Melting Pot or Salad Bowl: Cultural Distance and Housing Prices*

Maggie R. Hu
School of Banking and Finance, UNSW Business School, Australia
Tel: 61-2-93855623
Email: m.hu@unsw.edu.au

Adrian D. Lee
Finance Discipline Group, University of Technology Sydney
Tel: 61-2-95147765
Email: Adrian.Lee@uts.edu.au

Latest draft: January 20, 2016

*We thank Doug Foster, Marco Navone, Talis Putnins, Dugald Tinch, discussants and participants at the 2015 Asian Global Real Estate Summit and 2015 Australasian Financing and Banking Conference, and seminar participants at the University of Newcastle, University of Technology Sydney, University of Tasmania, UNSW Real Estate Symposium, for helpful comments and feedback. Hu thanks Centre for Applied Economic Research (CAER), University of New South Wales for supports. Lee acknowledges funding from the University of Technology Sydney Business School Research Grant.

Melting Pot or Salad Bowl: Cultural Distance and Housing Prices

Abstract:

We investigate whether the cultural distance between a homebuyer's ethnicity and that of the neighborhood affects a homebuyer's location choice of the home and the home's transaction price. Utilizing individual home sales data of culturally diverse Sydney, Australia, we find homebuyers are more likely to choose neighborhoods with a shorter cultural distance, and are willing to pay a premium for homes in those locations, consistent with homebuyers' preference for similar cultures. Home culture preference is stronger for ethnicities from recent migration waves, particularly Asia. Our results are robust to endogeneity and self-selection concerns. The findings have implications on the role of cultural and demographic shifts on housing prices.

Key words: *culture, cultural distance, hedonic pricing, housing price, information friction, home culture preference*
JEL Classification Code: *R00, O18, P22, R21*

1. Introduction

We analyze to what extent the cultural background of homebuyers affects housing transactions in Sydney, Australia's largest and most culturally diverse city.[1] In recent times the booming housing prices in cities like London, Singapore, Sydney[2] and Toronto have attracted global media attention. Such cities with culturally diverse citizens are becoming the norm yet little is known about how the interaction between a person and their host country's culture affects their personal and financial decisions, particularly in housing markets. In this ethnic stew with residents of different cultures and religions, will the "melting pot" theory[3] of culture assimilation dominate or will the "salad bowl" analogy holds where the immigrants choose to retain their cultural identities while integrating into a new society?

In our study we focus on the housing market decision, and examine whether the cultural background of the homebuyer, and inter-ethnicity cultural difference of the neighborhood influence their housing market decisions. Such a question is of paramount importance to diverse cities like Sydney where over half of the population are from an overseas background.[4] As housing constitutes around 60 to 70 percent of the total wealth of the household sector, "It is an extremely important asset class for most people", as emphasized by former Reserve Bank of Australia Governor Ian Macfarlane (2004).

[1] Throughout the 100 years since the first National Census in 1911, immigrants have made up a large component of the Australian population. Historically, the majority of migration has come from Europe, however, there are increasingly more Australians who were born in Asia and other parts of the world. This pattern of migration is evident in the make-up of the richly diverse society which has been recorded in the 2011 Census.

[2] Figure 4 shows the number and dollar volume of home sales by ethnic region of buyers.

[3] Melting pot is a term coined by a U.S. immigrant, Israel Zangwill. In the melting pot metaphor, people are combined so as to lose their discrete identities and yield a final product of uniform consistency and flavor. In the "salad bowl" metaphor, each culture retains its own distinct qualities, like the different ingredients in the salad. Different ingredients are all mixed together to make one dish, yet each ingredient also retains its own characteristics in the bowl.

[4] Based on the statistics from 2011 Australian Bureau of Statistics Census, 57% of Sydney urban respondents report non-Australian or British origins.

Culture is defined as "the collective programming of the mind that distinguishes the members of one group or category of people from another" (Hofstede (2001)). Despite culture being an important priced factor, the literature is scant on its effect on housing prices. Recent work suggests that buyers prefer to live with people from the same background or ethnicity. For example Li (2014) finds that neighborhoods with higher ethnicity concentration have higher housing prices in Toronto, Canada, and he argues it is due to buyers valuing social interactions with their own ethnicity more than with others. Wong (2013) finds preferences for own-ethnicity are inverted U-shaped so that 'once a neighbourhood has enough own-ethnic-neighbours, they would rather add a new neighbour from other groups'. Culture-related barriers such as language may also play an important role in explaining the nexus between housing prices and immigrant inflows. For example Fischer (2012) finds that non-common language immigrants to Switzerland are less price-sensitive to housing price changes than common language immigrants, suggesting language acting as a friction for integration.

Whether and how cultural distance affects housing transaction prices is ambiguous ex-ante. There are two competing hypotheses we evaluate in this study. On one hand, the ***information friction hypothesis*** posits that homebuyers who are more culturally distant from the culture of a property's neighborhood face higher search costs and greater information friction to access the local property market. As it is harder for them to arrive at the efficient price in the housing market, they may be forced to pay a higher price for their homes. On the other hand, the ***home culture preference hypothesis*** argues that homebuyers prefer home locations with greater cultural similarity to their culture of origin, and are willing to pay a premium for homes in those locations. There is evidence that people in general prefer to live in a

community or neighborhood with similar cultural background (e.g. Saiz (2007)). Ethnic clusters would be developed as ethnic group friends and family tend to live together in close proximity.

To empirically test the effect on the role of cultural distance on housing prices, we employ a transaction-level residential housing market dataset of the Sydney Metropolitan Area from 2006 to 2013. We adopt the widely used cultural framework from Hofstede (2001) as the measure of culture value, whereby the cultural distance is defined as the weighted average Euclidian distance between the culture value of the homebuyer's original ethnicity and that of the neighborhood of the home, using the ethnicity proportion as the weight. Homebuyer's ethnicity is inferred from their surnames based on a hand collected database of surnames and ethnicity from various internet sources, including surnamedb.com and Wikipedia. Suburb level demographic and ethnicity data for each suburb in Sydney is from Australian Bureau of Statistics Census 2006 and 2011, including a comprehensive list of variables on population size, ethnic composition by residents' birthplace and by residents' ancestry, etc. To ensure the robustness of our results, we construct four versions of the homebuyer-suburb pair's cultural distance measures using either ancestry or birthplace of the neighborhood and four-dimension or six-dimension Hofstede (2001) cultural scores.

The central finding is that cultural distance plays a determinant role on homebuyers' location choice and home transaction price. We find that homebuyers are more likely to choose homes located in neighborhoods which are more culturally similar to their own culture of origin. Further, they are willing to pay a premium for homes in locations with shorter cultural distance. The greater the cultural distance between the homebuyer's ethnicity and the ethnic composition of the home's neighborhood, the lower is the housing price in that transaction, ceteris paribus.

Specifically, if the cultural distance between a homebuyer with the neighborhood decreases by 1 unit, approximately the difference between the average Australian and average Chinese buyer's cultural distance in our sample, the housing price rises by 1.1 percent or A$7,382 based on the sample mean sales price of A$671,100. This amount is both economically sizeable and statistically significant, suggesting that homebuyers are willingly to pay higher prices for homes in neighborhoods which are closer to their culture of origin. This is in support of the home culture preference hypothesis, and confirms the salad bowl theory of culture integration.

Our hedonic housing price regression models control for a long list of housing characteristics, such as area size, property type, location, type of sale, in addition to buyer ethnicity fixed effect, year and month fixed effects, with robust standard errors clustered at the suburb level.

Besides using the baseline regression approach, we also check for robustness from endogeneity and sample selection concerns. For the endogeneity concern, we recognize it is possible that some unobserved homebuyer or suburb level characteristics or latent variables at the ethnicity level might affect the housing market decisions. To control for endogeneity, we use instrumental variable approach. Follow Guiso, Sapienza, and Zingales (2009) and Ahern et al (2015), to instrument for cultural differences in our estimations, we use genetic distance[5] as our IV. We are also aware of the fact that we only observe completed transactions in the housing market wherein the ask price from the seller is lower or equal to the offer price from the buyer. To rectify the sample selection bias, we use Heckman two-step selection models. Our results are robust to endogeneity and self-selection concerns.

[5] Genetic distance is a measure of the probability that two random alleles (DNA variations) from two populations will be different, based on the dominant population of a country.

Considering that different ethnicity groups in different part of the world may display varying degrees of home culture preference, we further explore the extent to which cultural distance affects housing prices for buyers across different regions of the world. There are several migration waves into Australia. Some ethnicities are more recent migrants into Australia, and may have stronger cultural bonds with their home country, and therefore may display stronger home culture preferences. If these migrants make their permanent home here, and end up buying property, a most natural impulse would be to congregate together in urban centres. Earlier immigrants, in general they are more neutralized and localized, after spending a long period of time in the local society. Our results show that Asian (including East Asian, South-East Asian and South Asian) ethnicities have the highest level of home culture preference, having the most negative coefficient of cultural distance on housing price. Other ethnicities (e.g. Australians and Europeans (including East Europe, North Europe, South Europe and West Europe)) display lesser effect of home cultural preference. Our interpretation of the result is that since Europeans came to Australia relatively early compared with Asian immigrants, they are better adapted to the local Australian society, and their ties to their home country are weaker.

The paper is related to several strands of the literature. Culture is known to play an important role in shaping the behavior and decision making of individuals (e.g., Hermalin (2001)). The significance of cultural distance in investment decisions is highlighted in prior works. Specifically, studies have shown that cultural distance provides important explanations for the magnitude of the cross-border flow of both debt (Aggarwal, Kearney and Lucey (2012)) and equity (Siegel, Licht and Schwartz (2013)), loan contract terms (Giannetti and Yafeh (2012)), the extent of investor home bias (Beugelsdijk and Frijns (2010); Anderson, Fedenia, Hirschey and Skiba (2011)),

and the degree of cross-border merger and acquisitions activity (Ahern, Daminelli and Fracassi (2015)). While these papers examine the impact of cultural distance on the investment decisions of investors and corporate managers, we apply it to individual homebuyers within a city made up of a culturally diverse population.

This study makes special contribution in several important ways. First, we quantify the effect of culture by constructing a buy-suburb pair, and for each pair we compute the cultural distance between the buyer and the suburb using cultural distance based on the six-dimension culture framework of Hofstede (2001), instead of just the buyer's own-ethnicity shares in the suburb. This novel approach differs from the prior work in the literature such as Wong (2013) which looks at own-ethnicity share, and does not explicitly consider that some ethnicities are more culturally compatible than others. For example, Chinese and Koreans share greater similarity in cultural background, compared with Chinese and Greeks. Our approach is able to capture this inter-ethnicity similarity, whereas the conventional approach using own-ethnicity share does not.

Second, we are able to infer the ethnicity of the homebuyer for each residential property transaction, and so measure an individual buyer's willingness to pay based on their cultural distance to a neighborhood. As such we have a more direct method of measuring buyer preferences than inferring flows from changes in the ethnic mix of residents as in Wong (2013) or using Census data as in Li (2014).

Third, we carefully address self-selection concerns using a Heckman two-step selection procedure, instrumenting the homebuyer's location choice using prior year's home purchase from the same ethnicity into that suburb. Furthermore, to address concerns of endogeneity and unobservable variables, we use two-stage least squares and instrument cultural distance using genetic distance of ethnicities. Our findings

remain robust to addressing endogeneity and self-selection. Being able to address endogeneity and self-selection in buyer preferences for housing is a non-trivial issue. For example, Wong (2013) addresses endogeneity of ethnic preferences by exploiting the unique ethnic quotas in Singapore housing blocks and seeing how marginally constrained blocks respond to ethnic quotas.

The paper is related to several strands of the literature. The significance of cultural distance in investment decisions is highlighted in prior works. Specifically, studies have shown that cultural distance provides important explanations for the magnitude of the cross-border flow of both debt (Aggarwal *et al.* (2012)) and equity (Siegel *et al.* (2013)), loan contract terms (Giannetti and Yafeh (2012)), the extent of investor home bias (Beugelsdijk and Frijns (2010); Anderson *et al.* (2011)), and the degree of cross-border merger and acquisitions activity (Ahern *et al.* (2015)). While these papers examine the impact of cultural distance on the investment decisions of investors and corporate managers, we apply it to individual homebuyers within a city made up of a culturally diverse population.

The remainder of the paper is organized as follows. Section 2 describes the institutional background on ethicality and immigration in Australia, and how culture relates to investment decisions. Section 3 describes data collection and sample construction. Section 4 elaborates on our empirical methodology. Section 5 presents the results and robustness checks and Section 6 concludes.

2. Background

2.1 Ethnicity and Immigration in Australia

In Australia, residents with foreign ancestry consist of more than half of the total population[6]. Historically, the majority of migration has come from Europe; however, there are increasingly more Australians who were born in Asia and other parts of the world. This pattern of migration is evident in the makeup of the richly diverse society which has been recorded in the 2011 Census. Over 300 ancestries were separately identified in the 2011 Census.

During the period from 1901 to 1958 with the enforcement of the "White Australia" policy, much of Australia's cultural diversity from its Asian neighbors, particularly China and India, was effectively extinguished. As a result, the predominant ethnicities were white Europeans, particularly Anglo-Saxons.

After the relaxation of White Australia policy and concurrent to the end of World War II, there were several migration waves. Appendix 1 provides a guide of the timeline of migrations from other countries between 1954 and 2011. We collect the top ten overseas countries of birthplace by percentage of the Australian population from each Australian Census between 1954 and 2011. The table reports for each top ten birthplace, the census year entry into the top ten and the census year and figure of when the birthplace was at the peak of its population.

Anglo-Saxons (i.e. Irish and UK) experienced peak population in 1954. In the 1950's and 60's, peak population occurred for Eastern and Western Europeans namely the Dutch, Germans and Polish. From the 1970's, the Southern Europeans (Greeks, Italians and Maltese) had population peaks. In the 1980's Lebanese migration peaked and in the 1990's it peaked for Yugoslavs. For both the Lebanese and Yugoslavs the peaks followed the outbreak of civil war in their respective countries. In the 2010's, Asian countries (China, India, Malaysia, Philippines and

[6] Appendix 2 Panel A reports the population rank of the top 20 ethnicity groups in Sydney based on 2011 census. Residents with Australian/British origin are consisted of 43% of the population, and the remaining 53% are of foreign orign.

Vietnam) experience their peak population as well as for neighboring Commonwealth countries New Zealand and South Africa. Asian countries entered the top ten in the 1990s, indicating that Asians are from the most recent wave of new migrants. In summary, the most established ethnicities in Australia are the Anglo-Saxons, followed by Western Europeans, Southern Europeans, Middle Eastern, Asian and New Zealanders. The sequencing is important as cultural distance sensitivity or home culture preference may be weaker for more established ethnicities than recent migrants.

In our multicultural and individualist age, will the immigrants choose to retain their distinctive culture from country of origin like the different "ingredients in the salad bowl", or will they surrender their cultural and ethnic identities and "get melted in a "melting pot"? The increasing migration has attracted a number of studies examining the social, economic and spatial implications for the host countries. With the various waves of immigrations into Australia, the homebuyers come from different countries and cultural backgrounds. In this study, we focus on the role of cultural interaction and integration between the homebuyers and the neighborhood on the local housing market.

2.2 Culture and Investment Decisions

There is a developing literature on the effect of culture on investment decisions. In general, these papers find consistent evidence that the cultural difference between two countries leads to a greater level of information asymmetries and agency problems, which negatively affect cross-country investment activities. For example, higher cultural distance between two countries is shown to be related to lower portfolio investment and direct investment (e.g. Guiso, Sapienza and Zingales (2009)),

lower allocation to foreign investments (e.g. Beugelsdijk and Frijns (2010) and Anderson *et al.* (2011)), smaller bank loans with higher interest rates (e.g. Giannetti and Yafeh (2012)) and lower cross-border merger volume (e.g. Ahern *et al.* (2015)). Overall, cultural difference between countries appears to act as a friction in the information flow between countries, affecting the size of investment and value generation between countries.

At the individual investors' level, literature documents that familiarity of investments is enhanced by cultural (e.g., language or religious) similarity, besides geographical closeness. For example, Grinblatt and Keloharju (2001) show that investors in Finland prefer to hold equity in firms whose CEOs have similar cultural origins even after controlling for the language of corporate reports and the physical distance from the company's headquarters. Bottazzi, Da Rin and Hellmann (2007), who provide evidence that venture capitalists are less likely to fund entrepreneurs in countries whose citizens they trust less and, if they do invest in these countries, the contracts they use are different from the contracts in use in high trust countries.

This study distinguishes from prior culture studies on investment decisions in several important aspects. Our paper employs a novel measure of cultural distance, which differs from the literature. First, we look at the buying behavior of ethnicities in one large city instead of cross-border transactions. Under this setting, cross-country differences such as in trade or legal frameworks do not factor into our analysis. Second, we construct homebuyer and neighborhood pair, and look at cultural differences between a buyer and the population-weighted ethnic mix in the neighborhood instead of direct cross-country pairs. Such analysis differs from the usual cross-country pair analysis in the literature and represents a novel method to quantity inter-ethnicity cultural distance. Third, we investigate the relationship of

cultural distance to housing prices at the individual housing transaction level whereas other studies focus on the country trade level (e.g. Guiso, Sapienza and Zingales (2006)) or on public company equity (e.g. Beugelsdijk and Frijns (2010)). As such, we investigate the relationship of cultural distance at a very granular level. While most studies in the prior literature considers ethnicity preferences of homebuyers to their own ethnicity, we are able to further analyze the preference of homebuyers to other ethnicities using the weighted average Euclidean distance of cultural scores derived from Hofstede (2001)'s along six dimensions.

3. Data

The principal dataset we use is the individual housing transaction data of the Sydney Metropolitan Area from 2006 to 2013 from Australian Property Monitors (APM).[7] The dataset includes a comprehensive list of variables, including the transaction price, transaction date, property address, buyer and seller names, whether the transaction is an auction sale, whether the home is a new development, number of bedrooms and bathrooms, whether the home has parking, area size of block of land, and other housing characteristics (garage, balcony, ocean views, etc.). Sales prices and area sizes at the 1st and 99th percentile are winsorized to remove outliers.

We adopt Hofstede (2001)'s six-dimension culture framework[8] as our main measure of culture. The Hofstede model of dimensions of national culture has been widely used in social science research.[9] In this framework, a score on the scale of 0-100 is assigned to each country along six distinctive dimensions, including

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

[8]The detailed culture data is obtained from this website http://www.geerthofstede.eu/research--vsm

[9] See Kirkman, Lowe and Gibson (2006) for a literature review on works using Hofstede's culture framework.

uncertainty avoidance, individualism, power distance, masculinity, and long-term orientation, and indulgence. The detailed definition of each dimension is explained in Figure 1, and the culture scores for the top 15 ethnicities in Australia population (2001 census) are shown in Table 2, along the six dimensions. Figure 2 offers a pairwise comparison of the culture values of Australia and China through the lens of the six-dimensional model. These two countries differ significantly on several dimensions, especially on individualism and indulgence.

[INSERT FIGURE 1 and 2 ABOUT HERE]

Individualism is the degree of interdependence a society maintains among its members. Australia, with a score of 90 on this dimension, is a highly individualist culture, whereas China scores only 20. This means that Australia is a relatively loosely-knit society compared with China, a more collective society. In individualistic societies, people tend to look after themselves and their direct family only. In Collectivist societies, people belong to 'in groups' that take care of them in exchange for loyalty. With a high score of 71 on indulgence compared with 24 for China, Australia is an indulgent country. People in societies classified by a high score in Indulgence generally exhibit a willingness to realize their impulses and desires with regard to enjoying life and having fun. They possess a positive attitude and have a tendency towards optimism. In addition, they place a higher degree of importance on leisure time, act as they please and spend money as they wish.

Other datasets we use include the Australian Bureau of Statistics (ABS) Census snapshots on the demographics of a suburb (i.e. ancestry, country of birth) in 2006 and 2011. Further, we use genetic distance between ethnicities from Spolaore and Wacziarg (2009). Genetic distance is a measure of the probability that two

random alleles (DNA variations) from two populations will be different, based on the dominant population of a country.

Table 1 reports the summary statistics for our entire sample of 217,452[10] sales and across the top twenty buyer ethnicities by sales.[11] The complete list of ethnicities and regions that we use is in Panel B of Appendix 2, and the full list of housing characteristic variables is in Appendix 3. The average housing price is A$671,100 with 60 percent of sales being free-standing houses with an average size of 4,010 square feet, 2.94 bedrooms and 1.61 bathrooms. 86 percent of homes have parking space and 16 percent were sold at an auction. Australian (Anglo-Saxon) homebuyers are the biggest group of homebuyers, making up about 35 percent of transactions in our sample. This is consistent with Australians being the majority ethnicity in Sydney, constituting 43 percent of the population. Australians on average paid $758,060 for a home, which is higher than the overall sample average. The housing characteristics of Australian buyers are similar to the overall sample average. The second largest buyers are Chinese making up about 18 percent of the sample, followed by Arabians making up about 10 percent of the sample. This is consistent with the recent wave of Chinese buyers entering the Australia housing markets. In general, Chinese buyers are less likely to buy a free-standing house compared with Australians. On average, Chinese buy smaller homes than Australians, with a size of 3,260 square feet. However, the number of bedrooms and number of bathrooms in their home are similar to that of Australian homes.

[INSERT TABLE 1 ABOUT HERE]

4. Methodology

[10] 218,869 when including Jewish and South African surnames.
[11] The top twenty buyer ethnicities is reported for conciseness.

In this section we first describe how we classify owner ethnicity from buyer surnames. We then show how we calculate cultural distance measures between buyers and neighborhoods. Next, we describe our hedonic housing price regression framework linking housing prices to cultural distance. Finally we describe how we address concerns on endogeneity and sample selection bias.

4.1 Buyer Ethnicity Classification

In order to calculate the cultural distance measures, we require the ethnicity of the homebuyers and ethnic composition of the suburb. We identify the ethnicity of homebuyers based on the surnames using a hand collected database of surnames and ethnicity from various free internet sources such as surnamedb.com and Wikipedia. For South African surnames, we use the list in Rosenthal (1965). For names unmatched by our database, we use the name to ethnicity classifier[12] from Ambekar, Ward, Mohammed, Male and Skiena (2009) and as Pool, Stoffman and Yonker (2015) use to match Arabic, British (Australian), French, Indian, Italian or Jewish names when the predicted probability of an ethnicity by the classifier is above 85 percent. Surnames with more than one ethnicity (e.g. the surname Lee could be Anglo-Saxon, Chinese or Korean) are dropped. We remove buyers with multiple owners of different ethnicities. We remove company owners which make up about 2.5 percent of the sample. Using these filters we are able to match about 54 percent of names in our sample of home sales to an ethnicity.

4.2 Cultural Distance Measures

[12] Available from http://www.textmap.com/ethnicity/

We calculate the cultural distance based on Hofstede (2001). Hofstede (2001) constructs culture scores on the basis of six dimensions: power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The cultural distance (CD) here is defined as the weighted average Euclidean distance between the culture value of the homebuyer's ethnicity and the culture value of the suburb (neighborhood) of the property, using the ethnic decomposition as the weight. Ethnicity of the suburb is based on the suburb's ancestry or birthplace from the ABS Census 2006 and 2011 records. For years between 2006 and 2011 where there is no census information we impute demographic information. For 2012 and 2013, we assume the demographic information is the same as for 2011.

Specifically, $CD_{i,s,t}$ measures the cultural distance between buyer *i*'s ethnicity and the culture of suburb *s* in year *t*:

$$CD_{i,s,t} = \sum_{j=1}^{J} w_{j,s,t} * CD_{i,j,s,t} = \sum_{j=1}^{J} w_{j,s,t} * \sqrt{\sum_{k=1}^{K} \left(C_{i,k} - C_{j,k,s}\right)^2 / V_k} \qquad \text{--- (1)}$$

where

$\boldsymbol{C_{i,k}}$ is buyer of sale *i*'s ethnicity culture value along the *k*-th culture dimension;

$\boldsymbol{C_{j,k,s}}$ ethnicity group *j*'s value on the *k*-th culture dimension in suburb *s*, *j*=1…J;

$\boldsymbol{V_k}$ is the variance of the culture value of the dimension *k*;

$\boldsymbol{w_{j,s,t}}$ is the percentage of ethnicity group *j*'s population in suburb *s*in year *t*;

There are in total *J* ethnicity groups and *K* culture dimensions.

Using this formula, we compute a weighted average measure of CD of a buyer to the home's suburb. The higher the score on the cultural distance measure, the greater is the cultural difference between buyer *i*'s culture and the cultural mix of the suburb. For robustness, we also measure a suburb's ethnicity by country of birthplace

instead of country of ancestry, and use all six dimensions or just four dimensions of Hofstede framework by excluding long-term orientation and indulgence versus restraint. By using only four dimensions we are able to include Israeli/Jewish and South African buyers as these ethnicities do not have cultural scores for long-term orientation and for indulgence versus restraint.

Figure 3 illustrates the cultural distance of homebuyers in various suburbs across Sydney. Cultural distance (CD) is measured for a representative Australian buyer in Figure 3 Panel A and an average Chinese buyer in Figure 3 Panel B. We use suburb demographic information on ancestry in the 2011 Census. The representative Australian buyer has the lowest CD generally, especially in the outer northern, outer western, outer southern parts of Sydney and some parts of eastern Sydney. Australian homebuyers' CD tends to be highest in inner Sydney. An average Chinese buyer in contrast faces the highest CD in the outer ring suburbs and lowest CD in some suburbs in inner Sydney.

[INSERT FIGURE 3 ABOUT HERE]

Table 2 reports summary statistics for cultural distance (CD) for the entire sample and across the top twenty ethnicities. Over the entire sample, average CD is 1.99. The statistic could be interpreted as the mean buyer's suburb is 1.99 standard deviations away from the buyer's ethnicity cultural dimension score. Australian buyers have the lowest average CD of 1.35 across all ethnicities consistent with most suburbs having a majority Anglo-Saxon demographic. Chinese and Vietnamese have the highest CD standard deviation, implying they tend to buy into suburbs with a broad range of CD to themselves. Other ethnicity groups have lower CD standard deviations, such as Japanese, suggesting that they tend to concentrate buying in fewer

suburbs. The wide variability of CD across buyers in our sample allows us to test the relationship between housing prices and homebuyers' CD to a suburb.

[INSERT TABLE 2 ABOUT HERE]

4.3 Hedonic Housing Price Model

After measuring CD we then estimate a hedonic housing price model using ordinary least squares based on the following empirical specification and variable definitions:

$$\ln(P_{ist}) = \alpha_t + \beta_k CD_{ist} + property\ char_i + \mu_s + \delta_i + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (2)}$$

Where

$\boldsymbol{ln(P_{ist})}$ denotes logarithm of housing prices paid by buyer of sale *i* at suburb *s* at time *t*;

$\boldsymbol{property\ char}$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type and area size[13];

$\boldsymbol{\mu_s}$ is the suburb location specific fixed effects;

$\boldsymbol{\delta_i}$ is buyer's ethnicity fixed effects;

$\boldsymbol{\gamma_t}$ is year/quarter fixed effects;

$\boldsymbol{\tau_t}$ is a monthly time trend.

A positive and statistically significant $\beta_k$would suggest that homebuyers pay higher prices when facing greater cultural distance to a suburb, consistent with the *information friction hypothesis*. On the other hand if we find a negative and statistically significant β_k, this is evidence consistent to the *home culture preference hypothesis* as it suggests that buyers pay higher prices for suburbs with lower CD. Our

[13]Appendix 3 shows the full list of housing characteristics that we use in the regression specification.

hedonic housing price regression models also control for a long list of housing characteristics, including area size, property type, location, type of sale, in addition to buyer ethnicity fixed effect, year and quarter fixed effects, with robust standard errors clustered at the suburb level. Suburb fixed effects controls for average suburb level characteristics which includes suburb specific amenities (e.g. schools, parks and public transport infrastructure). Buyer ethnicity fixed effect are used to control for average characteristics of ethnicities. Year and quarter fixed effects and monthly time trend control for time trends in housing prices.

There are two potential problems with the baseline specification estimate using ordinary least squares: sample selection bias and endogeneity. First, sample selection bias arises as we only observe completed deals in the housing transactions, wherein the offer price of the buyer is greater than or equal to the ask price of the seller. If the sample of completed deals in the housing market is systematically different from that of uncompleted deals, coefficients of cultural distance and other control variables on housing prices might be biased in our estimation.

To address this sample selection bias, we use a Heckman two-stage selection model. In the first stage we run the following Probit model across ethnicities at the suburb/quarter level:

$$\Pr(Buy_{jst} = 1|X) = F(\beta_1 CD_{jst} + \beta_2 lastyear\_buy_{jst} + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst}) \quad \text{--- (3)}$$

The dependent variable is a dummy which takes the value of 1 if a given ethnicity buys in a suburb in a given quarter and 0 otherwise. Our instrumental variable is $lastyear\_buy_{jst}$, a dummy which takes the value of 1 if there is any housing transaction by the buyer's ethnicity in the prior twelve months in suburb *s* and 0 otherwise. We obtain the Inverse Mills Ratio from the Probit estimation and use it as

an additional independent variable in Equation 2. The instrumental variable is motivated by the literature on peer group effects. It has been found that peer group effects such as within ethnicity groups strongly influences the behavior and decisions of an individual, controlling for other factors. See for example in car purchases (Grinblatt, Keloharju and Ikäheimo (2008), employment outcomes (Bayer, Ross and Topa (2008), Patacchini and Zenou (2012)) welfare participation (Bertrand, Luttmer and Mullainathan (2000), Betrand, Luttmer and Mullainathan (2000)) and worker productivity (Mas and Moretti (2009)). Importantly, this peer effect influences the decision and not the price paid by the individual. Consistent with this effect, Hvide and Östberg) find stock market decisions of individuals are positively correlated with those of co-workers and this positive correlation is not associated with positive future returns. As such we hypothesize that prior buying by the same ethnicity is a valid instrument as it increases the probability of home purchase by the same ethnicity group, and also in the meantime has no effect on the price paid.

Endogeneity might also be present in our baseline OLS regression approach as there may be an omitted variable bias where unobserved characteristics of the buyer, home or neighborhood may be correlated with both housing prices and CD. To address endogeneity, we use an instrumental variable strategy, and identify CD with genetic distance following Guiso *et al.* (2009) and Ahern *et al.* (2015). As Ahern *et al.* (2015) describes, genetic distance is a measure of the probability that two random alleles (DNA variations) from two populations will be different, based on the dominant population of a country. Genetic distance is correlated to CD as ethnicities that share common ancestors will tend to inherit both biological and cultural similarities (e.g. Spolaore and Wacziarg (2009)). However as genetic similarities take many generations to eventuate in ethnicity, it is unrelated to housing prices. The first

stage regression regresses the buyer's CD on all control variables including genetic distance as such:

$$CD_{ist} = \alpha_t + \beta_k GD_{ist} + property\ char_i + \mu_s + \delta_i + \gamma_t + \tau_t + \varepsilon_{it} \quad \text{--- (4)}$$

where GD_{ist} is the genetic distance of buyer *i* to the population weighted ethnicities in suburb *s*. We then use the estimated CD, $\widehat{CD_{ist}}$, from the first stage regression in the second stage instead of CD_{ist}.

5. Results

5.1 Cultural Distance and Homebuyers' Location Choice

To examine whether cultural distance affects homebuyers' location choice of their home, we estimate the Probit model specified in Equation 3. The Probit model also represents the first stage of the Heckman selection model that we use to adjust for sample selection bias. Table 3 reports the estimation result. The central finding is that cultural distance between the buyers' ethnicity and suburb lowers the probability of buyers buying into a particular suburb. Across the four different versions of cultural distance measures (CD, BCD, CD4, BCD4), the coefficients on cultural distance are negative and statistically significant, suggesting that the larger the cultural distance, the less likely an ethnicity group buying into a given suburb.

The finding that buyers display a preference for homes located in suburbs with similar cultures supports the home culture preference hypothesis. This is consistent with Ahern *et al.* (2015)'s first stage Probit results on cross-border mergers and acquisitions where higher CD between two countries reduces the probability of a merger occurring. Consistent with a peer effect in home buying, $lastyear\_buy_{jst}$ is positive and statistically significant across all four specifications with different measures of CD, suggesting that prior period's buying in a suburb by an ethnicity

increases the chances of the ethnicity buying in the current period. We use *lastyear_buy* as an instrument variable, and calculate the inverse mills ratio from this result to use in the Heckman Selection Model in Section 5.2.

[INSERT TABLE 3 ABOUT HERE]

To address the potential endogeneity in this simple Probit approach, we also use an instrumental variable approach to provide robustness result. We instrument the cultural distance with genetic distance, and estimate an IVProbit Model. Appendix 4 reports the IVProbit result which is consistent with our simple Probit result with Table 3. This confirms that homebuyer' cultural distance to a suburb is an important determinant in whether an ethnicity buys into a suburb or not.

5.2 Cultural Distance and Housing Prices

Table 4 reports coefficient estimates of cultural distance using ordinary least squares regression. Each panel represents the use of a different version of the cultural distance measure. We find the coefficient for CD is negative and statistically significant across all four versions of cultural distance measures in the baseline regressions. It is also robust when accounting for endogeneity in Table 5 as well as self-selection bias in Table 6. We control for a comprehensive list of housing characteristics including number of bedrooms, number of bathrooms, parking etc. We find that transaction prices are higher for homes with more rooms, larger size and other more positive features.

[INSERT TABLE 4 ABOUT HERE]

Generally across cultural distance measures in the three tables, the coefficient estimates become more negative when adjusting for self-selection or endogeneity

suggesting that the ordinary least squares estimates are biased downwards. The inverse mills ratio coefficients using the Heckman selection model across cultural distance measures is not statistically significant suggesting that selection bias is not a serious concern.

The coefficient estimate across measures and methods ranges from -0.009 (Table 4 Model 2, Heckman ordinary least squares) to -0.029 (Table 5 Panel B Model 3, two-stage least squares). This suggests that a one standard deviation increase in the CD of a buyer with the suburb reduces the buyer's price by between 0.9 to 2.9 percent. For example for our regression results using ordinary least squares in Table 4 Panel A, a one standard deviation move in CD reduces housing prices by 1.1 percent or A$7,382 given the mean sales price of A$671,100. As a buyer's CD may range from 0.59 to 4.37, this is an economically significant amount. Our results therefore support the effect of home culture preference of homebuyers.

5.3 Cultural Distance and Ethnicity Interaction, and Subsample Analysis for Each Region

Extending on our baseline results, this section investigates whether the CD coefficient is heterogeneous amongst ethnicities, particularly whether the effect is strongest for more recent migrants. There is evidence to believe this is the case. For example Fischer (2012) shows that immigrant inflows from countries with non-common language country into an area increases housing prices due to the valuing of immigrant specific amenities and networks. On the other hand, ethnicities that have been in Australia for many generations may not have such a strong home culture preference as they are more established locally.

We extend our baseline regression by interacting CD to a dummy of one if the ethnicity of the buyer is from a region as outlined in Appendix 2 Panel B (e.g. East

Asia), or otherwise. We interact with regions rather than each ethnicity for conciseness. The groupings also provide a rough guide of immigration arrival. For example as inferred from Appendix 1, the earliest mass migration waves were by Anglo-Saxons followed by Western and Northern Europeans. The most recent mass migrations were from Asia.

Table 7 reports results for cultural distance and buyer region interaction coefficient estimates. We provide baseline estimation result in Panel A, robustness check using 2SLS in Panel B and Heckman selection in Panel C. We use four different cultural distance measures for each table.

[INSERT TABLE 5 ABOUT HERE]

Consistent with more recent migrants being more sensitive to home culture preference, we find negative and statistically significant coefficients for CD across all cultural distance measures and all regression methods for East Asia and South East Asia. The CD coefficient estimates for these regions also tend to be much larger than for CD on the overall sample in Table 4. For example the coefficient estimates for CD*East Asia ranges from -0.019 to -0.048, suggesting while South East Asia ranges from -0.011 to -0.039.

Middle East, South Asia and Southern Europe also show negative and statistically significant coefficients for cultural distance measures although the result is not robust to controlling for endogeneity using two stage least squares. Australia, Eastern Europe, Northern Europe and Western Europe show negative and statistically insignificant or weakly significant CD coefficients, which suggests that these groups are less affected by CD, consistent with earlier migrants not being sensitive to CD. In summary, our results provide evidence that home culture preference appears related to

the migration time of the buyer's ethnicity, with the earliest migrants showing no evidence of home culture preference.

To further test whether our results are robust to within ethnicity region CD variation, Table 8 conducts subsample analysis, and reports the different degree of home culture preference for each of the three main regions: Asia, Australia, and Europe[14]. The coefficients on cultural distance can be considered as the degree of home culture preference. Asian and European homebuyers care about cultural distance of the suburb of the home, inferred from the negative coefficients on cultural distance in all four specifications in Table 8 Panel A and Panel C. Compare Asian and European buyers, Asian buyers display a higher degree of cultural distance. Asian buyers are willing to pay a higher price premium to live in culture similar suburbs than European buyers. Compared with Asian and European buyers, Australian homebuyers show the lowest degree of home culture preference, even having positive coefficients on cultural distance in all four specifications in Table 8 Panel B.

5.4 Home Cultural Preference and Individualism

In this section we focus on individualism, and examine whether buyers from countries high on individualism have lower degree of home culture preference. The intuition is that cultures high on individualism emphasize individual freedom, uniqueness, and individual personality; whereas cultures low on individualism emphasizes strong group cohesion, which represents a preference for a tightly-knit framework in society (Hofstede (2001)). In our regression model, we interact the culture score on standardized individualism score of the homebuyer's ethnicity with

[14] The aggregate number and dollar volume of housing transactions from these regions is illustrated in Figure 4. Asia consists of about 30% of dollar sales while Australia consists of about 40% over our sample period.

the cultural distance, and expect that the coefficient on the cultural distance to be negative and the interaction term to be positive.

Table 9 reports the estimation result of hedonic housing price model with this additional interaction term. As before, we have four versions of the cultural distance measure in the four model specifications, ACD6, BCD6, ACD4 and BCD4 in the Model (1) to Model (4), which correspond to two different versions of ethnicity identification (ancestry versus birthplace), and two different versions of Hofstede culture measures (6-Dimension or 4-Dimension). Across all four models, the coefficients on cultural distance measures are negative and significant, supporting the home culture preference for the entire sample. The interaction term of individualism and cultural distance is positive and significant, which is consistent with our prior that buyers high on individualism display lower home culture preference.

5.5 Additional Robustness Checks

In this section we conduct robustness checks on our baseline results by applying seller fixed effects and removing Chinese buyers from our sample.

5.5.1 Seller Fixed Effects

Implicit in our model is that we assume it is a buyers' market, and sellers accept the highest price available and so seller characteristics are not a determinant to the final transaction price. However sellers may still be able to influence prices in many ways. For example, it is possible that sellers might be willing to accept a lower price if they are in financial distress or they might ask for higher prices if they have more bargaining power in negotiation. Such omitted characteristics would be

controlled for using two-stage instrumental variable regression, as we have elaborated in earlier section.

In an attempt to further control for seller characteristics, we use seller ethnicity fixed effect based on seller surname. This is to alleviate the concern that sellers from certain ethnicity might prefer to sell to certain ethnicity groups of buyers more. Alternatively it might be easier for sellers from certain ethnicity to sell to buyers from the same ethnicity. Table 10 Panel A reports our coefficient estimates for CD including seller ethnicity fixed effects. Panel B uses the Heckman Selection Model approach, and Panel C uses the two-stage least square approach. The sample is reduced as we do not have seller name for all observations, and we delete transactions if the sellers had mixed ethnicities. In all three panels, we find the coefficient for CD remains negative and statistically significant across all measures of CD and regression models. This suggests that controlling for seller fixed effect does not affect our main results, which strengthens support for the home culture preference hypothesis.

[INSERT TABLE 6 ABOUT HERE]

5.5.2 Removing Chinese Buyers

As seen in the regional home culture preference results, we find the strongest cultural preference in East Asia which is dominated by Chinese buyers. The media suggests that wealthy foreign Chinese buyers are driving up housing prices globally, brought on by the increasing wealth of the middle class in China and their desire to emigrate (e.g. Patrick (2014)). To address for the concern that Chinese buyers might be driving our results, we remove Chinese buyers from our sample and redo our analysis. Table 11 reports our coefficient estimates for CD when removing Chinese

buyers from our sample. Panel A is OLS estimation result, Panel B is Heckman Selection Model result, and Panel C is two-stage least square estimation result.

From the results in all model specifications across the three panels, we can see that the coefficients on four versions of cultural distance measures are still mostly negative and significant, which means that the home culture preference result still holds after removing Chinese buyers from our sample. The result provides robustness to our result and confirms that removing Chinese buyers does not affect the main finding on home culture preference.

6. Conclusion

This paper shows that cultural distance matters in people's housing decisions. We develop two competing hypotheses, the home culture preference hypothesis and the information friction hypothesis. Our empirical evidence demonstrates that homebuyers are more likely to choose homes in locations that are closer in culture to their own culture background. And also they pay higher prices for those homes. These findings provide strong support for the home culture preference hypothesis. Homebuyers are willing to pay a premium for homes in locations with similar cultural, suggesting immigrants' preference to preserve their unique cultural traits in a foreign land, supporting the salad bowl theory of social integration.

Specifically, a one standard deviation increase of cultural distance between a homebuyer with the suburb causes a 1.1 percent decrease in home price paid. The effect of cultural distance is economically meaningful, given the sample mean sales price of A$671,100. The results are robust to different cultural distance measures, endogeneity and self-selection bias. We also find that different ethnicity groups display varying degrees of home culture preference. Homebuyers from Asia display a greater degree of home culture preference, compared with European and Australian

buyers. Taken together, our paper provides important insights in the area of culture and housing prices. We contribute to the literature on the role of cultural distance in housing market decisions and highlight the importance of culture in the residential housing market.

References

Aggarwal, R., C. Kearney, and B. Lucey, 2012, Gravity and culture in foreign portfolio investment, *Journal of Banking & Finance* 36, 525-538.

Ahern, K. R., D. Daminelli, and C. Fracassi, 2015, Lost in translation? The effect of cultural values on mergers around the world, *Journal of Financial Economics* 117, 165-189.

Ambekar, A., C. Ward, J. Mohammed, S. Male, and S. Skiena, 2009, Name-ethnicity classification from open sources, Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining (ACM, Paris, France).

Anderson, C. W., M. Fedenia, M. Hirschey, and H. Skiba, 2011, Cultural influences on home bias and international diversification by institutional investors, *Journal of Banking & Finance* 35, 916-934.

Bayer, P., S. L. Ross, and G. Topa, 2008, Place of work and place of residence: Informal hiring networks and labor market outcomes, *Journal of Political Economy* 116, 1150-1196.

Bertrand, M., E. F. P. Luttmer, and S. Mullainathan, 2000, Network effects and welfare cultures, *The Quarterly Journal of Economics* 115, 1019-1055.

Beugelsdijk, S., and B. Frijns, 2010, A cultural explanation of the foreign bias in international asset allocation, *Journal of Banking & Finance* 34, 2121-2131.

Fischer, A. M., 2012, Immigrant language barriers and house prices, *Regional Science and Urban Economics* 42, 389-395.

Giannetti, M., and Y. Yafeh, 2012, Do cultural differences between contracting parties matter? Evidence from syndicated bank loans, *Management Science* 58, 365-383.

Grinblatt, M., and M. Keloharju, 2001, How Distance, Language, and Culture Influence Stockholdings and Trades, *Journal of Finance* 56, 1053-1073.

Grinblatt, M., M. Keloharju, and S. Ikäheimo, 2008, Social influence and consumption: Evidence from the automobile purchases of neighbors, *Review of Economics and Statistics* 90, 735-753.

Guiso, L., P. Sapienza, and L. Zingales, 2006, Does culture affect economic outcomes?, *The Journal of Economic Perspectives* 20, 23-48.

Guiso, L., P. Sapienza, and L. Zingales, 2009, Cultural biases in economic exchange?, *Quarterly Journal of Economics* 124, 1095-1131.

Hermalin, B. E., 2001. *Economics and corporate culture* (John Wiley & Sons Chichester, England).

Hofstede, G. H., 2001. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (Sage).

Hvide, H. K., and P. Östberg, 2015, Social interaction at work, *Journal of Financial Economics, forthcoming*.

Kirkman, B. L., K. B. Lowe, and C. B. Gibson, 2006, A quarter century of culture's consequences: A review of empirical research incorporating Hofstede's cultural values framework, *Journal of international business studies* 37, 285-320.

Li, Q., 2014, Ethnic diversity and neighborhood house prices, *Regional Science and Urban Economics* 48, 21-38.

Mas, A., and E. Moretti, 2009, Peers at work, *American Economic Review* 99, 112-45.

Patacchini, E., and Y. Zenou, 2012, Ethnic networks and employment outcomes, *Regional Science and Urban Economics* 42, 938-949.

Patrick, A. O., 2014, For a century, the mansion sat above Sydney Harbor. Then China's nouveau riche arrived., Available at: (Last accessed

Pool, V. K., N. Stoffman, and S. E. Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *Journal of Finance* 70, 2679-2732.

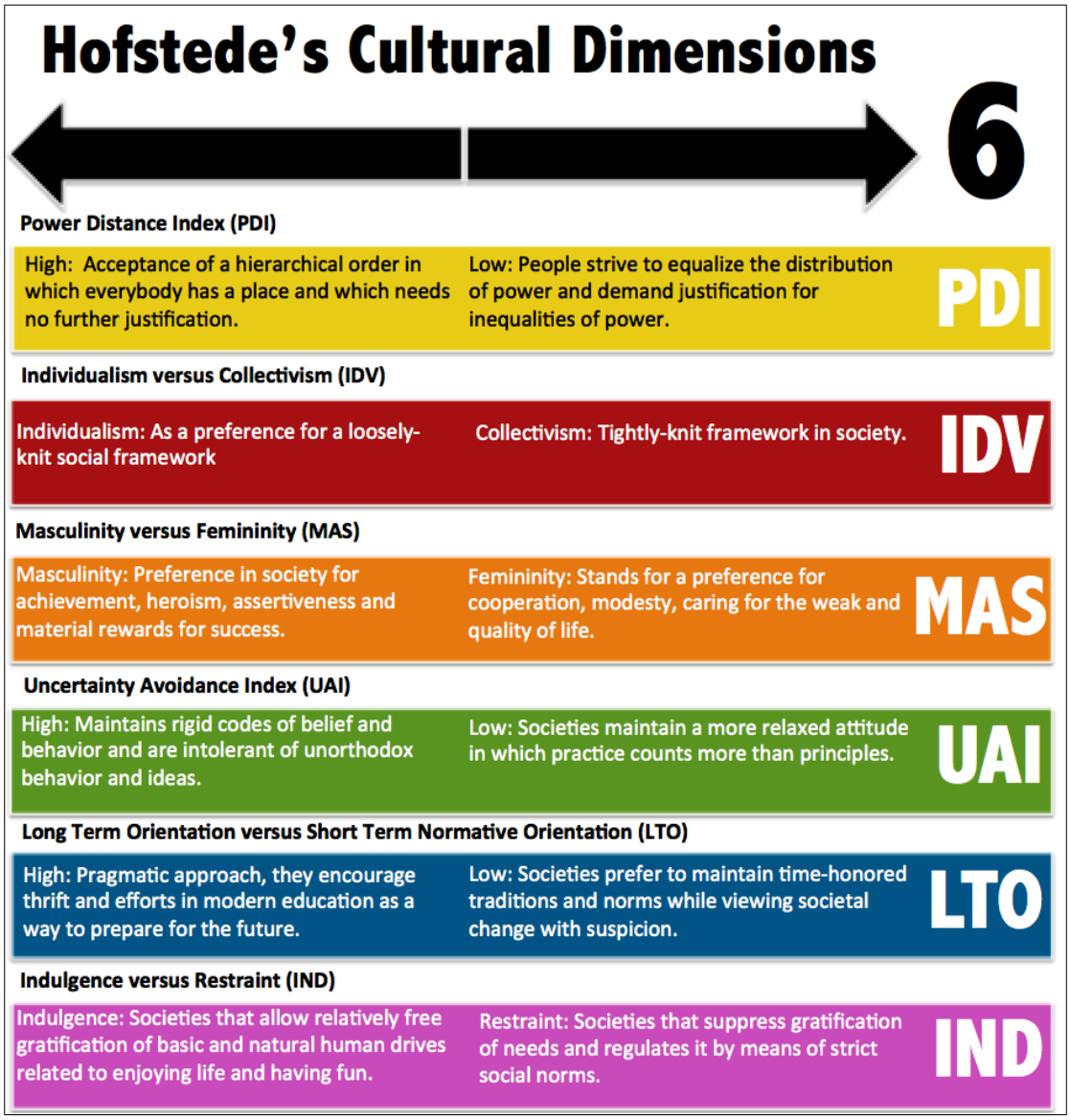
Rosenthal, E., 1965. *South African Surnames* (H. Timmins).

Saiz, A., 2007, Immigration and housing rents in American cities, *Journal of Urban Economics* 61, 345-371.

Siegel, J. I., A. N. Licht, and S. H. Schwartz, 2013, Egalitarianism, cultural distance, and foreign direct investment: A new approach, *Organization Science* 24, 1174-1194.
Spolaore, E., and R. Wacziarg, 2009, The diffusion of development, *Quarterly Journal of Economics* 124, 469-592.
Wong, M., 2013, Estimating ethnic preferences using ethnic housing quotas in Singapore, *Review of Economic Studies* 80, 1178-1214.

Figure 1
Hofstede (2001) Six cultural Dimensions

This figure shows the definition of six cultureal dimensions in the Hofstde (2001) framework. The six dimensions are power distance (PDI), individualism vs. collectivism (IDV), acceptance of gender difference (MAS), uncertainty avoidance (UAI), long-term orientation (LTOWVS) and indulgence vs. restraint (IVR).



Source: http://geert-hofstede.com/national-culture.html and http://www.skyword.com/contentstandard/wp-content/uploads/2015/10/hofstede-s-cultural-dimensions.png

Figure 2
Comparison of Culture Values of Australia and China

This figure presents the cultural scores of Australia in comparison with China, along six dimensions. Dark color bars represent the culture values of Australia and Light color bars are the culture values of China. .

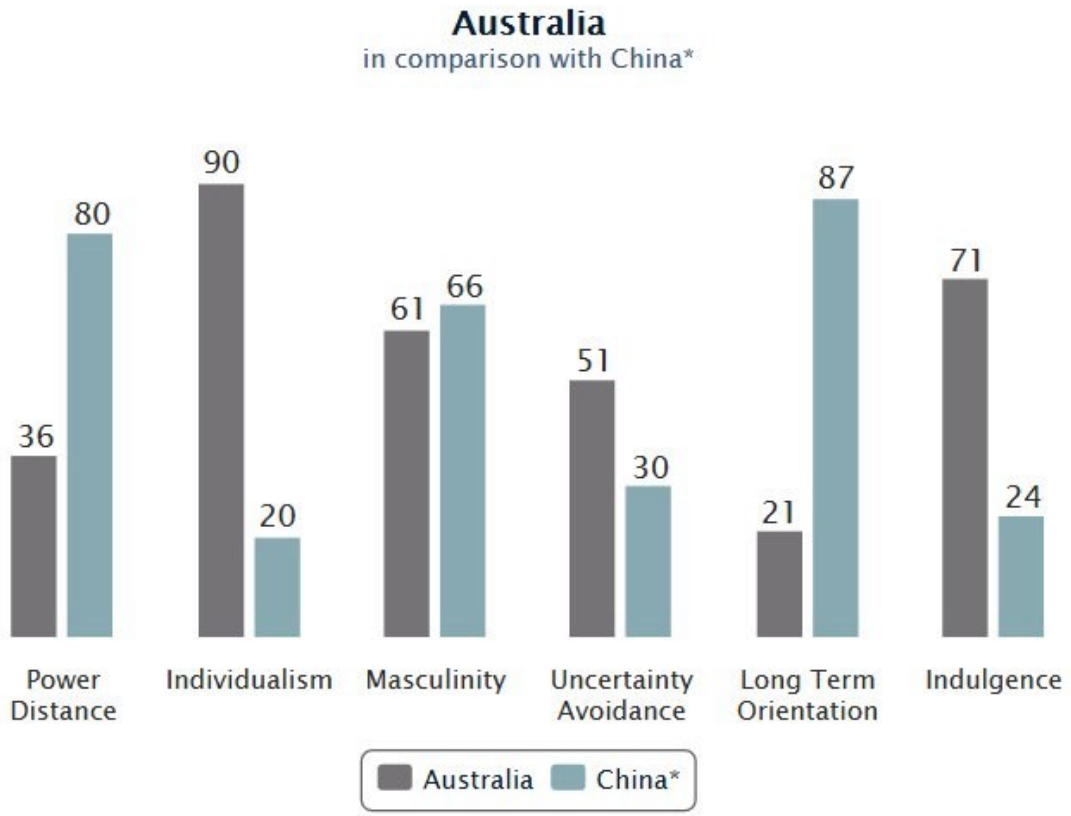
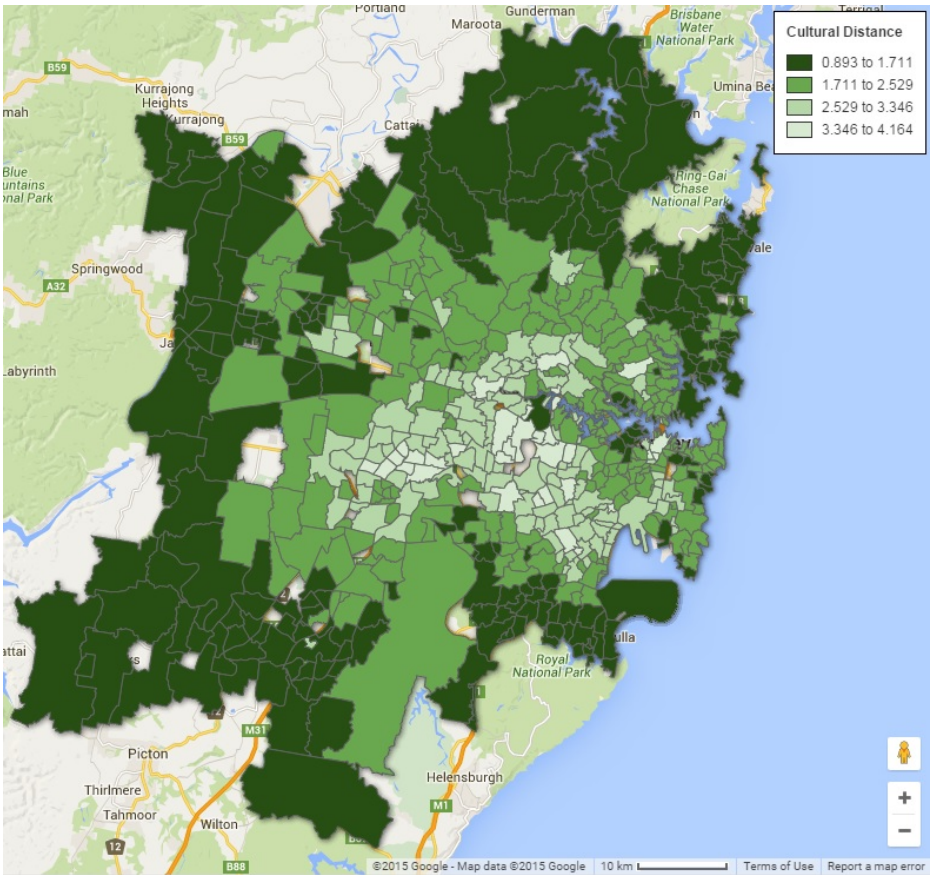


Figure 3
Heat Map of Buyer Cultural Distance across Sydney Suburbs

Panel A. Australian Homebuyers

This heat map shows the cultural distance between Australian homebuyers and suburb of the home for each suburb in the Sydney Metropolitan Area. Darker shades represent lower cultural distance, and lighter shades correspond to greater cultural distance.



Panel B. Chinese homebuyers

This heat map shows the cultural distance between Chinese homebuyers and the suburb of the home for each suburb in the Sydney Metropolitan Area. Darker shades represent lower cultural distance, and lighter shades correspond to greater cultural distance.

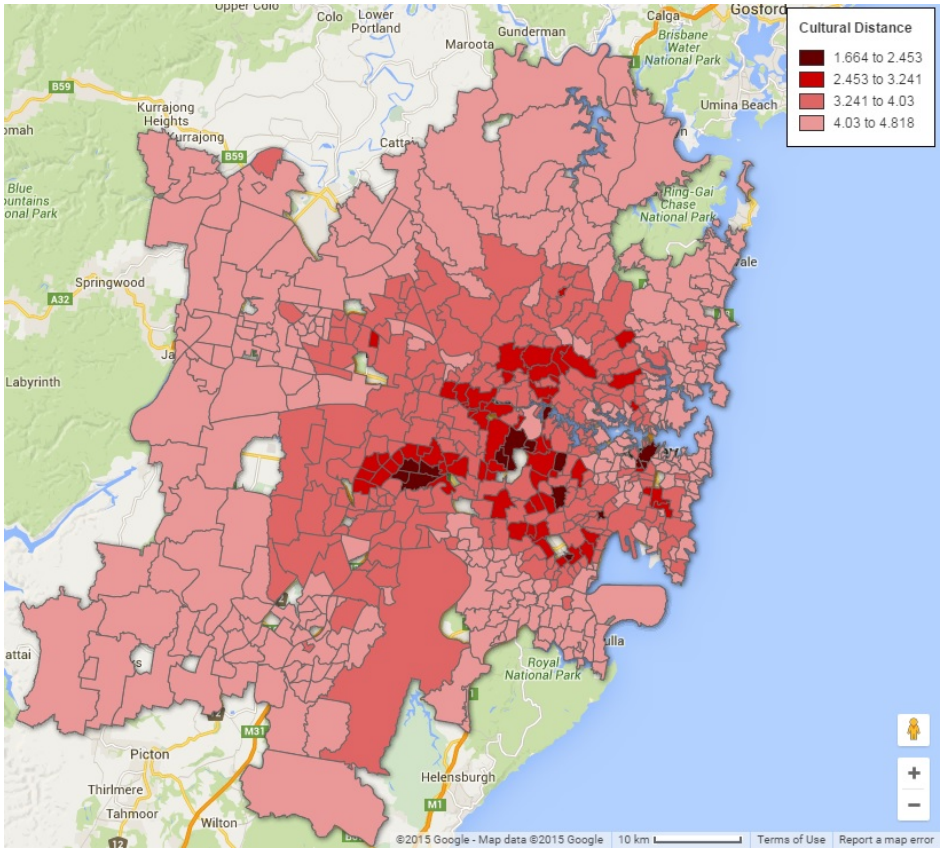
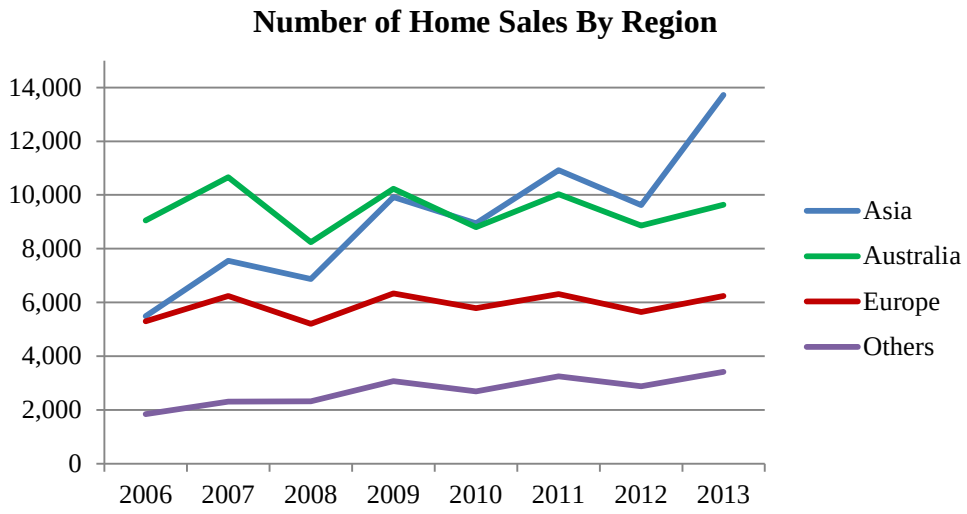


Figure 4
Housing Transactions over the Years

Panel A: Number of Home Sales By Region

This figure shows the number of home sales for the four ethinic regions in the housing transaction sample. There are 217,452 housing transactions from 2006 to 2013 in the entire sample. The four main ethnic regions include Asia, Australia, Europe and Others.



Panel B: Dollar Volume of Home Sales By Region

This figure shows the number of home sales for the four ethinic regions in the housing transaction sample. The total dollar volume of housing transactions is $145.9 billion Australian dollar from 2006 to 2013 in the entire sample. The four main ethnic regions include Asian, Australia, Europe and Others.

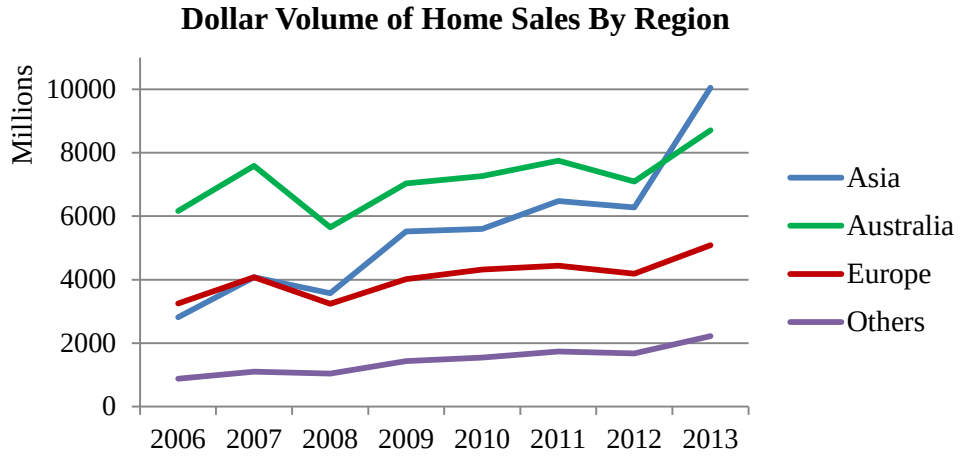


Table 1
Summary Statistics of Home Sales

The table reports mean summary statistics for home sales in the Sydney metropolitan area from 2006 to 2013 for the entire sample and for the top twenty buyer ethnicity groups by housing sales. The data is obtained from Australian Property Monitors. Buyer ethnicity is classified based on surname of the buyer. Price is in thousands of Australian dollars. House is a dummy variable equal to one for a free-standing house and zero otherwise (e.g. apartment (condominium), semi, studio, townhouse, villa, etc). Size is the area size of the home in 1,000 square feet. N_Bed is the number of bedrooms in the home. N_Bath is the number of bathrooms in the home. Parking is a dummy variable equal to one if the home has parking. Auction is a dummy variable equal to one if the home was sold at auction.

Ethnicity	Price	House	Size	N_Bed	N_Bath	Parking	Auction	N	%
Australian	758.06	0.6	4.08	2.93	1.62	0.83	0.16	75,529	34.7%
Chinese	670.25	0.48	3.26	2.89	1.7	0.89	0.15	39,840	18.3%
Arabic	534.98	0.73	4.9	3.07	1.51	0.88	0.18	21,429	9.9%
Indian	536.2	0.62	4.06	3.01	1.59	0.9	0.12	20,164	9.3%
Irish	754.15	0.58	3.91	2.90	1.60	0.83	0.16	14,920	6.9%
Italian	662.8	0.65	4.39	2.95	1.58	0.87	0.17	12,512	5.8%
Vietnamese	485.13	0.71	4.53	3.05	1.48	0.86	0.15	8,246	3.8%
Greek	718.62	0.65	4.1	2.93	1.55	0.85	0.24	5,313	2.4%
German	776.74	0.58	3.94	2.91	1.64	0.82	0.18	2,746	1.3%
French	696.5	0.62	4.14	2.94	1.6	0.83	0.17	2,211	1.0%
Spanish	508.07	0.58	3.59	2.92	1.52	0.87	0.1	1,975	0.9%
Korean	656.56	0.57	4.12	3.01	1.73	0.91	0.15	1,852	0.9%
Slovakian	543.9	0.59	3.77	2.90	1.51	0.88	0.14	1,440	0.7%
Portuguese	567.48	0.58	3.53	2.86	1.54	0.88	0.13	1,424	0.7%
Polish	630	0.54	3.55	2.82	1.53	0.85	0.14	1,211	0.6%
Maltese	532.55	0.72	5.27	3.07	1.52	0.88	0.1	927	0.4%
Indonesian	584.53	0.51	3.03	2.79	1.63	0.88	0.13	870	0.4%
Dutch	762.32	0.57	4.19	2.88	1.62	0.84	0.13	696	0.3%
Sri Lanka	580.65	0.65	4.42	3.02	1.61	0.9	0.11	612	0.3%
Japanese	670.94	0.5	3.17	2.64	1.53	0.84	0.11	553	0.3%
All	671.1	0.6	4.01	2.94	1.61	0.86	0.16	217,452	100.0%

Table 2
Culture Value and Cultural Distance by Ethnicity

Panel A: Six-Dimensional Culture Values based on Hofstede (2001)

This table shows the culture value based on Hoftede (2001)'s widely quoted theory of cultural dimensions for the top 15 ethnicities in Australia's population (2011 Census). In this framework, national cultures are placed in a six-dimensional space. The six culture dimensions are power distance (PDI), individualism vs. collectivism (IDV), acceptance of gender difference (MAS), uncertainty avoidance (UAI), long-term orientation (LTOWVS) and indulgence vs. restraint (IVR). Refer to Figure 1 for a detailed explanation for each dimension.

Ethnicity	Rank in population (2011 census)	PDI	IDV	MAS	UAI	LTOWVS	IVR
Australian/British	1	38	90	61	51	21	71
Chinese	2	80	20	66	30	87	24
Irish	3	28	70	68	35	24	65
Italian	4	50	76	70	75	61	30
Arabic	5	80	38	53	68	23	34
Indian	6	77	48	56	40	51	26
Greek	7	60	35	57	112	45	50
Vietnamese	8	70	20	40	30	57	35
Filipino	9	94	32	64	44	27	42
German	10	35	67	66	65	83	40
Korean	11	60	18	39	85	100	29
Maltese	12	56	59	47	96	47	66
Croatian	13	73	33	40	80	58	33
Polish	14	68	60	64	93	38	29
Serbian	15	86	25	43	92	52	28

Panel B: Summary Statistics of Cultural Distance between Buyer and Suburb

The table reports summary statistics for the cultural distance of the buyer to the home's suburb across housing sales in the Sydney metropolitan area from 2006 to 2013 for the top twenty buyer ethnicity groups in the sample by sales. Buyer ethnicity is classified based on surname of the buyer. Suburb demographic information is from the Australian Bureau of Statistics Census in 2006 and 2011. Cultural distance is measured as the Euclidean distance between the 6-D Hofstede cultural value of the buyer's ethnicity to the population weighted culture score in the home's suburb using the ancestry of the population.

Ethnicity	Mean	Std Dev	Q1	Median	Q3	Min	Max	N
Australian	1.35	0.47	1.00	1.25	1.58	0.59	3.30	75,529
Chinese	2.50	0.55	2.10	2.53	2.94	1.19	3.60	39,840
Arabic	2.29	0.47	1.95	2.29	2.68	1.35	3.44	21,429
Indian	2.13	0.32	1.95	2.15	2.37	0.81	2.71	20,164
Irish	1.63	0.40	1.31	1.52	1.83	0.87	3.03	14,920
Italian	2.03	0.25	1.87	1.99	2.16	1.17	3.14	12,512
Vietnamese	2.43	0.54	2.09	2.42	2.84	1.35	3.57	8,246
Greek	2.92	0.36	2.64	2.96	3.21	1.91	3.88	5,313
German	1.79	0.27	1.59	1.71	1.93	1.42	3.00	2,746
French	2.70	0.14	2.64	2.73	2.80	2.24	3.17	2,211
Spanish	2.60	0.19	2.48	2.64	2.75	2.05	2.94	1,975
Korean	3.05	0.38	2.73	3.09	3.36	2.27	3.86	1,852
Slovakian	3.67	0.18	3.54	3.68	3.83	3.01	4.10	1,440
Portuguese	3.46	0.37	3.20	3.51	3.75	2.59	4.17	1,424
Polish	2.73	0.17	2.63	2.76	2.86	2.20	3.12	1,211
Maltese	2.75	0.16	2.70	2.78	2.86	2.20	3.30	927
Indonesian	2.77	0.47	2.45	2.78	3.14	1.64	3.80	870
Dutch	3.08	0.19	2.94	3.04	3.18	2.80	3.74	696
Sri Lanka	2.75	0.36	2.53	2.79	3.03	1.81	3.42	612
Japanese	3.29	0.09	3.27	3.32	3.35	2.88	3.72	553
All	1.99	0.72	1.38	1.99	2.52	0.59	4.37	217,452

Table 3
Cultural distance and Homebuyers' Location Choice (Probit Estimation)

The table reports the estimation of the effect of cultural distance on the choice of home location, using Probit models in Equation 3 as:

$\Pr(Buy_{jst} = 1|X) = F(\beta_1 CD_{jst} + \beta_2 lagybuy_{jst} + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst})$ --- (3)

The sample is home location choice for each ethnicity group at each suburb in each year/quarter level. The dependent variable is a dummy which takes the value of 1 if a given ethnicity buys in a suburb in a given quarter and 0 otherwise. We control for suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. CD is Cultural distance, and is measured as the Euclidean distance between the Hofstede cultural dimensions of the buyer's ethnicity to the population weighted ethnicity cultural dimensions in the home's suburb. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede culture scores. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

	Dependent Variable: =1 if a ethnicity buys in a suburb in a given			
Variables	(1)	(2)	(3)	(4)
Lastyear_buy	0.287***	0.348***	0.309***	0.358***
	(0.007)	(0.007)	(0.007)	(0.007)
ACD6	-0.828***			
	(0.008)			
BCD6		-0.816***		
		(0.010)		
ACD4			-0.811***	
			(0.008)	
BCD4				-0.886***
				(0.012)
Constant	2.607***	1.498***	2.055***	1.405***
	(0.076)	(0.073)	(0.074)	(0.072)
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Pseudo R-square	0.2756	0.2718	0.2679	0.2648
Number of Observations	819,580	819,580	838,640	838,640

Table 4
Baseline Effect of Cultural distance on Housing Price

The table reports OLS regression result of cultural distance of homebuyer to a suburb on the housing transaction price, controling for property level, housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. Specifically we run the following regression from Equation 2:

$\ln(P_{ist}) = \alpha_t + \beta_k CD_{ist} + property\ char_i + \mu_s + \delta_i + \gamma_t + \tau_t + \varepsilon_{it}$ --- (2)

Where $ln(P_{ist})$ denotes logarithm of housing prices paid by buyer of sale *i* at suburb *s* at time *t*; $property\ char$ are various property characteristics such as number of bedrooms, number of bathrooms, parking, property type and area size (details of all variables are in Appendix 4); μ_s is the suburb location specific fixed effects; δ_i is buyer's ethnicity fixed effects. γ_t is year/quarter fixed effects and τ_t is a monthly time trend. Standard errors are clustered at the suburb level. The dependent variables are log home price in all the specifications. CD is Cultural distance, and is measured as the Euclidean distance between the Hofstede cultural dimensions of the buyer's ethnicity to the population weighted ethnicity cultural dimensions in the home's suburb. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.011***			
	(0.003)			
BCD6		-0.009*		
		(0.005)		
ACD4			-0.013***	
			(0.003)	
BCD4				-0.012**
				(0.006)
New Development	0.143***	0.143***	0.142***	0.142***
	(0.008)	(0.008)	(0.008)	(0.008)
Auction	0.055***	0.055***	0.055***	0.055***
	(0.003)	(0.003)	(0.003)	(0.003)
Number of Bedrooms	0.119***	0.119***	0.119***	0.119***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bathrooms	0.128***	0.128***	0.128***	0.128***
	(0.003)	(0.003)	(0.003)	(0.003)
Has Parking	0.083***	0.083***	0.083***	0.083***
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	7.878***	7.029***	8.332***	7.109***
	(0.896)	(0.904)	(0.903)	(0.896)
Housing Characteristics	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Clustered By ...	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8629	0.8628	0.8630	0.8629
Number of Observations	217,452	217,452	218,869	218,869

Table 5
Effect of Cultural distance on Housing Price
(Address Endogeneity Concern using 2SLS)

The table reports two stage least squares (2SLS) estimation result of cultural distance of homebuyer to a suburb on the housing transaction price, controling for housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend.
In the first stage, we obtain estimated CD using Genetic distance (GD) and other control variables. In the second stage, we estimate the hedonic housing price model using estimated CD, and other controls.

$CD_{jst} = F(\alpha_0 + \beta_1 GD_{jst} + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst})$ --- (stage 1)

$ln(price) = F(\alpha_0 + \beta_1 CD_{jst} + housing\ char + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst})$ --- (stage 2)

Genetic distance is used as an instrument for cultural distance in the first stage estimation The standard errors are clustered at the buyer ethnicity level. The dependent variables are log home price in all the specifications. CD is Cultural distance, and is measured as the Euclidean distance between the Hofstede cultural dimensions of the buyer's ethnicity to the population weighted ethnicity cultural dimensions in the home's suburb. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.017***			
	(0.004)			
BCD6		-0.014***		
		(0.003)		
ACD4			-0.029***	
			(0.004)	
BCD4				-0.018***
				(0.003)
New Development	0.088***	0.143***	0.143***	0.143***
	(0.008)	(0.008)	(0.008)	(0.008)
Auction	0.088***	0.056***	0.055***	0.055***
	(0.004)	(0.003)	(0.003)	(0.003)
Number of Bedrooms	0.182***	0.119***	0.119***	0.119***
	(0.006)	(0.004)	(0.004)	(0.004)
Number of Bathrooms	0.110***	0.128***	0.129***	0.129***
	(0.004)	(0.003)	(0.003)	(0.003)
Has Parking	0.067***	0.084***	0.085***	0.085***
	(0.007)	(0.007)	(0.007)	(0.007)
Constant	12.583***	13.182***	13.207***	13.184***
	(0.026)	(0.021)	(0.021)	(0.021)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8315	0.8616	0.8616	0.8617
Number of Observations	217,452	217,452	218,869	218,869

Table 6
Effect of Cultural Distance on Housing Price
(Address Selection Bias using Heckman Selection Model)

The table reports Heckman selection model estimation result of cultural distance of homebuyer to a suburb on the housing transaction price, controling for housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. ***lastyear_buy*** is used as an instrument for the suburb location choice dummy in the first stage estimation. Please refer to Table 3 for more details on the location choice model. The dependent variables are log home price in all the specifications. CD is Cultural distance, and is measured as the Euclidean distance between the Hofstede cultural dimensions of the buyer's ethnicity to the population weighted ethnicity cultural dimensions in the home's suburb. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede. The standard errors are clustered at the suburb level, and are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.015***			
	(0.004)			
BCD6		-0.012**		
		(0.005)		
ACD4			-0.016***	
			(0.005)	
BCD4				-0.014**
				(0.007)
New Development	0.140***	0.14***	0.14***	0.14***
	(0.008)	(0.008)	(0.008)	(0.008)
Auction	0.056***	0.056***	0.056***	0.056***
	(0.003)	(0.003)	(0.003)	(0.003)
Number of Bedrooms	0.121***	0.121***	0.122***	0.122***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bathrooms	0.129***	0.129***	0.129***	0.129***
	(0.003)	(0.003)	(0.003)	(0.003)
Has Parking	0.084***	0.084***	0.085***	0.085***
	(0.007)	(0.007)	(0.007)	(0.007)
Inverse Mills Ratio	0.010	0.007	0.007	0.004
	(0.006)	(0.006)	(0.006)	(0.006)
Constant	7.235***	7.21***	7.239***	7.223***
	(0.898)	(0.898)	(0.901)	(0.901)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8611	0.8610	0.8612	0.8612
Number of Observations	217,452	217,452	218,869	218,869

Table 7
Home Cultural Preference and Ethnicity Region

The table reports coefficient estimates regressing log sales price on cultural distance interacted with homebuyer's ethnicity region. The dependent variables are log home price in all the specifications. Control variables include housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. Standard errors are clustered at the suburb level. Model 1 uses ACD6 as the cultural distance measure. Model 2 uses BCD6 as the cultural distance measure. Model 3 uses ACD4 as the cultural distance measure. Model 4 uses BCD4 as the cultural distance measure. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using birthplace of the suburb's population and 4-dimension Hofstede culture scores. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: OLS Estimation

	Dependent variable: ln(price)			
	(1)	(2)	(3)	(4)
Variables	ACD6	BCD6	ACD4	BCD4
Intercept	7.203***	7.202***	7.212***	7.212***
	(0.898)	(0.898)	(0.901)	(0.902)
CD*Africa	0.055	0.711	0.038	0.042
	(0.072)	(0.451)	(0.108)	(0.157)
CD*Australia	-0.004	0.007	-0.003	0.013
	(0.005)	(0.007)	(0.005)	(0.009)
CD*East Asia	-0.022***	-0.029***	-0.029***	-0.048***
	(0.005)	(0.006)	(0.006)	(0.008)
CD*South Asia	-0.021***	-0.029***	-0.029***	-0.048***
	(0.006)	(0.008)	(0.007)	(0.011)
CD*South East Asia	-0.032***	-0.032***	-0.034***	-0.042***
	(0.006)	(0.006)	(0.006)	(0.007)
CD*Middle East	-0.002	-0.020	-0.008	-0.043***
	(0.006)	(0.017)	(0.006)	(0.013)
CD*Eastern Europe	-0.011	-0.015	-0.012	-0.001
	(0.012)	(0.018)	(0.015)	(0.025)
CD*Northern Europe	-0.023	-0.021	-0.072	-0.071
	(0.029)	(0.040)	(0.045)	(0.059)
CD*Southern Europe	-0.016**	-0.016	-0.015***	0.001
	(0.007)	(0.015)	(0.006)	(0.009)
CD*Western Europe	-0.006	0.007	-0.006	0.017
	(0.006)	(0.011)	(0.006)	(0.015)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8611	0.8611	0.8612	0.8612
Number of Observations	217,452	217,452	218,869	218,869

Panel B: 2SLS Estimation (to Address Endogeneity Concern)

	Dependent variable: ln(price)			
	(1)	(2)	(3)	(4)
Variables	ACD6	BCD6	ACD4	BCD4
Intercept	12.56***	13.16***	13.165***	13.161***
	(0.035)	(0.023)	(0.027)	(0.023)
CD*Africa	0.000	0.004	-0.023***	-0.022***
	(0.072)	(0.056)	(0.005)	(0.004)
CD*Australia	-0.009	-0.009	-0.009	-0.011
	(0.011)	(0.01)	(0.013)	(0.012)
CD*East Asia	-0.016***	-0.014**	-0.024***	-0.019**
	(0.006)	(0.006)	(0.008)	(0.008)
CD*South Asia	-0.013	-0.010	-0.015	-0.012
	(0.009)	(0.009)	(0.012)	(0.011)
CD*South East Asia	-0.012*	-0.011***	-0.016**	-0.013***
	(0.006)	(0.003)	(0.008)	(0.005)
CD*Middle East	-0.007	-0.007	-0.009	-0.006
	(0.008)	(0.005)	(0.012)	(0.008)
CD*Eastern Europe	-0.004	-0.002	-0.004	-0.002
	(0.016)	(0.009)	(0.016)	(0.010)
CD*Northern Europe	-0.029	-0.021	-0.025	-0.026
	(0.026)	(0.017)	(0.027)	(0.022)
CD*Southern Europe	-0.007	-0.004	-0.007	-0.005
	(0.008)	(0.005)	(0.010)	(0.007)
CD*Western Europe	-0.008	-0.005	-0.008	-0.006
	(0.010)	(0.008)	(0.013)	(0.010)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8317	0.8622	0.8623	0.8623
Number of Observations	217,452	217,452	218,869	218,869

Panel C: Heckman Selection Model Estimation (to Address Sample Selection Bias)

	Dependent variable: ln(price)			
	(1)	(2)	(3)	(4)
Variables	ACD6	BCD6	ACD4	BCD4
Intercept	7.195***	7.197***	7.206***	7.211***
	(0.898)	(0.898)	(0.901)	(0.902)
CD*Africa	0.044	0.710	0.031	0.041
	(0.073)	(0.450)	(0.109)	(0.157)
CD*Australia	-0.008	0.005	-0.006	0.013
	(0.005)	(0.007)	(0.006)	(0.010)
CD*East Asia	-0.025***	-0.031***	-0.032***	-0.048***
	(0.006)	(0.006)	(0.007)	(0.008)
CD*South Asia	-0.026***	-0.031***	-0.032***	-0.048***
	(0.006)	(0.008)	(0.007)	(0.011)
CD*South East Asia	-0.039***	-0.035***	-0.039***	-0.043***
	(0.006)	(0.007)	(0.007)	(0.008)
CD*Middle East	-0.006	-0.022	-0.012**	-0.043***
	(0.006)	(0.017)	(0.006)	(0.013)
CD*Eastern Europe	-0.02	-0.019	-0.018	-0.002
	(0.014)	(0.019)	(0.016)	(0.026)
CD*Northern Europe	-0.033	-0.025	-0.079*	-0.073
	(0.029)	(0.041)	(0.045)	(0.059)
CD*Southern Europe	-0.024***	-0.020	-0.020***	0.000
	(0.007)	(0.016)	(0.007)	(0.010)
CD*Western Europe	-0.014*	0.003	-0.011	0.017
	(0.008)	(0.012)	(0.008)	(0.016)
Inverse Mills Ratio	0.013**	0.006	0.009	0.001
	(0.006)	(0.006)	(0.006)	(0.006)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8611	0.8611	0.8612	0.8612
Number of Observations	217,452	217,452	218,869	218,869

Table 8
Regional Level of Home Culture Preference

The table reports the effect of cultural distance on housing price for each region subsample. The dependent variables are log home price in all the specifications. Control variables include housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. Standard errors are clustered at the suburb level. Model 1 uses ACD6 as the cultural distance measure. Model 2 uses BCD6 as the cultural distance measure. Model 3 uses ACD4 as the cultural distance measure. Model 4 uses BCD4 as the cultural distance measure. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using birthplace of the suburb's population and 4-dimension Hofstede culture scores.
Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A. Asia

	Dependent variable: log(price)			
Variable	(1)	(2)	(3)	(4)
Intercept	7.49***	7.529***	7.491***	7.544***
	(1.145)	(1.144)	(1.146)	(1.144)
ACD6	-0.022***			
	(0.006)			
BCD6		-0.03***		
		(0.007)		
ACD4			-0.035***	
			(0.007)	
BCD4				-0.044***
				(0.009)
New Development	0.142***	0.142***	0.142***	0.142***
	(0.009)	(0.009)	(0.009)	(0.009)
Auction	0.056***	0.056***	0.056***	0.056***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bedrooms	0.105***	0.105***	0.105***	0.105***
	(0.003)	(0.003)	(0.003)	(0.003)
Number of Bathrooms	0.114***	0.114***	0.114***	0.114***
	(0.003)	(0.003)	(0.003)	(0.003)
Has Parking	0.063***	0.063***	0.063***	0.063***
	(0.007)	(0.007)	(0.007)	(0.007)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8690	0.8690	0.8691	0.8691
Number of Observations	73,047	73,047	73,047	73,047

Panel B. Australia

	Dependent variable: log(price)			
Variable	(1)	(2)	(3)	(4)
Intercept	3.375**	3.376**	3.33**	3.372**
	(1.438)	(1.429)	(1.438)	(1.43)
ACD6	0.048			
	(0.043)			
BCD6		0.071**		
		(0.031)		
ACD4			0.085	
			(0.057)	
BCD4				0.093**
				(0.04)
New Development	0.146***	0.146***	0.146***	0.146***
	(0.011)	(0.011)	(0.011)	(0.011)
Auction	0.055***	0.055***	0.055***	0.055***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bedrooms	0.145***	0.145***	0.145***	0.145***
	(0.005)	(0.005)	(0.005)	(0.005)
Number of Bathrooms	0.143***	0.143***	0.143***	0.143***
	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.100***	0.100***	0.100***	0.100***
	(0.008)	(0.008)	(0.008)	(0.008)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8558	0.8559	0.8558	0.8559
Number of Observations	75,529	75,529	75,529	75,529

Panel C. Europe

	Dependent variable: log(price)			
Variable	(1)	(2)	(3)	(4)
Intercept	11.345***	11.362***	11.346***	11.363***
	(1.725)	(1.726)	(1.724)	(1.725)
ACD6	-0.013***			
	(0.004)			
BCD6		-0.006		
		(0.007)		
ACD4			-0.015***	
			(0.005)	
BCD4				-0.014
				(0.009)
New Development	0.129***	0.129***	0.129***	0.129***
	(0.01)	(0.01)	(0.01)	(0.01)
Auction	0.059***	0.059***	0.059***	0.059***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bedrooms	0.129***	0.129***	0.129***	0.129***
	(0.005)	(0.005)	(0.005)	(0.005)
Number of Bathrooms	0.134***	0.134***	0.134***	0.134***
	(0.004)	(0.004)	(0.004)	(0.004)
Has Parking	0.086***	0.086***	0.086***	0.086***
	(0.008)	(0.008)	(0.008)	(0.008)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8602	0.8601	0.8602	0.8601
Number of Observations	47,067	47,067	47,067	47,067

Table 9
Individualsim and Home Culture Preference

The table shows the effect of individualism on the degree of home culture preference, by interancting standardised individualism score of the homebuyer's ethinicity with the cultural distance mesure. The dependent variables are log home price in all the specifications. Control variables include housing characteristics, suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. Standard errors are clustered at the suburb level. Model 1 uses ACD6 as the cultural distance measure. Model 2 uses BCD6 as the cultural distance measure. Model 3 uses ACD4 as the cultural distance measure. Model 4 uses BCD4 as the cultural distance measure. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using birthplace of the suburb's population and 4-dimension Hofstede culture scores.Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

	Dependent variable: log(price)			
Variables	(1)	(2)	(3)	(4)
Intercept	7.261***	7.3***	7.283***	7.346***
	(0.897)	(0.897)	(0.901)	(0.901)
ACD6	-0.025***			
	(0.007)			
ACD6*Individualism	0.006***			
	(0.002)			
BCD6		-0.038***		
		(0.008)		
BCD6* Individualism		0.011***		
		(0.003)		
ACD4			-0.036***	
			(0.007)	
ACD4* Individualism			0.009***	
			(0.003)	
BCD4				-0.057***
				(0.009)
BCD4* Individualism				0.018***
				(0.003)
Individualism	-0.010	-0.026***	-0.017***	-0.037***
	(0.006)	(0.008)	(0.006)	(0.007)
New Development	0.14***	0.14***	0.14***	0.14***
	(0.008)	(0.008)	(0.008)	(0.008)
Auction	0.056***	0.056***	0.056***	0.056***
	(0.003)	(0.003)	(0.003)	(0.003)
Number of Bedrooms	0.121***	0.121***	0.122***	0.122***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Bathrooms	0.129***	0.129***	0.129***	0.129***
	(0.003)	(0.003)	(0.003)	(0.003)
Has Parking	0.084***	0.084***	0.085***	0.085***
	(0.007)	(0.007)	(0.007)	(0.007)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8611	0.8611	0.8612	0.8612
Number of Observations	217,452	217,452	218,869	218,869

Table 10
Robustness Check with Seller Ethnicity Fixed Effect

The table offers robustness check of the baseline result when controling for seller ethnicity fixed effects, by estimating the following Equation 2 and including seller ethnicity fixed effects. The dependent variable is logarithm of home price. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede culture scores. Standard errors are clustered by suburbs and are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: OLS Estimation with Seller Ethnicity Fixed Effect

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.011***			
	(0.003)			
BCD6		-0.009*		
		(0.005)		
ACD4			-0.014***	
			(0.003)	
BCD4				-0.013**
				(0.006)
Housing Characteristics	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Seller Ethnicity FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8586	0.8585	0.8586	0.8586
Number of Observations	162,079	162,079	163,775	163,775

Panel B: Heckman Selection Model Estimation with Seller Ethnicity Fixed Effect

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.016***			
	(0.005)			
BCD6		-0.013**		
		(0.006)		
ACD4			-0.018***	
			(0.005)	
BCD4				-0.016**
				(0.007)
Housing Characteristics	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Seller Ethnicity FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8586	0.8585	0.8586	0.8586
Number of Observations	162,079	162,079	163,775	163,775

Panel C: 2SLS Estimation with Seller Ethnicity Fixed Effect

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.014***			
	(0.004)			
BCD6		-0.008***		
		(0.003)		
ACD4			-0.016***	
			(0.004)	
BCD4				-0.010***
				(0.003)
Housing Characteristics	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Seller Ethnicity FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8298	0.8297	0.8298	0.8298
Number of Observations	162,079	162,079	163,775	163,775

Table 11
Subsample Robustness Check by Excluding Chinese Homebuyers

The table reports estimation result of the following equation by excluding Chinese Homebuyers from the sample using Equation 2:

$\ln(P_{ist}) = \alpha_t + \beta_k CD_{ist} + property\ char_i + \mu_s + \delta_i + \gamma_t + \tau_t + \varepsilon_{it}$ --- (2)

Ln(Home Price) ~ CD + Buyer Ethnicity FE + Year/Quarter FE + Monthly Time Trend

The dependent variable is logarithm of home price. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede culture scores. Standard errors are clustered by suburbs and are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: OLS Estimation Excluding Chinese Homebuyers from Sample

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.009***			
	(0.002)			
BCD6		-0.005		
		(0.004)		
ACD4			-0.011***	
			(0.002)	
BCD4				-0.008*
				(0.004)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8619	0.8618	0.8620	0.8620
Number of Observations	177,612	177,612	179,029	179,029

Panel B: Heckman Selection Model Excluding Chinese Homebuyers from Sample

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.009***			
	(0.003)			
BCD6		-0.003		
		(0.004)		
ACD4			-0.010***	
			(0.004)	
BCD4				-0.004
				(0.005)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8619	0.8618	0.8620	0.8620
Number of Observations	177,612	177,612	179,029	179,029

Panel C: 2SLS Estimation Excluding Chinese Homebuyers from Sample

	Dependent variable: ln(price)			
Variable	(1)	(2)	(3)	(4)
ACD6	-0.015***			
	(0.004)			
BCD6		-0.010***		
		(0.002)		
ACD4			-0.015***	
			(0.004)	
BCD4				-0.011***
				(0.003)
Housing Characteristics	Yes	Yes	Yes	Yes
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Monthly Time Trend	Yes	Yes	Yes	Yes
S.E. Cluster	Suburb	Suburb	Suburb	Suburb
Adjusted R-square	0.8324	0.8322	0.8326	0.8325
Number of Observations	177,612	177,612	179,029	179,029

Appendix 1
Ethnicity Composition of Australia Population

We collect top overseas countries of birth by percentage of the Australian population for each Australian Census between 1954 (just prior to the relaxation of the White Australia Policy in 1958) and 2011. The table reports for each overseas country of birthplace, the census year entry of the birthplace into the top ten countries (indicating the start year of migration wave for that country), the census year when the country reaches its peak percentage in the total population, and the total percentage of the Australian population when the birthplace was at its peak. ***Year of Entry***: the census year entry of the birthplace into the top ten countries. ***Peak Year***: the census year when the country reaches its peak percentage in the total population. ***%Pop in Peak Year***: the percentage of population with birthplace in this country when the country reaches its peak percentage in the total population. The census year of entry and census year at peak provide a guide of the most recent immigration wave from the birthplace country.

Overseas Country of Birth	Year of Entry	Peak Year	%Pop in Peak Year
Ireland	1954	1954	0.50
UK	1954	1954	6.86
Poland	1954	1954	0.63
Germany	1954	1961	1.04
Netherlands	1954	1961	0.97
Greece	1954	1971	1.26
Italy	1954	1971	2.27
Malta	1954	1971	0.42
Lebanon	1981	1981	0.28
Yugoslavia	1954	1991	0.96
China	1991	2011	1.48
India	2001	2011	1.37
Malaysia	2011	2011	0.54
Philippines	1991	2011	0.80
Vietnam	1991	2011	0.86
New Zealand	1954	2011	2.25
South Africa	2006	2011	0.68

Appendix 2
Distribution of Homebuyers by Region and Ethnicities

The table reports the percentage of population by region and ethnicities of their origin in our sample. Sydney Population statistics are from the 2006 and 2011 Australian Bureau of Statistics Census. Panel A ranks the top 20 ethnicities by population in the 2011 Census. Panel B provides a full list of ethnicities used in the paper. Ethnicities marked with an astericks only have four cultural dimensions measured.

Panel A: Top 20 ethnicities by population rank in 2011 Census

Ethnicity	Region	%(2006 Census)	%(2011 Census)	Rank (2011 census)
Australian/British	Australia	45.16	42.97	1
Chinese	East Asia	7.84	8.99	2
Irish	Western Europe	4.47	4.56	3
Italian	Southern Europe	3.95	3.82	4
Arabic	Middle East	3.52	3.72	5
Indian	South Asia	2.2	3.1	6
Greek	Southern Europe	2.57	2.43	7
Vietnamese	South East Asia	1.65	1.9	8
Filipino	South East Asia	1.19	1.38	9
German	Western Europe	1.41	1.37	10
Korean	East Asia	1.04	1.26	11
Maltese	Southern Europe	0.75	0.71	12
Croatian	Eastern Europe	0.71	0.68	13
Polish	Eastern Europe	0.58	0.53	14
Serbian	Eastern Europe	0.63	0.52	15
Sri Lankan	South Asia	0.46	0.51	16
Turkish	Middle East	0.49	0.51	17
Dutch	Western Europe	0.49	0.47	18
Bangladeshi	South Asia	0.28	0.44	19
South African	Africa	0.41	0.43	20

Panel B: Complete list of ethinicity group by region and percentage in population

Region	Ethnicity	Percent of Population in 2006 Census	Percent of Population in 2011 Census
Africa	Moroccan	0.01	0.01
	South African*	0.41	0.43
Australia	Australian/British	45.16	42.97
East Asia	Chinese	7.84	8.99
	Japanese	0.29	0.30
	Korean	1.04	1.26
	Taiwanese	0.05	0.05
South Asia	Bangladeshi	0.28	0.44
	Indian	2.20	3.10
	Nepalese	0.09	0.41
	Sri Lankan	0.46	0.51
South East Asia	Indonesian	0.25	0.31
	Filipino	1.19	1.38
	Malaysian	0.08	0.09
	Singaporean	0.02	0.02
	Thai	0.23	0.33
	Vietnamese	1.65	1.90
Middle East	Arabic	3.52	3.72
	Israeli/Jewish*	0.08	0.19
	Turkish	0.49	0.51
Eastern Europe	Croatian	0.71	0.68
	Czech	0.11	0.11
	Hungarian	0.31	0.28
	Polish	0.58	0.53
	Romanian	0.08	0.08
	Russian	0.39	0.40
	Serbian	0.63	0.52
	Slovak Republic	0.05	0.06
Northern Europe	Danish	0.07	0.07
	Estonian	0.03	0.03
	Finnish	0.05	0.05
	Latvian	0.06	0.05
	Lithuanian	0.05	0.05
	Norwegian	0.03	0.03
	Swedish	0.06	0.06
Southern Europe	Greek	2.57	2.43
	Italian	3.95	3.82
	Maltese	0.75	0.71
	Portuguese	0.33	0.33
Western Europe	French	0.26	0.28
	Dutch	0.49	0.47
	German	1.41	1.37
	Irish	4.47	4.56
	Swiss	0.07	0.07
Total		82.85	83.96

Appendix 3
List of Housing Characteristic Variables

Variable	Description
Beds	Number of beds
Baths	Number of bathrooms
Auction dummy	1 if the home was sold at auction, 0 otherwise
New development dummy	1 if the home was a new development, 0 otherwise
Multiple parking	1 if home has two or more parking spots, 0 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 metres)
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 4
Cultural distance and Homebuyers' Location choice (IVProbit)

The table reports the IV Probit estimation of the effect of cultural distance on the choice of home location, to provide robustness for the ordinary Probit model in Table 3.

In the first stage, we obtain estimated CD using Genetic distance (GD) and other control variables. In the second stage, we estimate the locaiton choice dummy using estimated CD, and other controls.

$CD_{jst} = F(\alpha_0 + \beta_1 GD_{jst} + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst})$ --- (stage 1)

$\Pr(Buy_{jst} = 1|X) = F(\alpha_0 + \beta_1 CD_{jst} + \mu_s + \delta_j + \gamma_t + \varepsilon_{jst})$ --- (stage 2)

The sample is home location choice for each ethnicity group at each suburb in each year/quarter level. The dependent variable is a dummy which takes the value of 1 if a given ethnicity buys in a suburb in a given quarter and 0 otherwise. We control for suburb fixed effect, buyer ethinicity fixed effect, year/quarter fixed effect, and monthly time trend. CD is Cultural distance, and is measured as the Euclidean distance between the Hofstede cultural dimensions of the buyer's ethnicity to the population weighted ethnicity cultural dimensions in the home's suburb. The 6-dimension cultural dimensions are power distance, uncertainty avoidance, individualism versus collectivism, masculinity versus femininity, long-term orientation and indulgence versus restraint. The 4-dimension cultural dimensions exclude long-term orientation and indulgence versus restraint. ACD6 is calculated using ancestry of the suburb's population and 6-dimension Hofstede culture scores. BCD6 is calculated using country of birthplace of the suburb's population and 6-dimension Hofstede culture scores. ACD4 is calculated using ancestry of the suburb's population and 4-dimension Hofstede culture scores. BCD4 is calculated using country of birthplace of the suburb's population and 4-dimension Hofstede culture scores. Standard errors are in parentheses. ***, **, * signifies statistical significance at the 1, 5 and 10 percent level, respectively.

Panel A: Probit Estimation using Instrument Variable Approach (First stage regression)

	Dependant Variable: 4 different version of cultural distance			
Variables	CD(Ancestry, 6D)	CD(Birthplace, 6D)	CD(Ancestry, 4D)	CD(Birthplace, 4D)
Genetic Distance	0.001***	0.001***	0.001***	0.001***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
Lastyear_buy	-0.114***	-0.060***	-0.1024***	-0.053***
	(0.011)	(0.001)	(0.001)	(0.001)
Constant	2.998***	3.107***	1.634***	2.607***
	(0.032)	(0.002)	(0.070)	(0.002)
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Adjusted R-square	0.8056	0.9372	0.8272	0.9377
Number of Obs	819,580	819,580	838,640	838,640

Panel B: Probit Estimation using Instrument Variable Approach (Second Stage)

	Dependent Variable: =1 if a ethnicity buys in a suburb in a given quarter			
Variables	(1)	(2)	(3)	(4)
Lastyear_buy	0.333***	0.374***	0.348***	0.389***
	(0.007)	(0.007)	(0.007)	(0.007)
$\widehat{CD}$ (Ancestry, 6D)	-0.828***			
	(0.024)			
$\widehat{CD}$ (Birthplace, 6D)		-1.049***		
		(0.300)		
$\widehat{CD}$ (Ancestry, 4D)			-0.854***	
			(0.024)	
$\widehat{CD}$ (Birthplace, 4D)				-1.061***
				(0.029)
Constant	2.038***	2.838***	1.633***	2.343***
	(0.086)	(0.104)	(0.070)	(0.088)
Buyer Ethnicity FE	Yes	Yes	Yes	Yes
Suburb FE	Yes	Yes	Yes	Yes
Year/Quarter FE	Yes	Yes	Yes	Yes
Wald Test P-value	<0.001	<0.001	<0.001	<0.001
Number of Obs	819,580	819,580	838,640	838,640