Street Name Fluency and Housing Prices[♦]

Sumit Agarwal
National University of Singapore
Email: ushakri@yahoo.com

Maggie R. Hu
Chinese University of Hong Kong
Email: maggiehu@cuhk.edu.hk

Adrian D. Lee
Deakin University
Email: Adrian.Lee@deakin.edu.au

Latest draft: October 15, 2019

[♦] We acknowledge funding from the UTS Business School research grant. We thank participants at the University of Technology Sydney, 2016 GCREC conference in Hangzhou, and the 2016 AREUEA International Conference in Alicante, Spain for helpful comments and feedback. We thank Mark Davies for use of the n-grams corpus from his website. Lee thanks Peter Wang for excellent research assistance.

Street Name Fluency and Housing Prices

Abstract

This study investigates whether and how street name fluency affects housing prices using a rich sample of housing sale transactions in Sydney, Australia. We find homes with shorter street names (only one word) are associated with a premium of 11.48% than those with three or more words in the street names, implying fluency preference. Meanwhile, street names with fewer letters are priced with a -0.6% discount, that is, street names with fewer words but more letters (longer words) are preferred. Moreover, homes with unique street names have statistically higher prices of 1.4% (or A$9,481) than those with more common names, suggesting disfluency preference. We argue this is consistent with the consumption context effect as homes are special occasion purchases where exclusivity and uniqueness are desired. In addition, homes with less fluent street names are valued more conditional on the street name is rare or the home is in the luxury price range, further confirming the disfluency premium given exclusivity preference. Our results are robust to a matched sample approach utilizing pairs of similar home sales on different streets. Overall, our findings shed light on understanding how name fluency affects the investment decision of special occasion goods such as real estate.

Key words: *Name fluency, hedonic pricing, real estate, behavioral economics*

JEL Classification Code: *R00, O18, P22, R21*

1. Introduction

A name certainly plays more of a part than we think. It has been documented that names are important considerations in employment opportunities for job seekers and in stock valuation for investors (e.g. Bertrand and Mullainathan (2004); Alter and Oppenheimer (2006); Green and Jame (2013)). It is reasoned that when making complex decisions, people simplify the task by relying on mental shortcuts (Tversky and Kahneman, 1973). One input shown to be influential in the decision-making process is fluency, or the ease with which people process information.

Psychology research has established that fluency has an impact on judgment that is independent of the content of the information. It is found that people perceive more fluent stimuli as more appealing (e.g. Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka and Simons (1991); Schwarz (2004)). For example, people tend to associate fluent sounding names with truth and disfluency with untruth (Schwarz (2004)), in large part because fluency implies frequency and familiarity, which in turn implies social consensus (Schwarz, et al. (2007)).

While there are extensive studies on the impact of physical features or physical attractiveness (e.g. Glaeser, Kincaid and Naik (2018)) on home prices, there has been no research to date on whether the street names of a home matter. The street name usually forms part of the address of the home, and is frequently referred to in identifying the specific residential building. Street names of a home can seem random or arbitrary as the street naming process is not an exact science. For example, In the United States or Australia, most streets are named after numbers, landscapes, trees (a combination of trees and landscapes such as "Oakhill" is used often in residential areas), or the surname of an important individual (in some instances, it is just a commonly held surname such as Smith). Streets are named in a myriad of ways, with little control over whether that name will add value to a property or not. Given home value is determined by features that appeal to homebuyers, the street name associated with a home is potentially a priced factor and making addresses not just some label or reference for a property, but influential factors in housing price valuation.

In this paper, we fill this gap in literature by examining whether and how street name fluency affects home price[1]. As name fluency is a multi-facet concept, to capture different dimensions of name appeal, we adopt six measures in total based on the following six aspect of a street name: 1) how Englishness a street name is; 2) number of words in the street name, 3) whether the name passes MS Word spell check, 4) how common a street name appear in other suburbs, 5) the number of syllables in a street name, and 6) the number of letters in a street name.

[1] We use street name as opposed to building or estate name as it is more ubiquitous and applies to every home.

Consistent with prior literature, we find that street name does play an influential role in home valuation. Specifically, in terms of word count, we find that homes with a single-word street names have about 11.8% higher prices than homes on streets named with two words or more, consistent with fluency preferences shown in Schwarz, Bless, Strack, Klumpp, Rittenauer-Schatka and Simons (1991), Schwarz (2004) and Green and Jame (2013). The magnitude is economically significant and comparable to other studies on fluency and asset pricing[2]. We also look at the number of letters in a street name, and find that street names with more letters are priced with a 0.6% premium, comparing least fluent letter group (street names with 1-6 letters) to least fluent letter group (street names with 8 or more letters). Taking together the findings on word count and letter count, we find homebuyers prefer homes where street names contain fewer words but more letters (longer words) in the street names. For example, our findings imply that the name Rosebridge is preferred over Wonga Wonga (same number of letters but more words).

Besides a fluency premium in terms of word count and letter count of street names, we also look at other dimensions of name fluency, such as how common or popular the street name shows up in other neighborhoods. Contrary to the commonness or familiarity preference, our result reveals that homes on streets with common names are priced lower, suggesting uniqueness preference. Specifically, unique street names (i.e. a street name used in only one neighborhood in Australia) are associated with 1.4% or A$9,481 higher prices than homes with common names (in six or more neighborhoods). We control for housing characteristics with street type, neighborhood, and time fixed effects. The results are also robust to using a matched sample approach of unique pairs of homes that are similar in characteristics, geographically proximate, and sold at similar times but on different streets.

Overall, we find three pieces of evidence on fluency preferences: uniqueness of street names carries an 11.8% positive premium, whereas street names that contain fewer but longer words on average are preferred. To interpret these findings, we explore whether fluency preferences varies with consumption contexts (Pocheptsova, Labroo and Dhar (2010)). Although fluency could improve a product's attractiveness to potential customers in the domain of everyday goods, Pocheptsova, Labroo and Dhar (2010) find that in the context of special occasion high-end goods, higher fluency is a negative cue because it makes the products feel less special, and hence less valuable, whereas the

[2] For example, Alter and Oppenheimer (2006) compare IPO return difference between stocks with the most fluent company name (proxied by pronounceability) and those with the least fluent name , and find a difference of 11.2%. Hence, result on fluency premium is of similar magnitude compared with differences in IPO returns. Further, Green and Jame (2013) use a five point scale to measure company name fluency and find a difference between 7.6% and 10.12% in firm value (proxied by market-to-book and Tobin's Q, respectively) when comparing companies with the most fluent names to the least fluent names.

lower fluency of a product name increases its uniqueness and makes the prodcut appear more exlcusive and desirable.

So disfluency may be preferred when choosing products that are uncommon or 'special-occasion' (i.e. occur only once or several times in a lifetime), reflecting the exclusivity of the product. As real estate properties are a large-ticket special-occasion purchase that also requires complex decision making, it explains why unique street names that appear less commonly across other suburbs could be a preferred feature by homebuyers and hence carry a price premium.

Further, we investigate whether disfluency premium is especially higher when the home is on a rare-named street or the home is a high-end property (to denote exclusivity). Consistent with a disfluency preference when the home is more exclusive, we find that homes with less fluent street names have higher prices, conditional on the home's street name being uncommon or the home being in the luxury price range. Our findings suggest that less fluent street names have higher prices due to a preference for uniqueness, particularly when conditions of exclusivity are met.

Next, we explore whether certain buyer attributes and housing features could affect the name fluency preference. It is possible that native English speakers have difference preference of name fluency from those buyers who only speak English as a second language. Thus, the preference for fluent street names would be stronger for homebuyers who speak English as a second language, as those street names are easier to pronounce and remember for them. We find this group of buyers indeed prefer higher fluency in terms of fewer letters in the street names, and are willing to pay 0.6% higher prices for this attribute.

This analysis on buyer language background is also related to prior work on the effect of culture and superstition on housing prices. For example, Deng, Hu and Lee (2019) find that homebuyers prefer culturally proximate neighborhood. Earlier studies find that lucky street numbers or floor numbers, like the number 8 for Chinese buyers, command higher prices (e.g. Chau, Ma and Ho (2001); Choy, Mak and Ho (2007); Fortin, Hill and Huang (2014); Agarwal, He, Liu, Png, Sing and Wong (2014)). The above studies relate to a specific Chinese cultural phenomenon, whereas it is unclear if name fluency requires English language competency. In our study, we test whether Asian buyers (whose first language is less likely to be English) differ to other buyers with respect to the preference on street name fluency and housing prices and find corroborating result.

In terms of property features, we also examine whether there is greater fluency premium for new homes. We hypothesize that street name fluency could play a more important role for new homes as the features of new homes are less known and more uncertain to potential buyers. In response, buyers formulate their price on other more salient features such as street name fluency. Indeed, we find new homes have especially higher prices of between 1.3% and 5.4% when the street names are more fluent in terms of having fewer syllables or fewer letters in the street name.

As an additional test, we examine whether homebuyers have a preference for royal names, and how royal names affect fluency premium. Association with royal names is usually perceived as higher social status and better recognition. Consistent with this notion, we find streets named after royalty are priced about 3% higher. Further, this royalty preference is stronger when the house is a luxury property. We then test whether buyers still care about street name fluency if the home has royal street names as a positive feature. We find that the fluency preference is reduced with royal names. Instead, unique names and names with shorter syllables are more favored for homes with royal words in the street names.

Besides royal names, words related to popular celebrities or trendy terms may be preferred and buyers could attach a premium for houses on streets with those trendy names. To measure how trendy a name is, we utilize Google Trends Search to create a popularity index based on search volume within our sample period. We find that homes on popular street names based on Google trend search are transacted at higher prices of 0.3%. We also find that the Google trend popularity has a separate price effect in addition to the fluency measures.

Next, we conduct further analysis to check the robustness of our result in the following. First, it is possible that certain latent features of a housing unit or the neighborhood could have a large impact on housing price, which may pose an endogeneity concern and hence bias our estimation. To address this, we utilize a matching approach and compare similar homes on different streets that are near each other and in the same suburb. In doing so, we ensure the homes are of similar quality and share the same amenity, so that we can focus on the differences between one-to-one matched homes on different streets. We start with matching each individual home to a list of candidate homes with the constraint that they are within the same suburb, of the same housing type (house or apartment), on different streets, within 100 meters (0.062 miles) of each other, and sold within 365 days apart, and retain only the most similar candidate home as control. Consistent results are obtained using the matched home sample for our baseline fluency group and categorical fluency tercile group regressions.

Second, we employ a special subsample of homes with multiple street names to investigate the importance of street name fluency. Occasionally a home may locate at the intersection of two or more streets. For this case, we look at the fluency scores of each of the street names and compute the maximum and minimum values. We find that the housing price is lower if any of the streets has very high English sounding word in the name. Moreover, having more popular name and fewer letters in the street name will increase the price.

Third, it is possible that street name fluency may be related to how central a street is. Anecdotal evidence suggests that major thoroughfares such as General Holmes Drive, Georges River Road tend to have more words in the street name. To more thoroughly capture the effect of street centrality in addition to our Long Street and Major Street measures, we construct a street centrality measure based on the geospatial data of all streets in Sydney and include it as a control variable in our baseline hedonic housing price regressions. The results are qualitatively similar to the baseline results with fluency measures being of similar magnitude and statistical significance. Street Centrality is negative and statistically significant of -0.012 across regressions, which implies that a one standard deviation increase in the street centrality measure reduces the price of the home by 1.2 percent.

We further investigate whether repeat home buyers tend to buy homes with fluent street names and also pay a premium for these homes. To study the buyer persistence effects, we identify a sample of homebuyers that have multiple home purchase records in the database. We then test whether these buyers show persistence in buying more fluent homes and in paying higher prices for homes with more fluent street names (fluency premium) based on whether the buyer's first purchase was with a fluent street name or had a higher fluency premium. We find evidence of both. Specifically, if a home buyer's first purchase was a fluent street name home or they paid a high fluency premium, their subsequent purchases will also be of a fluent street name home or have a high fluency premium. Our findings are consistent with a conspicuous consumption effect of buyers for fluent street name homes.

Our study on street name fluency and housing prices makes several key implications. First, we document the preferences for fluency regarding one of the largest investment decisions that a person makes in their lives, i.e., home purchase. Housing decisions can have substantial long-term consequences for household wealth accumulation, and almost two thirds of median U.S. household wealth is in housing wealth (e.g. Keys, Pope and Pope (2016)).[3] Compared to most psychology studies testing for fluency preferences, we offer evidence on the effect of name fluency for an important real-life decision where subconscious behavior should be less prevalent. We show that the street name of a home can determine whether that home is desirable, making addresses not just random labels for a property, but influential factors that could make or break a sale.

Second, an examination of fluency has implications concerning the marketing of homes. Norris (1999) finds that advertisements for developments in Rochester, New York tend to choose pleasant sounding names. While the effect of marketing tends to last for the length of the sales campaign, we document that the street name has a lasting effect on housing value. As names of locations are typically chosen due to non-financial reasons such as topography, cultural, or historical factors

[3] In Australia, housing wealth comprises 42% of household net wealth and is the largest component by far (Australian Bureau of Statistics (2019)).

(Cardoso and Meijers (2017)), we argue that urban planners also need to consider the economic consequence of naming decisions in the long run.

Moreover, as this study investigates real estate sales transactions, we are better able to investigate a larger cross-section of names. Prior studies in name fluency and asset prices find that fluent stock names tend to have better recognition and higher valuations by investors. Alter and Oppenheimer (2006) find that American Stock Exchange (AMEX) and New York Stock Exchange (NYSE) stocks that are easier to pronounce in English earn higher stock returns than those with unpronounceable names. Green and Jame (2013) find that U.S. stocks exhibiting increased fluency (e.g. less words, using common words) experience higher breadth of ownership, greater share turnover, and higher valuation ratios. They reason that fluent names are mentally easier to process and therefore more investors have an affinity towards them. In terms of sample comprehensiveness, our sample consists of 15,153 unique street names, whereas Alter and Oppenheimer (2004) use 781 (665 NYSE and 116 AMEX stocks) and Green and Jame (2013) use 4,600. We are also better able to compare similar housing investments with different name fluency, as homes may be geographically proximate and situated on different street names; a similar like-for-like comparison is difficult with stocks.

The rest of paper proceeds as follows. Section 2 describes data and method. Section 3 presents empirical results from baseline models. Section 4 presents further heterogeneity analysis. Section 5 conducts some robustness analysis and Section 6 concludes.

2. Data and Method

2.1 Data

We employ a large dataset with 958,408 individual housing sales transactions in the Sydney Metropolitan Area from January 2000 to June 2016, sourced from Australian Property Monitors (APM)[4], as our principal data source for street names and housing price. This dataset covers a comprehensive list of variables on property and sales, including the transaction price, transaction date, detailed property address (including street names and unit number), buyer and seller names, whether the transaction is an auction sale, and other housing characteristics. Appendix 1 provides a list of housing characteristics variables used in our hedonic housing price regression.

In order to measure how central a street is in a local neighborhood, we obtain the longitude and latitude of each property based on the home address from the Public Sector Mapping Agency

[4]APM is one of Australia's leading national suppliers of online property price information to banks, financial markets, professional real estate agents and consumers. See www.apm.com.au for further details.

Australia's (PSMA) Geocoded National Address File (G-NAF)[5] from the Australian government's data.gov.au website. The G-NAF database contains geocodes of exact addresses across Australia. According to the G-NAF product website, it is the most trusted source of geocoded addresses for Australian business and governments. Sales prices and land area sizes are winsorized at the 1st and 99th percentile to remove outliers.

2.2 Measures of Street Name Fluency

Street names are first separated from their street type (e.g. highway, road or street)[6] and any apostrophes removed (e.g. O'Dea becomes Odea) to calculate the measures. We adopt six measures of fluency defined below.

1) ***Englishness Group*** – measures how often a combination of letters appears in English media. We adopt an approach similar to Green and Jame (2013) to measure *Englishness* except with a modification for letter position. The occurrence rate of a specific word (using the Markov chain rule) without considering letter positioning is given by:

$$Pr(\#,l_1,l_2,l_3,l_4,...,l_n,\#) = Pr(\#)*Pr(l_1|\#)*Pr(l_2|\#,l_1)*...*Pr(l_n|l_{n-2},l_{n-1})*Pr(\#|L_{n-1},L_N) \quad \text{--- (1)}$$

Where l_n denotes letter l in position n of the word and '#' denotes a space at the beginning and at the end of a word. We estimate $Pr(l_k|l_{k-2},l_{k-1})$ as $F(l_{k-2},l_{k-1},l_k)/F(l_{k-2},l_{k-1})$, where $F(l_{k-2},l_{k-1},l_k)$ is the frequency count of the trigram l_{k-2},l_{k-1},l_k (bigram l_{k-2},l_{k-1}). We source our frequencies of bigrams and trigrams from word frequencies in Mark Davies' n-grams corpus of Historical American English for the 2000-2010 decade[7]. The database contains unique words with their frequency from the Corpus of Contemporary American English. The higher the word occurrence rate, the more commonly that combination of letters is found in English media.

For conciseness, we log transform the probabilities into a score as below:

$$E'(Pr(\#,l_1,l_2,l_3,l_4,...,l_n,\#)) = ln[F(\#,l_1)/F(\#)] + log[F(\#,l_1,l_2)/F(\#,l_1)] + log[l_1,l_2,l_3)/F(l_1,l_2) \quad \text{--- (2)}$$
$$+ ... + log[F(l_{n-1},l_n,\#)/F(l_{n-1},l_n)]$$

Where $F(\#)$ is the total frequency of all words in the corpus. One potential problem in the construction of the above score is that it does not consider the position of letters. For example, the

[5] The link is: https://data.gov.au/dataset/geocoded-national-address-file-g-naf. G-NAF database website is https://psma.com.au/product/gnaf.
[6] We follow official G-NAF address records which separate street name from street type when calculating fluency scores. For example, the street 'Avenue of Oceania' and other streets where the street type is in front of the name are recorded as having no street type. Therefore, the street type is considered part of the name. Similarly some street types may be part of the name. For example, for Highland Ridge Road, Middle Cove, the street name in G-NAF is 'Highland Ridge' even though 'ridge' is a street type.
[7] Available here: http://www.ngrams.info/download_coha.asp

letters 'THE' are more commonly found as the first three letters in an English word than as the last three letters. However, it is ignored in the above measure. To address this, in the spirit of Hamed and Zesch (2015), we use a modified probability score in which three scores are added together based on the prefix (first three letters), middle of the word (excluding start and end characters) and suffix (last three letters). This frequency count is position-specific. For example, for the name 'coogee', the score calculation is:

$$E^{*}(c,o,o,g,e,e) = \log(\Pr(c,o,o|prefix)) + E^{'}(o,o,g,e|middle) + \log(\Pr(g,e,e|suffix)) \quad \text{--- (3)}$$

Where *prefix*, *middle* and *suffix* relate to the frequency of trigrams at the prefix, the bigrams and trigrams in the middle of a word, and suffix of words (excluding start and end characters), respectively. Furthermore, as the scores are positively correlated to the number of letters in the word (by construction of the probability score), we regress the score E^* on the number of letters in the word and use the residual as our *Englishness* score, following Green and Jame (2013).

We measure *Englishness Group* by sorting street names into three equal groups based on the position-specific word probability score, with a measure of 1 if the street name is in the bottom group, 2 for the middle group, and 3 for the highest group. A street name in the lowest group means that this combination of letters appears less frequently in the English language and therefore has a low fluency score.

2) ***Words Group*** – Counts the number of words in the street name. We identify three groups: *Words Group* 3 for street names with one word only; *Words Group* 2 for street names with two words; and *Words Group* 1 for street names with three or more words. We adopt this definition in order to be consistent across measure that group 3 is the most fluent.

3) ***MS Word*** – A Microsoft Word spell check. This measure takes the value of 1 if all the words in the street name in lower case pass the Microsoft Word spell check, and 0 otherwise.

4) ***CommonName Group*** – is defined as the number of suburbs (neighborhoods) in Australia that share the same street name (regardless of street type), which measures how common a street name is. This measure takes a value of 1 if there is only one suburb with the street name, a value of 2 if there are two to five suburbs, and a value of 3 if there are six or more suburbs.[8]

5) ***Syllable Group*** – Counts the number of syllables in a street name. To do this we first make use of the word list from the Carnegie Mellon University (CMU) Sphinx website[9] which contains syllable counts of 134,000 words. For words not contained in the CMU Sphinx, we then hired two researchers

[8] In unreported results, we also use an alternative measure for only suburbs within Sydney and find qualitatively similar results.
[9] See http://www.speech.cs.cmu.edu/cgi-bin/cmudict

to manually count the syllables independently. Where there were discrepancies in the counts submitted between the two researchers, we then looked for other web sources to double check and validate the syllable counts. For words with discrepancies in the counts with no web sources and disagreement between the counters, we used the lower syllable count. We identify three syllable groups: syllable group 3 if the street name has one syllables; group 2 if the street name contains two to three syllables; and group 1 for street names with four or more syllable.

6) ***Letters Group*** – Counts the number of letters in a street name. Street names are ranked in three equal groups based on the number of letters in the street name. A value of 1 is given for the group with the most number of letters. A value of 2 is given for the middle group and a value 3 for the group of street names with the least number of letters.

The three measures, *Englishness Group, Words Group,* and *MS Word*, are similar to those used in Green and Jame (2013). *CommonName Group, Syllable Group* and *Letters Group* are extensions of word fluency. For example, Oppenheimer (2006) suggests that longer words are less fluent than shorter words. All fluency variables are described in ascending groups in order of perceived fluency; therefore the higher the measure, the more fluent is the street name.

There are 15,153 unique street names in our final sample. Appendix 2 shows some examples of street names and fluency measures. Streets with the lowest fluency scores across all measures tend to have several words in their name, have letter combinations not common in English, and are not commonly used as street names in other suburbs. For example, 'Avenue of Oceania' has a low score as it has three words. It fails to pass the Microsoft Word spell check in lower case completely and is ranked in the bottom trecile for syllables in a street name. Medium fluency street names tend to have more commonly expressed letter combinations in English, with one to two words in the name which are not used commonly as street names in Sydney. For example 'Charlie' ranks high for *Englishness Group* as it is a common name, but has a low overall ranking as its *Syllable Group* and *Letters Group* are not high. High fluency names usually contain common words used in English that are short and frequently used as street names. For example 'Cook'[10] and 'Spring' are common English words that are short and tend to be used as street names in Sydney.

Further, Appendix 3 shows the top 20 street names by sales in our sample. These names represent about 6% of total sales. Pacific is the most common street name (6,667) as the Pacific Highway is the longest street in NSW. Other street names such as Pittwater (2,522), Princes (2,189), Liverpool (2,173), Forest (2,132) and Anzac (2,074) are common due to being major thoroughfares.

[10] Although Cook may appear to be a very common street name thanks to Captain James Cook, only 1,741 or less than 0.2 percent of observations have the street name of 'Cook', 'Captain Cook', 'James Cook' or 'James Cook Island'.

Common English names (and also English monarch names) such as Victoria (6,109), George (3,005), William (2,876), Albert (2,439) and Elizabeth (2,059) are also popular. Finally common words that may denote the locality such as Park (4,209), Railway (3,309), Station (2,600), Bridge (2,211) and Church (2,201) round out the top 20.

Table 1 reports summary statistics for our 958,408 individual housing sales observations. These sales are spread across the 645 different suburbs in Sydney over the entire sample period from 2000 to 2016. Panel A reports various statistics on our housing related variables. The mean price is A$677,190 with 57 percent being free standing houses. Mean area size for free standing houses is 4,140 square feet. Homes on average have 2.89 bedrooms and 1.60 bathrooms with 75 percent also including a parking space. Five percent are newly developed homes, while 18 percent are sold at auction. 24 percent of homes sales are on a street over 1 kilometer in length while 7 percent are situated on a major street in a suburb[11].

[--- INSERT TABLE 1 ABOUT HERE ---]

Pertaining to the street name fluency measures, the mean *Englishness Group*, *Syllable Group*, and *Letters Group* are 2.20, 2.08, and 2.01, respectively. This means the average home sold has a street name that is in the middle fluency group for similarity to english words, the number of syllables, and the number of letters. 30 percent of housing sales have street names with all words passing the *MS Word* spell check. Most housing sales street names consist of one word as evidenced by the mean *Words Group* of 2.93, where group 3 denotes street names with one word. Similarly, most housing sales are in *CommonName Group 3* (street names that are used in six or more suburbs in Australia). This suggests a propensity towards fluent street names.

To compare housing characteristics against fluency measures, we calculate an aggregate fluency score as the sum of all six measures with a score of 5 being the least fluent (lowest scores across all fluency measures) and 16 the highest. Table 1 Panel B reports mean housing characteristics by the aggregated fluency score. Sales for the least fluent street name group (aggregate fluency score is between 5 and 6) make up the smallest group (less than one percent of the sample), while sales in medium fluency groups (aggregate fluency score is from 11 to 14) make up over half of house sales. Homes with low fluency street names (aggregate fluency score less than 10) tend to have higher housing prices than homes with high fluency street names. However, the higher prices appear to also reflect better housing characteristics, as homes with low fluency street names tend to be free standing houses (73 percent), exhibit larger area size, more number of bedrooms and bathrooms, and include

[11] Long Street is a dummy of 1 if the street on which the home is situated is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise. Major Street is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise.

parking compared to homes with high fluency street names. New homes account for four to five percent of the sample across fluency groups. Homes with street names of high fluency have more auction sales (19 percent) than those of low fluency (12 percent). No consistent relationship is found between fluency and whether the home is on a long or major street.

To visualise the relationship between prices and fluency, in Figure 1 we plot the mean housing prices[12] for each of the six fluency groups, with 95% confidence intervals. Similar to the aggregate fluency scores in Table 1, we find a monotonic declining relationship between housing prices and all six fluency measures. This suggests that homes tend to be of lower price on streets with more fluent names.

[--- INSERT FIGURE 1 ABOUT HERE ---]

As further analysis of housing price and fluency measures, we calculate their correlation in Table 1 Panel C. We find a very small negative to zero correlation of housing prices to fluency measures, similar to our findings in Table 1 Panel B. Fluency measures tend to be positively correlated with one another except for *Words Group* to *Englishness Group* and to *MS Word* where the correlation is negative. The positive (0.16) correlation between Words group and CommonName group suggests that street names with more words tend to utilize commonly used English words.

Table 1 Panel D reports frequency counts and percentages across fluency groups. The majority of sales are on street names with high fluency in terms of Englishness Group (44.09% of sales are in high fluency group), Words Group (93.71%) and CommonName Group (80.59%). Most sales are not on streets that pass MS Word (70.35%). In terms of syllables most sales are in the medium fluency group and for the number of letters it is roughly evenly divided.

We report the average fluency scores for the top 20 suburbs by sales in Table 1 Panel E. The top 20 suburbs represent about 15% of all sales across the 645 suburbs. The average fluency scores across the top 20 suburbs are comparable to the entire sample. The average fluency scores do not seem to vary a lot across suburbs with the exception of *Englishness* and *MS Word*. Englishness is lowest in Maroubra (1.94) and highest in Parramatta (2.58), while *MS Word* is lowest in Cronulla (0.15) and highest in Chatswood (0.49). Further, Figure 2 visually illustrates the mean fluency score of street names in all suburbs across Sydney using heat maps, whereby greener shades represent higher fluency scores, and browner shades correspond to lower fluency scores. The results for six different flucency measures are presented in Panel A to F. For example, Panel B on Words Group shows that most of the suburbs are of green color, which implies street names in most of the subursb

[12] We find qualitatively similar results using median housing prices.

are in Words Group 3 with only one word in the name, excluding the word "street" itself, and few suburbs have street names with two or more words.

Figure 2 Panel C on Microsoft Word spell check shows that most of the suburbs are brown or light brown, i.e., most of the suburbs have less than 45% of street names show up as errors in MS Word spell check. This suggests that most words used in street names are standard English words that pass the MS Word Spell check.

The pattern presented in Figure 2 Panel D on street name popularity is more balanced than the other measures as the green and brown shades are of similar size. We can see that suburbs in Inner Sydney, Estern Suburbs and North Shore tend to have more popular street names, whereas street names in areas such as Campbelltown, Fairfield and Liverpool are less common.

[--- INSERT FIGURE 2 ABOUT HERE ---]

2.3 Hedonic Housing Price Model using Fluency Group

We run the following hedonic housing price model across the full sample of individual housing transactions to test whether homes with more fluent street names have higher transaction prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k fluency_{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (4)}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices at sale *i* on name of street *j* in suburb *s* at time *t*;

$fluency_{ij}$ denotes one of the six street name fluency measures i for a home sold on street name *j*;

$property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, land area size, street type (e.g. street, road, highway, etc.) and other features;

$longstreet_{ij}$ is a dummy of 1 if the length of the street in the zip code[13] is greater than 1km (0.62 miles), 0 otherwise;

[13] We use zip codes instead of neighbourhood names from the G-NAF database, as we could more accurately merge our sales data by street and zip code than with street and neighbourhood name to G-NAF.

majorstreet$_{ij}$ is a dummy of 1 if the street in the zip code is in the top two longest streets in the zip code, 0 otherwise;

μ_s are the suburb location specific fixed effects;

γ_t are year/quarter fixed effects;

τ_t is a monthly time trend.

The variables *longstreet* and *majorstreet* are used to control for the fact that long streets are usually major thoroughfares and therefore homes on these streets tend to be sold for lower prices due to traffic externalities such as noise and local pollution (e.g Ossokina and Verweij (2015)). These streets typically also have more fluent sounding names. This is particularly true for our popular street name measure *CommonName Group*. For example, in the Sydney Central Business District, a major thoroughfare is George Street which is a street name commonly used in other suburbs as well and also a fluent name according to other fluency measures.

If the hypothesis that homebuyers are wiling to pay more for homes with more fluent street names, we expect the coefficient for $fluency_{ijt}$ to be positive and statistically significant, controlling for other characteristics of a home sale.

2.4 Hedonic Model with Categorical Fluency Measures

We use an alternative test for street name fluency by treating the fluency groups as categorical dummies rather than a continuous measure. We do this as it is unclear whether street name fluency has a linear relationship with housing prices. For example, if the housing price premium only exists for homes with high fluency street names and no effect for low or medium fluency names, then using continuous fluency measures would not capture the effect. To measure non-linearity of our fluency measures (except for *MS Word* as it is a dummy variable), we use the following model with categorical dummies for the fluency measures to redo the hedonic housing price model:

$$\ln(P_{ijst}) = \alpha_t + \beta_1 D(fluency_{ij} = 3) + \beta_2 D(fluency_{ij} = 2) + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (5)}$$

Where *D(fluency*$_{ij}$*=3)* is a dummy of 1 if the street name of the sold home belonged in the highest fluency group in the *Englishness*, *Words*, *CommonName*, *Syllable* or *Letters Groups*, zero otherwise. And *D(fluency*$_{ij}$*=2)* denotes the middle group (the omitted dummy being the lowest street name fluency group as the base case for comparison). For example, for *Englishness Group*, we use dummy variables for *Englishness Group=3* (low Englishness score) and *Englishness Group=2* and test whether the coefficients of the dummies are statistically different from the omitted dummy *English Group*=1. All other variables are the same as the baseline hedonic model.

2.5 Matched Home Analysis

It is possible that certain latent attributes of a housing unit could affect housing price and hence bias our regression result. For example, Beach St, Coogee may have homes of higher value not because the name 'beach' is more fluent but because the street is near Coogee Beach which is a favored amenity. Therefore a name could correlate with unobserved features, which may pose an endogeneity concern.

In order to control for such unobserved amenities, we compare similar homes on different streets that are near each other and in the same suburb. In doing so, we ensure the homes are of similar quality and share the same amenity, so that we can focus on the differences between name fluency for one-to-one matched homes on different streets.

The matching algorithm we use is adapted from Huang and Stoll (1996) and Davies and Kim (2009) and similar to a greedy algorithm (e.g. Rosenbaum (1989)) with constraints on matched homes). It is able to accommodate matching by geographic distance and time of sale. The details are as follows.

First we match individual homes to a list of candidate homes with the constraint that they are within the same suburb, of the same housing type (house or apartment), and most importantly, on different streets, within 100 meters (0.062 miles) of each other, and sold within 365 days apart. Duplicate pairs are then removed. For each pair of homes, we calculate the following score to capture the differences in sale time, physical distance, and housing features:

$$S_{hps} = salesdays_{hps}^2/v\_salesdays_{hs} + dist_{hps}^2/v\_dist_{hs} + \Delta bed_{hps}^2/v\_bed_{hs} + \Delta bath_{hps}^2/v\_bath_{hs} + \Delta areasize_{hps}^2/v\_areasize_{hs} \quad \text{--- (6)}$$

Where subscripts *h*, *p* and *s* denote property type (house or apartment), matched pair, and suburb, respectively. The variable *salesdays* is the number of days between the sales dates of the two home sales in the pair; *dist* is the geographic distance between the sales pair calculated using longitude and latitude; Δ*bed* is the difference in the number of bedrooms of the sales pair; Δ*bath* is the difference in the number of bathrooms of the sales pair; and Δ*areasize* is the difference in area size of the pair. *v_salesdays*, *v_dist*, *v_bed*, *v_bath and v_areasize* are the variances at the property type and suburb level across paired matches for Δ*salesdays*, Δ*dist*, Δ*bed*, Δ*bath* and Δ*areasize*, respectively. If there is no variability in the variance, we add 0.01 to the variance so that the denominatior is non-zero, and a score may be formed.

If a home is in one or more pairs, we sort by this score and take the pair with the lowest score (with a random number sort to break ties). Remaining pairs are discarded. Our sorting procedure ensures that all pairs are unique and that we only select the closest matching homes. We retain this smaller sample and employ it for our baseline fluency group and categorical fluency dummy regressions.

Our matching algorithm differs to a greedy algorithm (Rosenbaum (1989)) as there are no explicit treatment and control groups. Instead we are simply attempting to match similar proximate homes on different streets. We apply constraints to our algorithm as it is known that greedy algorithms make sub-optimal matches. The constraints therefore ensure matched pairs are not too disimilar from each other.

3. Empirical Results

3.1 Hedonic Model Results

In this section we report our results using the hedonic regression model with fluency group measures. As identified in our summary statistics, homes on streets with low name fluency tend to exhibit higher sales price and size. This means it is important to control for these factors to test the relationship between sales price and street name fluency in our regression. Table 2: Hedonic Regression with Street Name Fluency

reports the coefficient estimates of the hedonic regression. For each regression from columns 1 to 6 we separately test each fluency measure. In the last column we test all the measures together. Consistent with larger homes being sold for higher prices, the results show that homes that are new, sold at auction, of larger size with more bedrooms and bathrooms, and include parking have higher prices. In addition, homes on long streets or major streets in a suburb exhibit lower prices, potentilaly due to noise or pollution.

[--- INSERT TABLE 2 ABOUT HERE ---]

We find that fluency measures are either statistically insignificant or negatively related to housing prices, contrary to the hypothesis that homes with more fluent street names have higher prices. *Englishness Group, Words Group, MS Word* and *Syllable Group* are statistically insignificant. We find that *CommonName Group* and *Letters Group* are negative and statistically significant, suggesting that streets with popular names and with few letters have lower prices. For example, the coefficient estimate for *CommonName Group* of -0.007 implies that a highly popular street name has a 0.7

percent lower price than street names in the mid *CommonName Group,* or 1.4 percent lower than the low *CommonName Group* (i.e. a unique street name in Australia), controlling for other housing characteristics. Simarly, streets with high *Letters Group* (street names with six or fewer letters) have 0.3 percent lower prices than street names in the mid *Letters Group*. Our results suggest that there is an economically significant difference in housing prices based on street name fluency. For example, given the average housing price in our sample is A$677,190 (Table 1), highly popular street names exhibit a 1.4% lower price (or A$9,481) compared to the least popular street names.

3.2 Non-Linearity of Fluency Measures

In this section we test whether there is non-linearity in our fluency measures. In our baseline results we find that four out of six fluency variables are statistically insignificant. However, this may be due to preferences being priced for very high (or very low) fluency measures only. For example, fluency preferences may occur for street names with very high fluency compared with low fluency, but not for middle fluency compared with low fluency. To resolve this, we use categorial fluency tercile group dummies in Table 2: Hedonic Regression with Street Name Fluency

Panel B. We find some evidence of non-linearity. Although *Words Group* is statistically insignificant as a continous variable in our baseline results, using categorical variables we find both the mid and high fluency groups (single or two word street names) to have statistically higher prices than the low fluency group (street names with three or more words), suggesting that shorter naes with fewer words are preferred. Coefficients using categorical variables *CommonName Group* and *Letters Group* are consistent with our baseline results.

For *Words Group*, we find mid (*Words Group = 2)* and high (*Words Group = 3)* measures to be positive and statistically significant at 0.118 and 0.136, respectively. This suggests that homes on streets with one word names are 11.8 percent higher in price than homes with street names of three or more words. Similarly, street names with two words are 13.6 percent higher in price than homes with streets names of three words or more. These value differences are comparable to Alter and Oppenheimer (2006) for IPO first day return differences and Green and Jame (2013) for firm value differences. Alter and Oppenheimer (2006) find there is a difference of 11.2% in first-day IPO return between stocks with the most fluent company name (proxied by pronounceability) and the least fluent. A difference between 7.6% and 10.12% in firm value is also found in Green and Jame (2013) when comparing the most fluent company names to the least fluent.

For *CommonName Group,* we find the relationship is roughly linear. Compared with the low group, the mid group exhibits 0.4 percent lower prices (although statistically insignificant) and the high group exhibits 1.4 percent lower prices (statistically significant at five percent level). Similar to the baseline results for *CommonName Group*, we find homes on more unique street names have higher prices.

For *Letters Group*, homes in the mid Letters Group are 0.1 percent lower in price (although not statistically significant) and the high group is 0.6 percent lower in price (statistically significant at the five percent level) than the low group. As in our baseline results, we do not find *Englishness Group* or *Syllable Group* to be statistically significant. Our findings show that the inclusion of longer words in a street name is related to lower prices, consistent with fluency preferences,s whereas our results for street name popularity and letters show that buyers prefer more unique street names and names with more letters.

3.3 Matched Home Analysis

In this section we conduct a matched home analysis to compare housing prices of geographically proximate homes with similar housing characteristics but on different streets, which provides a clean setting and enables us to test the effect of street names on housing price. This serves as a robustness check of our baseline results in previous section. The matching algorithm is described in detail in Section 2.4. Note that the sample of 488,784 (244,392 pairs) matched homes is about half of the full sample.

Table 3 reports the multivariate regression results using the baseline model in Panel A and the categorical fluency measures model in Panel B. The results are consistent with the baseline full sample regression results. *CommonName Group* is negative and statistically significant (at the 5 percent level) using the baseline model at -0.4% and negative at -0.6% (and just beyond 10 percent statistical significance) for *CommonName Group =3* for the categorical dummy regression.

[--- INSERT TABLE 3 ABOUT HERE ---]

Words Group is positive and statistically significant for the categorical dummy model in Panel B, consistent with our full sample results. *Letters Group* is negative and statistically significant in the baseline regression of -0.3% and also for the *Letters Group = 3* dummy in the categorical dummy regression. The remaining fluency measures remain statistically insignificant. Combining all fluency measures in the last column of Table 3, Panel A, we find *CommonName Group* and *Letters Group* remain negative and statistically significant at the 5 percent level. Overall, our results are consistent with the full sample, although the matched sample exhibits reduced magnitude of fluency effects. The results suggest that the main findings are robust and not a result of unobservable or omitted variables related to a home's location.

4. Heterogeneity Analysis

4.1 Consumption Domain and Name Fluency

In our prior results we find evidence that rarely used street names and names with more letters have higher prices, inconsistent with street name fluency rendering homes more appealing. We also find that one-worded street names (i.e. more fluent street names) are sold at higher prices, consistent with our fluency hypothesis. These findings appear to be contradictory.

One possible explanation is that subjects' preference for fluency is dependent on the consumption domain, which is the context in which the buying decision is made. For example, Pocheptsova, Labroo and Dhar (2010) find that uncommon products are more desirable when the product's description is less fluent (manipulated by making the text font difficult to read), despite wide acceptance in the literature that fluency makes a product more appealing. The reasoning is that lower fluency makes the uncommon product appear more unique and therefore more desirable. In contrast, for everyday items the authors find that higher fluency makes products more appealing. Pocheptsova, Labroo and Dhar (2010) further posit that high-stakes purchases such as houses are difficult decisions and therefore lower fluency should make them more appealing.

We explore this consumption domain hypothesis as a vaiable explaination for our findings. We consider two consumption domains within housing which may create a preference for less fluent names: rare street names and expensive luxury homes. A rare street name denotes exclusivity and so a less fluent street name may be preferred. Similarly, a high-priced luxurious home is more exclusive, involving a more nuanced decision by the buyer (compared to a cheaper home) with higher metacognitive difficulty. As such, less fluent street names may be more desirable.

To test these hypotheses we estimate a hedonic regression with an interaction of fluency measure group dummies with either a rare street name dummy or luxury home dummy. The regression for rare street name interaction is:

$$\ln(P_{ijst}) = \alpha_t + \beta_1 D(fluency_{ij} = 3) + \beta_2 D(fluency_{ij} = 2) + \beta_3 D(fluency_{ij} = 3) * Rare_{ij} + \beta_4 D(fluency_{ij} = 2) * Rare_{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (9)}$$

And luxury home is:

$$\ln(P_{ijst}) = \alpha_t + \beta_1 D(fluency_{ij} = 3) + \beta_2 D(fluency_{ij} = 2) + \beta_3 D(fluency_{ij} = 3) * Exp_i + \beta_4 D(fluency_{ij} = 2) * Lux_i + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (10)}$$

The fluency measures dummies are the same as those used in the categorical dummy regression model. *Rare* is a dummy of 1 if the home's street name is used in less than five suburbs in Australia (i.e. *CommonName Group* 1 or 2), 0 otherwise. In unreported results, we also use an alternative definition of *Rare* as a dummy of 1 if the street name is found in only one suburb in Australia and find similar results. *Lux* is a dummy of 1 if the home's selling price is in the top quartile of sales prices in that year, zero otherwise. If the interaction of fluency measure dummies with *Rare* or *Lux* is negative and statistically significant, it implies that less fluent street names are more highly valued where exclusivity and/or more difficult decision making is required.

Table 4 reports our results for *Rare* and fluency measure interactions in Panel A (full sample) and Panel B (matched sample). *Lux* and fluency measure interaction regression results are reported in Panel C and Panel D for the full sample and matched sample, respectively. In Panel A, *Words Group=2(Mid) and Words Group=3(High)* remain positive and statistically significant, denoting buyer preferences for one-word street names. *Rare* is also positive and statistically significant, revealing a preference for uncommon street names. Consistent with a preference for less fluency when a street name is rare, *Words Group=2(Mid)*Rare* and *Words Group=3(High)*Rare* are negative and statistically significant in both the full and matched sample results. The results suggest that fewer words in a street name are usually preferred. However, given a rare street name, buyers prefer street names with more words. We also find that the interaction of *Rare* with *Syllable Group=2(Mid)* is negative and statistically significant in the full sample only (Panel A, column 4), consistent with a preference for more syllables given a rare street name. These findings imply that buyers prefer less street name fluency when there is exclusivity in the street name.

[--- INSERT TABLE 4 ABOUT HERE ---]

We find a similar consumption domain effect, according to the luxury home interaction results in Table 4 Panel C and Panel D. For luxury homes, homes with less fluent street names have higher prices than homes with more fluent street names. The interaction effects are also negative and statistically significant for *Lux* interactions with *Words Group=2* (full sample), *Words Group = 3* (both samples), *Syllable Group = 2* (full sample), *Syllable Group = 3*Rare* (full sample), *Letters Group =2* (matched sample), and *Letters Group =3* (matched sample) which shows further evidence of a preference for less fluent names. Overall, our results are consistent with the hypothesis that less fluent street name is preferred within the consumption domains of rare street names and luxury homes.

4.2 Street Name Fluency and English as Second Language Buyers

In this section we test whether street name fluency preferences are higher if a language barrier exists. We hypothesise that there is a preference for more fluency if English is not the buyer's first language, as street names are easier to pronounce and remember for this group of buyers. We do this by testing whether Asian buyers have statistically different fluency preferences from non-Asian buyers. We use Asian buyers as Asians as a group are more recent migrants to Australia compared with Europeans, and so are more likely to speak English as a second language. Asians are a large group in Sydney. In the 2011 Australian Bureau of Statistics Census, people of Asian ethnicity made up about 19 percent of the population in Sydney.

To test for reduced street name fluency effects we include an interaction effect for Asian Buyers in the baseline hedonic regression from Equation 4 as:

$$\ln\left(P_{ijst}\right) = \alpha_t + \beta_1 fluency_{ij} + \beta_2 fluency_{ij} * Asian\ Buyer_i + \beta_3 Asian\ Buyer_i + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (7)}$$

And we also include a similar interaction effect for the categorical dummy regression in Equation 5 as:

$$\ln\left(P_{ijst}\right) = \alpha_t + \beta_1 D(fluency_{ij} = 3) + \beta_2 D(fluency_{ij} = 3) * Asian\ Buyer_i + \beta_3 D(fluency_{ij} = 2) + \beta_4 D(fluency_{ij} = 2) * Asian\ Buyer_i + + \beta_a Asian\ Buyer_i + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (8)}$$

Where *Asian Buyer* is a dummy of 1 if the surname of the buyer(s) is Asian, 0 otherwise. We identify Asian buyers using the surname database of Deng, Deng, Hu and Lee (2019). Non-Asian surnames (e.g. Lee may be Chinese, Korean or Anglo-Saxon) are excluded. If there is more than one buyer in an individual housing sale, we require both buyers to have Asian surnames to be classified as an Asian buyer.

We hypothesise that Asian buyers prefer higher street name fluency due to a language barrier as it is harder for them to remember and pronounce unusual names. We expect the interaction effect between Asian buyers and fluency to be positive, meaning that Asian buyers prefer more fluent street names compared to non-Asian buyers. For example, as we find *CommonName* is negative and statistically significant in our main results, then *CommonName*Asian Buyer* is expected to be positive and statistically significant.

Note that the use of the *Asian Buyer* dummy is only a proxy for buyers where English is their second lanaguage. We use it as a proxy as we have no other information to identify buyers other than surname selection. Our interaction results therefore may be weaker if a component of the Asian buyers in our sample are second generation Asians in Australia and so speak English as their primary language. It may also be the case that street name fluency preferences in fact do not require English. For example, non-English speakers can easily count the number of letters or words in a street name (without knowing how to read it or what it means). Such effects would interfere with the accuracy of the results for the Asian Buyer interaction.

Table 5 reports our results for the baseline regression model with Asian buyer interaction in Panel A and using categorical dummy fluency measures in Panel B. In our sample, Asian buyers make up 151,596, almost 16 percent of the full sample. We do not report coefficient estimates for housing characteristics for conciseness of results. We find some evidence that Asian buyers prefer shorter words, with the *Asian Buyer* interaction for *Letters Group* being positive and statistically significant. While *Words Group* is negative and statistically significant, *Letters Group*Asian Buyer* is positive and statistically significant.

The same result exists for both the baseline regression and categorical dummy regression. In Table 5 Panel A Model (7), *Letters Group* is -0.005 while *Letters Group*Asian Buyer* is 0.003 (both statistically significant at five percent level). This suggests that Asian Buyers prefer fewer letters in a street name relative to non-Asian buyers, indicating a fluency preference. For the categorical dummy regression in Table 5, Panel B, we find a similar effect for the high fluency group interaction (*Letters Group=3(High)*Asian Buyer*). Overall, the results show evidence that non-English speakers prefer higher street name fluency regarding lower number of letters. However, we find no statistical difference in fluency preferences for other measures.

[--- INSERT TABLE 5 ABOUT HERE ---]

4.3 Street Name Fluency for New Homes

As new homes possess more features that are less familiar and more uncertain to potential buyers, name fluency could alleviate this uncertainty, and buyers may base their price on other features such as street name fluency. We use new home sale as the main sample of study and examine whether street name fluency plays a more important role for new homes as the features of new homes are less known to potential buyers. Table 6 presents analysis result using OLS regression. The dependent variable is price and the main explanatory variables are new home dummy and the interaction between new home dummy and the six fluency measures.

[--- INSERT TABLE 6 ABOUT HERE ---]

Consistent with earlier results, new homes are priced higher given the properties have better interior and exterior conditions. In addition, we find new homes indeed have higher price when the street names are more fluent in terms of having fewer syllables and fewer letters in the name. Specifically, when a new home has fewer syllable, such as going from four or five syllables (the least fluent group) to one (the most fluent group), in its street name, the transaction price will increase by 2.7%. When a new home has fewer letters in its name, from the first tercile group to the third tercile group, the price will increase by 2.6%.

This finding that new homes with fluent street names are priced higher than existing homes is consistent with the psychology literature that Processing fluency, or the subjective experience of ease with which people process information is more important when the information is less familiar. Zajonc (1968) provided early evidence that fluency influences liking judgments, when he showed that people prefer familiar stimuli to similar but novel alternatives. Jacoby et al (1992) document that previously-seen stimuli are easier to perceive, encode and process than are stimuli that have never been seen before. For old homes, it is easier for buyers to perceive and price, so the fluency effect will be less important, which is known as fluency discounting as coined by Bornstein and D'Agostino (1994). When information is made available that allowed subjects to correct (i.e., discount) fluency-based liking ratings of stimuli, subjects will lower an initial fluency-based liking rating, hence a discounting attribution.

4.4 Homes with Royal Names

As royalties are historically influential and well respected by the public, association with royal names is usually perceived as being of higher social status and with better recognition. For example, a study done by the Royal Mail Group in the UK finds that around 4,000 residents used Royal names when naming their own homes, out of the 312,000 names homes[14]. It is also found that the names that royals around the world choose for their babies can affect naming trends for years afterwards, as parents tend to name their babies following royal names[15].

Motivated by this notion, we conjecture that certain homebuyers may be willing to pay more for royal names. Specifically, we examine whether transaction prices are higher for homes located on streets named with royal words. We first choose a list of royal names. Royal names are selected if

[14] Source: “It’s all in the name: over 312,000 named homes in the United Kingdom”, Royal mail group, 2018 March 14, https://www.royalmailgroup.com/it%E2%80%99s-all-name-over-312000-named-homes-united-kingdom
[15] Source: “The Meanings Behind Popular Royal Baby Names” , Huffington Post, 2018 May 4th. https://www.huffingtonpost.ca/2018/04/05/royal-baby-names_a_23404002/

they refer to royal titles (e.g. King, Queen, Prince, Princess, etc.), or Buckingham Palace (the residence of the British monarchy) or royalty (e.g. Crown, Palace and Royal, etc.). We create an entire list of 28 words in total from all the street names in the sample. The complete list of royal names used in this study is shown in Appendix 4. “PRINCES”, “KING” and “QUEEN” are the top three royal words, each with occurrence frequency over 12% in the royal name sample.

We then create a royal dummy if the street name of a home contains one of the royal words in the list, and conduct regression analysis using this royal dummy. The results are presented in Table 7. We find that homes with royal names are priced 3.3% higher than those with ordinary non-royal names. We also look at detailed buyer characteristics such as whether buyers are local Australians who are more familiar with the royalty history or whether the buyer is from foreign countries such as Asian countries who do not speak English as a first language. We indeed find Australian buyers pay more for homes on royal name streets and Asians pay less, although the result lacks significance. As owner occupiers are going to stay in homes after purchase, the street name royalty carries more meaning to them. Our result in column 2 implies that owner-occupiers pay more for royal names, although lacking significance.

[--- INSERT TABLE 7 ABOUT HERE ---]

Further, we look at whether the home is luxury property with its price ranged in the top quartile. We find that luxury homes are priced higher with royal names. As buyers of luxury homes are in general wealthy without much financial constraints, they are more likely to pay a premium for positive attributes such as royal names.

As royal names have higher recognition by the public, they could be related with higher perceived popularity, and could affect buyers’ preference of name fluency. We interact royal name dummy with fluency measures in regression models to see whether royal names are priced even higher for fluent names. Table 7 Panel B presents this result. We find that royal names with higher street name fluency (higher in the *CommonName* group measure and hence higher in fluency and less unique) are not priced as high as unique royal names. This suggests that royalty related names could play a substitutive role for fluency. Buyers could substitute the higher recognition from royalty for fluency.

This could also imply buyers who value royalty place a higher emphasis on uniqueness, thus names with high fluency would actually be priced lower. Consistent with the substitution hypothesis, we also find that royal homes are priced lower when there are fewer syllables (Higher in the syllable group measure and hence more fluent and less unique) in the street names, compared with those with lower fluency. The evidence is also consistent with the notion that unique names are given more premiums in the housing market.

4.5 Homes with Trendy Words Based on Google Search

It is possible people prefer names that are popular or trendy in the current time and could have a higher willingness to pay for houses on streets with those trendy names. For example, when movie star Hugh Jackman became popular in the movies "The Wolverine" in 2013, and "Logan" in 2017, the street named after Jackman may have become more favorable than when he was not popular in other time periods.

We investigate this hypothesis by utilizing Google Trends search for each home's street name. For each street name, we collect the Google Trends monthly time series index for Australian region searches. The index ranges from 0 to 100 with 100 being when the street name search term has the highest search volume (as a percentage of total google searches in that month in Australia) over the extracted time period. For example, when plotting the Google Trends index for the street name 'Brock' in Figure 3 between January 2004 (when the Google Trends index starts) to September 2018, the peak search popularity is in September 2006, the month of racing car legend Peter Brock's death.

[--- INSERT FIGURE 3 ABOUT HERE ---]

We then run the following regression to test for the interaction effect of search term popularity, name fluency and housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k GTrend_{ij} + \beta_l GTrend_{ij} * fluency\ measure_{\ ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \qquad \text{--- (10)}$$

Where $GTrend_{ij}$ is a dummy of 1 in the year following the street name's peak search month based on Google Trends from 2004 to the end of our sample period in June 2016.

We indeed find that when Google Trends measure is high, homes on popular street names based on Google trend search are transacted at higher price. As shown in Table 8, the coefficients on *GTrend* range from 0.003 in model (1) to 0.015 in model (5). Given the mean housing price is \$677,000, a peaking of the google trend index for that street is associated with an increase in housing prices of between AUD\$2,000 and AUD\$10,000.

[--- INSERT TABLE 8 ABOUT HERE ---]

Next, we look at the interaction between fluency measure and popularity measure based on google trend. In general, we find Google trend popularity of street names has a complementary effect on name fluency. The result is significant for three fluency measures including Englishness, popular street name group in terms of frequency in suburbs, and letter groups. When google trend popularity is high, the effect of street name fluency is less important in pricing homes. Or put it another way, unique street names with higher Google trend popularity will have even higher price.

4.6 Buyer Fluency Persisitence Preferences

Buyer street name fluency preferences may potentially be stronger for buyers of multiple properties due to a conspicuous consumption effect. That is, they are willing to pay more for homes with more fluent names as a means of collecting more homes. To test this hypothesis, we firstly consider whether buyers do have fluency preferences when buying multiple properties.

We first take the sample of home buyers that have multiple purchases under their name. Homes with missing owner name, with only the surname registered and no first name, couples with the same surname or with common surname combinations (e.g. Kaur; Singh or Wang; Zhang), company owners (denoted with suffix Pty Ltd) and churches are excluded. We group multiple home owners in three equal groups (two for Word Group (single word or multiple word street name) and MS Word) based on their first home purchases' fluency premium or raw fluency measure (this is the fluency measure prior to sorting into groups). We then calculate each home owner's fluency measure over subsequent purchases as their average fluency measure for subsequent purchases (i.e. the fluency measure of their 2nd purchase if they only made 2 purchases and the average fluency measure of their 2nd and 3rd purchase if they made 3 purchases). We then take the mean of the average owner fluency measures for each group. The fluency premiums are the residuals of the baseline hedonic model with all fluency explanatory variables. Table 9 reports the mean owner first purchase fluency measure and mean owner subsequent purchase fluency measures. The difference between the high and low first purchase fluency groups for subsequent purchases is also reported. Panel A to G report sample statistics for fluency premium group, raw Englishness score group, raw words group, MS Word group, raw popularity group and raw syllable two-way sorts, respectively, with *t*-statistics in parentheses.

[--- INSERT TABLE 9 ABOUT HERE ---]

In Panel A for fluency premiums we find persistence in fluency premium based on the record of the homebuyers. For example, for the second purchase of buyers, the high fluency premium group pay 4% above the expected price from the hedonic model while the low group pay 2.1% less. This difference of about 6% is statistically significant at the 1% level. Looking at 3rd and subsequent purchases (all purchases after the first) we also find a positive difference between the high and low group, suggesting persistence in getting the fluency premium. Note while there is persistence, the magnitude of the fluency premium falls after the first purchase. For example, looking at the low group, the fluency premium for the first purchase is -21.6% but for subsequent purchases is -10.5%, a difference of 11.1%.

Looking at raw fluency groups in Panel B to Panel G we find similar patterns of persistence: high groups will buy higher fluency homes than low groups following the first purchase, with the difference shrinking after the first purchase. These differences are all positive and statistically significant. Our findings suggest that there is a fluency preference persistence by homebuyers.

We then examine whether this fluency preference persistence leads to higher fluency premiums in housing price. To do this we estimate the following regression:

$$\ln\left(P_{ijst}\right) = \alpha_t + \beta_k fluency_{ij} + \beta_l Hfluency_{ij} + \beta_m fluency_{ij} * Hfluency_{ij} * Next\ Buy_i + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (11)}$$

Hfluency$_{ij}$ is a dummy of 1 if a buyer's first purchase has a street name in the top tercile of our street name *fluency* raw measures[16], 0 otherwise. *Next Buy*$_i$ is a dummy of 1 if the purchase is the second or subsequent purchase made by the buyer, 0 otherwise. If the triple interaction $fluency_{ij} * Hfluency_{ij} * Next\ Buy_i$ is positive this suggests that those that tend to buy high fluency homes tend to also pay higher prices for such homes.

We report the coefficient estimates of our fluency persistence regression in Table 10. We find the triple interaction term is positive and statistically significant across fluency measures, except for *Words Group* suggesting a price premium paid by high fluency buyers for the homes. The coefficients for each *Hfluency*$_{ij}$ dummy are negative and statistically significant suggesting that although high fluency buyers pay more for high fluency, they pay less than other buyers for homes, all else being equal. The results provide support for multiple buyers of homes having a preference for name fluency.

[--- INSERT TABLE 10 ABOUT HERE ---]

[16] For *Words Group* it is if the street name has two or more words and for *MS Word group* it is if MS Word equals 1.

5. Robustness Checks

5.1 Homes with Multiple Street Names

In this section we focus on properties that locate on more than one streets, whose address can be denoted using multiple street names. Homes have multiple street names if their geocodes are matched to two or more addresses. To find homes with more than one address, we use geocode data from PSMA Australia's Geocoded National Address File (G-NAF) to identify geocodes with multiple addresses but with the same geocode. For example in Figure IA1, geocode (-33.9063,151.0792) is both 416 Punchbowl Road, Belfield and also 29 Bazentin St, Belfield. The G-NAF database contains the physical address records of locations in Australia and their respective geocodes. In our sample, we find 5,989 of home transactions with multiple streets out of the entire sample of 958,408 transactions.

For each property that is located on multiple streets, we calculate the highest and the lowest fluency scores of all the associated street names along the six fluency dimensions, including Englishness, number of words, etc. We investigate whether it is the most or least fluent street name that has a higher influence on the sales price.

Table 11 presents the result on homes with multiple street names. In model 1, we find that for homes on multiple streets, the street name with maximum Englishness score influence the price more, suggesting homes on street names that are English sounding words have lower prices. Model 2 presents result on Words Group. As a higher Words Group denotes fewer number of words in the street name, the positive and significant coefficent on *Max Words* suggests that shorter street names with fewer words get priced higher. Similary, the positive and significant coeffcient on *Max Popname* in model 4 posits that homes with more popular names are price higher, as a higher Popname group denotes higher popularity. In model 6, we find street name in *Min Letters Group* (i.e. street names with more letters) has a negative effect on sales price. Overall, we find that if there are multiple streets assciated with a home, the street name with English sounding word reduces the price, whereas a more popular name and fewer letters increases the price.

[--- INSERT TABLE 11 ABOUT HERE ---]

5.2 Street Centrality

A notable concern of our measures is that street name fluency may be related to how central a street is. Anecdotally, major thoroughfares tend to have more words in the street name such as General Holmes Drive, Georges River Road and James Ruse Drive. As such, our fluency measures may be picking up the effect of being on a busy street rather than the fluency of the street name. To more extensively capture the effect of street centrality in addition to our *Long Street* and *Major Street* measures, we construct a street centrality measure for all streets in Sydney and include it as a control variable in our baseline regressions.

To measure street centrality, we first collect geospatial data of all streets in Sydney from openstreetmap.org. We then apply network analysis to the street where every intersection (or the end of a street if it is a dead end) is a node and the streets to each node are edges. For every node, we calculate its degree centrality as:

$$CD(node) = deg(node)/E \qquad \text{--- (12)}$$

Where *deg(node)* is the number of edges that the node has. *E* is the number of edges in the entire network. *Street Centrality* is measured as the sum of *CD(node)* for all intersections (nodes) on a street, standardised (so *Street Centrality* has a mean of 0 and a standard deviation of 1 within the sample). Thus a central street such as a major thoroughfare will have high street centrality as it contain many nodes and each nodes has high degree centrality as it connects to side streets. In contrast a cul-de-sac will have low street centrality as it only contain two nodes with low degree centrality and thus have low street centrality.

Table 12 reports our regression results including street centrality. We lose 25,817 or 2.7% of observations due to street name/street type/suburb combinations from openstreetmap.org not matching with our sales database. Table 12 Panel A reports correlation statistics of the housing price, fluency measures and street centrality measure. Consistent to the anecdotal evidence, *Street Centrality* is negatively correlated to the housing price and to *Words, CommonName* and *Syllable* fluency measures. For example, the correlation of *Street Centrality* and *Words* is -0.15 so the higher the fluency (less words in a street name), the lower the street centrality, and vice versa. This is consistent to the anecdotal evidence that street names with more words are more central streets and therefore it is a legitimate concern to control for street centrality in our regressions.

[--- INSERT TABLE 12 ABOUT HERE ---]

The correlation to price however is only slightly negative of -6%. *Street Centrality* is highly positively correlated to *Long Street* and *Major Street* of 61% and 50%, respectively. This is consistent to our Street Centrality capturing how well connected a street is. To avoid multicollinearity issues in our regressions, we remove *Long Street* and *Major Street* from our control variables although our results remain qualitatively similar including them.

Panel B reports our baseline results using linear fluency measures with street centrality. Panel C reports regression results for categorical fluency measures with street centrality. For brevity, we do not report coefficient estimates for housing characteristics. The results are qualitatively similar to the baseline results in Table 2, with fluency measure coefficients being of similar magnitude and statistcal significance. *Street Centrality* is negative and statistically significant of -0.012 across regressions. The coefficient estimate implies that a one standard deviation increase in the street centrality measure reduces the price of the home by 1.2 percent. Thus more central streets that are likely to act as thoroughfares to other streets have lower housing prices. Overall, the street name fluency results remain robust to adjusting for the centrality of all streets to the neighborhood.

5.3 Homes with Suburb Name Changes

In this section, we consider the effect of suburb name changes in our sample period[17]. We are able to identify two suburb name changes in our sample.[18] Figure 5 depicts the regions. In the first case in Panel A, the suburb of Harbord in the Northern Beaches Local Government Area (LGA) officially changed its name on January 12th, 2008. In the second case in Panel B, a section of several streets in Moorebank in the western suburbs of Sydney changed its name to the adjoining suburb of Wattle Grove. This area is shaded in blue. This change is recorded by the NSW Government's spatial services on April 12th, 2012. To analyze the effect of these suburb name changes on housing price, we apply the following diff-in-diff regression:

$$\ln\left(P_{ijst}\right) = \alpha_t + \beta_k Post_t + \beta_l * Treatment\ Area_i + \beta_l Post_t * Treatment\ Area_i + property\ char_i + \mu_s + Y_t + \varepsilon_{it} \quad \text{--- (15)}$$

Where *Post* is a dummy of 1 if a home sale occurs after the announcement date, and 0 otherwise; *Treatment Area* is the area where a suburb name change occurs, 0 otherwise. For Harbord/Freshwater the treatment area is the suburb of Harbord/Freshwater and the control area is the Northern Suburbs LGA. For Moorebank/Wattle Grove the treatment area is the area which changed names to Wattle Grove and the control group are the remaining areas of Moorebank and Wattle Grove. The sample is two years before and after the announcement date excluding sales in the month of the announcement. If *Post*Treatment Area* is positive and statistically significant then the name change increased the value of the homes in the affected area.

[17] We also test for the effect of street name change by manually checking street name changes using google maps. We are able to identify 19 street name changes in our sample area and period. We then link these street name change pairs to our sales sample and find 47 observations for 5 street name changes. Internet Appendix Table IA1 reports the result. Panel A is the univariate test where we find price does increase after the name change though statistically insignificant. Street name fluency generally falls with Englishness Group, MS Word and CommonName Group differences being negative and statistically significant. Words Group fluency increases and is statistically significant. In Panel B of our hedonic model we do not find any fluency measure being statistically significant.

[18] Boundary changes are more common though affect few homes and to test for the name change would generally require the same home to transact before and after the boundary change which is rare.

Table 13 reports our results in Panel A for Harbord/Freshwater name change and Panel B for Moorebank/Wattle Grove. In Panel A column 1 we first test if the treated area increased in price univariately. We find a 11.2 percent statistically significant increase in price. Column 2 reports the diff-in-diff using the actual official name change date. *Post*Treatment Area* is positive and statistically significant of 0.029 which suggests that controlling for housing characteristics and surrounding homes in the same LGA, homes in Harbord/Freshwater increased by 2.9 percent after the name change. In column 3 we apply a falsification test using a date two years before the official name change and find the *Post*Treatment Area* coefficient is not statistically significant. This suggests that our main diff-in-diff result is not driven by a time trend in prices.

In Panel B we apply the diff-in-diff to the Moorebank/Wattle Grove suburbs, with the treatment area being the section of Moorebank that changed suburb names to Wattle Grove. Column 1 reports the univariate results that the changed name area increased by 14.6 percent. In column 2 for the diff-in-diff regression we find *Post*Treatment Area* coefficient of 0.022, statistically significant at the 1 percent level. This suggests that the area that changed to Wattle Grove increased by 2.2% after the name change and relative to the surrounding Moorebank and Wattle Grove suburb sales. In column 3's falsification test we do not find *Post*Treatment Area* coefficient statistically significant and thus the main diff-in-diff result is not due to a time trend in prices. Thus in our two areas where a suburb name change occurred, we find that it brought economically large gains of between 2.2 and 2.9% to home owners all because of a name change.

[--- INSERT TABLE 13 ABOUT HERE ---]

6. Conclusion

The economics and psychology literature documents that people derive higher utility from fluent sounding names. Fluent stimuli have been shown to appear more familiar and likeable than similar but less fluent stimuli, resulting in higher judgments of preference (see Alter and Oppenheimer, 2009 for a review). In this study, we examine whether the fluency of a street name is an important feature that influences household property investment decisions.

Utilizing individual residential housing trasnsaction data in Sydney, we investigate the fluency effect by testing the relationship between street name fluency and housing prices. Building on the literature in psychology, which finds that fluent stimuli appear more positive and familiar than nonfluent stimuli, we conjecture that investors will have a preference for homes with fluent street names.

Empolying hedonic housing price models, we find mixed evidence of fluency being priced for housing sales. Consistent with prior studies, we first find homebuyers in genral display a fluency preference for shorter street names with fewer words, and they are willing to pay higher prices for this feature of street name fluency. We then look at other dimensions of fluency and document a uniqueness preference whereby homes with unique street names are associated with statistically higher prices than homes with more common street names.

We conduct further heteroneity analysis on buyer and property charactersitics. Asian buyers, for whom English is more likely a second language, may have a fluency preference due to a language barrier. Our evidence indeed shows that Asian buyers prefer more fluent street names than non-Asian buyers. Preferences for fewer words and more unique street names remain prevalent. In addition, consistent with the consumption domain effect, we find that less fluent street names are preferred when the home is more exclusive, in terms of having a rare street name or in the luxury property price range. We also use a matched home analysis to control for unobserved spatial amenities to ensure the robustness of our results.

Our results reveal novel evidence on hombuyers' preference for street name fluency in six distintive fluency dimensions. We document both fluency preference and uniqueness preference in difference fluency dimensons. Overall, our findings contribute to understanding how name fluency affects the pricing of large investment decisions such as residental real estate.

References

Agarwal, Sumit, Jia He, Haoming Liu, Ivan PL Png, Tien-Foo Sing, and Wei-Kang Wong, 2014, Superstition and asset markets: Evidence from Singapore housing, *Working paper, National University of Singapore*.

Alter, Adam L, and Daniel M Oppenheimer, 2006, Predicting short-term stock fluctuations by using processing fluency, *Proceedings of the National Academy of Sciences* 103, 9369-9372.

Australian Bureau of Statistics, 2019, Household income and wealth, Australia, 2017-18, in ABS, ed.: Cat. no. 6523.0 (Canberra, Australia).

Bertrand, Marianne, and Sendhil Mullainathan, 2004, Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination, *American Economic Review* 94, 991-1013.

Bornstein, Robert F, and Paul R D'Agostino, 1994, The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect, *Social Cognition* 12, 103-128.

Cardoso, Rodrigo V, and Evert J Meijers, 2017, The metropolitan name game: The pathways to place naming shaping metropolitan regions, *Environment and Planning A* 49, 703-721.

Chau, K, Vincent Ma, and Daniel Ho, 2001, The pricing of 'luckiness' in the apartment market, *Journal of Real Estate Literature* 9, 29-40.

Choy, Lennon HT, Stephen WK Mak, and Winky KO Ho, 2007, Modeling Hong Kong real estate prices, *Journal of Housing and the Built Environment* 22, 359-368.

Davies, Ryan J, and Sang Soo Kim, 2009, Using matched samples to test for differences in trade execution costs, *Journal of Financial Markets* 12, 173-202.

Deng, Yongheng, Maggie Rong Hu, and Adrian D Lee, 2019, Melting pot or salad bowl: Cultural distance and housing investments Working paper, Chinese University of Hong Kong.

Fortin, Nicole M, Andrew J Hill, and Jeff Huang, 2014, Superstition in the housing market, *Economic Inquiry* 52, 974-993.

Glaeser, Edward L, Michael Scott Kincaid, and Nikhil Naik, 2018, Computer vision and real estate: Do looks matter and do incentives determine looks, *Working paper, National Bureau of Economic Research*.

Green, T Clifton, and Russell Jame, 2013, Company name fluency, investor recognition, and firm value, *Journal of Financial Economics* 109, 813-834.

Hamed, Osama, and Torsten Zesch, 2015, Generating nonwords for vocabulary proficiency testing, Proceeding of the 7th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics (Poznań, Poland).

Huang, RD, and HR Stoll, 1996, Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE, *Journal of Financial Economics* 41, 313-357.

Keys, Benjamin J., Devin G. Pope, and Jaren C. Pope, 2016, Failure to refinance, *Journal of Financial Economics* 122, 482-499.

Norris, Darrell A, 1999, Unreal estate: Words, names and allusions in suburban home advertising, *Names* 47, 365-380.

Oppenheimer, Daniel M, 2006, Consequences of erudite vernacular utilized irrespective of necessity: Problems with using long words needlessly, *Applied Cognitive Psychology* 20, 139-156.

Ossokina, Ioulia V., and Gerard Verweij, 2015, Urban traffic externalities: Quasi-experimental evidence from housing prices, *Regional Science and Urban Economics* 55, 1-13.

Pocheptsova, Anastasiya, Aparna A Labroo, and Ravi Dhar, 2010, Making products feel special: When metacognitive difficulty enhances evaluation, *Journal of Marketing Research* 47, 1059-1069.

Rosenbaum, Paul R, 1989, Optimal matching for observational studies, *Journal of the American Statistical Association* 84, 1024-1032.

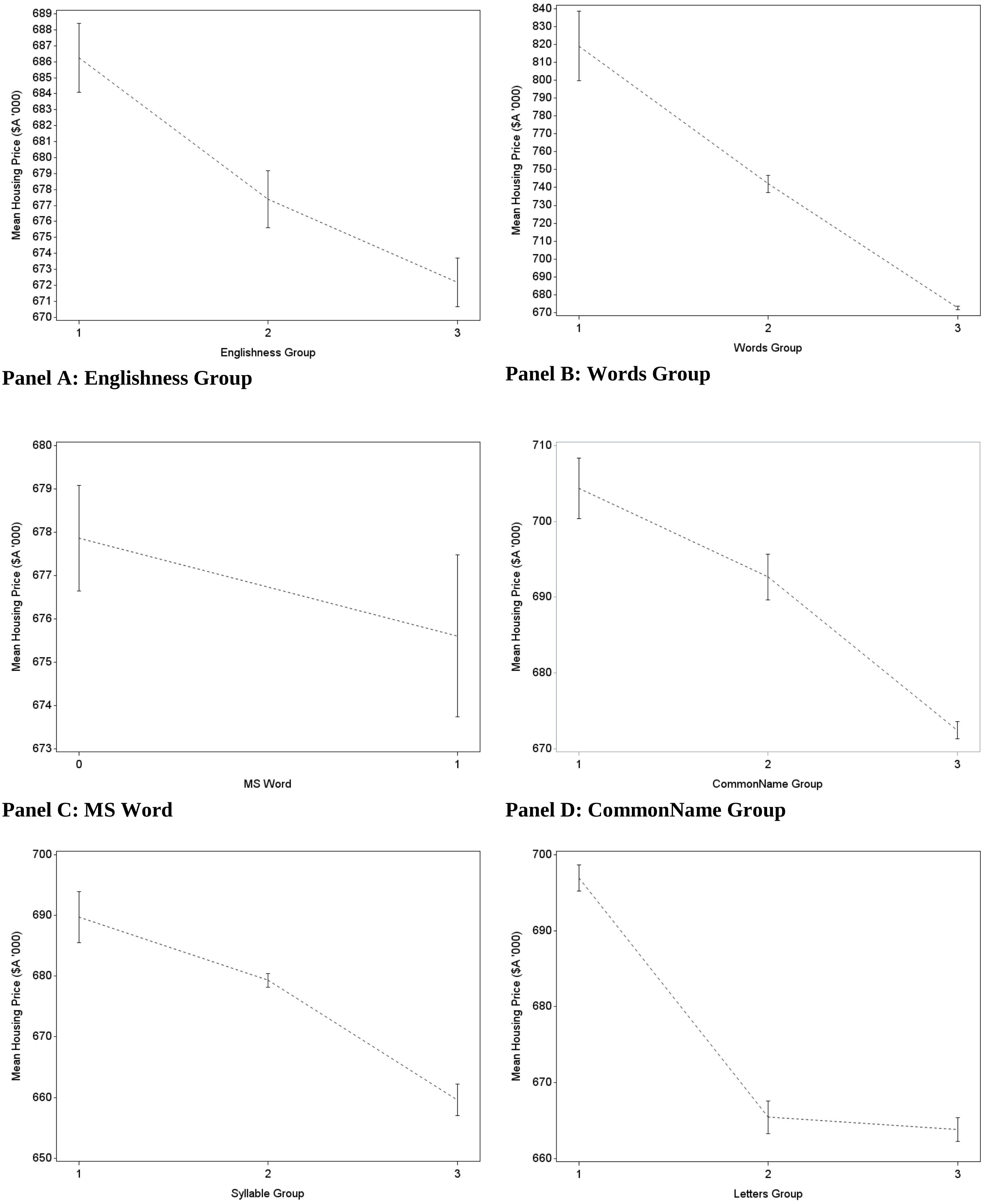
Schwarz, Norbert, 2004, Metacognitive experiences in consumer judgment and decision making, *Journal of Consumer Psychology* 14, 332-348.

Schwarz, Norbert, Herbert Bless, Fritz Strack, Gisela Klumpp, Helga Rittenauer-Schatka, and Annette Simons, 1991, Ease of retrieval as information: Another look at the availability heuristic, *Journal of Personality and Social psychology* 61, 195.

Zajonc, Robert B, 1968, Attitudinal effects of mere exposure, *Journal of Personality and Social Psychology* 9, 1-27.

Figure 1: Average Housing Price for Street Name Fluency Groups

This figure present the mean housing price for each fluency group using our six fluency measures. Higher ranked groups are more fluent. Bars represent 5% and 95% confidence intervals. Englishness Group measures how often a combination of letters appears in English media. Word group is the number of words in the street name. MS word indicates whehter a street name passes the MS Word spell check. CommonName group is the number of suburbs that share the same street name. Syllable Group is the number of syllables in a street name. See Section 2.2 for more details on fluency measures.



Panel A: Englishness Group

Panel B: Words Group

Panel C: MS Word

Panel D: CommonName Group

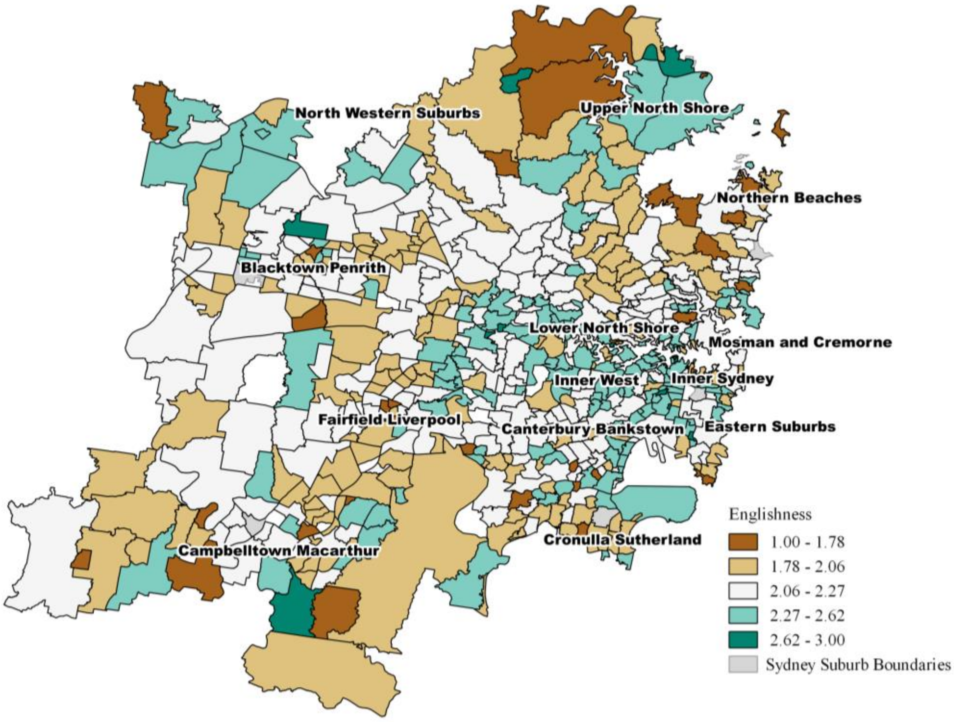
Panel E: Syllable Group

Panel F: Letters Group

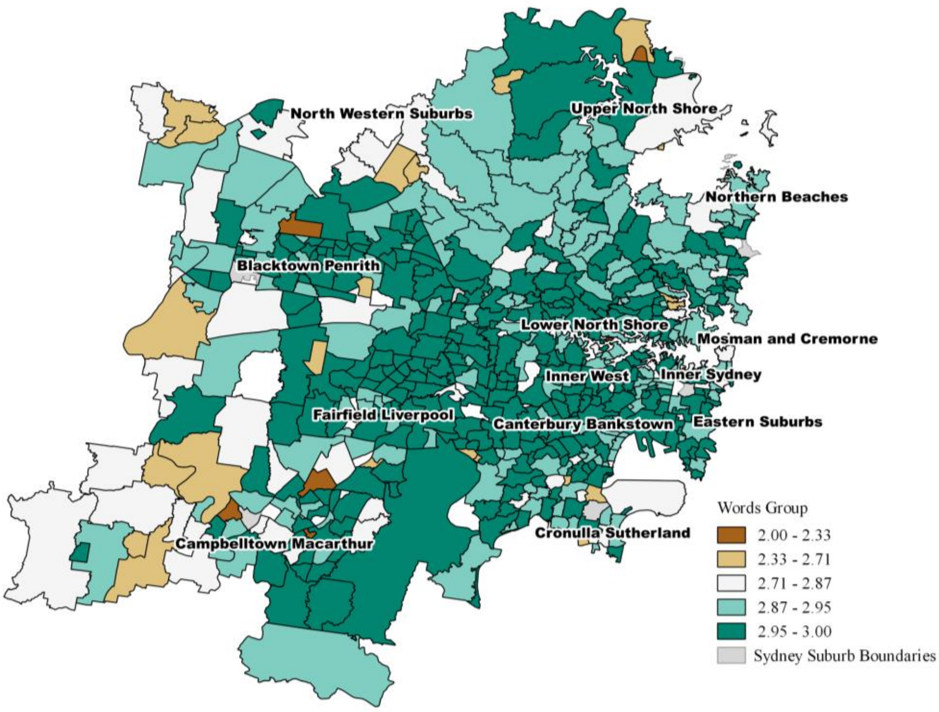
Figure 2: Fluency Heatmaps across Sydney Suburbs using Six Fluency Measures

This figure illustrates the fluency score of street names in various suburbs across Sydney using heat maps, whereby greener shades represent higher fluency scores, and browner shades correspond to lower fluency scores. The results for our six diffent fluency measures are presnted in Panels A to F. Englishness Group measures how often a combination of letters appears in English media. Word group is the number of words in the street name. MS word indicates whehter a street name passes the MS Word spell check. CommonName groop is the number of suburbs that share the same street name. Syllable Group is the number of syllables in a street name. See Section 2.2 for more details on fluency measures.

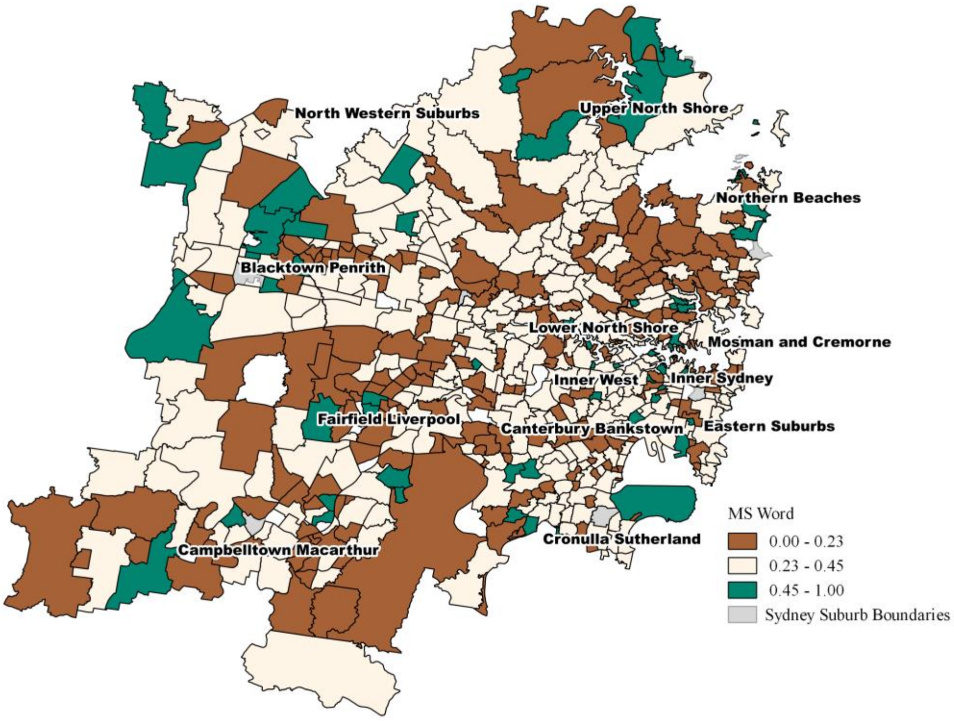
Panel A: Englishness Group



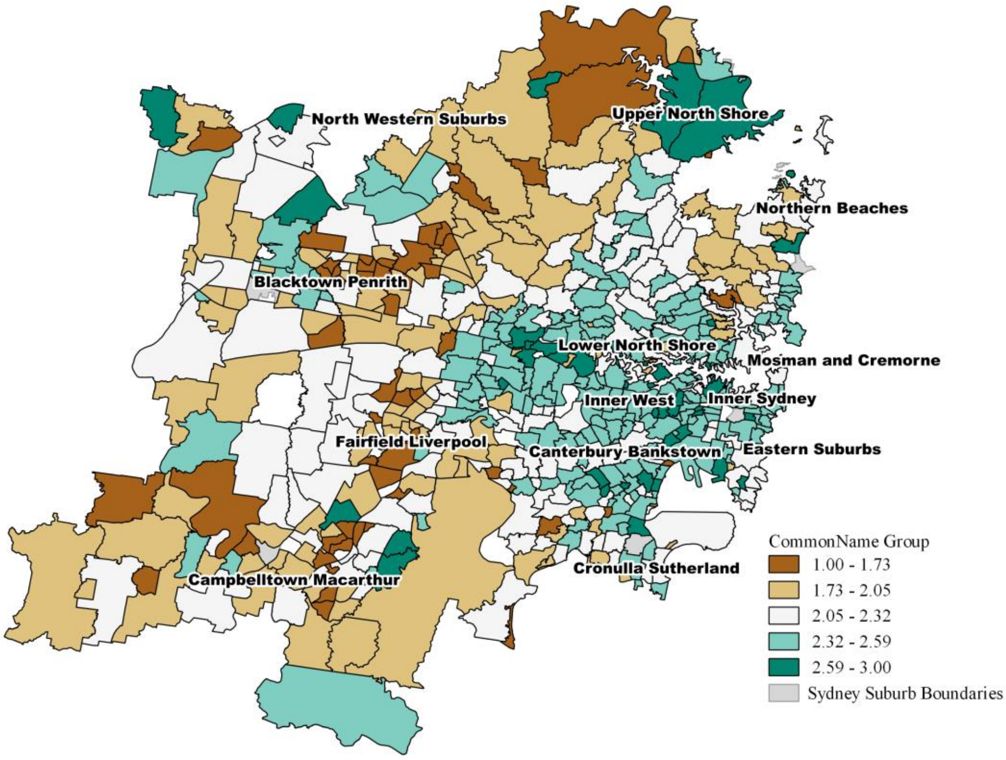
Panel B: Words Group



Panel C: MS Word



Panel D: CommonName Group



Panel E: Syllable Group

Panel F: Letters Group

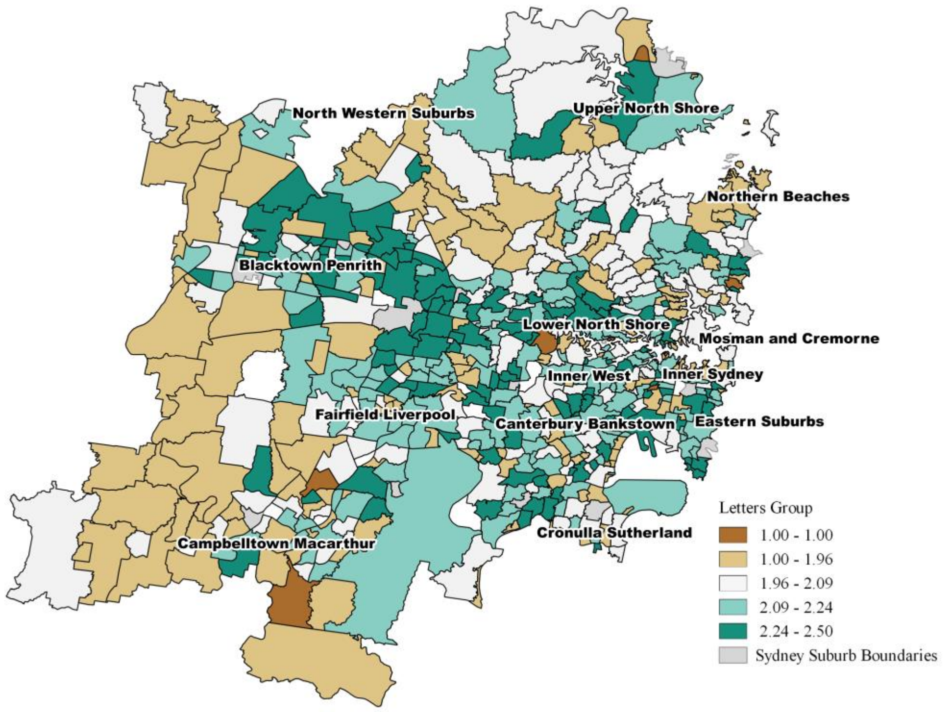


Figure 3: Google Trends Index for Australian Region Search of 'Brock'

The figure reports the Google Trends Index for the search term 'Brock' in the Australian region from 2014 Janury to 2018 August. The index spans from 0 to 100 and hits the maximum 100 in September 2006 when legendary Australian racecar driver Peter Brock passed away.

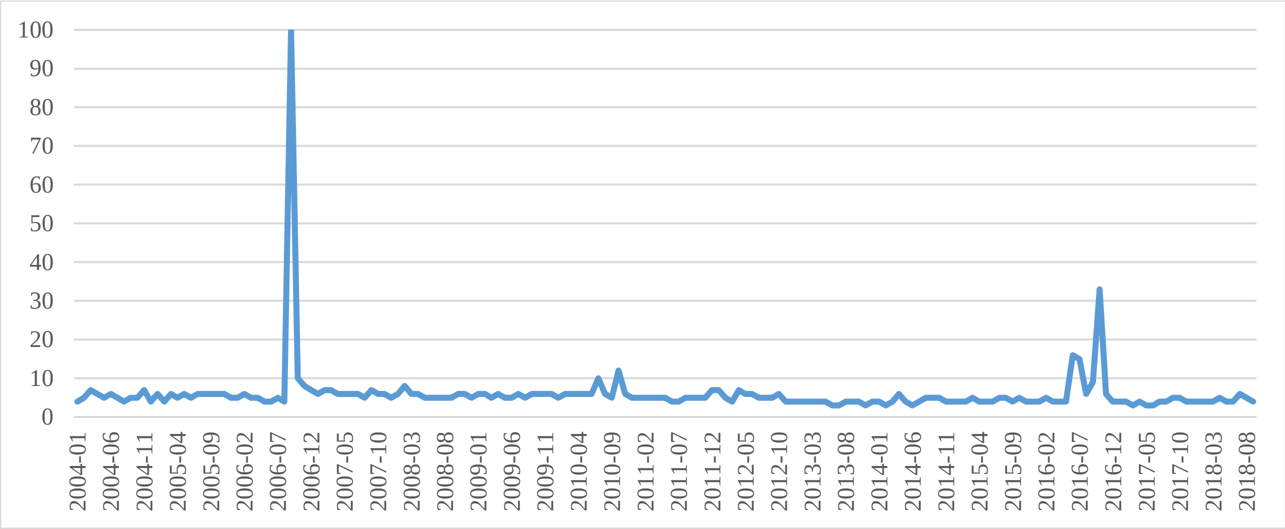
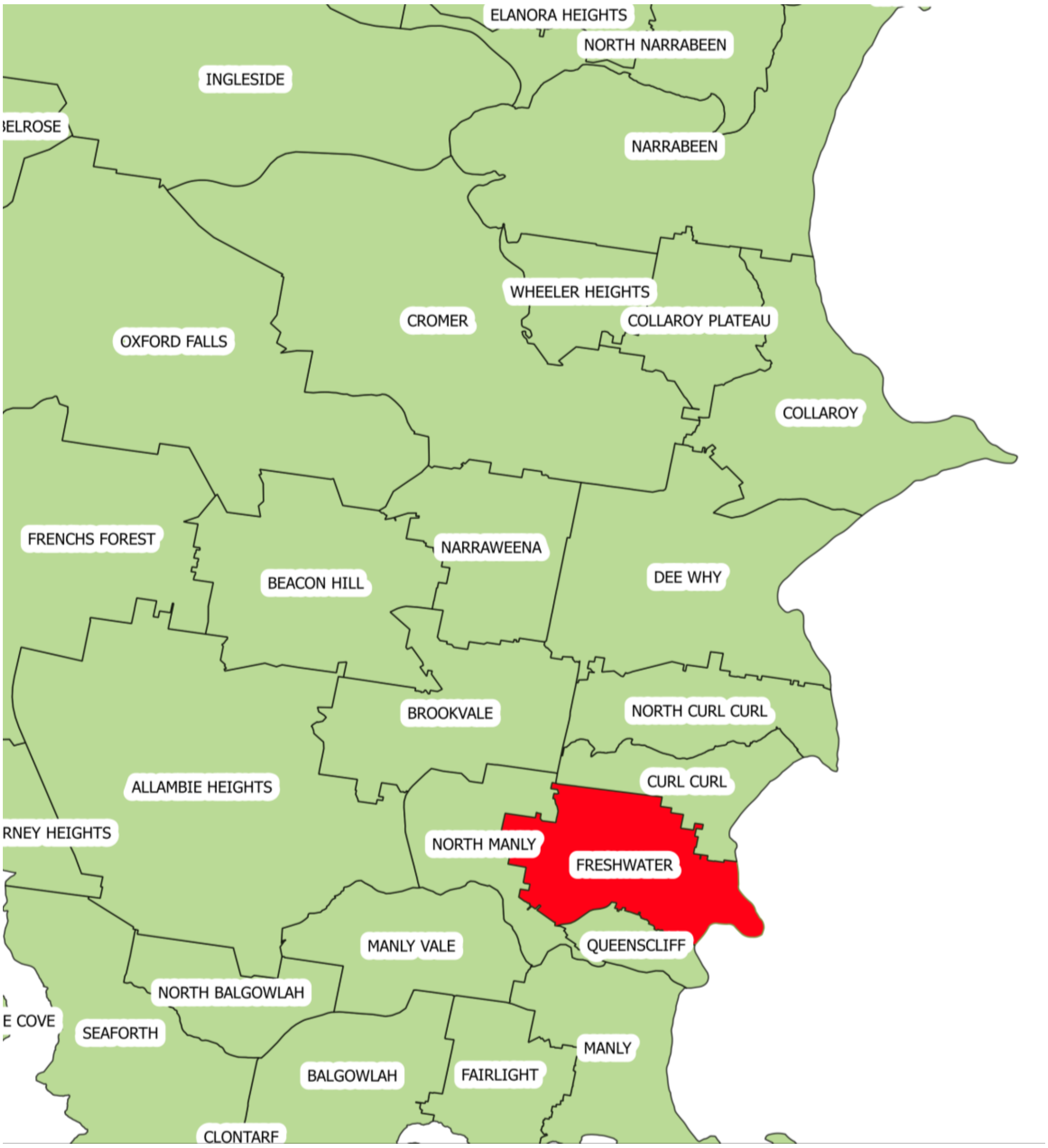


Figure 4: Changed Suburb Name Areas

Panel A: Harbord to Freshwater

The figure shows the treatment suburb Freshwater (formerly Harbord, in red) and surrounding suburbs in the Northern Beaches Local Government Area as the control area. Source: data.gov.au



Panel B: Moorebank to Wattle Grove Suburb Change

The figure shows Moorebank (bounded by the red border), Wattle Grove (bounded by the green border) and the area that changed from Moorebank to Wattle Grove (shaded in blue). The shaded area is the treatment area while Moorebank and the other parts of Wattle Grove are the control areas. Source: Google Maps

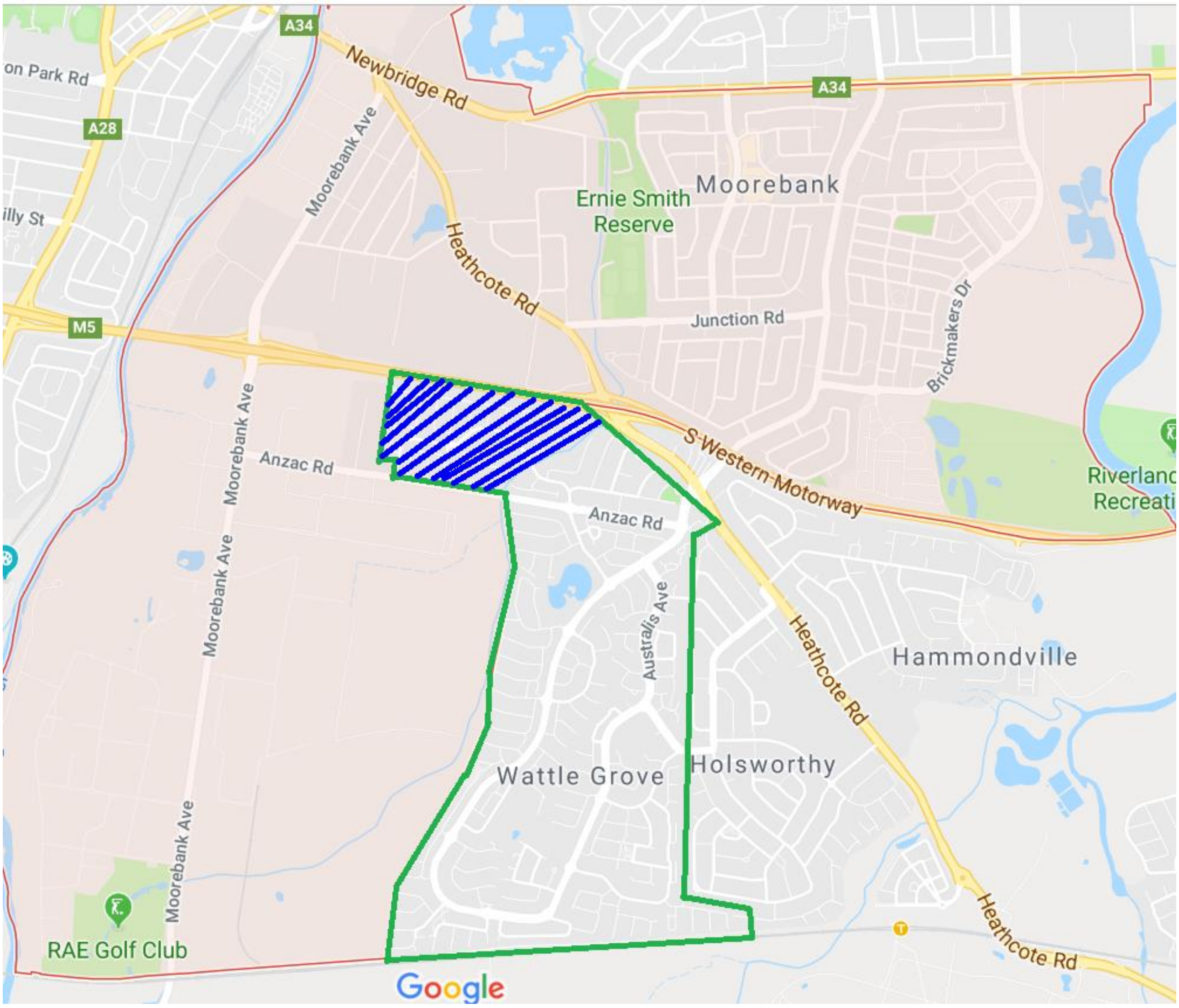


Table 1: Summary Statistics of Home Sales by Total Street Name Fluency Score

This table details various summary statistics for home sales in the Sydney metropolitan area from January 2000 to June 2016. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how street names are placed in the six fluency measure groups. Higher scores reflect higher fluency of street names. Price is documented in thousands of Australian dollars. House is a dummy variable equal to one for a freestanding house and zero. Size is the land area size of the home in 1,000 square feet. Beds is the number of bedrooms in the home. Baths is the number of bathrooms in the home. Parking is a dummy variable equal to one if the home has parking, zero otherwise. New is a dummy of 1 if the home is a new development sale, zero otherwise (i.e. a second hand sale). Auction is a dummy variable equal to one if the home was sold at auction. Long Street is a dummy of 1 if the street on which the home is situated is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise. Major Street is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise. Panel A reports mean median, first quartile, third quartile and standard deviation for each measure. Panel B reports mean summary statistics by street name fluency aggregate score (the sum of the six fluency measures). Panel C reports the correlation matrix of housing price and fluency measures. Panel D reports the frequency counts for fluency measure groups. Panel E reports mean sumarry statistcs of fluency scores in the top 20 suburbs by sales.

Panel A: Summary Statistics

Measure	Mean	Median	Std	Q1	Q3	N
Housing Chareacteistics						
Price	677.19	530.00	510.90	369.00	790.00	958,408
House	0.57	1.00	0.49	0.00	1.00	958,408
Size	4.14	3.23	7.15	0.00	6.58	958,408
Beds	2.89	3.00	1.06	2.00	4.00	958,408
Baths	1.60	1.00	0.72	1.00	2.00	958,408
Parking	0.75	1.00	0.43	0.00	1.00	958,408
New	0.05	0.00	0.21	0.00	0.00	958,408
Auction	0.18	0.00	0.38	0.00	0.00	958,408
Long Street	0.24	0.00	0.43	0.00	0.00	958,408
Major Street	0.07	0.00	0.26	0.00	0.00	958,408
Street Name Chareacteistics						
Englishness Group	2.20	2.00	0.80	2.00	3.00	958,408
Words Group	2.93	3.00	0.27	3.00	3.00	958,408
MS Word	0.30	0.00	0.46	0.00	1.00	958,408
CommonName Group	2.73	3.00	0.58	3.00	3.00	958,408
Syllable Group	2.08	2.00	0.45	2.00	2.00	958,408
Letters Group	2.01	2.00	0.89	1.00	3.00	958,408

Panel B: Mean Measures by Fluency Score

Aggregate Fluency Score	Price	House	Size	Beds	Baths	Parking	New	Auction	Long Street	Major Street	N
5 to 6 (low fluency)	698.56	0.73	5.66	3.33	1.90	0.87	0.04	0.12	0.34	0.09	1,799
7 to 8	728.98	0.65	5.35	3.06	1.71	0.76	0.04	0.16	0.25	0.06	28,014
9 to 10	711.22	0.62	4.78	3.00	1.65	0.75	0.04	0.17	0.28	0.11	154,583
11 to 12	673.89	0.59	4.35	2.95	1.61	0.76	0.04	0.17	0.25	0.08	351,709
13 to 14	660.55	0.55	3.74	2.83	1.57	0.75	0.05	0.18	0.21	0.06	288,555
15 to 16 (high fluency)	671.32	0.50	3.40	2.73	1.55	0.72	0.05	0.19	0.22	0.06	133,748
All Sales	677.19	0.57	4.14	2.89	1.60	0.75	0.05	0.18	0.24	0.07	958,408

Panel C: Correlation Matrix of Price and Fluency Measures

	Price	Englishness	Words	MS Word	CommonName	Syllable	Letters
Price	1						
Englishness Group	-0.01	1					
Words Group	-0.04	-0.06	1				
MS Word	0	0.29	-0.12	1			
CommonName Group	-0.02	0.16	0.26	0.14	1		
Syllable Group	-0.01	0.19	0.19	0.25	0.16	1	
Letters Group	-0.03	0.06	0.28	0.19	0.16	0.48	1

Panel D: Size and Proportion of Fluency Groups

	Low Fluency (Group = 1) or MS Word = 0		Medium Fluency (Group = 2)		High Fluency (Group = 3) or MS Word = 1		All	
	N	%	N	%	N	%	N	%
Englishness Group	226,641	23.65	309,162	32.26	422,605	44.09	958,408	100.00
Words Group	4,072	0.42	56,204	5.86	898,132	93.71	958,408	100.00
MS Word	674,284	70.35	-	-	284,124	29.65	958,408	100.00
CommonName	68,631	7.16	117,386	12.25	772,391	80.59	958,408	100.00
Syllable Group	61,862	6.45	761,135	79.42	135,411	14.13	958,408	100.00
Letters Group	377,241	39.36	199,094	20.77	382,073	39.87	958,408	100.00

Panel E: Mean Fluency Scores for Top 20 Suburbs by Sales Volume

Suburb	N	Englishness	Words	MS	Popname	Syllable	Letters
Mosman	10,3	2.21	2.92	0.45	2.60	2.06	2.10
Blacktown	9,91	2.14	2.98	0.30	2.72	2.15	2.09
Castle Hill	9,36	2.17	2.91	0.33	2.62	2.06	1.89
Dee Why	8,93	2.23	2.93	0.43	2.84	2.13	2.09
Baulkham Hills	8,47	2.31	2.90	0.27	2.66	2.05	2.10
Randwick	7,98	2.24	2.93	0.26	2.81	2.20	2.33
Cronulla	7,68	1.99	2.94	0.15	2.54	2.02	2.08
Maroubra	7,10	1.94	2.96	0.22	2.89	2.20	2.21
Parramatta	6,68	2.58	2.88	0.31	2.98	2.12	2.05
Auburn	6,65	2.34	2.95	0.36	2.79	2.06	2.00
Bankstown	6,59	2.29	2.92	0.23	2.75	2.07	2.05
Liverpool	6,49	2.29	2.99	0.21	2.90	2.16	1.83
Quakers Hill	6,38	1.97	2.99	0.26	2.65	2.03	2.02
Hornsby	6,37	2.43	2.97	0.46	2.91	2.10	2.22
Hurstville	6,27	2.45	2.96	0.40	2.87	2.18	2.19
Chatswood	6,26	2.40	2.99	0.49	2.90	2.12	2.10
Marrickville	6,08	2.16	3.00	0.23	2.84	2.07	1.99
Merrylands	5,91	2.30	2.98	0.21	2.81	2.10	2.00
Manly	5,80	2.31	2.93	0.38	2.83	2.16	2.06
Surry Hills	5,63	2.29	2.97	0.36	2.94	2.11	2.05
All Sales (Top 20)	144,	2.24	2.95	0.32	2.78	2.11	2.07

Table 2: Hedonic Regression with Street Name Fluency

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k fluency_{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; $fluency_{ij}$ denotes one of the six fluency street name measures for a home sold on street name *j*; $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. Other control variables are described in Appendix 1. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. Panel A uses continuous fluency group measures, as desceribed in Section 2.2. Panel B uses categorical fluency tercile group. A street name fluency is in the highest tercile 3 if the street name of the home belongs in the highest fluency group (i.e., Fluency Tercile Group==3), and zero otherwise. Fluency Tercile Group==2 denotes the middle group (the omitted dummy being the lowest street name fluency group). ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: using Fluency Score Group

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	0.000					
	(0.002)					
Words Group		0.000				
		(0.009)				
MS Word			0.003			
			(0.004)			
CommonName Group				-0.007***		
				(0.002)		
Syllable Group					-0.001	
					(0.003)	
Letters Group						-0.003**
						(0.001)
New	0.137***	0.137***	0.137***	0.137***	0.137***	0.137***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Auction	0.060***	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.125***	0.125***	0.125***	0.125***	0.125***	0.125***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bath	0.133***	0.133***	0.133***	0.132***	0.133***	0.132***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.040***	0.040***	0.040***	0.040***	0.040***	0.040***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Long Street	-0.011***	-0.011***	-0.012***	-0.011***	-0.011***	-0.012***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.021***	-0.021***	-0.021***	-0.021***	-0.021***	-0.021***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Intercept	3.819***	3.819***	3.82***	3.834***	3.82***	3.825***
	(0.457)	(0.456)	(0.457)	(0.457)	(0.456)	(0.456)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8493	0.8493	0.8493	0.8493	0.8493	0.8493
N	958,408	958,408	958,408	958,418	958,408	958,408

Panel B: using Categorical Fluency Tercile Group Dummy

	Five Fluency Measures				
	Englishness	Words	CommonName	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.001	0.118***	-0.014**	0.000	-0.006**
	(0.003)	(0.031)	(0.006)	(0.006)	(0.003)
Fluency Tercile = 2 (Mid)	0.002	0.136***	-0.004	0.001	-0.001
	(0.003)	(0.031)	(0.006)	(0.006)	(0.004)
New	0.137***	0.137***	0.137***	0.137***	0.137***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Auction	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.125***	0.125***	0.125***	0.125***	0.125***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bath	0.133***	0.132***	0.132***	0.133***	0.132***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.040***	0.040***	0.040***	0.040***	0.040***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Long Street	-0.012***	-0.011***	-0.011***	-0.011***	-0.012***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.021***	-0.02***	-0.021***	-0.021***	-0.021***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Intercept	3.309***	2.181***	1.271***	2.428***	3.038***
	(0.484)	(0.476)	(0.490)	(0.487)	(0.442)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8493	0.8494	0.8493	0.8493	0.8493
N	958,408	958,408	958,408	958,418	958,408

Table 3: Matched Home Hedonic Regressions

This table reports coefficient estimates using the baseline hedonic regression model in Equation 4 and categorical dummy regression model in Equation 5 using matched housing pairs only. To match homes, we find pairs of homes in the full sample of the same property type (house or apartment) that are within 100 meters of each other in the same suburb, on different streets, selling within one year of each other and with similar housing characteristics. Section 2.5 details the algorithm that we use. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. Panel A reports our coefficient estimates using the baseline regression model. Panel B reports coefficients using the categorical fluency measure regression model. The data is obtained from Australian Property Monitors. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Baseline Regression (Matched Pairs Only)

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	0.000					
	(0.001)					
Words Group		0.007				
		(0.008)				
MS Word			0.001			
			(0.004)			
CommonName Group				-0.004**		
				(0.002)		
Syllable Group					-0.002	
					(0.003)	
Letters Group						-0.003**
						(0.001)
New Development	0.131***	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Auction	0.057***	0.057***	0.057***	0.057***	0.057***	0.057***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.128***	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Bath	0.128***	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.051***	0.051***	0.051***	0.051***	0.051***	0.051***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Long Street	-0.013***	-0.013***	-0.013***	-0.013***	-0.013***	-0.013***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.024***	-0.023***	-0.024***	-0.024***	-0.024***	-0.024***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Intercept	3.477***	3.455***	3.477***	3.476***	3.48***	3.479***
	(0.537)	(0.537)	(0.537)	(0.537)	(0.536)	(0.536)
Other Housing	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8656	0.8656	0.8656	0.8656	0.8656	0.8656
N	488,784	488,784	488,784	488,784	488,784	488,784

Panel B: Categorical Regression (Matched Pairs Only)

	Five Fluency Measures				
	Englishness	Words	CommonName	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	-0.001	0.107***	-0.006	-0.003	-0.006**
	(0.003)	(0.038)	(0.004)	(0.006)	(0.003)
Fluency Tercile = 2 (Mid)	-0.002	0.114***	0.002	0.000	-0.001
	(0.003)	(0.038)	(0.005)	(0.005)	(0.004)
New Development	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Auction	0.057***	0.057***	0.057***	0.057***	0.057***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Bath	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.051***	0.051***	0.051***	0.051***	0.051***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Long Street	-0.013***	-0.012***	-0.013***	-0.013***	-0.013***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.024***	-0.022***	-0.024***	-0.024***	-0.024***
	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
Intercept	3.088***	1.458***	3.285***	1.984***	1.53***
	(0.582)	(0.555)	(0.559)	(0.582)	(0.583)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8471	0.8472	0.8472	0.8471	0.8472
N	488,784	488,784	488,784	488,784	488,784

Table 4: Consumption Domain and Fluency Measure Interactions

This table reports coefficient estimates of a hedonic model regression for the following model:

$$\ln(P_{ijst}) = \alpha_t + \beta_1 D(fluency_{ij} = 3) + \beta_2 D(fluency_{ij} = 2) + \beta_3 D(fluency_{ij} = 3) * Rare_{ij} + \beta_4 D(fluency_{ij} = 2) * Rare_{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; D(*fluency$_{ij}$=3)* is a dummy of 1 if the street name of the sold home belonged in the highest fluency group in either the *Englishness, Words, CommonName, Syllable* or *Letters Group*s, zero otherwise. D(*fluency$_{ij}$=2)* denotes the middle group (the omitted dummy being the lowest street name fluency group). *Rare* is a dummy of 1 if the home's street name is used in less than 5 suburbs (i.e. is in CommonName Group 1 or 2), 0 otherwise. $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. *longstreet* is a dummy of 1 if the home's street is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise. *majorstreet* is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise. μ_s are suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. Panel A and Panel B report estimates with the interaction for *Rare* using the full sample and matched sample, respectively. Panel C (full sample) and Panel D (matched sample) use *Lux* instead of *Rare* interaction. *Lux* is a dummy of 1 if the home is in the top quartile of prices for the year, zero othewise. The matched sample selection is described in Section 2.5. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Rare Street Name Interaction with Fluency (Full Sample)

	Five Fluency Measures				
	Englishness	Words	MS Word	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.001	0.142***	0.003	0.007	-0.005
	(0.004)	(0.035)	(0.004)	(0.007)	(0.004)
Fluency Tercile = 2 (Mid)	0.004	0.158***		0.008	-0.001
	(0.004)	(0.036)		(0.006)	(0.004)
Fluency Tercile = 3 (High)*Rare	0.001	-0.139***	0.007	-0.033	-0.007
	(0.006)	(0.044)	(0.008)	(0.027)	(0.007)
Fluency Tercile = 2 (Mid)*Rare	-0.010	-0.135***		-0.026**	0.000
	(0.007)	(0.047)		(0.013)	(0.007)
Rare	0.008	0.143***	0.004	0.029**	0.007
	(0.006)	(0.044)	(0.003)	(0.014)	(0.004)
New	0.137***	0.137***	0.137***	0.137***	0.137***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Auction	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.125***	0.125***	0.125***	0.125***	0.125***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bath	0.132***	0.132***	0.132***	0.132***	0.132***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.040***	0.040***	0.040***	0.040***	0.040***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Long Street	-0.011***	-0.010***	-0.011***	-0.011***	-0.011***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.020***	-0.019**	-0.021***	-0.020***	-0.021***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Intercept	3.438***	2.074***	4.400***	2.658***	2.246***
	(0.468)	(0.45)	(0.444)	(0.466)	(0.449)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8493	0.8495	0.8493	0.8493	0.8493
N	958,408	958,408	958,408	958,408	958,408

Panel B: Rare Street Name Interaction with Street Name Fluency (Matched Pairs Only)

	Five Fluency Measures				
	Englishness	Words	MSWord	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	-0.001	0.251***	0.000	0.003	-0.003
	(0.005)	(0.042)	(0.004)	(0.011)	(0.004)
Fluency Tercile = 2 (Mid)	0.001	0.249***		0.006	0.005
	(0.005)	(0.044)		(0.010)	(0.006)
Fluency Tercile = 3 (High)*Rare	0.002	-0.245***	0.005	-0.010	-0.004
	(0.007)	(0.062)	(0.005)	(0.015)	(0.005)
Fluency Tercile = 2 (Mid)*Rare	-0.004	-0.236***		-0.011	-0.011
	(0.006)	(0.063)		(0.011)	(0.007)
Rare	0.005	0.248***	0.004	0.014	0.008*
	(0.005)	(0.061)	(0.003)	(0.010)	(0.004)
New	0.131***	0.132***	0.131***	0.131***	0.131***
	(0.010)	(0.001)	(0.001)	(0.001)	(0.001)
Auction	0.057***	0.056***	0.057***	0.057***	0.057***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Bath	0.128***	0.128***	0.128***	0.128***	0.128***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.051***	0.051***	0.051***	0.051***	0.051***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Long Street	-0.012***	-0.012***	-0.012***	-0.012***	-0.013***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.023***	-0.020***	-0.023***	-0.023***	-0.024***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Intercept	3.225***	1.628***	2.224***	1.978***	1.862***
	(0.545)	(0.51)	(0.507)	(0.546)	(0.513)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8656	0.8657	0.8656	0.8656	0.8656
N	488,784	488,784	488,784	488,784	488,784

Panel C: Luxury Home Interaction with Street Name Fluency (Full Sample)

	Five Fluency Measures				
	Englishness	Words	MS Word	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.001	0.086***	0.003	0.002	-0.006**
	(0.003)	(0.025)	(0.003)	(0.006)	(0.003)
Fluency Tercile = 2 (Mid)	0.002	0.097***		0.002	0.000
	(0.003)	(0.025)		(0.005)	(0.003)
Fluency Tercile = 3 (High)*Lux	-0.013	-0.109*	0.012	-0.041*	-0.010
	(0.017)	(0.064)	(0.011)	(0.024)	(0.010)
Fluency Tercile = 2 (Mid)*Lux	-0.014	-0.042		-0.048**	-0.012
	(0.015)	(0.065)		(0.019)	(0.014)
Lux	0.562***	0.653***	0.548***	0.595***	0.557***
	(0.018)	(0.062)	(0.013)	(0.023)	(0.014)
New	0.131***	0.131***	0.131***	0.131***	0.131***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Auction	0.066***	0.066***	0.066***	0.066***	0.066***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
Bed	0.117***	0.117***	0.117***	0.117***	0.117***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bath	0.106***	0.106***	0.106***	0.106***	0.106***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Has Parking	0.033***	0.033***	0.033***	0.033***	0.032***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Long Street	-0.011***	-0.010***	-0.011***	-0.011***	-0.011***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Major Street	-0.016**	-0.016**	-0.016**	-0.016**	-0.016**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Intercept	4.515***	2.915***	3.035***	3.659***	3.028***
	(0.431)	(0.454)	(0.455)	(0.461)	(0.454)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8687	0.8688	0.8687	0.8687	0.8687
N	958,408	958,408	958,408	958,408	958,408

Panel D: Luxury Home Interaction with Street Name Fluency (Matched Pairs Only)

	Five Fluency Measures				
	Englishness	Words	MSWord	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.000	0.149***	-0.001	0.001	0.001
	(0.003)	(0.035)	(0.003)	(0.005)	(0.003)
Fluency Tercile = 2 (Mid)	0.003	0.141***		0.003	0.004
	(0.003)	(0.035)		(0.005)	(0.003)
Fluency Tercile = 3 (High)*Lux	-0.004	-0.160***	0.007	-0.022	-0.018**
	(0.010)	(0.051)	(0.007)	(0.016)	(0.007)
Fluency Tercile = 2 (Mid)*Lux	-0.011	-0.119**		-0.020	-0.012**
	(0.009)	(0.055)		(0.012)	(0.006)
Lux	0.367***	0.518***	0.360***	0.380***	0.371***
	(0.016)	(0.051)	(0.013)	(0.019)	(0.013)
New Development	0.113***	0.113***	0.113***	0.113***	0.113***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Auction	0.036***	0.036***	0.036***	0.036***	0.036***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Bed	0.097***	0.097***	0.097***	0.097***	0.097***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Bath	0.097***	0.097***	0.097***	0.097***	0.097***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Has Parking	0.042***	0.042***	0.042***	0.042***	0.042***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Long Street	-0.010***	-0.009***	-0.010***	-0.010***	-0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Major Street	-0.016***	-0.014**	-0.016***	-0.016***	-0.016***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Intercept	4.950***	3.613***	4.320***	4.290***	3.770***
	(0.542)	(0.554)	(0.544)	(0.542)	(0.551)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8944	0.8946	0.8944	0.8944	0.8944
N	488,784	488,784	488,784	488,784	488,784

Table 5: Interaction of Street Name Fluency and Asian Buyers

This table reports coefficient estimates for the regression model using Equation 7 and the categorical fluency regression model using Equation 8 that include interactions between fluency measures and Asian buyers. *Asian* is a dummy equal to one if the surname of the buyer(s) is Asian, zero otherwise. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. Panel A reports the coefficient estimates for the baseline regression with Asian Buyer interaction. Panel B reports the coefficient estimates for the categorical fluency regression with Asian buyer interaction. The data is obtained from Australian Property Monitors. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Baseline Regression with Asian Buyer Interaction

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	0.000					
	(0.002)					
Englishness Group*Asian Buyer	0.000					
	(0.002)					
Words Group		0.000				
		(0.01)				
Words Group*Asian Buyer		0.001				
		(0.007)				
MS Word			0.002			
			(0.004)			
MS Word*Asian Buyer			0.003			
			(0.003)			
CommonName Group				-0.008***		
				(0.003)		
CommonName Group*Asian Buyer				0.003		
				(0.002)		
Syllable Group					-0.001	
					(0.003)	
Syllable Group*Asian Buyer					0.002	
					(0.003)	
Letters Group						-0.004**
						(0.002)
Letters Group*Asian Buyer						0.003**
						(0.001)
Asian Buyer	-0.010**	-0.013	-0.011***	-0.014***	-0.015**	-0.016***
	(0.004)	(0.021)	(0.002)	(0.004)	(0.007)	(0.004)
Intercept	3.293***	3.293***	3.295***	3.293***	3.294***	3.291***
	(0.485)	(0.484)	(0.485)	(0.485)	(0.483)	(0.485)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8493	0.8493	0.8493	0.8493	0.8493	0.8493
N	958,408	958,408	958,408	958,408	958,408	958,408

Panel B: Categorical Dummy Regression with Asian Buyer Interaction

	Five Fluency Measures				
	Englishness	Words	CommonName	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.001	0.121***	-0.014**	-0.001	-0.007**
	(0.003)	(0.030)	(0.006)	(0.007)	(0.003)
Fluency Tercile = 3 (High)*Asian Buyer	-0.001	-0.032	0.003	0.002	0.006**
	(0.004)	(0.028)	(0.005)	(0.006)	(0.003)
Fluency Tercile = 2 (Mid)	0.003	0.140***	-0.004	0.002	-0.002
	(0.004)	(0.030)	(0.007)	(0.006)	(0.004)
Fluency Tercile = 2 (Mid)*Asian Buyer	-0.003	-0.044	-0.003	-0.004	0.006
	(0.004)	(0.029)	(0.006)	(0.005)	(0.004)
Asian Buyer	-0.009***	0.022	-0.013***	-0.007	-0.014***
	(0.003)	(0.028)	(0.005)	(0.005)	(0.002)
Intercept	3.295***	3.328***	2.766***	2.465***	3.290***
	(0.485)	(0.492)	(0.487)	(0.491)	(0.484)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8493	0.8495	0.8493	0.8493	0.8493
N	958,408	958,408	958,408	958,408	958,408

Table 6: Street Name Fluency and New Home Interaction

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k New_{ij} + \beta_l New_{ij} * fluency\ measure_{\ ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstree$$
$$+ \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; New_{ij} is a dummy of 1 if the home is a new development, 0 otherwise, *fluency measure*$_{ij}$ denotes a street name measure for a home sold on street name *j*; *property char* are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. μ_s is the suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	0.000					
	(0.002)					
Englishness Group*New	0.000					
	(0.006)					
Words Group		-0.001				
		(0.009)				
Words Group*New		0.023				
		(0.017)				
MS Word			0.002			
			(0.004)			
MS Word*New			0.014			
			(0.009)			
CommonName Group				-0.008***		
				(0.002)		
CommonName Group *New				0.012		
				(0.008)		
Syllable Group					-0.002	
					(0.003)	
Syllable Group *New					0.027***	
					(0.009)	
Letters Group						-0.004***
						(0.001)
Letters Group*New						0.013***
						(0.004)
New	0.136***	0.069	0.132***	0.105***	0.081***	0.11***
	(0.016)	(0.051)	(0.008)	(0.025)	(0.021)	(0.012)
Intercept	14.373***	14.37***	14.393***	14.374***	14.378***	14.376***
	(0.086)	(0.084)	(0.083)	(0.083)	(0.083)	(0.085)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8481	0.8481	0.8481	0.8482	0.8481	0.8482
N	958,408	958,408	958,408	958,408	958,408	958,408

Table 7: Street Name Fluency and Royal Name Interaction

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k Royal_{ij} + \beta_l Royal_{ij} * fluency\ measure_{\ ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; *Royal*$_{ij}$ is a dummy of 1 if the street name is royalty related name (see list in Appendix), *fluency measure*$_{ij}$ denotes a street name measure for a home sold on street name *j*; *property char* are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. *Aussie* is dummy of 1 if Australian buyer, *Chinese* is dummy of 1 if Chinese buyer. *Investor* is dummy of 1 if home is a rental property. *Luxury* is dummy of 1 if price is in top quartile for the year. μ_s is the suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Royal Names and Other Buyer Characteristics Interactions

Dep Var: Log(Price)	(1)	(2)	(3)	(4)
Aussie	0.005***			
	(0.001)			
Royal*Aussie	0.008			
	(0.01)			
Chinese		-0.003		
		(0.002)		
Royal*Chinese		-0.020		
		(0.014)		
Investor			-0.003**	
			(0.001)	
Royal*Investor			-0.009	
			(0.008)	
Lux				0.39***
				(0.011)
Royal*Lux				0.035*
				(0.02)
Royal	0.031**	0.035***	0.034***	0.008
	(0.012)	(0.013)	(0.012)	(0.009)
Intercept	13.889***	13.52***	13.889***	13.964***
	(0.03)	(0.081)	(0.029)	(0.023)
Housing Characteristics	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8482	0.8482	0.8482	0.8826
N	958,408	958,408	958,408	958,408

Panel B: Royal Names and Fluency Measures Interactions

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Englishness Group		0.000					
		(0.002)					
Royal*Englishness Group		0.008					
		(0.015)					
Words Group			0.002				
			(0.009)				
Royal*Words Group			-0.032				
			(0.032)				
MS Word				0.002			
				(0.004)			
Royal*MS Word				0.009			
				(0.029)			
CommonName Group					-0.007***		
					(0.002)		
Royal*CommonName					-0.044***		
					(0.016)		
Syllable Group						-0.001	
						(0.003)	
Royal*Syllable Group						-0.044*	
						(0.023)	
Letters Group							-0.003**
							(0.002)
Royal*Letters Group							-0.024
							(0.015)
Royal	0.033***	0.01	0.122	0.024	0.163***	0.145**	0.092**
	(0.012)	(0.039)	(0.092)	(0.025)	(0.042)	(0.06)	(0.038)
Intercept	13.518***	13.89***	13.513***	13.517***	13.544***	13.519**	13.527**
	(0.081)	(0.029)	(0.084)	(0.082)	(0.082)	(0.081)	(0.081)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8482	0.8482	0.8482	0.8482	0.8482	0.8482	0.8482
N	958,408	958,408	958,408	958,408	958,408	958,408	958,408

Table 8: Street Name Fluency and Trendy Words based on Google Search

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k GTrend_{ij} + \beta_l GTrend_{ij} * fluency\ measure\ _{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; $GTrend_{ij}$ is a dummy of 1 in the year following the street name's peak search month based on google trends data during our sample period (from 2004 when google trends started collecting search data), $fluency\ measure_{ij}$ denotes a street name measure for a home sold on street name *j*; *property char* are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. μ_s is the suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2004 to June 2016. The data is obtained from Australian Property Monitors. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GTrend	0.003**	0.013***	0.001	0.004**	0.015*	0.007	0.004
	(0.002)	(0.004)	(0.033)	(0.002)	(0.008)	(0.007)	(0.004)
		(0.002)					
Englishness Group		0.001					
		(0.002)					
Englishness Group*GTrend		-0.004**					
		(0.002)					
Words Group			-0.002				
			(0.009)				
Words Group*GTrend			0.001				
			(0.011)				
MS Word				0.004			
				(0.004)			
MS Word*GTrend				-0.003			
				(0.003)			
CommonName Group					-0.007***		
					(0.002)		
CommonName Group *GTrend					-0.004		
					(0.003)		
Syllable Group						0.000	
						(0.003)	
Syllable Group *GTrend						-0.002	
						(0.003)	
Letters Group							-0.003*
							(0.002)
Letters Group*GTrend							0.000
							(0.002)
Intercept	14.329***	14.326***	14.336***	14.329***	14.349***	14.328***	14.334***
	(0.102)	(0.102)	(0.103)	(0.103)	(0.102)	(0.102)	(0.100)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8462	0.8462	0.8462	0.8462	0.8463	0.8462	0.8462
N	762,484	762,484	762,484	762,484	762,484	762,484	762,484

Table 9: Persistence of Buyers in Fluency Premium and Fluency Measures

The sample includes home buyers that have multiple purchases under their names. Homes with missing owner name, with only the surname registered and no first name, couples with the same surname or with common surname combinations (e.g. Kaur; Singh or Wang; Zhang), company owners (denoted with suffix Pty Ltd) and churches are excluded. First we group multiple home owners in three equal groups (two for Word Group and MS Word) based on their first home purchases' fluency premium or raw measure. We then calculate each home owner's fluency measure over subsequent purchases as their average fluency measure for subsequent purchases (i.e. the fluency measure of their 2nd purchase if they only made 2 purchases and the average fluency measure of their 2nd and 3rd purchase if they made 3 purchases). We then take the mean of the average owner fluency measures for each group. Fluency premiums are calculated as the residual of the baseline hedonic model with all fluency explanatory variables. The table reports the mean owner first purchase fluency measure and mean owner subsequent purchase fluency measures. The difference between the high and low groups for subsequent purchases is also reported. Panel A to G report sample statistics for fluency premium group, raw Englishness score group, raw words group, MS Word group, raw popularity group and raw syllable two-way sorts, respectively. *t*-stat is in parentheses.

Panel A: Fluency Premium Groups

Purchase Time	Agg Ave	1 (Low Group)	2	3 (High Group)	High–Low	T-stat	N
First Purchase	-0.012	-0.216	-0.015	0.194	0.41	(217.01)***	33,663
2nd	0.006	-0.021	0.001	0.04	0.06	(18.58)***	26,336
3rd	0.003	-0.013	-0.006	0.029	0.043	(6.25)***	5,848
Subsequent Purchases (2nd and after)	0.006	-0.105	-0.007	0.109	0.213	(19.53)***	39,934
Subsequent – 1st	0.013	0.112	0.008	-0.084	-0.196	(-107.92)***	

Panel B: Raw Englishness Score Groups

Purchase Time	Agg Ave	1 (Low Group)	2	3 (High Group)	High–Low	T-stat	N
First Purchase	1.663	-3.723	2.032	6.665	10.388	(273.69)***	33,663
2nd	1.511	1.212	1.563	1.758	0.547	(7.36)***	26,336
3rd	1.619	1.365	1.570	1.924	0.558	(3.49)***	5,848
Subsequent Purchases (2nd and after)	1.560	-0.982	1.801	3.990	4.972	(7.37)***	39,934
Subsequent – 1st	-0.070	2.750	-0.235	-2.719	-5.469	(-123.56)***	

Panel C: Raw Words Group

Purchase Time	Agg Ave	1 (Low Group)	2 (High Group)	High–Low	T-stat	N
First Purchase	1.053	1	2.046	1.046	(891.71)***	33,663
2nd	1.059	1.056	1.102	0.046	(6.51)***	26,336
3rd	1.059	1.057	1.104	0.047	(3.09)***	5,848
Subsequent Purchases (2nd and after)	1.059	1.031	1.535	0.504	(6.62)***	39,934
First Purchase	0.02	0.06	-0.433	-0.492	(-90.6)***	

Panel D: MS Word Group

Time	Agg Ave	1 (Low Group)	2 (High Group)	High–Low	T-stat	N
First Purchase	0.298	0.000	1.000	1.000	-	33,663
2nd	0.294	0.283	0.321	0.038	(6.16)***	26,336
3rd	0.293	0.289	0.303	0.013	(1.04)***	5,848
Subsequent Purchases (2^{nd} and after)	0.295	0.156	0.628	0.472	(5.97)***	39,934
Subsequent – 1st	0.015	0.215	-0.365	-0.58	(-135.35)***	

Panel E: Raw Popularity Group

Time	Agg Ave	1 (Low Group)	2	3 (High Group)	High–Low	T-stat	N
First Purchase	89.673	6.346	45.134	217.894	211.548	(158.43)***	33,663
2nd	84.746	78.319	81.14	94.648	16.329	(8.88)***	26,336
3rd	88.778	83.244	87.946	94.799	11.555	(2.85)***	5,848
Subsequent Purchases (2^{nd} and after)	85.894	46.469	66.154	149.601	103.132	(9.14)***	39,934
Subsequent – 1st	-2.368	41.251	21.316	-68.592	-109.843	(-95.06)***	

Panel F: Raw Syllables Group

Time	Agg Ave	1 (Low Group)	2	3 (High Group)	High–Low	T-stat	N
First Purchase	2.218	1.000	2.000	3.195	2.195	(369.73)***	33,663
2nd	2.241	2.182	2.235	2.283	0.101	(6.64)***	26,336
3rd	2.234	2.163	2.217	2.302	0.139	(4.38)***	5,848
Subsequent Purchases (2^{nd} and after)	2.24	1.648	2.127	2.697	1.049	(7.27)***	39,934
Subsequent – 1st	0.027	0.651	0.162	-0.496	-1.148	(-107.71)***	

Panel G: Raw Letters Group

Time	Agg Ave	1 (Low Group)	2	3 (High Group)	High–Low	T-stat	N
First Purchase	7.121	5.274	7.000	9.232	3.957	(261.15)***	33,663
2nd	7.156	7.059	7.143	7.273	0.215	(7.3)***	26,336
3rd	7.175	7.089	7.202	7.258	0.169	(2.78)***	5,848
Subsequent Purchases (2^{nd} and after)	7.165	6.259	7.090	8.161	1.902	(7.81)***	39,934
Subsequent – 1st	0.051	1.037	0.109	-1.054	-2.091	(-113.09)***	

Table 10: Buyer Fluency Persistence and Housing Prices

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k fluency_{ij} + \beta_l Hfluency_{ij} + \beta_m fluency_{ij} * Hfluency_{ij} * Next\ Buy_i + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; $fluency_{ij}$ denotes one of the six fluency street name measures for a home sold on street name *j*; $Hfluency_{ij}$ is a dummy of 1 if a buyer's first purchase has a street name in the top third of street name *fluency* measures, 0 otherwise. $Next\ Buy_i$ is a dummy of 1 if the purchase is the second or subsequent purchase made by the buyer, 0 otherwise. Buyers are tracked by their owner name. $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. *Long Street* is a dummy of 1 if the street on which the home is located is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise. *Major Street* is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise. μ_s is the suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	-0.004**					
	(0.002)					
HEnglishness Buyer	-0.014***					
	(0.003)					
Englishness*HEnglishness Buyer*Next Buy	0.01***					
	(0.001)					
Words Group		0.002				
		(0.01)				
HWords Buyer		-0.005				
		(0.008)				
Words*HWords Buyer*Next Buy		0.004				
		(0.003)				
MS Word Group			-0.002			
			(0.004)			
HMS Word Buyer			0.001			
			(0.002)			
MS Word*HMS Word Buyer*Next Buy			0.020***			
			(0.005)			
CommonName Group				-0.01***		
				(0.003)		
HCommonName Buyer				-0.012***		
				(0.003)		
CommonName*HCommonName Buyer*Next Buy				0.012***		
				(0.001)		
Syllable Group					-0.003	
					(0.004)	
HSyllable Buyer					-0.008***	
					(0.003)	
Syllable*HSyllable Buyer*Next Buy					0.011***	
					(0.002)	
Letters Group						-0.004**
						(0.002)
HLetters Buyer						-0.009***
						(0.003)
Letters*HLetters Buyer*Next Buy						0.011***
						(0.002)
Intercept	12.731***	12.727***	12.735***	12.744***	12.739***	12.745***
	(0.024)	(0.036)	(0.023)	(0.024)	(0.023)	(0.023)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Additional Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8018	0.8018	0.8018	0.8019	0.8018	0.8018
N	958,408	958,408	958,408	958,418	958,408	958,408

Table 11: Street Name Fluency for Homes with Multiple Street Names

This table reports coefficient estimates for the following hedonic model across individual housing prices for homes located at the intersection of multiple street names:

$$\ln(P_{ijst}) = \alpha_t + \beta_1 minfluency_{ij} + \beta_2 maxfluency_{ij} + property\ char_i + \beta_l longstreet_{ij} + \beta_m majorstreet_{ij} + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ist})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; for homes located on at the intersection of multiple streets, $minfluency_{ij}$ denotes the minimum of the fluency measures of all the street names; $maxfluency_{ij}$ denotes the maximum of fluency measures of all the street names; $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. *Long Street* is a dummy of 1 if the street on which the home is located is more than 1 kilometre (0.62 miles) in the zip code, zero otherwise. *Major Street* is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise. μ_s are suburb location specific fixed effects; γ_t are year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. Homes have multiple street names if their geocode matches to two or more addresses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Min Englishness Group	-0.005					
	(0.016)					
Max Englishness Group	-0.043*					
	(0.024)					
Min Words Group		-0.036				
		(0.025)				
Max Words Group		0.17***				
		(0.056)				
Min MSWord			-0.014			
			(0.077)			
Max MSWord			-0.047			
			(0.042)			
Min Popname Group				-0.01		
				(0.018)		
Max Popname Group				0.076*		
				(0.04)		
Min Syllable Group					-0.032	
					(0.023)	
Max Syllable Group					-0.048	
					(0.04)	
Min Letters Group						-0.028***
						(0.01)
Max Letters Group						0.009
						(0.009)
New	0.064***	0.068***	0.068***	0.067***	0.065***	0.067***
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)
Auction	0.028***	0.029***	0.028***	0.028***	0.029***	0.03***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Bed	0.21***	0.214***	0.213***	0.216***	0.216***	0.214***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Bath	0.102***	0.1***	0.103***	0.099***	0.098***	0.099***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
Parking	0.056***	0.054***	0.053***	0.054***	0.056***	0.054***
	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)
Major Street	-0.126*	-0.135*	-0.131*	-0.131*	-0.133*	-0.137*
	(0.072)	(0.078)	(0.068)	(0.073)	(0.072)	(0.079)
Intercept	13.489***	12.914***	13.293***	13.183***	13.466***	13.388***
	(0.135)	(0.211)	(0.118)	(0.141)	(0.154)	(0.127)
Other Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Clustered SE	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.866	0.865	0.866	0.865	0.866	0.865
Number of Observations	5,989	5,989	5,989	5,989	5,989	5,989

Table 12: Robustness Check on Street Centrality

This table reports coefficient estimates for the following hedonic model across the full sample of individual housing prices:

$$\ln(P_{ijst}) = \alpha_t + \beta_k fluency_{ij} + property\ char_i + \beta_c Street\ Centrality + \mu_s + \gamma_t + \tau_t + \varepsilon_{it}$$

Where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; *fluency*$_{ij}$ denotes one of the six fluency street name measures for a home sold on street name *j*; $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. To measure *Street Centrality*, first we collect geospatial data of streets in Sydney from openstreetmap.org We then apply network analysis to the street where every street intersection (or end of a street if a dead end) is a node and the streets to each node being edges. For every node, we calculate its degree centrality as:

CD(node) = deg(node)/E

Where *deg(node)* is the number of edges that the node has. E is the number of edges in the entire network. *Street Centrality* is measured as the sum of *CD(node)* for all intersections on a street, standardized. *longstreet* is a dummy of 1 if the home's street is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise. *majorstreet* is a dummy of 1 if the home's street is in the top two longest streets in the zip code, zero otherwise. μ_s is the suburb location specific fixed effects; γ_t is year/quarter fixed effects; and τ_t is a monthly time trend. Other control variables are described in Appendix 1. Street names are separated from their street type (e.g. highway, road or street) and any apostrophes are removed (e.g. O'Dea) to calculate the fluency measures. Section 2.2 describes how we construct the street name fluency measures. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. Panel A reports the correlation matrix of housing price, fluency measures and street centrality measures. Panel B reports regression results for linear fluency measures and street centrality. Panel C reports regression results for categorical fluency measures and street centrality. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Correlation Statistics

	Price	Street Centrality	Englishness	Words	MS Word	Common Name	Syllable	Long Street	Major Street
Price	1.00								
Street Centrality	-0.06	1.00							
Englishness	-0.01	0.04	1.00						
Words	-0.03	-0.15	-0.06	1.00					
MS Word	-0.02	0.05	0.16	0.26	1.00				
CommonName	-0.01	-0.06	0.19	0.19	0.16	1.00			
Syllable	-0.03	-0.09	0.06	0.27	0.15	0.48	1.00		
Long Street	-0.05	0.61	0.02	-0.16	0.04	-0.07	-0.11	1.00	
Major Street	0.00	0.50	0.02	-0.18	0.02	-0.07	-0.08	0.47	1.00

Panel B: Linear Fluency Measure Regression with Centrality

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)	(6)
Englishness Group	0.000					
	(0.002)					
Words Group		0.005				
		(0.009)				
MS Word			0.003			
			(0.004)			
CommonName Group				-0.007***		
				(0.002)		
Syllable Group					-0.001	
					(0.003)	
Letters Group						-0.003*
						(0.002)
Street Centrality	-0.012***	-0.012***	-0.012***	-0.012***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intercept	13.74***	13.726***	13.74***	13.755***	13.743***	13.750***
	(0.091)	(0.093)	(0.091)	(0.091)	(0.09)	(0.09)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8499	0.8499	0.8499	0.8499	0.8499	0.8499
N	932,591	932,591	932,591	932,591	932,591	932,591

Panel C: Categorical Fluency Measure Regression with Street Centrality

	Five Fluency Measures				
	Englishness	Words	CommonName	Syllable	Letters
Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Fluency Tercile = 3 (High)	0.000	0.131***	-0.013**	-0.001	-0.006*
	(0.003)	(0.034)	(0.005)	(0.007)	(0.003)
Fluency Tercile = 2 (Mid)	0.002	0.145***	-0.005	0.000	0.000
	(0.003)	(0.034)	(0.006)	(0.006)	(0.004)
Street Centrality	-0.012***	-0.012***	-0.012***	-0.012***	-0.012***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intercept	13.739***	14.175***	14.334***	14.022***	14.021***
	(0.091)	(0.105)	(0.095)	(0.114)	(0.114)
Housing Characteristics	Yes	Yes	Yes	Yes	Yes
Suburb Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year/Quarter Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes	Yes
Cluster Error	Suburb	Suburb	Suburb	Suburb	Suburb
Adj Rsq	0.8499	0.8501	0.8499	0.8499	0.8499
N	932,591	932,591	932,591	932,591	932,591

Table 13: Suburb Name Change Regression

The table reports coefficient estimates for the diff-in-diff model to test for the effect of suburb name changes in two areas. The general regression is:

$$\ln(P_{ijst}) = \alpha_t + \beta_k Post_t + \beta_l * Treatment\ Area_i + \beta_l Post_t * Treatment\ Area_i + property\ char_i + \mu_s + Y_t + \varepsilon_{it}$$

Where *Post* is a dummy of 1 if a sales is made after the announcement date, 0 otherwise. *Treatment Area* is a dummy of 1 if a home sells in the area where a suburb name change occurs, 0 otherwise. $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type, and land area size. μ_s is the suburb location specific fixed effects; Y_t are year fixed effects. Panel A reports the diff in diff regressions for the suburb name change from Harbord to Freshwater using the official name change date, approved name change date by local council and a false date two years before the official name change. Panel B reports the diff-in-diff regression for when a section of a suburb changed suburb names from Moorebank to Wattle Grove using the name change date as recorded by the NSW Government Spatial Services and a false date two years before the recorded name change. Sales in the month of the announcement date are removed. The sample comprises home sales in the Sydney metropolitan area from January 2000 to June 2016. The data is obtained from Australian Property Monitors. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Harbord to Freshwater Suburb Name Change

Dep Var: Log(Price)	(1)	(2)	(3)
Post	0.112***	0.159***	0.018
	(0.036)	(0.033)	(0.045)
Treatment Area		-0.587***	-0.577***
		(0.03)	(0.033)
Post*Treatment Area		0.029***	0.002
		(0.009)	(0.007)
New Development		0.094***	0.124***
		(0.03)	(0.033)
Auction		0.099***	0.14***
		(0.013)	(0.017)
Bed		0.135***	0.124***
		(0.019)	(0.019)
Bath		0.176***	0.179***
		(0.01)	(0.012)
Parkings		0.026	0.009
		(0.027)	(0.017)
Intercept		-0.016	-0.009
		(0.017)	(0.017)
Sample Time	Two years before and after official suburb name change on Jan 12 2008	Two years before and after official name change Jan 12 2008	Two years before and after false date Jan 12 2006
Sample Area	Harbord/Freshwater Only	Freshwater and Northern Beaches Local Government Area	Freshwater and Northern Beaches Local Government Area
Area Fixed Effects	No	Yes	Yes
Year Fixed Effects	No	Yes	Yes
Clustered Errors	None	Suburb	Suburb
Adj Rsq	0.0096	0.7535	0.7456
N	905	16,256	14,567

Panel B: Moorebank to Wattle Grove part Suburb Name Change

Dep Var: Log(Price)	(1)	(2)	(3)
Post	0.146***	0.029	0.032
	(0.041)	(0.035)	(0.054)
Treatment Area		-0.033*	0.008
		(0.009)	(0.018)
Post*Treatment Area		0.022***	-0.081
		(0.001)	(0.038)
New Development		0.103**	0.021
		(0.017)	(0.011)
Auction		-0.013	-0.026
		(0.026)	(0.058)
Bed		0.062***	0.091**
		(0.005)	(0.01)
Bath		0.082***	0.081**
		(0.007)	(0.008)
Parkings		0.046*	0.044**
		(0.014)	(0.005)
Intercept	13.065***	13.049***	12.16***
	(0.032)	(0.064)	(0.011)
Sample Time	Two years before and after official suburb name change on Apr 12 2012	Two years before and after official name change Apr 12 2012	Two years before and after false date Apr 12 2010
Sample Area	Changed Suburb Area Only	Moorebank and Wattle Grove	Moorebank and Wattle Grove
Area Fixed Effects	No	Yes	Yes
Year Fixed Effects	No	Yes	Yes
Cluster Error	None	Suburb	Suburb
Adj Rsq	0.1854	0.6950	0.6450
N	53	995	989

Appendix 1
List of Housing Characteristic Variables

Variable	Description
Asian Buyer	1 if the home buyer has an Asian surname, 0 otherwise.
Bed	Number of beds.
Bath	Number of bathrooms.
Auction	1 if the home was sold at auction, 0 otherwise.
New	1 if the home was a new development, 0 otherwise.
Has parking	1 if home has one or more parking spots, 0 otherwise.
Long street	1 if the home's street is more than 1 kilometer (0.62 miles) in the zip code, zero otherwise.
Major street	1 if the home's street is in the top two longest streets in the zip code, zero otherwise.
Street type dummies	1 if a certain street type (e.g. avenue, highway, lane, street, road, etc.), 0 otherwise.
Housing type dummies	1 if a certain housing type (e.g. apartment/condominium, house, semi, studio, townhouse, villa, etc.), 0 otherwise.
Area size	Land area size of home (square meters).
HasAirConditioning	1 if home has air conditioning, 0 otherwise.
HasAlarm	1 if home has alarm system, 0 otherwise.
HasBalcony	1 if home has balcony, 0 otherwise.
HasBarbeque	1 if home has barbeque, 0 otherwise.
HasBeenRenovated	1 if home has been renovated, 0 otherwise.
HasBilliardRoom	1 if home has billiard room, 0 otherwise.
HasCourtyard	1 if home has courtyard, 0 otherwise.
HasEnsuite	1 if home has ensuite, 0 otherwise.
HasFamilyRoom	1 if home has family room, 0 otherwise.
HasFireplace	1 if home has fire place, 0 otherwise.
HasGarage	1 if home has garage, 0 otherwise.
HasHeating	1 if home has heating, 0 otherwise.
HasInternalLaundry	1 if home has internal laundry, 0 otherwise.
HasLockUpGarage	1 if home has lock up garage, 0 otherwise.
HasPolishedTimberFloor	1 if home has polished timber floors, 0 otherwise.
HasPool	1 if home has swimming pool, 0 otherwise.
HasRumpusRoom	1 if home has rumpus room, 0 otherwise.
HasSauna	1 if home has sauna, 0 otherwise.
HasSeparateDining	1 if home has separate dining room, 0 otherwise.
HasSpa	1 if home has spa, 0 otherwise.
HasStudy	1 if home has study room, 0 otherwise.
HasSunroom	1 if home has sunroom, 0 otherwise.
HasTennisCourt	1 if home has tennis court, 0 otherwise.
HasWalkInWardrobe	1 if home has walk in wardrobe, 0 otherwise.
View dummies	1 if home has a certain view (e.g. bush, city, district, harbour, ocean, park, river, etc.), 0 otherwise.

Appendix 2
Examples of Street Names and Fluency Scores

Street Name	Englishness Group	Words Group	MS Word	CommonName Group	Syllable Group	Letters Group	Total
Low Fluency (Score <= 6)							
AVENUE OF OCEANIA	1	1	0	1	1	1	5
SIR JOHN JAMISON	1	1	0	1	1	1	5
SIR WARWICK FAIRFAX	1	1	0	1	1	1	5
ABBE RECEVEUR	1	2	0	1	1	1	6
LILLI PILLI POINT	2	1	0	1	1	1	6
YUNGA BURRA	1	2	0	1	1	1	6
Medium Fluency (Score = 11)							
ABIGAIL	2	3	0	2	2	2	11
BANDICOOT	2	3	1	2	2	1	11
CHARLIE	3	3	0	1	2	2	11
EXCELSIOR	2	3	1	3	1	1	11
GARFIELD	2	3	0	3	2	1	11
HIGHLAND RIDGE	3	2	1	2	2	1	11
High Fluency (Score = 16)							
COOK	3	3	1	3	3	3	16
HOOD	3	3	1	3	3	3	16
SPRING	3	3	1	3	3	3	16
VIEW	3	3	1	3	3	3	16
WHITE	3	3	1	3	3	3	16
YOUNG	3	3	1	3	3	3	16

Appendix 3
Top 20 Street Names by Sales

Street Name	Frequency
PACIFIC	6,667
VICTORIA	6,109
PARK	4,209
RAILWAY	3,309
GEORGE	3,005
WILLIAM	2,876
STATION	2,600
CAMPBELL	2,524
PITTWATER	2,522
ALBERT	2,439
OCEAN	2,320
BRIDGE	2,211
CHURCH	2,201
PRINCES	2,189
LIVERPOOL	2,173
FOREST	2,132
WENTWORTH	2,094
ANZAC	2,074
ELIZABETH	2,059
HAMPDEN	1,961
All Streets (Top 20)	57,674

Appendix 4
List of Royal Names

List of Royal Names (28 in total)	*Number in Sample*	*Percent in Royal Names*	*Percent in Sample*
PRINCES	2,190	16.673%	0.229%
KING	1,851	14.092%	0.193%
QUEEN	1,650	12.562%	0.172%
CROWN	1,159	8.824%	0.121%
QUEENS	1,153	8.778%	0.120%
KINGS	675	5.139%	0.070%
KING GEORGES	604	4.598%	0.063%
PRINCE	598	4.553%	0.062%
DUKE	406	3.091%	0.042%
PRINCESS	352	2.680%	0.037%
QUEEN VICTORIA	317	2.413%	0.033%
OLD PRINCES	284	2.162%	0.030%
LORD	272	2.071%	0.028%
PRINCE EDWARD	255	1.941%	0.027%
BUCKINGHAM	224	1.705%	0.023%
PALACE	177	1.348%	0.018%
PRINCE ALBERT	142	1.081%	0.015%
PRINCE ALFRED	134	1.020%	0.014%
PRINCE CHARLES	114	0.868%	0.012%
KING EDWARD	106	0.807%	0.011%
PRINCE EDWARD PARK	101	0.769%	0.011%
ROYAL	101	0.769%	0.011%
GREAT BUCKINGHAM	82	0.624%	0.009%
ROYAL GEORGE	67	0.510%	0.007%
KING WILLIAM	38	0.289%	0.004%
DUCHESS	36	0.274%	0.004%
PRINCESS MARY	28	0.213%	0.003%
KING GEORGE	19	0.145%	0.002%
Total	**13,135**	**100%**	**1.371%**

Internet Appendix to

Street Name Fluency and Housing Prices

Figure IA1: Dual Street Name Home Example

The figure shows the layout of streets surronding a housing unit 416 Punchbowl Road, Belfield, which is also 29 Bazentin St, Belfield, as seen from the map. Source: Google Maps

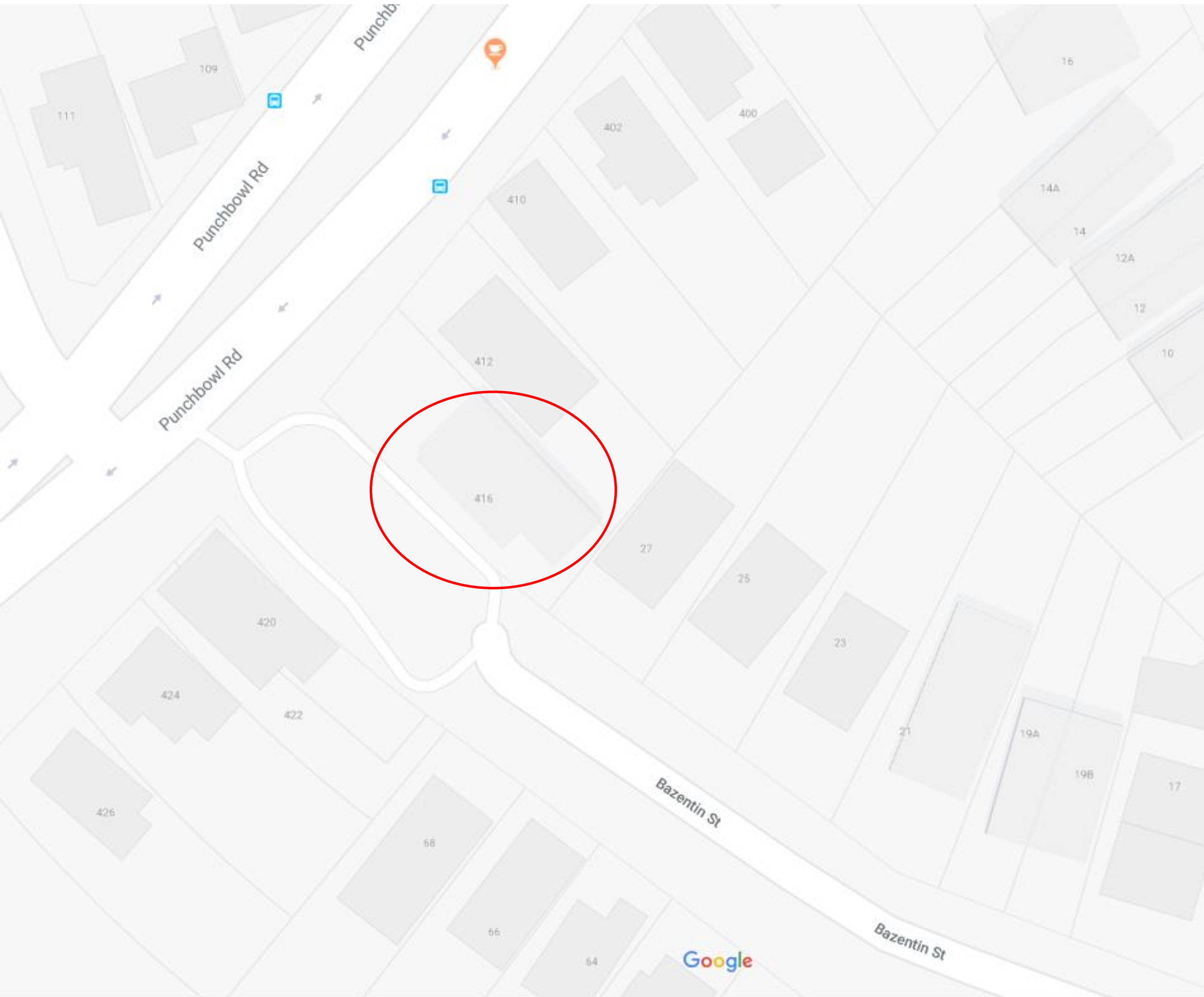


Table IA1: Street Name Change Regression

Panel A reports univariate differences in price and fluency measures before and after the street name change (T-stats in parenthesis). Panel B reports estimation result of the following hedonic model using the sample of individual housing price sales on streets where the street name has changed:

$$\ln(P_{ijst}) = \alpha_t + \beta_k fluency_{ij} + \beta_m bed_{ij} + \mu_{street} + Y_t + \varepsilon_{it}$$

where $ln(P_{ijst})$ denotes the natural logarithm of housing prices for sale *i* on street name *j* in suburb *s* at time *t*; $fluency_{ij}$ denotes one of the five fluency street name measures (excluding Words Group as all streets only had one word) for a home sold on street name *j*; bed_{ij} is the number of bedrooms; μ_{street} are street fixed effects, Y_t are year fixed effects. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Univariate Results

	Sales Price (AUD$'00,000)	Englishness Group	MS Word	CommonName Group	Syllable Group	Letters Group
Before	4.929	1.733	0.400	3.000	2.067	2.067
After	5.366	1.406	0.063	2.813	1.906	2.688
After − Before	0.437	-0.327*	-0.338***	-0.188	-0.160*	0.621**
T-stat	(0.712)	(-1.718)	(-3.092)	(-1.219)	(-1.800)	(2.524)
N	47	47	47	47	47	47

Panel B: Hedonic Model

Dep Var: Log(Price)	(1)	(2)	(3)	(4)	(5)
Englishness Group	-0.036				
	(0.110)				
MS Word		-0.11			
		(0.145)			
CommonName Group			0.167		
			(0.102)		
Syllable Group				0.167	
				(0.102)	
Letters Group					0.056
					(0.094)
Bed	0.05	0.04	0.108	0.108	0.072
	(0.043)	(0.049)	(0.09)	(0.09)	(0.055)
Intercept	12.173***	12.227***	11.533***	11.701***	11.878***
	(0.055)	(0.13)	(0.405)	(0.31)	(0.513)
Additional Housing Characteristics	No	No	No	No	No
Street Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Monthly Time Trend	No	No	No	No	No
Cluster Error	Street	Street	Street	Street	Street
Adj Rsq	0.7873	0.7895	0.7973	0.7973	0.7928
N	47	47	47	47	47