ARTICLE IN PRESS

JID: APM [m3Gsc;June 14, 2016;18:18]

Applied Mathematical Modelling 000 (2016) 1–15

Contents lists available at ScienceDirect

Applied Mathematical Modelling

journal homepage: www.elsevier.com/locate/apm

A probabilistic method for forensic cost estimating of infrastructure projects

Regis Signor [a], Peter E.D. Love [b,*], Oluwole Olatunji [c], Fernanda F. Marchiori [d], William G. Gripp [e]

[a] *Federal Police Department, Florianópolis, SC, Brazil*
[b] *Department of Civil Engineering, Curtin University, GPO Box U1987, GPO Box U1985, Perth, WA, Australia*
[c] *Department of Construction Management, Curtin University, GPO Box U1987,GPO Box U1985, Perth, WA, Australia*
[d] *Federal University of Santa Catarina, Florianópolis, SC, Brazil*
[e] *Federal Police Department, Londrina, PR, Brazil*

A R T I C L E I N F O

Article history:
Received 3 October 2015
Revised 2 May 2016
Accepted 17 May 2016
Available online xxx

Keywords:
Brazil
Corruption
Infrastructure
Forensics cost estimating
Overpricing
Probability distribution

A B S T R A C T

To determine the existence of overpricing in infrastructure projects a deterministic approach has been typically used by the Brazilian Federal Police. Yet, Judges have often found it too difficult to determine if overpricing had occurred as contractors have tended to claim their projects were 'unique' and therefore subjected to price increases. Consequently, Judges have had to consider the likelihood of an error occurring, which has more often led to application of the *in dubio pro reo* principle. To address the pervasive issue of overpricing in Brazilian infrastructure projects, a novel and robust probabilistic method that utilizes 'distribution fitting' using empirical data to forensically determine its occurrence is presented. Considering the limited research that has examined forensic cost estimating in construction, it is promulgated that the proposed approach can be used to support the legal fraternities worldwide to obtain a criminal conviction for overpricing of public infrastructure projects.

1. Introduction

Corruption is the abuse of bestowed power or position to acquire a personal benefit. It includes activities such as overpricing, extortion, bribery, theft, and embezzlement [1]. According to Kenny [2] the construction industry is the most corrupt sector around the world, particularly in the case of public infrastructure procurement where large payments to gain or alter contracts and circumvent regulations are common. Consequences of corruption is far-reaching: it can result in poor economic returns, reduced funding for maintenance and poor quality construction, and can potentially jeopardize work's safety. The bigger spoils are imposed to the most vulnerable people, who have no access to welfare and often to basic services as health and safety [3].

Addressing corruption in construction is a pervasive challenge and on-going problem; it is a function of human nature and thus can only be minimized. It is, however, a criminal act and therefore the consequences of being found guilty of

* Corresponding author. Tel.: +61 864601272.
E-mail addresses: regis.rs@dpf.gov.br (R. Signor), p.love@curtin.edu.au, plove@iinet.net.au (P.E.D. Love), oluwole.olatunji@curtin.edu.au (O. Olatunji), fernanda.marchiori@ufsc.br (F.F. Marchiori), william.wgg@dpf.gov.br (W.G. Gripp).

http://dx.doi.org/10.1016/j.apm.2016.05.025
0307-904X/

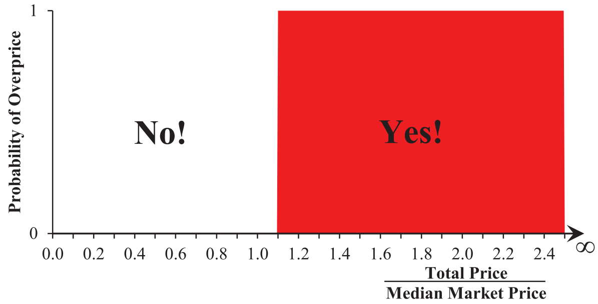

Fig. 1. Deterministic approach.

corruption can be severe (e.g., imprisonment, fines, and loss of reputation). In the context of the construction industry, it is important that authorities are equipped with the appropriate investigative techniques to address corruption [4].

In Brazil, for example, corruption is an endemic problem within the construction industry, particularly in the case of public works. Like many governments worldwide, Brazilian law requires firms to tender for the work required to delivery infrastructure projects. Yet, during the bidding it has been revealed by the Brazilian Federal Police that contractors are able to circumvent this process and overprice their contracts [5–7]. However, it is often difficult to detect and prove how overpricing has occurred; as a result to obtain a criminal conviction has been a problematic issue for the Brazilian Federal Police.

This paper builds upon the deterministic approach that has been used and reported in Signor et al. [8], and presents a novel and robust probabilistic alternative to forensically determine if overpricing has occurred in infrastructure projects. Considering the limited research that has examined forensic cost estimating in construction, it is promulgated that the proposed approach can be used to support the legal fraternities worldwide to obtain a criminal conviction for overpricing of public infrastructure projects. There have been a limited number of studies that have specifically examined nature of overpricing in infrastructure projects (e.g., [9–12]), though an extensive review of the guises of corruption can be found in Sohail and Cavill [13]. Moreover, specific details of corruption in the Brazilian construction are presented in Lopes [6].

In the next section of this paper, the approach that had been traditionally used by the Brazilian Federal Police to determine if overpricing was occurring in infrastructure projects is examined and the case for a probabilistic approach based on the use of 'distribution fitting' is then made. This novel approach is being developed by the Brazilian Police Forensic Experts through extensive workshops and will form an integral part of its arsenal of tools that can be applied to obtain a criminal conviction.

2. Moving beyond a deterministic approach

Under the existing system, experts that integrate the Brazilian legal system are required to produce court evidence to identify that overpricing has occurred. Several forms of assessment can be used such as a simple parametric method, which estimates a unitary price for each project depending on its type and characteristics. While this method may provide an initial indication that overpricing has occurred, the unitary values that are derived are not considered as valid evidence in court as they cannot be used to prove the defendants' misconducts. Thus, experts are required to re-calculate the specific point where overpricing occurred taking into account externalities that could have influenced the prices to avoid incorrect decisions being made. Accordingly, experts calculate the price and quantities of materials and labor using the median market values derived from the *National System of Costs Survey and Indexes of Construction* (SINAPI), which has been indoctrinated into Brazilian Law. The system is maintained by the Brazilian government's financing agency, Caixa Econômica Federal (CAIXA).

The law established the SINAPI's median as a clear and rigid parameter. However, contractors' defendants often argue that their client's project was unique and there were externalities not considered by SINAPI, suggesting that despite the work's price being above the legal limit a conviction would be unfair. Judges have limited engineering knowledge and hence reasonable doubt is created (Poder [14]). Acknowledging this problem forensic experts have established an agreed margin of 10% over and above SINAPI's median price (total budget) due to the variability that may exist in a market (Fig. 1). Notably, for a single item such as cement, forensic experts may allow a 30% increase over the median.

The deterministic approach provides a basis for identifying overpricing, although it does not consider underpricing. In addition, defendants can question its validity and reliability essentially because it lacks scientific rigor [8]. The deterministic approach's reliability is dependent on the *Law of Large Numbers* (LLN) (i.e. as a sample size grows, its estimate will converge to the true value). While this may be appropriate for items that are well balanced within a given budget, it is unsuitable for those that are not. For example, if a contractor's bid for a school project is 30% lower than SINAPI's median, then all legal requirements are deemed to have been met and a profit may still be obtained. Such conclusions are based on the assumption that there is a high degree of variability for quantities and unit prices available from suppliers. However, if the same contractor is hired to construct a metal structure 30% lower than SINAPI's median, then there is a likelihood that he will lose money or fail to meet his contractual obligations; the quantities (i.e., productivity) of this highly industrialized

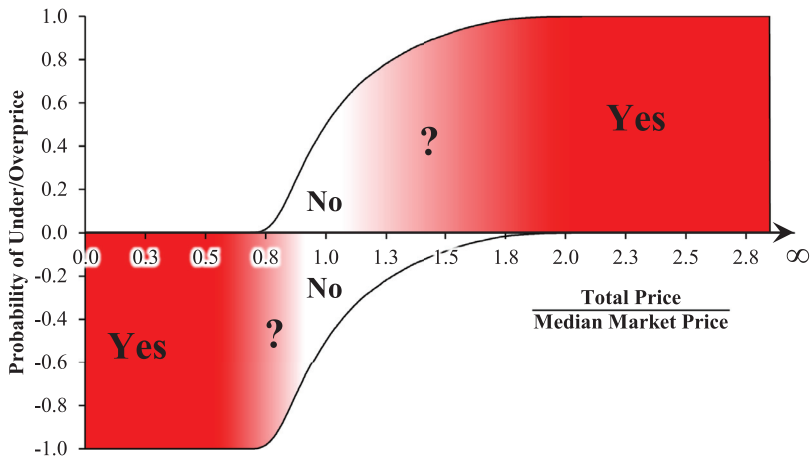

Fig. 2. Proposed probabilistic approach.

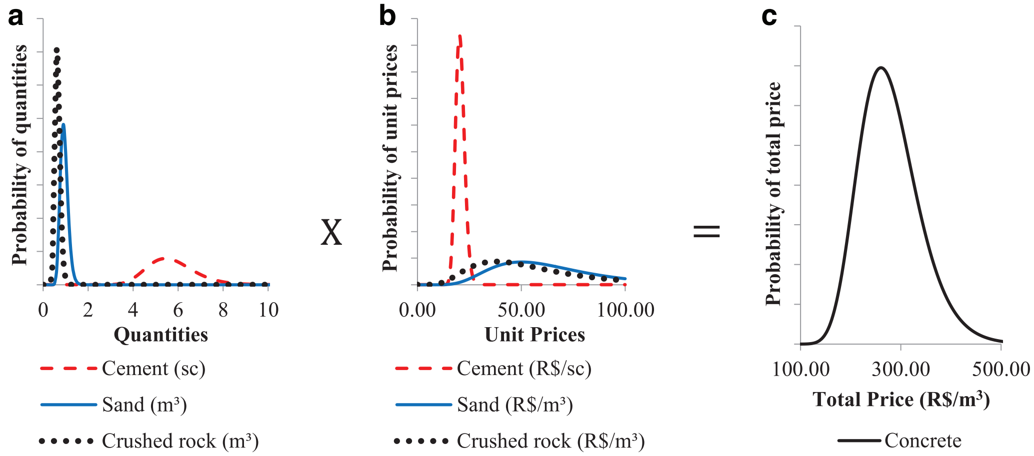

Fig. 3. Probability density function for quantities, unit prices and the total price for concrete.

process are rigid and there is a limited number of suppliers for this specific item and the price of steel fluctuates slightly, in accordance with demand in world markets. With this in mind, to improve the reliability of SINAPI there is a need to move from a deterministic to a probabilistic approach to accommodate its inherent limitations. The proposed probabilistic approach is presented in Fig. 2.

3. Propagation of a probabilistic method

To develop a probabilistic model for a budget of an infrastructure project, it is necessary to consider the variability that may exist for quantity and unit price of resources. Consequently, the variation in supplies' quantities and their unit prices need to be examined by computing their individual Probability Distributions and combining them to create a total price such as for concrete presented in Fig. 3 – it shows that the probabilistic cost of the concrete (Fig. 3c) can be assessed by multiplying the probabilistic quantities of its resources (only the raw materials are represented in Fig. 3a) by their probabilistic unit prices (Fig. 3b). Determining the distributions for quantities and unit prices enables forensic experts to ascertain the frequency of occurrence of the magnitude of total prices at specific intervals.

Using the SINAPI database, which was provided by CAIXA, the construction of 1m^2 of concrete block masonry wall with net area smaller than 6m^2, without openings, is examined as example for São Paulo in March 2015. The individual elements that are required to construct a masonry wall are identified in Table 1 as well as their quantities and unit prices. The probabilistic approach that will be presented in this paper can be extrapolated to an entire project, but for the purposes of brevity it is confined to demonstrating the occurrence of overpricing for the construction of this wall only.

The raw data for the bricklayer and unskilled worker productivity rates, blocks' wastage and the quantity of mortar required to construct a masonry wall was provided by a worksite survey hired by CAIXA; their extreme values, median and quartiles were derived and are presented in Table 2. Information about the quantity of welded steel mesh and steel pin

Table 1
Quantities and unit prices for masonry according SINAPI.

SINAPI's code	Element	Unit	Quantity	Unit prices Q1	Median	Q3
87,292	Laying mortar 1:2:8	m^3	0.0103000	288.15	317.87	344.23
88,309	Bricklayer with social charges	h	0.9900000	17.84	18.12	19.27
88,316	Unskilled worker with social charges	h	0.4950000	15.40	15.65	16.52
651	Concrete hollow block $14 \times 19 \times 39\,cm^3$	unit	13.5000000	1.57	1.76	2.07
34,547	Welded steel mesh	m	0.7850000	1.70	2.12	2.60
37,395	Steel pin 27 mm	unit	1.8900000	0.45	0.46	0.55
Totals (R$/$m^2$ – São Paulo, March/2015)				51.63	55.25	61.83

Table 2
Descriptive statistics of the surveyed values for labor, wastage and mortar.

Descriptive statistics	Bricklayer (h/m^2)	Unskilled worker (h/$h_{bricklayer}$)	Blocks waste (%)	Mortar (l/m)
N	30	6	45	29
Minimum	0.12	0.25	0.30	0.87
Q1	0.29	0.44	1.85	1.28
Median	0.53	0.54	4.70	1.55
Q3	1.38	0.87	8.85	1.98
Maximum	3.39	1.00	39.50	5.62

Table 3
Descriptive statistics of the surveyed values for the unit prices for the city of São Paulo (March 2015).

Descriptive statistics	Medium sand (m^3)	Hydrated lime (kg)	Portland cement (bag 50 kg)	Concrete block 9 × 19 × 39 (thousand)	Steel bar 16 mm (kg)	Mixer 400 l (unit)	Grubbing hoe (unit)	Leather work gloves (pair)	Bricklayer (h)	Unskilled worker (h)
N	10	16	9	9	8	10	5	11	6	6
Minimum	35.00	0.33	20.90	1200.00	3.00	2400.00	12.20	6.44	6.33	5.21
Q1	38.75	0.39	21.90	1250.00	3.17	2899.00	12.90	7.34	6.33	5.21
Median	50.25	0.46	22.60	1400.00	3.79	3088.00	16.28	8.90	6.34	5.21
Q3	59.25	0.47	26.39	1650.00	4.20	3350.00	25.69	10.80	7.41	5.43
Maximum	73.00	0.53	28.00	1665.19	4.61	4800.00	26.62	19.00	9.60	5.44

27 mm required are not available. Thus, for the purpose of this example fixed values are assigned with no variation and wastage expected. Similarly, the quantities for the raw components of mortar are also fixed values, as the raw data are still unavailable. The Brazilian Institute of Geography and Statistics (IBGE) provided raw data for unit prices collected in São Paulo, in March of 2015 – the descriptive statistics are shown in Table 3. This data accommodates the range of materials, types of labor and equipment required to construct the masonry wall.

3.1. Determining the 'best fit' probability distribution

To ascertain the 'best fit' probability distribution from the empirical distributions of quantities and unit prices shown in Tables 2 and 3, the procedure advocated by Love et al. [15] and Love and Sing [16] is adopted. For the purpose of brevity the raw data is not presented in the tables, however, it is available upon request. Descriptive statistics such as the extreme values, median and quartiles were calculated for resource quantities and unit prices. The Probability and Cumulative Density Functions (PDF and CDF) for each variable were computed using the software *EasyFit 5*. The 'best fit' distributions were then determined using the following 'Goodness of Fit' tests, which measure the compatibility of a random sample with a theoretical probability distribution: *Kolmogorov-Smirnov statistic (D); Anderson-Darling statistic (A^2); Chi-squared statistic* (χ^2).

The 'Goodness of Fit' tests were used to test the null (H_0) and alternative hypotheses (H_1) that the datasets: H_0 – follow the specified distribution; and H_1 – do not follow the specified distribution. The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the statistics D, A^2 and χ^2 are greater than the critical value. For the purposes of this research, a 0.05 significance level was used to evaluate the null hypothesis.

Once the 'best fit' distributions for each variable were identified, their compatibility was examined to determine if they could be practically applied to material quantities and unit prices. For example, when distributions with negative values were identified they were discarded and attention was given to Kolmogorov-Smirnov statistic to determine an alternative. After the definition of the probabilistic distributions, which describe the quantities and the unit prices of the elements of the masonry wall, it is possible to use the *Monte Carlo Method* to assess the probabilistic price of the masonry wall itself. To achieve this objective, 50,000 values were generated at random for resource quantities and their unit price, using the 'best

Table 4
'Best fit' distributions for quantities and unit prices.

Type	Element	Distribution and parameters	Goodness of fit
Quantities	Bricklayer (h/m^2)	Frechet (α=1.2259 β=0.39283 γ=0)	D=0.10309 A^2=0.37268 χ^2=0.16299
	Unskilled worker (h/h$_{bricklayer}$)	Wakeby (α=3.1509 β=7.005 γ=0.18061 δ=0.15886 ξ=0)	D=0.18589 A^2=0.27085 χ^2=*
	Blocks waste (%)	Exponential (λ 0.15121 γ=0)	D=0.09361 A^2=0.45181 χ^2=2.5272
	Mortar (L/m)	Burr (4P) (k=0.46811 α=4.3035 β=0.76214 γ=0.55112)	D=0.1006 A^2=0.42908 χ^2=0.55395
Unit prices	Medium sand (m^3)	Rice (ν=48.968 σ=11.547)	D=0.111348 A^2=0.22852 χ^2=0.08647
	Hydrated lime (kg)	Burr (k=3.6597E+5 α=6.9012 β=2.2877 γ=0.10456)	D=0.1246 A^2=0.30578 χ^2=0.08108
	Portland cement (bag 50 kg)	Burr (k=0.14149 α=63.78 β=21.302 γ=0)	D=0.13999 A^2=0.24686 χ^2=*
	Concrete block 9 × 19 × 39 (thousand)	Gen Pareto (k=−0.87606 σ=586.14 μ=1122.6)	D=0.16336 A^2=0.39098 χ^2=*
	Steel bar 16 mm (kg)	Gen extreme value (k=−0.30158 σ=0.61251 μ=3.5525)	D=0.12183 A^2=0.17397 χ^2=*
	Mixer 400 l (unit)	Burr (k=0.44981 α=15.844 β=2848.0 γ=0)	D=0.20802 A^2=0.39843 χ^2=1.84 × 10^{-4}
	Grubbing hoe (unit)	Gen Pareto (k=−0.42319 σ=13.795 μ=8.9952)	D=0.21694 A^2=0.30979 χ^2=*
	Leather work gloves (pair)	Log-Logistic (3P) (α=1.6723 β=2.5701 γ=6.1992)	D=0.1079 A^2=0.20427 χ^2=0.02307
	Bricklayer (h)	Gen Pareto (k=0.9131 σ=0.05482 μ=6.3042)	D=0.32369 A^2=0.96975 χ^2=*
	Unskilled worker (h)	Gen Pareto (k=0.38235 σ=0.05928 μ=5.1874)	D=0.36674 A^2=1.0284 χ^2=*

*Some Chi-squared statistics were not performed due to small number of available data.

Table 5
Descriptive statistics for the random values generated for the quantities (N=50,000).

Descriptive statistics	Bricklayer (h/m^2)	Unskilled worker (h/h$_{bricklayer}$)	Blocks waste (%)	Mortar (l/m)
Minimum	0.05	0.00	0.00	0.60
Q1	0.30	0.47	1.92	1.28
Median	0.53	0.55	4.57	1.56
Q3	1.09	0.69	9.11	2.04
Maximum	1968.90	6.40	65.58	345.94

Table 6
Descriptive statistics for the random values generated for the unit prices (N=50,000).

Descriptive statistics	Medium sand (m^3)	Hydrated lime (kg)	Portland cement (bag 50 kg)	Concrete block 9 × 19 × 39 (thousand)	Steel bar 16 mm (kg)	Mixer 400 l (unit)	Grubbing hoe (unit)	Leather work gloves (pair)	Bricklayer (h)	Unskilled worker (h)
Minimum	6.06	0.16	18.95	1122.60	1.44	1495.39	9.00	6.20	6.30	5.19
Q1	42.59	0.40	21.95	1271.07	3.34	2828.20	12.71	7.54	6.32	5.21
Median	50.20	0.44	23.01	1426.56	3.76	3093.80	17.27	8.78	6.36	5.23
Q3	57.94	0.48	24.85	1592.46	4.18	3452.99	23.49	11.15	6.46	5.30
Maximum	96.49	0.61	58.87	1791.65	5.50	13,508.90	41.40	1989.89	75,223.44	15.04

fit' distributions presented in Table 4. Tables 5 and 6 provide a summary of descriptive statistics for the random values that were generated for material quantities and unit prices.

4. Delimiters

Considering the dominion range of the 'best fit' continuous distributions presented in Table 4 and the random values that were generated (Tables 5 and 6), it is evident that delimiters (i.e. the establishment of boundaries) are required. In establishing a lower delimiter, it needs to be recognized that a contractor may provide a low bid, which may initially be considered acceptable. Yet, often with low bids there is a propensity for a mistake to have been made or its acceptance may lead to poor quality work being undertaken. Accordingly, the Brazilian Federal Police have observed that with low bids there is a proclivity for the specified materials to be substituted for a lower quality and/or for defective work to occur; it is illegal to submit excessively low bids in Brazil and thus a delimiter is required to determine the maximum level of underpricing below SINAPI's median value. Similarly, a delimiter is required to establish a 'ceiling price', as the government has an obligation to taxpayers to obtain the most possible effective price in procuring their infrastructure projects. According

Table 7
Inferior and superior boundaries for the quantities.

Element	Unit	Boundaries	
		Inferior	Superior
Laying mortar 1:2:8	l/m	0.9600	2.4750
Bricklayer with social charges	h	0.2175	1.7250
Unskilled worker with social charges	h/h	0.3300	1.0875
Concrete hollow block $14 \times 19 \times 39\,cm^3$	% waste	0.0000	11.0625
Welded steel mesh	m	0.7850	0.7850
Steel pin 27 mm	unit	1.8900	1.8900

to the Brazilian law, the general procedure is to limit the maximum global price for an infrastructure project by the median prices provided by SINAPI [17].

4.1. Quantities delimiters

Studies examining the statistical nature of quantities have commenced in Brazil, which have been initiated by CAIXA, though no formal analysis been has made public at the time of writing. While the discussion about what is an acceptable range for overpricing in public projects needs to be made in Brazil, inter alia, the approach being proposed in this paper can accommodate any changes required anytime without changing the nature of the method.

In Brazil, it is common for a ‘10 mm thick’ mortar joint to be specified for walls; though this thickness can vary from range 3–50 mm. Considering the worksites from where the quantities data were collected followed quality standards, the delimiters are adopted at 75% and 125% of the first and third quartiles, which results in delimiters being set at 0.96 and 2.475 l/m, respectively.

For the bricklayer, issues surrounding productivity levels need to be considered. As the random values that were generated in Table 5 suggest, the maximum man-hours allocated to a bricklayer-gang for constructing a square meter of masonry could reach 1968 h. Clearly, this is extremely high; a more reasonable delimiter that reflects realistic practice needs to be determined. Considering the variability in levels of skill that may exist, and still looking for keeping a simple rule as example, the lower limit is again determined at 75% of the first quartile and the upper limit is 125% of the third quartile of values provided in Table 2. Thus, this results in a lower delimiter of 0.2175 and an upper of 1.725 h/m^2. The unskilled labor needed to assist the bricklayer also needs to be considered. Accordingly, their productivity is expressed as a ratio of the bricklayer’s quantity. There is a possibility that this variable could be ‘zero’, which means that no unskilled labor is required. Such a situation needs to be avoided as inefficiencies at the workface may manifest, thus having a negative impact of productivity levels. Akin to the bricklayers, the lower limit for the ratio unskilled worker/bricklayer is determined at 75% of the first quartile and the upper limit at 125% of the third quartile, which results in delimiters being set at 0.33 and 1.0875 respectively.

The delimiters for block wastage are influenced by the quality of the design and product that has been specified. In Brazil, a great deal of effort is being expanded to provide ‘zero waste’ to clients. Thus, lower delimiter for block wastage is zero in this instance. Conversely, the upper delimiter is 125% of the third quartile as a result of the values derived from Table 2, which results in a maximum waste of 11.06%. The raw data for welded steel mesh and the steel pin are not available. As a result, constant values are assumed. Table 7 summarizes the boundaries for the quantities of the surveyed components for the masonry wall.

4.2. Unit prices delimiters

As in the case for the quantities, boundaries need to be established to limit the unit prices variations to ensure acceptable values can be attained. For materials and equipment, the minimums are determined at 75% of the first quartile values shown in Table 3. In the case of the workforce, represented by the bricklayer and the unskilled worker, their minimum salaries and social benefits are protected by law. In São Paulo, in March of 2015, the minimum salaries were R$6.33 per hour for bricklayers and R$5.21 per hour for unskilled workers, values which will be considered as minimum boundaries for these components of the budget.

SINAPI’s median values are the maximum acceptable unit prices for materials and equipment. The law limits these prices and they are the most controllable variables for a contractor and the most reliable data for forensic experts (IBGE surveys the market for prices changes in all Brazilian capital cities each month). Regarding to the unit costs of the labour, Brazilian construction workers usually receive the minimum wage, which has been agreed in the form of bargaining agreements between contractors and labour unions. As these minimum wages are really low and the possibility of better payments for these workers cannot be restricted, their upper limit is extended to 125% of the third quartile in this study.

Table 8
Inferior and superior boundaries for the unit prices.

Element	Unit	Boundaries	
		Inferior	Superior
Medium sand	m^3	29.06	50.25
Hydrated lime	kg	0.29	0.46
Portland cement	bag 50 kg	16.43	22.60
Concrete block 9 × 19 × 39	thousand	937.50	1,400.00
Steel bar 16 mm	kg	2.38	3.79
Mixer 400 l	unit	2,174.25	3,088.00
Grubbing hoe	unit	9.68	16.28
Leather work gloves	pair	5.51	8.90
Bricklayer	h	6.33	9.26
Unskilled worker	h	5.21	6.79

Table 9
Basic descriptive statistics for the bounded random quantities' values (N=20,198).

Descriptive statistics	Bricklayer (h/m^2)	Unskilled worker ($h/h_{bricklayer}$)	Blocks waste (%)	Mortar (l/m)
Minimum	0.22	0.33	0.00	0.96
Q1	0.35	0.48	1.53	1.27
Median	0.52	0.55	3.45	1.49
Q3	0.82	0.67	6.17	1.77
Maximum	1.72	1.09	11.06	2.47

Table 10
Descriptive statistics for the bounded random unit prices' values (N=20,198).

Descriptive statistics	Medium sand (m^3)	Hydrated lime (kg)	Portland cement (bag 50 kg)	Concrete block 9 × 19 × 39 (thousand)	Steel bar 16 mm (kg)	Mixer 400 l (unit)	Grubbing hoe (unit)	Leather work gloves (pair)	Bricklayer (h)	Unskilled worker (h)
Minimum	29.06	0.29	18.95	1122.60	2.38	2175.19	9.68	6.20	6.33	5.21
Q1	38.45	0.39	21.36	1189.05	3.09	2669.97	11.15	7.02	6.35	5.23
Median	43.17	0.42	21.78	1257.29	3.37	2829.79	12.71	7.58	6.39	5.26
Q3	46.86	0.44	22.18	1327.67	3.59	2957.75	14.43	8.18	6.52	5.33
Maximum	50.25	0.46	22.60	1399.98	3.79	3087.99	16.28	8.90	9.26	6.77

5. Probabilistic masonry prices

As reported in Tables 5 and 6, resource quantities and the unit prices of the components necessary to build one square meter of a masonry wall have 50,000 random values each. Adopting the representative statistical distributions from data derived from construction sites generated these values. In addition, boundary limits were determined both to the quantities and the unit prices (Tables 7 and 8 refer). As a way of simulating the expected behavior of contractors, random values that are outside the boundaries were eliminated. For example, in certain samples of specialized operational costs, a larger number of discharges were noticed (e.g. the unit prices for the grubbing hoe). In these, the surveyed values presented the greatest dispersion. As result of the boundary conditions, 20,198 random sets of data were considered valid, and these are within the acceptable limits for the quantities and unit prices of each resource element (Tables 9 and 10). Fig. 4 presents the PDF for both quantities and unit prices.

Notably, the units for the surveyed components' quantities presented in Table 9 are different from those used for SINAPI shown in Table 1. To make them compatible, several arithmetic operations were needed. The quantity of unskilled worker in h/m^2 is given by the multiplication of the quantities of bricklayer by the ratio of unskilled worker/bricklayer provided in the second and third columns of Table 9 respectively. The quantities of blocks are given by the theoretical quantity of 12.5 units/m^2 added from the percentage of waste specified in the fourth column of Table 9. Finally, the quantity of laying mortar in m^3/m^2 is given by the multiplication of the volume in l/m specified in the fifth column of Table 9 by the constant of 7.38×10^{-3}, which represents the length of joints in each square meter of wall (12.5 units/$m^2 \times 0.59$ m/unit $\times 10^{-3}$ m^3/l).

It is also necessary to transform the surveyed unit prices available in Table 10 to the units shown in Table 1. This way, the unit price of the mortar will be given by the sum of the product between the quantities of its components by its random prices shown in Table 10. To perform this calculation, SINAPI's composition standards were considered e.g. for mortar (1:2:8), components include mixer operators' charges; mixer 400 l – operating time and idle time; and protection equipment, as well the prices of sand, hydrated lime and cement, statutory levies and charges and other cost items. The unit price of the laying

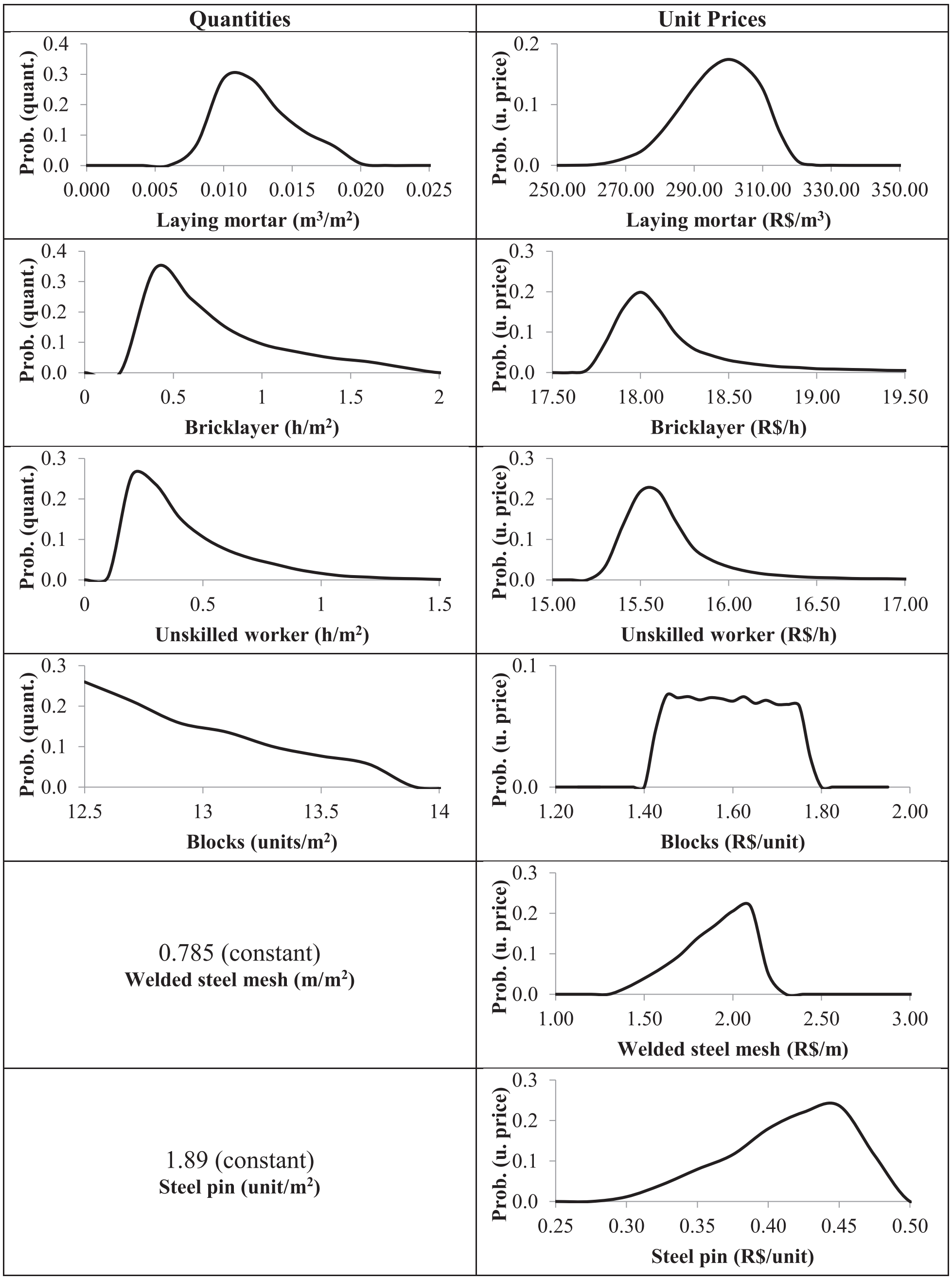

Fig. 4. PDF for quantities and unit prices.

mortar is summarized in the Eq. (1).

$$LM_{UP} = 11.41 \times UW_{UP} + 0.60 \times LG_{UP} + 4.65 \times 10^{-4} \times M_{UP} + 1.29 \times S_{UP} + 193.70 \times HL_{UP} + 3.71 \times PC_{UP} + 13.90, \quad (1)$$

where:

LM_{UP} = laying mortar unit price;
UW_{UP} = unskilled worker unit price;
LG_{UP} = leather gloves unit price;
M_{UP} = mixer 400 L unit price;
S_{UP} = sand unit price;

Table 11
Descriptive statistics for the random values generated for total prices ($N = 20{,}198$).

Descriptive statistics	$Q_i \times P_i$						Total price ($R\$/m^2$)
	Laying mortar 1:2:8	Bricklayer with social charges	Unskilled worker with social charges	Concrete hollow block $14 \times 19 \times 39$ cm^3	Welded steel mesh	Steel pin 27 mm	
Minimum	1.92	3.87	1.13	17.65	1.05	0.55	27.83
Median	3.23	9.46	4.67	20.54	1.48	0.77	40.43
Q3	3.86	15.06	7.59	21.67	1.58	0.82	48.92
Maximum	5.70	39.45	28.88	24.38	1.66	0.87	89.71

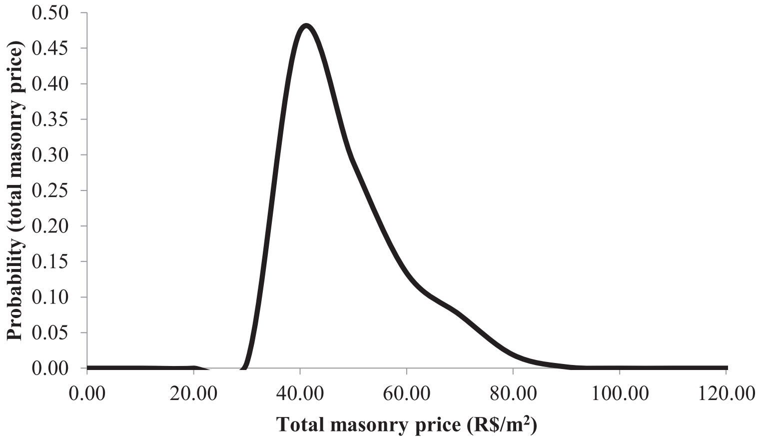

Fig. 5. PDF for masonry.

HL_{UP} = hydrated lime unit price;
PC_{UP} = Portland cement unit price.

'Bricklayers' charges' and 'unskilled workers' charges' unit prices are made up of wages to workers, costs of tools, protection equipment and social charges (feeding, transport, medical examinations and insurance). Equations 2 and 3 represent 'bricklayers' charges' and 'unskilled workers' charges' respectively.

$$BSC_{UP} = 2.18 \times B_{UP} + 0.02 \times GH_{UP} + 0.13 \times LG_{UP} + 2.81, \tag{2}$$

where:

BSC_{UP} = bricklayer with social charges unit price;
B_{UP} = bricklayer unit price;
GH_{UP} = grubbing hoe unit price;
LG_{UP} = leather gloves unit price.

$$UWSC_{UP} = 2.18 \times UW_{UP} + 0.02 \times GH_{UP} + 0.13 \times LG_{UP} + 2.81, \tag{3}$$

where:

$UWSC_{UP}$ = unskilled worker with social charges unit price;
UW_{UP} = unskilled worker unit price;
GH_{UP} = grubbing hoe unit price;
LG_{UP} = leather gloves unit price.

The size of the block used in the masonry wall adopted to illustrate this study is $14 \times 19 \times 39$ cm (width × height × length). Considering that the surveyed block is $9 \times 19 \times 39$ cm (thousand), it is necessary to apply a correction factor (this is part of the SINAPI's methodology that surveys representative items and obtains the prices of similar items using these correction factors). In this case, the unit price of a $14 \times 19 \times 39$ block is given by the multiplication of the random prices shown in Table 10 by the correction factor 1.26×10^{-3}. Similarly, the unit prices for the welded steel mesh and the steel pin 27 mm is approximated by the multiplication of the random prices for the steel 16 mm shown in Table 10 by the correction factors 0.56 and 0.12 respectively. Once all the random quantities and unit prices in their correct units are available, 20,198 random

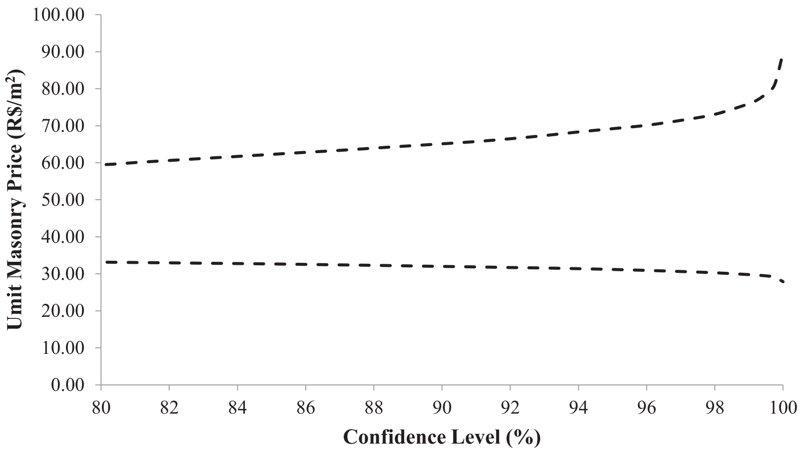

Fig. 6. Range of possible values depending on the required confidence level.

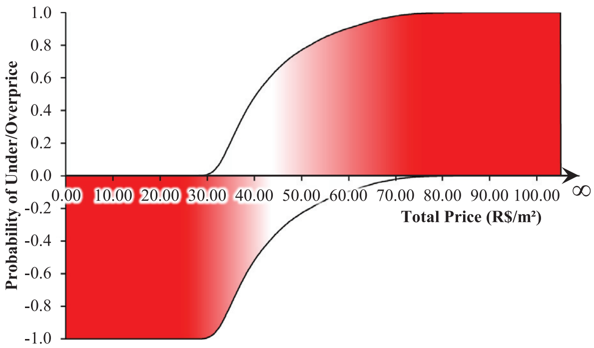

Fig. 7. Probabilities of ‘under/overpricing’ for masonry.

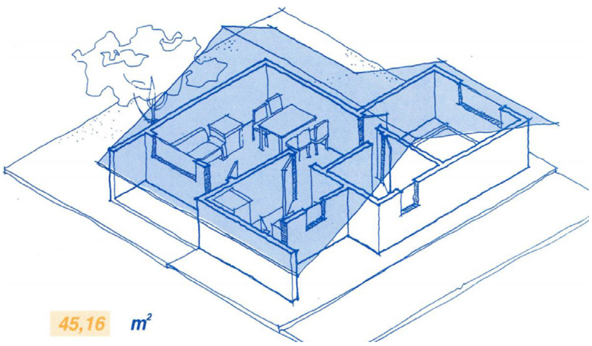

Fig. 8. Schematic perspective of social house used as case example.

total prices were calculated according to Eq. (4), and an extract of these values are shown in Table 11.

$$TP = \sum_{i=1}^{n} Q_i \times P_i, \tag{4}$$

where:

TP = expected budget’s total price;
Q_i = random quantity for each supply;
P_i = random unit price for each supply.

The general behavior of the analyzed data can be obtained, and a PDF for the expected probabilistic price was obtained for the masonry wall (Fig. 5). These probabilistic calculations allow the forensic experts to access the probability that a given price can (or is expected to) occur. As Fig. 6 shows, the prices can be accessed for each confidence interval, and through

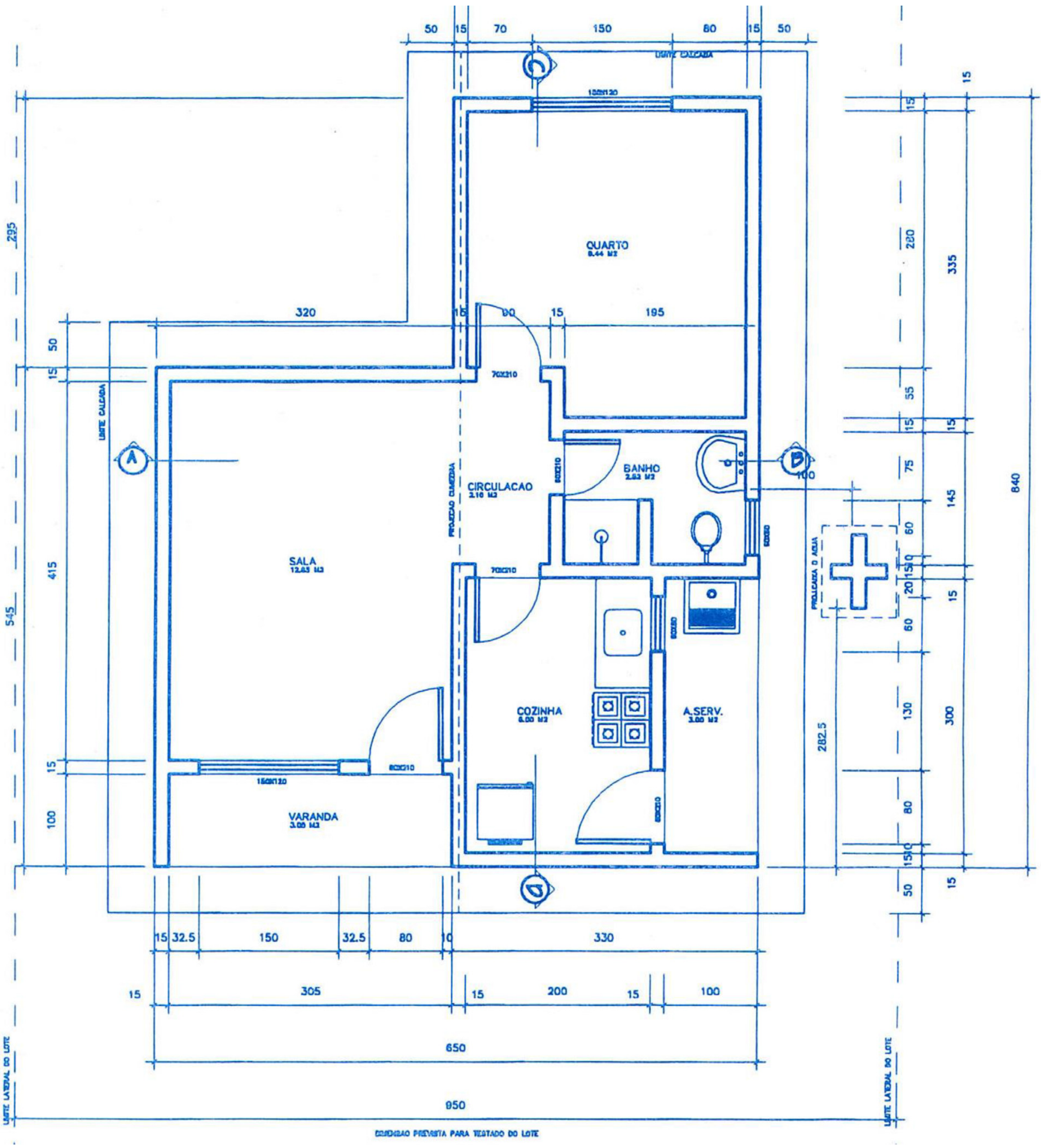

Fig. 9. Floor plan for the house.

this it is possible to estimate the probability of occurrence for each price range, or vice versa. For example, instead of fixing masonry price under R$ 60.78/m^2 as 'adequate' (i.e. prices within 10% over SINAPI's median price), the forensic experts can conclude that, with 80% confidence, the expected unit price for the same item (masonry) could be between R$ 33.13 and R$ 59.43. This will rank as under or overpricing the values outside this range. Or rather, accordingly to the assumptions of the calculations (and in the absence of extraordinary events), the price of masonry used in the illustration presented earlier could be between R$ 27.83/m^2 and R$ 89.71/m^2, at 100% of confidence.

The method introduced in this paper can enable forensic experts to scientifically to determine if an infrastructure has been under or overpriced. For example, if masonry (tagged as #87449 in SINAPI's database) was contracted for R$6500/m^2 in São Paulo in March 2015, there is a 94% probability that the contract was overpriced (and is underpriced with the complementary 6% of probability, excluded the negligible probability of the questioned price be exactly the price of reproduction of the work). Fig. 7 allows the visualization of the general behavior of the phenomenon. In the next section of this paper,

Table 12
Social house project budget (in SINAPI's order of appearance).

SINAPI's code	Element	Unit	Quant.	Unit prices (R$)		
				Q1	Q2	Q3
23,726/2	Roof with ceramic tile including wooden structure	m²	58.78	89.05	95.12	109.92
23,717/1	Wooden door (including frame and painting)	m²	5.04	155.79	166.94	202.75
23,720/1	Iron door (venetian, including frame and painting)	m²	3.36	221.33	242.76	304.08
23,739/2	Tilting iron window, 1.2 × 1.2 m, glass 4 mm, primer and glaze painting	m²	0.84	368.44	375.77	408.71
23,740/1	Sliding iron window, 1.2 × 1.0 m, glass 4 mm, primer and glaze painting	m²	3.60	424.96	432.89	472.17
23,732/1	Lock and hinge in iron nickel (supply and installation)	un.	3.00	93.88	106.44	116.29
24,260/3	Foundation beams in cyclopean concrete (including excavation)	m³	5.12	355.02	391.68	472.82
23,733/1	Reinforced concrete belt fck 15 MPa including steel and formwork	m³	0.59	1.201.39	1306.55	1584.31
23,748/1	Precast concrete lintel fck 20 MPa including steel and formwork	m	11.80	11.06	12.23	14.42
25,573/2	Wooden support for water tank	un.	1.00	17.98	19.94	21.89
23,711/1	Mortar 1:3 (cement:sand) with waterproofing additive (2 cm thick)	m²	11.97	25.41	26.88	33.95
39,629	Electrical installations for social house	un.	1.00	841.98	964.96	1158.64
41,598	Provisory connection to the electrical energy network (three-phase 40A on wood pole)	un.	0.50	625.73	713.48	846.76
23,728/1	Laundry tub (synthetic marble 22 L) including metal valve (supply and installation)	un.	1.00	164.98	202.35	237.14
23,729/1	Kitchen sink (synthetic marble 1.2 × 0.6 m) including tap (supply and installation)	un.	1.00	186.86	227.36	264.99
23,742/1	Paper dispenser, soap dispenser and hang in white porcelain (supply and installation)	un.	1.00	103.70	110.43	129.19
23,743/1	Toilet in white porcelain with plastic water tank (supply and installation)	un.	1.00	198.15	211.15	260.88
23,745/1	Bathroom sink (white porcelain) including metal tap (supply and installation)	un.	1.00	126.58	139.79	173.15
39,706	Protective case for water meter	un.	1.00	24.12	26.83	31.93
23,706/1	Sewage installations for social house	un.	1.00	544.29	596.39	792.16
23,706/4	Water installations for social house	un.	1.00	550.07	629.02	764.37
23,746/1	Connection to the water and sewage network	un.	1.00	144.43	164.72	195.37
23,776/1	Clay block masonry 10 cm tick	m²	105.50	28.93	31.29	38.09
23,714/1	Primer and latex PVA paint (two coats)	m²	209.91	11.16	11.63	16.22
23,741/2	Cemented floor 1:4 (sand:cement) 1.5 cm thick	m²	39.62	26.26	27.54	35.39
23,744/1	Sidewalk in lean concrete 5.0 cm tick, regularized with mortar	m²	12.45	31.08	33.29	40.70
23,719/2	Lean concrete blinding layer	m³	1.98	396.27	433.09	520.11
23,708/1	Slurry mortar (external)	m²	94.03	3.40	3.54	4.32
23,708/2	Slurry mortar (internal)	m²	122.50	3.40	3.54	4.32
23,710/1	Fine mortar plaster 2 cm tick (internal)	m²	122.50	16.68	17.35	21.77
23,712/1	Mortar plaster 2 cm tick (external)	m²	94.03	17.46	18.17	22.64
23,713/1	White tile 15 × 15 cm (supply and laying)	m²	6.62	38.62	42.43	52.14
23,735/1	Ground manual scraping and cleaning	m²	200.00	1.99	2.00	2.66
23,738/1	Location frame	m²	45.16	4.99	6.20	7.91
Totals (R$ – Florianópolis, Dec/2012)				27968.36	30000.35	36609.25

two examples are presented to demonstrate how the proposed probabilistic approach can be used in practice. The approach is being studied by the Brazilian Federal Police and will serve as method to be used in obtain convictions for overpricing of public sector projects.

6. Case example

A house constructed in Florianópolis in December 2012 with an area of 45.16 m² (registered as project #1878 in the SINAPI database) is selected to demonstrate the application of the proposed method (Figs. 8 and 9). The derived data from SINAPI appears in its original form in Table 12. Noteworthy, SINPAI does not have price and quantities data for the elements. Moreover, only the quartiles for unit prices have been released by IBGE. Thus, the 'best fit' distribution cannot be computed. The characteristics of all components have positive values and the data generally skewed to the right; in this instance, a Lognormal distribution is selected. The means for all variables are assumed to be equal for the fixed quantities and the median unit prices. The standard deviation for quantities were determined to be 1.2, assuming that both materials and workmanship have combined variability's that can accommodate this value. For instances were an item is represented by supplied material, the quantities' standard deviations are determined to be 1.05. Thus, the chance of an error being made and the variability are reduced. For example, in the case of the supply and installation of a kitchen sink, the materials are well defined and leave no chance for errors in quantities. The workmanship component, however, imposes some degree of uncertainty. For unit prices the standard deviations are calculated from the interquartile range released by SINAPI, using the following equation:

$$\sigma \approx \left(\frac{Q3}{Q1}\right)^{\frac{1}{1349}}, \tag{5}$$

Table 13
Lognormal distributions properties for quantities and unit prices of the social house project.

Element	Unit	Quantities' properties		Unit prices' properties	
		μ	σ	μ	σ
Roof with ceramic tile including wooden structure	m²	58.78	1.20	95.12	1.17
Wooden door (including frame and painting)	m²	5.04	1.05	166.94	1.22
Iron door (venetian, including frame and painting)	m²	3.36	1.05	242.76	1.27
Tilting iron window, 1.2 × 1.2 m², glass 4 mm, primer and glaze painting	m²	0.84	1.05	375.77	1.08
Sliding iron window, 1.2 × 1.0 m², glass 4 mm, primer and glaze painting	m²	3.60	1.05	432.89	1.08
Lock and hinge in iron nickel (supply and installation)	un.	3.00	1.05	106.44	1.17
Foundation beams in cyclopean concrete (including excavation)	m³	5.12	1.20	391.68	1.24
Reinforced concrete belt fck 15 MPa including steel and formwork	m³	0.59	1.20	1306.55	1.23
Precast concrete lintel fck 20 MPa including steel and formwork	m	11.80	1.20	12.23	1.22
Wooden support for water tank	un.	1.00	1.20	19.94	1.16
Mortar 1:3 (cement:sand) with waterproofing additive (2 cm thick)	m²	11.97	1.20	26.88	1.24
Electrical installations for social house	un.	1.00	1.20	964.96	1.27
Provisory connection to the electrical energy network (three-phase 40A on wood pole)	un.	0.50	1.05	713.48	1.25
Laundry tub (synthetic marble 22 l) including metal valve (supply and installation)	un.	1.00	1.05	202.35	1.31
Kitchen sink (synthetic marble 1.2 × 0.6 m²) including tap (supply and installation)	un.	1.00	1.05	227.36	1.30
Paper dispenser, soap dispenser and hang in white porcelain (supply and installation)	un.	1.00	1.05	110.43	1.18
Toilet in white porcelain with plastic water tank (supply and installation)	un.	1.00	1.05	211.15	1.23
Bathroom sink (white porcelain) including metal tap (supply and installation)	un.	1.00	1.05	139.79	1.26
Protective case for water meter	un.	1.00	1.20	26.83	1.23
Sewage installations for social house	un.	1.00	1.20	596.39	1.32
Water installations for social house	un.	1.00	1.20	629.02	1.28
Connection to the water and sewage network	un.	1.00	1.20	164.72	1.25
Clay block masonry 10 cm tick	m²	105.50	1.20	31.29	1.23
Primer and latex PVA paint (two coats)	m²	209.91	1.20	11.63	1.32
Cemented floor 1:4 (sand:cement) 1.5 cm thick	m²	39.62	1.20	27.54	1.25
Sidewalk in lean concrete 5.0 cm tick, regularized with mortar	m²	12.45	1.20	33.29	1.22
Lean concrete blinding layer	m³	1.98	1.20	433.09	1.22
Slurry mortar (external)	m²	94.03	1.20	3.54	1.19
Slurry mortar (internal)	m²	122.50	1.20	3.54	1.19
Fine mortar plaster 2 cm tick (internal)	m²	122.50	1.20	17.35	1.22
Mortar plaster 2 cm tick (external)	m²	94.03	1.20	18.17	1.21
White tile 15 × 15 cm² (supply and laying)	m²	6.62	1.20	42.43	1.25
Ground manual scraping and cleaning	m²	200.00	1.20	2.00	1.24
Location frame	m²	45.16	1.20	6.20	1.41

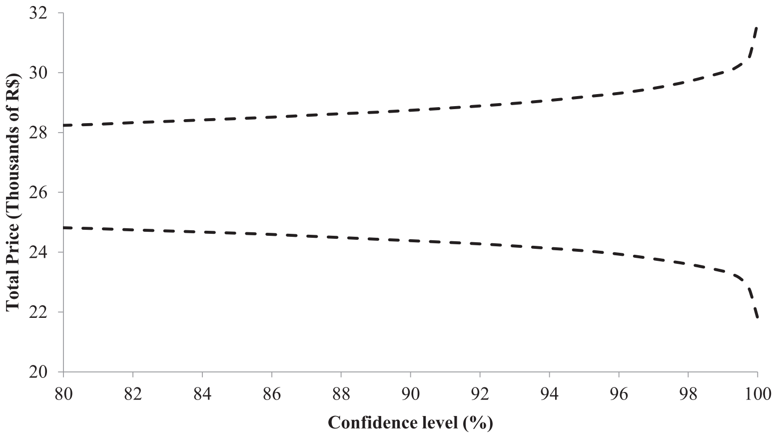

Fig. 10. Range of prices and desired confidence level.

where,

σ = expected standard deviation for the unit price of each component;
Q3 = third quartile price of each component;
Q1 = first quartile price of each component.

Using these considerations, the Lognormal distributions properties for quantities and unit prices are provided in Table 13. With these properties, each step of the described process was undertaken. The results for the entire project are

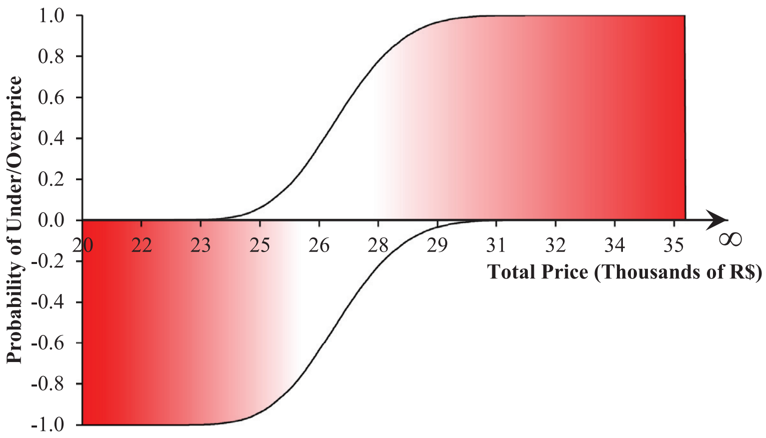

Fig. 11. Probabilities of 'under/overpricing' for the house.

provided in Figs. 10 and 11, which leads to the conclusion that the method is applicable to projects of any size, requiring only the existence of raw data, which in Brazil is being collected by the SINAPI's review process.

7. Conclusion

Fraudulent overpricing in infrastructure projects is common in Brazil and is also globally a ubiquitous problem. Its determination in Brazil has been based upon the utilization of a deterministic approach, which has sometimes been ineffective in proving beyond reasonable doubt that overpricing has occurred. Consequently, it has been difficult for Judges to apply criminal convictions for this practice. To address this pervasive problem a new probabilistic method for determining overpricing is presented.

The new approach enables forensic experts to establish the probability of overpricing occurring in accordance with the variability that exists for quantities and the unit prices resources needed to construct a project. Once it is adopted, the budget for public works will no longer be represented by a fixed value, but by a range of possible values associated with certain level of confidence.

The reliability of the method is assured by the LLN principle and depends on data on quantities and unit prices. In Brazil, the data on unit prices are monthly surveyed by IBGE, and the data on quantities are currently being surveyed in the review process of an important official database, the SINAPI. Once this data available, the decisions on infrastructure projects' prices can be made probabilistically, instead of deterministically. A major limitation of the proposed method is its dependence upon the use of reliable empirical data that has been collated in standardized format; this may not be available in many other countries. In this case of Brail, SINAPI forms part of its legislative framework but in other jurisdictions such databases of this nature may have an innate bias and be subjective due to their lack of standardization rendering the proposed method to be obsolete.

Acknowledgments

The authors would like to acknowledge the *Caixa Econômica Federal*, SINAPI's manager, and the *Instituto Brasileiro de Geografia e Estatística*, for their support and for the data released for this study.

References

[1] M. Shan, A.P.C. Chan, Y. Le, B. Xia, Y. Hu, Measuring corruption in public construction projects in China, ASCE J. Prof. Issues Eng. Educ. (2015), doi:10.1061/(ASCE)EI.1943-5541.0000241.
[2] C. Kenny, Construction, Corruption, and Developing Countries, World Bank, Washington, DC, 2007 https://wdronline.worldbank.com/handle/10986/7451 License: CC BY 3.0 Unreported.
[3] Transparency International. Corruption Perceptions Index 2014. (Available at www.transparency.org/whatwedo/publication/cpi2014, Accessed June 10th 2015).
[4] United Nations, (2004). United Nations Convention against Corruption. (Available at: www.unodc.org/documents/treaties/UNCAC/Publications/Convention/08-50026_E.pdf , Accessed June 10th, 2015).
[5] M.C. Lima, Comparação de custos referenciais do DNIT e licitações bem sucedidas, in: Proceedings of the XIII Simpósio Nacional de Auditoria de Obras Públicas, SINAOP – Porto Alegre, RS, 2010 (In Portuguese).
[6] A.O. Lopes, Superfaturamento de Obras Públicas, Livro Pronto, São Paulo, 2011 (In Portuguese).
[7] Signor, R., Oliveira, P.S. Jr., Silva, M.A., Oliveira, D.A.R. (2006). O Potencial Das Perícias De Engenharia Em Rodovias. *Perícia Federal*, January, n. 23, pp. 6–10. (Available at http://www.apcf.org.br/Portals/0/revistaAPCF/23.pdf , Accessed Sep 12th 2015) (In Portuguese).
[8] R. Signor, P.E.D. Love, O. Olatunji, Determining overpricing in Brazilian infrastructure projects. A forensic approach, ASCE J. Constr. Eng. Manag. (2016), doi:10.1061/(ASCE)CO.1943-7862.0001156.
[9] Boudreaux, C.J. and Coats, R.M. and Karahan, G., (2016). Bend it Like FIFA: Corruption on and off the Pitch (Accessed at 14th February, Available at SSRN: http://dx.doi.org/10.2139/ssrn.2719718)

[10] A. Carrión, Megaprojects and the restructuring of urban governance: the case of the New Quito international airport, Lat. Am. Perspect. 43 (1) (2016) 252–265.
[11] X. Deng, Y. Wang, Q. Zhang, X.J. Huang, Cui Jingjing, Analysis of fraud risk in public construction projects in China, Public Money Manag. 34 (1) (2014) 51–58.
[12] K. Mochtar, D. Arditi, Pricing strategy in the US construction industry, Constr. Manag. Econ. 19 (4) (2001) 405–415.
[13] M. Sohail, S. Cavill, Accountability to prevent corruption in construction Projects, ASCE J. Constr. Eng. Manag. 134 (9) (2008) 729–738.
[14] Poder Judiciário (2009). Justiça Federal de Primeira Instância. Seção Judiciária do Estado de Sergipe. 2a Vara Federal. Processo n° 2009.85.00.005336-0. (Available at http://consulta.jfse.jus.br/Consulta/resconsproc , Accessed May 28th, 2015) (In Portuguese).
[15] P.E.D. Love, C.-P. Sing, X. Wang, R. Tiong, Determining the probability of cost overruns in Australian construction and engineering projects, ASCE J. Constr. Eng. Manag. 139 (3) (2013) 321–330.
[16] P.E.D. Love, C.-P. Sing, Determining the Probability Distribution of Rework costs in construction and engineering projects, Struct. Infrast. Eng. 9 (11) (2013) 1136–1148.
[17] Brazil (2013) Decreto no 7.983 (Available at: http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2013/Decreto/D7983 , accessed July 15th 2015) (In Portuguese).

Please cite this article as: R. Signor et al., A probabilistic method for forensic cost estimating of infrastructure projects, Applied Mathematical Modelling (2016), http://dx.doi.org/10.1016/j.apm.2016.05.025