# Rhetoric in the Language of Real Estate Marketing*

*Gwilym Pryce*
*Department of Urban Studies,*
*University of Glasgow,*
*g.pryce@lbss.gla.ac.uk*

*Sarah Oates*
*Department of Politics,*
*University of Glasgow*
*s.oates@socsci.gla.ac.uk*

**Abstract:**

"Des. Res.", "rarely available", "viewing essential" – these are all part of the peculiar parlance of housing advertisements which contain a heady mix of euphemism, hyperbole and superlative. Of interest is whether the selling agent's penchant for rhetoric is spatially uniform or whether there are variations across the urban system. We are also interested in how the use of superlatives varies over the market cycle and over the selling season. For example, are estate agents more inclined to use hyperbole when the market is buoyant or when it is flat, and does it matter whether a house is marketed in the summer or winter? This paper attempts to answer these questions by applying textual analysis to a unique dataset of 49,926 records of real estate transactions in the Strathclyde conurbation over the period 1999 to 2006. The analysis opens up a new avenue of research into the use of real estate rhetoric and its interaction with agency behaviour and market dynamics.

* A revised version of this paper has been published in Housing Studies: Pryce, G. and Oates, S. (2008) 'Rhetoric in the Language of Real Estate Marketing', *Housing Studies*, 23(2), pp.319-348.

# 1. Introduction

Analysis of the transactions process has traditionally focussed on pricing strategies (Smith et al 2006; Levin and Pryce 2007), time on the market (Haurin 1988; Pryce and Gibb 2006), the bidding/bargaining process (Levin and Pryce 2007; Merlo and Ortalo-Magné 2004) and broker behaviour in response to financial incentives (Munneke and Yavas 2001). Much of this literature assumes that the data disseminated by agents is informative or neutral, rather than manipulative or emotive. Realtors, in economic models at least, are typically assumed to be dispassionate profit maximisers – market intermediaries lubricating the dynamics of the market by mitigating information imperfections. It is only very recently that the textual content of property advertisements has itself been considered worthy of research (Levitt and Syverson 2005) and, even then, its sphere of influence is assumed to be peripheral.

Yet, the notion of estate agents as impartial information disseminators contrasts strongly with their common perception by the media and the general public. The idiom used by estate agents is perhaps the single most important determinant of their popular characterisation. “Des. Res.”, “rarely available”, “viewing essential” – these are all part of the peculiar parlance of housing advertisements that contain a readily identifiable combination of euphemism, hyperbole and superlative. Indeed, it is the realtors’ idiosyncratic use of language that has marked them out as objects of ridicule. Many of the jokes about estate agents would be devoid of meaning if there were not an accepted assumption about their "flexible" use of language, as the following extracts from humorous “dictionaries” of estate agent euphemism demonstrate:

| | |
|---|---|
| *Benefits From*: | Contains a feature you may expect to be the bare minimum for the extraordinary price you are paying. Example: "Benefits from roof, floors, walls". (BBC News Online 2002) |
| *Bijou*: | Would suit contortionist with growth hormone deficiency. (*Ibid*) |

| | |
|---|---|
| *Compact*: | See Bijou, then divide by two. (*Ibid*) |
| *In Need of Modernisation*: | In need of demolition. (*Ibid*) |
| *Internal Viewing Recommended*: | Looks awful on the outside. (*Ibid*) |
| *Original Features*: | Water tank still contains cholera bacterium. (*Ibid*) |
| *Studio*: | You can wash the dishes, watch the telly, and answer the front door without getting up from the toilet. (*Ibid*) |
| *Secluded location:* | It was in the middle-of-nowhere - barren and desolate. Suitable film set for Mad Max 5. (Houseweb, 2006) |

It is beyond the scope of this study to verify the extent to which these prejudices about estate agents are justified. That would entail systematic comparison of estate agent descriptions, combining independent physical assessments of each property with an evaluation of how the typical use of words in estate agent descriptions contrasts with their everyday meaning. Even if such an evaluation were feasible, it is debatable whether it would be worthwhile. After all, why should we even be concerned that estate agents tend to exaggerate? In principle, consumers will simply adjust their interpretation and expectations. The humorous dictionary of estate agent speak in the BBC Online article cited above is, in one sense, an acknowledgement that this filtering process is already ubiquitous. Such dictionaries represent a tongue-in-cheek articulation of the unspoken acknowledgement that the realtor cannot help but converse in optimistic euphemism. But it is, nevertheless, a widely recognised and legitimate language. One does not actually expect a 'stunning lounge' to render one unconscious or an 'exclusive neighbourhood' to literally screen out undesirable people who want to move to the area. Rather, the hyperbole of estate agency forms an internally consistent dialect in which words take on significance within the context of house-advertising. It seems there is an understood dialogue of real estate, but it is one that moves beyond a mere description of the physical state of a property (or even a rather one-sided version of the attributes). As in many other forms of modern marketing, an appeal is being made to the human tendency to invest emotional capital

in inanimate objects. A house is 'seen as an expression of our taste and as an extension of our personality. It's a sophisticated language, but one we all understand' (Sweet, 1999, p. 15).

It follows that, if estate agents are consistent in their use of hyperbole and euphemism, their rhetoric will form a means of communication that can potentially capture the subtle dialogue of aspiration and promise underpinning the true nature of supply and demand. The apparent failure of the agent to be embarrassed by gushing property descriptions only serves to liberate potential buyers to indulge in the fantasy of lifestyle-real-estate. If the rhetoric and colloquialism of house marketing is consistent, it becomes a stable and useful medium for communication, and there is no need for concern. A handful of property viewing excursions will provide the average house hunter with the Rosetta Stone they need to make the necessary translation of all subsequent property descriptions.

But what if agents are *not* consistent in their use of language? To what extent does the pattern of exaggeration and misrepresentation *vary*? Code-breaking becomes considerably more complex when the process of decoding is itself subject to change. This was the primary innovation of the Second World War code-making machines such as Enigma, and it is the principle that underpins modern encryption. Of course, the notion that the parlance of urban property peddling varies over time and space presupposes the existence of forces sufficient to catalyse change in the evolution of realtor dialect over very short intervals of time and across relatively small distances.

This brings us to the primary focus of our paper: to consider why and whether we might expect spatial and temporal variation, and to investigate those arguments using data on the Strathclyde housing market. A variety of theories are considered, but there is at least one common

implication: if the language of selling is itself the product of market forces, then the analysis of that language has the potential to provide insights into the structural, seasonal and cyclical dynamics of market behaviour. Charting the variation of language over time and space may tell us something about the way in which the market is working and about the character and definition of local submarkets. The counter-argument is that we should expect no variation in realtor dialect, or that any such changes are merely white noise. Immutability and stochasticity thus form our null hypotheses.

All this is rather unexplored territory. As such, our paper should be viewed as an attempt at making limited headway on selected fronts rather than achieving comprehensive advancement across the board. With this caveat in mind, the remainder of the paper proceeds as follows. In section 2, we present a brief summary of the relevant qualitative literature. This is followed by an outline of possible theoretical explanations of why the language of selling might vary (section 3), and a summary of our methods and data (section 4). In section 5, the empirical results are presented (qualitative, bivariate and multiple regression). The paper concludes with a brief summary of our findings (section 6).

## 2. Qualitative Literature

While the vernacular of realtors appears to have thus far escaped serious quantitative analysis (see the Introduction and more detailed review in Oates and Pryce 2007), we should note that there have been *qualitative* investigations in related fields, such as an analysis of television property programmes (Lorenzo-Dus 2006) and in the broader literature on the nature of persuasion (Pardo 2001). Pardo's work is particularly relevant here as she articulates the common strategies used by those who attempt to persuade:

> "Persuasion is in some respects a linguistic phenomenon (persuasion may be achieved in various ways that do not involve language). In relation to argumentation it is characterized by an increase in linguistic

> resources and strategies in general (hierarchical presentation of information, tonalization, evidentialness markers, etc.). Its communicative function is to try to convince another of something. Like any other language element it is necessarily linked to power and therefore it always entails some degree of it." (Pardo 2001, p.99)

Summarising the work of van Dijk (1998) and Pardo (2001), Lorenzo-Dus explains that:

> *"the* communicative goal of persuasive texts is to convince others of something. Persuasive discourse is also a form of power… [P]ower is connected to people's minds, specifically to our wanting to control the minds of others so that they may see things as we do and act as we want them to. Giving orders is one way to achieve power. Trying to convince others – persuading them – is a more complex and subtle, yet often more effective, alternative. Moreover, for persuasion to work, persuader and persuadee must agree that the implications of non-persuasion, as it were, are worse than those of persuasion. This agreement, which is grounded on an 'implicit threat' (van Dijk, 1998), therefore lends further support to Pardo's view above that persuasion and power are connected." (Lorenzo-Dus, 2006, p.741).

While it is not obvious how the act of persuasion required in the marketing role of estate agents could entail any direct "threat", agents can in principle draw on the kind of implicit threat suggested by Lorenzo-Dus by emphasising, or at least hinting at, the negative implications of non-persuasion. They may, for example, claim that a property of a particular type is "rarely available", that it is an opportunity "not to be missed". More subtly, estate agents may select marketing phrases that draw on the lifestyle aspirations readily propagated in the property media, with an implicit threat that failure to achieve particular set of lifestyle characteristics will reflect a failure to achieve in life *per se*, or will lead to "looser connections between material and symbolic choices, and lack of tangible identity markers" (Lorenzo-Dus 2006, p.758).

Although the work of Lorenzo-Dus is potentially useful in helping us to understand the act of persuasion embodied in real estate marketing, it is clear from even a cursory reading of estate agent advertising material that there are aspects to persuasion other than the deployment of implicit threats. There are other ways to appeal to emotion. A broader framework is therefore needed if we are to develop a meaningful categorisation of the words used by estate agents to market properties. (We shall attend to this in section 4).

## 3. Why might the language of selling vary?

We argued in the Introduction that the question of greatest interest was not whether estate agent stereotypes were valid, but the extent to which realtors' use of rhetoric varies over time and space. Such variation may reveal insights into the structure and dynamics of private housing systems, and may further identify a source of market failure (on the basis that variation frustrates decoding).

We present below a number of theories that lead us to expect language variation. First, however, let us posit our null hypothesis – one that counters the notion that there will exist variation in the parlance of property marketing:

1. *Drivers of Uniformity*: realtors are sufficiently well-established as a profession to have arrived at a common set of communication norms which, in turn have led naturally to a widely accepted and stable form of marketing language. This uniformity is likely to be reinforced by household mobility and the widespread use of the internet. In the same way that television has been blamed for the cross-fertilisation of regional accents (Stuart-Smith, Timmins and Pryce 2005), the explosion of web-based property advertising has all but eradicated temporal, regional and intra-urban variation in realtor dialect.

***Theories of Temporal Variation in the Use of Language***

Consider now the arguments for non-uniformity in the use of marketing terminology:

2. *White Noise:* Sentences, whether in speech or written form, do not contain a rigid composition of word-types even when the use of language, in general, is static. Random selection (from the population of words contained in common vocabulary) leads to

3. *Lagged Response to Legislation*:[1] Following the introduction of the Property Misdescriptions Act in 1991, estate agents went through a period of excessive caution. However, eight years elapsed between the introduction of this legislation (1991) and the commencement of our data (1999), and so it is unlikely that it would have much bearing on our particular sample – it is mentioned here because of its relevance to research in this area more generally.

4. *Property Characteristics*: Use of hyperbole and emotive language is likely to vary between properties for sale because of real differences in the characteristics of dwellings, many of which cannot easily be captured through quantitative measurement. This is the rationale behind the inclusion of property descriptions in the regression analysis of Levitt and Syverson (2005). While this might lead us to anticipate variations the use of language across space due to the clustering of properties of particular types in particular areas (one of the basic motivations behind submarket analysis – see Rothenburg *et al*, 1991), it would not lead to seasonal or cyclical variations unless there were systematic changes over time in the characteristics of properties coming onto the market.

[1] Suggested by delegates at the National Association of Estate Agents Conference, March 2007, in response to a presentation of an earlier draft of this work.

5. *Cycles in Staff Composition*:[2] As the market booms, new staff are needed to cope with the rising turnover of properties. New employees are typically less experienced and more prone to hyperbole. (Experienced staff know that buyers are not easily duped, and that a more judicious approach is more profitable in the long run.) When the market slows, there are insufficient sales to maintain the expanded workforce. Staff are laid-off on a LIFO (last-in-first-out) basis, increasing the share of experienced agents, and leading to a pro-cyclical pattern in the language of selling. This process would not cause regular seasonal or secular patterns in the use of language.

6. *Irrational Exuberance:* Market value is not an intrinsic constant, and neither is the propensity to enthuse about property features. A mid-terraced house described as "well-maintained" during the dark valley of a market slump, may become "truly fantastic" at the dizzy heights of a boom. Once the zenith has passed, however, the property is seen for what it is, and more restrained descriptions once again become the norm. This theory would suggest that particular types of emotive expression – those less grounded in reason – will be more volatile and more sensitive to market swings, but will not lead to secular or seasonal movements.

7. *Strategy to Market Difficult to Sell Properties:* when a property is difficult to sell – either because the market is experiencing a downturn or because there are rarely many buyers for houses in that location or of that type – sellers may adopt a strategy of exaggerating a property's attributes in an attempt to attract interest. During a hot market, properties "sell themselves" so there is less need for strained descriptions. This theory suggests that we should anticipate the incidence of effusive language to be *counter*-cyclical – to fall during a boom and rise in a slump. It would not cause seasonal or secular patterns in the data.

[2] Suggested by delegates at the National Association of Estate Agents Conference, March 2007.

8. *Opportunity Cost of Viewing:* No-one bids without viewing, so why should the parlance of property adverts have an affect on whether (and what) buyers are willing to bid? The answer may lie in the opportunity cost of viewing a property. Buyers have a fixed (or at least optimal) window of time within which to secure a new home. Even if viewing is something of a disappointment in comparison with the agent's description, house-hunters still have a strong incentive to submit an offer. Turning down a property after viewing introduces the risk that a continued search will not yield a superior alternative within the buyer's timeframe.

   As far as the estate agent is concerned, viewing is all important because it shifts the probability of a buyer submitting a bid from zero to some positive value, and the greater the number of bidders, the greater the expected selling price, *cet par* (see Levin and Pryce 2007). Agents know that property adverts are not the basis on which purchases are made – that is not their purpose. Their purpose is simply to attract viewers.

   Language variation occurs when there are changes in the opportunity cost of viewing. While agents may not understand the theory behind the strategic manipulation of language, they will be aware that it is more profitable, at certain times of the year and in certain phases of the cycle, to use emotive terminology. When properties are selling very quickly, the effective choice-set facing a buyer may be very small, even though there are many properties coming onto the market. A buyer might view *x* properties over a particular period, but by the end of that period, only a small proportion of those properties may still be available for sale. Consequently, there is a very strong incentive during such periods to bid for a property once it is viewed, and a very strong incentive for estate agents to use any means possible to get potential buyers to view. In contrast,

during phases when selling times are long but there remains a continued stream of new properties being offered for sale (a "buyers market"), agents may have little to gain from exaggerating a property's attributes – disappointed viewers can simply go elsewhere, most notably to more trustworthy agents.

The implication of this theory is a pro-cyclical pattern in the use of hyperbole. There will also be a seasonal aspect due to pronounced changes in the opportunity cost of viewing during the course of the year. There is, for example, a traditional aversion to moving or transacting over the Christmas period (indeed, in Strathclyde, the number of transactions drops virtually to zero during the festive season), imposing a fixed horizon for many buyers. The start of the school year is also another important horizon, as is the deadline for being eligible for particular school catchments.

**Theories of Spatial Variation in the Use of Language**

With the exception of 1, 2 and 3, each of the above theories of temporal variation also have spatial implications. When differences across submarkets in the timing of the market cycle (Pryce and Gibb 2006), in long-term levels of demand, and in the quality of stock, are combined with one or more of theories 4 to 8, the corollary is geographical variation in marketing vernacular. We can add a further theory that pertains exclusively to spatial variation:

9. *Local conventions*: Given the tendency for local conventions to occur in accents, pronunciation, idiom, and terminology generally, it would be surprising if such developments did not occur in the language of selling. Local moves (which the majority tend to be) foster and preserve a common dialect between estate agents and those in the surrounding community. Elaborate idiosyncrasies are able to evolve

**Hypotheses**

How shall we choose between these theories? Where two or more lead to mutually exclusive outcomes, there is the potential to analyse our data in such a way as to reject one in favour of another. On the other hand, where theories are not mutually exclusive, no such clarification will be achievable. For example, if the incidence of emotive language is positively correlated with market buoyancy, we shall be able to reject theories 1 and 7, but that finding on its own will not allow us to choose between theories 5, 6, and 8. One important question relates to the existence of seasonal variation. Since only theory 8 predicts this outcome, the existence of seasonal variation might lead us to prefer it. However, such a finding would not preclude the veracity of other theories – it is conceivable that theory 8 could be the exclusive cause of seasonal variation but one of many drivers of cyclical variation, for example.

The following hypotheses have been constructed in an attempt to maximise the potential of our data to distinguish between theories (we highlight in bold the key phrases in each hypothesis to help the reader summarise):

*Hypothesis A:* Marketing **language is uniform** over time and space.

*Hypothesis B:* Variation in marketing language across time and space is **stationary**.

*Hypothesis C:* Variation in marketing language over time and space is positively related to variation in **housing characteristics** (this presupposes that the housing characteristics of marketed properties varies accordingly – i.e. cyclically, seasonally, spatially).

*Hypothesis D:* Variation in marketing language is related to **market buoyancy**, independent of variations in housing characteristics.

*Hypothesis E:* Variation in marketing language is **seasonal**, independent of variations in housing characteristics.

*Hypothesis F:* Particularly **emotive words are more volatile** & more strongly correlated to market buoyancy.

*Hypothesis G:* There is **spatial persistence** in the pattern of marketing language, independent of variations in housing characteristics.

These hypotheses and their implications for our nine theories are combined sequentially in Figure 1 in the form of a decision tree. Where two hexagonal boxes (denoting a hypothesis) emerge directly from the same branch (such as hypotheses D and G) then the hypotheses are not mutually exclusive. This helps us to see that it is possible, for example, for theories 6, 8 and 9 to be simultaneously true. A more complex and comprehensive diagram is possible but we have decided against this for the sake of clarity and simplicity. We also recognise that

there may be questions raised by our theories that are broader than the range our data can address – our empirical investigation will, inevitably, be less than comprehensive in its sweep (*Theory 3*, for example, is not considered at all; neither are secular trends).

[**Figure 1 Hypothesis Tests Decision Tree** <TO GO HERE>]

## 4. Methods

Before we can test these hypotheses, we need to devise an appropriate method of linguistic classification. How can one identify whether the propensity to use a particular category of language varies across space and time if there is no rationale for categorising language in the first place? We assume that property promotion entails an attempt to persuade potential buyers to view and bid for the property. The fact that the estate agent's use of language is motivated by the desire to persuade (rather than simply to disseminate) links it to the wider discussion on the analysis of *rhetoric* and indeed to Aristotle's seminal work on the subject. Aristotle decomposed the act of persuasion into three components: *ethos* (reliability of the speaker), *pathos* (the manipulation of the emotional predisposition of the audience) and *logos* (logical argument).

We apply this characterisation of the act of persuasion to around 49,926 written property descriptions published by estate agents. We find no evidence of *ethos* in our data (though we acknowledge that the use of the generic GSPC brand to market properties may represent an implicit attempt to construct a broader sense of trustworthiness and reliability). Moreover, *logos* – the listing of facts about the house – takes up the majority of words in these descriptions and there is little of interest or surprise in these particular aspects of the language of selling. Of far greater interest is the extent to which *pathos* is used and the different types of

*pathos* that the agent employs. We extend Aristotle's classification, therefore, to include the following sub-categorisation of pathos: (i) *originality*, (ii) *ambience*, (iii) *prestige* and (iv) *excitement.* We develop these categories using both our own knowledge of the Glasgow real estate market as well as an examination of many of the property descriptions from the dataset. Our plan is not only to identify which words denote pathos, but also to place every *pathos*-word into one of these four categories. This categorization process is applied to each of the 49,926 published descriptions in our data. We then measure the incidence of each category of pathos as the proportion of words in each property description that fall into each sub-category. Finally we consider how these proportions vary over time and space.

Inevitably, one of the difficulties associated with this kind of categorisation is subjectivity. We attempt to mitigate this problem by considering both broad and narrow definitions of pathos – if both yield similar results then we might tentatively conclude that subjectivity in the selection of words has no material affect on our findings. We, therefore, construct an additional generic category based on a much narrower definition of pathos which we call *Core Pathos*. We now have a total of six categories of pathos, which we detail below:

*All Pathos:* includes all words identified as being potentially emotive. This broad definition was subdivided into four mutually exclusive sub-classifications :

> ***Pathos Type I: Originality:*** These are words and phrases that evoke feelings of uniqueness as well as the prospect of being able to break from the anonymity and uniformity that characterises mass production. Such language taps into the urge to assert one's personality and individuality, to be "more than a number". The *Pathos*

Type I classification includes words such as: "character", "bespoke", "natural", "individual", "imaginative", "innovative", "original", "unique", "unusual", and "rare".

***Pathos Type II: Ambience:*** This is language that taps into particular lifestyle fantasies and 'nesting' instincts. It includes words such as "bright", "fresh", "charming", "attractive décor", "deluxe", "fashionable", "elegant", "stylish", "pleasant", and "mature".

***Pathos Type III: Prestige:*** This type of rhetoric appeals to our desire for respect, status and admiration. The agent is attempting to suggest that to live in this property and/or locality is a signal that the owner has achieved a certain status in society. This suggests that with ownership comes the perception of success (see de Boton's 2004 "Status Anxiety"). The *Pathos* Type III classification includes words such as: "exclusive", "executive", "enviable", "prestigious", "up-market", and "successful".

***Pathos Type IV : Excitement:*** Such words are a consequence of (or an attempt to foster and exploit) the excitement and giddiness that comes from the purchasing process itself – the "retail therapy" element of house purchase. So agents use superlative adjectives to evoke excitement about a property. However, the employment of these words or phrases may betray the difficulty of using more precise and informative description because, in reality, the property has little going for it. Examples of this kind of description include: "!", "amazing", "breathtaking", "deceptively", "fantastic", "generous", "immaculate", "incredible", "too many features to", "well", and "wow".

***Core Pathos:*** This is our second generic measure of pathos (the other being *All Pathos*). It is based on a narrower selection of words, including only those identified as being unambiguously emotive ("Preferred", "Lovely", "Exceptional", "Prime", "Generous", "Outst", "Fant", "Excl", "Beautiful ", "Charm", "Impress", "Sought after", "Superb", "Stun", "Del", "Magnif", "Pleas", "Unique", "Sunny", "Professional", "Enviable", "Prestig", "Splend", "Prestigous", "Smart", "Character", "Executive", and "Eleg").

We use the PATHOS program (see Oates and Pryce 2007, Appendix 1) to count how many times each of these words occurred in each of the 49,926 property adverts, which in turn allows us to calculate the incidence of each category of pathos in each of those adverts.

## Textual Analysis

Having established a framework for categorising language, our next step is to conduct a detailed textual examination of a selection of descriptions with a view to framing the subsequent quantitative investigation. We employ a modified version of qualitative analysis of texts used primarily in the context of political persuasion in party manifestoes, political advertising, candidate statements and election news broadcasts. This means that we examine both words and phrases to look for trends and patterns. However, as with work by Budge *et al* (1987) on political party manifestoes, we attempt to go beyond merely counting words. We look at both how often the word appears (to establish which words were the most common in the language of real estate pathos) and how these words are used. Just as one can track and identify the construction of particular political themes around particular words and phrases (Oates 2006), one can find a real estate rhetoric that is measurable across time and space. This qualitative analysis is aimed at discussing the meanings of the words within the context of the adverts. While the quantitative analysis accounts for the presence of the word, the qualitative

element will attempt to uncover any trends that would be missed by quantitative analysis. For example, are some words now so ubiquitous that they are devoid of meaning? Are some used in surprising or unexpected ways? Are some frequently paired together? This widens and deepens our understanding of the pathos of real estate and helps us to define the *Core Pathos* category.

## Statistical Analysis

The final stage in our research involves applying bivariate and multiple regression analysis in an attempt to test the hypotheses presented above. Bivariate analysis allows us to plot summary measures of pathos over time and across space to confirm whether variation does indeed occur. For example, we conduct equality of means t-tests to investigate apparent differences in the incidence of pathos across submarkets, and compute the coefficient of variation of Type IV pathos and of the other subcategories of pathos to investigate whether Type IV pathos is indeed more volatile.

The disadvantage of bivariate analysis is that it does not hold constant other factors. We remedy this by applying multiple regression. Because the dependent variable of interest – the incidence of pathos in the language of selling – is a proportion, it is bounded at zero and one and therefore violates the assumptions of ordinary least squares regression (OLS). Consequently, we use Fractional Logit Regression (FLR) to estimate the proportionate change in odds caused by each of the independent variables (see Papke and Wooldridge, 1996, for details of the FLR method; see Hendershott and Pryce, 2006, for a recent application of FLR in the real estate literature).

For sake of brevity, we generally avoid reporting the sample sizes of each test and associated significance levels in the text if sig. ≤ 0.01. In other words, unless the null hypothesis has at least one in a hundred chance of being rejected incorrectly, we shall avoid listing the significance level (full details are recorded in Oates and Pryce, 2007, however).

## Data

Our analysis is based on information extracted from 49,926 property transactions in the Strathclyde conurbation (for details of the Scottish house selling system see Smith *et al*., 2006, and Pryce and Gibb, 2006). The data, supplied by GSPC (Glasgow Solicitors Property Centre – a consortium of estate agents in the West of Scotland), covers the period 1999-2006. The data includes the text used to describe each property sold by GSPC member firms, along with basic property attribute and location information. At the start of this period, the market was relatively stagnant and properties were taking more than 150 days on average to sell (see Figure 2). A boom period then ensued. By 2004, selling times had plummeted to around thirty days and annual house price inflation rose to over thirty percent in some areas. By 2005 the market had started to slow, but remained significantly more buoyant than it was in 1999.

Table 1 provides basic summary information on our data. We can see from the information on Pathos_n in Table 1 that there were, on average, around two *pathos* words, and 0.4 Core Pathos words used in each description. The average total number of words in each description was 32. Thus, the proportion of words in each description that are classified as *pathos* words (Pathos_p) and *Core Pathos* words (Pathos_Core_p) was around 6% and 1% respectively. While *pathos* words only comprise a relatively small proportion of the words used – most of the property description is typically devoted to listing of attributes – there were relatively few

properties (14%) that had a property description that did not include any pathos words (noPathos). In contrast, around 67% did not have *Core Pathos* words (noCOREPathos).

Around half of the properties in the dataset are flats and half are houses of various types (6% are bungalows, 10% are detached, 8% are terraced). 12% of properties are made of stone, 16% have bay windows, 29% have a garage and 70% have a garden. Our data also include information on the location of the property, including the deprivation score (supplied by Communities Scotland) which ranges between 2.0 and 16.2, where the higher the score the greater the deprivation. We have also calculated the distance to the centre of Glasgow from each of the properties in the data. We find that on average properties are located 12.7 km from the city centre.

**[Figure 2 Average Marketing Time Since 1999 <TO GO HERE>]**

**[Table 1 Descriptive Statistics <TO GO HERE>]**

## 5. Results

***Results of Qualitative Analysis***

From the GSPC house descriptions, we identified the most popular words and word fragments that could be construed as denoting pathos (see Table 2 of Oates and Pryce 2007). We examined every word that appeared more than 100 times in the database. Each word was studied in a sample of the adverts to look at how the words were used in context. This allowed us to consider in more detail which words and fragments seemed to have pathos-type resonance within the context of the description, and which ones seemed to be banal 'filler' or rote phrases.

Many adjectives that could reflect pathos are frequently used in conjunction with other words, such as 'bright' in 'bright and spacious' or 'mature' in 'mature gardens'. In order for a word to qualify unambiguously as pathos, it needs to be used as a relatively flexible adjective instead of as part of a 'canned' phrase with little meaning. Although the authors had to be somewhat subjective about judging the relative pathos resonance of a word, we attempted to be as scientific as possible by eliminating words that have fallen into a sort of estate-agent jargon and identifying those with emotional content in the context of house adverts. To qualify as true real estate pathos, the word had to have an elusive and somewhat flexible meaning, to function beyond the rather dry and trite phrases ('must view' etc.) found in many adverts.

This analysis led to the definition of *Core Pathos* as a smaller subset of the wide number of rhetorical words that estate agents use in their house descriptions (see section 4, Methods, above). In the thicket of hackneyed phrases, some language still seems to hold a fairly emotive and somewhat distinct sense. It may be these particular words that can captivate the buyer. It is interesting to note that only a few brave estate agents venture into unusual language. For

example, in all of the adverts, there is only one house that is described as having ‘tremendous’ proportions. Artistic references also are rare, although those who follow the debate over the relative merits of Glasgow architects Alexander ‘Greek’ Thomson and Charles Rennie Mackintosh may be interested to note that there are 20 references to Thomson in the adverts and only two to Mackintosh (and one misspelled) in the 22,613 GSPC adverts from the Glasgow Local Authority area.

In terms of what would be the most appropriate measure for use in our quantitative analysis, there is a case for using a measure of pathos that is as broad as possible. This is because the incidence of pathos is generally so low that omission of a potentially relevant word could cause disproportionately large distortions in the regression results, while the inclusion of words that turn out to be irrelevant (i.e. words that are really just "fillers") would simply increase the white noise of the regressions and not actually cause bias or inconsistency. On the other hand, inclusion of “filler” words that comprise the relatively meaningless bulk of generic estate agent speak could muddy the meaning of our dependent variable and lead to dampened estimates of the responsiveness of pathos language to market cycles and spatial variation. As a result, we present regression results based on both our broad definition of pathos (along with it’s four subcategories) and also our narrow definition (which recognises only the core words as being truly pathos).

***Results of Bivariate and Graphical Analysis***

The first two stages in our decision tree (boxes A and B in Figure 1) require us to consider whether there is any variation at all across time and space (Hypothesis A) and whether or not this variation is stationary white noise (Hypothesis B). Both these hypotheses can be rejected from even a cursory examination of the data (see Figures 3 to 9 below). If we run equality of means t-test for properties coming onto the market in 1999 compared with 2006, we find that null hypothesis of equal average incidence of pathos is conclusively rejected for both *All Pathos* and *Core Pathos*. Similarly, an equality of means t-test for the City of Glasgow compared with the rest of Strathclyde, conclusively rejects the null hypothesis of equal average incidence of *All Pathos*. The same is true if we compare submarkets within the City of Glasgow. For example, the null of equal pathos is rejected if we compare the West End with North Glasgow, and if we compare the West End with the South Side. Similar rejections of uniformity arise if we compare North Lanarkshire with South Lanarkshire, Renfrewshire with East Renfrewshire, and North Ayrshire with East Dunbartonshire. All these tests yielded equivalent results when we used *Core Pathos* instead of *All Pathos*. While we shall consider the nature of this variation in more detail below, it is already clear that there exists non-stationary variation in the incidence of pathos both over time and across space (and so we can reject Hypotheses A and B).

The next step in our decision tree is to ascertain whether the incidence of pathos is related to property attributes (Hypothesis C). One might anticipate, for example, that there is more to boast about when marketing larger, more expensive dwellings. And indeed this appears to be

the case. Comparing 1 and 2 bedroom properties, we find that the incidence of pathos is higher for the latter. The same is true when comparing 2 and 3 bedroom properties (one-tail sig. *All Pathos* = 0.013; two-tail sig. *All Pathos* = 0.026; one-tail sig.) and when comparing 3 and 4 bedroom properties. Although the average incidence of pathos is higher for detached villas compared with semi-detached villas, the difference is not significant for *All Pathos* (one-tail sig. = 0.266; two-tail sig. = 0.532; $n_{\text{detached}}$ = 3,397; $n_{\text{semi}}$ = 7,742), though it is significant for *Core Pathos.* The same was true when comparing semi-detached villas with detached bungalows.

The ultimate single measure of the quality and size of a property is its selling price so we attempt to verify Hypothesis C by comparing the incidence of pathos across price bands. Because the threshold for expensive properties shifts significantly over time, we have to find a way of defining price bands that incorporates this movement. Our approach allocates each property in the sample to one of five bands based on its relative selling price at the time of sale (see Oates and Pryce 2007 for details). Comparing the incidence of pathos between price bands is potentially problematic because it is possible that pathos has an effect on price independent of true housing quantity/quality effects. The wider the price difference between bands being compared, however, the less likely it is that any observed pathos differences are due to endogeneity. For example, even if pathos did have a material affect on price, it is very unlikely to shift the price from band 1 to band 5. Comparison of means tests reject the null of homogenous pathos between price bands: the higher the price band, the higher the average incidence of pathos in each case (sig. = 0.000 in every instance for both 1 and 2 tailed tests, for both *Core Pathos* and *All Pathos*). In some cases the difference is very large – for example, the incidence of *Core Pathos* was nearly four times greater in price band 5 compared with price band 1.

Given that Pryce and Gibb (2006) report significant variation in dwelling characteristics across submarkets, it is likely that the correlation between pathos and housing attributes might explain some of the observed variation in language across space. However, it is far less clear that changes in dwelling attributes would explain variation in the language of selling over time, particularly any seasonal variation that we might observe. For example, there is no significant difference in the average number of bedrooms of properties that sell during the summer, compared with those that sell during the autumn (sig. = 0.2000) or spring (sig. = 0.4286). Nevertheless, taken together, these findings suggest that we should accept Hypothesis C (Theory 4), and control for housing attributes when testing the remaining hypotheses. Holding attribute effects constant is best attempted using multiple regression, which we apply as the final step in our analysis. We shall, however, continue with our application of bivariate and graphical analysis as a precursor.

To test Hypothesis D, we need to choose a measure of buoyancy. Three are immediately obvious and readily available from our data: (i) prices; (ii) time-on-the-market; and (iii) number of GSPC sales. We plot the incidence of pathos against each of these in Figures 3, 4 and 5. The incidence of pathos is measured as the proportion of words in each new description of a house issued in that quarter. In other words, in each quarter we use the descriptions of properties just coming onto the market in that quarter to ascertain the incidence of pathos. Time on the market and house price data are based on properties just leaving the market in each quarter. We have calculated the measurements in this way in an attempt to isolate the response of estate agents to the market (rather than the other way round). Because time-on-the-market (TOM) falls as the market becomes more buoyant, we have plotted the inverse (1/TOM) to make the correlation (or lack of correlation) easier to identify. We would expect

house prices and number of sales to rise as the market becomes more buoyant (see Stein 1995 for an explanation of why this is likely to be the case).

Note that there is no perfect correlation even among the measures of market buoyancy. It is therefore highly unlikely that the incidence of pathos will be perfectly correlated with any of these measures. Nevertheless, it is clear from these graphs that there is indeed a strong correlation between the incidence of *pathos* and market buoyancy. Generally, the incidence of *pathos* in the language of selling rose as the market boomed (2001-2004) and has declined as the market slowed (2005-2006). In a simple quarterly time-series linear regression of the incidence of *pathos* on each of these three measures of market buoyancy we find the following $R^2$ results (slope coefficient is positive – the correct sign – in each case): $R^2$ = 77.92 % for the house sales regression; $R^2$ = 57.93% for the time-on-the-market regression; and $R^2$ = 59.48% for the price regression. If we run a quarterly time-series multiple regression of the incidence of *pathos* on all three variables, the Adjusted $R^2$ comes out at 87.18%, with the following t-ratios: 4.50 for the price variable, -4.83 for the time on the market variable and 0.80 for the number of sales variable (based on White's standard errors; n = 30).

If we had to choose a single measure of market buoyancy from these three alternatives, we would do well to choose time-on-the market, partly because it had the highest t-ratio in the time-series multiple regression, and partly because it is free of the significant measurement issues associated with the other two indicators. House prices, for example, particularly at the level of individual transactions, are complicated by the heterogeneity of dwelling size, quality and location. Number of sales would also be problematic if analysed at the micro level because our data are drawn exclusively from properties that were sold through the GSPC

consortium of estate agents and so we would have to grapple with the possibility of the GSPC market share shifting over time in particular areas.

The results of these simple time-series graphs and regressions are useful, however, in that they provide an initial indicator of how estate agent rhetoric varies over time. Taken together, the results indicate that the incidence of pathos does indeed vary pro-cyclically (i.e. rises as the market rises and falls as the market falls) which suggests that we should reject theory 7 – *Strategy to Market Difficult to Sell Properties* – which predicted counter-cyclical variation in pathos.

But what of the *type* of pathos? Does this also vary over time? To investigate we calculated the number of Type I, II, III and IV words as a proportion of the total number of pathos words in each description. We then calculated the average incidence of each of these types (as % of pathos words) for the whole of Strathclyde for each quarter. The results are presented in Figure 6 and suggest that, while the incidence of pathos words change pro-cyclically over the course of the market cycle, the relative shares of pathos words that fall into each of our four categories do not change radically over time. There is some indication that the proportion of Type IV words does follow a pro-cyclical pattern, and that the proportions of Type II and Type III words converge as the market booms and then diverge as it slows. Note that if we were to plot each line on a separate graph the variation would look more pronounced. Nevertheless, the relative ordering of the size of each of our four categories does not change (or only briefly in the case of Type II and Type III words).

**[Figure 3 House Prices and the Incidence of Pathos in the Language of Selling <TO GO HERE>]**

**[Figure 4 Time on the Market and the Incidence of Pathos in New Descriptions <TO GO HERE>]**

**[Figure 5 Number of GSPC Sales and the Incidence of Pathos in New Descriptions <TO GO HERE>]**

**[Figure 6 Variation in the Type of Pathos Over Time <TO GO HERE>]**

Consider, now, Hypothesis G, that the use of *pathos*, and the type of *pathos*, will vary over space due to local language conventions and selling practice. Examination of the two contour maps (one for 1999 and one for 2005 presented in Figures 7 and 8 respectively) demonstrates unequivocally that there is significant spatial variation in the use of pathos across the Strathclyde conurbation. As expected, there has been an upward shift in the incidence of pathos right across the city over the intervening period. Although there are notable differences between the two years in the relative incidence of pathos in different areas (such as the area to the north east of Bearsden), comparison of the maps also suggests a degree of persistence over time in the spatial patterns. The area to the south west of Barrhead, for example, appears to have above average levels of pathos in the language of property sales both in 1999 and in 2005. Similarly for the area to the south east of East Kilbride, and for Bearsden and Bishopbriggs.

The persistence over time in the spatial variation of the incidence of pathos is illustrated further in the cross-sections of the two pathos surfaces plotted in Figure 9, drawn as the crow flies from Bearsden to Renfrew (i.e. Bearsden is located at zero metres on the horizontal axis). The variation across space in both periods is enormous, and although the two lines are certainly not parallel, there are a number of common peaks (at 2.4km, 4.2km and 5.5km from Bearsden) and troughs (at 1.8km, 4.0km, 4.8km, and 5.8km from Bearsden).

As a formal test of spatial persistence, we calculated $Ave\_Pathos_{9900,i}$ the average incidence of pathos in 1999/2000 for each post sector *i* where the number of observations, $n_i$, was greater than 30. We did the same for 2005/2006 and ran a simple regression of $Ave\_Pathos_{0506}$ on $Ave\_Pathos_{9900,i}$. If no spatial persistence existed then there would be no relationship between the two variables and the slope would be zero. The procedure was executed using both *All Pathos* and *Core Pathos*. In the event, our results conclusively rejected the null of zero slope coefficients, both for *All Pathos* ($b$ = 0.676; sig. = 0.000 using White's Standard Errors; $R^2$ = 0.343) and for *Core Pathos* ($b$ = 0.826; sig. = 0.000 using White's Standard Errors; $R^2$ = 0.4273).

Similarly, there is evidence of spatial variation in the *type* of pathos. In fact, the spatial variation is much more volatile across space, though the persistence over time in the spatial pattern is less obvious (see the maps in Figures 10 and 11 of Oates and Pryce 2007 which plot the contours for Type IV as % of pathos words in 1999/2000 and 2004/2005 respectively, and also from the cross-sections from Bearsden to Renfrew in Figure 12 of Oates and Pryce 2007). Once again there is a need to establish whether geographic patters in language would persist independently of the spatial variation in housing attributes. We attempt to do this by including submarket dummies and other spatial variables in the regression analysis below.

**[Figure 7 Spatial Variation of Pathos as % of No. Words in Each Property Description (1999) TO GO HERE]**

**[Figure 8 Spatial Variation of Pathos as % of No. Words in Each Property Description (2005) TO GO HERE]**

**[Figure 9 Cross-Section of the Pathos Surfaces from Bearsden to Renfrew TO GO HERE]**

Before we proceed to the multiple regressions, consider first the *bivariate* evidence for Hypotheses E and F. If we run a t-test of equal means between seasons (Hypothesis E) we find that we could reject the null of homogenous average pathos between winter and spring[3], with the incidence of pathos being slightly higher in spring, though the difference was not significant when we used the *Core Pathos* variable (one-tail sig. = 0.365; two-tail sig. = 0.730). Comparing spring and summer, we find that the average incidence of pathos is significantly higher in summer both for *All Pathos* (one-tail sig. = 0.017; two-tail sig. = 0.034; $n_{summer}$ = 14,257) and for *Core Pathos*. Use of pathos tends to fall in the autumn, which is true for *All Pathos,* and for *Core Pathos*. The incidence of pathos falls again in winter (one-tail sig. = 0.008; two-tail sig. = 0.016), though the reduction is not significantly different from zero when we use *Core Pathos* (one-tail sig. = 0.162; two-tail sig. = 0.323).

Finally, consider *Hypothesis F*, that *Excitement-inducing* superlatives will be more volatile than the other types of pathos terminology and more susceptible to particular market conditions. To investigate we calculated the average Type IV pathos incidence in each quarter of our data (30 quarters in total) and then did the same for the incidence of non-Type IV pathos. The crucial question was whether the variation in the quarterly average was greater for the incidence of Type IV pathos words than for the incidence of other types of pathos. This amounted to testing for the equality of variances of these incidences over time. We applied three tests: Levene's robust test statistic for the equality of variances plus Brown and Forsythe's two tests (the W_50 test and the W_10 test) that replace the mean in Levene's formula with the median and the 10 percent trimmed mean respectively. The results for these tests are reported in Table 2 along with the average incidence for each quarter and the coefficient of variation. The null of equal variances was rejected in all three tests at the 5%

[3] We define the seasons as follows: winter comprises December, January and February; spring comprises March, April and May; summer comprises June, July and August; autumn comprises September, October and November.

significance level (the Levene test rejected it at the 1% significance level). These results confirmed that the difference in the standard deviations of the incidence of Type IV (sd = .0059) and other pathos types (sd = .0034) over time was not due to sampling variation alone, but real difference in fact.

We also report the Coefficient of Variation which measures the standard deviations as a proportion of the mean (which allows us to compare the variation of variables measured in different units). For Type IV pathos, the Coefficient of Variation results reveal that the standard deviation was 16.55% of the mean; whereas for other types of pathos, the standard deviation over time was only 12.56% of the mean, which again confirms our hypothesis that there is greater variability in Type IV pathos.

**[Table 2 Variation in the Average Incidence of Type IV Pathos vs. Other Types of Pathos]**

***Results of Fractional Logit Multiple Regression Analysis***

How do we know whether the rise and fall in the incidence of pathos across space is caused by local conventions in language or by other factors? It is conceivable, for example, that variation in property type would be the main driver of spatial variation in the incidence of pathos since property types are both spatially fixed and spatially clustered. In other words, if we were able to hold property attributes constant, would we detect any significant degree of spatial clustering of pathos in the language of selling? A similar question could be asked with regard to the hypothesis that the incidence of pathos will vary over time due to changes in the buoyancy of the market. Although the make-up of the housing stock will have changed very little over the course of seven years, it is possible that the mix of property types that come onto

the market varies between phases of the business cycle and across space. So the question is would we be able to identify any significant variation in the incidence of pathos if we were able to hold property attributes and other factors constant?

Multiple regression analysis allows us to estimate the impact of the variables of interest while holding everything else constant. As discussed in the methods section, the fact that the dependent variable is bounded between zero and one makes Ordinary Least Squares inappropriate and so the regressions reported in Table 3 are computed using the Fraction Logit methodology. Regression [1] estimates the sensitivity of the incidence of pathos (all types) to a range of independent variables. Regressions [2], [3], [4] and [5] applies the Fractional Logit estimation to each of the four subcategories. Regression [6] uses as its dependent variable our narrower definition of pathos, *Core Pathos*.

Following the sequence set out in Figure 1, consider first Hypotheses A, that language is uniform across space and time, and Hypothesis B that the language of selling is not always homogenous but that the variation is not systematic. Both these are rejected by the regression analysis. We include a range of independent variables that capture systematic drivers across time and/or space (average postcode sector selling time for the quarter when a property comes onto the market, and average postcode sector pathos, deprivation, distance to Glasgow city centre, price bands, seasonal dummies, and submarket dummies). Length of description is included as a control variable. If there were no variation in the dependent variable (incidence of pathos), or if the variation were entirely white noise, all slope coefficients would be zero – the null hypothesis tested by the $Chi^2$ statistic. Given that the $Chi^2$ figure is very large for all regressions reported in Table 3 (sig. = 0.000 in each case) we can reject Hypotheses A and B (Theories 1 and 2).

We test Hypothesis C by including a range of dwelling attributes (number of bedrooms, flat, bungalow, detached, terraced, stone, stone flat, bay window, conservatory, garage, parking, garden). We also include price band dummies to capture unmeasured location and attribute effects. As noted earlier, we use band dummies rather than the price variable itself to help mitigate the endogeneity problem – while conceivably pathos could affect price it is only likely to do so at the margin and will not be sufficient to make a property shift from one price band to another, and certainly not cause it to shift two or more price bands. Most attribute coefficients are statistically different from zero (sig. < 0.05) in most of the regressions. Coefficients on the price band dummies are generally as anticipated – that is, they are greater, the higher the price band. In regression [1], for example, we find that properties in price band 2 have 13.3% higher odds of pathos than price band 1 (the reference category), while the odds are 20.9% higher for price band 3, 24.4% higher for price band 4 and 31.5% higher for price band 5 (all with sig. < 0.05). Interestingly, the effect is noticeably less pronounced for Type IV pathos where even price band 5 properties only have 16.0% higher odds of pathos than price band 1 properties. The opposite is true for Type I pathos the odds of which are 283.6% higher for price band 5 properties than for price band 1 properties, and also for *Core Pathos* where the odds are 161.5% higher for price band 5 properties than for price band 1 properties. These findings support Theory 4 – that use of hyperbole and emotive language is likely to vary between properties for sale because of real differences in the characteristics of dwellings – and verify the need to control for dwelling type when considering the subsequent hypotheses, which we do by retaining these variables in the regression.

In order to test Hypothesis D – the relationship between market buoyancy and pathos – we

estimate the effect on pathos of the average selling time in the postcode sector of the dwelling, in the quarter that the property comes onto the market. If the relationship between pathos and selling time is negative (proportional change in odds > 1), then pathos will be positively related to market buoyancy (pro-cyclical) and we can reject theory 7 (*Strategy to Market Difficult to Sell Properties*) in favour of theories 5 (*Cycles in Staff Composition*), 6 (*Irrational Exuberance*) and 8 (*Opportunity Cost of Viewing*). Though the effect is relatively small, we find that pathos falls as selling time rises. If time-on-the-market rises by one month, the odds of pathos are 97.9% of what they were before that rise in selling time (sig. = 0.000). The effect is slightly stronger when we use the narrow definition of pathos (regression [6]) – the odds of *Core Pathos* are only 95.6% of what they were before a rise of one month in selling duration. So Hypothesis D appears to be confirmed by the Fractional Logit Model for *All Pathos* and for *Core Pathos* words, and also appears to hold true for Type III and Type IV pathos (regressions [4] and [5]). For Types I and II, the effect is not significantly different from zero (sig. > 0.05).

Hypothesis F (Theory 6: *Irrational Exuberance*) suggests that Type IV pathos will be more sensitive to local market conditions than other types of pathos. Though the bivariate results appeared to support this theory, the Fractional Logit models do not provide strong evidence for it. In the Type IV regression, the coefficient for average time-on-the-market was not significantly greater than the coefficients estimated for the other pathos types. In fact, the largest effect is actually for Type III pathos (percentage change in odds in regression [4] = 94.8%, compared with percentage change in odds in regression [5] of 97.9%). So it seems that, when other factors are held constant, there is little evidence to support Hypothesis F and we must reject Theory 6 (or question our method of verifying it).

Out of our nine theories, the only one that predicted both pro-cyclical *and* seasonal variation in

the language of selling is theory 8: *Opportunity Cost of Viewing*. We have already discussed how the negative relationship with selling time would appear to verify pro-cyclicality. The bivariate analysis suggested that there was a seasonal effect, but this did not tell us whether the effect holds when we control for selling time and other factors. Looking at the fractional logit coefficients on the seasonal dummies, it seems that there is some evidence to support the notion of seasonality in the language of selling for the *All Pathos* variable, but not for the other dependent variables. Compared with autumn/winter, the odds of pathos is 1.7% higher during the spring (sig. = 0.028), and 1.5% higher during the summer (sig. = 0.057), *cet par*.

Finally, we turn to the question of spatial variation in the language of selling due to factors other than dwelling type (Hypothesis G, Theory 9). We attempt to capture the impact of spatial variation due to local conventions by including the average incidence of pathos in the locality (the second independent variable in Table 3). If Hypothesis G is valid, we would expect the odds of pathos to be positively correlated with the average pathos in the locality even when selling times, property types and seasonal variations are held constant. Conversely, if there is no spatial effect, the correlation will be negative or non-existent. In the event, we found a strong spatial effect, particularly for *All Pathos* (regression [1]) and for *Core Pathos* words (regression [6]) where the odds of pathos being used in a particular property description rose by 12.4% and 19.6% respectively for every 1% rise in the average level of pathos in the area.

We included three further sets of spatial variables to capture spatial patters: deprivation index, distance to the centre of Glasgow, and a number of submarket dummies. The deprivation index appeared to have an ambiguous effect (positive for some measures of pathos and negative for others) and the magnitude of the effect was negligible. The same is true for distance from the

city centre. A number of submarket dummies were, however, significant. *All Pathos* tended to be slightly lower in the West End (97.6%, sig. = 0.014), for example, than in most other areas, and similarly for *Core Pathos* (96.2%, sig. = 0.063), Type I pathos, (86.8%, sig. = 0.007), and for Type II and Type IV pathos. Other things being equal, pathos tended to be lower in East Dunbartonshire and East Renfrewshire, but higher in the East End. Taken together, these results suggest that there are marked and persistent spatial effects in the pattern of pathos, but that there is no simple explanation in terms of deprivation or distance from the centre of the city. Such a finding seems consistent with idiosyncratic local conventions in language, as predicted in Theory 9.

**["Table 3 Fractional Logit Regressions for the Incidence of Pathos"<TO GO HERE>]**

Note that these results were generally insensitive to changes to the model specification – the implications remained the same when we altered the list of independent variables (e.g. included year dummies, dropped seasonal dummies, dropped submarket dummies etc.).

# Conclusion

In this article we have sought to choose between theories of language variation by constructing a series of hypotheses that exploit, where possible, the incompatibilities between theories. We then tested these hypotheses using house transactions data from Strathclyde, Scotland. Our research has uncovered strong evidence that the verbal construction of house adverts varies systematically, both spatially and temporally. We found that this variation was partly due to changes in the mix of properties being sold, which led us to conclude that there was a need to

control for dwelling attributes when testing subsequent hypotheses. To this end, we employed fractional logit regression methods to help us investigate our hypotheses in a multiple-causation estimation framework.

Controlling for property type we sought to establish, in particular, whether the deployment of euphemistic dialect was pro- (rising as the market rises) or counter- (falling as the market rises) cyclical. On this point, a number of our theories were in conflict. The theory that agents utilise more effusive descriptions when dealing with difficult to sell properties suggested that the incidence of *pathos* (emotive language) will be counter-cyclical, whereas theories based on cycles in *staff composition*, *irrational exuberance* and changes in the *opportunity cost of viewing*, all predicted pro-cyclicality. Though the effect proved to be relatively small, we found that the incidence of pathos tended to vary with market buoyancy (pro-cyclical), even when holding everything else constant (including property type). We rejected the counter-cyclical theory on this basis but, in actual fact, the small net effect may be evidence that opposing forces are at work, with one force dominating on-balance. Of the pro-cyclical theories, we rejected the irrational exuberance explanation on the basis that excitement-inducing pathos words did not seem to be any more pro-cyclical than the others, though of course this may have reflected deficiencies in our method of pathos categorisation.

We then sought to test the theory that changes in language occurred due to shifts in the opportunity cost of viewing over the course of the year. Examining the evidence for seasonal variation we found that the incidence of pathos was slightly (but significantly) higher during the spring and summer for our broad definition of pathos, but not for the narrow definition. Our results are somewhat ambiguous, therefore, with regard to opportunity cost of viewing. Cycles in staff composition would plausibly offer a complementary explanation for the cyclical

variation, though data on employee characteristics should really be considered before embracing this theory.

Our final theory was one of spatial variation in the parlance of property peddlers. We hypothesised that local conventions might emerge that lead to persistent differences in the way dwellings were marketed in different submarkets. We found strong evidence that the use of pathos in property adverts varies across geographical space and this finding appeared to be independent of property type, area deprivation, distance from the city centre and market buoyancy. We could not, therefore, reject the theory that there exist local conventions in estate agent dialect.

It is possible that these variations over time and space in the rhetoric of selling have the potential to hinder the attempts of house-buyers to decipher the euphemism of estate agent advertisements, particularly if they are moving between areas. This has implications for information dissemination and the efficiency of local housing markets, and invites further research into these ramifications. It may be that buyers are adept at adjusting to these changes and there is no material consequence as a result. Perhaps the most important implication of our findings, therefore, is that fluctuations and patterns in the language of selling appear to reveal aspects of market structures and dynamics and, in that sense, may hold out the prospect of offering additional insights into the machinations of the market.

More generally, our investigation has emphasised the importance of considering the emotional issues associated with the real estate process, and the intrinsic link between the psychology of the house purchase decision and the dynamics of the market itself. We have attempted to link the powerful emotional and economic aspects of the buying process, and to achieve this a

mixture of qualitative and quantitative analysis. While studies of the housing market have brought many answers to the puzzles of real estate markets, the acknowledged visceral side is somewhat harder to quantify and remains relatively unexplored for economists. Hopefully, we have made some headway in establishing a methodological foundation that will encourage future work and facilitate a more rounded understanding of housing markets.[4]

## References


Aristotle (350 B.C.) *Rhetoric*, http://www.greektexts.com.

BBC News Online (2002) "Estate Agent speak: A dictionary" 11 October, 2002. http://news.bbc.co.uk

Budge, I., Robertson, D. and Hearl D. (eds). (1987) *Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies*. Cambridge: Cambridge University Press.

de Boton, A. (2004) *Status Anxiety*, (London: Penguin/Hamish Hamilton).

Haurin, D. (1988) The Duration of Marketing Time of Residential Housing. *AREUEA Journal* 16(4) pp. 396–410.

Hendershott, P. and Pryce, G. (2006) The Sensitivity of Homeowner Leverage to the Deductibility of Home Mortgage Interest, *Journal of Urban Economics* 60 (1) pp. 50-68.

Houseweb (2006) "Estate Agent Speak" http://www.houseweb.co.uk/.html

Levin, E.J. and Pryce, G. (2007) 'A Statistical Explanation for Extreme Bids in the House Market', *Urban Studies*, 44:12, pp. 2339-2355.

Levitt, S. and Syverson, C. (2005) *Market Distortions when Agents are Better Informed: The Value of Information in Real Estate*, NBER Working Paper 11053.

Lorenzo-Dus, N. (2006) "Buying and selling: mediating persuasion in British property shows" *Media Culture & Society* 28 (5) pp. 740-761.

Merlo, A. and Ortalo-Magné, F. (2004) "Bargaining over residential real estate: evidence from England", *Journal of Urban Economics*, vol. 56, pp. 192-216.

Munneke H. and Yavas, A. (2001) Incentives and performance in real estate brokerage, *Journal of Real Estate Finance and Economics* 22 (1) pp. 5-21.

Oates, S. (2006) *Television, Elections and Democracy in Russia*. London: Routledge.


[4] See Oates and Pryce 2007 for suggested avenues of future research.

Oates, S. and Pryce, G. (2007) *Pathos and Patter in Real Estate Parlance*, SHEFRN Discussion Paper, available from www.gpryce.com and www.shefrn.org.

Papke, L. and Wooldridge, J. (1996) Econometric methods for fractional response variables with an application to 401(k) plan participation rates, *Journal of Applied Econometrics* pp. 11619–632.

Pardo, M. (2001) 'Linguistic Persuasion as an Essential Political Factor in Current Democracies: Critical Analysis of the Globalization Discourse in Argentina at the Turn and at the End of the Century', *Discourse and Society* 12(1): pp. 91–118.

Propertypanther.com (2006) "Understanding Estate Agents Speak", Parts 1 & 2, www.propertypanther.com

Pryce, G. and Gibb, K. (2006) Submarket Dynamics of Time to Sale, *Real Estate Economics*, 34(3), pp. 377–415.

Rothenberg J., Galster, G., Butler, R., and Pitkin, J. (1991) *The Maze of Urban Housing Markets: Theory, Evidence, and Policy*, Chicago: University of Chicago Press.

Smith, S., Munro, M., and Christie, H. (2006) "Performing (housing) markets", *Urban Studies*, vol. 43, pp. 81-98.

Stein, J. (1995) Prices and Trading Volume in the Housing Market: A Model with Downpayment Constraints, *Quarterly Journal of Economics* 110(2) pp. 379–406.

Stuart-Smith, J., Timmins, C. and Pryce, G. (2005) '*Explaining phonological variation and change in an urban vernacular: does television play a role?*', Conference Paper, 34th Conference of New ways of analyzing variation, New York, July 2005.

Sweet, F. (1999) *Space: Reshaping Your Home for the Way You Want to Live*. (London: Conran Octopus Ltd.).

Van Dijk, T. (1998) *Ideology: A Multidisciplinary Approach* (London: Sage).

## *Acknowledgements*

Grateful thanks to 5 anonymous reviewers for their enthusiasm for the paper and their excellent suggestions and multitude of ideas. Thanks also to participants at the ENHR conference in Edinburgh (February 2007) and the National Association of Estate Agents conference (March 2007) for their helpful comments. All remaining errors are our own.

# Figure 1 Decision Tree of Hypothesis Tests

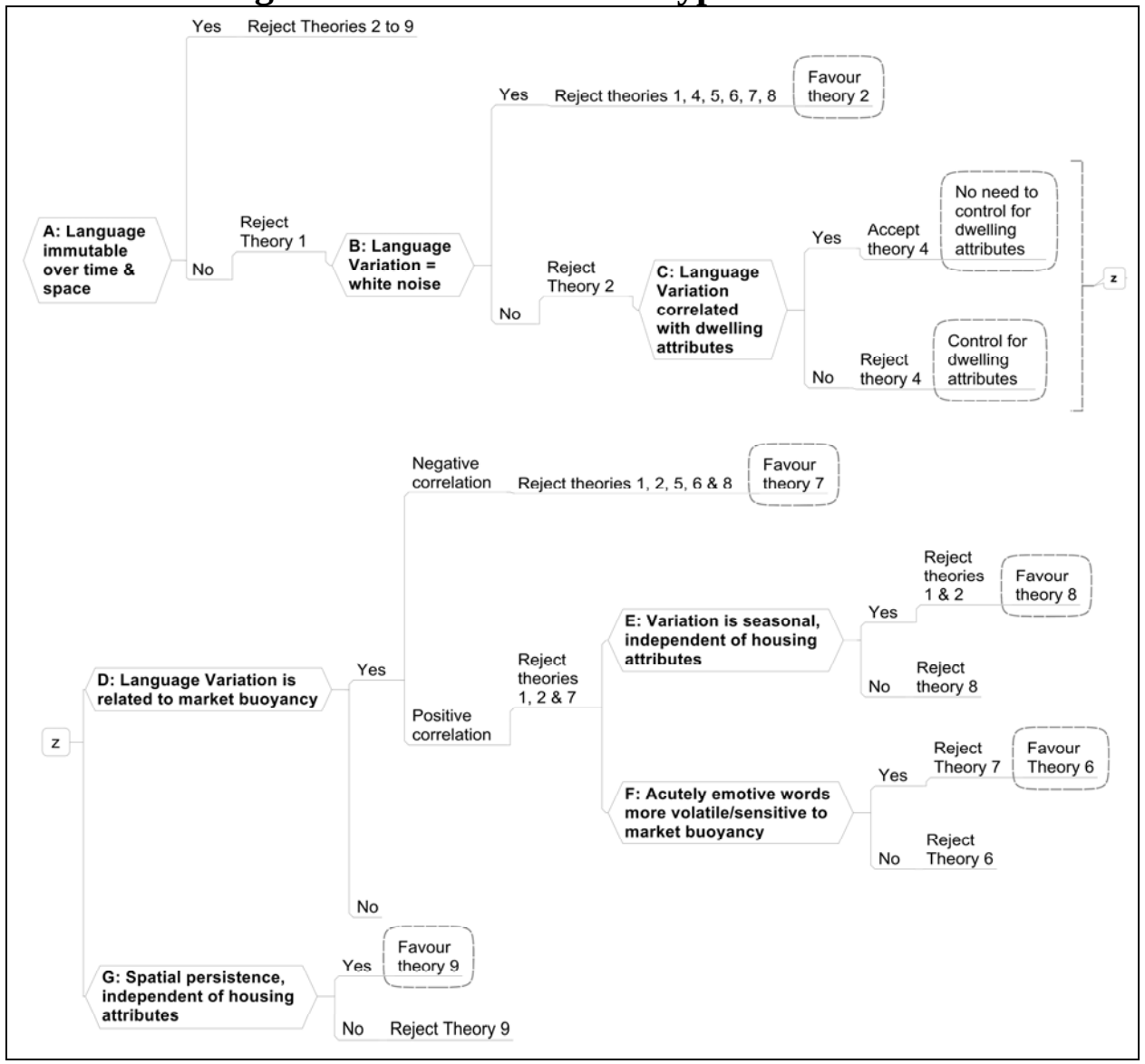

**Figure 2 Average Marketing Time Since 1999**

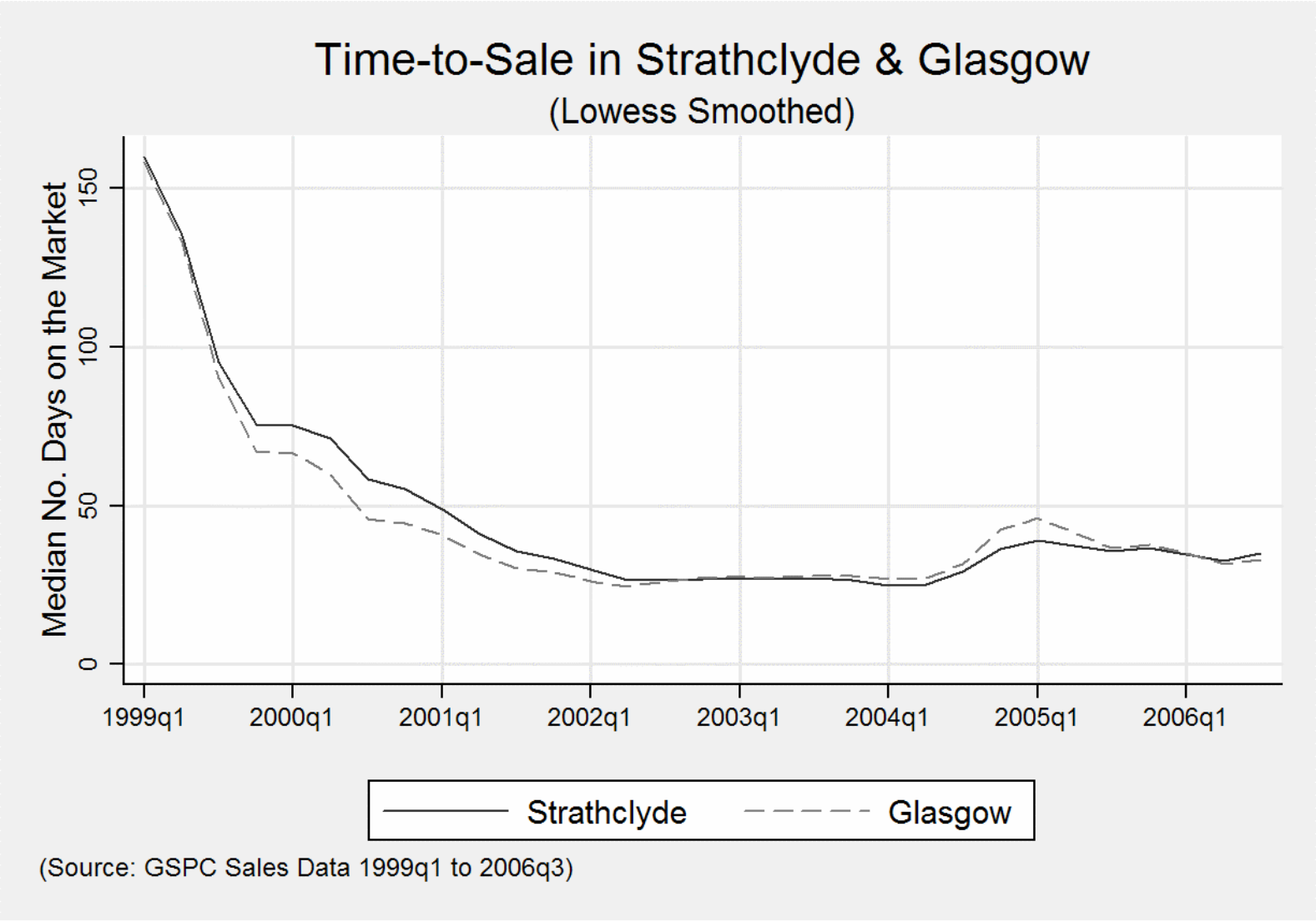

**Figure 3 House Prices and the Incidence of Pathos in the Language of Selling**

**Figure 4 Time on the Market and the Incidence of Pathos in New Descriptions**

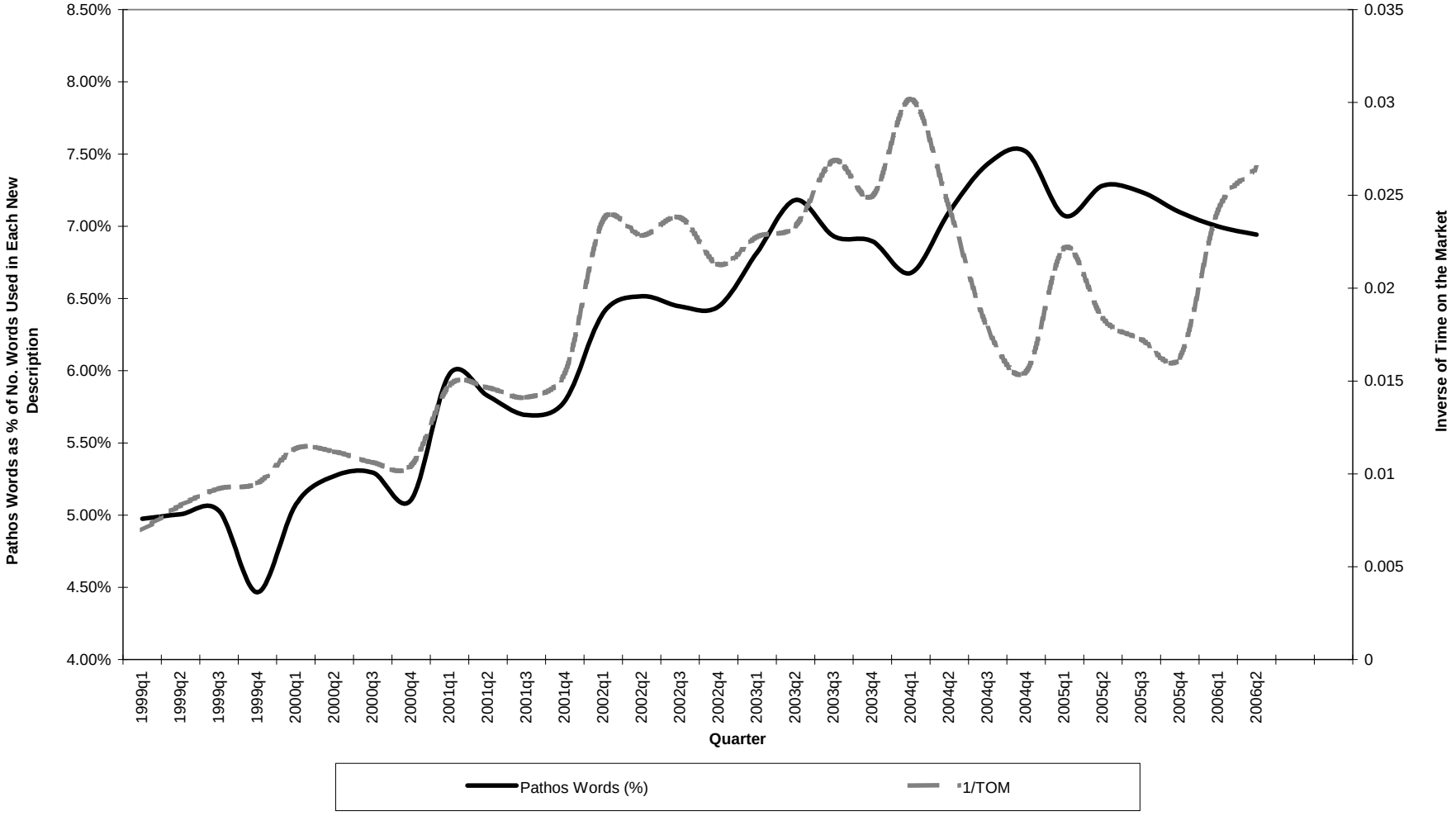

**Figure 5 Number of GSPC Sales and the Incidence of Pathos in New Descriptions**

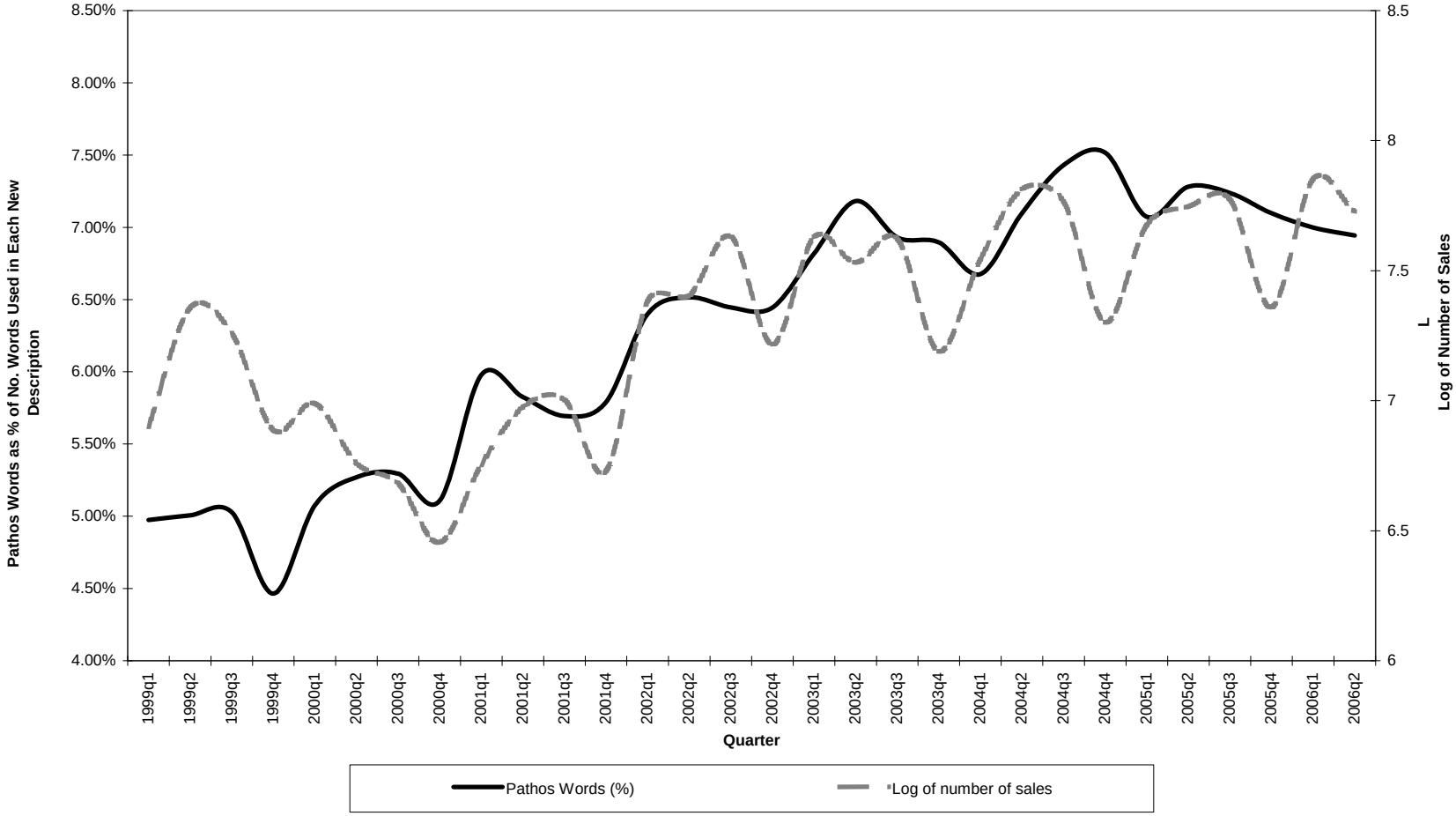

**Figure 6 Variation in the Type of Pathos Over Time**

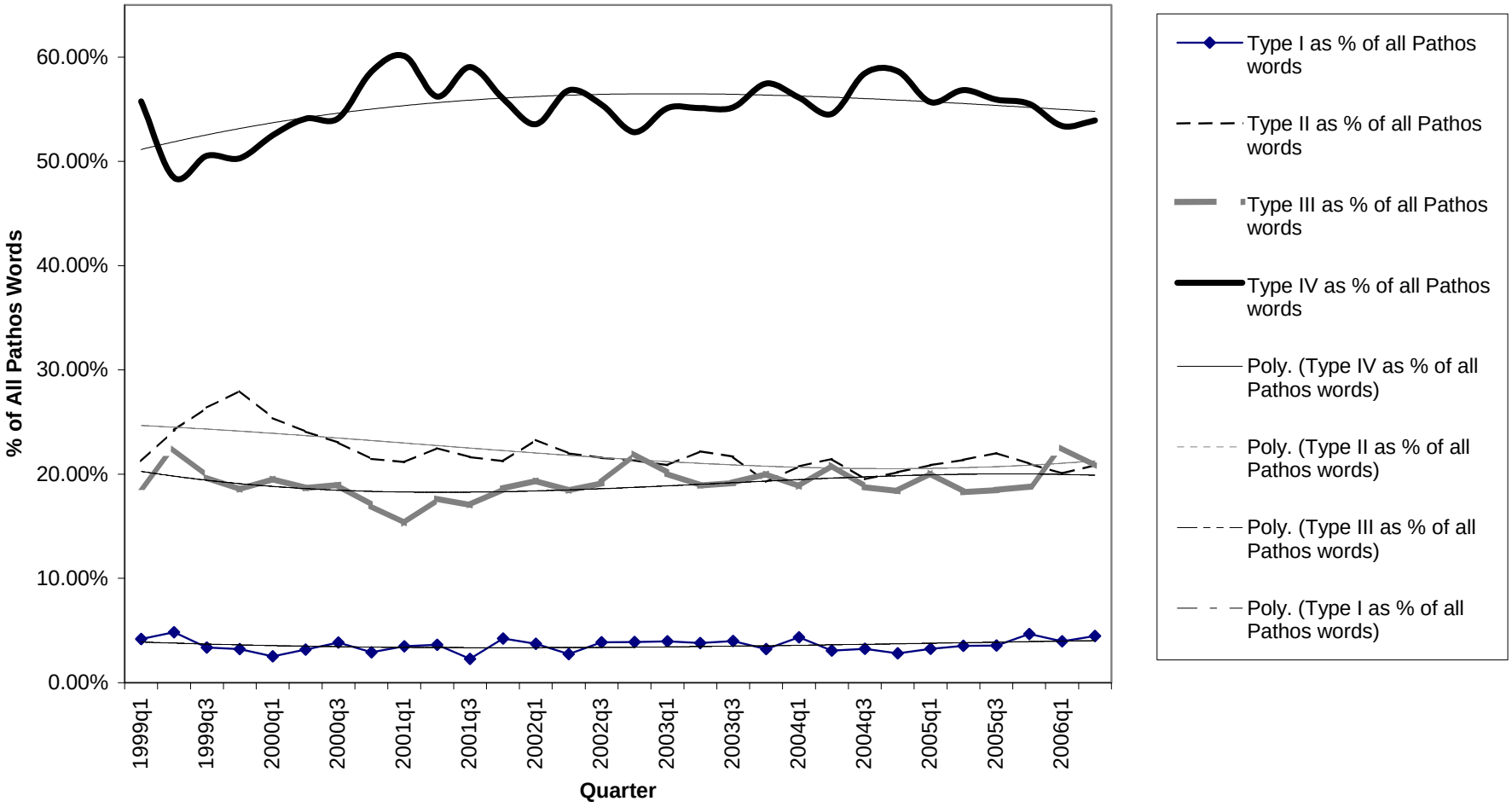

**Figure 7 Spatial Variation of Pathos as % of No. Words in Each Property Description (1999)**


Duntocher
Bearsden
Kilmalcolm
Clydebank
Bishopbriggs
Stepps
Renfrew
Bridge of Weir
Glasgow
Paisley
Johnstone
Rutherglen
Cambuslang
Lochwinnoch
Barrhead
Neilston
Newton Mearns
Uplawmoor
Beith
East Kilbride
Eaglesham
Dunlop

**Figure 8 Spatial Variation of Pathos as % of No. Words in Each Property Description (2005)**

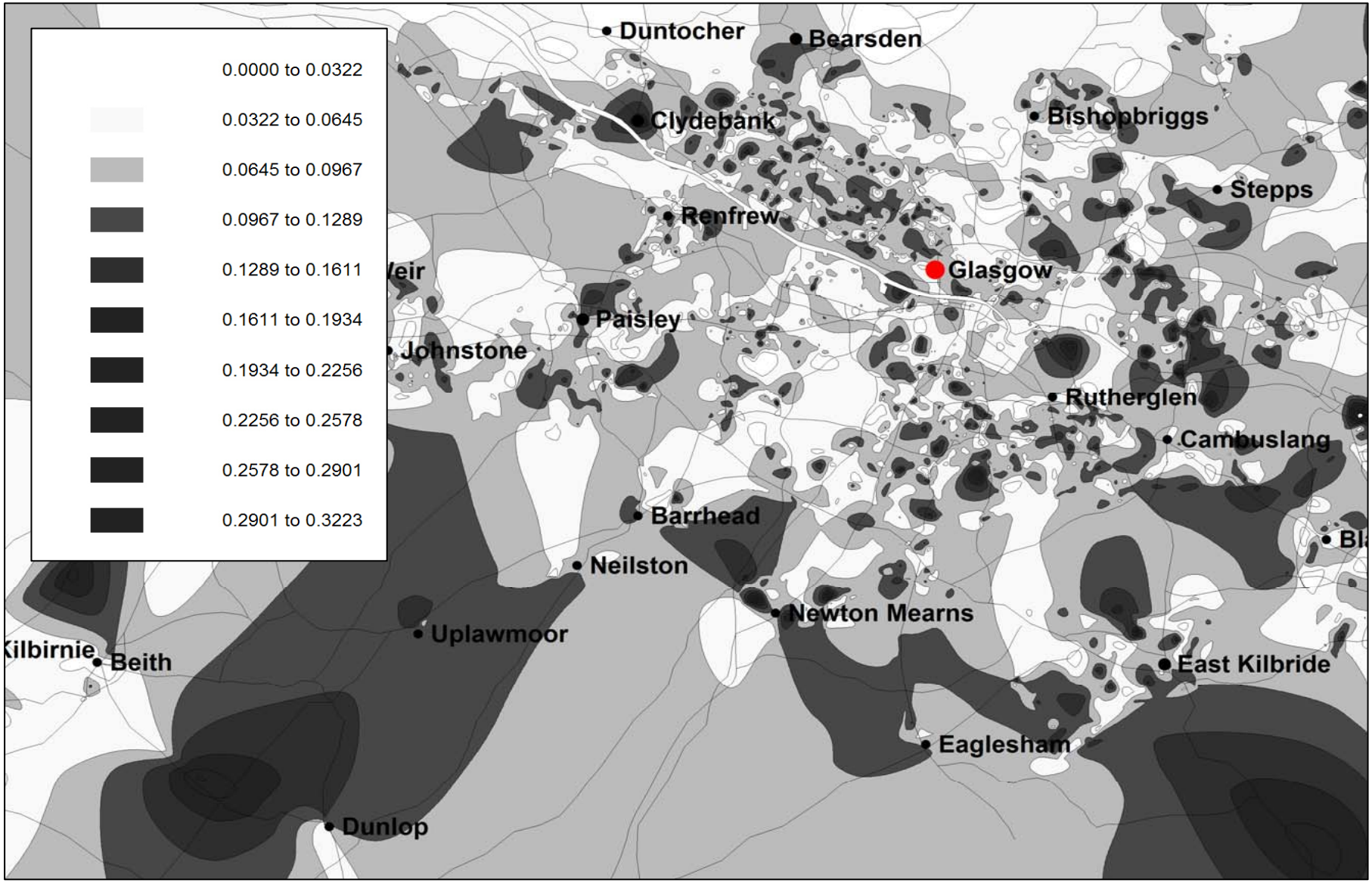

**Figure 9 Cross-Section of the Pathos Surfaces from Bearsden to Renfrew**

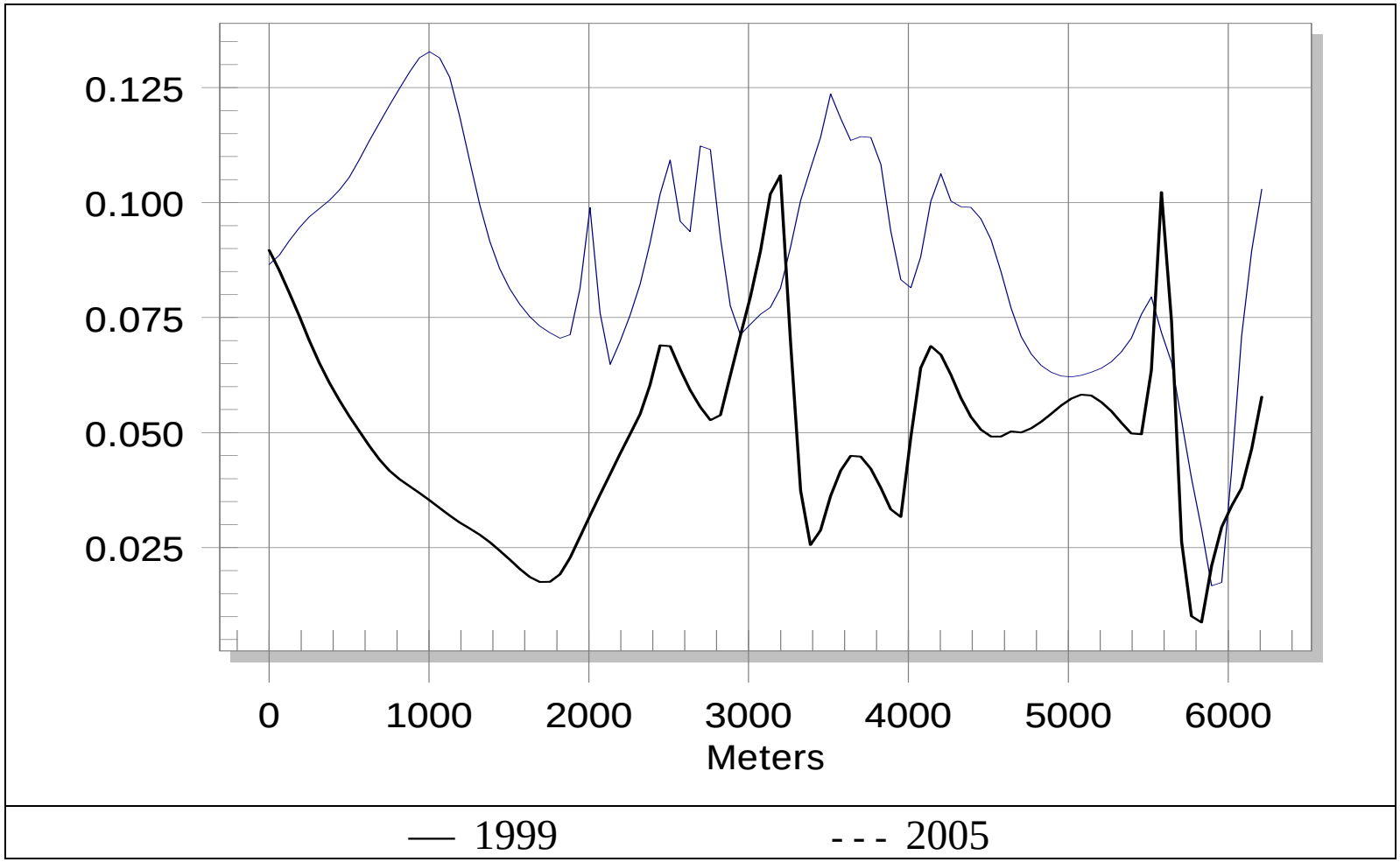

**Table 1 Descriptive Statistics**

| **Continuous Variables and Proportions:** | | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **Description** | **mean** | **median** | **min** | **max** | **n** |
| Pathos_n | Number of pathos words in each description | 2.08 | 2.00 | 0.00 | 11.00 | 49,926 |
| Pathos_Core_n | Number of Core pathos words in each description | 0.41 | 0.00 | 0.00 | 6.00 | 49,926 |
| Type_I_n | Number of Type I pathos words in each description | 0.08 | 0.00 | 0.00 | 3.00 | 49,926 |
| Type_II_n | Number of Type II pathos words in each description | 0.44 | 0.00 | 0.00 | 5.00 | 49,926 |
| Type_III_n | Number of Type III pathos words in each description | 0.39 | 0.00 | 0.00 | 4.00 | 49,926 |
| Type_IV_n | Number of Type IV pathos words in each description | 1.18 | 1.00 | 0.00 | 9.00 | 49,926 |
| Pathos_p | Proportion of all words in each description that are pathos words | 0.06 | 0.06 | 0.00 | 0.50 | 49,926 |
| Pathos_Core_p | Proportion of all words in each description that are Core pathos words | 0.01 | 0.00 | 0.00 | 0.19 | 49,926 |
| Type_I_p_P | Proportion of pathos words in each description that are Type I | 0.04 | 0.00 | 0.00 | 1.00 | 42,778 |
| Type_II_p_P | Proportion of pathos words in each description that are Type II | 0.22 | 0.00 | 0.00 | 1.00 | 42,778 |
| Type_III_p_P | Proportion of pathos words in each description that are Type III | 0.19 | 0.00 | 0.00 | 1.00 | 42,778 |
| Type_IV_p_P | Proportion of pathos words in each description that are Type IV | 0.55 | 0.50 | 0.00 | 2.00 | 42,778 |
| dscrptn_wordcount | Word count for each description | 31.61 | 32.00 | 1.00 | 51.00 | 49,926 |
| dscrptn_charcount | Character count for each description | 196.66 | 205.00 | 2.00 | 244.00 | 49,926 |
| deprivtn | Deprivation score | 5.78 | 4.48 | 2.03 | 16.24 | 49,926 |
| cbd_glas_km | Distance to City Centre (km) | 12.73 | 8.06 | 0.32 | 519.64 | 49,926 |
| allrooms | Number of rooms (bedrooms + public rooms) | 3.53 | 3.00 | 0.00 | 24.00 | 49,926 |
| **Binary Variables:** | | | | | | |
| **Variable** | **Description** | | **% of cases that = 1** | | | **n** |
| noPathos | = 1 if no pathos words the property description; = 0 otherwise | | 14.3% | | | 49,926 |
| noCOREPathos | = 1 if no Core pathos words in the property description; = 0 otherwise | | 67.2% | | | 49,926 |
| flat_all | = 1 if the property is a flat; = 0 otherwise | | 48.7% | | | 49,926 |
| bung_ALL | = 1 if the property is a bungalow; = 0 otherwise | | 6.2% | | | 49,926 |
| detached | = 1 if the property is detached; = 0 otherwise | | 10.4% | | | 49,926 |
| terraced | = 1 if the property is terraced; = 0 otherwise | | 8.3% | | | 49,926 |
| stone | = 1 if the property is constructed of stone; = 0 otherwise | | 12.2% | | | 49,926 |
| stone_flat | = 1 if the property is a flat constructed of stone; = 0 otherwise | | 9.8% | | | 49,926 |
| bay | = 1 if the property has a bay window; = 0 otherwise | | 16.0% | | | 49,926 |
| conservy | = 1 if the property has a conservatory; = 0 otherwise | | 2.7% | | | 49,926 |
| garage_d | = 1 if the property has a garage; = 0 otherwise | | 28.8% | | | 49,926 |
| parking | = 1 if the property has parking facilities; = 0 otherwise | | 12.1% | | | 49,926 |
| garden_d | = 1 if the property has a garden; = 0 otherwise | | 70.3% | | | 49,926 |

**Table 2 Variation in the Average Incidence of Type IV Pathos vs Other Types of Pathos**

| Quarter | Pathos Type IV | All other Pathos Types |
|---|---|---|
| 1999q1 | 0.0277 | 0.0220 |
| 1999q2 | 0.0251 | 0.0250 |
| 1999q3 | 0.0260 | 0.0243 |
| 1999q4 | 0.0230 | 0.0216 |
| 2000q1 | 0.0274 | 0.0234 |
| 2000q2 | 0.0290 | 0.0237 |
| 2000q3 | 0.0287 | 0.0242 |
| 2000q4 | 0.0300 | 0.0211 |
| 2001q1 | 0.0363 | 0.0235 |
| 2001q2 | 0.0332 | 0.0250 |
| 2001q3 | 0.0337 | 0.0233 |
| 2001q4 | 0.0342 | 0.0237 |
| 2002q1 | 0.0354 | 0.0286 |
| 2002q2 | 0.0374 | 0.0278 |
| 2002q3 | 0.0360 | 0.0285 |
| 2002q4 | 0.0351 | 0.0294 |
| 2003q1 | 0.0384 | 0.0298 |
| 2003q2 | 0.0407 | 0.0312 |
| 2003q3 | 0.0393 | 0.0300 |
| 2003q4 | 0.0400 | 0.0290 |
| 2004q1 | 0.0384 | 0.0284 |
| 2004q2 | 0.0398 | 0.0312 |
| 2004q3 | 0.0439 | 0.0304 |
| 2004q4 | 0.0457 | 0.0295 |
| 2005q1 | 0.0402 | 0.0305 |
| 2005q2 | 0.0421 | 0.0307 |
| 2005q3 | 0.0417 | 0.0307 |
| 2005q4 | 0.0402 | 0.0308 |
| 2006q1 | 0.0392 | 0.0308 |
| 2006q2 | 0.0389 | 0.0305 |
| 2006q3 | 0.0381 | 0.0318 |
| **Summary Statistics:** | | |
| Mean of all quarterly means: | 0.0356 | 0.0274 |
| Standard deviation of means: | 0.0059 | 0.0034 |
| Coefficient of Variation for the quarterly means: | 16.55% | 12.56% |
| **Equality of Variances Test:** | | |
| Levenes Test | F=7.151 df(1, 60) sig. = .0096 | |
| Brown & Forsyth W50 Test | F=4.973 df(1, 60) sig. = .0295 | |
| Brown & Forsyth W10 Test | F= 6.620 df(1, 60) sig. = .0126 | |

**Table 3 Fractional Logit Regressions for the Incidence of Pathos**

| | *Dependent Variable* [‡] | | | | | |
|---|---|---|---|---|---|---|
| | Pathos (all) | Type I | Type II | Type III | Type IV | Core |
| *Independent Variables:* | [1] | [2] | [3] | [4] | [5] | [6] |
| Average selling time | 0.979 [§] | 0.987 | 1.002 | 0.948 | 0.979 | 0.956 |
| | (0.000) [†] | (0.169) | (0.567) | (0.000) | (0.000) | (0.000) |
| Average Pathos in the area | 1.124 | 1.051 | 1.116 | 1.134 | 1.117 | 1.196 |
| | (0.000) | (0.004) | (0.000) | (0.000) | (0.000) | (0.000) |
| deprivtn | 1.004 | 0.979 | 1.006 | 0.989 | 1.009 | 1.001 |
| | (0.016) | (0.007) | (0.045) | (0.000) | (0.000) | (0.878) |
| cbd_glas_km | 0.998 | 1.003 | 1.001 | 0.996 | 0.997 | 1.002 |
| | (0.000) | (0.001) | (0.030) | (0.000) | (0.000) | (0.000) |
| dscrptn_charcount | 1.006 | 1.010 | 1.003 | 1.003 | 1.007 | 1.009 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| allrooms | 0.974 | 1.072 | 0.942 | 0.967 | 0.983 | 1.023 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.002) |
| flat_all | 0.988 | 1.116 | 0.960 | 0.777 | 1.080 | 1.094 |
| | (0.203) | (0.060) | (0.040) | (0.000) | (0.000) | (0.000) |
| bung_ALL | 1.049 | 1.661 | 1.097 | 1.142 | 0.938 | 1.009 |
| | (0.001) | (0.000) | (0.002) | (0.000) | (0.002) | (0.747) |
| detached | 0.922 | 0.763 | 0.910 | 0.869 | 0.979 | 0.892 |
| | (0.000) | (0.000) | (0.001) | (0.000) | (0.245) | (0.000) |
| terraced | 0.958 | 0.833 | 1.001 | 0.827 | 1.009 | 0.871 |
| | (0.001) | (0.014) | (0.959) | (0.000) | (0.623) | (0.000) |
| stone | 0.851 | 1.453 | 0.904 | 0.925 | 0.750 | 0.917 |
| | (0.000) | (0.000) | (0.027) | (0.056) | (0.000) | (0.044) |
| stone_flat | 0.876 | 0.576 | 0.767 | 0.672 | 1.074 | 0.922 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.049) | (0.106) |
| bay | 0.931 | 0.965 | 0.985 | 0.941 | 0.911 | 0.955 |
| | (0.000) | (0.478) | (0.437) | (0.003) | (0.000) | (0.024) |
| conservy | 1.065 | 1.186 | 1.188 | 0.918 | 1.055 | 1.253 |
| | (0.001) | (0.025) | (0.000) | (0.030) | (0.046) | (0.000) |
| garage_d | 0.988 | 0.867 | 1.052 | 1.063 | 0.948 | 1.065 |
| | (0.170) | (0.001) | (0.005) | (0.000) | (0.000) | (0.001) |
| parking | 0.901 | 0.882 | 1.041 | 0.887 | 0.862 | 0.973 |
| | (0.000) | (0.035) | (0.052) | (0.000) | (0.000) | (0.239) |
| garden_d | 0.968 | 1.069 | 0.869 | 1.207 | 0.947 | 0.863 |
| | (0.000) | (0.215) | (0.000) | (0.000) | (0.000) | (0.000) |
| Price band 2 | 1.133 | 1.417 | 1.203 | 1.208 | 1.074 | 1.312 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Price band 3 | 1.209 | 1.771 | 1.370 | 1.241 | 1.118 | 1.641 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Price band 4 | 1.245 | 2.314 | 1.439 | 1.302 | 1.117 | 1.971 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Price band 5 | 1.315 | 3.836 | 1.460 | 1.326 | 1.160 | 2.615 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Spring | 1.017 | 1.007 | 1.010 | 1.029 | 1.015 | 0.987 |
| | (0.028) | (0.857) | (0.546) | (0.072) | (0.159) | (0.458) |

| | *Dependent Variable* ‡ | | | | | |
|---|---|---|---|---|---|---|
| | Pathos (all) | Type I | Type II | Type III | Type IV | Core |
| *Independent Variables:* | [1] | [2] | [3] | [4] | [5] | [6] |
| Summer | 1.015 | 0.983 | 1.025 | 1.011 | 1.013 | 1.009 |
| | (0.057) | (0.677) | (0.124) | (0.506) | (0.227) | (0.569) |
| West End | 0.976 | 0.868 | 0.978 | 1.046 | 0.964 | 0.962 |
| | (0.014) | (0.007) | (0.288) | (0.032) | (0.006) | (0.063) |
| East End | 1.073 | 1.143 | 1.026 | 1.155 | 1.051 | 1.185 |
| | (0.000) | (0.055) | (0.352) | (0.000) | (0.003) | (0.000) |
| East Dunbartonshire | 0.922 | 0.740 | 1.009 | 1.036 | 0.863 | 0.922 |
| | (0.000) | (0.000) | (0.742) | (0.174) | (0.000) | (0.005) |
| East Renfrewshire | 0.959 | 0.716 | 0.952 | 1.028 | 0.957 | 0.841 |
| | (0.024) | (0.001) | (0.219) | (0.422) | (0.100) | (0.000) |
| Intercept | 0.011 | 0.000 | 0.004 | 0.003 | 0.005 | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| *n* | 49,926 | 49,926 | 49,926 | 49,926 | 49,926 | 49,926 |
| Log likelihood | -8,960 | -680 | -2,960 | -2,621 | -6,079 | -2,685 |
| $Chi^2$ | 6,758 | 1,713 | 1,289 | 3,352 | 3,647 | 5,991 |

‡ The dependent variable measures the number of pathos words as a proportion of all words (or all pathos words) used in the marketing description.

§ Coefficients are in exponential form to measure the proportionate change in odds of pathos due to a unit increase in the explanatory variable, holding all other variables constant.

† Figures in brackets are significance levels calculated using Papke and Wooldridge (1996) robust standard errors.