TOPICS IN REAL ESTATE RESEARCH, 1973–2010: A LATENT SEMANTIC ANALYSIS

Kimberly Winson-Geideman
University of North Texas

Nicholas Evangelopoulos
University of North Texas

Abstract
With the development and evolution of real estate literature a number of studies, primarily qualitative, have attempted to build structure into the body of published research. This research uses Latent Semantic Analysis to evaluate the content of abstracts published in *Real Estate Economics*, the *Journal of Real Estate Finance and Economics*, and the *Journal of Real Estate Research* from 1973 through 2010. Two main themes are identified and results are disaggregated, creating a pool of 25 topics within the main themes. Findings show a relatively stable pattern of research for some topics and the influence of market conditions on others. The extracted research topics comprise the intellectual core of real estate research.

The study of real estate as an academic discipline has been conducted in many different ways. Several studies have used the number of citations a given article or journal generates as the basis for analysis, while others have surveyed faculty members from a pedagogical perspective, inquiring about course offerings and undergraduate or graduate programs. Still others have used faculty publication records or university affiliations as a framework. This paper takes a different approach. Latent Semantic Analysis (LSA) is used to examine word patterns in a large body of published real estate research in order to establish a topical structure that defines the intellectual core of the real estate discipline. Specifically, the abstracts of all research papers published between 1973 and 2010 in the three primary real estate journals—*Real Estate Economics*, the *Journal of Real Estate Finance and Economics*, and the *Journal of Real Estate Research*—are analyzed at various alternative levels of semantic aggregation. At one such level of aggregation, results identify two core research areas that have remained relatively stable over the study period: (1) real property markets and (2) real estate portfolios. The two areas are relatively balanced across the three journals with 53% and 47% of published articles, respectively. In addition to the core, LSA is used to regulate the data aggregation at more detailed levels, suggesting different themes within those two areas.

BACKGROUND

BRIEF HISTORY OF REAL ESTATE ACADEMIC ORGANIZATIONS AND JOURNALS

On December 29, 1964, a group of about 20 real estate academics created the American Real Estate and Urban Economics Association (AREUEA). Prior to this,

real estate was not widely accepted as an academic discipline separate from the social sciences and there was no academic organization catering to the field. The only existing forum for academics was the National Association of Real Estate Boards (NAREB, later the National Association of Realtors, NAR) education committee, whose focus was to make recommendations regarding educational policies and practice for NAREB members. AREUEA was founded with the chief objectives of encouraging research in real estate and promoting the development of real estate as an academic discipline (Hendershott, Thibodeau, and Smith, 2009).

In 1973 the inaugural issue of the *Journal of the American Real Estate and Urban Economics Association* (*AREUEA Journal*) was published, marking the beginning of a distinct, multi- and interdisciplinary research field in real estate. Over the years, the topics covered within the journal differed, but the most highly referenced papers fell into four distinct categories: housing, urban, commercial, and finance. Housing topics ranged from tenure choice to subsidies to price determination. Urban papers covered environmental issues, transportation, and rent gradients. Commercial topics include portfolio analysis and land development, and the finance papers were focused on issues relevant to both the commercial and residential markets. In 1989 the AREUEA Journal became known as *Real Estate Economics*.

While AREUEA served as a strong advocate for real estate as an academic discipline, it became obvious to some members that a second organization focused on bridging the gap between the profession and the academy was needed. The American Real Estate Society (ARES) was founded in 1985 to address the integration of real estate academics with industry research practitioners, striving to include all real estate thought leaders. The corresponding journal, the *Journal of Real Estate Research*, was founded in 1986. AREUEA and ARES are the primary professional organizations for real estate academics in the United States today. A third journal, the *Journal of Real Estate Finance and Economics* (*JREFE*), was first printed in 1988. While not officially associated with either of the national organizations, *JREFE* is considered one of the premier core real estate journals (along with *REE* and *JRER*). Several important studies exist that delineate these core journals from other real estate-related journals, as will be shown in the following sections.

TOPICS IN THE REAL ESTATE DISCIPLINE

Since real estate is an integral part of and is affected by a number of different disciplines, one would expect a vast array of research topics to be conducted in the name of real estate. Evidence of this is shown in a special issue of *JRER*, where Johnson, Roulac, and Followill (1996) identify the 16 most frequently presented topics in 998 papers over the first ten years of ARES annual meetings using paper and session titles. Topics are presented in rank order based on total number of papers presented from 1985 through 1994 and include urban economics/development (18.54%), valuation (14.63%), mortgage applications (9.02%), international (6.71%), corporate real estate/asset management (6.61%), brokerage (6.41%), investments (6.01%), risk/portfolio management (5.61%), taxes (5.31%), securitization (4.71%), body of knowledge (3.31%), cases (3.21%), legal issues (3.11%), computer

applications (3.11%), markets and institutions (2%), and government programs (1.7%). Jud (1996) conducts a similar analysis, providing a retrospective of articles published in *JRER* between 1986 and 1994. The author identifies 10 distinct subject areas that are addressed in 189 out of the 257 papers published over the nine-year period, including investment (18.52%), appraisal (13.23%), corporate real estate (12.7%), housing values (12.17%), brokerage (11.11%), REITs (7.94%), housing markets (6.35%), regional (6.35%), environment (5.82%), and mortgages (5.82%).

Dombrow and Turnbull (2004) develop a topical analysis of real estate research similar to the one proposed here by evaluating each paper published in *JREFE* and *REE* from 1988 through 2001. The authors develop nine topics—appraisal, brokerage, housing, institutions, investments, mortgages, nonresidential, public policy, and other—and then categorize each paper according to the applicable research technique: empirical, method, literature review, simulation, or theory. They also look at trends within three subsets of the sample period to identify the relative increase or decrease in popularity over time for each of the nine topic areas.

Research conducted by Harrison and Manning (2008) categorizes real estate articles published in 119 academic journals based on a classification system used by the *Journal of Real Estate Literature*. The authors investigate the level of interest shown in various topics and attempt to evaluate future trends within the discipline. Results show that 70% of real estate research is concentrated in four categories: real estate business/industry, type of decision, government policy/planning, and type of real estate. Furthermore, the popularity of the 10 broad topic categories does not appear to follow any appreciable trend, although there is some variability that occurs over the seven years that are analyzed.

The results of these papers show the depth and breadth of real estate research and confirm its multi- and inter-disciplinary nature. While this makes some real estate research suitable for publication in finance, law or technology journals, those journals focused exclusively on real estate are *REE*, *JREFE*, and *JRER*, the "core" of the real estate discipline (Dombrow and Turnbull, 2002). The importance and efficacy of the core journals has been substantiated in numerous publications and in various ways. Some research focuses on the number of citations a given journal or article generates (Redman, Manakyan, and Tanner, 1998, 1999; Hardin, Liano, and Chan, 2007a, b), and others focus on the productivity of authors (Sa-Aadu and Shilling, 1988; Clauretie and Daneshvary, 1993; Dombrow and Turnbull, 2000; Urbancic, 2007) or editorial board members (Beauchamp, Hardin, Hill, and Liano, 2008). Surveys of real estate faculty and researchers (Diaz, Black, and Rabianski, 1996; Gibler and Ziobrowski, 2002; Gibler and Zhou, 2010) or academics in related fields such as finance are also used (Webb and Albert, 1995).

The preceding studies provide evidence that real estate is an academic discipline complete with prolific authors, heavily-cited journals and research, and a distinct identity. However, most extant studies possess two characteristics of potential concern: the use of predefined topic categories and a somewhat subjective assignment of individual articles to those categories. The present study seeks to address the first issue by letting topics grow organically in a bottom-up fashion, rather than adopting

Exhibit 1
Article Data (1973–2010)

Journal	Number of Articles				
	1973–1980	1981–1990	1991–2000	2001–2010	Total
Journal of Real Estate Finance & Economics		68	324	413	805
Journal of Real Estate Research		117	404	227	748
Real Estate Economics	166	284	269	254	973
Total	166	469	997	894	2,526

Notes: The article counts for *Real Estate Economics* include counts for its predecessors, *AREUEA Journal* in the period 1973–1976, and *Journal of the American Real Estate and Urban Economics Association* in the period 1977–1994. Our analysis excluded editorials and book reviews.

pre-defined categories. The second issue is addressed by the use of a quantitative analysis approach that extracts the topics as language patterns in the abstracts provided by the RE authors who contributed one or more of the 2,526 papers included in our study. In the sections that follow, the research methodology, results, and conclusions are presented.

METHOD

DATA COLLECTION

The articles included in our study were published in the 1973–2010 period in *JREFE*, *JRER*, and *REE*. Certain decades had non-existent data, as *JREFE* was first published in 1988 and *JRER* in 1986. The article counts for *REE* include counts for its predecessors, *AREUEA Journal* in 1973–1976 and the *Journal of the America Real Estate and Urban Economics Association* in 1977–1988. Our analysis excluded editorials and book reviews, yielding a total sample of 2,526 research articles. See Exhibit 1 for a summary of the articles included in our sample, broken down by journal and decade. For each one of the articles, full bibliographical information was obtained from the electronic library EBSCO. Specifically for *JRER*, since EBSCO contained a number of omissions, article information was verified manually through the journal's website. Our analysis was performed on the article abstracts, which were expected to offer the best description of each article's research topic.

Following best practices in analysis of textual data, we represented our collection of documents in matrix format using the Vector Space Model (Salton, 1975). We followed standard text processing operations, including tokenization, stemming, filtering, term weighting, and dimensionality reduction, as presented in Coussement and Van Den Poel (2008). Trivial English words, such as *the*, *and*, *of*, etc. that carry no meaning associated with research topics were excluded. Terms were stemmed. As an example of term stemming, *tax*, *taxation*, *taxes*, etc., were all represented by the stemmed term *tax-*. Numerical representation (coding) of our textual data started with

the compilation of a 3644-by-2526 matrix that recorded the occurrence (frequency count) of 3,644 stemmed terms in the 2,526 article abstracts (documents). The term frequencies were then transformed using a common term frequency transformation scheme, known as *TF-IDF* transformation.[1] The effect of this transformation is to promote the occurrence of relatively rare terms, such as auction-, which appears 81 times in the entire collection of abstracts, and discount the influence of more common terms, such as properti-, which appears 1,463 times in the abstracts.[2]

LATENT SEMANTIC ANALYSIS

Our main analysis step involves Latent Semantic Analysis (Deerwester et al., 1990), where Singular Value Decomposition (SVD) is used in order to extract socially constructed components of meaning (Landauer, 2007). SVD is a matrix operation that extends Principal Component Analysis by extracting simultaneous least-squares principal components of two sets of variables, the set of terms, and the set of documents. Our term frequency matrix $\mathbf{A}$, the Vector Space Model representation of our collection of abstracts, was thus decomposed into term eigenvectors $\mathbf{U}$, document eigenvectors $\mathbf{V}$, and the diagonal matrix of singular values $\mathbf{\Sigma}$, as $\mathbf{A} = \hat{\mathbf{U}}\mathbf{\Sigma}\mathbf{V}^{\mathbf{T}}$. Following the factor analysis approach to LSA (Evangelopoulos, Zhang, and Prybutok, 2012), these matrices were combined to produce factor loadings for terms $\mathbf{U\Sigma}$ and factor loadings for documents $\mathbf{V\Sigma}$. The loadings were then rotated through varimax rotations of the term loadings, in order to produce interpretable content factors that are conceptually distinct. Factor rotations were reciprocated by the document loadings, using the same rotation matrix for both term and document loadings. Since we were interested in exploring the semantic structure of the real estate discipline, as represented by our set of 2,526 abstracts, we produced a number of factor solutions that correspond to a high level of semantic granularity (2- to 6-factor solutions), as well as a number of factor solutions corresponding to a medium level of semantic granularity (20, 25, 30, 40, and 50-factor solutions). Note that traditional dimensionality selection criteria such as keeping principal components that have eigenvalues greater than $\bar{\lambda}$ (which would produce 355 factors), or those that account for 85% of variance explained (which would produce 547), are not suitable for the purpose of summarizing a large collections of documents. Therefore, we followed the recommendation in Evangelopoulos, Zhang, and Prybutok (2012) and examined a number of alternative factor solutions qualitatively. After examining 10, 20, 30, and 50 factors, we determined the range of 20 to 30 to be the most descriptive of RE research. Then we examined the 25-factor solution. After a comparison of the 20-, 25-, and 30-factor solutions, we concluded with the 25-factor solution as our final choice.

For each one of the examined solutions, in a way similar to what is typically done in factor analysis of numerical data, we obtained high-loading terms and high-loading documents for each factor, and used them to label the factors. The rotated factors were interpreted and labeled through a co-examination of each factor's high-loading terms and documents. For example, Exhibit A1 in the Appendix presents the top-loading terms for factors F25.1 and F25.2. Exhibit A2 in the Appendix presents the top high-loading documents for the same two factors. It is obvious that the first factor

discusses research methods used in RE research, such as hedonic regression. Our label for factor F25.1 is "Research Methods." Similarly, it is obvious that the second factor primarily discusses real estate investment trusts (REITs). Our label for F25.2 is "REITs." Following a similar examination of high-loading terms and documents, all factors in each solution were labeled. Each of the co-authors produced independent sets of labels, which were then compared and reconciled without controversy.

RESULTS

SEMANTIC DISAGGREGATION OF RESEARCH IN REAL ESTATE: A HIGH-LEVEL VIEW

Results for the high-level aggregation factors are presented in Exhibit 2. The variance explained by each latent semantic factor was computed as that factor's sum of squared loadings over all 1,195 terms or, equivalently, all 2,526 articles. The figures shown in Exhibit 2 are the corresponding percentages out of the total variance explained by the total number of factors in each solution (2 to 6, depending on the solution). The high-loading articles are defined as articles whose loading exceeds a certain threshold. The threshold was computed empirically based on the heuristic assumption that each article loads on average on one factor. The actual threshold values used in the factor solutions with 2–6 factors presented in the table were 0.1181, 0.1159, 0.1182, 0.1273, and 0.1327, respectively. The total number of high-loading articles in the second column was, therefore, fixed to be 2,526, equal to the total number of articles in our sample. Small discrepancy in certain factor solutions is due to rounding. Individual articles, however, may not load on exactly one factor. A large number of articles cross-load on two or more factors. Those double and triple counts were offset by a corresponding number of weakly loadings articles, which failed to exceed the applicable loading threshold on any of the factors.

The two-factor solution represents the broadest themes represented in RE research over the study period. The five highest loading articles in the first factor, *Real Property Markets* (F2.1), specifically address topics such as hedonic pricing and rental amenities in apartments, cash assistance programs, residential property taxes, and office market rents. The 25 highest loading articles are almost exclusively focused on issues related to house prices, taxes, apartment rent or office rent. This is in contrast with the 25 highest loading articles in the second factor, *Real Estate Portfolios* (F2.2), which represent publications focused on REITs, portfolio diversification, and real estate returns.

In the three-factor solution, the topics in the first two factors, *Real Property Markets* (F3.1) and *Real Estate Portfolios* (F3.2) remain stable, with high-loading articles representing the same or similar publications as the two-factor solution. The five highest loading articles in the third factor, *Real Estate Finance* (F3.3), are specifically focused on mortgage termination, prepayment, and default, and the 25 highest loading articles include those topics, as well as financing shopping centers, mortgage pricing, credit, and adjustable-rate mortgages.

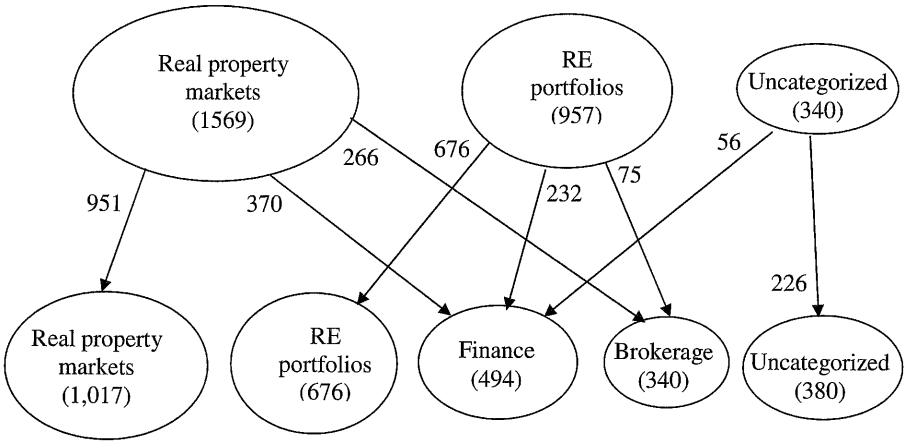
Exhibit 2
High-level Semantic Factors in Real Estate Research

Factor	Label	% Var. Expl.	High-loading Article Count			
			3 Journals	*JREFE*	*JRER*	*REE*
F2.1	Real property markets	52.98	1,569	476	444	649
F2.2	RE portfolios	47.02	957	320	274	363
	Total	100.00	2,526	796	718	1,012
F3.1	Real property markets	36.67	1,198	340	376	482
F3.2	RE portfolios	33.60	758	251	245	262
F3.3	Finance	29.74	572	211	97	264
	Total	100.00	2,528	802	718	1,008
F4.1	Real property markets	30.12	1,017	296	302	419
F4.2	RE portfolios	27.01	676	224	226	226
F4.3	Finance	24.00	494	187	79	228
F4.4	Brokerage	18.87	340	96	136	108
	Total	100.00	2,527	803	743	981
F5.1	RE portfolios	21.88	574	190	198	186
F5.2	Real property markets	21.86	680	174	187	319
F5.3	Research, pricing & valuation methods	21.68	611	200	192	219
F5.4	Finance	19.90	428	160	67	201
F5.5	Brokerage	14.69	234	68	94	172
	Total	100.00	2,527	792	738	997
F6.1	RE portfolios	19.78	537	182	183	172
F6.2	Research, pricing & valuation methods	18.66	515	167	167	181
F6.3	Finance	17.37	394	152	60	182
F6.4	Land use, policies and regulations	17.16	564	150	155	259
F6.5	Real property markets	14.10	312	82	115	115
F6.6	Brokerage	12.92	208	60	84	64
	Total	100.00	2,530	793	764	973

Notes: The variance explained by each latent semantic factor was computed as that factor's sum of squared loadings over all 1,195 terms or, equivalently, all 2,526 articles. The figures shown here are the corresponding percentages out of the total variance explained by the total number of factors in each solution (2 to 6, depending on the solution). The high-loading articles are defined as articles whose loading exceeds a certain threshold. The threshold was computed empirically based on the heuristic assumption that each articles loads on average on one factor.

The *Brokerage* (F4.4) and *Research, Pricing & Valuation Methods* (F5.3) factors result from the addition of a fourth and fifth factor, respectively. *Brokerage* is represented by high loading terms such as broker, buyer, seller, list, and agent, and high loading documents that discuss the use of buyers' brokers, agency, and the brokerage process. High loading terms in *Research, Pricing, & Valuation Methods* include index, estimate, appraise, method, and hedonic. Documents contain information on housing price and transaction-based indices, repeat sales models, and hedonics. The six-factor solution introduces *Land Use, Policies, and Regulation* (F6.4), which addresses the impact of various regulatory schemes on the RE. High-loading terms including land,

Exhibit 3
Semantic Disaggregation Diagram



Notes: The upper layer of article clusters shows articles loading on the factors of the two-factor solution, as well as articles that loaded on those two factors only weakly ("uncategorized" articles). The lower layer of article clusters shows articles loadings on the factors of the four-factor solution, as well as articles that load on those four factors only weakly. The numbers in parentheses show the high-loading article counts and the sizes of the oval shapes reflect those cluster sizes. All counts include cross-loading articles. The figures next to the arrows indicate article groups that move from one topic cluster to another, as the article collection is disaggregated from the two-factor level to the four-factor level.

tax, income, development, and urban represent topics such as urban land prices, land development, and property taxes.

Exhibit 3 illustrates the distribution of various articles from the two- to the four-factor solutions. The upper layer of the diagram shows the results of the two-factor solution, *Real Property Markets* (1,569 articles) and *RE Portfolios* (957 articles). The third article cluster, *Uncategorized* (340 articles), represents papers that fail to load on either of the other two factors. The arrows and corresponding numbers show the amount of articles that load from a specific category in the two-factor solution to the specific factor in the four-factor solution. For example, 266 articles from the *Real Property Markets* category in the two-factor solution load onto the *Brokerage* category in the four-factor solution. *Brokerage* also draws 75 articles from the *RE Portfolios* category, for a total of 341. The total differs somewhat from the *Brokerage* total (340 articles), due to cross-loading articles which result in double-counts.

RESEARCH TOPICS IN REAL ESTATE: A MORE DETAILED VIEW

Exhibit 4 shows the topics generated by the 25-factor solution. The variance explained by each latent semantic factor was computed as each factor's sum of squared loadings over all 1,195 terms or, equivalently, all 2,526 articles. The figures shown in Exhibit 4 are the corresponding percentages out of the total variance explained by the 25

Exhibit 4
Research Topics in Real Estate: 25-Factor Solution

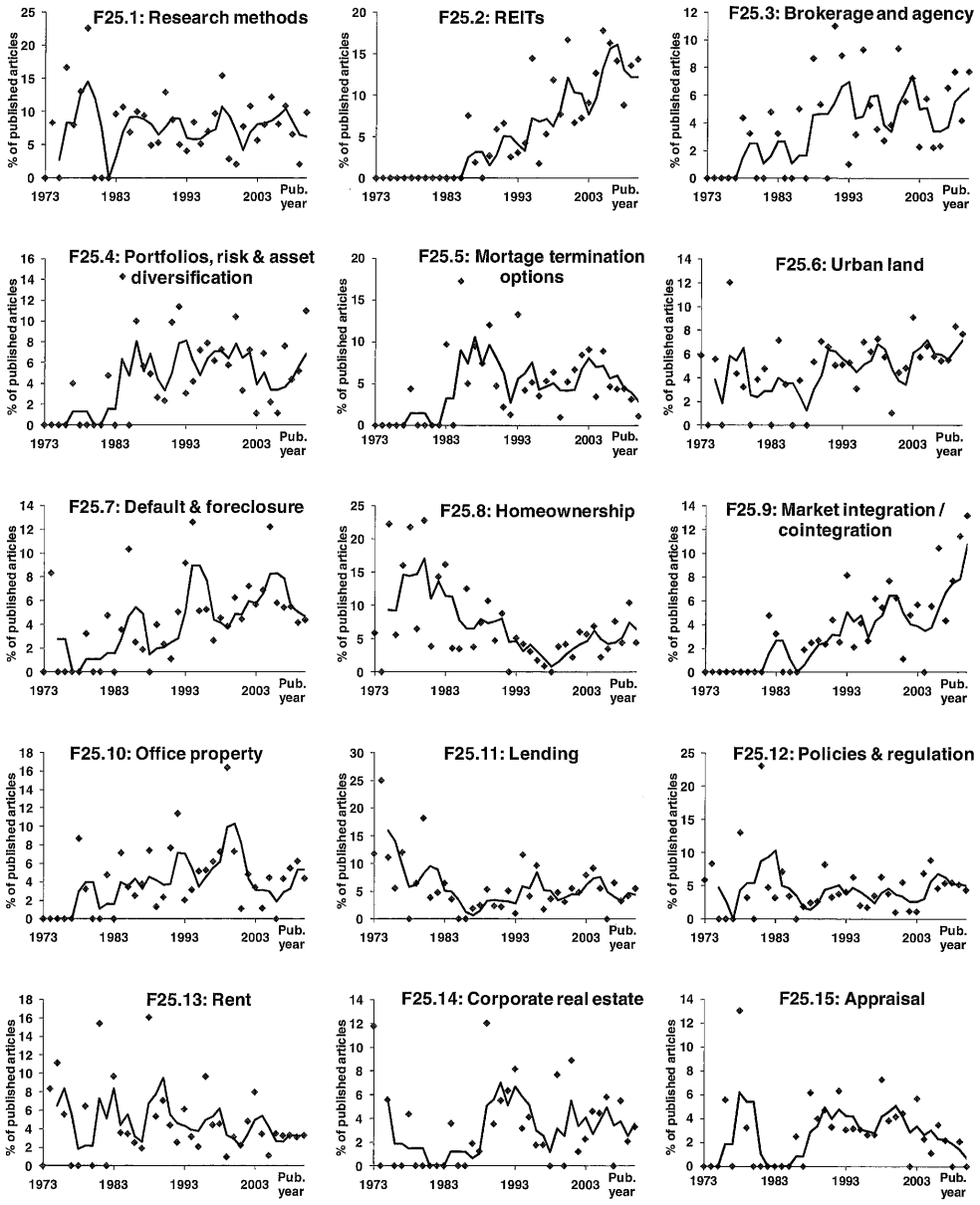
			High-loading Article Count			
Factor	Label	% Var. Expl.	3 Journals	*JREFE*	*JRER*	*REE*
F25.1	Research Methods	6.80	195	71	60	64
F25.2	REITs	5.63	193	71	60	62
F25.3	Brokerage and agency	5.23	119	33	45	41
F25.4	Portfolios, risk & asset diversification	4.97	138	36	59	43
F25.5	Mortgage termination options	4.70	131	58	16	57
F25.6	Urban land	4.58	134	45	42	47
F25.7	Default and foreclosure	4.56	124	45	19	60
F25.8	Homeownership	4.54	137	43	16	78
F25.9	Market integration/cointegration	4.49	117	60	27	30
F25.10	Office property	4.36	122	25	52	45
F25.11	Lending	4.09	125	50	26	49
F25.12	Policies and regulation	3.75	109	52	18	39
F25.13	Rent	3.73	114	33	40	41
F25.14	Corporate real estate	3.66	90	15	57	18
F25.15	Appraisal	3.47	76	19	24	33
F25.16	Tax	3.35	81	19	23	39
F25.17	Indices	3.35	70	31	13	26
F25.18	Capital and leverage	3.28	95	24	25	46
F25.19	Adjustable-rate mortgages	3.26	48	19	7	22
F25.20	Inflation	3.24	63	15	15	33
F25.21	Brokerage industry	3.22	69	15	39	15
F25.22	Journals, publications and authors	3.21	40	12	17	11
F25.23	Retail	2.98	45	10	30	5
F25.24	Auctions	2.84	46	13	16	17
F25.25	Leases	2.71	45	12	18	15
	Total	100.00	2,526	826	764	936

Notes: The variance explained by each latent semantic factor was computed as each factor's sum of squared loadings over all 1,195 terms or, equivalently, all 2,526 articles. The figures shown here are the corresponding percentages out of the total variance explained by the 25 factors.

factors. The high-loading articles are defined as articles whose loading exceeds a certain threshold. The total number of high-loading articles in the second column is, therefore, fixed to be equal to the total number of articles in our sample, but individual articles are allowed to cross-load on two or more factors.

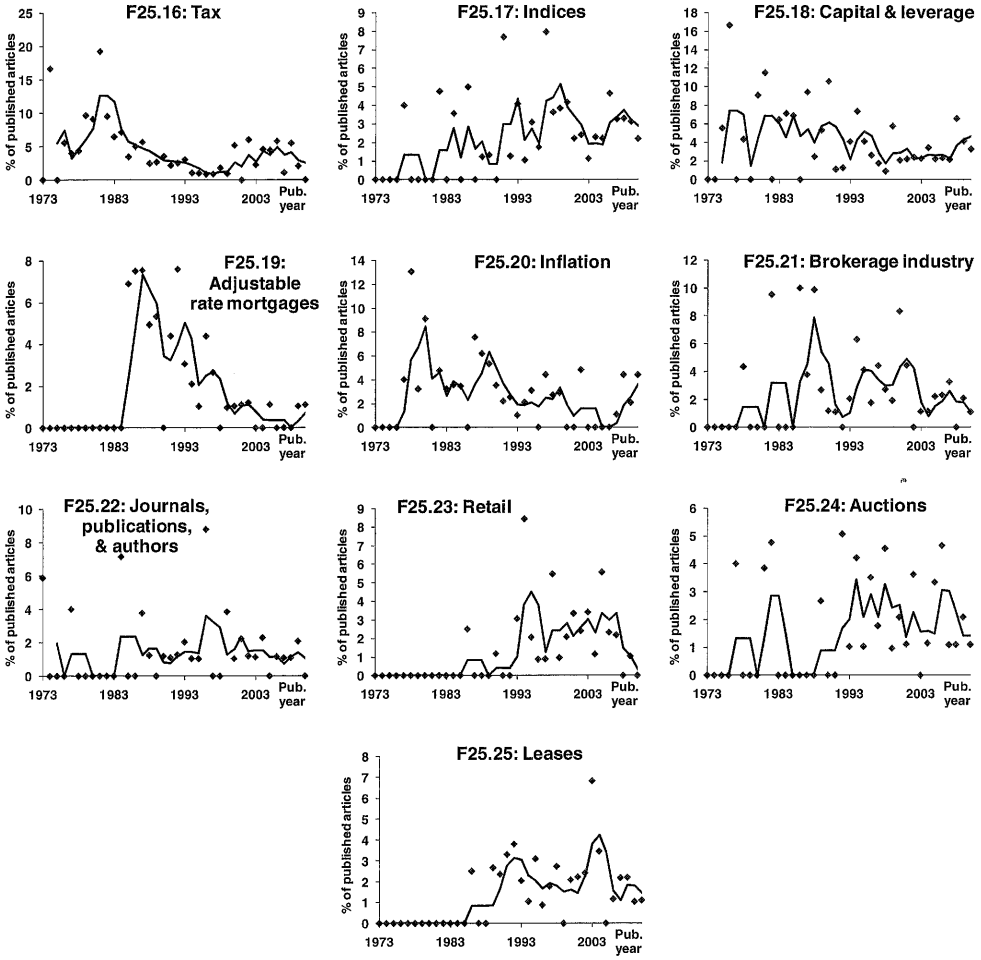
Exhibit 5 shows the percentage of articles published for each of the 25 topics between 1973 and 2010. The time series plots show the counts of articles that load high on

Exhibit 5
Trends in Real Estate Topics: 25-Factor Solution



each of the 25 factors presented in Exhibit 4, expressed as a percentage of all articles published per year in the three journals. The solid lines are the three-year moving averages of those percentages. The first chart in Exhibit 5 (F25.1) shows that *Research Methods*, including papers that address hedonic modeling, regression techniques, house price prediction, spatial analysis, and appraisal methods, have been consistently

Exhibit 5 (continued)
Trends in Real Estate Topics: 25-Factor Solution



Notes: The time series plots show the counts of articles that load high on each of the 25 factors presented in Exhibit 3, expressed as a percentage of all articles published each year in the three journals in the study. The solid lines are the three-year moving averages of those percentages.

studied and published since the inception of *REE*. The *REIT* chart (F25.2), however, tells quite a different story. Interest in REITs from a research perspective has increased dramatically since the first article was published in the mid-1980s and appears to possess continued momentum, suggesting the need for more research in the area. *Brokerage & Agency* (F25.3) shows a similar trend, although not quite as dramatic as that of the *REIT* factor.

There has been a sustained interest in *Portfolios, Risk & Asset Diversification* publications since the mid-1980s (F25.4). Publications in mortgage prepayment and termination options are not nearly as prolific now as they were in the mid-1980s to

early 1990s (F25.5). Publications in *Urban Land* (F25.6), which include papers on economic issues such as population growth, employment, economic development, and location, have remained relatively steady, with a slight upward momentum throughout the study period.

Interest in *Default & Foreclosure* (F25.7) as a research topic follows a consistent pattern, peaking about every 10 years, with an upward trend over the life of the study. One might surmise that, given the recent credit crisis, corresponding collapse in RS prices, and increases in default and foreclosure, that this topic may show additional popularity in the future. *Homeownership* (F25.8) topics including demand, household composition, pricing, and tenure steadily declined from the early 1970s through the late 1990s, reviving somewhat beginning in 2000. The subsequent increase may be related to the policies and market trends that resulted in record rates of homeownership from 2000 to 2008.

The percentage of *Market Integration/Co-integration* (F25.9) papers has consistently increased since the first article was published in 1982, addressing issues such as the relationships between equities and REITs, house values, and interest rates, and securitized real estate and stocks. Interest in *Office Property* (F25.10) research has remained relatively stable over the years. It peaked at about 17% of published articles in 1999. Research on *Lending* (F25.11), which covers subjects such as credit availability, bank regulation, and subprime mortgages, has been inconsistently published, ranging from close to 25% in the early 1970s to 0% in the late 1980s to 12% of articles published in the mid-1990s. Interest in *Policies & Regulation* (F25.12) topics such as zoning issues, investment incentives, covenants, and other land use controls remained relatively steady over the course of the study period after peaking near 25% in the early 1980s.

Studies on issues associated with *Rent* (F25.13) such as rent control, determinants, and concessions have decreased fairly regularly since the late 1980s, leveling off more recently. Interest in *Corporate Real Estate* (F25.14) surged in the late 1980s, moving from 0% to 12% of published articles in just a few years. The number of publications has dropped since then, but it is consistently addressed, averaging about 4% of published articles annually since 2000. Articles on *Appraisal* (F25.15) show a somewhat different trend, peaking in the late 1970s at about 13% and moving on a relatively steady downward trajectory since the late 1990s.

After peaking in the early 1980s, studies of property and RE-related income tax (*Tax*, F25.16) have remained between 1% and 5% annually. Frequent changes in tax law provide a continuous stream of potential research topics and consequently, publishable papers. The *Indices* factor (F25.17) includes papers written on house price indices, repeat sales indices, and real estate return indices. A number of studies of indices have been published over the time period and interest as a research topic has generally followed an upward pattern, although declining more recently. Studies of *Capital & Leverage* (F25.18) were more frequently published during the early years of the analysis rather than the latter, but the trend line does show a slight resurgence in recent years that is likely related to the liberal flow of capital into RE in the 2000s.

Not surprisingly, *Adjustable-Rate Mortgages* (F25.19) were heavily studied in the mid-1980s, a clear result of the creation of ARMs and other alternative mortgage instruments to address the extraordinarily high interest rates of the early 1980s. Interest in ARM research shows a steady decline since then, however, like *Capital and Leverage*, the upward trajectory in recent years is likely to continue as the substantial increase in ARM originations between 2000 and 2008 is widely considered to have contributed to the collapse of the U.S. real estate market. Just as is the case with ARMs, studies of *Inflation* (F25.20) also peaked in the early to mid-1980s, a time when inflation expectations and Federal Reserve monetary policy caused the rate on the standard, fixed-rate mortgage to hover near 20%.

Studies of the *Brokerage Industry* (F25.21) have been published steadily since the inception of *REE*, at about 4% of the total articles published in a given year. There are, however, a few notable exceptions to this pattern as the number of published articles spiked in the 1980s at about 10%–12% of published research, and in 2000 at 8%. There has also been a relatively steady stream of research on *Journals, Publications & Authors* (F25.22) at about 1% of published research, with three notable exceptions: in 1973 at the inauguration of the *Journal of AREUEA* with 6% of the published research, at 7% in the mid-80s when the *Journal of American Real Estate and Urban Economics Association* celebrated its 10th anniversary, and at 9% in the mid-1990s when it was *JRER*'s turn to celebrate its first 10 years.

Very little research was published on *Retail* (F25.23) through the 1970s and 1980s. It was not until the 1990s that this trend changed, which may be a factor of data availability. Interest in retail has steadily waned since the mid-2000s. *Auction* (F25.24) articles have been published throughout the course of the study period, attracting about 2%–3% of published research on a three-year moving average since the early 1990s. Finally, research on *Leases* (F25.25) followed a very similar pattern to what was observed for retail research, emerging in the mid-1980s and comprising 1%–3% of published articles on a fairly consistent basis since the early 1990s.

CONCLUSION

This research uses Latent Semantic Analysis to analyze a large body of research from the three core RE journals: *REE*, *JREFE*, and *JRER*. Specifically, word patterns within all of the abstracts for research papers published between 1973 and 2010 (excluding editorial and book reviews) are examined to determine what topics constitute the intellectual foundation of RE research. Topics are assessed at the highest level, a two-factor solution that includes *Real Property Markets* and *Real Estate Portfolios*. Further disaggregation of the data identifies sub-topics within the two main factors, concluding with a 25-factor solution. Time series plots show how these topics have performed over time.

The time series graphs clearly illustrate three trends within the data. First, published research on several topics directly reflects changing conditions within the RE market. The clearest example of this is in the ARM factor (F25.19). The introduction of this

alternative mortgage product in the early 1980s quickly resulted in a number of publications on a topic that had never before been addressed in the RE literature. Then, as the lending market adjusted and ARM originations became more common, interest in the subject diminished and fewer articles were published.

The second trend addresses topics that are published consistently throughout the course of the time period analyzed. *Papers on Research Methods* (F25.1), for instance, comprise 5%–10% of all articles on a three-year moving average for nearly the entire study period. This is not surprising since new methods are constantly developed, critiqued, and used to research various aspects of RE.

The last trend (and possibly the most interesting) speaks to future trends and opportunities in RE research. For example, note the relationship between *REITs* (F25.2) and *Market Integration/Co-integration* (F25.9). Both topics are of specific interest to institutional investors and both have consistently grown in popularity as research topics. In addition, *Portfolios, Risk & Asset Diversification* (F25.4), an area also of interest to institutional investors, has consistently remained one of the most published. The combination of these three topics shows that substantial research opportunity may lie in the institutional arena for the near future. That is not to say, however, that other topics will not trend upward. Both *Inflation* (F25.20) and *Adjustable-Rate Mortgages* (F25.19) show recent gains that may be the result of the close relationship between these two topics to the credit crisis and corresponding reduction in home values that has occurred since 2008.

We recognize that our analysis has some limitations as there are a number of factors that may influence the topical structure of RE publications. First of all, some topics may receive minimal coverage in top, peer-reviewed journals because the data necessary to produce meaningful research may not be readily available. Conversely, data for other topics (such as REITs) may be easily acquired, resulting in a large number of high-quality papers. Furthermore, it is possible that tenure and promotion requirements may influence the direction of research as authors pursue topics with the greatest potential for publication in top journals or that the more heavily favored topics reflect the preferences of journal editors. Future research could investigate these issues.

APPENDIX

In this Appendix, we present some technical details on our Latent Semantic Analysis of the RE research corpus under study.

In order to avoid overfitting the semantic space and focus more on important concepts related to research topics, rather than spurious language patterns, we followed Sidorova, Evangelopoulos, Valacich, and Ramakrishnan (2008) and performed communality filtering: an initial round of Singular Value Decomposition performed on our term frequency matrix identified the set of terms that account for 95% of

variability explained by the top 100 principal components. For example *tax-*, appearing 449 times among the 2,526 abstracts is the top term, explaining about 0.8% of variability, whereas *real*, appearing in almost every abstract (2,492 times among the 2,526 abstracts) is near the bottom, explaining only 0.0001% of variability in term usage patterns. The communality filtering step reduced our vocabulary to 1,195 terms.

Exhibit A1
High-loading Terms for the First Two Factors in the 25-Factor Solution

	Factor Loading	
Term	F25.1	F25.2
method	1.090	
estim	0.997	
regres	0.952	
techniqu	0.874	
hedon	0.822	
spatial	0.755	
approach	0.691	
reit		2.073
trust		1.134
stock		1.050
invest		0.868
equiti		0.793
perform		0.745
institu		0.684
liquid		0.659

Exhibit A2
High-loading Articles for the First Two Factors in the 25-Factor Solution

Author(s)	Title	Journal	Factor Loading F25.1	Factor Loading F25.2
Panel A: First Factor				
Kang & Reichert	An Empirical Analysis of Hedonic Regression...	*REE*	0.431	
Dubin	Predicting House Prices Using...	*JREFE*	0.418	
Richardson & Thalheimer	Alternative Methods of Variable Selection...	*REE*	0.417	
Han & Reichert	An Evaluation of Alternative Estimation Techniques...	*JRER*	0.403	
Bourassa et al.	Predicting House Prices with Spatial Dependence...	*JRER*	0.401	
Isakson	The Nearest Neighbors Appraisal Technique...	*REE*	0.397	
Panel B: Second Factor				
Below et al.	An Examination [...] of Real Estate Investment Trusts	*JRER*		0.5307
Chan et al.	Institutional Investment in REITs...	*JRER*		0.4978
Ciochetti et al.	Institutional Investors' Preferences for REIT Stocks	*REE*		0.4734
Chan et al.	Changes in REIT Structure and Stock Performance...	*REE*		0.4539
Glascock et al.	Analysis of REIT IPOs Using...	*JREFE*		0.4455

ENDNOTES

1. The TF-IDF transformation replaces the raw term frequencies by the product $w_{ij} = (tf_{ij})(idf_i)$, where $idf_i = \log_2(N/n_i)$, N is the number of documents in the collection, tf_{ij} is the raw term frequency of term i in document j, n_i is the term frequency of term i in the entire collection of documents, and the Inverse Document Frequency (IDF) idf_i serves as a metric of rarity of term i in the entire collection of documents.
2. The practice of promoting rare term occurrences while discounting common terms is well-established in the information retrieval and text mining literature (e.g., Harman, 1992).

REFERENCES

Beauchamp, C.F., W.G. Hardin III, M. Hill, and K. Liano. The Finance and Real Estate Publications of Real Estate Editorial Board Members. *Journal of Real Estate Literature*, 2008, 16:1, 23–32.

Clauretie, T.M. and N. Daneshvary. A Note on the Ranking of Real Estate Authors: Where Else Do They Publish and Who Cares? *Journal of Real Estate Research*, 1993, 8:3, 445–53.

Coussement, K. and D. Van Den Poel. Improving Customer Complaint Management by Automatic Email Classification using Linguistic Style Features as Predictors. *Decision Support Systems*, 2008, 44:4, 870–82.

Deerwester, S., S. Dumais, G. Furnas, T. Landauer, and R. Harshman. Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 1990, 41:6, 391–407.

Diaz III, J.D., R.T. Black and J.S. Rabianski. A Note on Ranking Real Estate Research Journals, *Real Estate Economics*, 1996, 24:4: 551–563.

Dombrow, J. and G. Turnbull. Individual and Institutional Contributors to the Journal of Real Estate Finance and Economics: 1988–1999. *Journal of Real Estate Finance and Economics*, 2000, 21:2, 203–14.

——. Individuals and Institutions Publishing Research in Real Estate—1989–1998. *Journal of Real Estate Literature*, 2002, 10:1, 45–92.

——. Trends in Real Estate Research, 1988–2001: What's Hot and What's Not. *Journal of Real Estate Finance and Economics*, 2004, 29:1, 47–70.

Evangelopoulos, N., X. Zhang, and V. Prybutok. Latent Semantic Analysis: Five Methodological Recommendations. *European Journal of Information Systems*, 2012, 21:1, 70–86.

Harman, D. Ranking Algorithms. In W. Frakes and R. Baeza-Yates (eds.), *Information Retrieval: Data Structures & Algorithms*. Upper Saddle River, NJ: Prentice-Hall, 1992, 363–92.

Gibler, K.M. and X. Zhou. Changing Real Estate Journal Perceptions over Time. *Journal of Real Estate Practice and Education*, 2010, 12:2, 107–30.

Gibler, K.M. and A.J. Ziobrowski. Authors' Perceptions and Preferences among Real Estate Journal. *Real Estate Economics*, 2002, 30:1, 137–57.

Hardin, W.G., K. Liano, and K.C. Chan. A Citation Proportions Evaluation of Real Estate Research. *Journal of Real Estate Literature*, 2007a, 15:3, 383–96.

——. Influential Journals, Institutions, and Researchers in Real Estate. *Real Estate Economics*, 2007b, 34:3, 457–78.

Harrison, D.M. and C. Manning. Characteristics of Recent Real Estate Research: 2000–2006. *Journal of Real Estate Practice and Education*, 2008, 11:2, 109–26.

Hendershott, P.H., T.G. Thibodeau, and H.C. Smith. Evolution of the American Real Estate and Urban Economics Association. *Real Estate Economics*, 2009, 37:4, 559–58.

Johnson, L.L., S.E. Roulac, and R.A. Followill. American Real Estate Society Annual Meeting Paper Presentations: The First Decade (1985–94). *Journal of Real Estate Research*, 1996, 12: 2, 195–207.

Jud, G.D. The Journal of Real Estate Research: A Ten-Year Review. *Journal of Real Estate Research*, 1996, 12:2, 249–313.

Landauer, T. LSA as a Theory of Meaning. In T. Landauer, D. McNamara, S. Dennis, and W. Kintsch (eds.), *Handbook of Latent Semantic Analysis*, Mahwah, NJ: Lawrence Erlbaum Associates, 2007, 3–32.

Redman, A.L., H. Manakyan, and J.R. Tanner. The Ranking of Real Estate Journals: A Citation Analysis Approach. *Financial Practice and Education*, 1998, 8:2, 59–69.

——. A Normalized Citation Analysis of Real Estate Journals. *Real Estate Economics*, 1999, 27:1, 169–82.

Sa-Aadu, J. and J. Shilling. Rankings of Contributing Authors to the AREUEA Journal by Doctoral Origin and Employer: 1973–1987. *Journal of the American Real Estate and Urban Economics Association*, 1988, 16:3, 908–10.

Salton, G. A Vector Space Model for Automatic Indexing. *Communications of the ACM*, 1975, 18:11, 613–20.

Sidorova, A., N. Evangelopoulos, J. Valacich, and T. Ramakrishnan. Uncovering the Intellectual Core of the Information Systems Discipline. *MIS Quarterly*, 2008, 32:3, 467–82 and A1–A20.

Urbancic, F.B. Contributors to the Journal of Real Estate Research: The First Twenty Years. Journal of Real Estate Practice and Education, 2007, 10:1, 81–106.

Webb, J.R. and J.D. Albert. Evaluating the Real Estate Journals: The Mainstream Finance Perspective. *Journal of Real Estate Research,* 1995, 10:2, 217–26.

Kimberly Winson-Geideman, University of North Texas, Denton, TX 76201 or geideman@unt.edu.

Nicholas Evangelopoulos, University of North Texas, Denton, TX 76201 or Nick.Evangelopoulos@unt.edu.