

Pathos and Patter in Real Estate Parlance

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

Gwilym Pryce
Department of Urban Studies,
University of Glasgow,
g@gpryce.com

Discussion Paper

23rd October 2007

Please do not quote without permission

Abstract

This paper presents the first systematic analysis of estate agent language and employs Aristotle's ponderings on the art of persuasion as a means of classifying the peculiar parlance of property peddlers. "Des. Res.", "rarely available", "viewing essential" – these are all part of the peculiar parlance of housing advertisements. The question is whether the selling agent's penchant for rhetoric is uniform across a single urban system or whether there are variations, even within a relatively limited geographical area. We are also interested in how the use of superlatives varies over the market cycle. For example, are estate agents more inclined to use hyperbole when the market is buoyant or when it is flat? 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 West of Scotland over the period 1999 to 2006. Our analysis has implications for our understanding of the agency behaviour of housing market professionals and endeavours to open up a new avenue of research into the market-impact of rhetoric in the language of selling.

Key words: real estate brokerage, textual analysis, marketing, estate agent, Aristotle.

Non-Technical Summary

Despite ubiquitous humour about estate agent hyperbole, and not infrequent references to it in the press (see, for example, the extensive press coverage of estate agent Julian Bending who has become famous for his frank, and often shocking, property descriptions), there has been no systematic study of realtor language patterns. Curiously, much of the academic research on housing markets (particularly that carried out by economists) assumes that estate agents are impartial disseminators of information, which of course, lies in stark contrast to their popular characterisation.

Although estate agents are notorious for using euphemisms and hyperbole, we question whether this really matters. So long as estate agents are reasonably consistent in their use of language, house buyers can quickly learn how to translate the peculiar parlance of property peddlers into everyday English. So the crucial question is whether this is actually the case -- whether estate agent rhetoric follows similar patterns across different parts of a city and whether there is much variation over time. If such variation does occur, then it is more difficult for house-hunters to make sense of property advertisements.

And we did indeed find evidence of volatility in the use of language, both over time and across space. Estate agents in certain areas and at certain times were much more likely than those in other parts of the city to use "*pathos*" – the term used by Aristotle to describe the use of emotive language when attempting to persuade.

In some ways these variations over time and across areas in the pattern of estate agent parlance is all the more surprising when you think that the internet now plays a very prominent role in the marketing of properties. Rather like the argument that television would dilute regional accents, one might expect the internet to have had a similar effect on the marketing of properties: that it would result in greater uniformity in the language of property adverts. This has not been the case -- in fact, over the period of our study (1999 to 2005) estate agents appear to have become more likely to use emotive language and the variation across the city has not diminished.

1. Introduction

Analysis of the transactions process has traditionally focussed on buyer search behaviour (Yavas 1992), 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 language used in property advertisements has itself been considered worthy of research (Levitt and Syverson 2005), and even then, its assumed sphere of influence has not extended beyond the peripheral.

Yet the notion of estate agents as impartial information disseminators contrasts strongly with the common perception of them by the media and the general public. Indeed, the language that estate agents employ 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 set of assumptions about their ‘flexible’ use of language, as the following extracts from humorous “dictionaries” of estate agent euphemism demonstrate:

<i>‘Benefits From:</i>	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)
<i>‘Bijou:</i>	Would suit contortionist with growth hormone deficiency.’	(Ibid)
<i>‘Compact:</i>	See Bijou, then divide by two.’	(Ibid)
<i>‘Convenient For:</i>	A deceptive term with two possible definitions depending on the object of the phrase: Eg "Convenient For A40" means your garden doubles as the hard shoulder. Whereas "Convenient For local amenities" means you can run to the shops. If you are Paula Radcliffe.’	(Ibid)
<i>‘In Need of Modernisation:</i>	In need of demolition.’	(Ibid)
<i>‘Internal Viewing Recommended:</i>	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)

This alleged abuse of language has fostered a general cynicism towards the profession and fuelled an accompanying brand of humour centred round the notion of the estate agent as outright charlatan:

‘Question:	How can you tell when an estate agent is lying?	
Answer:	His lips move.’	(Booth, 2006)
‘Question:	Why won't a shark bite an estate agent?	
Answer:	Professional courtesy!’	(ibid)

It is beyond the scope of this study to verify the extent to which these prejudices about estate agents are justified. That noble endeavour 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. But should we really 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 entail a tongue-in-cheek articulation of the unspoken acknowledgement that the realtor cannot help but converse in the language of 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 idiom 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 her gushing property descriptions only serves to liberate potential buyers to indulge in the fantasy of lifestyle-real-estate. If the language 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 language 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 the language of selling, 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. Our first step is to identify changes in the use of language, and this requires us to find a way of categorizing the use of language. Our methodological framework uses, as its starting point, Aristotle's theory of rhetoric which divides the act of persuasion into three categories: 1) *ethos*, which is appeal based on the character of the speaker; 2) *logos* (appeal based on logic or reason); and 3) *pathos*, i.e. appeal based on emotion. We then apply this rhetorical theory to house ads. An examination of our dataset of house descriptions (as well as experience in observing the Glasgow housing market over several years) quickly shows that *ethos* does not play a significant role in the language of house descriptions. For example, we found no examples of the type, "the trusted firm of John Smith Realtors brings this property to the market, etc."¹ This is not to say that estate agents do not engage in *ethos* activities in other settings (advertising campaigns, leafleting, branding, blurb on website-homepages, for example, may well contain claims of integrity), just that, in the short property descriptions that make up our data, no such incidents were evident. In contrast, *pathos* occurs frequently in property descriptions and covers those elements that aren't usefully describing a tangible, central feature of a property. The majority of property descriptions, however, are dominated by the language of *logos*. It is, unfortunately, sometimes a matter of judgement whether a description amounts to logos or pathos. For example, references to "original features" might be construed as logos, but the phrase also has elements of pathos in its link to historical appeal. We attempt to mitigate the problem of subjectivity in the categorisation of words 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

¹ There is some reference to particular homebuilders, such as Bett, Cala, Miller, etc., but arguably this is strictly descriptive rather than status-seeking

our findings. We also identify a number of sub-categories of pathos, based on the various facets of consumer desire and emotional vulnerability. Our empirical analysis is primarily concerned with the propensity of these different definitions of pathos to vary over time and space.

The full account of our investigation adheres to the following structure. In section 2, we present a brief summary of the relevant 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. Literature Review

As noted in the introduction, the language used by estate agents in the course of selling properties has attracted relatively little attention in the academic literature. Traditionally, the role of estate agents is assumed to be neutral, though the assumption is typically implicit rather than overt. In recent years, a number of housing economists have relaxed this assumption and investigated the possible perverse incentives faced by estate agents. This work, however, has focussed on agent effort, fees and advice and has tended to overlook the particular role of language. Sirmans et al (1995), for example, considered the possibility that estate agents have an incentive “to urge the seller to accept suboptimal offers” (p.230). However, they find no significant difference between the selling prices of properties that sold quickly compared with those that sold after the normal time on the market. This contrasts with the findings of Levitt and Syverson (2005), who compare the prices and selling-times of properties owned by estate agents with the selling price of properties owned by their clients. They find that homes owned by estate agents sell for around 4% more than other houses and are typically on the market for approximately ten days longer, suggesting that estate agents push harder for better prices on their own properties. Other authors also have compared the prices of properties sold by estate agents with those sold by owners. Elder et al (2000) found that real estate brokers have no independent effect on selling prices. In addition, it would appear that offering more commission does not lead to estate agents obtaining higher prices for houses (Munneke and Yavas 2000 compare selling times and prices of houses sold through full-commission agents and traditional agents).

To our knowledge, the only paper in these various strands of literature to have attempted any form of textual analysis of the language used by estate agents is the working paper by Levitt and Syverson (2005). Their primary concern, however, is the difference in selling time and sale price of properties owned by estate agents compared to properties owned by clients, rather than the use of language *per se*. Levitt and Syverson enter a number of words (such as “beautiful”, “appealing” and “wonderful”) as independent variables in a house price regression in an attempt to control for unobserved differences in quality between dwellings owned by the agents themselves and all other properties. They conclude that, “Systematic quality differences appear to be responsible for part of the gap between agent-owned and non-agent-owned homes sales prices and times-on-market” (p.11). However, “most of the difference is explained by broad indicators of age and style rather than by more subtle characteristics picked up by the agents’ descriptions” (p. 12).”

In seeking to identify the role that realtor language plays in determining the marketing time and price, we believe that a fundamental stage in the analysis first has to be thoroughly established in order for subsequent attempts at causal reasoning to be meaningful. The preliminary stage to which we refer is a rigorous understanding of the use of language by estate agents and how it varies over the market cycle and across submarkets. This is the missing element in the literature and is the area to which the current paper hopes to contribute.

While the language of realtors appears to have thus far escaped serious quantitative analysis, we should note that there have been qualitative investigations in the related field of television property programmes. Lorenzo-Dus (2006), for example, conducted a textual analysis of 45 episodes of British primetime television property shows. Her goal was to examine the strategies utilised by the lifestyle media to persuade viewers to pursue specific lifestyles. Although the analysis cannot give us direct insight into the parlance of estate agents, it does deepen our understanding of the ambient use of language that conditions and articulates the lifestyle aspirations underpinning housing demand. It is not infeasible that estate agents may themselves draw upon the language cues of property shows and the broader property media in their bid to persuade potential buyers that a particular property is worth viewing (or even worth paying more money for). So the role of the media is certainly key in establishing and refining the common terms of dialogue between estate agents and house buyers.

The discussion by Lorenzo-Dus about the nature of persuasion is of particular interest to our inquiry. Drawing on Pardo's (2001, p. 99) study of persuasion in relation to the discourse of globalization in Argentina, she highlights 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 return to this task 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 as to the structures and dynamics of housing markets that enrich our understanding of how markets operate and may further identify a source of market failure (on the basis that variation frustrates decoding).

To help us structure our analysis, we summarise below a range of possible theories of why we might expect the language of selling to vary and the nature of the variation one would expect from each theory. Before we do that, let us first posit a null hypothesis – a theory that counters the notion that we should expect variation in the use of language:

1. *Drivers of Uniformity*: It is possible that 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 degree of uniformity and stability in the nature 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 may have all but eradicated temporal, regional and intra-urban variation in the language of selling houses.

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, but involve a degree of random selection from the population of words contained in common vocabulary. The same is true of collections of sentences and, indeed, of property descriptions. Just how many emotive words are used is likely to vary from advert to advert due to purely random factors that affect the predisposition of the estate agent and the set of words that come to mind on the day of writing. Lack of formal training or professional qualifications may exacerbate the random variation in advert-writing styles between agents. This theory suggests a volatile but stationary time series of language variation. Depending on the amplitude of the white noise, it has the capacity, nonetheless, to frustrate communication. It is unlikely, however, to cause secular, cyclical or seasonal variations.

3. *Lagged Response to Legislation*: it has been suggested² that, following the introduction of the Property Misdescriptions Act in 1991, estate agents went through a prolonged period of being excessively cautious in their use of language in property ads. Caution gradually gave way to the resurgence of effusive language as it became clear that prosecutions were rare. This theory would not, of course, explain seasonal or cyclical variations. It seems likely to be more relevant to the care taken in listing the technical specifications of a property than the use of emotive language (it is improbable, for example, that an agent could be successfully prosecuted for the use of a subjective adjective such as “beautiful”). More importantly, 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 offer any explanatory power in 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 variations over time in the characteristics of properties coming onto the market.
5. *Cycles in Staff Composition*: it has been suggested³ that the composition of staff at a firm of realtors changes over the market cycle. As the market booms, new staff are needed to cope with the rising turnover over of properties. These new employees are typically less experienced and more prone to elaborate description. Experienced staff know that buyers are not easily duped, however, and have learned that a more judicious approach is to be preferred. When the market slows, there are insufficient sales to maintain the expanded workforce, and staff are typically laid-off on a LIFO (last-in-first-out) basis, increasing the share of experienced agents. This theory would suggest a pro-cyclical pattern in the language of selling. It would not, however, lead to regular seasonal or secular patterns in the use of language.
6. *Irrational Exuberance*: Value, as determined by the market, is not an intrinsic constant. So at the moment when the price of housing is rising rapidly relative to other goods, it may seem odd to apply the same bland description penned when a house was worth half its current value. 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 irrationality of agent exuberance is seen for what it is, and more restrained descriptions become the norm. Alternatively – or additionally – agents may utilise such words at points in the market cycle when they are most likely to foster such excitement in potential buyers. This theory would suggest that particular types of emotive language – those less grounded in reason – will be more volatile than others, and be more sensitive to market swings. It would not, however, lead us to expect particular seasonal or secular movements.

² by delegates at the National Association of Estate Agents Conference, March 2007, in response to a presentation of an earlier draft of this work.

³ again by delegates at the National Association of Estate Agents Conference, March 2007.

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, however, lead to particular seasonal or secular patterns in the data.
8. *Opportunity Cost of Viewing:* A problem facing any theory that posits the strategic manipulation of language by estate agents (as in theory 7) is the assumption of buyer naivety. It is implausible that a buyer would make such a large purchase on the basis of a property description, particularly when the profession providing that description is notorious for its propensity to exaggerate. No-one bids without viewing, so why should the parlance of property ads have any affect at all on whether (and what) buyers are willing to bid? The answer may lie in the opportunity cost of viewing a property. This cost arises from the fact that 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 glowing description, house-hunters still have a strong incentive to submit an offer. Turning down a property after viewing introduces the risk that continued search will not yield a superior alternative within the buyer’s timeframe.

As far as the estate agent is concerned, viewing is 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 ads are not the basis on which purchases are made – that is not their purpose. Their purpose is simply to attract viewers. Once this is achieved, the opportunity costs associated with buyer search behaviour will hopefully more than compensate for any disgruntlement arising from exaggerated claims.

But why should this theory imply variation in the use of language? Variation in language occurs when there is variation in the opportunity cost of viewing. This influences the relative returns to marketing strategies that succeed in encouraging 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 at certain phases of the cycle, to use emotive language. When properties are selling very quickly, the effective choice-set facing a buyer may be very small at any given moment, even though there are many properties coming onto the market. For example, 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, by extension, 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.

This explanation suggests pro-cyclical patterns in the use of hyperbole in the language of selling. It also suggests that there may be seasonal aspects to the use of such language because of 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 school catchment.

Theories of Spatial Variation in the Use of Language

With the exception of 2 and 3, each of the above theories of temporal variation in the use of language potentially also implies variation across space because of differences in the timing of the market cycle across submarkets (Pryce and Gibb 2006), different long-term levels of demand, and differences in the quality of stock across space. We can add a further theory that pertains exclusively to spatial variation:

9. *Local conventions*: Given the tendency for local conventions to occur in language generally – the propensity for differences in accents, pronunciation, idiom and terminology occur and persist at small spatial scales in the wider population – it would not be surprising if such developments occur in the language of selling. Most moves are local moves, so conceivably differences in marketing language could still persist despite the innovations in communication technology. Local moves could foster and preserve a common dialect between estate agents and those in the surrounding community, one that is only comprehensible because of persistence in the spatial and cultural proximity of the parties involved in the majority of transactions. While moves are relatively infrequent, the interaction of punters and agents may not be (part of the goal of estate agents is to tempt owners to move, and the propensity of consumers to “window shop” facilitates this ongoing dialogue). This theory implies that there will be persistence over time in spatial patterns of language.

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:

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.

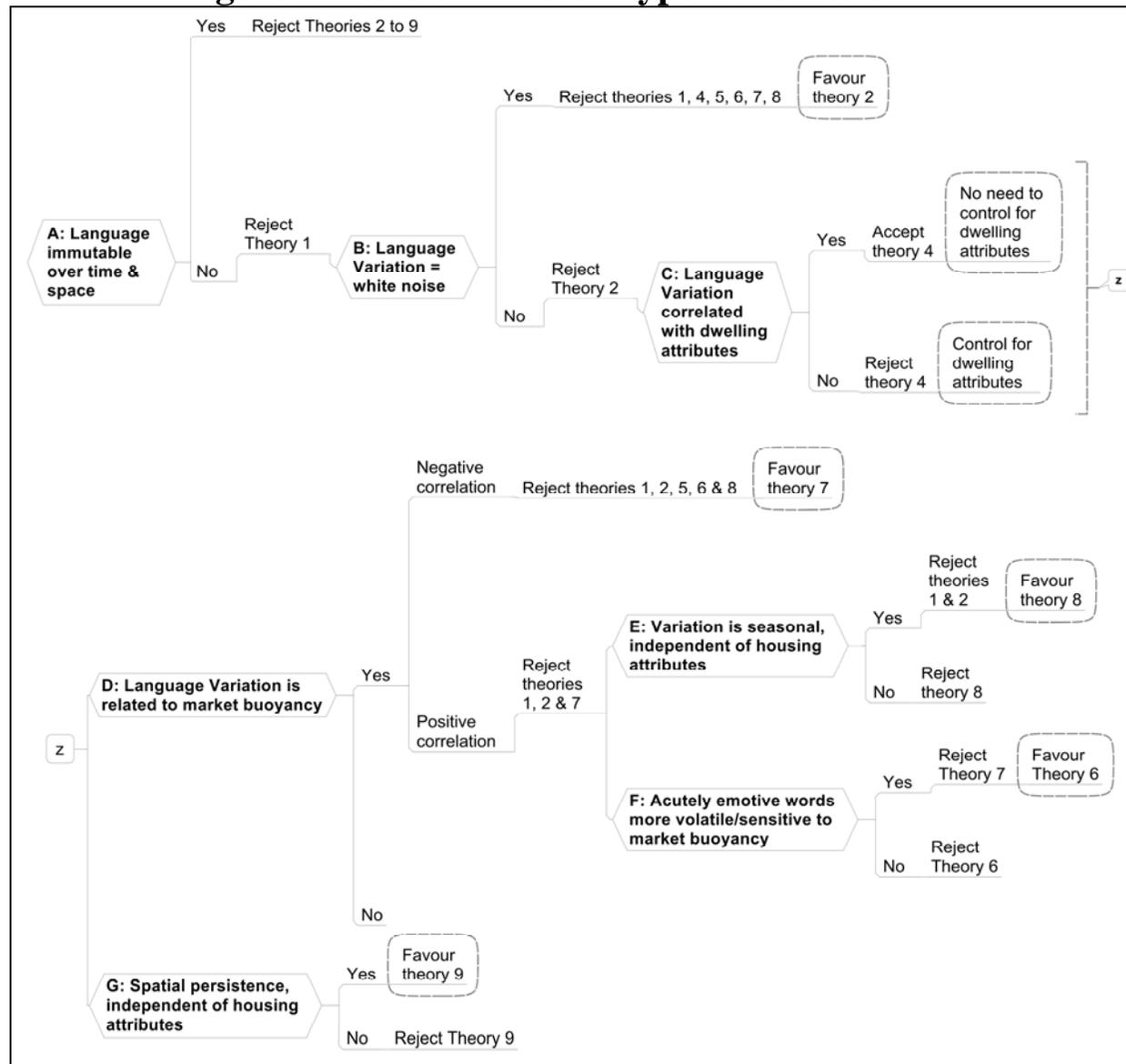
Hypothesis E: Variation in marketing language is seasonal, independent of variations in housing characteristics.

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 hypotheses) 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 sake of clarity and simplicity. We also recognise that there may be questions raised by our theories that are broader than the range that 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 Decision Tree of Hypothesis Tests



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? Without categorization, there is no measurement of variation. In this paper we assume that property promotion entails an attempt to persuade potential buyers to view and bid for the property. This assumption is something of a tautology as estate agency can be defined as the marketing of homes for sale, and promotion is one of the four ‘Ps’ of the “marketing mix” – Price, Product, Place and Promotion – the most common way of defining the activities of a marketer. The fact that the estate agent use of language is motivated by the desire to persuade (rather than simply 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).

Our approach, expounded in the pages that follow, is to apply this characterisation of the act of persuasion to around 49,926 written property descriptions published by estate agents. As noted in the introduction, we found no evidence of *ethos* in our data, though 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 developed these categories using both our own knowledge of the Glasgow real estate market as well as an examination of many of the ad 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 must be applied to each of the 49,926 published descriptions in our data. We can then measure the incidence of each category of *pathos* as the proportion of words in each property description that fall into each sub-category. Once this has been achieved we shall be able to consider how these proportions vary over time and space.

One of the difficulties associated with this kind of categorisation is its inevitable 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. As a result we construct an additional category based on a much narrower definition of *pathos* which we call *Core Pathos*. We now have a total of six categories of *pathos*, descriptions of which are given 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 will come 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 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*) but 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", "Enviably", "Prestig", "Splend", "Prestigious", "Smart", "Character", "Executive", and "Eleg").

Textual Analysis

Having established a framework for categorising marketing language, our next step was to conduct a detailed textual examination of a selection of descriptions with a view to framing the subsequent quantitative investigation. We employed 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 examined both words and phrases to look for trends and patterns. However, as with work by Ian Budge *et al* (1987) on political party manifestoes, we attempted to go beyond merely counting words. We look at both how often the word appears (to establish which words were the most common in 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 ads. 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, while it is left to the quantitative analysis to do the task of weighing our pathos results in a way that holds the measurement of pathos constant across a range of factors.

Statistical Analysis

The final stage in our research involved applying bivariate and multiple regression analysis in an attempt to verify or falsify the hypotheses presented above. Bivariate analysis entailed plotting summary measures of pathos over time and across space to confirm whether variation does indeed occur. For example, we conducted equality of means t-tests to investigate apparent differences in the incidence of pathos across submarkets, and computed 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 techniques. 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 one of the assumptions of ordinary least squares regression (OLS). There is nothing in the standard algorithm for OLS to tell it that the dependent variable is bounded and so it will potentially predict outside the zero-one range. The problem is a similar to that of modelling mortgage debt as a proportion of house value (Hendershott and Pryce 2006) or the proportion of employees participating in pension plans (Papke and Wooldridge 1996), or in fact any situation when the dependent variable is continuous but restricted to the interval $[c, d]$.⁴ One solution to the problem of modelling variables bounded between zero and one is to apply the log-odds transformation to the dependent variable ($\log[y/(1-y)]$) which allows OLS to be applied to the estimation of $\mathbf{x}\beta$. According to Wooldridge (2002) this approach has two major drawbacks, however:

“First, it cannot be used directly if y takes on the boundary values, zero and one. While we can always use adjustments for the boundary values, such adjustments are necessarily arbitrary. Second, even if y is strictly inside the unit interval, β is difficult to interpret: without further assumptions, it is not possible to recover an estimate of $E(y|\mathbf{x})$, and with further assumptions, it is still nontrivial to estimate $E(y|\mathbf{x})$.” (Wooldridge, 2002, p.662).

Papke and Wooldridge (1996) and Wooldridge (2002) suggest modelling $E(y|\mathbf{x})$ as a logistic function, as in [4], which ensures that “predicted values for y are in $(0,1)$ and that the effect of any x_i on $E(y|\mathbf{x})$ diminishes as $\mathbf{x} \rightarrow \infty$.” (Wooldridge, 2002, p.662). The technique, labelled Fractional Logit Regression (FLR), has recently been used in the real estate literature by Hendershott and Pryce (2006) to model loan-to-value ratios and it is the method we apply here to investigate the determinants of the incidence of pathos in the language of selling.

We also make use of a feature common to all logit models that the exponent of the coefficient equals the proportionate change in odds caused by the variable in question – a measure that holds all other effects constant. If the proportionate change in odds is greater than one, then as the explanatory variable increases, the odds of the outcome will increase (i.e. the explanatory

⁴ Where c and d do not equal 0 and 1 respectively, fractional logit estimation can be applied by transforming y to ensure that it lies in the $[0,1]$ range. Wooldridge (2002, p. 661) suggests the following simple transformation: $(y_2 - c)/(d - c)$.

variable in question has a positive effect on the dependent variable). The larger the value, the more sensitive the predicted probability is to unit changes in the variable. Conversely, if the proportionate change in odds is less than one, then as the explanatory variable increases, the predicted probability declines (i.e. the variable has a *negative* effect on the dependent variable), and the *smaller* the value of exponent of the coefficient, $\text{Exp}(B)$, the more sensitive the predicted probability is to unit changes in the variable.

Data

Our analysis is based on information extracted from 49,926 property transactions in the Strathclyde conurbation. 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 below). 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. Variables that describe the textual composition of the property description were created using the PATHOS program (see Appendix 1). This is a routine written in the Stata programming language that counts the number of times a word from each *pathos* category occurs in each description.

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 simply lists 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

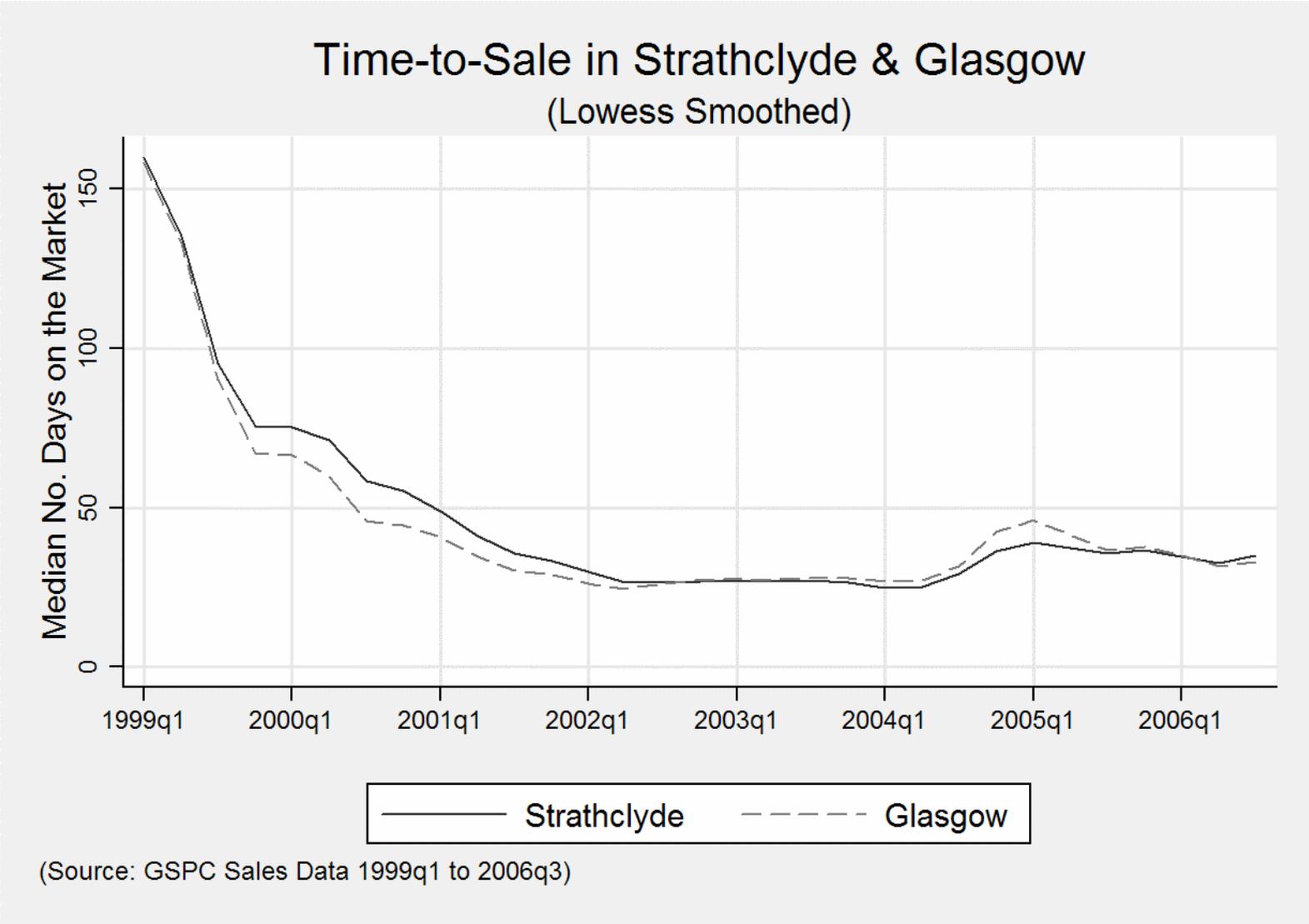


Table 1 Descriptive Statistics

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
tom	Time on the market (in days)	70.66	35.00	-79.00	2917.00	49,919
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

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). We examined every word that appeared more than 100 times in the data base. Each word was studied in a sample of the ads 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.

It is interesting to note that 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.’ Table 2 below defines which of these most common words appear to best denote pathos and displays these words in bold. In order for a word to qualify as the best sort of 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 ads. To qualify as true real-estate pathos, the word had to have an elusive and somewhat flexible meaning, to function beyond the somewhat dry and trite phrases (‘must view’ etc.) found in many ads.

Here, we see that one might conceive of *Core Pathos* as a smaller subset of the wide number of rhetorical words that estate agents use in their house descriptions. 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 ads, 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 ads and only two to Mackintosh (and one misspelled) in the 22,613 GSPC ads 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 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 its four subcategories) and also our narrow definition (which recognises only the “core” words as being truly pathos).

Table 2 Most common words in GSPC house ads: Where is the pathos?

<i>Word/ word string</i>	<i>Word count</i>	<i>Comment</i>
!	1372	Used in a variety of contexts (<i>pathos</i> and <i>logos</i>) but seems to be somewhat random. For example: "Upgrading a must!" and "Fantastic views over golf course!".
Abs	248	Generic estate-agent speak with little emotive resonance.
Appeal	481	Used as a modifier in less meaningful phrases.
Apprec	685	Descended into estate-agent speak, little emotive resonance, "must view to appreciate" would appear to be filler.
Att	7025	Word string possibly too flexible – included in too broad a range of words (e.g. attic, flatted).
Beaut	3422	Used frequently in the phrase "beautifully appointed", which is not a particularly emotive phrase.
Beautiful	457	Pathos -- when not part of phrase "beautifully appointed"
Bright	3134	Generic estate-agent speak, so little emotive value. Typically used in phrases like "bright and spacious" and "bright lounge".
Character	106	Pathos
Charm	410	Pathos – indeed what makes a bathroom "charming"?
Class	138	Generic estate-agent speak, little emotive resonance; typically found phrases like "first class order throughout." The interpretation of "first-class" could be quite broad so not a core pathos word.
Decep	612	Used a lot, but what does it really mean? If it indicates "better than it seems" as in the phrase "deceptively spacious" than it is the opposite of pathos. Again, essentially meaningless.
Del	1101	Pathos
Des	2658	Word string possibly too flexible, e.g. "resides" etc.
Eleg	101	Pathos
Enjoy	1674	Generic estate-agent speak, little emotive resonance.
Enviabl	199	Pathos
Essent	2811	Typically used in "viewing essential," no emotional meaning.
Estab	579	It is unclear what this is signalling. This is generally found in the phrase "established neighbourhood."
Exceptional	342	Pathos.
Excl	305	Pathos, as in "exclusive"
Executive	102	Pathos
Extr	5408	Usually just lists extras, essentially meaningless as all homes offer something "extra".
Fab	1540	Generic estate-agent speak, little emotive resonance.
Fant	225	Pathos, as in "fantastic"
Flex	726	This appears to have little meaning – it is not clear what "flexible" means. Arguably any home bigger than one room has "flexible" layout.
Fresh	971	Possibly descriptive (i.e. <i>logos</i>) as opposed to pathos –freshly decorated, which tells buyers about the physical condition (as opposed to more ephemeral qualities of the property).
Generous	1988	Pathos, interesting use in place of "large"
Great	487	Often used in phrases like, "Greatly reduced". Also used as a modifier in less meaningful phrases. "Greatly reduced" does sound a signal, but it is one of a bargain and possibly <i>logos</i> .
Handsome	126	Generally just used to describe buildings.
High	4052	Possibly too many meanings, high floor etc to capture the emotive use of this word which seems only to occur in a minority of cases.
Ideal	5045	Usually in phrase "ideal for first-time buyer" – which is not really an emotive phrase
Imaginative	145	Double edged, as "imaginative" can signal something non-standard that can hold lower market value
Immac	3556	This seems to denote clean, tidy. Hence perhaps closer to <i>logos</i> than <i>pathos</i> .
Immed	267	"Immediately available" signals a buyer keen to sell – and perhaps at a lower price. This is perhaps more typically <i>logos</i> than <i>pathos</i> .
Imp	2647	Generic estate-agent speak with little emotive resonance.
Impress	1664	Pathos

Individ	242	Double edged, as “individ” can signal something non-standard that can hold lower market value. “Unique” can also have this meaning, but perhaps less so.
Lovely	570	Pathos. Interesting to note that this word is popular while more extreme words such as “opulent” or “lavish” are not. This word is generally used to describe gardens.
Lux	1540	Possibly more of a <i>logos</i> word, as relatively common phrases “luxury bathroom” or “luxury kitchen” abound. Arguably this is used to describe more expensive fittings.
Mag	871	Word string possibly too flexible – included in too broad a range of words (Magnet, magazine etc).
Magnif	637	Pathos
Mature	952	Perhaps more descriptive (<i>logos</i>) than emotive (<i>pathos</i>). This generally refers to gardens.
Most	1173	Generic estate-agent speak, little emotive resonance.
Much	1257	Generic estate-agent speak, little emotive resonance.
Must	1129	Generic estate-agent speak as in “must view”, little emotive resonance.
Of	783	Word string possibly too flexible – included in too broad a range of words (loft etc).
Orig	792	Bridges <i>pathos</i> and <i>logos</i> – although the notion of “original features” is emotion, it is also valuable in today’s mark.
Outst	587	Pathos, generally used to describe views from the property.
Pleas	871	Pathos
Pop	7692	Essentially meaningless and so not particularly potent as a <i>pathos</i> word.
Preferred	550	Pathos. Usually refers to “preferred” first floor but not always, so the word is just a signal that we are supposed to value this attribute of a flat, sometimes used in a strange way
Prestig	172	Pathos
Prestigious	149	Pathos
Prime	1141	Pathos, often used to discuss position/location
Professional	221	Pathos
Quality	1393	Descended into estate-agent speak, little emotive resonance.
Rare	1917	Descended into estate-agent speak, little emotive resonance. Generally part of phrase “rarely available” but this is clearly over-used.
Seldom	1410	Descended into estate-agent speak, little emotive resonance. Generally part of phrase “seldom available” but this is clearly over-used.
Smart	140	Pathos
Sought after	3933	Pathos
Splend	161	Pathos
Stun	1390	Pathos
Stylish	930	Descended into estate-agent speak, little emotive resonance. The word seems dated.
Sunny	230	Pathos
Sup	3137	Better to use “superb”.
Superb	2863	Pathos
Taste	581	Descended into estate-agent speak, little emotive resonance, usually in phrase “tastefully decorated,” although this could mean a very wide range of tastes.
Trem	1328	Doesn’t work as it’s in “extremely”, only one use of “tremendous”
True	293	Used as a modifier in less meaningful phrases.
Truly	298	Generic estate-agent speak, little emotive resonance.
Unique	292	Pathos
Versatile	142	Double edged, as “versatile” can signal something non-standard that can hold lower market value. As with “flexible”, it’s a bit meaningless.
Very	1315	Generic estate-agent speak, little emotive resonance.
Well	9484	Possibly too many uses and too vague to count as <i>pathos</i>

Highlighted cells indicate that the word has been evaluated as “Core” *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 2 to 12 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 (sig. = 0.000 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 pathos (sig. = 0.000, $n_{\text{Glasgow}} = 22,613$; $n_{\text{non-Glasgow}} = 27,869$). The same is true if we compare submarkets within City of Glasgow. For example, the null of equal pathos is rejected if we compare the West End with all other submarkets (sig. = 0.000, $n_{\text{WEnd}} = 8,171$, $n_{\text{Other}} = 42,311$), with North Glasgow (sig. = 0.000, $n_{\text{Nglasgow}} = 1,427$), and with the South Side (sig. = 0.000, $n_{\text{Side}} = 8,423$). Similar rejections of uniformity arise if we compare North Lanarkshire with South Lanarkshire, Renfrewshire with East Renfrewshire, and North Ayrshire with East Dunbartonshire (sig. = 0.000 in each). All these tests yielded similar results using *Core Pathos*, our narrower definition of emotive words (the only noticeable difference was that sig. = 0.008 was slightly higher for the comparison of North Lanarkshire and South Lanarkshire, but still less than 0.05). 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.

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 (sig. = 0.000 for both one- and two-tail tests run on both *All Pathos* and *Core Pathos*; $n_{1\text{bed}} = 9,931$; $n_{2\text{bed}} = 21,495$). 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. *Core Pathos* = 0.000; two-tail sig. *Core Pathos* = 0.000; $n_{3\text{bed}} = 14,523$) and when comparing 3 and 4 bedroom properties (one-tail sig. *All Pathos* = 0.001; two-tail sig. *All Pathos* = 0.003; one-tail sig. *Core Pathos* = 0.000; two-tail sig. *Core Pathos* = 0.000; $n_{4\text{bed}} = 3,171$). 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 for *Core Pathos* (one-tail sig. = 0.000; two-tail sig. = 0.000). 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 the time period considered, 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:

- Price band 1 (the reference category) includes those properties whose selling price was in the lowest two deciles in the year that it sold

- Price band 2 includes those properties whose selling price was between the second and fourth decile in the year that it was sold.
- Price band 3 includes those properties whose selling price was between the fourth and sixth decile in the year that it was sold.
- Price band 4 includes those properties whose selling price was between the sixth and eighth decile in the year that it was sold.
- Price band 5 includes those properties whose selling price was in the top two deciles in the year that it was sold.

Comparing the incidence of pathos between price band 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 (and 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 in 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 so far 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 overall change procyclically 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



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

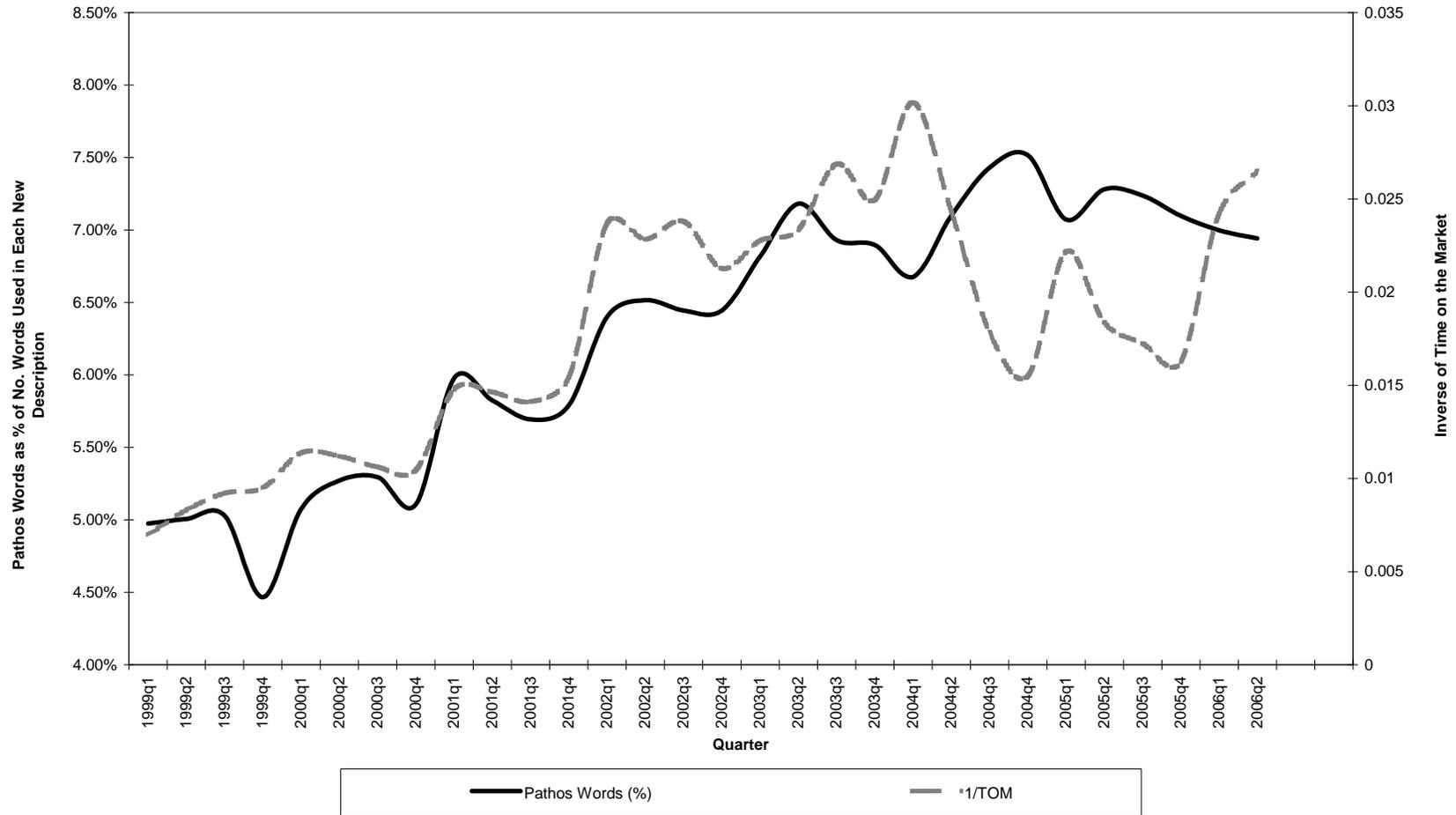


Figure 5 Number of GSPC Sales and the Incidence of Pathos in New Descriptions

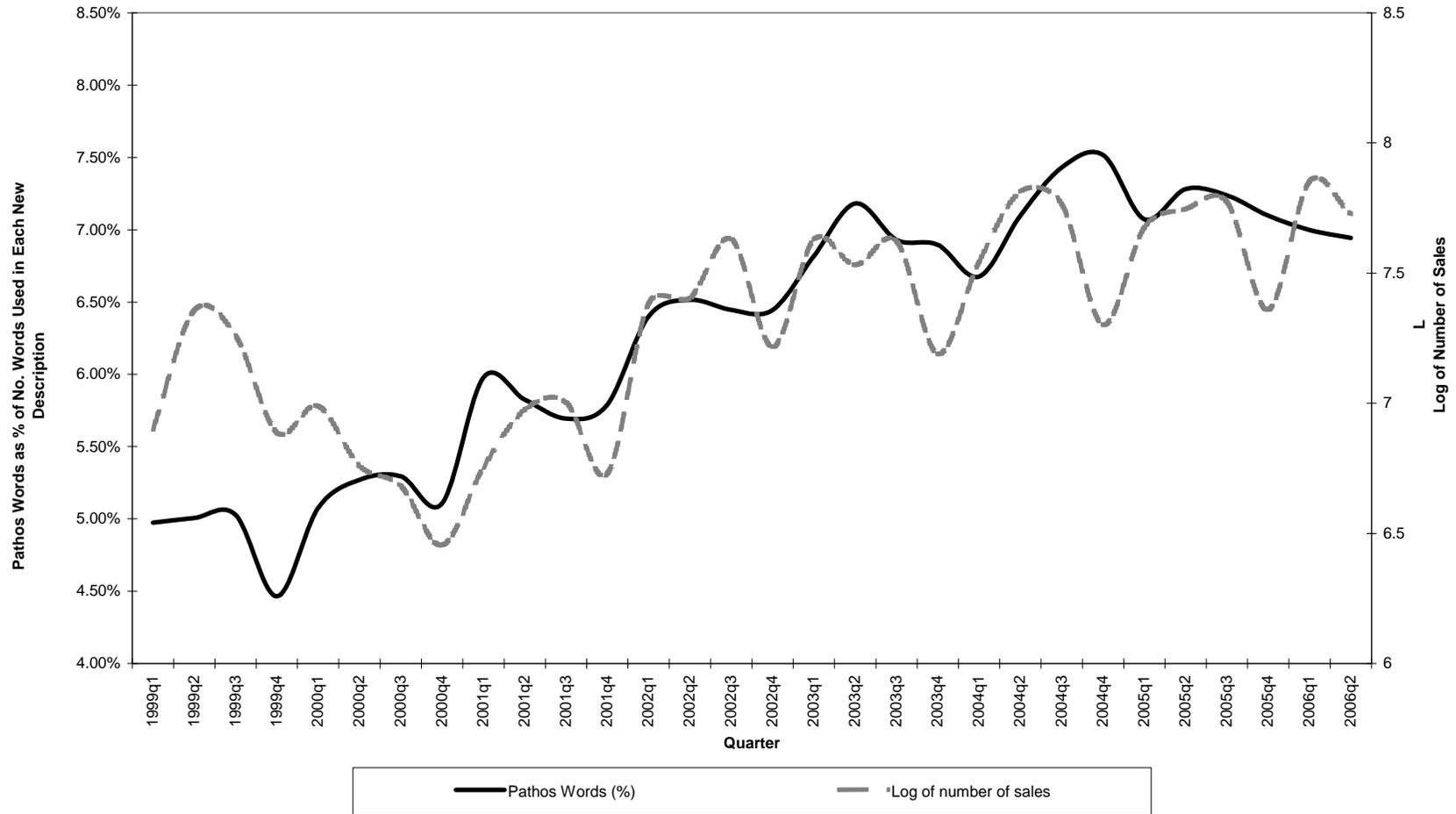
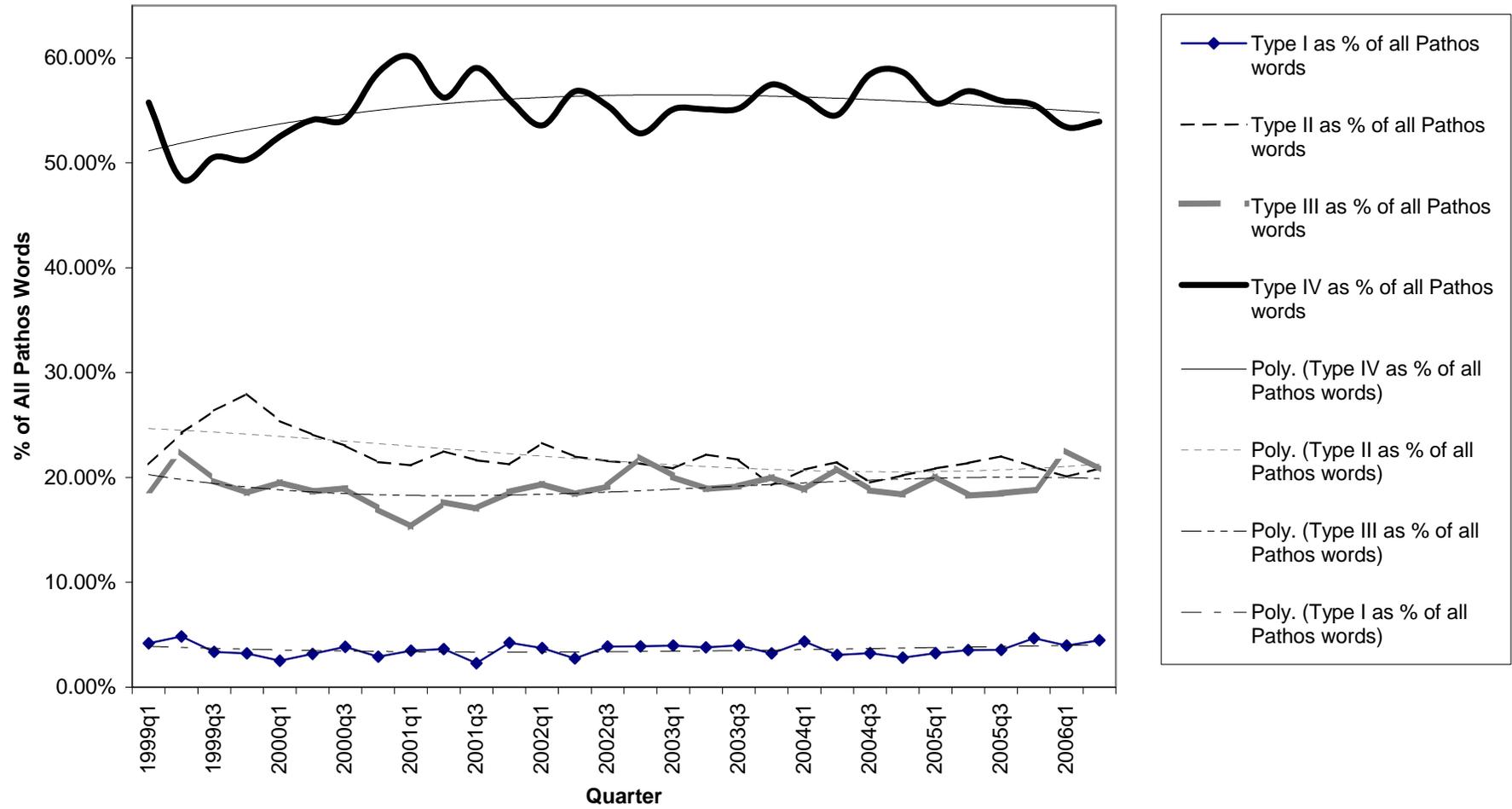


Figure 6 Variation in the Type of Pathos Over Time



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 **Error! Reference source not found.** 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 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 **Error! Reference source not found.**, 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 both spatial variation and a degree of temporal inertia in that spatial variation in the *type* of *pathos*. In fact, the spatial variation is much more volatile across space and the inertia less pronounced. This can be seen from the maps in Figure 10 and Figure 11 (which plot the contours for Type IV as % of *pathos* words in 1999/2000 and 2004/2005⁵ respectively) and also from the cross-sections (from Bearsden to Renfrew) in Figure 12. Note the different scale on the vertical axis of the cross-section graph compared to the previous cross-section graph– the first set of cross-sections vary by a factor of 5 (0.25 to 2.25), whereas the second set of cross-sections vary by a factor of 10 (0.1 to 1.0). In other words, the Type IV incidence of *pathos* as a proportion of all *pathos* words is twice as spatially volatile as the incidence of *pathos* generally. Also, the correlation between the two cross-sections in Figure 12 is less obvious, suggesting a lower degree of inertia. However, once again there is a need to establish whether this inertia 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.

⁵ Years had to be combined for the Type IV analysis because only observations could be included where the incidence of *pathos* was greater than zero.

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

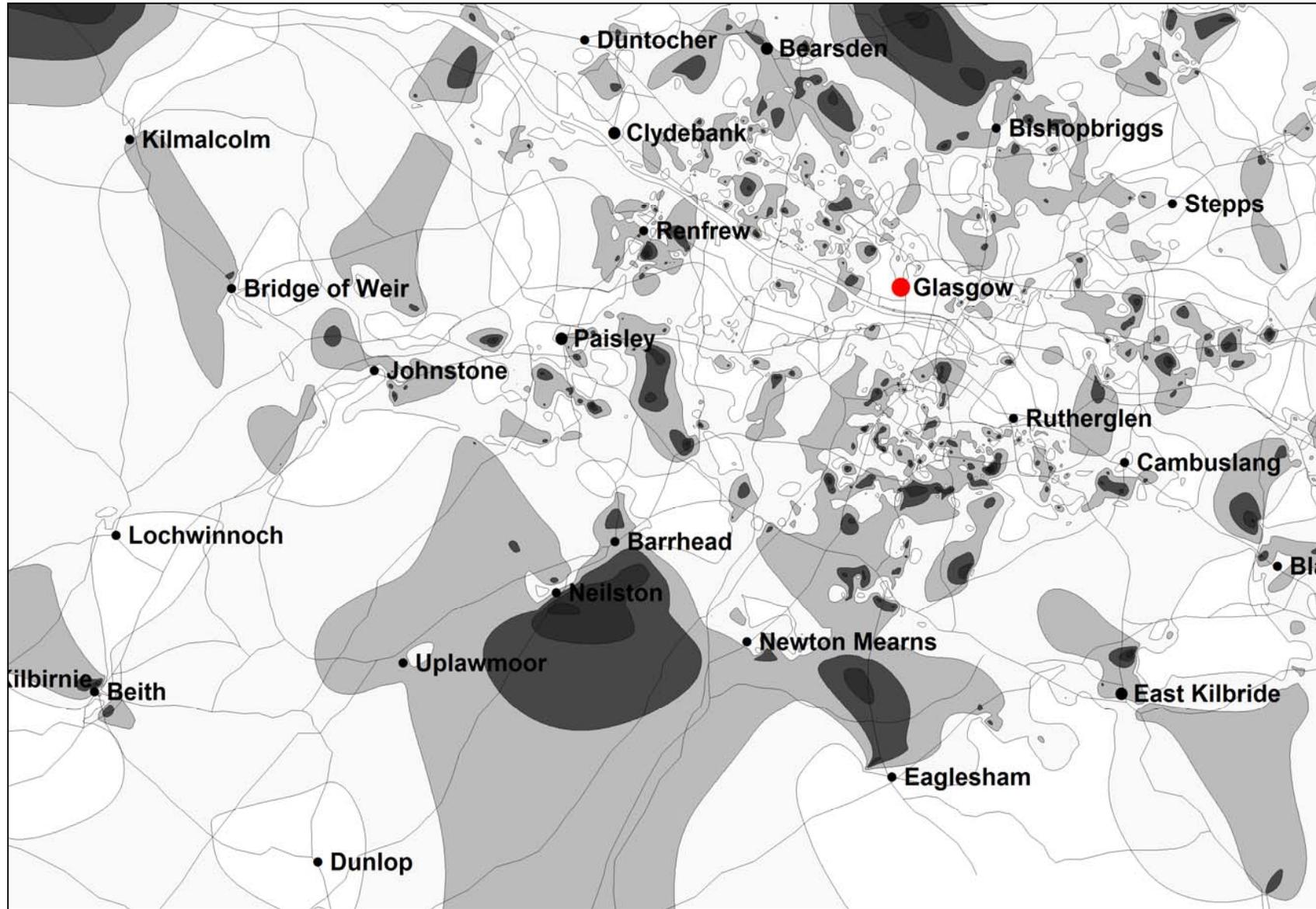


Figure 8 Spatial Variation of Pathos as % of No. Words in Each Property Description (2005)

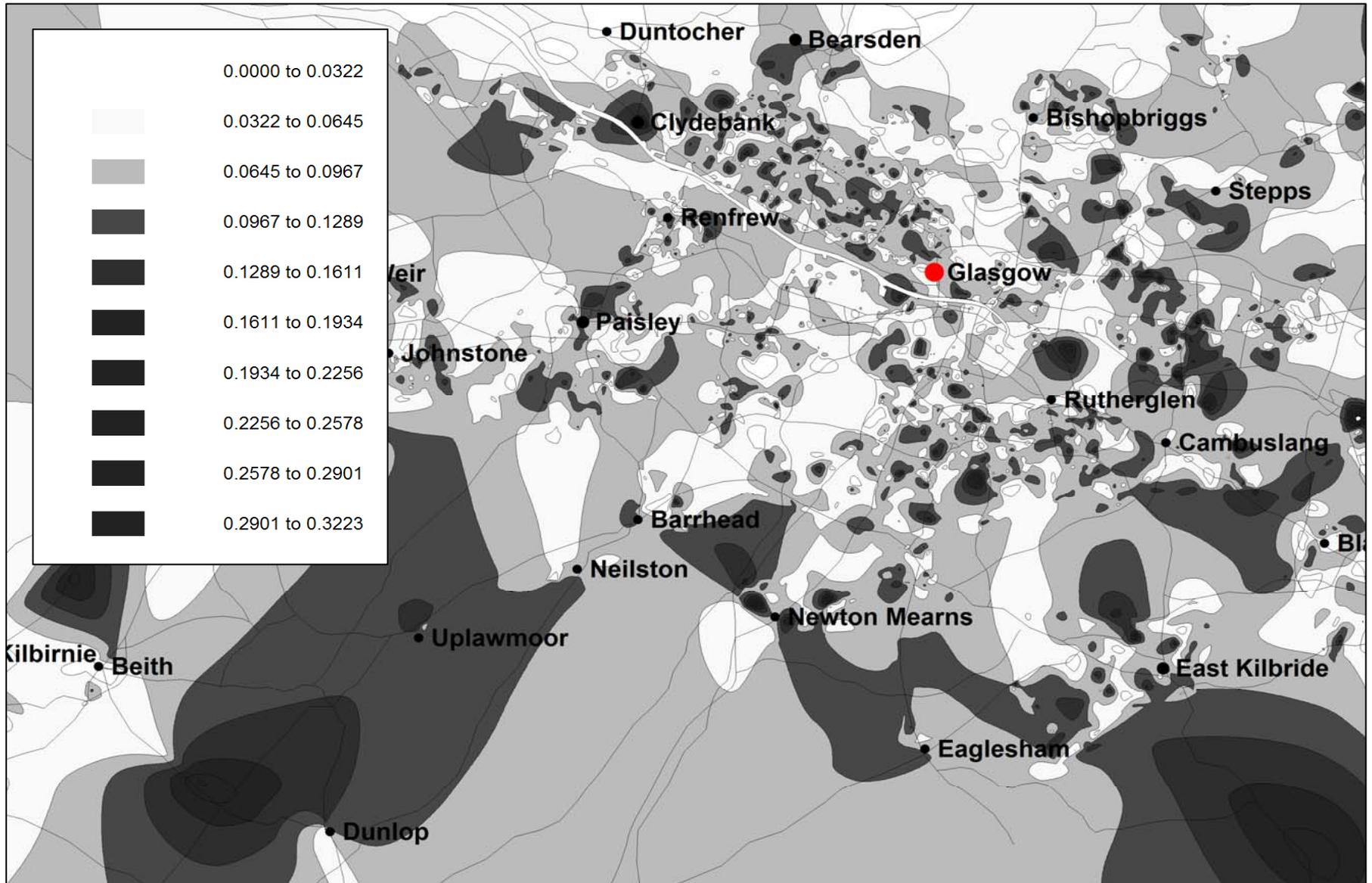


Figure 9 Cross-Section of the Pathos Surfaces from Bearsden to Renfrew

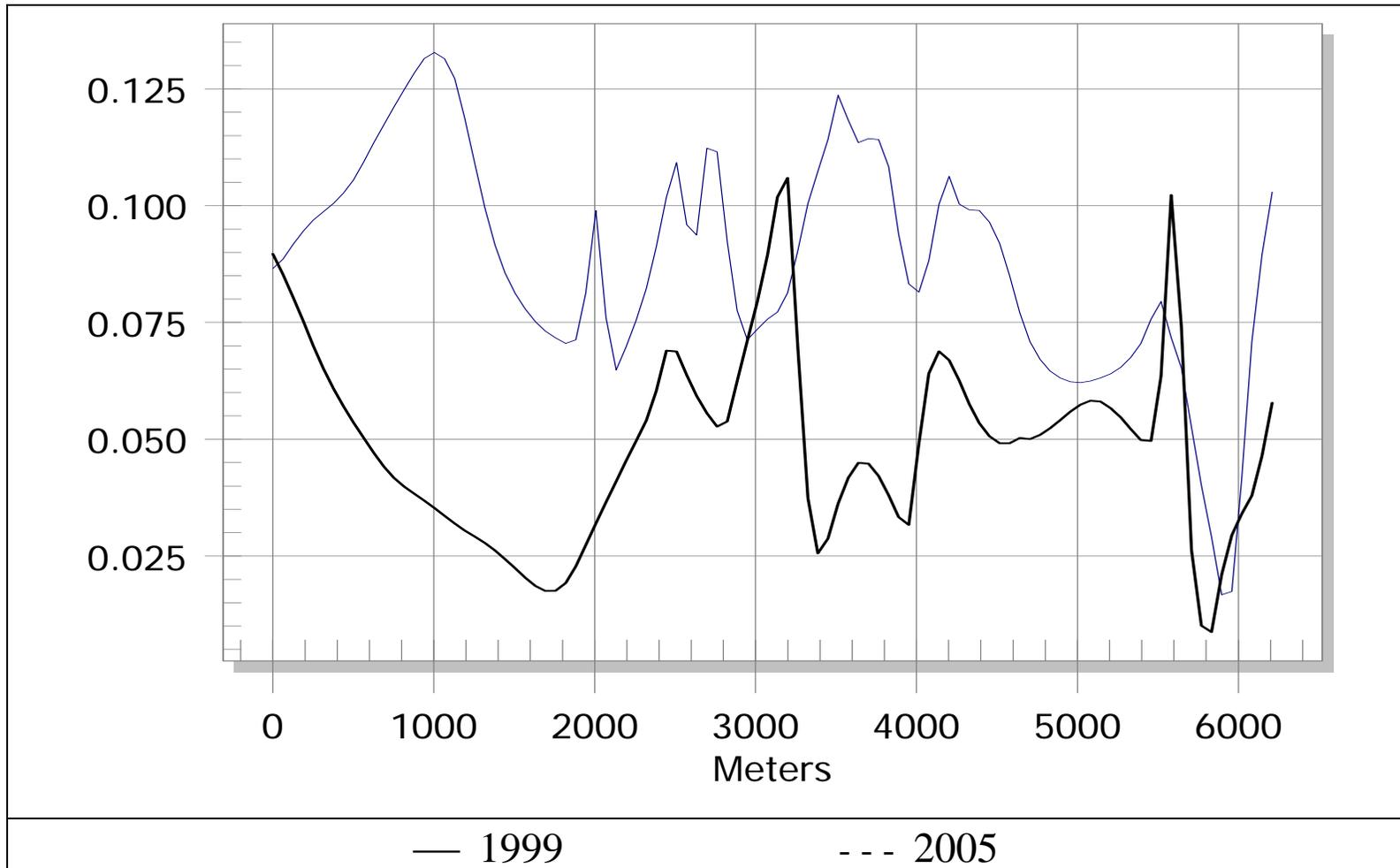


Figure 11 Spatial Variation of Type IV as % of No. of Pathos Words in 2004/2005

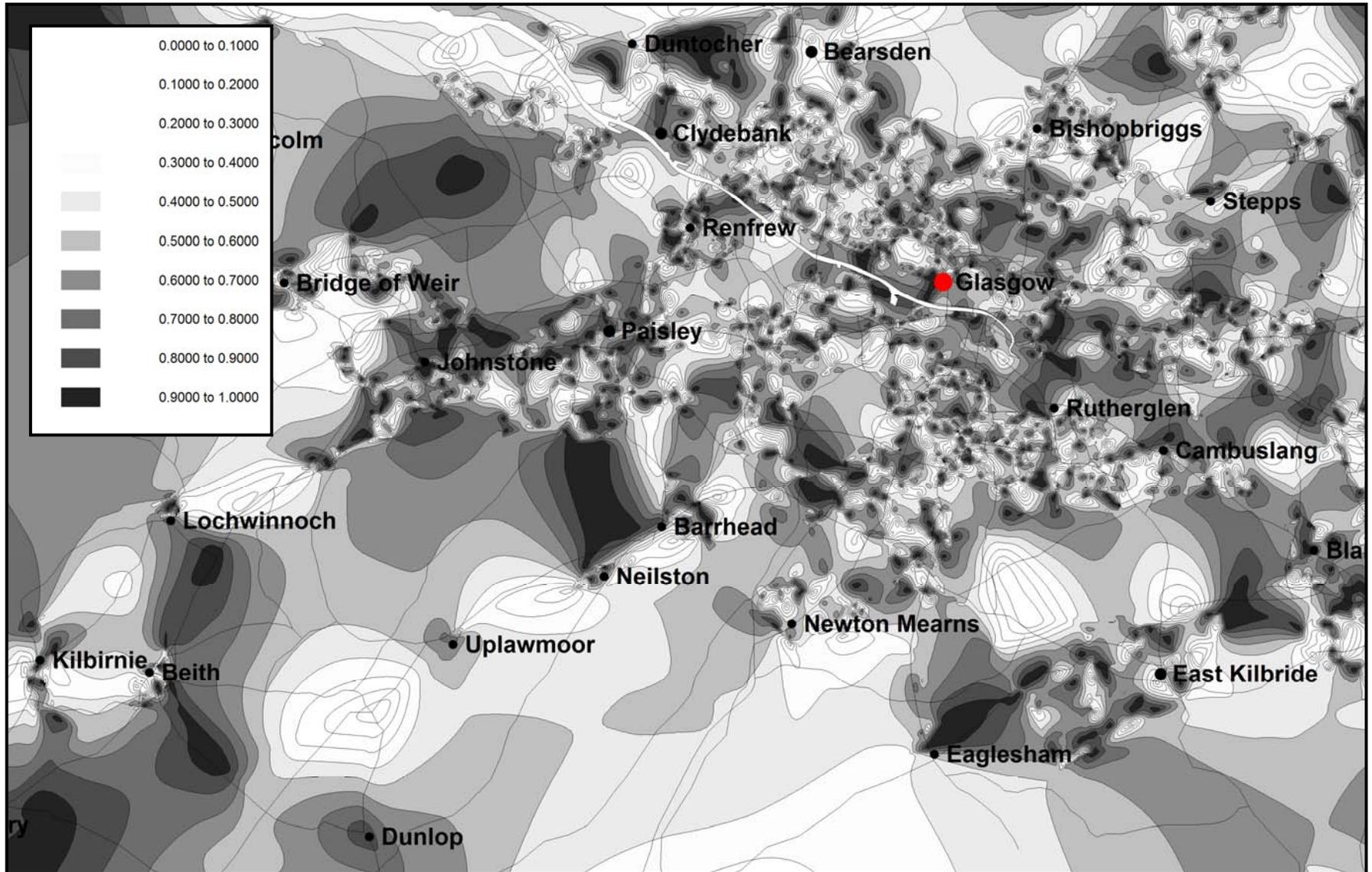
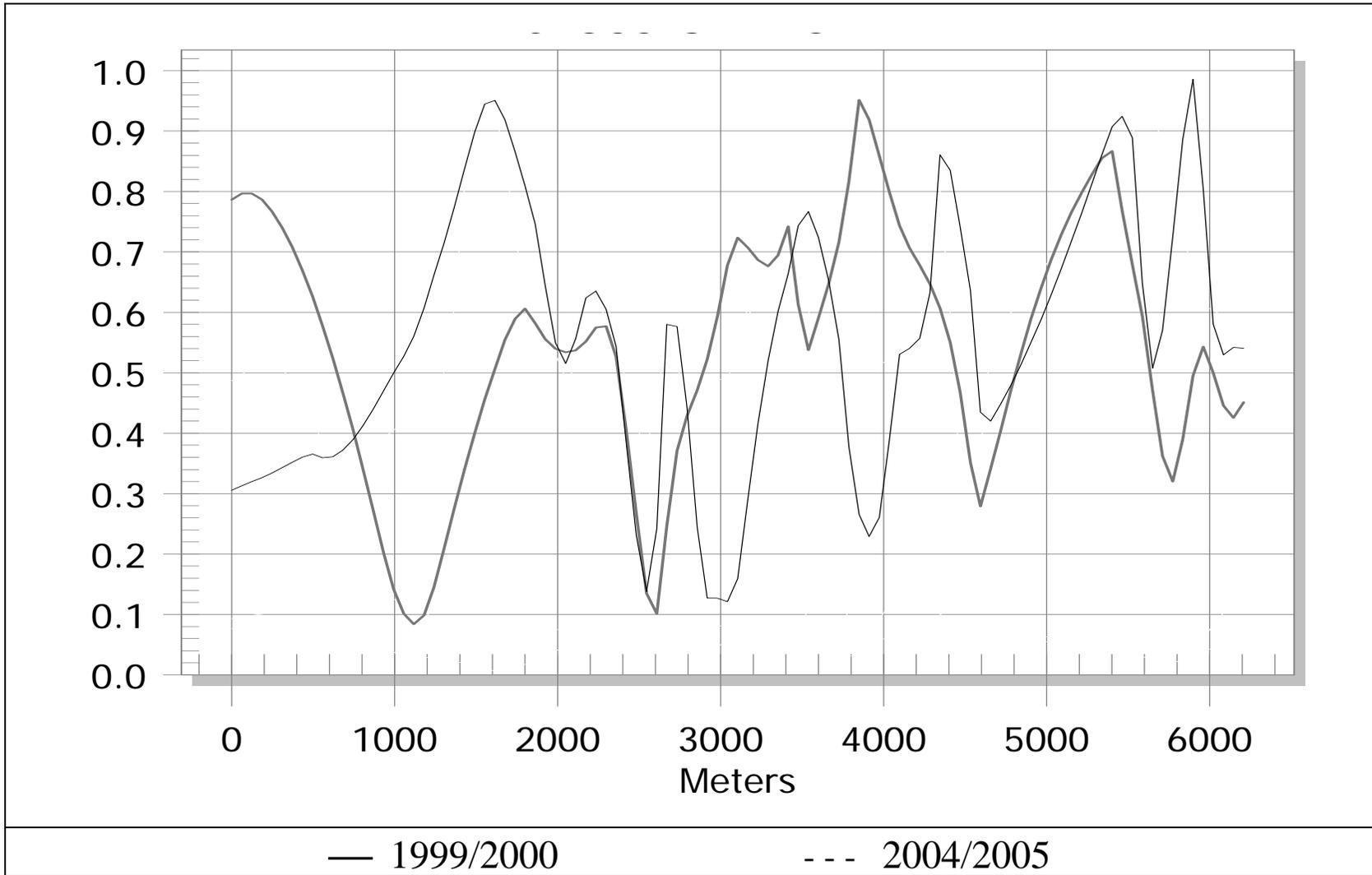


Figure 12 Cross-Section of the Type IV as % Pathos Surfaces from Bearsden to Renfrew



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 Spring⁶, with the incidence of pathos being slightly higher in Spring (one-tail sig. = 0.001; two-tail sig. = 0.001; $n_{\text{Winter}} = 9,663$; $n_{\text{Spring}} = 14,357$) 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; $n_{\text{Summer}} = 14,257$) and for *Core Pathos* (one-tail sig. = 0.000; two-tail sig. = 0.001). Use of pathos tends to fall in the Autumn, which is true for *All Pathos* (one-tail sig. = 0.002; two-tail sig. = 0.004; $n_{\text{Autumn}} = 12,205$) and for *Core Pathos* (one-tail sig. = 0.004; two-tail sig. = 0.009). 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 3 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% 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.

⁶ 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.

Table 3 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	

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 whether we would be able to identify any significant variation of the incidence of pathos if we were able to hold property attributes and other factors constant?

The corollary of these lines of questioning is to use multiple regression analysis. This will allow 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 4 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 which we have labeled “*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 without systematic component. Both these are rejected by the regression analysis. We include a range of independent variables that capture systematic drivers across either time and/or space (average postcode sector selling time for the quarter that a property comes onto the market, and average postcode sector pathos, deprivation, distance to Glasgow 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) the regression would not run, so Hypothesis A and Theory 1 can immediately be rejected. If the variation were entirely white noise, we would expect all the coefficients to be zero – the null hypothesis tested by the Chi² statistic. Given that the Chi² figure is large for all regressions reported in Table 4 (sig. = 0.000 in each case) we can reject Hypothesis B and Theory 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 likely only 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 a 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 a 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%; sig. = 0.000), compared with percentage change in odds in regression [5] of 97.9%; sig. = 0.000). 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-cyclicity. 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 4). 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% (sig. = 0.000) and 19.6% (sig. = 0.000) 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 the notion of idiosyncratic local conventions in language predicted in Theory 9.

We should note that these results were generally insensitive to changes to the model specification – the implications (if not the precise estimates) remained the same when we altered the list of independent variables (e.g. included year dummies, dropped seasonal dummies, dropped submarket dummies etc.).

Table 4 Fractional Logit Regressions for the Incidence of Pathos

<i>Independent Variables:</i>	<i>Dependent Variable ‡</i>					
	Pathos (all)	Type I	Type II	Type III	Type IV	Core
	[1]	[2]	[3]	[4]	[5]	[6]
Average selling time	0.979 § (0.000) †	0.987 (0.169)	1.002 (0.567)	0.948 (0.000)	0.979 (0.000)	0.956 (0.000)
Average Pathos in the area	1.124 (0.000)	1.051 (0.004)	1.116 (0.000)	1.134 (0.000)	1.117 (0.000)	1.196 (0.000)
deprivtn	1.004 (0.016)	0.979 (0.007)	1.006 (0.045)	0.989 (0.000)	1.009 (0.000)	1.001 (0.878)
cbd_glas_km	0.998 (0.000)	1.003 (0.001)	1.001 (0.030)	0.996 (0.000)	0.997 (0.000)	1.002 (0.000)
dscrptn_charcount	1.006 (0.000)	1.010 (0.000)	1.003 (0.000)	1.003 (0.000)	1.007 (0.000)	1.009 (0.000)
allrooms	0.974 (0.000)	1.072 (0.000)	0.942 (0.000)	0.967 (0.000)	0.983 (0.000)	1.023 (0.002)
flat_all	0.988 (0.203)	1.116 (0.060)	0.960 (0.040)	0.777 (0.000)	1.080 (0.000)	1.094 (0.000)
bung_ALL	1.049 (0.001)	1.661 (0.000)	1.097 (0.002)	1.142 (0.000)	0.938 (0.002)	1.009 (0.747)
detached	0.922 (0.000)	0.763 (0.000)	0.910 (0.001)	0.869 (0.000)	0.979 (0.245)	0.892 (0.000)
terraced	0.958 (0.001)	0.833 (0.014)	1.001 (0.959)	0.827 (0.000)	1.009 (0.623)	0.871 (0.000)
stone	0.851 (0.000)	1.453 (0.000)	0.904 (0.027)	0.925 (0.056)	0.750 (0.000)	0.917 (0.044)
stone_flat	0.876 (0.000)	0.576 (0.000)	0.767 (0.000)	0.672 (0.000)	1.074 (0.049)	0.922 (0.106)
bay	0.931 (0.000)	0.965 (0.478)	0.985 (0.437)	0.941 (0.003)	0.911 (0.000)	0.955 (0.024)
conservy	1.065 (0.001)	1.186 (0.025)	1.188 (0.000)	0.918 (0.030)	1.055 (0.046)	1.253 (0.000)
garage_d	0.988 (0.170)	0.867 (0.001)	1.052 (0.005)	1.063 (0.000)	0.948 (0.000)	1.065 (0.001)
parking	0.901 (0.000)	0.882 (0.035)	1.041 (0.052)	0.887 (0.000)	0.862 (0.000)	0.973 (0.239)
garden_d	0.968 (0.000)	1.069 (0.215)	0.869 (0.000)	1.207 (0.000)	0.947 (0.000)	0.863 (0.000)
Price band 2	1.133 (0.000)	1.417 (0.000)	1.203 (0.000)	1.208 (0.000)	1.074 (0.000)	1.312 (0.000)
Price band 3	1.209 (0.000)	1.771 (0.000)	1.370 (0.000)	1.241 (0.000)	1.118 (0.000)	1.641 (0.000)
Price band 4	1.245 (0.000)	2.314 (0.000)	1.439 (0.000)	1.302 (0.000)	1.117 (0.000)	1.971 (0.000)
Price band 5	1.315 (0.000)	3.836 (0.000)	1.460 (0.000)	1.326 (0.000)	1.160 (0.000)	2.615 (0.000)
Spring	1.017 (0.028)	1.007 (0.857)	1.010 (0.546)	1.029 (0.072)	1.015 (0.159)	0.987 (0.458)
Summer	1.015	0.983	1.025	1.011	1.013	1.009

Independent Variables:	Dependent Variable ‡					
	Pathos (all)	Type I	Type II	Type III	Type IV	Core
	[1]	[2]	[3]	[4]	[5]	[6]
West End	(0.057) 0.976	(0.677) 0.868	(0.124) 0.978	(0.506) 1.046	(0.227) 0.964	(0.569) 0.962
East End	(0.014) 1.073	(0.007) 1.143	(0.288) 1.026	(0.032) 1.155	(0.006) 1.051	(0.063) 1.185
East Dunbartonshire	(0.000) 0.922	(0.055) 0.740	(0.352) 1.009	(0.000) 1.036	(0.003) 0.863	(0.000) 0.922
East Renfrewshire	(0.000) 0.959	(0.000) 0.716	(0.742) 0.952	(0.174) 1.028	(0.000) 0.957	(0.005) 0.841
Intercept	(0.024) 0.011	(0.001) 0.000	(0.219) 0.004	(0.422) 0.003	(0.100) 0.005	(0.000) 0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
<i>n</i>	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.

Conclusion

This article has considered the arguments for and against systematic variation in the language of real estate marketing. We have sought to choose between theories by constructing a series of nested 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 fairly substantial evidence that the verbal construction of house ads varies systematically, both across space and over time. 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-cyclical (falling as the market rises). 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-cyclical. 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 side 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 ad language varies across geographical space and this finding appeared to be independent of deprivation, property type, distance from the 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. However, 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. What we hoped to achieve with this study is a means of linking 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 emotional side is somewhat harder to quantify and remains relatively unexplored for economists. Hopefully, therefore, we have made some headway in establishing a methodological foundation that will encourage future work and facilitate a more rounded understanding of housing markets. In addition to further work on the particular theories offered here, the following stand out as particularly fruitful avenues for future research:

- *Objective classification of words*: while our results were generally robust to the definition of pathos (narrow and broad definitions yielded similar findings), it would be interesting to investigate systematically how house buyers respond to particular words and phrases. Research methods developed in psychology and experimental economics are obvious ways to pursue this, as are survey based approaches.
- *Analysis of Longer Descriptions*: our analysis has focused on a very large number of very short property ads. A complementary approach would be to look at the more detailed property schedules provided by estate agents. These would offer a much richer source of raw data.
- *Withdrawn Properties*: we have only considered properties that were eventually sold, but what of those dwellings that were removed from the market by sellers? Consideration of property descriptions of withdrawn properties would offer a way of controlling for transactions bias and, more interestingly, establish whether the language used actually affects the chances of selling.
- *Endogeneity*: a related point is to consider the price effect of pathos. Disentangling cause and effect is no mean feat (we employed the relatively crude method of including price bands) but it may be possible to use simultaneous equation techniques to establish whether and why more elaborately described properties achieve a higher (or lower) price.
- *Logos*: we have made only passing reference to what is in fact the largest category of words in our sample of property descriptions – words and phrases that appeal to reason – and so a systematic textual analysis of logos is an obvious candidate for further work.
- *Ethos*: one might similarly explore the means by which estate agents persuade on the basis of their claimed integrity, though data sources other than property descriptions would need to be identified.
- *Submarkets*: we have only scratched the surface of how the language of selling varies across space. There maybe considerable scope for investigating the use of language to “frame” properties in particular localities and for deepening our understanding of how agents adopt different marketing strategies in different neighbourhood contexts.

References

- Aristotle, (350 B.C.) *Rhetoric*, <http://www.greektxts.com>.
- BBC News Online (2002) "Estate Agent speak: A dictionary" 11 October, 2002. <http://news.bbc.co.uk>
- Booth, T. (2006) “An Author with an eye for property: The Worst Estate Agent Jokes”, <http://www.tonybooth.info/6.htm>
- 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)

- Elder H.W., Zumpano L.V., & Baryla E.A. (2000) Buyer brokers: Do they make a difference? Their influence on selling price and search duration, *Real Estate Economics* 28 (2) pp. 337-362
- 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.
- 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: A National Archive” <http://www.houseweb.co.uk/house/pros/estate/speak.html>
- Levin, E. and Pryce, G. (2007) A Statistical Explanation for Extreme Bids in the House Market, *Urban Studies*, Volume 44, No.12.
- Levitt, S. and Syverson, C. (2005) *Market Distortions when Agents are Better Informed: The Value of Information in Real Estate*, National Bureau of Economic Research 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.
- 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
- Sirmans C., Turnbull G., and Dombrow, J. (1995) Quick House Sales - Seller Mistake or Luck, *Journal of Housing Economics* 4 (3): pp. 230-243

- 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).
- Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data* (Cambridge and London, MIT Press).

Appendix 1 PATHOS Program

The following Stata program counts the occurrences of words (from a pre-specified list) in the observations of a text variable, such as property descriptions.

```
*-----  
*PATHOS © Gwilym Pryce 2006  
*Stata program written to count the occurrence of particular words.  
*-----  
capture program drop PATHOS  
program PATHOS  
*-----  
version 9.0  
*-----  
syntax [varlist] [if] [in] [, npathos(real 1) occurrences(real 1)  
description(string)]  
*-----  
foreach var of varlist `varlist' {  
    local i = 1 // index for PATHOS  
    local m = `occurrences' -1  
    capture drop `description'  
    gen `description'1 = `description'  
    gen `var'_all = 0  
    while `i' <= `npathos' {  
        local j = 1 // index for occurrence number  
        local x`i' = `var'[`i']  
        display "=====  
        display "**First run index to get first occurrence"  
        display "=====  
        gen `var'`i'_o`j' = index(`description'`j', "`x`i'")  
        recode `var'`i'_o`j' (0=0) (1/max=1)  
        label variable `var'`i'_o`j' "Occurrence `j' of `var' `i': `x`i' "  
        tab `var'`i'_o`j'  
        replace `var'_all = `var'_all + `var'`i'_o`j'  
        forvalues h = 1(1)`m' {  
            local j = `j' +1  
            local k = `j' - 1  
            display "=====  
            display "**Then delete first occurrence and re-run index"  
            display "=====  
            capture drop `description'`j'  
            gen `description'`j' = substr(`description'`k', "`x`i'", "#", 1)  
            gen `var'`i'_o`j'_ = index(`description'`j', "`x`i'")  
            recode `var'`i'_o`j'_ (0=0) (1/max=1)  
            label variable `var'`i'_o`j' "Occurrence `j' of `var' `i': `x`i' "  
            tab `var'`i'_o`j'  
            replace `var'_all = `var'_all + `var'`i'_o`j'  
        }  
        local i = `i' +1  
    }  
} *-----  
end  
*-----
```

Having run the above syntax, we were then able to run the following command:

```
PATHOS [varlist], npathos([n1]) occurrences([n2]) description([s1])
```

where *varlist* is the variable(s) which contains as its observations the words we were interested in, *n1* is the number of words in our word list, *n2* is the upper limit on the number of occurrences we were interested in searching for in any one description, and *s1* is the name of the string variable containing the text descriptions.

For example, suppose we have a list of four ($n1 = 4$) words and/or phrases (e.g. “sought after”, “exclusive”, “fabulous” and “beautiful”) entered as the first four observations of a variable we have called “**Xwords**” and we want to know how many times at least one of these words occurs in each observation of “**Advert**” – a variable in the same dataset as **Xwords** that lists, say, 50,000 property descriptions. Suppose also that we are fairly sure that no one word in **Xwords** will occur more than two or three times in any single property description, but to be certain, we want the program to count up to five occurrences ($n2 = 5$) in any one description. We would then enter the following command,

```
PATHOS Xwords, npathos(4) occurrences(5) description(Advert)
```

One thing to note is that the **Xwords** variable – the list of words we are interested in – must be created in such a way that each of its rows contains only one of the words (or phrases) we are interested in. Spaces can be included after and before each word entered in this list if we want the program to search for whole words (rather than combinations of letters within words). Note that the search procedure is case sensitive, so if we wanted the search process to be blind to whether a word contains capital or lowercase letters (which was indeed the case), we can transform the *varlist* and *s1* variables so that they are in lower case (we used the `gen x2 = lower(x1)` Stata command to achieve this, where *x1* is the original string variable with observations that contain a mixture of upper and lower case letters, and *x2* is the transformed variable which is entirely made up of lower case strings). A useful way to enter the word lists is to use the `input` command. So, in the above example, we might have created the **Xwords** variable by first opening up the dataset which contains the **Adverts** variable, and then entered the following syntax:

```
input str25 Adverts  
"sought after"  
"exclusive"  
"fabulous"  
"beautiful "  
end
```

which creates a new variable in the same dataset as the **Adverts** variable. This new variable has just four rows. Note that we included a space after the fourth word to ensure that the program does not search for longer words like “beautifully”.

Appendix 2: Word Lists

We searched for 181 *pathos* words in total. We also searched for different subsets of this list, including our four subcategories of *pathos* (Types I, II, III and IV) plus the narrower definition which we labelled *Core Pathos Words*:

Pathos Type I: Originality

"architect", "authentic", "bespoke", "character", "commissioned", "crafted", "custom", "designer", "earliest", "hand carved", "hand made", "hand-carved", "hand-made", "imaginative", "individ", "inimitable", "innov", "irreplaceable", "natural", "one of a kind", "one off", "one-of-a-kind", "one-off", "orig", "purp b", "purp-b", "purpose b", "purpose-b", "rare", "specially", "tailor made", "tailor-made", "to order", "to-order", "unique", "unusual".

Pathos Type II: Ambience

"atr", "att", "beaut", "bright", "buzz", "charm", "chic", "del", "eleg", "fashionable", "fresh", "latest", "lovely", "lux", "mature", "picturesque", "pleas", "smart", "striking", "stylish", "sunny", "taste", "trendy", "ambien",

Type III: Prestige

"cachet", "choice", "class", "des", "elite", "enviable", "estab", "excl", "executive", "first class", "first rate", "first-class", "first-rate", "high regard", "high-class", "high-regard", "kudos", "pop", "preferred", "prestig", "prestigious", "prime", "professional", "reputation", "seldom", "select", "sought after", "status", "successful", "up market", "up-market"

Type IV: Excitement

"!", "abs", "actually", "amaz", "appeal", "apprec", "astonish", "beyond", "breathtak'g", "breathtaking", "brill", "b'thtak", "categorically", "clearly", "dazzling", "decep", "definitely", "dream", "enjoy", "esp", "essent", "exceed", "exceptional", "excit", "exq", "extr", "fab", "fant", "flex", "for sure", "generous", "glorious", "good deal of", "great", "h. stand", "handsome", "high", "ideal", "idyl", "immac", "immed", "imp", "incontestably", "increasing", "incred", "indeed", "indisputably", "indubitably", "jaw dropping", "jaw-dropping", "knockout", "knock-out", "mag", "marv", "most", "much", "must", "oft", "o'standing", "outst", "particularly", "perfect", "positively", "quality", "really", "remark", "requested", "sensation", "seriously", "splend", "stun", "sup", "surely", "terrif", "too many features to", "trem", "triumph", "true", "truly", "unbelievable", "undeniably", "undoubtedly", "unquestionably", "vastly", "versatile", "very", "well", "without doubt", "wonderful", "wow".

Core Pathos Words

"preferred", "lovely", "exceptional", "prime", "generous", "outst", "fant", "excl", "beautiful", "charm", "impress", "sought after", "superb", "stun", "del", "magnif", "pleas", "unique", "sunny", "professional", "enviable", "prestig", "splend", "prestigious", "smart", "character", "executive", "eleg".