difference* not equal to zero while the purple stars represent contract *differences* above the dotted line, i.e. actual spend exceeds contracted spend. Additionally, each scatter describes the proportion of *difference*s greater than 10 per cent scaled by the total number of non-zero differences. The results indicate that 2018 was the year with the highest proportion of *differences* greater than 10 per cent, with 2020 being the lowest. Years 2020 and 2019 were economically similar, achieving proportions of 8.43 per cent and 8.31 per cent respectively.

Figure 10 describes the *difference* scatters stratified by government entity. Each scatter depicts non-zero actual versus contracted spend (white dots), positive *difference*s related to non-COVID-19 expenditure (blue stars) and positive to COVID-19 procurement (olive crosses). The top left text in each scatter describes positive *difference*s scaled by all procurement of the respective government entity across COVID and non-COVID-19 procurement. First, the results indicate that Limpopo and KwaZulu-Natal did not procure any COVID-19-related goods and services at a provincial level. Second, both the Eastern and Western Cape experienced a relative increase in the proportion of instances of actual spend amounts exceeding contracted procurement values during COVID-19. Conversely, the Free State, Northern Cape, Mpumalanga as well as procurement at a national level all experienced a decrease across COVID-19-related procurement contracts.

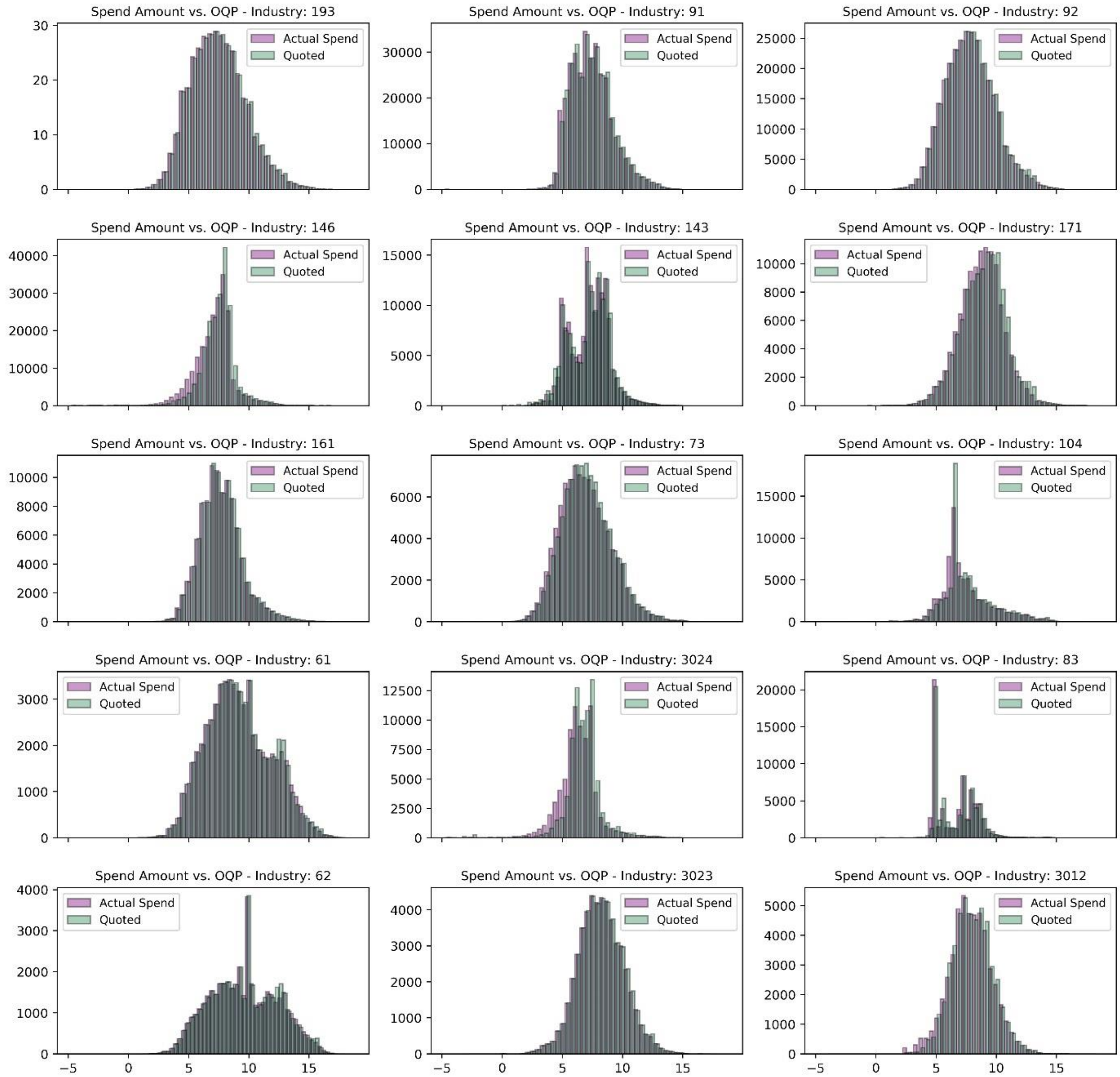
Figure 10: Quoted vs. actual spend scatter per government entity




Source: authors' calculations based on data presented in Section 2.

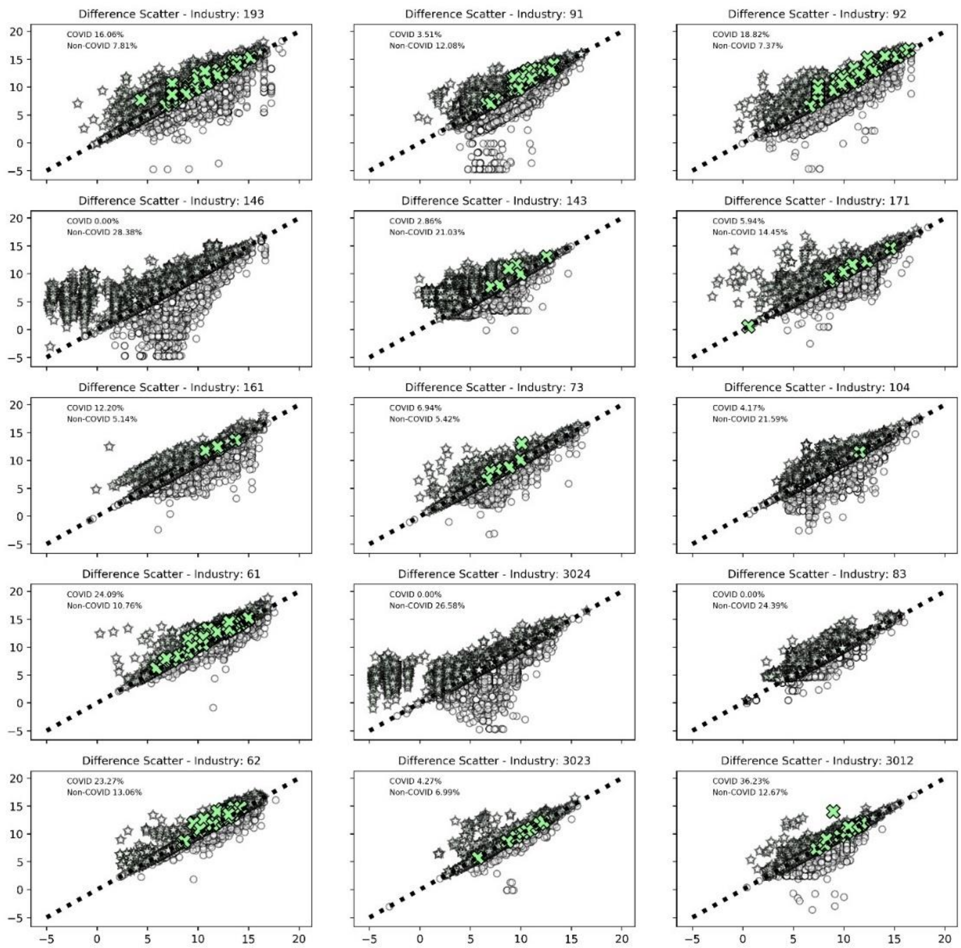
Figure 11 describes the distributional differences in actual amounts expended versus contracted spend across 15 supplier industries with the highest amount of aggregate spend amount over the sample period. The results show that *difference* split according to supplier industry varies dramatically on a distributional basis, yet once again show a high level of consistency between actual against contracted spend over the full sample period.

Figure 11: Quoted vs. actual spend distributions across industry




Source: authors' calculations based on data presented in Section 2.

Figure 12: Quoted vs. actual spend scatters across industries




Source: authors' calculations based on data presented in Section 2.

Figure 12 describes the *difference* scatters stratified according to industry code. Once again, each scatter displays all non-zero *difference*s, positive non-COVID-19 procurement *difference*s (grey star) and positive COVID-19 procurement *difference*s (lime cross). Within each sub-scatter, the proportion of positive *differences* between actual amount against contracted amount is described as a percentage over total procurement, split between COVID and non-COVID-19 spending.

Based on the analysis per Figure 12, the results indicate that contracts where suppliers fall within the industry codes 193 (other personal service activities), 92 (food and beverage services), 61 (construction of buildings), 161 (education), 62 (civil engineering) and 3012 (pharmaceutical manufacturing) experienced significant increases in actual amounts exceeding contracted amounts specific to COVID-19-related procurement. Conversely, industries such as 73 (retail trade), 91 (accommodation), 104 (telecommunications), 171 (human health and social work activities), 143 (travel and tourism) and 3,023 (other manufacturing) experienced significant drops in actual versus contracted differences in terms of COVID-19-related procurement.

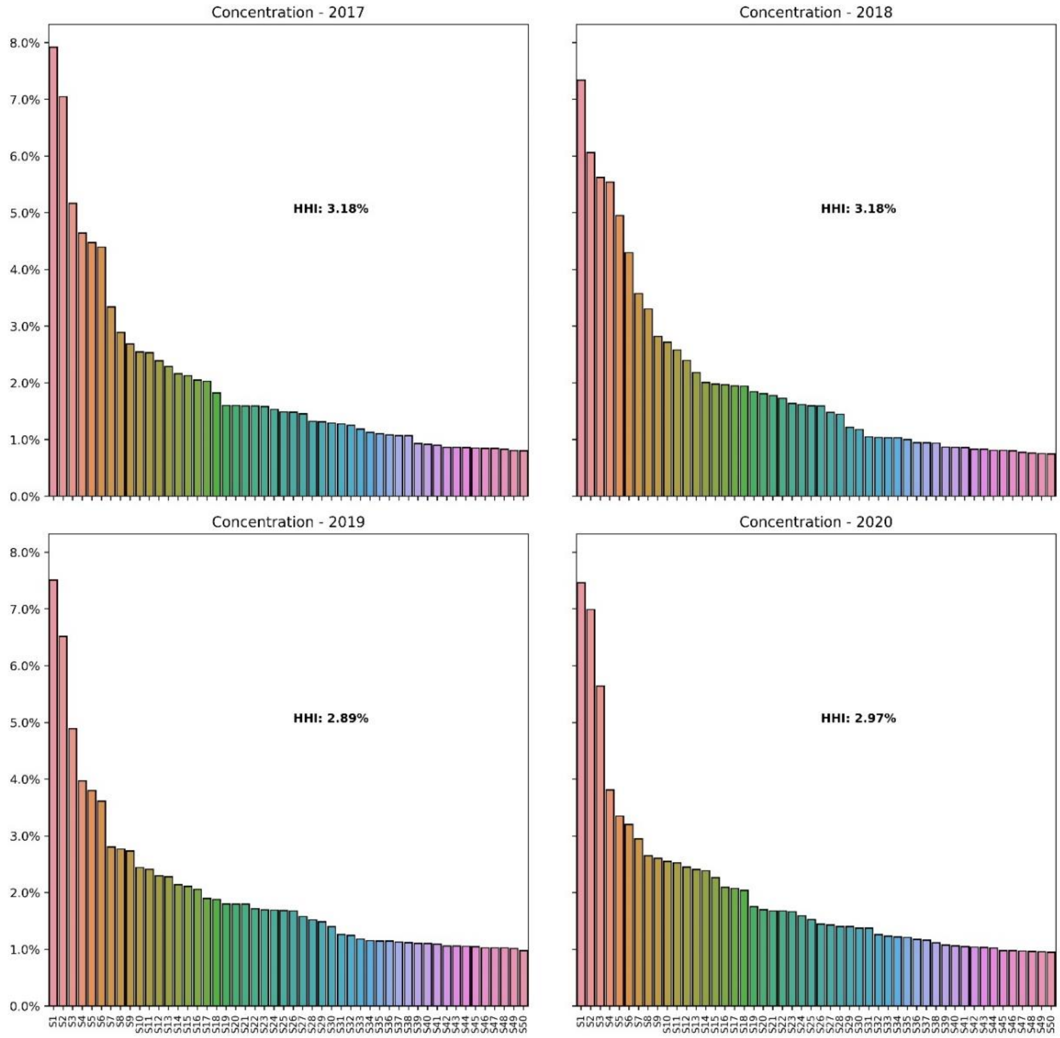
## 6 Concentration analysis

The Herfindahl-Hirschman Index is applied as a proxy for the level of competition in a market economy where high index values imply high levels of concentration or monopolistic characteristics across a subset of suppliers. Similarly, low index measures indicate high levels of spread and equitable competition (Rhoades 1995; Pavic, Galetic and Piplica 2016; Carril and Duggan 2020). The HHI formula is described below

$$HHI = \sum_{i=1}^{n} s_i^2$$

And is the sum of squared market shares across *n* suppliers. The analysis conducted focuses on real *spend amount* of the top 50 suppliers[4] stratified according to industry and province (see appendix A3 for a simple 10 supplier example and the impact of concentration on HHI). *industry_classification_code* of the supplier industry was taken to denote the industry to which a procurement contract belongs. Importantly, there is the possibility that suppliers were contracted to supply goods other than those which fall in their industry.

Figure 13: Concentration through time




Source: authors' calculations based on data presented in Section 2.

Government entity is less considered in literature; however, such analysis may prove useful if government entities responsible for procurement are treated as micro-economies or markets. Like

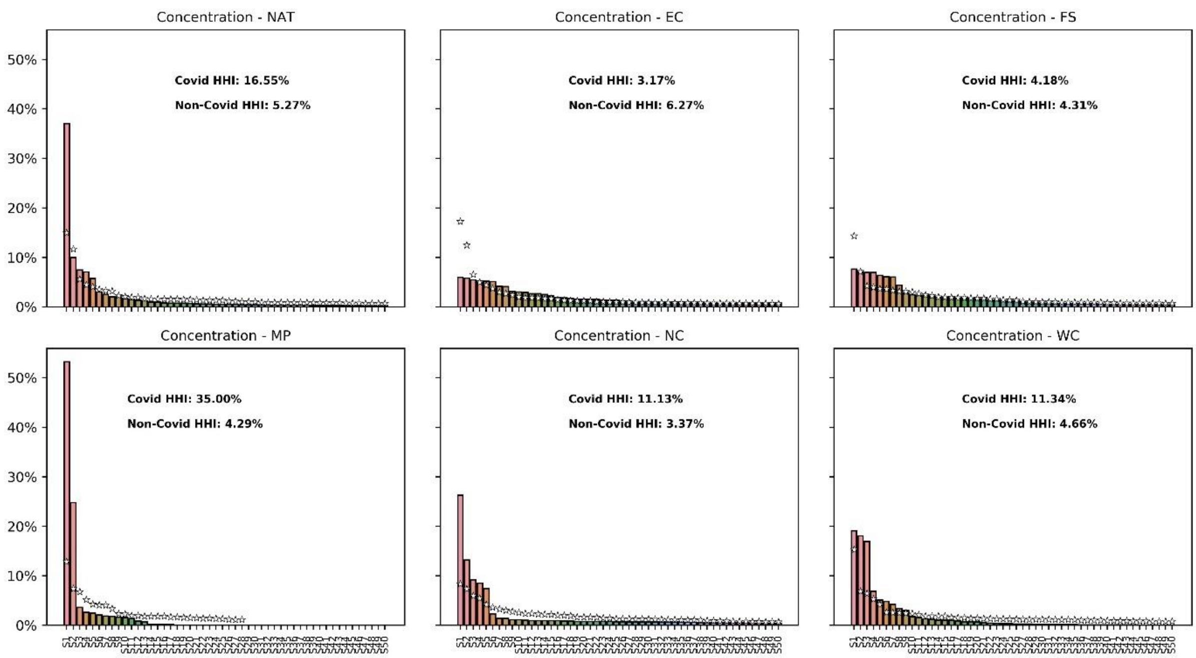
[4] Per literature, analysis typically limited to top 50 firms/suppliers.

the contract deviation analysis, the objective is to determine the impact of COVID-19 on concentration. If concentration experiences a dramatic increase, it may signal a higher propensity for and occurrence of illicit or corrupt procurement practices.

Figure 13 describes annual concentration over the period analysed. The bar chart describes the top 50 suppliers by *spend_amount* scaled by cumulative *spend_amount* as well as the HHI in each of the respective years. The result indicates that when looking at overall procurement spend across all suppliers within a particular year, the level of concentration remains both low and relatively constant, implying high levels of competition and low levels of monopolistic tendencies across the top 50 suppliers.

Figure 14 describes concentration analysis conducted across the six of the eight government entities, with each bar representing the COVID-19 *spend amount* of the top 50 suppliers scaled by their cumulative COVID-19 spend amount. Each sub-figure also displays white stars which represent the level of concentration across the top 50 suppliers related to non-COVID-19 procurement as well as the COVID-19 versus non-COVID 19 HHI. The results indicate that procurement conducted by Mpumalanga, Northern Cape, Western Cape and on a National level have higher levels of concentration associated with COVID-19-related procurement. Further, both the Eastern Cape and Free State, the two largest procuring government entities of COVID-19-related goods and services, display lowered levels of concentration (the same however cannot be said about the third largest procurer, national government).

Figure 14: Concentration across provinces

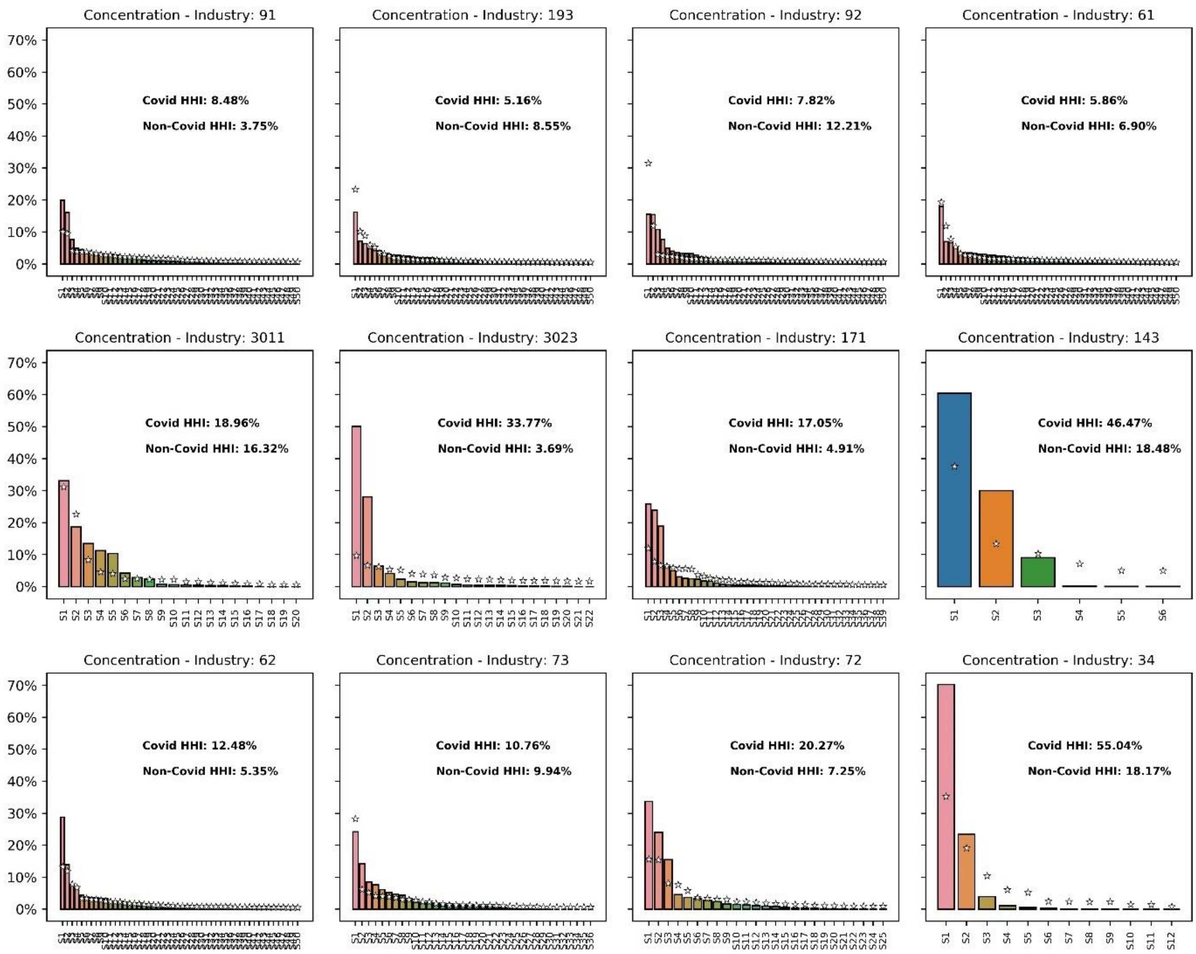



Source: authors' calculations based on data presented in Section 2.

More importantly, when considering the difference in concentration between COVID and non-COVID-19 procurement, the results indicate that concentration increased for the Western Cape, KwaZulu-Natal, Limpopo, Northern Cape, North West and Mpumalanga. Conversely, Gauteng, Eastern Cape and the Free State all experienced drops in concentration with respect to COVID-19-related expenditure. In terms of the magnitude of change, the North West experienced the highest increase in concentration, producing an increase of over 1400 per cent when compared to the HHI measured for non-COVID-19 procurement. All other province increases in HHI averaged around 100 per cent barring the Western Cape which experienced a 59 per cent increase in concentration. Anecdotally, the HHI results for the Western Cape are surprising given that the Western Cape is generally considered less prone to frivolous expenditure and procurement related corruption.

Figure 15 that follows describes the change in the level of concentration across industries. Given the relatively narrow period as well as lowered focus on non-essential procurement associated with COVID-19-related procurement, several industries had less than 50 suppliers. For comparative purposes, the number of applicable suppliers that engaged in COVID-19-related procurement was applied as *n* when calculating and comparing HHI for the said industries non-COVID-19-related procurement.

Figure 15: Concentration across industries




Source: authors' calculations based on data presented in Section 2.

The results of the analysis indicate that seven of twelve industries considered experienced dramatic increases in concentration associated with COVID-19 procurement, namely: 91 (accommodation), 62 (civil engineering), 72 (wholesale trade), 171 (human health and social work activities), 143 (travel agency), 3023 (other manufacturing) and 34 (textile manufacturing). Several industries experienced modest increases in concentration, namely 91 (accommodation), 3011 (Manufacturing of chemical products), 62 (civil engineering), 73 (retail) and 72 (wholesale trade). As with the government entity analysis, several industries experienced decreases in concentration, implying that competition increased, namely 193 (other personal services), 92 (food and beverage services) and 61 (construction). The results of the analysis indicate that on both a provincial and industry level, the impact of COVID-19 on procurement competition levels varied dramatically, potentially being a symptom of the unique nature of the pandemic and government's subsequent reaction. However, instances of extreme increases could possibly indicate higher levels of corruption, specifically procurement activity that may have flouted processes prescribed by the PFMA.

# 7 Outlier analysis

The final analysis considered focuses on outlier detection and applies the popular Mahalanobis Distance ('MD' hereafter) conceptualized by Mahalanobis (1936). MD is a measure of distance between a point *P* and a distribution *D*. MD differs to conventional Euclidian distance as it allows for the determination of distance in a multi-variate space where the distribution of independent variables is applied to determines *P*'s distance to the mean of *D*. Put differently, MD is effectively the calculation of a multi-variate z-score that offers a simplistic framework for classifying data points as outliers and statistically determining the extent of an outlier. The requirements of calculating MD are like that of typical regression analysis, namely that independent variables are correlated with the dependent variable but uncorrelated with each other. Lastly, critical values can be calculated using a chi-square with degrees of freedom equal to the *n* independent variables applied.

In reference to COVID-19 corrupt procurement detection, MD was applied to:

a) Determine whether COVID-19 pricing for specific and discernible goods and services can be considered outliers.
b) Evaluate whether COVID-19 *difference* instances (actual versus contracted) can be considered outliers.

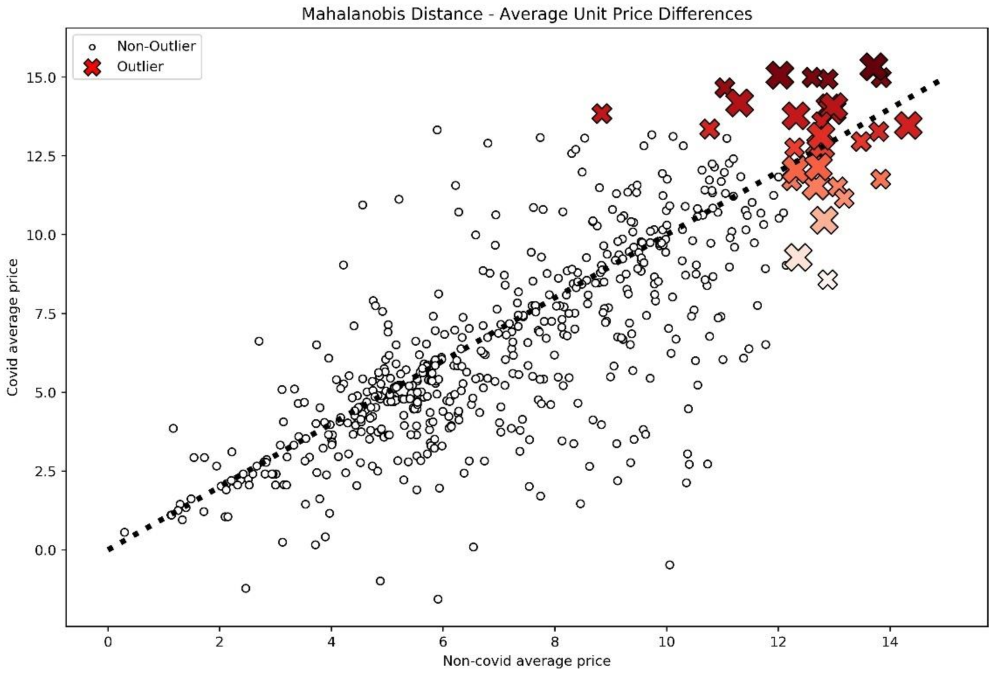
## 7.1 Payment unit price outlier analysis

To test for outliers in COVID-19 pricing of goods and services, the underlying procurement data was limited to instances where *payment_unit_price* was equal to *spend_amount* scaled by *payment_quantity* and that the *unit_of_issue* variable was labelled *EA,* representing singular units ('EA' equates to 'each') and not groupings of items.[5] The *icn_description* was then applied as a means of determining the item or service specific to the respective contract. The empirical design of the MD test applied the COVID-19 real price of the good or service and applied the historical (real) average non-COVID-19 prices as the explanatory (independent) variable.

Figure 16 describes the output of the MD test conducted on the average COVID-19 unit price against the average non-COVID-19 unit price. Within the scatter, each dot represents a specific good or service while an observation is presented as a red cross if the chi-squared p-value associated with the MD statistic is less than 10 per cent. The size of the cross indicates the respective points significance, while the hue represents whether the outlier is above (higher price entails darker red) or below (lower price entails lighter red) the diagonal dotted line. Consistent with the contract difference and concentration analysis presented above, outlier analysis is conducted on government entity where the proportion of significantly positive outlier instances are explored.

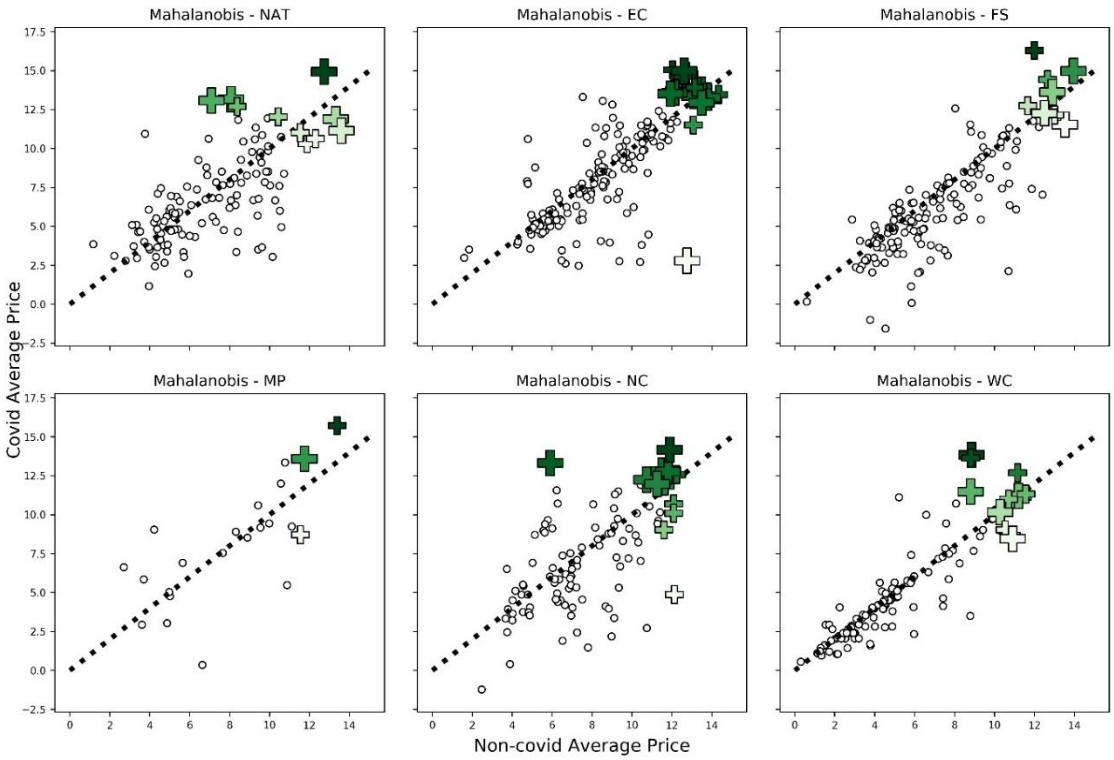
[5] Limiting the data on this basis only resulted in a 2.2% attrition of the original combined datasets.

Figure 16: Mahalanobis distance calculated on average payment unit price




Source: authors' calculations based on data presented in Section 2.

Figure 17: Mahalanobis distance calculated on average payment unit price per province




Source: authors' calculations based on data presented in Section 2.

Figure 17 describes MD analysis conducted on a provincial level and for the purpose of comparison, limits the analysis to the specific goods and services procured during COVID-19. This allows for a like-for-like test of COVID-19 real prices against aggregate non-COVID-19 real prices. The results indicate that the Eastern Cape, Northern Cape and Free State experience a higher number of positive and significant outliers across COVID-19-specific goods and services. The table below describes the top 10 goods and services price outliers (*icn_description*) per government entity.

Table 1: Top 10 outlier procured goods and services per province

| Rank | National | Eastern Cape | Free State | Mpumalanga | Northern Cape | Western Cape |
|---|---|---|---|---|---|---|
| 1 | Consulting*** | Professional fees*** | File Server*** | Services*** | Training*** | Security*** |
| 2 | Printing*** | Construction/Repair*** | Services*** | Cleaning*** | Clothing*** | Cleaning*** |
| 3 | Service *** | Construction*** | Building Repair*** | Fumigation*** | Drill*** | Advertisement*** |
| 4 | Service *** | Fees*** | Project*** | Security | Motor Vehicle*** | Hygiene*** |
| 5 | Courier*** | Gas supply information*** | Printing*** | Paint | Sign board*** | Advertisement*** |
| 6 | Digital Thermometer*** | Building alterations and additions*** | Contracted Services*** | Delivery | Disinfection*** | Cleaning*** |
| 7 | Cleaning Services*** | Service – Payment of water*** | Ventilators** | Poverty Relief | Provincial Road*** | Printing*** |
| 8 | Disinfectant** | Building Renovations*** | Fencing* | Delivery of PPE | Blood Gas Analyser*** | Advertising Production** |
| 9 | Scanner** | Building Renovations*** | Services | Masks | Electrical Maintenance** | Gas hire** |
| 10 | Advertisement* | Emergency building work | Blood Gas Analyser | Masking Tape | Laboratory** | Doctors fees* |

Note: ***, **, * indicate significance at the 1%, 5%, and 10% level.

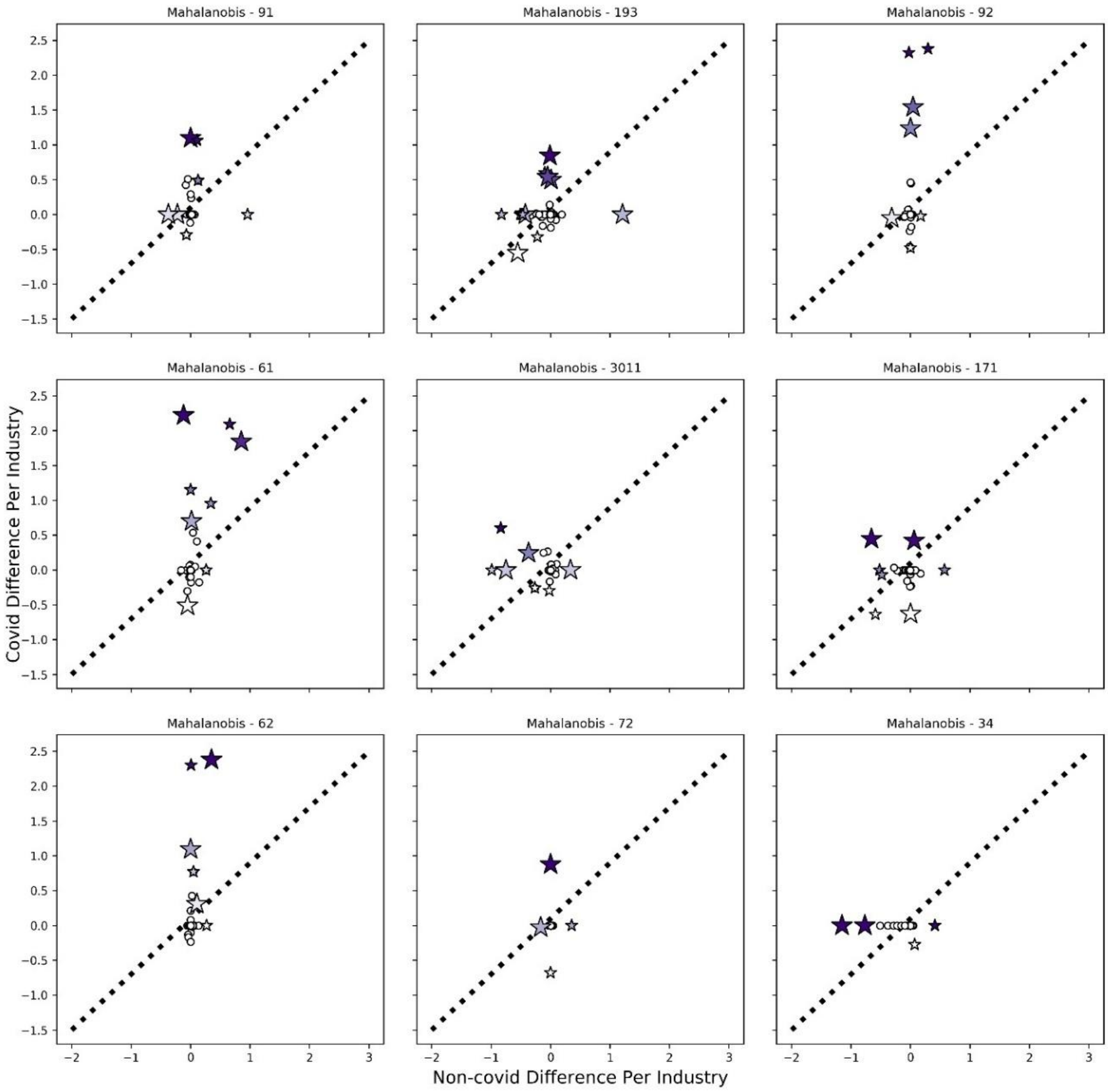
Source: authors' calculations based on data presented in Section 2.

The result of the government entity analysis displayed in Table 1 indicates that COVID-19-specific pricing spiked for the goods and services described in the table above. It is possible that the said outliers are reasonable given the possible shifts in supply and demand for certain goods caused by the pandemic, however, all outliers should be investigated to determine whether such differences are possibly explained by illicit or corrupt procurement pricing. As expected, numerous goods and services relate to expected items such as medical equipment, sanitisers, and fumigation, however, a number of generalized services are present across the top 10 per province. The application of generalized codes or labels for either goods or services i.e. 'consulting', 'service', 'building repairs' can itself become an additional identifier that forms part of the analytical framework, given that broad, opaque definitions may be an additional indication of illicit or corrupt procurement activity.

## 7.2 Contract deviation outlier analysis

To explore potential outliers on a supplier industry level, MD analysis is conducted on calculated *difference* between average actual versus contracted spend grouped by *icn_description* and supplier industry. Therefore, the MD test can be viewed as an extension of the contract deviation analysis but focuses on average *differences* across specific goods and services associated with COVID and non-COVID-19 procurement. Scatters per Figure 18 now describe statistically significant outlier *difference*s as stars where size indicates statistical significance and hue describes the sign of the outlier in terms of whether it sits above or below the diagonal dotted line (higher (lower) difference equates to darker (lighter) purple). The results show that industries 193 (other personal services), 92 (Food and beverage services) and 61 (construction of buildings) indicate a relatively higher proportion of contract deviation outliers. As with the provincial analysis above, the table that follows describes the top 10 COVID-19-specific *difference* outliers across industries based on *icn_description*.

Figure 18: Mahalanobis distance calculated on difference between actual and contracted spend amount per industry




Source: authors' calculations based on data presented in Section 2.

Table 2 shows several goods and services that would have been expected to experience dramatic increases in demand as several of the top 10 goods and services across industries relate to medical and hygiene related items. However, one should note that unlike the price analysis, the dependent variable is the log difference between actual and contracted spend. The fact that the procurement of such items deviated so dramatically from historical differences may only be partially explained by the increased demand of COVID-19 and hence deserve further scrutiny.

Table 2: Top 10 outlier procured goods and services per industry

| Rank | Accommodation (91) | Other services (193) | Food Services (92) | Construction (61) | Chemical Manufact. (3011) | Other Manufact. (3023) |
|---|---|---|---|---|---|---|
| 1 | Dishes*** | Disposal*** | Training*** | Construction*** | Medical gas*** | Linen saver*** |
| 2 | Sealer*** | Laryngoscope*** | Construction repairs*** | Waterpump repair *** | Medical oxygen*** | ID band*** |
| 3 | Emergency repairs*** | Advertising*** | Services*** | Drill*** | Medical oxygen*** | Medical filter*** |
| 4 | Renovation*** | Water payment*** | Waterpump repair*** | Cleaning compound*** | Soap*** | Medical mask |
| 5 | Building** | Scanner*** | Building repairs*** | Emergency repairs*** | Gas cylinder hire*** | Trolley |
| 6 | Detergent* | Medical*** | Soap liquid*** | Drill borehole** | Gas cylinder hire** | Scissors |
| 7 | Services* | Cleaning*** | Cleaning Services* | Medical equipment service** | Refilling gas** | Locker |
| 8 | Delivery | Construction repairs*** | Cleaning and building maintenance | Construction repairs** | Oxygen | Mattress |
| 9 | Polish removal | Hypodermic needle*** | Alcohol | Delivery | Gas installation | Stool |
| 10 | Hygiene cleaning | Hypodermic needle** | Services | Building repairs | Medical oxygen | Medical tray |
| **Rank** | **Human Health (171)** | **Travel (143)** | **Civil engineering (62)** | **Retail (73)** | **Wholesale (72)** | **Textiles (34)** |
| 1 | Personnel agency*** | Accommodation | Sanitizer*** | Disposable gloves*** | Medical clamp*** | Sanitary pad*** |
| 2 | Container hire*** | Conference and seminar venue | Waterpump repair*** | Medical*** | Cleaning*** | Night dress (L)*** |
| 3 | Locum*** | | Delivery*** | Pen*** | Bio indicator*** | Delivery*** |
| 4 | Thermometer*** | | Construction repairs** | Markers*** | Pulpit*** | Draw sheet** |
| 5 | Sponge*** | | Emergency repairs** | Trolley** | Medical dressing | Night dress (M) |
| 6 | Dosimeter*** | | Building repairs* | Sanitizer** | Medical gauze | Mattress cover |
| 7 | Medical*** | | Porridge | Ballpoint pen* | Crutch | Night dress (XL) |
| 8 | Locum services | | Fencing | ECG monitor* | Catering | Pillowcase |
| 9 | Humidifier | | Cleaning compound | Ruler | Tailoring | Shirt |
| 10 | Cleaning | | Construction | Staple remover | Delivery | Service clothing |

Note: ***, **, * indicates significance at the 1%, 5%, and 10% level.

Source: authors' calculations based on data presented in Section 2.

## 8 Recommended price analysis

A final set of analysis was conducted on the recommended price list published by National Treasury related to COVID-19 personal protective equipment (PPE) as at 20 November 2020. The price list contained recommended unit costs as well as minimum and maximum price ranges associated with 16 PPE items. A simple comparison of the recommended prices to the *order_quoted_price* per the combined dataset offers an additional avenue of illicit procurement identification. To test the feasibility of conducting any meaningful analysis associated with the recommended price list, a matching algorithm was applied to determine whether any of the listed product labels could be broadly matched to the *icn_description* per the dataset.

Table 3: Matching algorithm output

| Product | Words to Match | No Match | 1 word | 2 words | 3 words | >3 words | Matching icn_description |
|---|---|---|---|---|---|---|---|
| Surgical Mask Patient | 3 | 754 | 50 | 4 | 0 | 0 | - MASK,DUST / SURGICAL / CLOTH MASKS ALL TYPES<br>- MASK,MEDICAL SURGICAL FACE MASK<br>- MASK,MEDICAL SURGICAL AS PER QUOTE<br>- MASK,MEDICAL ,FACE,SURGICAL,RE-OPERATOR 3M |
| Surgical Mask Health Care Worker | 5 | 757 | 46 | 5 | 0 | 0 | - MASK,SAFETY HEALTH PROTECTOR WHITE SOFT PAPER<br>- MASK,DUST / SURGICAL / CLOTH MASKS ALL TYPES<br>- MASK,MEDICAL SURGICAL FACE MASK<br>- MASK,MEDICAL SURGICAL AS PER QUOTE<br>- MASK,MEDICAL ,FACE,SURGICAL,RE-OPERATOR 3M |
| Mask Respirator | 2 | 770 | 33 | 5 | 0 | 0 | - MASK,DUST MOULDER RESPIRATOR MASK<br>- MASK,DUST MOULDER RESPIRATOR MASK<br>- MASK,MEDICAL RESPIRATOR - N95 PARTICULATE - 1860<br>- MASK,SAFETY , SPRAYPAINTING, RESPIRATOR, HALF<br>- MASK,MEDICAL RESPIRATOR N95 HALYARD |
| Apron | 1 | 803 | 5 | 0 | 0 | 0 | - APRON<br>- MEDICAL EQUIPMENT SET , PPE KIT COVID-19 (SURG...<br>- APRON , DISPOSABLE, POLYETHYLENE, RED, 127CM X.<br>- APRON BOX 100 PIECES<br>- APRON , DISPOSABLE, POLYETHYLENE, BLUE, 127CM |
| Eye Protection | 2 | 802 | 6 | 0 | 0 | 0 | - TUBE,ENDOTRACHEAL , CUFFED, ORAL/NASAL, MAGILL...<br>- TUBE,ENDOTRACHEAL , ARMOURED, CUFFED, MURPHY<br>- APRON , DISPOSABLE, POLYETHYLENE, RED, 127CM X.<br>- TUBE,ENDOTRACHEAL , CUFFED, ORAL/NASAL, MAGILL...<br>- TUBE,ENDOTRACHEAL , CUFFED, ORAL/NASAL, MAGILL...<br>- TUBE,ENDOTRACHEAL , CUFFED, ORAL/NASAL, MAGILL... |
| Visor Face Shield | 3 | 795 | 11 | 2 | 0 | 0 | - SHIELD,MEDICAL FACE CLEAR LENSES - SF2510<br>- SHIELD,SAFETY CLEAR FOR FACE |
| Gowns | 1 | 806 | 2 | 0 | 0 | 0 | - GOWN DISPOSABLE GOWNS<br>- GOWN DISP. GOWNS HOSPITAL LARGE |
| Coveralls | 1 | 808 | 0 | 0 | 0 | 0 | N/A |
| Boot Covers | 2 | 808 | 0 | 0 | 0 | 0 | N/A |
| Digital Thermometer | 2 | 794 | 8 | 6 | 0 | 0 | - THERMOMETER COVID 45 DIGITAL THERMOMETER<br>- THERMOMETER,DIGITAL<br>- THERMOMETER THERMOMETER DIGITAL;AMAREL<br>- TTHERMOMETER,DIGITAL FOR COVID-19<br>- THERMOMETER,MEDICAL DIGITAL<br>- THERMOMETER,MEDICAL DIGITAL INFRARED |
| Sanitisers and Disinfectants | 3 | 787 | 21 | 0 | 0 | 0 | N/A (all related to AND) |
| Biohazard bags | 2 | 806 | 2 | 0 | 0 | 0 | - BAG CLEAR REFUSE BAGS 25MIC (20S)<br>- BAG PLASTIC BAGS #800+180X1800X60 |
| Body Bags | 2 | 803 | 5 | 0 | 0 | 0 | - BODY BAG PLASTIC BODY BAG WITH ZIP<br>- BODY BAG CHILD(INFANT)<br>- BODY BAG<br>- BAG CLEAR REFUSE BAGS 25MIC (20S)<br>- BAG PLASTIC BAGS #800+180X1800X60 |
| Examination Gloves non-sterile | 3 | 782 | 23 | 3 | 0 | 0 | - GLOVES EXAMINATION SMALL (COVID-19)<br>- GLOVES EXAMINATION MEDIUM (COVID-19)<br>- GLOVES EXAMINATION LARGE (COVID-19) |
| Gloves examination or surgical sterile | 5 | 750 | 51 | 7 | 0 | 0 | - GLOVES LATEX NON STERILE MEDIUM<br>-GLOVES ,SURGICAL,NATURAL,RUBBER,LATEX,NONSTERI..<br>- GLOVES LATEX NON STERILE LARGE<br>- GLOVES SURGICAL GLOVE MEDIUM<br>- GLOVES EXAMINATION SMALL (COVID-19)<br>- GLOVES EXAMINATION MEDIUM (COVID-19)<br>- GLOVES EXAMINATION LARGE (COVID-19) |
| Cloth Mask | 2 | 764 | 39 | 5 | 0 | 0 | - MASK,DUST / SURGICAL / CLOTH MASKS ALL TYPES<br>- MASK,MEDICAL FACE CLOTH MASK<br>- MASK,SAFETY 3 LAYER CLOTH MASK<br>- MASK,SAFETY CLOTH<br>- CLOTH CLOTH MASK |

Source: authors' calculations based on data presented in Section 2.

Table 3 describes the results of the matching algorithm run on the recommended price list product labels. The far-left column describes the recommended price list item label and the second column (light green) the number of words to match. Columns 3-7 describe the number of words successfully matched while the final column describes the *icn_description* labels that matched. The results of the analysis indicates that the disparity in item labelling would render any analysis using the recommended PPE price list inconclusive and inconsequential. As an example, consider Mask Respirator. The recommended price label matched successfully with 5 *icn_description*'s, yet there are subtle to extreme differences within the *icn_description*'s that may result in incorrect conclusions. As shown in the final column, two possible *icn*_description's include 'medical' but are differentiated by the terms 'particulate', 'halyard' and '1860'. If such differentiators have a significant impact on the underlying price per item, a simple aggregation of pricing across may flag a supplier or contract that was licit and legitimate (Type I error). In summary, without a standardized index/labelling across products, recommended PPE price lists cannot be meaningfully applied as an additional test in identifying potentially illicit procurement activity associated with COVID-19 spending.

## 9 Summary and conclusion

The purpose of this study is to analyse COVID-19-specific procurement and attempt to identify instances of illicit or corrupt procurement practices. An initial analysis of spending purely related to COVID-19 is described, where outlier spending was identified in Sundry payments for COVID-19-related spending. However, to confirm outlier spending behaviour, an analysis of COVID-19-related spending and non-COVID-19-related spending was undertaken, and a framework developed to identify potentially illicit spending patterns.

This analysis to create the framework is split into three core components (contract deviation, concentration and outlier tests) where analysis is conducted on a government entity and supplier-industry level. Contract deviation analysis focuses on the determination of whether differences in actual and contracted spend amounts increased or decreased across COVID-19-related procurement. Concentration analysis considers the changes in supplier concentration between COVID-19 and non-COVID-19 procurement using HHI while outlier analysis, using MD, attempts to determine whether COVID-19 pricing and contract deviation differences can be considered outliers against their non-COVID-19 counterparts. The results of the analysis indicate that under each measure there are numerous possible instances of corruption and illicit procurement, however, the accuracy of each measure is unexplored. Importantly, each analytical component/test satisfies an *a priori* theoretical framework for determining whether a contract is potentially illicit or corrupt. Further, since each of the analysis components are independent and focus on different aspects of the data, they can be combined into a multi-stage ensemble of tests able to isolate specific contracts and/or suppliers that require further investigation. An area of necessary further research would be the assessment of the independent ability of each of the measures in detecting illicit contracts and whether some combination or permutation results in greater levels of accuracy.

## References


Benford, F. (1938). 'The Law of Anomalous Numbers'. Proceedings of the American Philosophical Society, 551–72.

Carril, R., and M. Duggan (2020). 'The Impact of Industry Consolidation on Government Procurement: Evidence from Department of Defense Contracting'. *Journal of Public Economics*, 184: 104141. https://doi.org/10.1016/j.jpubeco.2020.104141

Goodman, W. (2016). 'The Promises and Pitfalls of Benford's Law'. *Significance*, 13(3): 38–41. https://doi.org/10.1111/j.1740-9713.2016.00919.x

Koch, C., and K. Okamura (2020). 'Benford's Law and COVID-19 Reporting'. Economics Letters, 196: 109573. https://doi.org/10.1016/j.econlet.2020.109573

Mahalanobis, P.C. (1936). 'On the Generalized Distance in Statistics'. Delhi: National Institute of Science of India.

McClelland, D., A. Oyenubi, G. Bridgman, D. Page, and U. Kollamparambil (2021). 'Introduction to the Government Procurement Data and Historical Assessment of Procurement Trends'. SA-TIED Technical Note 2021/3. Helsinki: UNU-WIDER.

Michalski, T., and G. Stoltz (2013). 'Do Countries Falsify Economic Data Strategically? Some Evidence that they Might'. *The Review of Economics and Statistics,* 95(2): 591–616. https://doi.org/10.1162/REST_a_00274

Pavic, I., F. Galetic, and D. Piplica (2016). 'Similarities and Differences between the CR and HHI as an Indicator of Market Concentration and Market Power'. *British Journal of Economics, Management & Trade*, 13(1): 1–8. https://doi.org/10.9734/BJEMT/2016/23193

Rhoades, S.A. (1995). 'Market Share Inequality, the HHI, and Other Measures of the Firm-Composition of a Market'. *Review of Industrial Organization,* 10(6): 657–74. https://doi.org/10.1007/BF01024300

Sambridge, M., and A. Jackson (2020). 'National COVID Numbers: Benford's Law Looks for Errors'. *Nature*, 581(7809): 384. https://doi.org/10.1038/d41586-020-01565-5

Varian, H.R. (1972). 'Benfords Law'. *American Statistician,* 26(3): 65.

## Appendix A1:Benford's law analysis

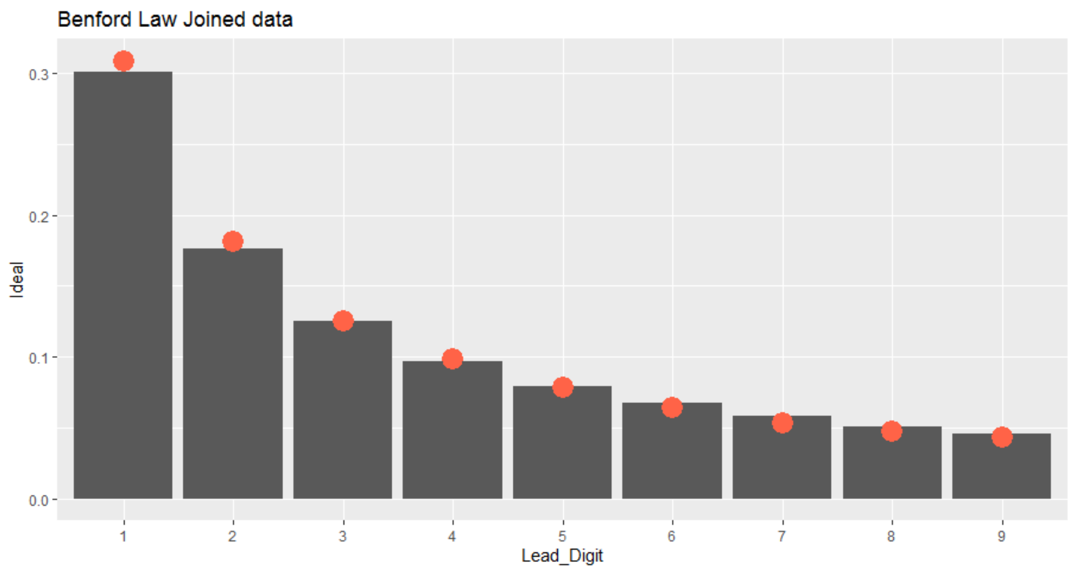
Benford's law is used to detect fraud or flaws in data collection based on the distribution of the first digits of observed data (Koch and Okamura 2020). Benford's law distribution of first digits occurs naturally for exponential process with multiple changes in magnitude (Michalski and Stoltz, 2013). This approach has been used to detect fraud in the economics and accounting literature (Varian, 1972), and more recently it has been used to check COVID-19 mortality and reported cases of COVID-19 (Sambridge and Jackson 2020; Koch and Okamura 2020). Following this approach is therefore appropriate in analysing procurement spending related to COVID-19.

Under Benford's law the expected proportions of numbers in a data set having particular first digits can be calculated as:

$$Prob(first\ digit) = log_{10}\left(1 + \frac{1}{d}\right)$$

where $d$ is the particular first digit. For example, the probability that the first digit is 1 is 0.031. For Benford's law to be applicable the data should satisfy certain requirements (i) sufficient sample size (ii) large span of number values (iii) positively skewed distribution (iv) Not human assigned numbers (Goodman 2016). One can argue that the South Africa procurement data satisfy all these conditions.[6] Figure A1 compares the distribution of *spend amount* in the joined dataset with what is expected under Benford's law. The result shows that the data obey Benford's law (red dots are the observed frequencies of first digits for the spend amount in the procurement data while the bars represent the observed frequency under Benford's law).

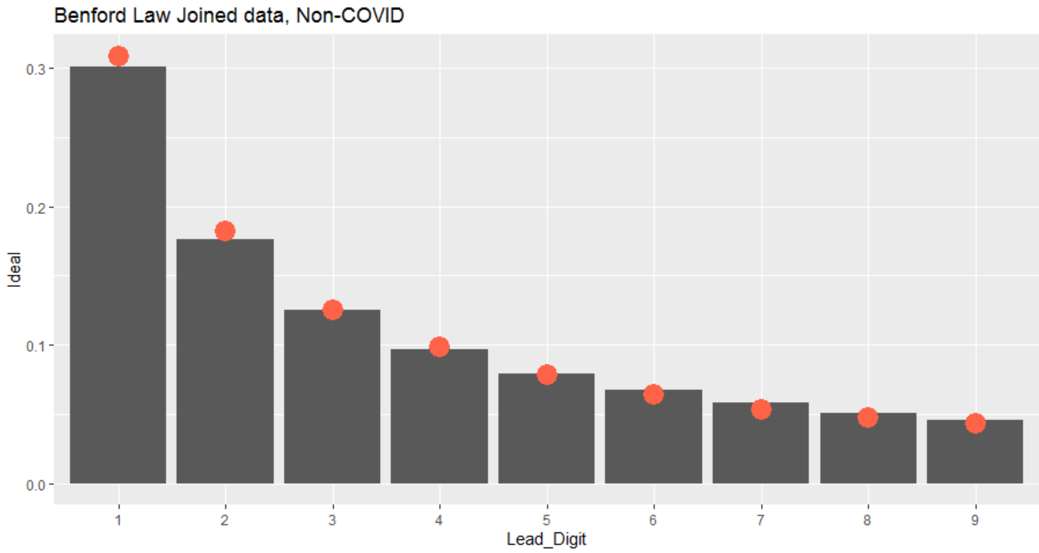
Figure A1.1: Benford's law for the joined data




Source: authors' calculations based on data presented in Section 2.

[6] See Figures A1 and A2 in the appendix of McClelland et al. (2021).

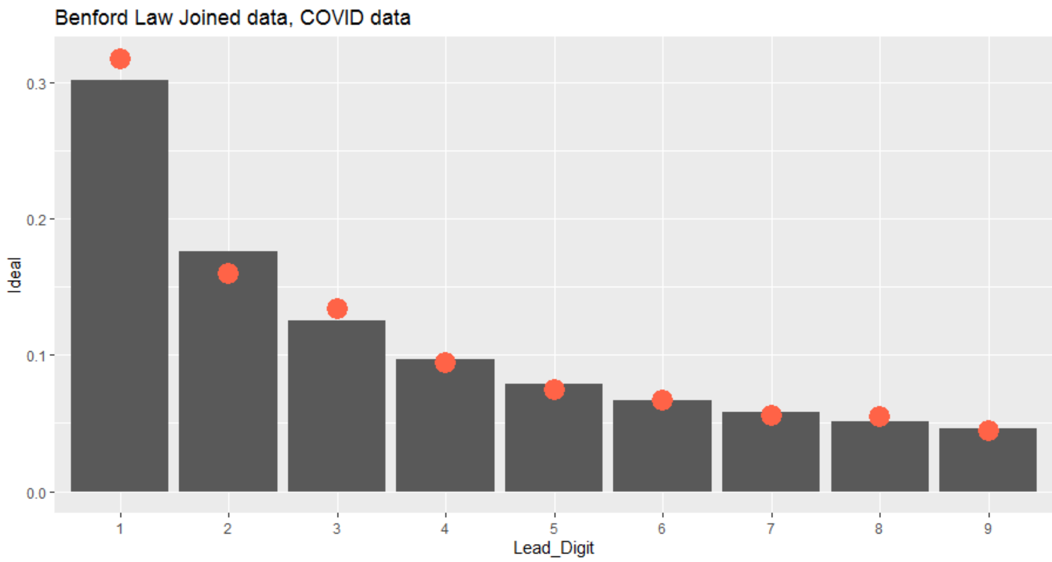
Figure A1.2: Benford's law for the joined data (non-COVID transactions)




Source: authors' calculations based on data presented in Section 2.

Figure A1.2 indicates that Benford's law is not violated for non-COVID-related expenditures. Again, the data conforms to theoretical expectation under Benford's law. However, when we look at the distribution for COVID-related expenditure for the joined data in Figure A1.3 there are some differences mostly for the first digits 1, 2 and 3. These differences are exacerbated when the analysis focuses on BAS data alone (i.e., figure A1.4 considers COVID spending in BAS data) with all of the probabilities being different when compared to expectation.

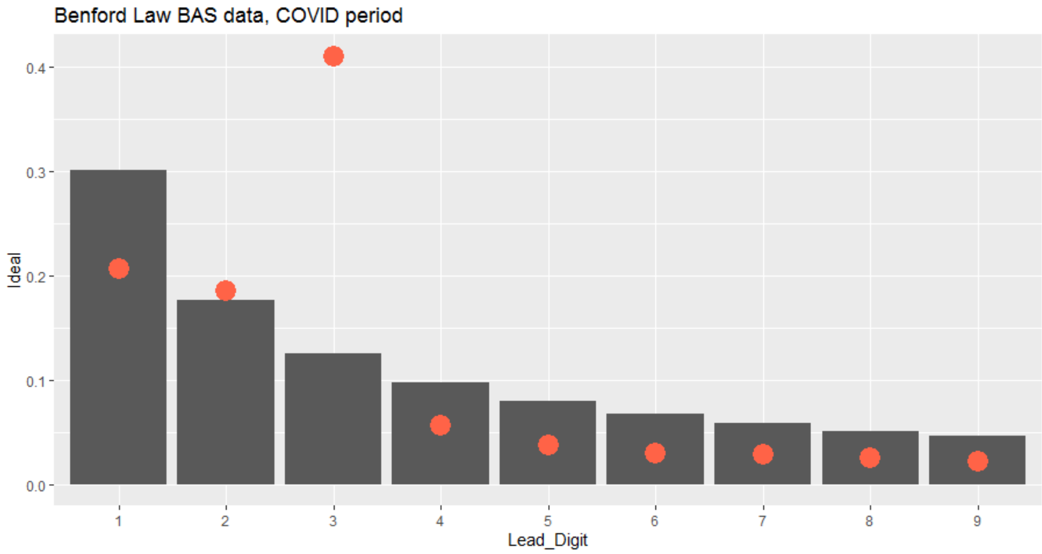
Figure A1.3: Benford's law for the joined data (COVID transactions)




Source: authors' calculations based on data presented in Section 2.

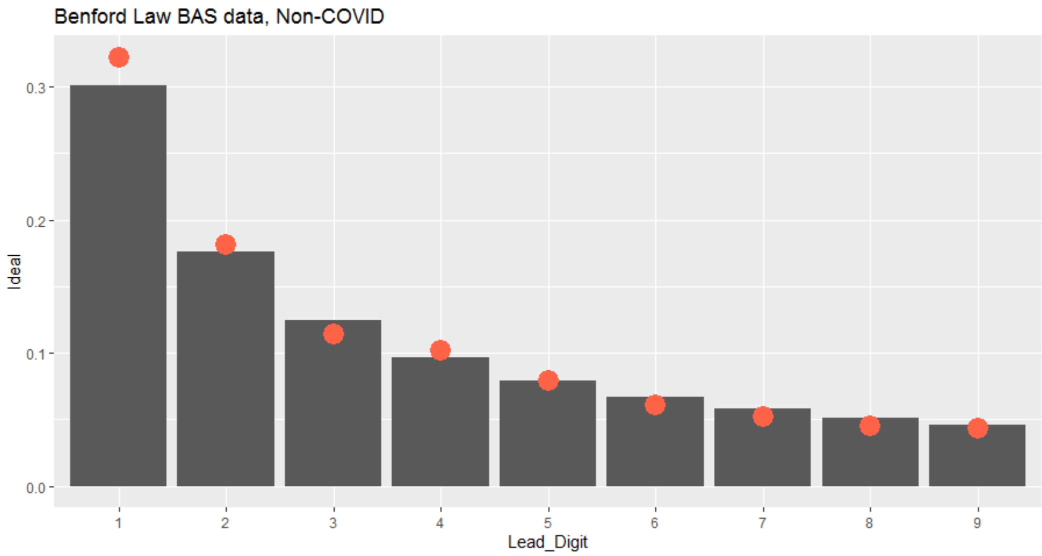
Figures A1.5 presents similar analysis for the complete BAS data. While there are some departures these are not as apparent as when the analysis is restricted to COVID spending (i.e., compare figure A1.4 with A1.5).

Figure A1.4: Benford's Law for BAS data (COVID-Spending)




Source: authors' calculations based on data presented in Section 2.

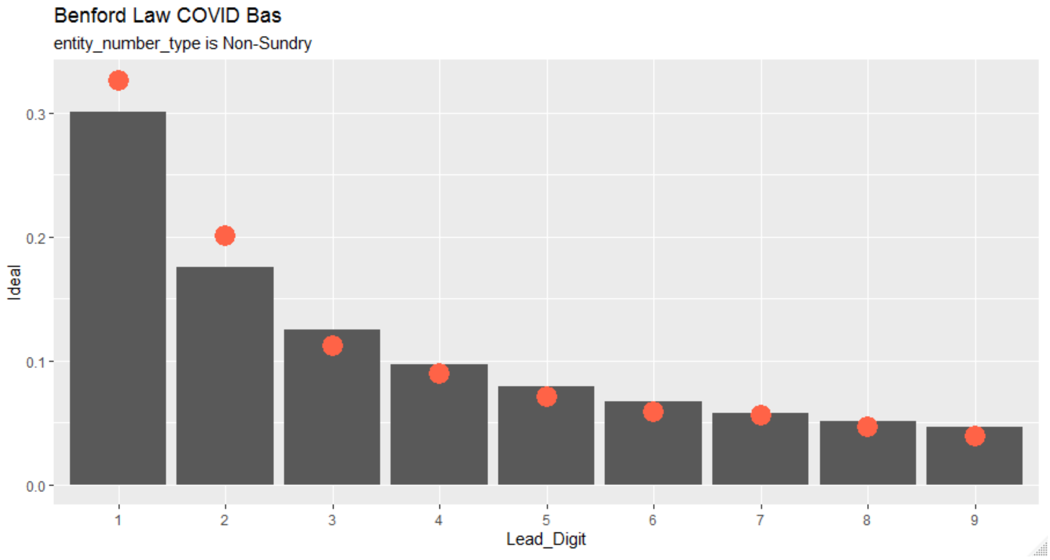
Figure A1.5: Benford's Law for BAS data




Source: authors' calculations based on data presented in Section 2.

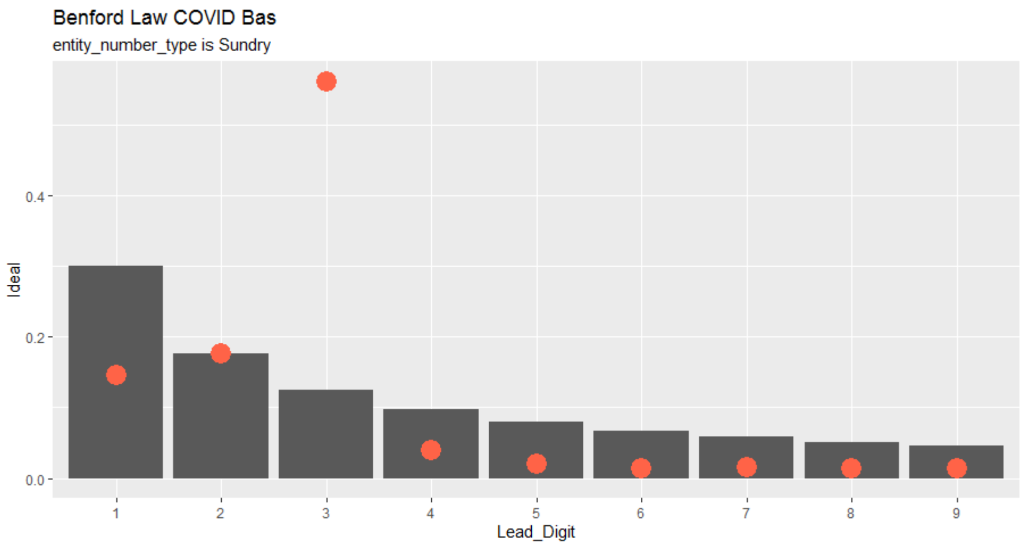
To dig deeper the analysis is run to compare sundry and non-sundry payments. The results are shown in figures A1.6 and A1.7 below. It does appear that most of the departures from the theoretical distribution can be explained by the COVID or/and sundry payments.

Figure A1.6: Benford's Law for BAS data (sundry payments)




Source: authors' calculations based on data presented in Section 2.

Figure A1.7: Benford's Law for BAS data (COVID sundry payments)




Source: authors' calculations based on data presented in Section 2.

This analysis finds that Benford's law is most violated for Sundry payments which are related to COVID-19 spending. One should exercise extreme caution in interpreting these findings, specifically Benford's law is not an automatic fraud detector (Goodman, 2016). There may be reasonable reasons why such patterns emerge. However, Benford's law provides a way to pick a point or points for further investigation in what is otherwise a very large dataset. Therefore, the remainder of this note develops a framework for identifying potentially corrupt procurement spending using empirical techniques for identifying outliers and potential monopolistic behaviour.

## Appendix A2: Industry classification codes

| industry_classification_code | industry_classification_name |
|---|---|
| 3024 | Repair and installation of machinery and equipment |
| 146 | Office administrative, office support and other business support activities |
| 73 | Retail trade, except of motor vehicles and motorcycles |
| 193 | Other personal  service activities |
| 92 | Food and beverage service activities |
| 61 | Construction of buildings (for erection of complete prefabricated constructions from self-manufactured parts not of concrete, see divisions 16 and 25) |
| 36 | Manufacturing of leather and related products |
| 3022 | Manufacture of furniture (for manufacture of furniture of ceramics, concrete and stone, see 2393, 2395, 2396) |
| 91 | Accommodation |
| 7 | Wholesale and retail trade, repair of motor vehicles and motorcycles |
| 161 | Education |
| 105 | Computer programming, consultancy and related activities |
| 202 | Undifferentiated goods- and services-producing activities of private households for own use |
| 136 | Other professional, scientific and technical activities |
| 104 | Telecommunications (for telecommunications resellers, see 6190) |
| 12 | Forestry and logging |
| 71 | Wholesale and retail trade and repair of motor vehicles and motorcycles |
| 11 | Crop and animal production, hunting and related service activities |
| 151 | Public administration and defence, compulsory social security |
| 62 | Civil engineering |
| 38 | Manufacturing of paper and paper products |
| 144 | Security and investigation activities |
| 41 | Electricity, gas, steam and air conditioning supply |
| 182 | Libraries, archives, museums and other cultural activities |
| 137 | Veterinary activities |
| 106 | Information service activities |
| 35 | Manufacturing of wearing apparel |
| 191 | Activities of membership organizations |
| 54 | Remediation activities and other waste management services |
| 19 | Other service activities |
| 3012 | Manufacture of basic pharmaceutical products and pharmaceutical preparations |
| 3011 | Manufacture of chemicals and chemical products |
| 31 | Manufacturing of food products |
| 72 | Wholesale trade, except of motor vehicles and motorcycles (for activities of commission agents for motor vehicles, see 4510) |
| 171 | Human health and social work activities |
| 63 | Specialised construction activities |
| 111 | Financial service activities, except insurance and pension funding |
| 21 | Mining of coal and lignite |
| 85 | Postal and courier activities |
| 143 | Travel agency, tour operator, reservation service and related activities |

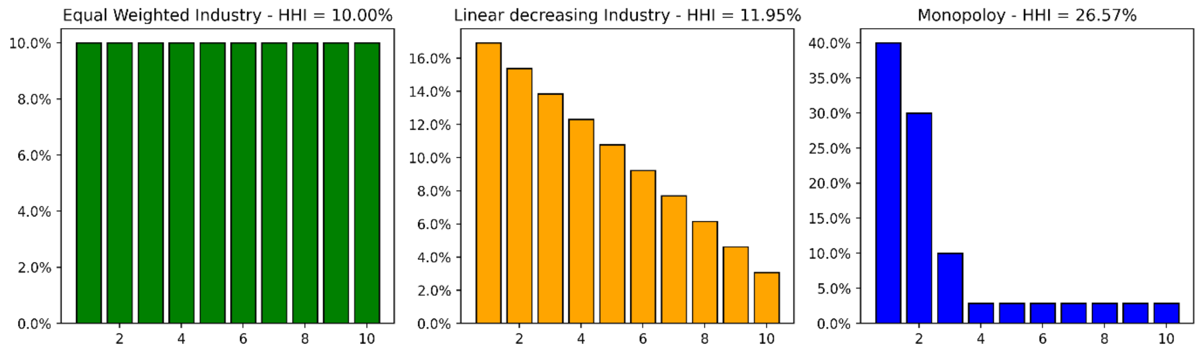
| | |
|---|---|
| 145 | Services to buildings and landscape activities |
| 81 | Land transport and transport via pipelines |
| 3019 | Manufacture of machinery and equipment n.e.c. |
| 3013 | Manufacture of rubber and plastic products |
| 3 | Manufacturing |
| 53 | Waste collection, treatment and disposal activities; materials recovery |
| 52 | Sewerage |
| 192 | Repair of computers and personal and household goods |
| 3016 | Manufacture of fabricated metal products, except machinery and equipment |
| 6 | Construction |
| 112 | Insurance, reinsurance and pension funding, except compulsory social security |
| 3010 | Manufacture of coke and refined petroleum products |
| 3023 | Other manufacturing |
| 181 | Creative, arts and entertainment activities |
| 51 | Water collection, treatment and supply |
| 9 | Accommodation and food service activities |
| 3017 | Manufacture of computer, electronic and optical products |
| 18 | Arts, entertainment and recreation |
| 102 | Motion picture, video and television program production, sound recording and music publishing activities |
| 10 | Information and communication |
| 84 | Warehousing and support activities for transportation |
| 8 | Transportation and storage |
| 17 | Human health and social work activities |
| 172 | Residential care activities |
| 34 | Manufacturing of textiles |
| 121 | Real estate activities |
| 113 | Activities auxiliary to financial service and insurance activities |
| 39 | Printing and reproduction of recorded media |
| 201 | Activities of households as employers of domestic personnel |
| 25 | Mining support service activities |
| 141 | Rental and leasing activities |
| 142 | Employment activities |
| 3015 | Manufacture of basic metals |
| 82 | Water transport |
| 133 | Architectural and engineering activities; technical testing and analysis |
| 132 | Activities of head offices; management consultancy activities |
| 3018 | Manufacture of electrical equipment |
| 135 | Advertising and market research |
| 103 | Programming and broadcasting activities |
| 131 | Legal and accounting activities |
| 37 | Manufacturing of wood and of product of wood and cork, except furniture; manufacture of articles of straw and plaiting materials |
| 101 | Publishing activities |
| 1 | Agriculture, forestry and fishing |
| 83 | Air transport |

| | |
|---|---|
| 184 | Sports activities and amusement and recreation activities |
| 20 | Activities of households as employers, undifferentiated goods- and services producing activities of households for own use |
| 4 | Electricity, gas, steam and air conditioning supply |
| 16 | Education |
| 13 | Fishing and aquaculture |
| 23 | Mining of metal ores |
| 13 | Professional, scientific and technical activities |
| 3020 | Manufacture of motor vehicles, trailers and semi-trailers |
| 3021 | Manufacture of other transport equipment |
| 5 | Water supply, sewerage, waste management and remediation activities |
| 24 | Other mining and quarrying |
| 14 | Administrative and support activities |
| 12 | Real estate activities |
| 11 | Financial and insurance activities |
| 134 | Scientific research and development |
| 15 | Public administration and defence, compulsory social security |
| 173 | Social work activities without accommodation |
| 22 | Extraction of crude petroleum and natural gas |
| 183 | Gambling and betting activities |
| 3014 | Manufacture of other non-metallic mineral products |
| 32 | Manufacturing of beverages |
| 2 | Mining and quarrying |

Source: LOGIS dataset.

## Appendix A3: HHI Example

Figure A3.1 HHI Simulations




Source: authors' calculations based on data presented in Section 2.

Figure A3.1 describes the HHI under three scenarios within a 10-supplier market. The far-left figure shows that the HHI will equal the 1/n when each supplier has a perfectly equal share of the market. The middle and right-hand figure show the impact on HHI when concentration decreases linearly and is concentrated in the top tertile. Under both, HHI begins deviating from the 1/n optimum and increases as concentration increases.