Tracking the price of almonds in Spain[1]

Francisco Martínez-Sánchez[2]
Universidad de Murcia

October 6, 2020

Abstract

I look for signs of manipulation of the price of almonds in the markets of Albacete, Murcia and Reus in Spain. To that end, I use Benford's Law to analyze the distribution of the second and third digits of the price of almonds. I find that prices in the Albacete and Reus markets deviate from Benford's Law, especially for the third digit, and those in the Murcia market deviate from Benford's Law for the second digit but not for the third digit. I therefore find signs that the price of almonds in the Spain's main markets is being manipulated.

JEL classification: C46; L13; L40; N54
Keywords: Benford's Law; Fraud detection; Price manipulation; Data quality

---

[1] I acknowledge financial support from the Spanish Ministerio de Ciencia, Innovación y Universidades and Spanish Agencia Estatal de Investigación under projects ECO2016-76178-P and PID2019-107192GB-I00, which are co-financed by FEDER funds. Any remaining errors are mine alone.

[2] Departamento de Métodos Cuantitativos para la Economía y la Empresa, Universidad de Murcia, 30100 Murcia, Spain, E-mail: fms@um.es

1. Introduction

Farmers in developed countries (e.g. France[3], Germany[4], Italy[5], Japan[6], Spain[7], the UK[8], and the USA[9]) complain about unfair global competition and supermarket chains lowering domestic food prices. In general, they complain about the pricing process, which is important in determining the investment decisions of farmers and the consumption decisions of buyers. It is a complex process that sets the price of many products, which can be substitutes or complements. Furthermore, it is delocalized into many local markets where local and foreign agents participate. This process could be corrupted by manipulations by the participants, which would cast doubt on the integrity of the price of foodstuffs.

In this paper, I look for signs of manipulation. Given the large quantity of agricultural products, I focus on markets for almonds. As can be seen in Table 1, the main almond producing country is the United States of America, which in 2018 produced 57.9% of the world's output. It is followed at some distance by Spain, which accounts for 10.5% of world output, and then by Iran with 4.3%. Ideally, this analysis would cover all markets, but in practice I look for signs of price manipulation in the main almond markets in Spain: The Reus market in the region of Catalonia, the Murcia market in the region of the same name, and the Albacete market in Castilla La Mancha.

Table 1: Production of Almonds (in shells) in 2018

Country	Output (tonnes)	Production (%)
USA	1,872,500	57.9
Spain	339,033	10.5
Iran	139,029	4.3
Morocco	117,270	3.6
Turkey	100,000	3.1
Italy	79,801	2.5
Australia	69,880	2.2
Tunisia	66,733	2.1
Algeria	57,213	1.8
China	49,879	1.5
World	**3,232,780**	**100**

Source: UN FAOSTAT[10]

To look for signs of price manipulation I use Benford's Law, which was reported first by Newcomb (1881) and then independently by Benford (1938). Although this law originally analyzed the distribution of the first digit of a set of numbers, I use extensions of it that describe the distribution of the second and third digits (Berger and Hill (2011)) because prices seem somewhat sticky over time, so there are numbers that do not appear as the first digit in the set of numbers considered here. This method is used by Abrantes-Metz et al. (2011) to monitor the dynamic integrity of the daily Libor from 2005 to 2008, although it only considers the

[3]See https://www.canadianbusiness.com/business-news/french-farmers-protest-stagnant-revenues-unfair-competition/and https://www.france24.com/en/20150720-france-meat-industry-protest-prices-livestock
[4]https://www.dw.com/en/germany-thousands-of-farmers-protest-in-berlin/a-52039202
[5]https://www.bbc.com/news/av/world-europe-47199286
[6]https://www.reuters.com/article/japan-trade-farmers-idAFL3E7LQ06Y20111026
[7]https://english.elpais.com/economy_and_business/2020-02-05/hundreds-of-farmers-rally-in-madrid-to-demand-fairer-prices.html
[8]https://www.bbc.com/news/uk-33777075
[9]https://time.com/5736789/small-american-farmers-debt-crisis-extinction/
[10] "Almonds (in shells) production in 2018, Crops/Regions/World list/Production Quantity (pick lists)". UN Food and Agriculture Organization, Corporate Statistical Database (FAOSTAT). 2019. Retrieved 13 August 2020.

distribution of the second digit. They find that the Libor deviates significantly from Benford's Law, raising suspicions of Libor manipulation and collusion between banks.

Benford's Law emerges in many fields, including economic systems, as demonstrated by Villas-Boas et al. (2017). For example, the series of 1-day returns in the Dow-Jones Industrial Average Index and Standard and Poor's Index reasonably follow Benford's Law, as shown by Ley (1996). Its use to detect fraud in economic data was suggested by Varian (1972), but Nigrini (1996) shows that it can be used to detect fraud in company accounting. This has given rise to a wave of studies that use Benford's Law to detect manipulations in economic data[11] of all types, particularly studies that apply it to investigate the quality of macroeconomic data. For example, González-García and Pastor (2009) find that failure to comply with Benford's Law should not be interpreted as a reliable indication of poor quality macroeconomic data. However, looking at the quality of the macroeconomic data relevant to the deficit criteria reported to Eurostat by EU member states, Rauch et al. (2011) find that the data reported by Greece show the largest deviation from Benford's Law of any Euro state. Michalski and Stoltz (2013) find that countries sometimes misreport their economic data strategically. China's economic data has also been assessed. Holz (2014) assesses the quality of China's GDP statistics and finds that official GDP data show few statistical anomalies according to Benford's Law; Huang et al. (2020) assess the quality of Chinese industrial census data and show that the Benford's Law method is effective in uncovering irregularities. There are, of course, other important uses of this statistical law: For example as a tool for selecting audits as in Watrin et al. (2008). However, auditors must be cautious to ensure that Benford’s Law can be expected to apply to unmanipulated data on the prospective audit target. Ausloos et al. (2017) show that there is manipulation of income reports in various regions and municipalities of Italy and that some accountants are aware of Benford's Law and can avoid non-compliance at individual level but not at aggregate city level. Kaiser (2019) assesses the quality of six widely used survey datasets and finds that income generally obeys Benford's Law, although almost all datasets show substantial discrepancies. Barabesi et al. (2018) suggest a new way of testing Benford's Law that is particularly attractive for detecting fraud in customs data collected from international trade. Benford’s Law has also been used as an indicator of fraud in economic research: Günnel and Tödter (2009) show that economic research is broadly consistent with it, but there are doubts about forecast information. Tödter (2009) shows, in a sample of published articles, that a large proportion deviate from Benford’s Law for the first digit, but not for the second.

In this paper I look for signs of manipulation of the price of almonds in the markets of Albacete, Murcia and Reus in Spain. To that end, I use Benford's Law to analyze the distribution of the second and third digits of the price of almonds because this law arises in self-organizing dynamic economic behavior systems (Villas-Boas et al. (2017)). I find that prices in the Albacete and Reus markets deviate from it, especially for the third digit, and those in the Murcia market deviate for the second digit but not for the third digit. I therefore find signs of manipulation of the price of almonds in the main markets of Spain.

The rest of the paper is organized as follows: Section 2 describes Benford’s Law. Section 3 presents and describes the dataset. Section 4 analyses the dataset and reports the main results. Section 5 concludes.

2. Benford’s Law

Seeking to check for manipulation in the price of the varieties of almonds listed in the markets of Albacete, Murcia and Reus, I use Benford's Law, which was first reported by Newcomb (1881) and later independently by Benford (1938). This law describes the distribution of the first digit of a set of numbers. Thus, digit 1 appears more frequently (30.1%), digit 2 less frequently (17.6%), and the frequency decreases as the value of the first digit increases. However, there are numbers that do not appear as the first digit in the price of almonds. This is because the prices of different varieties of almonds have not changed much.

[11] See Nigrini (2012) for more information on Benford's Law.

To overcome this drawback, I use extensions of Benford's Law that describe the distribution of the second and third digits (Berger and Hill (2011)), given by

$$P(D_1 = d_1) = log_{10}\left(1 + \frac{1}{d_1}\right),\ \text{d}_1 = 1, 2, \ldots, 9, \qquad (1)$$

$$P(D_2 = d_2) = \sum_{i=1}^{9} log_{10}\left(1 + \frac{1}{10i + d_2}\right),\ \text{d}_2 = 0, 1, 2, \ldots, 9, \qquad (2)$$

$$P(D_3 = d_3) = \sum_{j=1}^{9}\sum_{i=0}^{9} log_{10}\left(1 + \frac{1}{100j + 10i + d_3}\right),\ \text{d}_3 = 0, 1, 2, \ldots, 9, \qquad (3)$$

where D_1 represents the first digit, D_2 the second digit, and D_3 the third digit. As can be seen in Figure 1, the digit distribution becomes more uniform as the order of the digit increases.

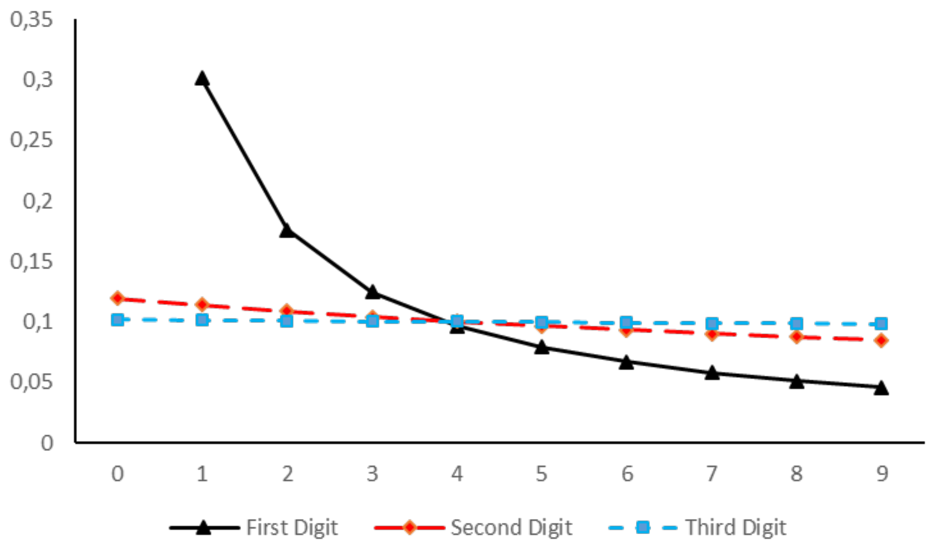

Figure 1: Benford's Law

To check whether an empirical distribution follows Benford's Law I use the Pearson's Chi-square test, which uses χ_D^2 as the statistic for the second and third digits:

$$\chi_D^2 = \sum_{i=0}^{9} \frac{(O_i - E_i)^2}{E_i} \sim \chi_9^2 \qquad (4)$$

where O_i is the observed frequency of (second or third) digit i = 0, 1..., 9 and E_i is the expected frequency of (second or third) digit i according to Benford's Law. High values of the statistic thus indicate large deviations between the observed and expected frequencies. Notice that χ_D^2 follows a Chi-square distribution with 9 degrees of freedom. Thus, the critical values are 14.68 for the 10% significance level, 16.92 for the 5% level and 21.67 for the 1% level.

3. Data

This paper draws on several data sources. Murcia market prices are obtained directly from the market itself. Reus market prices are also obtained directly from the market, but through its website.[12] Finally, Albacete market prices are obtained from the Spanish website www.precioalmendra.es.[13]

Before analyzing the distribution of the second and third digits of the price of almonds, I briefly describe the trend in the prices of almonds in the markets of Albacete, Murcia, and Reus from 2016 to 2019. I focus mainly on the almond varieties listed in all three markets, i.e. Commune, Largueta, Marcona, and Ecological. As shown in Figure 2, the trend in these prices is simila in all three markets, although there are

[12] http://www.llotjadereus.org/?go=e6598a7e63ddfde8a9557f334d3f9063f1a92bcf0018cd07976fe59ff9f4f7d6fe76da618ed9577903f00115a92e55ae73c0e76916090a1f

[13] In particular from the website: https://www.precioalmendra.es/lonja-albacete/

some differences. I also note that each variety of almond is normally priced lowest in Reus and highest in Albacete.

Table 2 shows that the Albacete market offers the highest minimum prices and the Reus market the lowest. The Reus market also offers the lowest maximum prices and the Murcia market the highest, except for the Marcona variety. Note that the Albacete market does not provide almond prices for the full year in 2016, as shown in Figure 2. If it had done so, the Albacete market might have provided the highest maximum prices for all varieties given the trend of prices in all three markets (see Figure 2).

Table 2: Descriptive statistics for common varieties

Variety	Commune			Largueta		
Market	Albacete	Murcia	Reus	Albacete	Murcia	Reus
Min.	4.35	4.16	3.9	4.9	4.62	4.4
Q1	4.763	4.6	4.25	5.2	5.06	4.7
Median	5.175	5.01	4.7	5.6	5.575	5.025
Mean	5.175	5.212	4.862	6.044	6.069	5.812
Q3	5.55	5.655	5.35	7.2	7.143	7.125
Max.	6.3	7.69	7.2	7.95	8.21	7.8
SD.	0.503	0.824	0.784	1.017	1.144	1.262
Obs.	162	208	172	156	208	180
Variety	Marcona			Ecological		
Min.	5.00	4.830	4.600	6.450	6.280	5.750
Q1	5.388	5.455	4.900	6.750	6.600	6.100
Median	6.225	6.270	5.950	7.200	6.930	6.300
Mean	6.671	6.810	6.438	7.305	7.377	6.425
Q3	8.200	8.303	8.100	7.850	7.960	6.450
Max.	9.200	8.950	8.800	8.750	9.940	8.300
SD.	1.430	1.415	1.504	0.622	0.971	0.598
Obs.	156	208	182	162	208	125

To complete the information on all the varieties of almonds considered, Table 3 gives descriptive statistics for local varieties in each market. On average, these varieties are cheaper than the varieties marketed in all three markets.

Table 3: Descriptive statistics for varieties specific to each market

Market	Albacete	Murcia				Reus		
Variety	Guara	Guara	Ferragnes	Garrigues	Ramillete	Mollar	Mallorca	Pelona
Min.	4.450	4.320	4.290	4.380	4.410	3.800	3.800	2.800
Q1	4.900	4.760	4.838	4.787	4.890	4.200	4.188	3.000
Median	5.275	5.100	5.255	5.165	5.270	4.600	4.600	3.400
Mean	5.293	5.337	5.413	5.394	5.469	4.769	4.763	3.471
Q3	5.688	5.803	5.870	5.855	5.893	5.250	5.250	3.800
Max.	6.450	7.820	7.840	7.820	7.820	7.350	7.350	5.300
SD.	0.521	0.820	0.787	0.800	0.780	0.811	0.809	0.457
Obs.	162	208	208	208	208	171	172	180

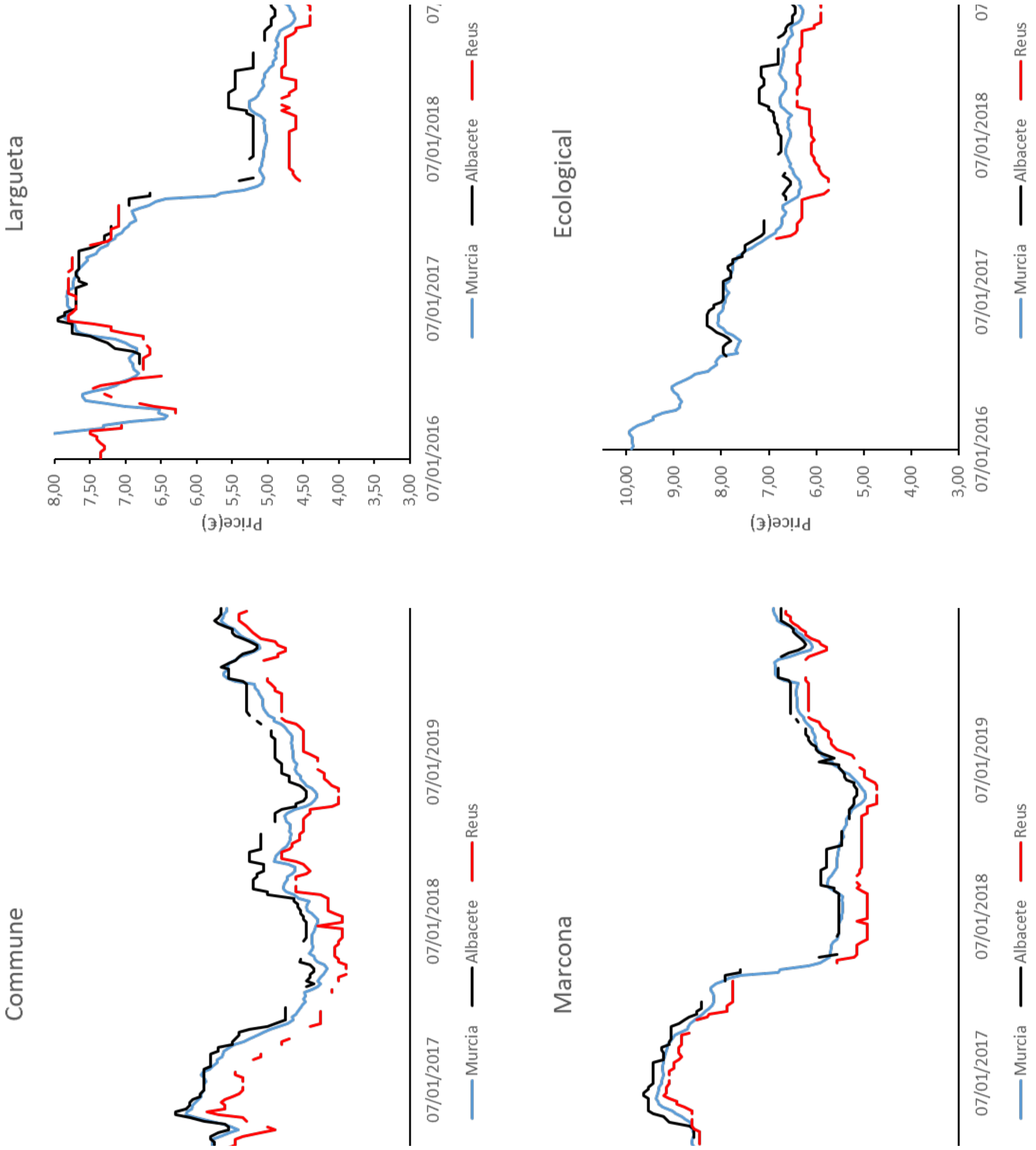

Figure 2: Price of almonds

4. Results

This section compares the empirical distributions of the second and third digits of the price of almonds with the corresponding version of Benford's Law. The process runs from the general to the particular, so I first analyze each market as a whole, then distinguish between common and local almond varieties in each market. Finally, I analyze each common variety by market.

4.1. Analysis for each market as a whole

Figure 3 reveals that Benford's Law for the third digit is not complied with in the Albacete and Reus markets. This is because the third digit is the last digit of the price of almonds. At least 97% of prices thus end in zero or five, so Benford's law is a long way from being met. However, prices in the Murcia market comply with Benford's Law for the third digit, as can be seen in Figure 3.

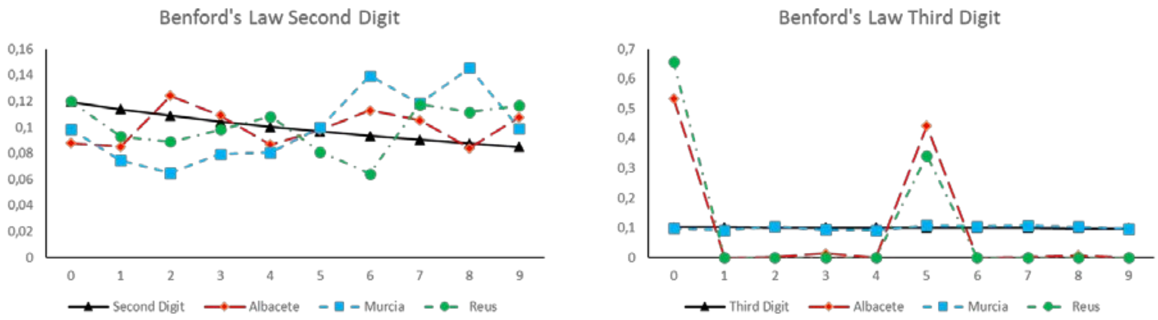

Figure 3: Benford's Law

For the second digit, all three markets deviate from Benford's Law, but by a lesser extent than for the third digit: The exception is the Murcia market, where there is no deviation for the third digit and the deviations for the second digit are greater than those in Albacete and Reus. I also observe a predominance of the high values of the second digit in the Murcia and Reus markets with respect to Benford's Law (see Figure 3). However, in the Albacete market two, six, seven and nine predominate as the second digit.

By applying Pearson's Chi-square test to this dataset (see Table 4), I confirm the conclusions obtained from Figure 3 and find that the price patterns are similar in the Albacete and Reus markets. In those markets the degrees of noncompliance with Benford's Law for the second and third digits are similar because the statistical values for the second digit are relatively close and those for the third digit are extremely high. However, the deviation from Benford's Law for the second digit is greater in the Murcia market

Table 4: Chi-square test statistic values

Market	χ^2_D second digit	χ^2_D third digit	Observations
Albacete	26.148***	3003.1***	798
Murcia	194.78***	7.175	1664
Reus	55.229***	5226.8***	1182

*p-value < 0.10, **p-value < 0.05, ***p-value < 0.01

As mentioned in the previous section, there are many different varieties of almonds and not all of them are sold on all three markets, which could explain the deviations from Benford's Law. This would mean that there are no indications of price manipulation. To overcome this drawback, in the following subsection I extend the analysis to distinguish between the varieties of almonds which are sold on all three markets and those which are not.

4.2. Analysis distinguishing between varieties which are present on all three markets and this which are not

The varieties that appear on all three markets are Commune, Largueta, Marcona and Ecological, which I refer to here as the "common varieties". The rest of the varieties are specific to each market, although the Guara variety is found in both Albacete and Murcia.

As Table 5 shows, I draw the same conclusions as in the previous subsection: i) Prices in the Albacete and Reus markets deviate from Benford's Law for the second and third digits but the deviations for the third digit are very high; and ii) prices in the Murcia market deviate from Benford's Law for the second digit but not for the third digit.

Table 5: Chi-square test statistic values

Market	Albacete		Murcia		Reus	
	Common	Non-common	Common	Non-common	Common	Non-common
χ_D^2 second digit	32.701***	17.841**	64.102***	161.23***	45.881***	35.15***
χ_D^2 third digit	2438.2***	575.22***	3.9791	12.865	2824***	2420.7***
Observations	636	162	832	832	659	523

*p-value < 0.10, **p-value < 0.05, ***p-value < 0.01

It might be thought that the varieties listed in all three markets would be under more competition, and that there would thus be less price manipulation. This reasoning holds only for the Murcia market, where the deviations from Benford's Law are greater in non-common or local varieties (high statistics). However, it does not hold for Albacete and Reus, though in Reus there are no great differences in the deviations between the common and non-common (or local) varieties. In the Albacete market there is only one non-common variety, Guara, which is also found on the Murcia market. These two markets are only 150 km apart, so in Albacete there is no local variety of almonds and the results obtained for this market are not solid.

4.3. Analysis for each variety in each market

Table 6 shows the results of the analysis for each variety in each market. I confirm the results obtained in the previous subsections: i) Prices in the Albacete and Reus markets deviate from Benford's Law, especially for the third digit; and ii) prices in the Murcia market deviate from Benford's Law for the second digit but not for the third. Moreover, I find that for the second digit Benford's Law is not fulfilled anywhere except in the Reus market for the Commune, Mollar, and Mallorca varieties. For the third digit Benford's Law is only fulfilled in the Murcia market for all its varieties.

Table 6: Chi-square test statistic values

Market	Variety	χ_D^2 second digit	χ_D^2 third digit	Observations
Albacete	Commune	29.089***	595.37***	162
	Largueta	55.91***	620.19***	156
	Marcona	56.325***	651.42***	156
	Ecological	43.065***	593.46***	162
	Guara	17.841**	575.22***	162
Murcia	Commune	52.437***	11.675	208
	Largueta	65.401***	10.311	208
	Marcona	28.543***	6.2803	208
	Ecological	73.347***	9.5644	208
	Guara	59.812***	14.213	208
	Ramillete	51.001***	7.1891	208
	Garrigues	60.362***	12.741	208
	Ferragnes	27.471***	14.607	208
Reus	Commune	8.0704	743.12***	172
	Largueta	157.65***	788.14***	180
	Marcona	78.795***	774.15***	182
	Ecological	65.433***	521.47***	125
	Mollar	9.9957	713.5***	171
	Mallorca	12.214	715.22***	172
	Pelona	27.609***	1124.7***	180

*p-value < 0.10, **p-value < 0.05, ***p-value < 0.01

Comparing the three markets, I find that the biggest deviations from Benford's Law for the second digit are in the Murcia market in all cases except the Largueta and Marcona varieties, for which the biggest differences are in Reus.

5. Conclusions

In this paper, I look for signs of manipulation of the price of almonds in the markets of Albacete, Murcia, and Reus in Spain. To that end, I use Benford's Law to analyze the distribution of the second and third digits of the price of almonds because this law arises in self-organizing dynamic economic behavior systems (Villas-Boas et al. (2017)).

I find that prices in the Albacete and Reus markets deviate from Benford's Law, especially for the third digit, and prices in the Murcia market deviate from Benford's Law for the second digit but not for the third. I also find that for the second digit Benford's Law is not fulfilled in any market, except in Reus for the cases of the Commune, Mollar and Mallorca varieties. However, Benford's Law for the third digit is only fulfilled in the Murcia market for all its varieties. I therefore find signs of manipulation of the price of almonds in the main markets of Spain.

6. References

Abrantes-Metz, R.M., Villas-Boas, S.B. and Judge, G. (2011). Tracking the libor rate. *Applied Economics Letters* 18(10): 893–9.

Ausloos, M., Cerqueti, R. and Mir, T.A. (2017). Data science for assessing possible tax income manipulation: The case of Italy. *Chaos, Solitons & Fractals*. 104: 238–256. https://doi.org/10.1016/j.chaos.2017.08.012

Barabesi, L., Cerasa, A., Cerioli, A. and Perrotta, D. (2018). Goodness-of-Fit Testing for the Newcomb-Benford law with application to the detection of customs fraud.*Journal of Business & Economic Statistics*. 36(2): 346-358. DOI: 10.1080/07350015.2016.1172014

Benford, F. (1938) The law of anomalous numbers, *Proceedings of the American Philosophical Society*. 78 (4): 551—572.

Berger, A, Hill, T.P. (2011). A basic theory of Benford's Law. *Probability Surveys*. 8, 1—126.

Gonzalez-Garcia, J., and Pastor, G. (2009). Benford's Law and MacroeconomicData Quality.*International Monetary Fund workingpaper WP/09/10*.

Günnel, S. and Tödter, K.-H. (2009). Does Benford's Law hold in economic research and forecasting? *Empirica*. 36: 273–292.

Holz, C.A. (2014). The quality of China's GDP statistics. *China Economic Review*.30. 309–338.

Huang, Y., Niu, Z. and Yang, C. (2020). Testing firm-level data quality in China against Benford's Law.*Economics Letters*. 192: 1—4.

Kaiser, M. (2019) Benford's Law as an indicator of survey reliability—can we trust our data? *Journal of Economic Surveys*. 33 (5): 1602–1618.

Ley, E. (1996). On the peculiar distribution of the U.S. stock indexes' digits.*The American Statistician*, 50(4): 311-313.

Michalski, T. and Stoltz, G. (2013). Do Countries Falsify Economic Data Strategically? SomeEvidence That They Might.*Review of Economics and Statistics*. 95(2):591—616.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*. 4: 39—40.

Nigrini, M. (1996). A taxpayer compliance application of Benford's Law. *The Journal of the American Taxation Association*. 18(1): 72—91.

Nigrini, M. (2012).*Benford's law: Applications for forensic accounting, auditing, and fraud detection* (Vol. 586). Hoboken, NJ: John Wiley & Sons.

Rauch, B., M.Göttsche, G.Brähler and S. Engel. (2011). Fact and Fiction in EU-GovernmentalEconomic Data. *German Economic Review*. 12(3): 243–255.

Tödter, K.-H. (2009). Benford's Law as an Indicator of Fraud in Economics. *German Economic Review*. 10(3): 339–351.

Varian, H. (1972).Benford's Law (Letters to the Editor). *The American Statistician*. 26(3): 65—66.

Villas-Boas, S.B., Fu, Q. and Judge, G. (2017).Benford's law and the FSD distribution of economic behavioral micro data. *Physica A: Statistical Mechanics and its Applications*. 486: 711–719.

Watrin, C., R. Struffert, and R. Ullmann, (2008), Benford's law: An instrument for selecting tax audit targets? *Review of Managerial Science*. 2: 219–237.