The analysis of pre-tender building price forecasting performance: A case study

Skitmore, Martin; Drew, Derek

Published in:
Engineering, Construction and Architectural Management

DOI:
10.1108/09699980310466532

THE ANALYSIS OF PRE-TENDER BUILDING PRICE FORECASTING PERFORMANCE: A CASE STUDY

Martin Skitmore (corresponding author)
School of Construction Management and Property
Queensland University of Technology
Gardens Point
Brisbane Q4001
Australia

Derek Drew
Department of Building and Real Estate
The Hong Kong Polytechnic University
Kowloon
Hong Kong

Paper for *Engineering, Construction and Architectural Management*

21 May 2002
(version 2)

THE ANALYSIS OF PRE-TENDER BUILDING PRICE FORECASTING PERFORMANCE: A CASE STUDY

Abstract: The financial management of the construction procurement process is dependent upon on the performance of the managers involved. This paper describes an analysis of pre-tender building price forecasts (estimates) made by a Hong Kong consulting organisation for a series of 89 building projects from 1995 to 1997 to identify factors influencing the accuracy of the forecasts made for possible improvement in performance. This involved the consideration of two distinct sets of models the purpose of which was (1) to identify and explain the underlying systematic causes of errors and (2) to assist in improving the predictive ability of the forecasts.

The analysis for (1) used ANOVA to detect significant differences between the errors grouped according to building size (value), building size (floor area), forecasting (estimating) method (approximate quantities and superficial), nature of the work (new build and alteration work), type of client and type of project. This was followed by a Gunner-Skitmore Price Intensity (PI) theoretic analysis. For (2), MRA was used with using cross-validation analysis to simulate the *ex-post* errors.

Keywords: Pretender estimating, performance, accuracy, price intensity theory.

INTRODUCTION

The analysis of pre-tender forecasting (estimating) accuracy has long been a topic of academic interest and has developed over the years into a wide-ranging set of approaches. Although benefiting the field of study, in terms of sheer volume of analyses reported, the lack of a clear consensus on analytical method to use (Gunner and Skitmore, 1999) has caused difficulties in (1) comparing the results of the reported studies and (2) gaining the confidence of practitioners for serious implementation in practice.

The causes of this situation are multifarious. A major issue has been that most researchers in the topic have been illequipped to do the work – the analysis is essentially statistical in nature and the researchers, almost without exception are not statisticians. Another equally important issue is that, until recently, there has been a complete absence of any theoretical explanation for the underlying causes of forecasting accuracy (Gunner and Skitmore, 1999). The result of this has been the undirected collection and analysis of data, relying on retrospective, *ad hoc*, data held by practitioners.

Gunner and Skitmore's (1999) recent Price Intensity (PI) theory, however, promises to change this situation. As yet, PI theory has had little direct empirical confirmation except for the analysis of one set of Singapore data (although it is also supported indirectly through replication of all previous work with no notable discrepancy in results).

In this paper, an analysis is described of pre-tender forecasts (estimates) made by a Hong Kong consulting organisation for a series of 89 building projects from 1995 to 1997, the main objective being to identify factors influencing the accuracy of the forecasts made for possible improvement in performance. This involved the consideration of two distinct sets of models the purpose of which was (1) to identify and explain the underlying systematic causes of errors and (2) to assist in improving the predictive ability of the forecasts. The analysis for (1) first follows the conventional approach of summarising the distributional characteristics of the errors for subgroups of each of the variables available. The statistical results of this analysis show the only significant effect, in both bias and consistency, to be the forecasting (estimating) method used. This was followed by a PI theoretic analysis, which simultaneously removed the confounding effects of the conventional treatment and confirms the applicability of PI theory to the case study.

The analysis for (2) involved regressing all the independent variables, and a variety of transformations, on an additive as well as multiplicative version of the dependent variable, using cross-validation analysis to simulate the *ex-post* errors.

ANALYSIS

Data

Pretender forecasting (estimating) data for a complete set of 89 building projects for the period January 1995 to October 1998 were collected from a Hong Kong private quantity surveying consultant (Appendix A). For each of the projects, the forecasting (estimating) method (approximate quantities or superficial), tender date, gross floor area, forecast price, lowest bid, type of client, type of project and nature of the work (new build or alterations) were recorded. For comparison purposes, all the monetary values were deflated to a common base date of

March 1999 by means of the Levett and Bailey (1999) local tender price index. The floor areas of three projects were not known.

All the projects were carried out in Hong Kong and therefore, as Hong Kong building prices are known to be homogeneous (Drew and Skitmore, 1997) no adjustment was made for geographical price differences. All of the forecasts were carried out by professional, certified trained, surveyors.

All significance tests were made at the 5% level.

Explanatory models

Conventional analysis

Table 1 summarises the results of the conventional analysis. Column 1 describes the grouping variable, with sub-groups inset; column 2 gives the number of relevant projects involved; column 3 gives the mean percentage difference between the forecast bid price and the lowest bid price (a positive value indicates an overestimate); column 4 gives the standard deviation around the mean; and column 6 gives this expressed as the coefficient of variation.

The overall distribution of errors has a mean of –1.78% with a standard deviation of 12.95 (13.19% coefficient of variation[1]). The mean is not significantly different from 0%, which indicates the forecasts to be unbiased overall.

[1] The coefficient of variation of the ratio, ie., 100(0.1295/0.9822)

The only sub-groups to record significant differences in bias (mean errors) or consistency (standard deviations) to each other are those for 'Method', with significant differences in both means (ANOVA F=5.501, df=1,87, p=0.021) and standard deviations (Bartlett's χ^2=4.73, p=0.030). This is very much a surprise - the Approximate Quantities method, as it utilises more data, being expected to have a lower coefficient of variation than the Superficial method instead of *vice versa*. Even more suprisingly however, apart from James' (1954) early comparison of the accuracy of the Superficial, Cube and Storey Enclosure methods, which showed the Superficial and Storey Enclosure methods produce significantly more accurate results than the Cube method (χ^2=5.99, df=2) (Skitmore et al, 1990:15), no other empirical studies comparing traditional methods have been reported. It is possible, therefore, that the result found here – that the Superficial method is significantly more accurate than the relatively more resource consuming Approximate Quantities method – is correct. Of course, this does not imply that the Superficial method will always outperform the Approximate Quantities method. It is possible, for example, that the use of the Superficial method has deliberately been restricted to simpler, more typical, projects, with the Approximate Quantities method being reserved for only those that are more complex and less typical. No data, however, were available to verify this possibility.

PI analysis

Gunner and Skitmore's (1999) approach to PI analysis was to conduct a series of trivariate regression analyses of the ratio error (forecast/low bid value) on price intensity ratios (low bid value/gross floor area) plus a further independent variable. As the theory holds that price intensity, and price intensity alone, correlates with forecasting error (bias), the expectation is that the price intensity variable will always be significant, irrespective of the additional independent variable and that the additional independent variable will never be significant. Trivariate regression, however, estimates the partial coefficients of the price intensity variable and the additional variable simultaneously, with each allowing for the influence of the other. This rather contradicts the theory itself, which maintains that the additional variable can have no such influence. In other words, the theory implies that whatever correlation is detected by the regression for the additional variable must be spurious. To avoid this contradiction, the regressions were instead approached in hierarchical manner by first regressing the error variable on the price intensity variable and then testing the ensuing residuals with the additional variables for any possible residual correlation or heteroscedasticity (there should be none). Fig 1 provides a plot of price intensity and against the percentage error of forecast. Table 2 summarises the regression results. Only 86 cases could be analysed due to the three projects with unknown floor areas. As can be seen, the model is significant with an r^2 of 0.145 and the price intensity variable (INTENSIT) together with the intercept have a significant t values.

As predicted by PI theory, none of the sub-groups now has significantly different means or standard deviations.

Predictive models

Introduction

PI theory, whilst accounting for all significant systematic forecasting errors, has relatively little to offer as a predictive theory. Clearly, it is one thing to know that expensive projects, in terms of price intensity, are systematically underestimated and inexpensive projects are systematically overestimated (Fig 1) but quite another to be able to predict expensive and inexpensive projects **in advance**. Inspection of the regression equation of percentage error on price intensity indicates why this is the case – both dependent and independent variables contain what is in predictive mode the unknown value of the lowest bid. The obvious solution to this is just to replace the unknown true value of the lowest bid with the forecaster's estimated value of the lowest bid. However, a moment's reflection will show why this is not appropriate. The forecaster's estimated value of the lowest bid is, as illustrated in Fig 1, biased towards the mean price intensity of all the projects and, as this bias is a major aspect that we are trying to correct (we would also like to reduce the spread or standard deviation of the errors), it is unlikely to be of much use in identifying the nature of the correction.

It is also clear that, assuming PI theory is correct, whenever the price intensity error is removed only purely random 'noise' can remain. It follows therefore that the only correction needed is to the price intensity forecast. This suggests the need for a model that either has (1) the actual price intensity as the dependent variable with the forecasted price intensity included among the independent variables or (2) the actual intensity forecast error as the dependent variable with or without the forecasted price intensity included among the independent variables. In fact using arithmetic differences (between actual and forecasted price intensity) as the dependent variable produces proportionally identical results for (1) and (2), so there are really only two basic approaches available depending on whether a multiplicative (percentage error) or additive (difference) dependent variable is used.

Finally, the cross-validation method was used to measure the performance of the predictive models as this gives the closest simulation to *ex post* results available.

Results

The independent variables chosen for the analysis were the raw, log and inverse forecast value, raw, log and inverse gross floor area, the forecasted price intensity value and dummy variables representing the forecasting method, nature of the work, project type and client type. A forward stepwise-like procedure was used in which each independent variable was entered into the regression equation, the deleted residual calculated for each project, and the standard deviation of the deleted residuals calculated for the independent variable. The independent variable was then removed and replaced by a different independent variable from the above list and the process repeated. The results are summarised in Table 3 for both the multiplicative (percentage) and additive (difference) error dependent variable. The first row of the results gives the standard deviation of the original error term (before adjustment) for comparison with the regression results. As Table 3 shows, such improvements in standard deviation that do occur as trivial. In view of this, the intended forward stepwise-like approach was terminated as having failed to achieve enough improvement to allow entry of the first variable into the equation.

CONCLUSIONS

The work described in this paper was aimed at developing a statistical method for improving pre-contract price forecasting accuracy. This was demonstrated via the analysis of a set of forecasts from a Hong Kong firm of consultants for 89 building projects. Using a modified Price Intensity (PI) theoretic approach, the significant findings of a conventional analysis were disconfirmed. The use of PI theory for predictive purposes was then examined and two approaches applied to the prediction of the forecasting errors by a cross validation forward stepwise-like technique.

ACKNOWLEDGEMENTS

We are indebted to H K Yu for kindly granting permission, and Chan Suk Wan (whose dissertation provided the basis for this paper) and Anslem Chow, for collecting the data.

REFERENCES

Ashworth, A., Skitmore, M., 1983, Accuracy in estimating, CIOB, *Occasional Paper* 27, The Chartered Institute of Building, Englemere, Kings Ride, Ascot, UK, ISBN 0 906600 57 X

Drew, D., Skitmore, M., 1997, The effect of contract type and contract size on competitiveness in construction contract bidding. *Construction Management and Economics* **15**(5) 469-89.

Gunner, J., Skitmore, M., 1999, Building contract price forecasting: price intensity theory. *Engineering, Construction and Architectural Management* **6**(3) 267-75.

James, W., 1954, A new approach to single price-rate approximate estimating. *The Chartered Surveyor*, May.

Levett and Bailey, 1999, Tender price indices and cost trends, June 1998, Levett and Bailey Quantity Surveyors, Hong Kong.

Skitmore, M., Stradling, S.G., Tuohy, A.P., Mkwezalamba, H., 1990, The accuracy of construction price forecasts: a study of quantity surveyors' performance in early stage estimating. University of Salford, UK, ISBN 0 901025 12 7.

APPENDIX A: Data

Proj	Forecast	Lowest bid	Meth	GFA	sec	type	nature
1	79241250	82888103	2	8760	3	32	1
2	43471015	41618240	1	4390	3	721	1
3	80493383	89250034	2	5195	1	721	1
4	183939361	214178580	1	19985	1	816	1
5	11628199	13268606	2	915	2	816	1
6	42960772	40726449	1	5288	2	442	1
7	4058621825	5274295620	1	295417	1	32	2
8	29569708	31339221	1	1950	3	412	2
9	129290146	135400955	2	20000	1	32	1
10	69903285	77979898	2	4100	3	721	2
11	88273179	80922259	1	8940	3	442	1
12	107814107	87754775	1	8070	3	442	1
13	106748679	104342104	1	7590	3	852	1
14	91408929	87171429	2	5720	3	852	1
15	273016071	231894093	1	20233	1	816	1
16	90004500	89981696	2	7956	1	32	1
17	131096143	130389206	1	9770	3	442	1
18	143663357	119127192	1	13416	3	442	1
19	91893214	89981696	1	6519	1	32	2
20	581506071	617222143	2	33490	1	816	1
21	1240431818	1146066545	2	63000	3	852	1
22	298201469	276770979	2	31072	1	713	1
23	188642832	224336213	1	10753	2	713	1
24	169179965	167282039	1	16350	3	721	1
25	79605734	73008626	1	12893	1	282	1
26	2821048951	3761003497	1	175400	1	32	2
27	35879476	28053749	1	158900	4	981	2
28	1153311189	998245766	2	134468	2	981	1
29	1124863636	994479021	1	148090	4	816	1
30	159424826	184302210	2	21930	2	816	1
31	574344231	581259165	1	47882	1	816	1
32	150666667	118659657	1	17535	2	981	1
33	619787879	668167744	2	6200	3	721	2
34	253393939	281472727	2	31096	2	522	2
35	128671616	124642424	1	7168	1	32	1
36	902858586	836329546	1	61370	1	816	1
37	67434747	81993798	1	3705	1	816	1
38	31149192	32020059	2	31096	2	522	1
39	196505859	216027465	1	20889	2	713	1
40	100395296	98575559	1	10876	2	442	1
41	45720395	52231029	1	6975	1	447	1
42	84192434	80289474	2	5710	3	412	1
43	132867928	211053148	2	8864	2	713	1
44	98321151	96104570	1	7500	2	412	1
45	1148586	1417245	2	1550	1	816	1
46	9179905	9347925	2	470	2	543	1
47	213623810	204515505	2	19350	2	852	2
48	13377048	12825850	1	800	2	852	1
49	171996762	142815443	1	18756	3	412	1
50	201247619	204515605	2	13155	2	852	1
51	519800000	570447434	1	28002	1	816	1
52	768432286	858800000	2	102440	1	816	1
53	353905238	328572439	1	45480	2	981	1
54	81446095	91956294	2	5080	4	412	1
55	224364190	213085714	1	11760	3	412	1
56	258522476	247590291	2	23227	4	713	1

57	118079619	131994762	2	9760	2	442	1
58	319079714	328572439	1	45400	2	981	1
59	25290476	27709752	1	238500	4	114	1
60	113107619	131437924	2	7828	1	32	1
61	99805905	99975480	2	9860	2	442	1
62	163139714	147438095	2	20488	4	32	1
63	280675663	311747946	1	43710	1	282	1
64	22463855	23374229	2	1441	2	534	2
65	63051958	68964095	2	4130	2	154	1
66	11970938	10900744	1		2	816	2
67	18876136	23820698	2		2	32	2
68	62599432	79200725	1	3650	4	114	1
69	66615426	69804935	1	8050	2	442	1
70	92550852	69804935	1	7100	2	442	1
71	62599432	57898118	2	2400	2	816	1
72	7242273	7121754	2		2	342	2
73	65552486	69200538	1	5000	1	816	2
74	86847680	86995238	2	8140	2	442	1
75	638359475	747298343	2	93452	4	816	1
76	125636022	157379841	1	16620	2	442	1
77	73390691	89900552	2	2132	2	852	1
78	7988039	6261316	1	470	2	534	2
79	188987818	181013465	1	18756	3	721	1
80	263408619	271229594	1	27594	1	32	1
81	468232044	496383817	1	28000	1	816	1
82	51791667	64076323	1	4440	3	721	1
83	172861190	189310974	2	12750	1	816	1
84	148238074	141997658	2	9325	1	816	1
85	657400000	658550000	2	28350	2	32	1
86	41900000	39188000	2	2527	1	816	1
87	650000000	658550000	2	28240	1	32	1
88	515460000	658550000	2	28610	1	32	1
89	170000000	197129126	2	11700	4	816	1

Variable	N	mean error (%)	standard deviation	coefficient of variation
Total	*89*	*-1.78*	*12.95*	*13.19*
Method				
Approx Q	46	1.25	14.45	14.27
Superficial	43	-5.04	10.32	10.87
Project type				
Commercial	20	-1.24	14.90	15.09
Health	18	4.07	13.29	12.77
Apartment	21	-3.95	11.00	11.45
Education	12	-7.20	12.89	13.89
Other	18	-2.13	11.39	11.64
Project size (value)				
<$60m	18	0.50	13.39	12.75
$60-100m	23	-1.80	12.92	13.15
$100-250m	26	-1.39	14.37	14.57
>$250m	22	-4.12	11.25	11.73
Project size (area)				
<5000m2	17	-5.40	12.86	13.59
5000-10000m2	23	-0.43	13.29	13.34
10000-30000m2	27	-0.55	12.66	12.73
>30000m2	19	-1.76	13.42	13.80
Nature				
New work	74	-1.64	12.42	12.63
Alterations	15	-2.51	15.79	16.20
Client type				
Private, experienced	29	-5.60	10.57	11.20
Private, inexperienced	33	-1.39	14.35	14.55
Public, primary	18	3.46	10.84	10.47
Public, secondary	9	-1.49	16.24	16.49

Table 1: Forecasting errors summarised

STAT. MULTIPLE REGRESS.	Regression Summary for Dependent Variable: PCTERR R= .38122202 R2= .14533023 Adjusted R2= .13515559 F(1,84)=14.284 p<.00029 Std.Error of estimate: 12.039					
N=86	BETA	St. Err. of BETA	B	St. Err. of B	t(84)	p-level
Intercpt			8.381499	2.977280	2.81515	.006073
INTENSIT	-.381222	.100869	-.000772	.000204	-3.77936	.000293

Table 2: Price intensity regression

Independent variable	**Dependent variable (sd deleted residual)**	
	Percentage	**Difference**
None	*12.95*	*2030.7*
Raw forecast price	12.99	2025.1
Log forecast price	13.33	2065.3
Inverse forecast price	15.96	2381.1
Raw floor area	12.91	2084.1
Log floor area	13.33	2088.9
Inverse floor area	13.43	2090.9
Forecasted price intensity	13.17	2029.6
Nature	13.35	2079.1
Method	12.89	2051.9
Private, experienced	12.95	2057.8
Private, inexperienced	13.26	2085.3
Public, primary	12.95	2022.9
Commercial	13.25	2077.4
Health	12.92	2019.8
Apartment	13.16	2065.3

Table 3: Results of cross validation analysis

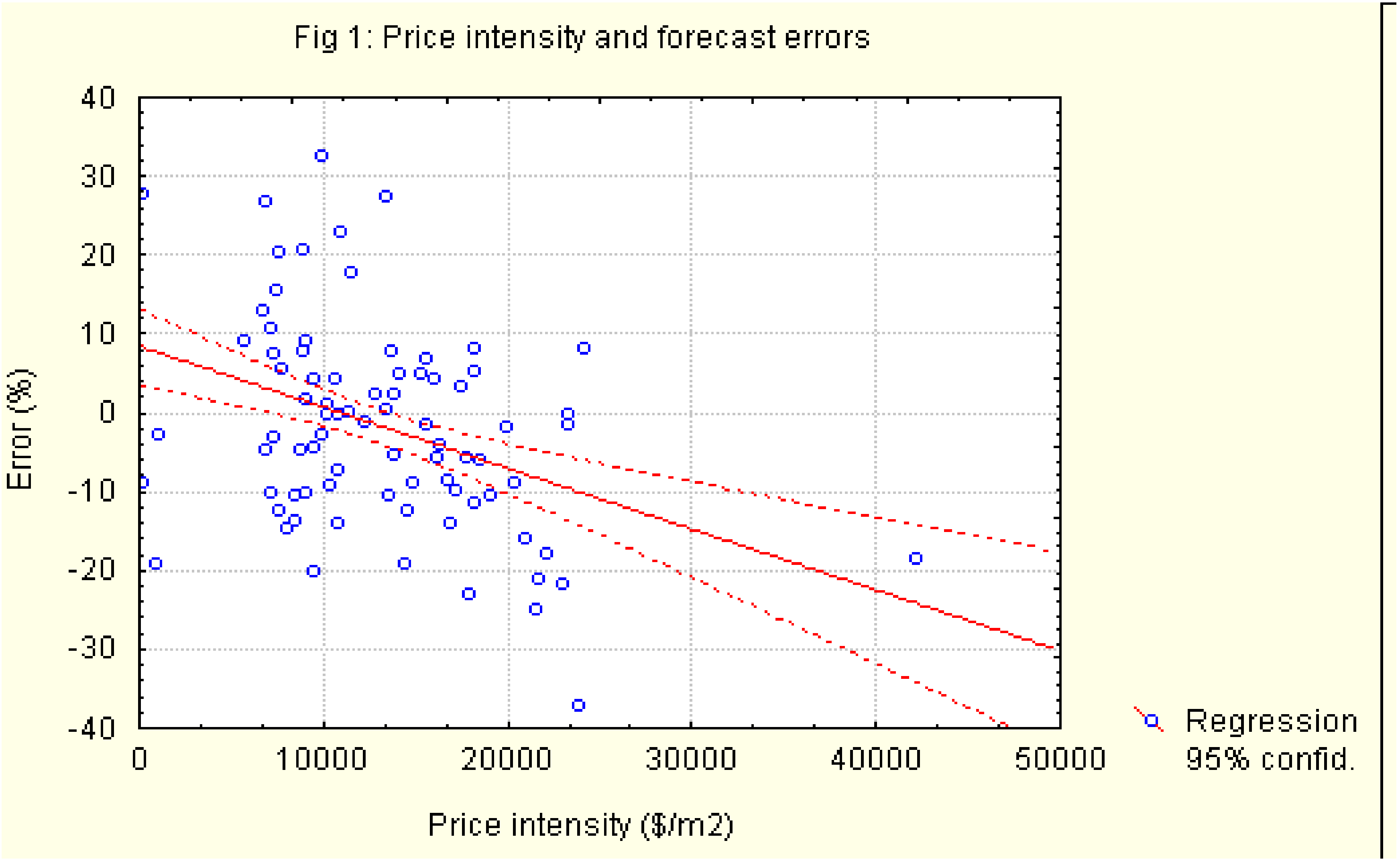
Fig 1: Price intensity and forecast errors
40
30
20
10
0
-10
-20
-30
-40
Error (%)
0
10000
20000
30000
40000
50000
Price intensity ($/m2)
Regression
95% confid.