VGTU
LEIDYKLA
TECHNIKA

A TAXONOMIC FIELD INVESTIGATION INTO INDUCED BIAS IN RESIDENTIAL REAL ESTATE APPRAISALS

Julian DIAZ III [1] ✉ and J. Andrew HANSZ [2]

[1] *Richard Bowers and Company Professor of Real Estate, Chair and Professor, Department of Real Estate, J. Mack Robinson College of Business, Georgia State University, University Plaza, Atlanta, GA 30303-3083, USA*
E-mail: jdiaz@gsu.edu ; Tel.: (404) 413-7723

[2] *Professor and Director of the Gazarian Real Estate Center, Department of Finance and Business Law, Craig School of Business, California State University, Fresno, 5245 N. Backer Avenue PB7, Fresno, CA 93740-8001, USA*
E-mail: hansz@gazarian.info ; Tel.: (559) 325-4732

ABSTRACT. A taxonomic approach to field research was developed and utilized to support empirical and experimental research findings into the impact that incentives/pressures to over-value have on systematic valuation bias. An expected no-bias population was defined and valuation judgments from actual, real-world appraisals were statistically tested against it. The judgments of appraisers presented with no incentive/pressure to over-value were consistent with the no-bias population, while the judgments of appraisers presented with incentive/pressure to over-value were significantly incompatible with the defined no bias population.

1. INTRODUCTION

This research, conducted and documented in 2004, demonstrated the existence of incentives/pressures for residential appraisers to provide favorable valuations. Residential mortgage originators, whose compensation is contingent upon originating loans, have an incentive to influence appraisers, and this study provided strong evidence and loud warning well in advance of the Liquidity Crisis of 2008 when such irresponsible agency behavior exacted terrible costs on worldwide financial markets. Because appraisals are used to estimate borrowers' equity, over-valuation of collateral results in under-estimation of default risk. Originators typically pass on this inflated default risk by selling residential originations to secondary mortgage market agents, such as Fannie Mae and Freddie Mac. By reaping the benefits while bearing little of the costs, loan originators have a strong motivation to quietly maintain a system of appraisal incentives/pressures that frustrates underwriting standards, overprices mortgage-backed assets, fuels speculative bubbles, and misrepresents the level of risk assumed by the investing community thereby setting markets up for the next financial disaster.

The question of whether or not incentives/pressures to over-value result in systematic valuation bias has been studied both empirically

and experimentally. Evidence of bias emerged, but methods are not above question. Empirical studies rely upon databases that are left-tail truncated. They contain no information on appraisals associated with rejected loans. Experimental studies are always vulnerable to the criticism that no matter the level of fealty to real world conditions, laboratory settings are artificial and render generalization of results problematic.

This paper introduces a taxonomic field study approach that does not suffer from the shortcomings described above. An expected, no-bias population of subjective value judgments is defined, and the results of actual, real-world appraisals are statistically tested against the constructed no-bias population. While the findings of this work are dated, the innovative method outlined is as relevant today as it was when the paper was circulated within the academic publishing community in 2004 and 2005 and rejected. Its utility for critical research and essential oversight is the motivation behind its resurrection here.

The balance of this paper continues with a brief review of the relevant literature that stimulates the driving research hypothesis. The specification of a model of subjective value judgment precedes the development of data and a statistical test to examine the research hypotheses. Finally results and conclusions are offered.

2. LITERATURE REVIEW AND RESEARCH HYPOTHESIS

The issue of client pressure and bias in valuation judgment has been addressed both empirically and experimentally. For example, using a database of 600,000 residential mortgages purchased by Fannie Mae in 1993, Chinloy, Cho and Megbolugbe (1997) and Cho and Megbolugbe (1996) discovered that 95% of the appraised values were greater than or equal to the pending sale price. This result was perhaps

less than surprising given that low appraisals result in loan rejection or renegotiation and rejected loans were not represented in this database (the truncated distribution problem). Noordewier, Harrison and Ramagopal (2001) investigated the relationship between default risk and over-valued collateral with a database of 1,428 residential loans from the portfolio of a national mortgage lender and concluded that loans on properties valued above the sale price of all comparable properties used in the appraisal exhibited increased default risk. These studies suggested that over-valuation may be chronic and associated with increased default risk, but they provide no link between appraisal judgment and bias from incentives/pressures to do business.

Experimental research has been introduced into real estate to help explain the behavior of market participants such as appraisers and hence is sometimes called behavioral research. One stream of this literature, reviewed in Diaz (2002), concluded that appraisers can be subject to many potentially biasing influences including the value opinions of others (Diaz, 1997), their own previous value opinions (Diaz and Wolverton, 1998), unclosed contract prices on subject and comparable properties (Diaz and Hansz, 2001), and the pressures of clients. Of particular relevance to the development of a research hypothesis for the present study was the research into client influences and pressures.

Appraisers revealed on the postal survey/experiment by Kinnard, Lenk and Worzala (1997) that they felt pervasive client pressure and had a tendency to succumb to it especially when exerted by important clients and regardless of the size of the desired adjustment. Levy and Schuck (1999) found that both sophisticated pressure, based on the use of property and market information, and unsophisticated pressure, based on the threat of withholding fee payments or future assignments, were applied to appraisers. In a survey conducted by

Wolverton and Gallimore (1999), appraisers responded that while they viewed their own role as estimating market value, they believed that clients (lenders) viewed the appraiser's role as validating a pending sale price. Hansz and Diaz (2001) uncovered an asymmetric response to transaction price feedback. When appraisers were told that their value estimate on a previous appraisal was "too high" (regardless of whether it really was or not), appraisers did not adjust their next valuation on an unrelated assignment. However, when appraisers were told that their value estimate on a previous appraisal was "too low" (regardless of whether it really was or not), appraisers responded by adjusting upward their next valuation on an unrelated assignment. While acknowledging the possibility of other causes, the authors interpreted this result as a routinized response to pervasive agent-client concerns.

These investigations, just described, were surveys or experiments conducted under laboratory conditions and, as argued earlier, were subject to the criticism that their conclusions may or may not reflect actual behavior in real world settings. Nevertheless they remain highly suggestive and stimulated the research hypothesis that the judgment of an appraiser can be influenced by agent-client concerns to the extent that the produced value estimate is not recognizable as coming from a population of uninfluenced value estimates. To further the investigation of this research hypothesis, a model of subjective value judgment was specified.

3. A MODEL OF SUBJECTIVE VALUE JUDGMENT

A model of subjective value judgment was initiated by defining the following normally distributed random variable, subjective value judgment, J, of property p at period n by appraiser i:

$$J_{pni} \mid I_{n-t} \tag{3.1}$$

where: $E(J_{pni} \mid I_{n-t}) = \mu_J$; $\sigma = \sigma_J$; I_{n-t} = historical information available at period n

A potentially upward biasing contaminant, C, was introduce into this random variable.

$$J_{pni} \mid I_{n-t}, C \tag{3.2}$$

where: if $C = 0$ (has no contaminating impact) then $E(J_{pn}i \mid I_{n-t}, C) = \mu_J$; $\sigma = \sigma_J$; but if $E(C) > 0$ (has a contaminating impact) then $E(J_{pni} \mid I_{n-t}, C) > \mu_J$; $\sigma \geq \sigma_J$;

The objective market value of property p at time n given historical information I_{n-t} was represented as

$$E(V_{pn} \mid I_{n-t}) \tag{3.3}$$

A population of residuals, R, was defined as

$$J_{pni} \mid I_{n-t} - E(V_{pn} \mid I_{n-t}) = R_{pni} \tag{3.4}$$

where: R_{pni} = a normally distributed random variable with $E(R_{pni}) = 0$; $\sigma = \sigma_J$.

Similarly, a population of potentially biased residuals, B, was defined as

$$J_{pni} \mid I_{n-t}, C - E(V_{pn} \mid I_{n-t}) = B_{pni} \tag{3.5}$$

where: B_{pni} = a random variable of unknown functional form and if $C = 0$ (has no contaminating impact) then $E(B_{pni}) = 0$; $\sigma = \sigma_J$; but if $E(C) > 0$ (has a contaminating impact) then $E(B_{pni}) > 0$; $\sigma \geq \sigma_J$.

The research hypothesis that appraisers can be influenced by agent-client concerns to the extent that produced value estimates were not recognizable as coming from a population of uncontaminated value estimates was stated in terms of the empirical model.

$$E(J_{pni} \mid I_{n-t}, C) > E(V_{pn} \mid I_{n-t}) =$$
$$E(J_{pni} \mid I_{n-t}) \tag{3.6}$$

or equally

$$E(B_{pni}) > 0 = E(R_{pni}) \tag{3.7}$$

This research hypothesis was tested by sampling from the random variable B_{pni}. If the observation was significantly different from 0, the research hypothesis was supported. Similarly sampling from R_{pni} should yield results not significantly different from 0. The research hypotheses and supporting test hypotheses for B_{pni} and R_{pni} were constructed below. Note that research hypothesis 3.7a is supported by rejecting equation 3.8 while research hypothesis 3.7b is supported by failing to reject equation 3.10. Also note that research hypothesis 3.7a was examined using a 1-tailed test, whereas research hypothesis 3.7b was examined using a 2-tailed test.

$$\text{Research Hypothesis: } E(B_{pni}) > 0 \qquad (3.7a)$$

$$H_o: \ E(B_{pni}) \leq 0 \qquad\qquad\qquad (3.8)$$

$$H_a: \ E(B_{pni}) > 0 \qquad\qquad\qquad (3.9)$$

$$\text{Research Hypothesis: } E(R_{pni}) = 0 \qquad (3.7b)$$

$$H_o: \ E(R_{pni}) = 0 \qquad\qquad\qquad (3.10)$$

$$H_a: \ E(R_{pni}) \neq 0 \qquad\qquad\qquad (3.11)$$

4. DATA

4.1. Observations from appraisal populations

The property selected was a 1900 square foot, one-story residence located in Arlington, Texas. Neighborhood boundaries were well defined and the general area was known as a mature and stable residential market. The subject dwelling, built in 1960, was typical of the residential properties in the area. Property ownership most recently transferred in January 2002 for $122,000 ($64.21 per square foot). At time of purchase, the property was appraised for $123,000. After settlement, the property received cleaning and some maintenance (most notably, partial new interior paint).

To obtain observations of $(J_{pni} \,|\, I_{n-t})$ and $(J_{pni} \,|\, I_{n-t}, C)$ six appraisers were hired, in pairs, over a five month period and asked to appraise the subject property. The first pair inspected/valued the property on the 7th and 8th of October 2003, the next pair inspected/valued the property on the 22nd and 23rd of December 2003, and the final pair inspected/valued the property on 1st and 2nd of March 2004. All appraisers were contacted by the homeowner and randomly selected from a list of local appraisers. For each pair, one appraiser (hereafter referred to as a "no pressure" or "control" appraiser) was hired to estimate the value of the property for decision making purposes. No further directions were given to the control appraiser. The second appraiser (hereafter referred to as the "pressured" or "treatment" appraiser) was informed that the homeowner required an appraised value of at least $150,000 to secure a home-equity loan. To avoid overlapping inspections, appraisers in a pairing inspected the subject property on consecutive days. Inspections lasted between 30 and 60 minutes and included interior and exterior examinations, exterior measurement, and photographs. The property was in identical physical condition for each pairing.

In summary, each appraiser in an appraiser pairing valued the same property, at the same point in time (within one day of each other), and was unaware of the other appraisers. The only structured difference was an agent-client concern as only the treatment appraisers had knowledge that the homeowner required a value of at least $150,000. Hard copies of all appraisals were obtained and analyzed and Table 1 provides an overview of each report.

Table 1. Overview of appraisals

	Valuation pair 1		Valuation pair 2		Valuation pair 3	
	Control (no pressure)	Treatment (pressure)	Control (no pressure)	Treatment (pressure)	Control (no pressure)	Treatment (pressure)
Pressure treatment point	None	$150,000+	None	$150,000+	None	$150,000+
Inspection/valuation date	October 7, 2003	October 8, 2003	December 22, 2003	December 23, 2003	March 2, 2004	March 1, 2004
Appraisal fee	$325	$325	$325	$350	$325	$350
Value estimate	$128,000 ($67.37/sf)	$156,000 ($82.11/sf)	$124,000 ($65.26/sf)	$150,000 ($78.95/sf)	$137,000 ($72.11/sf)	$150,000 ($78.95/sf)
Oldest comp. sale	2/2003	5/30/2002	8/2003	5/2003	8/2003	10/2003
Most recent comp. sale	8/2003	9/26/2003	10/2003	10/2003	1/2004	2/2004
Number of comp. sales	4	3	3	3	3	3
Highest comp. sale	$150,000	$165,500	$136,972	$175,000	$136,972	$175,000
Lowest comp. sale	$123,900	$150,500	$120,000	$136,972	$106,000	$135,000
Range (comp. sales)	$26,100	$15,000	$16,972	$38,028	$30,972	$40,000
Avg. comp. sale price	$131,562	$158,000	$126,957	$156,657	$122,291	$149,333
St. deviation of comps	$12,351	$7,500	$8,890	$19,050	$15,549	$22,279
Avg. # of adjustments	1.5/sale	1.7/sale	5.0/sale	2.0/sale	3.7/sale	4.7/sale
Avg. gross adjustments*	$10,078/sale	$11,733/sale	$21,100/sale	$5,850/sale	$10,917/sale	$23,133/sale
Avg. net adjustments**	-$3,348/sale	-$5,200/sale	-$2,900/sale	+$4,517/sale	+$1,783/sale	+$200/sale
Highest adj. comp. sale	$130,650	$157,700	$124,972	$160,950	$137,472	$155,100
Lowest adj. comp. sale	$125,500	$155,100	$123,100	$136,472	$113,970	$146,500
Range (adjusted sales)	$5,150	$2,600	$1,872	$24,478	$23,502	$8,600
Avg. of adj. comp. sales	$128,215	$156,267	$124,057	$152,141	$124,074	$149,533
St. dev. of adj. comp. sales	$2,655	$1,320	$937	$13,604	$12,092	$4,827

* For each appraisal, the absolute values of all adjustments were summed and divided by the number of comparable sales.

** For each appraisal, all adjustments were summed and divided by the number of comparable sales.

4.2. Overview of valuation pairs

4.2.1. Valuation pair 1 (October 2003)

The difference between the two value estimates (control and treatment) was $28,000 ($14.74 per square foot). The control appraiser's value estimate was $128,000 ($67 per square foot), a 4.9% increase over the January 2002 sales price of $122,000. The pressure treatment appraiser's value estimate was $156,000 ($82 per square foot), a 27.9% increase over the January sale price. In the sales comparison approach, the control appraiser used four comparable sales ranging in value from $123,900 to $150,000. The treatment appraiser used three comparable sales with a higher price range from $150,500 to $165,500. Both appraisers bracketed their value estimate, that is, each appraiser selected a set of comparable sales whose transaction price range contained his value judgment.

4.2.2. Valuation pair 2 (December 2003)

The difference between the two values was $26,000 ($13.68 per square foot). The control appraiser's value estimate was $124,000 ($65 per square foot), a 1.6% increase over the January 2002 sales price of $122,000. The pressure treatment appraiser's value estimate was $150,000 ($79 per square foot), a 23.0% increase over the January sale price. In the sales comparison approach, the control appraiser used three comparable sales ranging in value from $120,000 to $136,972. The treatment appraiser used three comparable sales with a higher price range from $136,972 to $175,000. Again, both appraisers bracketed their value estimate.

4.2.3. Valuation pair 3 (March 2004)

The difference between the two values was $13,000 ($6.84 per square foot). The control appraiser's value estimate was $137,000 ($72 per square foot), a 12.3% increase over the January 2002 sales price of $122,000. The pressure treatment appraiser's value estimate was $150,000 ($79 per square foot), a 23.0%

increase over the January sale price. In the sales comparison approach, the control appraiser used three comparable sales ranging in value from $106,000 to $136,972. The treatment appraiser used three comparable sales with a higher price range from $135,000 to $175,000. The control appraiser's final value estimate was slightly above the unadjusted comparable sale price range and the treatment appraiser did bracket his value estimate.

In summary, six valuations, in three pairings, produced a range in appraised values from $124,000 to $156,000 and it appeared that the pressure treatment did have an influence on appraiser judgment. However, each valuation estimates was an appraiser's personal opinion. To evaluate these subjective judgments it is first necessary to estimate the characteristics of a population of uncontaminated value estimates by examining the characteristics of the population of objective market value estimates (equation 3.6).

4.3. Objective market value estimate

Rosen (1974) argued that the value of any asset was the sum of the value of the asset's components. He has been credited with contributing to early hedonic pricing theory and regression based hedonic modeling has been a dominate research paradigm in real estate research for three decades (see Cho, 1996 for a more contemporary survey of theoretical and empirical issues in housing price estimation).

In this present study, the objective market value of property p at time n, $E(V_{pn} | I_{n-t})$, was hedonically defined as a function of the property's unique bundle of characteristics. Regressing a vector of transaction prices against the set of property characteristics produces an estimation model as specified in equation (4.1).

$$Y_{pn} = \beta_0 + \sum \beta_k X_{kpn} + \varepsilon_{pn} \qquad (4.1)$$

where: Y_{pn} was the transaction price of property p at time n; β_k was a vector of estimated

partial regression coefficients on the property's structural, quality, and site characteristics (independent variables), X_{kpn}; and ε_{pn} was the normally distributed random error term with mean 0 and variance σ^2.

A data set of 321 single-family residential sales was collected from the local multiple listing service (MLS). Because Texas is a non-disclosure state and property sale prices are not reported in public records, residential appraisers rely exclusively on the MLS for comparable sale data. All MLS sales collected were completed between March 3, 2003 and March 2, 2004. This time period was selected because it represents one-year prior to the most recent appraiser pairing on March 2, 2004. Appraisers search for comparable sales, typically, within a six-month period from the valuation date and usually not more than one-year. The data was also limited by sale prices ranging from $75,000 to $250,000, which was the most common neighborhood price range defined by the appraisers, and a geographic area judged to be the subject property's neighborhood, also as described in the appraisal reports. Therefore, this MLS data set represented a broad universe of comparable sale data available to the appraisers.

Independent variables were selected based on interviews with knowledgeable market participants, including brokers and sales agents, in this area and also from adjustments made in the submitted appraisal reports. The dependent variable and nine independent variables were coded and summarized in Table 2.

Table 2. Variable definitions and descriptive statistics

Symbol	Variable	Description	n	Minimum	Maximum	Mean	Std. deviation
Y	Dependent variable: Sale Price	Sale price in dollars	321	75000	250000	124802.24	38874.20
	Independent variables:	*Structure characteristics*					
X_1	Size	Dwelling size in square feet	321	906	3593	1823.31	563.99
X_2	No garage	Binary: 1 if no garage; 0 otherwise	321	0	1	.1246	.3308
X_3	2 stories	Binary: 1 if two stories; 0 otherwise	321	0	1	.1526	.3602
		Quality characteristics					
X_4	Time	Month of sale coded from 0 (most recent) to -12 (oldest)	321	−12	0	−6.0997	3.4218
X_5	Age	Age of improvements in years	321	0	66	36.07	12.91
X_6	Age squared	Age of improvements in years squared	321	0	1467.12	1467.12	789.52
X_7	Lamar school district	Binary: 1 if property located in Lamar High School district; 0 otherwise	321	0	1	.1838	.3879
		Site characteristics					
X_8	Large lot	Binary: 1 if lot is greater than or equal to .5 acre; 0 otherwise	321	0	1	.0748	.2634
X_9	Pool	Binary: 1 if lot has a pool; 0 otherwise	321	0	1	.1246	.3308

Size (X_1) in square feet was expected to be positively related to sale prices. Other size related variables, including the number of bedrooms, bathrooms, dinning rooms, and living areas, were considered, however, these variables were highly correlated with "size." The majority of the properties had either a two car garage or no garage with just a few one and three-car garages. Because a two-car garage was the market standard or baseline, a binary variable indicating properties with "no garage" (X_2) was coded and expected to be negatively related to sale prices. The dominate house style in this market was one-story (85%), similar to the subject property. Above average marketing periods for two-story properties were evident from MLS statistics and a potential two-story (X_3) property price discount was suspected.

The data were arrayed based on month of sale and coded from 0 (the most recent sale month) to –12 (the oldest sale month). In interpreting the "time" (X_4) variable, a positive coefficient would indicate increasing property values and a negative coefficient would indicate declining values. Local brokers reported relatively flat to slightly increasing property values in this neighborhood and the "time" variable expectation was positive but modest.

The age (X_5) of the property, in years, was a proxy for property condition. It was anticipated that age would be negatively associated with sale prices, as older properties were typically in inferior condition. Age squared (X_6) was included to model the decreasing marginal impact as age increases. Most sales (82%) were located in the Arlington High School district with the remaining sales located in the Lamar High School district. Local brokers and property owners indicated a slight preference for the Lamar school district (X_7), all other factors being equal, and a modest price premium was anticipated.

About 7.5% of the sales had lot sizes of a half acre or more. These "large lot" sales may potentially sell at premiums due to the benefit of excess land. Therefore, a positive relationship

was anticipated between the "large lot" variable (X_8) and sale prices. Finally, pools (X_9) were considered a desirable amenity in this sultry climate and a price premium was anticipated. The regression estimates were as follows (probability values reported in parenthesis):

$$\hat{y}_{pn} = 50957.43 + 57.47X_1 - 10998.58X_2 -$$
$$\quad (.000) \qquad (.000) \qquad\qquad (.000)$$

$$5059.65X_3 + 86.89X_4 - 1525.17X_5 +$$
$$\ (.054) \qquad\quad (.730) \qquad\quad (.000)$$

$$17.27X_6 - 2162.37X_7 + 6353.05X_8 +$$
$$\ (.000) \qquad\quad (.341) \qquad\quad (.055)$$

$$10626.66X_9 \qquad\qquad\qquad\qquad\qquad (4.2)$$
$$\ (.000)$$

Overall model characteristics were acceptable with an F-statistic of 197.463 (p-value of .000), an R^2 of .851, an adjusted-R^2 of .847,[1] and a standard error of the estimate of 15,217.86.[2] Examinations of the model residuals revealed no evidence of hedonic assumption violations. A vital concern in structural modeling is multicollinearity. Because the purpose of this model was mean estimation rather than structural modeling, correlation among the independent variables was a secondary concern. Regardless, multicollinearity did not appear to have a strong influence on variable coefficients.[3]

---

[1] For a point of reference, Mark Linne, M. Steven Kane, and George Dell in *A Guide to Appraisal Valuation Modeling* published by the Appraisal Institute in 2000 explain that a regression model used for real estate valuation with an R2 greater than 60% is a powerful model as a general rule of thumb (see page 50).

[2] The Durbin-Watson test statistic was 2.014. A test statistic of 2 indicates no autocorrelation in the residuals (see page 762 of *Basic Business Statistics* 7th *Edition* by Mark Berenson and David Levine).

[3] All variance inflation factors were low (ranging from 1.02 for "time" to 1.51 for "size") with the anticipated exceptions of the age and age squared combination (13.99 and 14.43, respectively).

The model coefficients were generally as anticipated and statistically significant with two notable exceptions. The market conditions trend variable X_4 "time" and the location variable X_7 "Lamar school district" were not statistically different from zero (with p-values of .730 and .341, respectively). The insignificance of the "time" variable confirmed anecdotal evidence of a trend-less market during the study period. Therefore, the same objective value estimate (at $X_4 = 0$) was used for the statistical evaluation of each pairing. Also noteworthy, variables X_8 "large lot" and X_3 "2 stories" were marginally significant with p-values of .055 and .054, respectively. The derived multiple regression equation was used to calculate the estimate of the objective market value of the subject property,[4] $E(V_{pn}|I_{n-t})$, at 126,503.74 with a standard deviation ($s\{\hat{y}_{pn}\}$) of 2,036.35. The next section describes the statistical tests.

5. STATISTICAL TESTS

Given the population characteristics for R specified in equation (3.4), the following random variable, used as a test statistic, was defined:

$$z = R_{pri}/\sigma_J. \tag{5.1}$$

Note that $R_{pri} = (J_{pni}|I_{n-t}) - E(V_{pn}|I_{n-t})$. Since $E(V_{pn}|I_{n-t})$ was unknown, it must be estimated using $\hat{y}_{pn}$. This introduces an additional element of variability estimated by $s\{\hat{y}_{pn}\}$, the standard deviation of $\hat{y}_{pn}$. Also note that with σ_J unknown, an estimate, s_J, was

used and the random variable, defined below, becomes t distributed.

$$t = R_{pri}/((s_J^2 + s\{\hat{y}_{pn}\}^2)^{.5}) \tag{5.2}$$

With estimates of $\hat{y}_{pn}$ and $s\{\hat{y}_{pn}\}$ provided by the regression analysis, only s_J, the estimate of the standard deviation of subjective value estimates was needed to fully define the test statistic. Geltner and Miller (2001) reviewed several studies investigating appraisal dispersion in both residential and commercial appraisals. They found an average error magnitude (standard deviation as a percent of average estimate of property value) on the order of 5% to 10% and noted that the higher end of the range would be appropriate for thinly traded, unique assets. Variation in residential valuations should therefore be located in the lower end of this range with more thinly traded commercial markets representing the higher end of the range. This argument was supported by the valuation data from this present study. The average error under the conditions of the null, the three control (no pressure) appraisals, was 5.1%.[5] Based on the Geltner and Miller argument and the valuation data from this present study, the lower estimate of standard deviation was more likely. However, both the upper and lower limits of the estimate were used to examine test hypotheses. This estimated range of 5% to 10% for the magnitude of the standard deviation suggested a range for s_J from 6,325.19 to 12,650.37.

The test statistic, equation (5.2), was fully defined to operationalize both sets of test hypotheses (3.8), (3.9) and (3.10), (3.11) and examine research hypotheses (3.7a) and (3.7b). The test statistic calculations are shown for all three valuation pairings.

---

[4] The subject property is 1900 square feet in size ($X_1 = 1900$), has a garage ($X_2 = 0$), one-story ($X_3 = 0$), 43 years old ($X_5 = 43$, $X_6 = 1849$), located in Arlington High School district ($X_7 = 0$), does not have a large lot ($X_8 = 0$), and does not have a pool ($X_9 = 0$). Note that the coefficients in equation 4.2 are rounded.

[5] Standard deviation of 6,658 divided by mean of 129,667.

5.1. Paired valuation test statistics

5.1.1. Paired valuations 1 (October 2003)

To examine research hypothesis (3.7a): Conservative test statistic:

$29,496.26/((12,650.37)^2 + (2,036.35)^2)^{.5} = 2.302$

p-value = .011

More realistic test statistic:

$29,496.26/((6,325.19)^2 + (2,036.35)^2)^{.5} = 4.439$

p-value = .000

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was remote. The null hypothesis was rejected. The value opinion of the pressured appraiser was biased by the introduced influence.

To examine research hypothesis (3.7b): Conservative test statistic:

$1,496.26/((12,650.37)^2 + (2,036.35)^2)^{.5} = .117$

p-value = .907

More realistic test statistic:

$1,496.26/((6,325.19)^2 + (2,036.35)^2)^{.5} = .225$

p-value = .822

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was very high. The null hypothesis was not rejected. The value opinion of the control (no pressure) appraiser came from the population of uninfluenced value opinions.

5.1.2. Paired valuations 2 (December 2003)

To examine research hypothesis (3.7a): Conservative test statistic:

$23,496.26/((12,650.37)^2 + (2,036.35)^2)^{.5} = 1.834$

p-value = .034

More realistic test statistic:

$23,496.26/((6,325.19)^2 + (2,036.35)^2)^{.5} = 3.536$

p-value = .000

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was remote. The null hypothesis was rejected. The value opinion of the pressured appraiser was biased by the introduced influence.

To examine research hypothesis (3.7b): Conservative test statistic:

$-2,503.74/((12,650.37)^2 + (2,036.35)^2)^{.5} = -.195$

p-value = .845

More realistic test statistic:

$-2,503.74/((6,325.19)^2 + (2,036.35)^2)^{.5} = -.377$

p-value = .707

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was very high. The null hypothesis was not rejected. The value opinion of the control (no pressure) appraiser came from the population of uninfluenced value opinions.

5.1.3. Paired valuations 3 (March 2004)

To examine research hypothesis (3.7a): Conservative test statistic:

$23,496.26/((12,650.37)^2 + (2,036.35)^2)^{.5} = 1.834$

p-value = .034

More realistic test statistic:

$23,496.26/((6,325.19)^2 + (2,036.35)^2)^{.5} = 3.536$

p-value = .000

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was remote. The null hypothesis was rejected. The value opinion of the pressured appraiser was biased by the introduced influence.

To examine research hypothesis (3.7b):
Conservative test statistic:

$10,496.26/((12,650.37)^2 + (2,036.35)^2)^{.5} =$ .819

p-value = .413

More realistic test statistic:

$10,496.26/((6,325.19)^2 + (2,036.35)^2)^{.5} =$ 1.580

p-value = .115

Conclusion: The probability that the observed result could occur under the conditions of the null (no influence on the appraiser) was high. The null hypothesis was not rejected. The value opinion of the control (no pressure) appraiser came from the population of uninfluenced value opinions.

5.2. Examination of joint events

Each of the three pressure treatment appraisals can be conceptualized as a Bernoulli trial with two possible mutually exclusive outcomes, reject as coming from an unbiased population of valuation judgments versus fail to reject. Because each trial is independent of the outcome of the other trials and because the probability of rejecting the outcome of any one trial, p, can be held constant from one trial to the next, the series of three pressure treatment appraisals is modeled as a Bernoulli process whose probabilities are given by the binomial distribution. Setting p equal to .05 and .01 results in the following probability distributions for a process of three trials:

Probability of rejecting a single trial (p)
.05   .01

Probability of three rejects in three trials:
.000125   .000001

Probability of two rejects in three trials:
.007125   .000297

Probability of one reject in three trials:
.135375   .029473

Probability of no rejects in three trials:
.857375   .970299

Total
1.000000   1.000000

Using the conservative test statistic (high variability) and assuming the conditions of the null, that is that the pressure treatment has no biasing impact on valuations, the probability of an appraised value estimate as extreme or more extreme than the first pressure treatment estimate is .011, the second .034, and the third .034. The conservative joint probability of three valuation estimates at least as extreme as the three actually obtained is therefore .000013. With realistic individual probabilities of .000006, .0002, and .0002, the realistic joint probability is $2.4E - 13$. Comparing these results to the probability distributions above offers strong evidence that the pressure treatment valuation estimates did not come from an unbiased population.

5.3. Recapitulation

Table 3 provides a summary of the obtained value estimates and selected statistics. Over a five month period the same single family residence was valued by three pairs of appraisers. The date of valuation varied between pairs but was constant within pairs. Significant effort was extended to prevent any one appraiser discovering that the property was being valued by other appraisers. The only notable difference between paired appraisers was an incentive/pressure treatment. One appraiser per pair, selected at random, was informed that the homeowner required a value estimate of at least $150,000. The second appraiser (control) in each valuation pair was simply asked to value the property for decision making purposes.

Table 3. Summary of object value estimate and statistical tests

Summary of objective value estimate and standard deviation			
Objective value estimate ($\hat{y}_{pn}$):			
Observations	321	$E(V_{pn} \mid I_{n-t})$	126,503.74
adj-R^2	.847	$s\{y_{pn}\}$	2,036.35
Magnitude of standard deviation:			
Realistic (5%)	6,325.19	Conservative (10%)	12,650.37
Summary of statistical tests			
Valuation pair 1 (October 2003):			
Control (no pressure)		**Treatment (pressure)**	
r_{pn1}	1,496.26	b_{pn1}	29,496.26
Conservative stat.	0.1168	Conservative stat.	2.3020
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.9071	p-value	0.0110
Realistic stat.	0.2252	Realistic stat.	4.4389
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.8220	p-value	0.000006
Valuation pair 2 (December 2003):			
Control (no pressure)		**Treatment (pressure)**	
r_{pn1}	−2,503.74	b_{pn1}	23,496.26
Conservative stat.	−0.1954	Conservative stat.	1.8338
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.8452	p-value	0.0338
Realistic stat.	−0.3768	Realistic stat.	3.5360
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.7066	p-value	0.0002
Valuation pair 3 (March 2004):			
Control (no pressure)		**Treatment (pressure)**	
r_{pn1}	10,496.26	b_{pn1}	23,496.26
Conservative stat.	0.8192	Conservative stat.	1.8338
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.4133	p-value	0.0338
Realistic stat.	1.5796	Realistic stat.	3.5360
Deg of freedom	320	Deg of freedom	320
Tailed test	2	Tailed test	1
p-value	0.1152	p-value	0.0002

Not only were the value estimates provided by the appraisers receiving the pressure treatment numerically different from estimates provided by the non-pressured, control appraisers, but the value estimates from the pressured appraisers were statistically different from a defined population of unbiased, objective values while the control valuations were not. Considering the three independent pressure treatment valuations as a Bernoulli process also produces strong statistical evidence that the pressure treatment valuations do not come from a population of unbiased valuations estimates and that there is a significant biasing contaminant.

It is also worth noting that the treatment applied to the appraisers in the present study, a clear request by the homeowner/client, is less severe than other possible pressure options available to purchasers of appraisal services. Levy and Schuck (1999) discussed several techniques clients may use to pressure appraisers ranging from withholding fees to threatening to withhold future business. To conclude that stronger incentives/pressures would evoke results similar to those in this study is tempting.

6. CONCLUSIONS

The case that appraisers are subject to the influences of their clients and that under certain conditions these influences result in biased appraisals is established. This study presents statistical evidence that agent-client concerns influenced the reported value estimates offered by independent appraisers hired to provide unbiased valuation judgments. Conversely, hired appraisers who were not exposed to agent-client concerns provided estimates that appeared not to be biased. These results support the findings of empirical studies that suffer from reliance upon questionable databases as well as the findings of an extensive experimental literature subjected to the criticism of uncertain generalizability. The concern

that under these conditions, default risk will be significantly underestimated and passed on to the investment community via the wholesaling channel and secondary mortgage market operations is no longer questioned. Future research is needed to understand better the role of originators in perpetuating a system that circumvents underwriting standards designed to protect the investing public, but based on the emerging evidence, the gate-keeping role of the originator-appraiser relationship is dangerously ineffectual.

Inflated or misrepresented collateral is one of the most crucial issues facing mortgage investors, purchasers of mortgage derivatives, and regulators and policy makers seeking to repair a damaged financial system and ensure its continued viability. A moral hazard certainly did exist leading up to the Liquidity Crisis of 2008 as appraisers were influenced by clients with incentives/pressures to complete mortgage financing transactions. The moral hazard will certainly not disappear and its terrible consequences will continue to be felt until regulators and policy makers find effective ways to discourage and detect the agency-client behavior revealed in this study. The method demonstrated here represents a research and policing tool to advance this quest.

A comment on methodology is also warranted. A traditional research design would require 30 or more treatment observations and 30 or more control observations from which group statistics would be calculated and compared. The ability to collect observations in some real estate market field settings is limited. The taxonomic approach offers an alternative design when data poverty exists but population characteristics are known or can be estimated. Rejection of this study when previously offered for publication in 2004 and 2005 was uniformly on the grounds that a larger sample size was required to support the results, that three paired samples represent insufficient numbers to substantiate conclusions. This criticism is

simply in error and derives from a misunderstanding of the procedure employed. The relevant sample size is the number of observations used to create the hedonic model for the population of uncontaminated (unbiased) value estimates. The 321 observations actually used are clearly sufficient to provide the procedure with adequate statistical power. Once the unbiased appraisal population is established, no further observations are needed to draw powerful statistical conclusions about whether or not a new observation belongs to it.

The procedure we employed in this study is absolutely analogous to a biologist testing the hypothesis that wolves are in a remote valley by collecting DNA from animal hair samples in the field. The biologist knows the DNA structure of the wolf population as well as for other candidate animal species. When subjected to DNA analysis, some of the collected hair may prove to be from coyotes or lynx or bear, but if one sample of collected hair proves to have the DNA structure of a wolf, no reviewer will say to this researcher that you need more samples of wolf hair before you can conclude that the hair you collected is from a wolf. Yet this was the exact criticism that prevented the timely publication of this research. Now the questions becomes, with hair samples discovered and documented to be from three separate individual wolves, what is the probability that there are no wolves in this valley? In our study of agency behavior, we demonstrated that the probability of no biased appraisal behavior was dismissingly trivial.

Finally, the results of this study, consistent with the findings of previous experimental research, extend laboratory conclusions into field settings and therefore tend to support the fecundity of experimental methods in real estate. Coupled with laboratory work, field study, such as the demonstrated taxonomic approach, offers noteworthy promise for investigation into economic questions that have proved somewhat resistant to more traditional approaches.

REFERENCES

Chinloy, P., M. Cho, M. and Megbolugbe, I. (1997) Appraisals, transaction incentives, and smoothing, *Journal of Real Estate Finance and Economics*, 14(1/2), pp. 89–111. doi:10.1023/A:1007772018106

Cho, M. (1996) House price dynamics: A survey of theoretical and empirical issues, *Journal of Housing Research*, 7(2), pp. 145–172.

Cho, M. and Megbolugbe, I. (1996) An empirical analysis of property appraisal and mortgage redlining, *Journal of Real Estate Finance and Economics*, 13(1), pp. 45–55. doi:10.1007/BF00174550

Diaz, J. (1997) An investigation into the impact of previous expert value estimates on appraisal judgment, *Journal of Real Estate Research*, 13(1), pp. 57–66.

Diaz, J. (2002) *Behavioural Research in Appraisal and Some Perspectives on Implications for Practice*, London; RICS Foundation.

Diaz, J. and Hansz, J. (2001) The use of reference points in valuation judgment, *Journal of Property Research*, 18(2), pp. 141–148. doi:10.1080/09599910110039897

Diaz, J. and Wolverton, M. (1998) A longitudinal examination of the appraisal smoothing hypothesis, *Real Estate Economics*, 26(2), pp. 349–358. doi:10.1111/1540-6229.00749

Geltner, D. and Miller, N. (2001) *Commercial Real Estate Analysis and Investments*, Prentice Hall: New Jersey.

Hansz, J. and Diaz, J. (2001) Valuation bias in commercial appraisal: A transaction price feedback experiment, *Real Estate Economics*, 29(4), pp. 553–565. doi:10.1111/1080-8620.00022

Kinnard, W., Lenk, M. and Worzala, E. (1997) Client pressure in the commercial appraisal industry: how prevalent is it?, *Journal of Property Valuation and Investment*, 15(3), pp. 233–244. doi:10.1108/14635789710184952

Levy, D. and Schuck, E. (1999) The influence of clients on valuations, *Journal of Property Investment & Finance*, 17(4), pp. 380–400. doi:10.1108/14635789910271773

Noordewier, T., Harrison, D. and Ramagopal, K. (2001) Semivariance of property value estimates as a determinant of default risk, *Real Estate Economics*, 29(1), pp. 127–159. doi:10.1111/1080-8620.00005

Rosen, S. (1974) Hedonic prices and implicit markets: product differentiation in pure competition, *Journal of Political Economy*, 82(1), pp. 34–55. doi:10.1086/260169

Wolverton, M. and Gallimore, P. (1999) Client feedback and the role of the appraiser, *Journal of Real Estate Research*, 18(3), pp. 415–432.

SANTRAUKA

GYVENAMOSIOS PASKIRTIES NEKILNOJAMOJO TURTO VERTINIMŲ TAKSONOMINIS TYRIMAS SIEKIANT NUSTATYTI VERTINIMO PAKLAIDAS

Julian DIAZ III, J. Andrew HANSZ

Taksonomijos metodas buvo išplėtotas ir pritaikytas taikomajam tyrimui siekiant patvirtinti empirinius ir eksperimentinius tyrimų rezultatus, kad paskata arba spaudimas pervertinti yra sisteminė paklaida. Buvo apibrėžta tikėtinoji nešališka imtis ir nustatytos statistinės realios vertės. Vertinimai, kuriuose nebuvo paskatos ar spaudimo pervertinti, buvo suderinti su imtimi be sisteminės paklaidos, o vertinimai su paskata ar spaudimu pervertinti buvo visiškai nesuderinti su apibrėžta imtimi be sisteminės paklaidos.