부동산학연구 제18집 제2호, 2012. 6, pp. 125～142

Journal of the Korea Real Estate Analysts Association
Vol.18, No.2, 2012. 6, pp. 125～142

Newspaper content and home prices: perception, reasoning and affect*

신문기사 내용과 주택가격: 인식, 사유, 그리고 투자심리

Jin, Changha (진 창 하)**
Gallimore, Paul***

< Abstract >

부동산 시장에 있어 시장 참여자는 다른 자산시장과는 상이한 특징을 나타낸다. 특히 수익형 부동산 시장과 달리 비전문적인 시장참여자들에 의해 시장의 행태가 영향을 받게 된다. 의사결정자들은 객관적인 정보에 의해 영향을 받을 뿐아니라 직관과 투자심리에 따라 어느정도 영향을 받게 된다. 우리는 애틀랜타 CMSA 지역을 대상으로 신문기사의 내용과 부동산 가격에 대한 관계를 연구하였다. 본 연구는 공적분 검정을 수행하였고, 그랜저 인과관계 검정(Granger's Causality Test) , 그리고 오차수정모형(Error Correction Modeling)을 이용하여 주택시장과 신문기사와의 관계성에 대해 연구를 하였다. 본 연구에서 시사하는 점은 특히 부정적인 용어의 사용이 주택시장의 가격변화에 대해 연관성이 있는 것으로 나타났다.

주 제 어 : 주택가격, 신문기사, 인지화과정, 오차수정모형, 투자심리
Keywords : Residential Home Prices, Newspaper Content, Cognitive Processing, Effect Heuristics, Error Correction Model.

* 본 논문은 2009년 미국부동산학회(America Real Estate Society)에서 발표된 박사과정 논문을 수정 보완한 연구임.

** Department of Economics, College of Economics and Business Administration, Hanyang University, 55 Hanyangdaehak-ro, Sangnok-gu, South Korea , (82)400-5602, FAX (82) 31-436-8180, e-mail cjin@hanyang.ac.kr

*** Department of Real Estate, J. Mack Robinson College of Business, Georgia State University, P.O. Box 4020, Atlanta, Georgia 30302-4020, (404) 413-7724, FAX (404) 413-7736, e-mail pgallimore1@gsu.edu

I. Introduction

This study investigates the relationship between real estate newspaper content and the behavior of individuals in residential real estate markets. We argue that newspaper articles do not merely reflect contemporaneous market fundamentals but convey, through particular kinds of words, messages that both impact individuals' perceptions of the world and are impounded into their contemporaneous decisions. We locate our study broadly in the literature of "behavioral" explanations of financial decision-making: explanations that may in some instances, or to some degree, appear anomalous or biased in relation to rational models of judgment. These explanations have been widely studied at both the institutional and individual levels. Barberis and Thaler (2003) provide a review of this field. We focus on the decisions of individuals, exploring the interplay of perception and reason, to which we add the role of "feeling" or "affect" response.

Although numerous investigations have been made of behavioral bias in stock-investment decisions, relatively few have covered decisions about transactions or pricing in residential real estate. Notable exceptions include Case and Shiller(1988) and Genesove and Mayer(2001). Likewise, although there has been much attention paid to the role of information in the pricing of securitized real estate, it has generally focused on information asymmetries(e.g. Downs, Güner and Patterson, 2000; McDonald, Nixon and Slawson, 2000) What we pursue in this paper is the conjecture that words and emotions in information publicly available in the media affect individuals' perceptions of the world and alter financial decision-making: specifically, becoming manifest in residential transaction prices. There is good reason to expect that emotive media effects exert such an influence. Relative to other markets, and notwithstanding that most buyers and sellers are guided in some respects by agents, residential real estate is dominated by occasional decision-makers relying upon unsophisticated information-evaluation strategies. Given this, it is not surprising that, for reasons that may include emotive effects, housing market efficiency is eroded(Clayton, 1998).

The two key theoretical underpinnings to the question in this study are first that people's decisions and judgments reflect not solely their reasoning, but also have an intuitive, perception-based dimension; and second that the intuitive component of people's decisions is strongly influenced by their "feeling" towards the stimuli upon which their perceptions are based. The contention that there is a perception-based dimension of judgment, that is distinct from reasoning, has been extensively studied. A relatively recent summary of the treatment these two "systems" of understanding and thinking is provided by Kahneman(2003). As described by Kahneman, cognitive processing in one system − perception/intuition − generally operates in an effortless, fast, automatic, associative, implicit and often emotionally charged manner. Cognitive

processing in the other system activates an effortful, controlled, slower, serial, more deliberate, and consciously monitored response.

Importantly, for our study, the role of intuition in judgment reflects the formulation of impressions and can be evoked by language. Thus, the second aspect of theory that informs the study is that of the "affect heuristic", as developed by Slovic and colleagues(e.g. Slovic, Finucane, Peters and MacGregor, 2002). The affect heuristic refers to reliance in decision-making on one's feeling – positively or negatively – towards a stimulus. While the stimulus of imagery has been found to show a strong relationship with decision outcomes (Finucane, Alhakami, Slovic and Johnson, 2000), language is clearly an alternative means of communicating such stimuli. While the affective response occurs rapidly and automatically, its lasting imprint may be drawn upon over a longer period in either the intuitive or reasoning mode of judgment: or both (Kahneman, 2003).

Newspapers, particularly local ones, play a central role in providing information on an ongoing basis about activity in residential real estate markets. The transient, relatively superficial nature of newspaper content, however, makes it likely to be utilized more in the intuitive component of decision-making than in controlled reasoning; and it is in the nature of journalism to stimulate readers' interest by evoking automatic, emotional response rather than slow, deliberate evaluation (Soroka, 2006). Therefore, although do we not allege that emotive word usage is in any way the only factor that influences market behavior, we expect newspaper content, and specifically the direction of its emotional "charge", to impact the behavior of buyers and sellers; and posit the existence of a relationship between newspaper content and prices that is consistent with this. We operationalize this hypothesis by investigating the predictive relationship between newspaper usage of optimistic and pessimistic words and, as a key indicator of market activity, home prices. While we choose newspapers as the "affect" instrument for our study, we do not maintain that other media sources do not have similar potential. Nor do we contend that our approach in this paper is able to test for causality in this relationship or determine the relative emotive impact of newspapers against other media sources. Our purpose is to establish the presence or not of the relationship, as a prior condition for any subsequent investigation of causality.

There is existing empirical support for the expectation that the information content of news media influences people's decisions. In its broad sense, Erikson, Mackuen and Stimson(2002) assert that news coverage about the economy has an effect on the public's perceptions of the economy, which in turn has an effect on public support of government. In a study more specifically focused on people's financial decisions, Tetlock (2007) examines the relationship between the content of a Wall Street Journal(WSJ) daily

column and stock market prices. He finds evidence that news reports can predict movements in broad indicators of stock market activity and concludes that high levels of media pessimism robustly predict downward pressure on market prices, followed by a reversion to fundamentals.

These findings are consistent with theories of noise traders and "noninformational" traders(De Long, Shleifer, Summers and Waldmann, 1990; Campbell, Grossman and Wang, 1993), with news content acting as a proxy for investor sentiment. Using a narrower channel of media news content, namely Barron's "Ground Floor" column, Downs, Güner, Hartzell and Torres(2001) show that focused market commentaries provide information that significantly impacts prices of REIT stocks. Most recently, Engelberg and Parsons (2011) investigate the relation between media coverage of firm earnings announcements and contemporaneous trading activity in the firm's stock. Segmenting their study into 19 cities, they focus on the impact of local daily newspaper reporting on activity within the boundaries of the newspaper's influence. They find that local media coverage strongly predicts local trading, and by varying only the media exposure of investors they argue that this effect can be treated as causal. Overall, however, it is perhaps easier to believe that institutional and individual investors read and are affected by the daily financial pages than it is to believe that intermittent newspaper stories influence less frequent residential real estate transactions. We therefore expect the detection of relationships in our study to be more difficult than in the contexts investigated by Tetlock(2007), Downs et al.(2001), and Engelberg and Parsons(2011).

The stronger impact of pessimistic information, highlighted by Tetlock(2007), reflects in general the role of loss aversion(Kahneman and Tversky, 1979) but more specifically the contention that an individual's response to positive and negative information is asymmetric. This is confirmed by Soroka(2006). In the context of how public response to economic data might influence government policymaking, Soroka finds not only that negative economic information impacts much more than positive economic information, but that this asymmetry in public responsiveness is enhanced by news content in mass media.

Prior studies of "investment" reaction to news have centered on stocks, with the observed information impacts being generally present for only days rather than several weeks or longer. An important distinction in our study is that an individual's participation in the residential real estate market is either very infrequent or sometimes one-time only. This, coupled with the substantially lower volatility in residential prices compared to stock prices, and corresponding lower liquidity, leads us to expect that relations between residential news coverage and prices will resonate for weeks, longer than has been found in stock studies. We contend that the language and words used in newspaper articles that report upon the residential market will percolate into buyers' and sellers' decision-making. Because we

treat this process as having an "affect" component, we focus upon the emotive content of the language rather than the actual arguments contained in articles. Specifically, we posit a long-term relationship, persisting over the life of our series, between newspaper use of pessimistic and optimistic words, and home prices; a relationship predictable from the former to the latter. We anticipate, in line with prior findings, that the relationship for negative words will be stronger than for positive words.

We use real estate articles in The Atlanta Journal Constitution to construct word count time-series and for prices we use the S&P/Case-Shiller Home Price Index. Our study period is December 2002 through October 2008. We test for cointegration of the data series and investigate both short- and long-run models. We find evidence, in the direction expected, of a long-run equilibrium relationship between pessimistic word usage and residential home prices and also in the short-run adjustments within that relationship. The next section explains the source, nature and application of our data. The three sections following that respectively describe our method of analysis, our findings, and conclusions.

Ⅱ. Data

In order to test for our posited relationship, the first step is to convert qualitative news content into a quantitative research variable. The news content we use is that found in real estate articles in The Atlanta Journal-Constitution(AJC.). The AJC is the largest daily newspaper in the Southeast US, and the "local" nature of its circulation reflects the localized nature of the region's residential real estate market. It is the most reputable publication in Atlanta's Consolidated Metropolitan Statistical Areas(CMSA) with the largest number (1.9 million) of readers. Most influential regional forecasters make use of this newspaper to deliver their findings. An online search was conducted for real estate articles in the daily editions of the AJC from December 2002 to October 2008, identified with key words such as real estate, real estate price or home price(appearing anywhere in the article, including the headline). Six hundred and five relevant articles were captured (the successive annual totals being 32 in 2002, then 75, 78, 75, 88, 102, and finally 155 in 2008).

The content analysis was performed using General Inquirer(GI), a widely used text analysis software program. To count the frequency of optimistic and pessimistic word usage, we utilize word classification categories provided within GI, and based on the Harvard Psychological Dictionary. It should be noted that the use of this source has come under some scrutiny recently, on the basis that some words mutate from their general meanings in specific contexts. Loughran and McDonald (2011), for example, argue that this is true of documents produced for financial

purposes (e.g. 10-K filing returns). While the question of general versus custom dictionaries has not been settled, we do not in any event believe this argument holds for our study, in which we can expect that words in newspapers are used in their general context and likely to be so read.

We amalgamate all articles appearing in a calendar month so as to produce monthly optimistic and pessimistic word frequency measures that are converted into percentages, relative to the total number of words in the articles in which they appear[1]. It is likely that negative qualifiers will sometimes be used, such as “not fall” in an article whose general tone might be “positive”. Since we would code “fall” in “not fall” as negative, this may appear as a threat to our approach. We do not believe this threat is strong, for two reasons. First, the use of negative qualifiers is small in our sample. Second, more importantly, we investigate the emotive power of individual word usage, rather than the overall message in an article.[2] (In fact, the latter would be difficult or impossible to transform into a quantitative variable and thus not be amenable to the econometric technique that we adopt.)

To track prices, we draw from the S&P/Case-Shiller Home Price Index, using monthly data for the Atlanta MSA for the period from December 2002 through October 2008. Although our expectation is that language in newspapers will have a contemporaneous fairly immediate impact on readers’ feelings and emotions and that this will be reflected in contemporaneous price decisions, we take into account that price data only become available to the Index compilers after transaction closing and its subsequent recording. We therefore conjecture a lag in the relationship between our time series. We expect this lag to be approximately two months and use this in our modeling.[3] As a robustness check, we also repeat the analysis with

1) For classification of words into our pessimistic words category, we group together four categories from the Harvard IV-4 classification. These are ‘ngtv’ (as termed in Harvard IV-4), ‘weak’, ‘fail’ and ‘fall’ words. For classification of optimistic words, we use the general ‘positiv’ (sic) category in General Inquirer. This provides relatively balanced word content in each category (2094 pessimistic words and 1915 optimistic words), although other category constructions are possible. Tetlock (2007), for instance, uses ‘negative’ and ‘weak’ to capture pessimism.

2) A similar, though somewhat different concern, might be that our article-set contains reports on conditions in other metro markets whose momentum might at times be at odds with metro Atlanta. Our intention, however, is to demonstrate the predictive power of the words used, rather than the appropriateness, to their decisions, of what readers read.

3) The lag is between the month in which words appear in the AJC and the month of the Index (and not the month in which the Index is published, which, coincidentally, is done with a two month lag). We are also aware that the Case-Shiller Index uses a three month moving average but do not believe that this undermines our conjecture about the relationship between the series.

Federal Housing Finance Agency (FHFA, formerly OFHEO) Home Price Index for the Atlanta MSA. This index is based on transactions involving conforming, conventional mortgages. It uses a similar repeat-sales methodology as used in the Case-Shiller index but is available only on a quarterly basis. Accordingly, given that we are constrained in our lags to quarters, we adopt a one quarter lag with this index, reasoning this as the closest approximation to the lag we adopt for the Case-Shiller index.

Ⅲ. Method

1. Cointegration Test and Vector Error Correction Model

Regression of time series data may produce spurious results where the series itself(i.e.level) is non-stationary.[4] Nevertheless, under certain conditions, we may expect non-spurious results of regression of such level series. These conditions are where the first differences of non-stationary level series are all stationary, and where it is possible for linear combinations of such first difference series to produce a stationary series. Variables to which this applies are viewed as having a co-integrating relationship. In order to investigate our posited relationship we apply the cointegration test suggested Johansen and Juselius(1990). The cointegration analysis first requires a unit root test (Dickey and Fuller, 1981) to determine if the first differences of the variables are stationary. If such a stationary linear combination is the case, the Johansen test can confirm that the non-stationary level time series are cointegrated. The stationary linear combination - the cointegrating equation - may be interpreted as a long-run equilibrium relationship among the variables.

As explained below, the presence of a cointegrating relationship forms the basis of the Vector Error Correction specification. We implement this cointegration test using the methodology developed by Johansen (1991, 1995).

Consider the order:

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + Bx_t + \varepsilon_t \quad (1)$$

Where y_t is a k-vector of non-stationary I(1) variables, x_t is a d-vector of deterministic variables, and ε_t is a vector of innovations. We may rewrite this as,

$$\triangle y_t = \Pi y_{t-1} + \sum_{t=1}^{p-1} \Gamma_i \triangle y_{t-1} + Bx_t + \varepsilon_t \quad (2)$$

where:

$$\Pi = \sum_{i=1}^{p} A_i - I, \qquad \Gamma_i = \sum_{j=i+1}^{p} A_j \qquad (3)$$

4) A stationary time series is one whose statistical properties (e.g. mean and variance) are constant over time.

Granger's representation theorem asserts that if the coefficient matrix Π_j has reduced rank $r < k$, then there exist $k \times r$ matrices α and β each with rank r such that $\Pi = \alpha\beta'$ and $\beta' y_t$ is I(0). r is the number of cointegrating relations (the cointegrating rank) and each column of β is the cointegrating vector. As explained below, the elements of α are known as the adjustment parameters in the VEC model. Johansen's cointegration test is to estimate the Π matrix from an unrestricted VAR and to test whether we can reject the restrictions implied by the reduced rank of Π.

2. Structural modeling for long- run and short- run models

We adopt the Engle-Granger two-step method in which a long-run model is specified in levels and a short-run adjustment model is specified in first differences but includes a long-run error correction term derived from the estimation of the long-run equilibrium model. The equilibrium relationship is built into the specification so that it restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationships, while allowing for short-run adjustment dynamics. The cointegration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments.

In the first-stage, a long-run equilibrium relation (i.e. a cointegrating regression) can be specified in levels as:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta X_{it} + \nu_t \qquad (4)$$

Where y_t is the dependent research variable and X_{it} are theoretical based research variables i at time t. From this regression, residuals canbe estimated as the difference between the actual and estimated equilibrium between t esting factors, residential home price as the dependent variable, and causal factors of optimistic and pessimistic language. If the residuals from equation (4) are stationary, they may be used as an error correction term in the short run media model as follows:

$$\triangle y_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_{it} - \gamma \widehat{\nu_{t-1}} + \varepsilon_t \qquad (5)$$

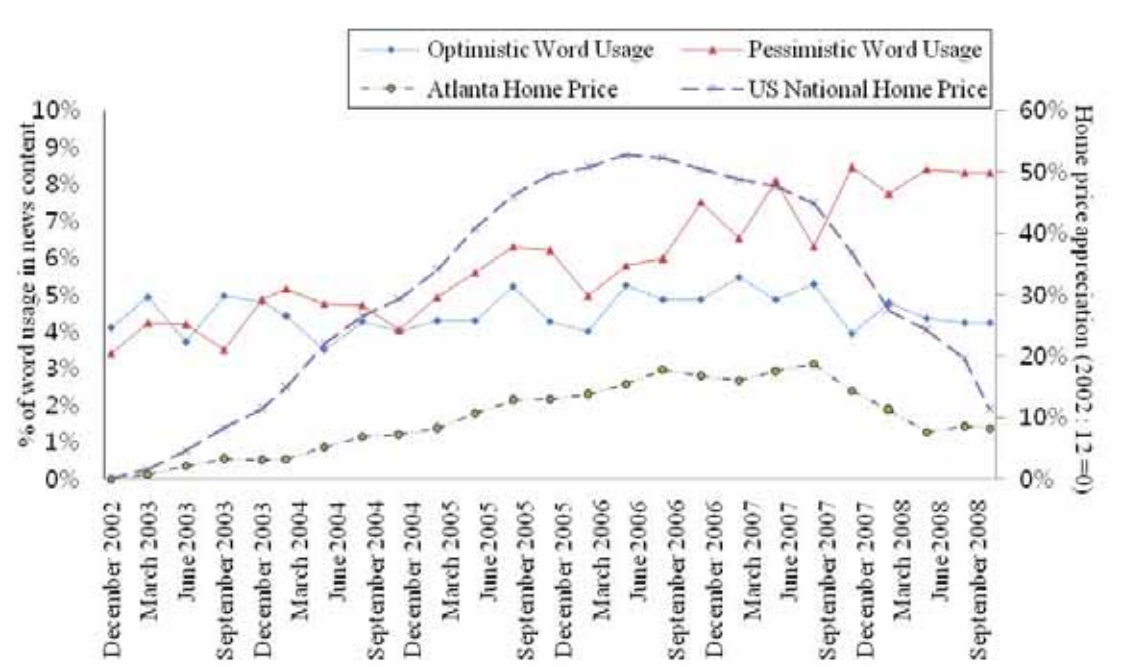
Where $\triangle y_t = y_t - y_{t-1}$ is the first difference of the dependent variable for the study, residential home prices, $\triangle X_{it}$ are the first differences of the explanatory variables, and $\widehat{\nu_{t-1}}$ is the error correction term (the lagged residuals from the long-run regression). Estimation of equation (5) provides evidence of a short-run residential market dynamic and adjustments to the previous disequilibrium in the long-run relation, γ (the speed of adjustment parameter). If $\gamma = 1$, there is full adjustment,

while $\gamma = 0$ suggests no adjustment. A more general specification of the short-run model may also include multiple lags of the explanatory and dependent variables.

Ⅳ. Results

Results for the extent of optimistic and pessimistic word usage are shown in Figure 1. The use of pessimistic words increased significantly from 2006 to 2008 while optimistic word usage remained generally flat, declining slightly. The peak in pessimistic words occurs at the end of 2007. Optimistic word usage and pessimistic word usage diverge significantly starting in September 2006. For purpose of comparison with pessimistic and optimistic word usage, the normalized Case-Shiller Atlanta home-price and national home-price indices are also shown in Figure 1. Although Atlanta home prices stabilize in June 2006, they peak in September 2007. The divergent pattern between pessimistic word usage and optimistic word usage begins in 2006 when the national home price level begins to decline. Beginning mid-year 2007, just before home prices

〈Figure 1〉 Percentage of pessimistic and optimistic word usage in newspaper



Note : A quarterly average of the percentage of pssimistic and optimitic words used in The Atlanta Journal Constitution from January 2003 - October 2008. The S&P/Case-Shiller 20- City Composite and S&P Case Shiller Atlanta Home Price Index are used as a proxy for home price of U.S. national and Altanta level, respectively. The reference point to normalize the begining point is Deceber 2002. Thus, the home price graph shows relative index change from the reference point, December 2002.

〈Table 1〉 Unit Roots Test for Newspaper Content and Residential Home Prices

	Level		First Difference	
Research Variable	t-statistic	p-value	t-statistic	p-value
Home Price($R_{t,HP}$)	2.4	1.00	-6.4	0.00
Pessimisticwords($R_{t,Pessm\_word}$)	-1.0	0.75	-6.6	0.00
Optimistic words($R_{t,Optism\_word}$)	-2.1	0.25	-7.3	0.00

Note: The Augmented Dickey-Fuller test statistic (ADF) is used. The specification of $R_{t,HP}$ is the change of home price, $R_{t,Pessm\_word}$ is the percentage of pessimistic words, and $R_{t,Optism\_word}$ is the percentage of optimistic words.

start falling in Atlanta, the divergence between optimistic word

usage and pessimistic word usage clearly accelerates, in lagged synchronization with deteriorating local market conditions.

1. Stationarity Test

The data generated from the content analysis are qualitative in nature. After the data are transformed into quantitative measures, however, they become, like home prices, time-dependent data series or more specifically quantitative time series. As the first stage in investigating the relationship between these time series, a unit root test is conducted, to check whether the research variables meet the stationarity condition. The unit root results are reported in Table 1.

The test shows evidence of non-stationary conditions in the level series, in that for each series we cannot reject the non-stationary hypothesis. There is, however, strong evidence of stationary conditions in the first difference series, with all three variables highly significant in the ADF test. This allows that the research variables may have a cointegrated, non-spurious long-run relationship, which may deviate in the short run, but return to the underlying association in the long run.

2. Cointegration Test

To test for cointegration, we adopt Johansen's procedure. Since there are three variables in the cointegration test, there can be at most two linearly independent cointegrating vectors, i.e., $\gamma \leq 2$. The results are presented in Table 2.

The first column contains the number of cointegrating vectors under the null hypothesis (below), the second column contains the ordered eigenvalues, the third column displays the trace statistics, and the last two columns are the 5% critical values and p-values. Trace statistics are used to test for the number of cointegrating relationships. The maximum eigenvalue statistic is used to test the null hypothesis of cointegrating relationships

against the alternative of $\gamma + 1$ cointegrating

〈Table 2〉 Johansen's Cointegration Test Newspaper Content and Residential Home Prices

Number of Cointergrating Vertors(γ) at Null Hypothesis	Eigenvalue	Trace Statistics	Critical Value at 5%	p-value
$\gamma = 0$	0.44	59.29	29.79	0.00
$\gamma \leq 1$	0.24	21.18	15.49	0.00
$\gamma \leq 2$	0.04	2.90	3.84	0.08

Note: Trace test indicates 2 cointegrating eigenvalues. The parameter estimates are from the Johansen's test using monthly data over the time period from December 2002 to October. 2008. The variables are home prices, pessimistic and optimistic word usage.

relationships. The trace statistic for the null hypotheses of $\gamma = 0$ and $\gamma \leq 1$ is larger than the critical values so we can reject these hypotheses. The corresponding statistic for $\gamma \leq 2$ (2.90) is smaller than the critical value(3.84), so the null hypothesis that $\gamma \leq 2$ cannot be rejected, suggesting there are two linearly independent cointegrating vectors representing the empirical long-run equilibrium relationship among the research variables.

3. Controlling for Inverse Causality

It is possible of course that pessimistic and optimistic word usage in news content may be an "advance" reaction to other residential market indicators that only subsequently get impounded into reported prices. In other words, changes in newspaper content may predict (in the sense of preceding) later changes in prices. To investigate this, we test the relative responsiveness of changes in pessimistic and optimistic word usage to the changes in home prices by adopting the Granger causality test for inverse causality. The Granger approach to the question of whether a variable x causes y is to see how much of the current y can be explained by past values of y and then to see whether adding lagged values of x can improve the explanation. Thus, y is said to be Granger-caused by x, if x helps in the prediction of y[5], or equivalently if the coefficients on the lagged s are statistically significant. The Granger Causality test can be expressed as

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + ... + \alpha_l \Delta y_{t-l} + \beta_1 \Delta x_{t-1} + ... + \beta_l \Delta x_{t-l} + \varepsilon_t \quad (6)$$

$$\Delta x_t = \alpha_0 + \alpha_1 \Delta x_{t-1} + ... + \alpha_l \Delta x_{t-l} + \beta_1 \Delta y_{t-1} + ... + \beta_l \Delta y_{t-l} + u_t \quad (7)$$

for all possible pairs of series in the group. The reported p-value results from the Wald statistics for the joint hypothesis:

5) Thus Granger causality is not inconsistent with true causality but is not itself a test for true causality.

$$\beta_1 = \beta_2 = \beta_3 = ... = \beta_l = 0 \quad (8)$$

The null hypothesis is that does not Granger-cause in the first regression and that does not Granger-cause in the second regression.

In Table 3, Panel 1 shows that the change in home prices relates significantly to pessimistic word usage and to optimistic word usage: both of which therefore "Granger cause" changes in home prices. In Panel 2, pessimistic words respond significantly to optimistic word usage but do not respond to the change in home price variables. Optimistic words do not respond to other variables, as reported in Panel 3. We find that both optimistic word usage and pessimistic word usage granger cause home price and also optimistic word affect pssimistic word usage as shown in panel 2.

4. Long-Run Equilibrium Relationship

As previously outlined, if the data series are cointegrated, a long-run equilibrium relationship (i.e. cointegrating regression) can be specified by the Engle-Granger two-step method in levels such that change in home prices is related to pessimistic and optimistic word usage. That is:

$$R_{t,HP} = \beta_0 + \beta_1 R_{t-2,Pessm\_word} + \beta_2 R_{t-2,Optism\_word} + v_t \quad (9)$$

where $R_{t,HP}$ is the change in monthly home prices; $R_{t-2,Pessm_word}$is the monthly number of pessimistic words as a percentage of the total monthly number of words used, lagged two months; and $R_{t-2,Optism_word}$is the

〈Table 3〉 Granger Causality Tests for the Change in Newspaper Content and Residential Home Prices

	Independent Variable	Dependent Variable	p-value
Panel 1			
	$R_{t,Optism\_word}$	$R_{t,HP}$	0.04*
	$R_{t,Pessm\_word}$	$R_{t,HP}$	0.05*
Panel 2			
	$R_{t,HP}$	$R_{t,Pessm\_word}$	0.07
	$R_{t,Optism\_word}$	$R_{t,Pessm\_word}$	0.02*
Panel 3			
	$R_{t,HP}$	$R_{t,Optism\_word}$	0.38
	$R_{t,Pessm\_word}$	$R_{t,Optism\_word}$	0.38

Note: $R_{t,HP}$ is the change in home price, $R_{t,Pessm\_word}$ is the percentage of pessimistic word usage, and $R_{t,Optism\_word}$ is the percentage of optimistic word usage. The parameter estimates are from the time period December 2002 to October 2008. **and*denote p-value significantat 1%and 5%, respectively. Each dependent variable is tested within dependent variables in two month lags.

〈Table 4〉 Long-run Equilibrium Model for the Relationship between Newspaper Content and Residential Home Prices. December 2002 - October 2008

Long-run Equilibrium Relationship Dependent Variable $R_{t,HP}$	β_0	$R_{t-2,\,Pessm\_word}$	$R_{t-2,\,Optism\_word}$	$Adj\,R^2$
S&P C-S index (monthly data)	0.011**	-0.12*	-0.06	0.14
	(3.20)	(-2.89)	(-1.02)	
	β_0	$R_{q-1,\,Pessm\_word}$	$R_{q-1,\,Optism\_word}$	$Adj\,R^2$
FHFA HPI (quarterly data)	0.012**	-0.15**	-0.05	0.29
	(2.54)	(-3.34)	(-0.74)	

Note: The parameter estimates are from Equation (9).
t-statistics are in parentheses.**and* denote significance at 1% and 5%, respectively.
Lags are two months for the C-S index and one quarter for the FHFA HPI.

corresponding percentage for optimistic words with the same lag. β_0 is a constant and v_t is the error term.

One condition for long-run equilibrium is that the relationship between home prices and news content is not changing. An Error Correction Model can address this interpretation problem by using a combination of first differenced and lagged levels of cointegrated variables. Here, the Error Correction Model is specified as follows.

$$\triangle R_{t,HP} = \alpha_0 + \alpha_1 \triangle R_{t-2,Pessm\_word} + \alpha_2 \triangle R_{t-2\,Optism\_word} + \gamma \widehat{v_{t-1}} \quad (10)$$

where $\triangle$ denotes the first difference of each research factor and γ is the speed

adjustment parameter. The change in home price is associated with prior changes in pessimistic and optimistic word usage but also acts in part to "correct" for any disequilibrium that existed in the previous period. Broadly, γ in equation (10) describes the speed adjustment back to the equilibrium measuring the proportion of the last period's equilibrium error that is corrected for.

Table 4 contains parameter estimates and t statistics for our long-run model (Equation 9) for the time period December. 2002 to October 2008. (It also displays the corresponding statistics using the FHFA HPI, the use of which we report upon at the end of this Section). The estimated coefficient on $R_{t-2,\,Pessm\_word}$, pessimistic news content two months prior to reported house price, is negative and statistically highly significant. The estimated coefficient on optimistic words, $R_{t-2,\,Optism\_word}$ is also negative but not statistically significant. The

〈Table 5〉 Error Correction Model for the Relationship between Newspaper Content and Residential Home Prices. December 2002 - October 2008

Short-run Equilibrium Relationship Dependent Variable $\triangle R_{t,HP}$	α_0	$\triangle R_{t-2,Pessm\_wor}$	$\triangle R_{t-2,Optism\_wor}$	γ	$Adj\,R^2$
S&P C-S index (monthly data)	0.00	-0.04	0.04	-0.20*	0.12
	(0.27)	(-2.04)	(1.28)	(-2.64)	
	α_0	$\triangle R_{q-1,Pessm\_wor}$	$\triangle R_{q-1,Optism\_wor}$	γ	$Adj\,R^2$
FHFA HPI (quarterly data)	0.00	-0.13**	-0.05	-0.85**	0.27
	(2.54)	(-5.74)	(1.24)	(-6.71)	

Note: The parameter estimates are from Equation (10).
t-statistics are in parentheses.**and* denote significance at 1% and 5%, respectively.
Lags are two months for the C-S index and one quarter for the FHFA HPI.

adjusted R^2 for the long-run equilibrium model is 0.14. The extent of relatively pessimistic news content in articles about the residential real estate market is inversely associated with house price levels as recorded two months later (and made available in the Case-Shiller index after a further two month delay). To the extent that it is statistically significant, optimistic news content does not display the same association.[6]

5. Short-Run Error Correction

The short-run error correction equilibrium relationships are shown in Table 5 (which also displays the statistics relating to the FHFA HPI).

As noted, a short-run error correction model is based on the idea that two or more time series exhibit a long-run time-varying equilibrium to which the systems tend to converge. The long-run equilibrium influence in the error correction model is achieved through an error correction term which is speed adjustment γ. The results from December 2002 to October 2008 show that the estimated coefficient on $\triangle R_{t-2,Pessm\_word}$, a change in pessimistic words usage is at the 5% level. Mirroring the

6) We observe (Figure 1) a decline in home price in the period where the difference between optimistic and pessimistic word usages diverges. We therefore also investigate the long-run equilibrium model by controlling for the difference between optimistic and pessimistic word usage. None of the word parameter estimates are significant at our test levels of significance. (Results available from authors)

〈Table 6〉 Ordinary Least Square(OLS) results for the Relationship between Newspaper Content and Residential Home Prices along with Macro Economic Variables. December 2002 - October 2008

$\triangle R_{t,HP}$	β_0	$\triangle R_{t,Pessm}$	$\triangle R_{t,Optism\_wo}$	△Interest	△Income	△Unem.	△Building Permit	$Adj\,R^2$
S&P C-S index	0.019**	-0.028	0.062*	-0.328**	-0.155	-0.06	0.001	0.17
	(3.20)	(-1.24)	(-1.89)	(-2.18)	(-1.22)	(-1.02)	(0.42)	

Note: Dependent variable is the S&P C-S index. t-statistics are in parentheses.**and* denote significance at 5% and 10%, respectively. Lags are two months for the C-S index. The dependent variable is the monthly home price index. The independent variables are pessimistic word($R_{t,Pessm\_word}$), optimistic word($R_{t,Optism\_word}$), 30 year fixed rate mortgage rate(Interest), income, unemployment(Unem) rate, building permit.

results of the long term model, the estimated coefficient of a change in optimistic word usage is statistically not significant. The error correction term γ, a difference between actual and predicted home price from a long run model, has a statistically significant coefficient estimate of -0.20. A relative change in pessimistic word usage from the previous month has negative impact on incremental change in home price (as later recorded and reported) but this does not happen with optimistic words. The model also confirms the short-term dynamic adjustment toward long-run equilibrium.[7)]

Our check for robustness, using the alternative of the FHFA HP index, reveals results consistent with the Case-Shiller index when we test for stationarity, cointegration and Granger causality; although in respect of cointegration, the failure to reject the hypothesis of two cointegrating vectors is shifted to the 10% level.[8)] As shown in Tables 4 and 5, the long run equilibrium and short-run error correction results are also in line with the Case-Shiller results.

In Table 6, we adopt a set of macro economic variables such as 30 year fixed mortgage interest rate, building permit and unemployment rate. We assume this result can be compared with the short-run model in statistical reference, although it does not perfectly match including error correction term. However, at this stage, we consider this result can be used an alternative way to unfold our hypothesis with macro economics variables. As we expect, the magnitude of coefficient for pessimistic word usage was diminished and optimistic word usage has remained a similar level and marginally significant at 10 percent level. Especially, the interest rate for 30-year fixed rate mortgage

7) To further test for robustness, we repeat the Case-Shiller analysis by examining both three month and four month lags but find results (not reported here) that are similar to those for our conjectured two month lags.

8) The statistics for these tests using the FHFA data are not reported here but are available from the authors.

shows significant coefficient at 5 percent level. We can conclude that the pessimistic word usages has been diminished as we include macro economics variables. Still the optimistic word usage has been significant interpreted as during the last decade, the optimisitc word usages and interest rate has been critical variables on change of home price.

V. Conclusion

This study examines the underlying links between the content of real estate news articles and residential home prices. We build from the idea that decision-making is dependent on intuitive response, as well as reasoning, and that judgment can biased by "affect heuristics". The human mind does not functionally separate perception, feeling and intuition from reasoning. Our contention is that language - in this case through newspaper content - will therefore impact an individual's decisions and be embodied in the pricing records of those decisions as people internalize and rely on these depictions. Evidence consistent with these conjectures would be observable in the relationship between prices and the time series representing the balance of pessimistic and optimistic language. This study supplies evidence of a long-run equilibrium relationship, in the direction expected, between pessimistic word usage and residential home prices in the period (2002 – 2008) studied, and also in the short-run adjustments within that relationship.

The ebb and flow of the affective presence of pessimism in news articles contributes to "prediction" what is happening contemporaneously in the residential market, but which only emerges into the public domain months later. That this relationship is absent for optimistic language is further evidence of the asymmetry in this "mood" dichotomy, which has been found in studies across a range of contexts. This work is a first step in applying in a real estate context what was previously known about the interaction of perception, reasoning and emotion to major financial decisions. Time has shown, however, that our period of analysis contained an almost unprecedented phase in US housing history. We believe our findings should be tested for their presence in less turbulent market times. Conversely, we also believe it worth exploring whether our findings are replicated in residential markets that exhibit greater momentum dynamics than that of the Atlanta CMSA, which, despite experiencing its own housing bubble and bust, was subject to substantially more muted swings than other US markets. This cross-city comparison may also enable a research design that moves towards the testing of causality, as performed by Engelberg and Parson (2011); although it may be challenging to devise the necessary controls, not least because the timing of the evidence of the home-purchase decision (i.e. the index "date") is

considerably more remote from the media occurrence than in studies of stock-trades. A third improvement would be to encompass testing of other market indicators, most notably market sales volume, which has been successfully used in non-residential studies (De Long etal.,1990;Tetlock,2007). For this study we were unable to gain sanctioned access to such a data series.

논문접수일 : 2012년 3월 28일
논문심사일 : 2012년 4월 20일
게재확정일 : 2012년 6월 10일

References

1. Barberis, N. and R. Thaler, *A survey of behavioral finance*, in Handbook of the Economics of Finance, North-Holland, Amsterdam. 2003
2. Campbell,Y., S. Grossman and J. Wang., "Trading Volume and Serial Correlation in Stock Returns." *The Quarterly Journal of Economics*, Vol. 108 No. 4, 1193, pp. 905-939
3. Case, K. and R. Shiller. "The behavior of home buyers in boom and post-boom markets." *New England Economic Review*, November/ December, 1988. pp. 29-46.
4. Clayton, J. "Further evidence on real estate market efficiency." *Journal of Real Estate Research*, Vol. 15, 1998. pp. 41-57
5. De Long, J., A. Shleifer, L. Summers and R. Waldmann, "Noise trader risk in financial markets." *Journal of Political Economy*, Vol 98, 1990. pp. 703–738
6. Dickey, D. and W. Fuller. "Likelihood ratio statistics for autogressive time series with a unit root." *Econometrica*, Vol. 49 No. 4, 1981. pp. 1057-1072.
7. Downs, D., N. Güner and G. Patterson. " Capital distribution policy and information asymmetry: A real estate market perspective." *Journal of Real Estate Finance and Economics*, Vol. 21, 2000. pp. 235–250.
8. Downs, D., N. Güner, D. Hartzell and M. Torres. "Why do REIT prices change? The information content of Barron's ``The Ground Floor''. *Journal of Real Estate Finance and Economics*, Vol. 22,

2001. pp. 63-80.

9. Engelberg, J. and C. Parsons. "The causal impact of media in financial markets." *Journal of Finance*, Vol. 66 No. 1, 2011, pp. 67-97
10. Erikson, R., M. MacKuen and J. Stimson. *The Macro Polity,* Cambridge University Press, Cambridge. 2002.
11. Finucane, M., A. Alhakami, P. Slovic and S. Johnson. "The affect heuristic in judgments of risks and benefits." *Journal of Behavioral Decision Making*, Vol. 13 No 1, 2000. pp. 1-17.
12. Genesove, D. and C. Mayer., "Loss aversion and seller behavior: Evidence from the housing market." *The Quarterly Journal of Economics*, Vol. 116 No. 4, 2001, pp. 1233-1260.
13. Johansen, S. and K. Juselius. "Maximum likelihood estimation and inference on cointegration – with applications to the demand for money." *Oxford Bulletin of Economics and Statistics*, Vol. 52, 1990. pp. 169-210.
14. Johansen, S. "Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models." *Econometrica*, Vol. 59, 1991 pp. 1551-1580.
15. Johansen, S. *Likelihood-based inference in cointegrated vector autoregressive models.* Oxford University Press, Oxford. 1995.
16. Kahneman, D and A. Tversky. "Prospect theory: An analysis of decision under risk." *Econometrica,* Vol 47, 1979, pp. 263-291
17. Kahneman, D. A "perspective on judgement and choice." *American Psychologist*, Vol 58, No 9, 2003. pp. 697-720.
18. Loughran, T. and B. McDonald. "When Is a Liability Not a Liability?" *Textual Analysis, Dictionaries, and 10-Ks Journal of Finance*, Vol 66 No 1, 2011 pp. 35-65
19. McDonald, C., T. Nixon and V. Slawson. "The changing asymmetric information component of REIT spreads: a study of anticipated announcements." *Journal of Real Estate Finance and Economics*, Vol 20, 2000, pp. 195–210.
20. Slovic, P., M. Finucane, E. Peters and D. MacGregor. "Rational actors or rational fools: implications of the affect heuristic for behavioral economics." *Journal of Socio-Economics*, Vol 31 No. 4, 2002. pp. 329-342.
21. Soroka, S. "Good news and bad news: Asymmetric responses to economic information." *Journal of Politics*, Vol. 68 No. 2, 2006. pp. 372-385.
22. Tetlock, P. "Giving content to investor sentiment: The role of media in the stock market." *Journal of Finance*, Vol 62 No. 3, 2007. pp. 1139-1168.