

Estimating Racial Price Differentials in the Housing Market*

Patrick Bayer
Duke University and NBER

Marcus Casey
University of Illinois at Chicago

Fernando Ferreira
The Wharton School, University of Pennsylvania and NBER

Robert McMillan
University of Toronto and NBER

Abstract

This paper uses unique panel data covering over two million repeat-sales housing transactions from four metropolitan areas to test for the presence of racial price differentials in the housing market. Drawing on the strengths of these data, our research design controls carefully for unobserved differences in the quality of neighborhoods and the homes purchased by buyers of each race. We find that black and Hispanic homebuyers pay premiums of about three percent on average across the four cities, differences that are not explained by variation in buyer income, wealth or access to credit. Further, the estimated premiums do not vary significantly with the racial composition of the neighborhood; nor, strikingly, do they vary with the race of the seller. This latter finding suggests that racial prejudice on the part of sellers is not the primary explanation for the robust premiums we uncover. The results have implications for the evolution of racial differences in wealth and home ownership and the persistence of residential segregation.

JEL Codes: H0, J15, K4, R2, R3

Keywords: Price Differentials, Housing Market, Home Sales, Racial Differences, Racial Prejudice, Repeat Sales

* The authors would like to thank Jeff Brown, Hanming Fang, Robert Kaestner, Ron Laschever, Darren Lubotsky, Marie Mora, Ben Ost, Steve Rivkin, Jake Vigdor and seminar participants at the Harris School of Public Policy, Purdue University, the University of Illinois, the University of Windsor, and the AEA and IEA Meetings for helpful comments. Financial support is gratefully acknowledged from the National Science Foundation through grant SES-0721136, the Research Sponsor Program of the Zell/Lurie Real Estate Center at Wharton, and SSHRC. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

I. Introduction

For much of the twentieth century, overt forms of centralized racial discrimination were widespread and deeply-ingrained in US housing markets. Restrictive covenants written into the deeds of homes prevented the re-sale of properties to minority buyers; redlining in the mortgage industry substantially raised the costs of borrowing in predominantly black neighborhoods; and blockbusting and other racially-motivated tactics by realtors reinforced strict racial segregation in many cities.¹ While these institutionalized discriminatory practices had largely disappeared by the end of the twentieth century,² important questions remain about the role of race as a determinant of prices in the housing market. At a basic level, do different-race homebuyers pay similar prices for comparable housing? And to what extent does decentralized discrimination on the part of individual home sellers or other market agents continue to play a significant role in the market? While these questions have important implications for the dynamics of racial inequality and residential segregation, the difficulty of designing and implementing studies that compare market outcomes for different-race households in essentially identical circumstances has limited our ability to provide clear and convincing answers.

This paper seeks to shed new light on these questions. It has two primary goals: (i) to estimate the extent of any racial price differentials in the housing market – differentials that may be indicative of discriminatory behavior, and (ii) to test whether race prejudice on the part of sellers³ can provide the primary explanation for any measured price differences.

Identifying the extent and nature of any racial price differentials in most market settings is challenging because a variety of other factors might generate correlations between sales prices and buyer race. An obvious reason why households of different races pay different prices for housing is that they purchase houses of varying quality, both in terms of structural and neighborhood amenities. To properly detect price differentials, therefore, one needs to compare the price that different-race buyers pay for essentially identical properties.⁴ The previous housing literature has pursued this strategy almost uniformly by comparing the prices of homes purchased in the same neighborhood at roughly the same time. But these traditional studies implicitly ignore unobserved differences in house quality within neighborhoods and, unsurprisingly, typically find

¹ See Yinger (1995, 1998) for a detailed discussion. Hirsch (1998) provides an account focusing on Chicago.

² See Cutler, Glaeser, and Vigdor (1999) for a characterization of the rise and decline of centralized discrimination in American cities over the course of the twentieth century.

³ In the discussion and empirical work that follows, we use the term “seller” to refer to the actual owner or the owner’s selling agent.

⁴ Controlling for differences in product attributes is potentially more straightforward in other contexts such as new automobile sales – see, for example, Ayres (1991), Ayres and Siegelman (1995), and Goldberg (1996) – where researchers can compare sales prices for cars of the same make, model, and options.

that white buyers, who have systematically higher levels of income and wealth, pay a ‘premium’ for housing.⁵

In this paper, we extend the existing literature by exploiting a unique data set that provides information on transaction prices and buyer and seller attributes for a large, representative sample of home sales in order to isolate variation in sales prices for comparable properties. Our data set is based on *all* housing transactions that occurred in four large metropolitan areas (Chicago, Baltimore/Washington DC (‘Maryland’ for shorthand), Los Angeles, and San Francisco) over a period spanning two decades, 1990-2008.⁶ For each transaction, we observe the sales price as well as detailed information about the house location and housing attributes. To these transactions data, we merge information about the race and income of the buyer by matching mortgage records, creating a rich data set featuring over two million transactions. For houses in our data set that sell more than once, we can also pin down the race of the seller.

Our research design capitalizes on the geographic and temporal specificity of these data, examining differences in prices for houses that are far more comparable, in terms of housing and neighborhood characteristics, than those studied in prior work. In particular, we estimate how the sales price varies with the race and ethnicity of the buyer in a repeat-sales specification that controls explicitly for the unobserved quality of the individual house as well as time-varying neighborhood effects. In essence, a repeat-sales approach provides a measure of the expected sales price of a house, given a prior sale and the estimated rate of appreciation in the neighborhood since it last sold. By comparing the price paid by buyers of different races relative to this expected sales price, we isolate racial differences in sales prices for comparable homes.

Our initial results indicate that black and Hispanic buyers pay a statistically significant premium for comparable housing in all four major city markets, ranging from 1 to 4 percent relative to their white counterparts; the average estimated premium pooling across all four markets is more than 2 percent for both black and Hispanic buyers in this initial analysis. The estimated premiums for black buyers are highest in the Chicago and Maryland areas, where they represent a reasonably large fraction of the population and segregation levels have historically been very high. There is weaker evidence for price differentials for black buyers in Los Angeles and San Francisco, where the estimated premiums are 1.1-1.2 percent. The estimated premiums for Hispanic buyers are fairly comparable in Chicago, Maryland, and San Francisco (around 2.5-

⁵ An exception is the study by Myers (2004). We discuss the previous literature in detail in Section III.

⁶ We observe data for the two California metropolitan areas from 1990 to 2008, and for Chicago and Maryland from 1997 to 2007.

3.0 percent), but only 1.2 percent in Los Angeles. Due to the large sample sizes involved, all of these estimates are statistically significantly different from zero.

While these initial results provide strong statistical evidence that black and Hispanic buyers pay more for comparable housing than whites, it is unclear whether these price differences are directly related to race. Instead, race might simply be correlated with other buyer attributes affecting the home sales process. In practice, the typical housing transaction is complicated by the need of the buyer to secure a mortgage and engage in a formal inspection of the property. This means that there is often some uncertainty for the seller at the time the sales price is set about the likelihood that the sale will go through. It would not be surprising, therefore, if sellers were willing to accept a lower transaction price from a buyer whom they deemed more likely to be able to secure financing and close the deal. On the flipside, wealthier buyers may have higher search costs and therefore make higher initial offers when they find a house that they like in order to make sure to get it. In this way, the estimated price premiums for black and Hispanic buyers may simply reflect differences in the buyer's income, wealth, or ability to secure mortgage financing.

The novel features of our data set allow us to address these issues directly. In particular, for each transaction, we observe the buyer's income as well as the down-payment associated with the mortgage and the lender's name. When we include controls for buyer's income and down-payment in the repeat-sales specification described above, the estimated premiums for black and Hispanic buyers increase significantly, to an average of 2.8 percent in each case (over 4 percent for black buyers in Chicago and Maryland). When we add a complete set of lender fixed effects, which might proxy for the buyer's access to credit, the results in each city remain largely unchanged, with the average premiums remaining at 2.8-2.9 percent. Thus it appears that, if anything, the correlation of race with financial considerations leads to an underestimate of the racial premium in our initial analysis.

Having established the existence of robust racial differentials in sales prices, we turn our attention to the second main goal of our analysis: testing whether racial animosity/prejudice on the part of sellers can explain these premiums. To this end, we present a series of results that characterize how these differentials vary across sellers and neighborhoods. We begin by decomposing the estimated price differences by the race of both the buyer and seller. If racial animosity or prejudice is an important factor driving seller behavior, we would expect sellers to systematically favor buyers of their own race or ethnicity. Instead, we find that the average price premium paid by black and Hispanic buyers is about the same regardless of the race of the seller. That is, black and Hispanic buyers pay significantly more than their white counterparts not only when they buy from white sellers but also from black and Hispanic sellers. Our analysis also

indicates that there is little difference in the estimated price differentials across different neighborhood compositions. Taken together, these results provide strong support for the conclusion that racial prejudice – or any explanation that would lead sellers to favor buyers of their own race – is not the primary explanation for the observed racial price differentials.⁷

A limitation of our analysis is that we cannot isolate a single alternative explanation for the estimated price premiums paid by black and Hispanic buyers in these markets. Statistical discrimination might be motivated, for example, by the correlation of race with search costs or experience with real estate bargaining; and further research is needed to determine the extent to which a particular channel predominates. Regardless of the ultimate explanation, though, our results show that black and Hispanic buyers pay more than their white counterparts in almost every purchase setting and that robust racial differences in the price paid to buy a home – upwards of 3-3.5 percent *on average* in multiple major US markets – persist to the present day, long after many of the most overt forms of institutional discrimination have been eliminated. Price differences of this magnitude have clear implications for the levels and persistence of racial differences in home ownership, neighborhood segregation, and the dynamics of wealth accumulation.

The remainder of the paper is organized as follows: Section II describes the data set that we have constructed for our analysis. Section III introduces our primary research design and relates it to those used in previous studies that attempt to infer racial price differentials in housing markets. We present our main estimates in Section IV, and test whether racial prejudice provides the primary explanation for these results in Section V. Section VI concludes, drawing out the implications of our findings.

II. Data

The data set we use is constructed by combining information from proprietary transactions data collected by a real estate monitoring service, *DataQuick*, with publicly-available loan application registry information gathered under the Home Mortgage Disclosure Act (HMDA).⁸ The transactions data set includes a complete census of housing transactions and is available for the San Francisco and Los Angeles metropolitan areas from 1988-2008, and for the Chicago metropolitan area and Baltimore-Washington DC corridor in Maryland from 1997-2008. These cities are particularly useful for studying racial price differences because each has a large

⁷ Note that these results for home sales, where buyer and seller engage in a single interaction, do not imply that racial animosity or prejudice is not an issue in the rental market, where the rental contract generates a dynamic relationship between landlord and tenant.

⁸ Additional information about the HMDA legislation and the public-use data can be found at <http://www.ffiec.gov>.

and heterogeneous population of homebuyers. Data on the transaction price, date of sale, loan amount, lender name, and location of the home are provided for each transaction. In addition – most crucial to our analysis – each property is characterized by a unique identifier that makes it possible to track the longitudinal transactions history of each home.

We match demographic information to individual homes by using the HMDA application registry files. The HMDA legislation was enacted to monitor potential redlining and discriminatory lending behavior. An important feature of this legislation is the requirement that race and other pertinent demographic information be recorded for each mortgage application. Hence, from these data we are able to retrieve demographic information relevant to our study: race and income of the buyer, transactions date, and the census tract of the home in question for linkage purposes.

Our primary data set for each city is then constructed by matching the transactions and HMDA data sets on the basis of (i) census tract, (ii) loan amount, (iii) transaction year, and (iv) lender name.⁹ An observation in the matched data set consists of the full set of information provided for each transaction record by *DataQuick*, along with demographic and economic information about the buyer drawn from HMDA. Finally, tract-level Census data are merged into the *DataQuick*-HMDA matched data set to augment the neighborhood level characteristics available for empirical analysis.

As we describe below, our main estimating equation includes house fixed effects. Thus, the effective sample for our analysis, which uses a repeat-sales approach, includes only houses that sell at least twice over the period. Our sample includes over a quarter million transactions for both Chicago and Maryland, over half a million transactions for San Francisco, and close to one million for Los Angeles, giving an overall sample size of over two million. Table 1 presents descriptive statistics for the attributes of buyers and sellers, as well as transaction prices for the observations in our repeat-sales sample. For comparison, Appendix Table 1 reports analogous statistics for homes that transacted only once during our sample period; the differences are, for the most part, small to negligible.¹⁰ The columns of Table 1 report statistics separately for each metropolitan area, and the top panel of the table reports the number of observations in the repeat-sales data set and the frequency with which homes in each area sold 2-4 or more times,

⁹ Overall, the matching procedure works well. Due to missing, misreported, or other idiosyncratic differences in the values of the matching variables, a unique match was not achieved for all observations. However, these unmatched homes account for a small percentage of transactions – in most cases, less than 5 percent of the data.

¹⁰ The two distributions of buyer race are essentially identical, comparing the repeat-sales sample to the full sample; mean and median incomes are close to one another; and median transactions prices are slightly higher in the repeat-sales sample for all but Chicago.

respectively (The three categories are exhaustive.). In all four cities, houses that sold exactly twice make up the majority of observations.

The second panel of the table reports the distribution of buyer race and ethnicity. While the majority of buyers in each city are white, there is considerable variation in buyer race and ethnicity across the cities in our sample. For example, while only 5 percent of buyers are black in Los Angeles and 8 percent in Chicago, blacks constitute 25 percent of buyers in Maryland. The fraction of Hispanic buyers is contrastingly low there, at 7 percent, and much higher, at 25 percent, in Los Angeles, while the fraction of Asian (including ‘other’ race/ethnicity) buyers is highest in San Francisco, at 25 percent.

For homes that sell multiple times, we are able to characterize the race of the seller, beginning with the second transaction observed in the data set.¹¹ The next panel of Table 1 reports the distribution of sales in each metropolitan area jointly by the race of the buyer and seller for all transactions where both buyer and seller race is recorded. In each metropolitan area, white-to-white transactions make up the majority of sales – between 43 to 60 percent of all transactions. There is, however, considerable variation in the composition of the remaining transactions in each metropolitan area. Maryland has the highest fraction of interracial transactions: white-to-black and black-to-white transactions make up 12 and 7 percent of repeat sales, respectively. Chicago also has a considerable fraction of white-to-black sales but far fewer black-to-white transactions. Transactions between white and Hispanic buyers, running in each direction, are reasonably prevalent in Chicago, higher in San Francisco and still higher in Los Angeles.

Finally, the bottom panel of Table 1 reports some basic summary statistics by metropolitan area: Home prices are significantly higher in the two Californian metropolitan areas, not surprisingly, and so too are the average incomes and median down-payments of homebuyers.

III. Research Design: Identification of Racial Price Differentials

The research design that we develop exploits the unique structure of our data set to shed light on racial differences in the price paid to purchase comparable housing. Our primary methodology can be summarized using the following regression equation, which relates the log price of transaction i for house j in neighborhood n at time t over the sample period to the buyer’s race and a set of controls:

¹¹ The seller’s race is unobserved for the first transaction for each house in our repeat sales sample.

$$(1) \ln(p_{ijt}) = \text{race_buyer}_{it}\gamma + X_{ijt}\beta + \mu_j + \theta_{nt} + \varepsilon_{ijt}.$$

Here, race_buyer_{it} is a vector of indicators for the race/ethnicity of the buyer at time t ; X denotes a vector of observable house buyer characteristics for transaction i at time t ; μ_j denotes a house-specific fixed effect, and θ_{nt} denotes a set of neighborhood-by-time fixed effects. As reported below, the coefficient γ measures the average premium, if any, paid by minorities (blacks and Hispanics) relative to whites.¹²

Our main goal is to measure any difference in the prices that buyers of different races/ethnicities pay for comparable housing. As we make clear in what follows, the inclusion of house fixed effects, μ_j , in equation (1) ensures that the race parameters are identified by comparing prices for houses for which buyer race changes over successive transactions to those for which buyer race remains constant. The inclusion of neighborhood-by-time fixed effects ensures that these comparisons are made within the same neighborhood during the same time period.

To see how the effect of buyer's race is identified in this model, it is helpful to re-write equation (1) by differencing observations for a consecutive pair of transactions (i, t) and (i', t') involving house j :

$$(2) \ln(p_{ijt}) - \ln(p_{i'jt'}) = (\text{race_buyer}_{it} - \text{race_buyer}_{i't'})\gamma + (X_{ijt} - X_{i'jt'})\beta + (\theta_{nt} - \theta_{nt'}) + \omega_{ijt}$$

Notice that the house fixed effect drops out of equation (2) and that γ now multiplies the difference in buyer race for transaction i and i' . This term is, of course, equal to zero if the buyer's race does not change between transactions, and thus those transactions contribute exclusively to the identification of the θ parameters, which characterized the pattern of price appreciation in the neighborhood n . When the buyer's race *does* change between transactions, the vector of differenced buyer race variables is non-zero and γ is identified by comparing the price appreciation of these transactions relative to the baseline rate of appreciation in the neighborhood.

By way of a concrete example, consider neighboring houses that both sold initially to white buyers in 1999 and that both sell again in 2006, this time to a black and white buyer respectively. In this case, the estimated racial difference in prices would be identified by the difference in price appreciation between the two sales. If, for example, the price paid by the white

¹² We also control for Asian and 'Other' races. Interpretation of the corresponding coefficient is less clear because the large set of ethnicities that fall under this umbrella. These results are available from the authors upon request.

buyer in 2006 implied an appreciation rate of 50 percent and the price paid by the black buyer implied an appreciation rate of 55 percent, we would infer that the black buyer paid a 5 percent premium. Averaging these differences across the full set of comparable houses in the data set provides the basis for identifying the buyer race parameters in equations (1) or (2).

A potential concern with the research design captured in equation (1), and one that applies to the entire existing literature, is that the homes purchased by buyers of different races may have undergone differential amounts of renovation or maintenance over the holding period of the previous owner. If buyers of one race, for example, tend to buy houses that are more likely to have been improved over the previous holding period, this would bias the buyer race parameters. In essence, the appreciation associated with these improvements would be mistakenly attributed to the race of the buyer. Because white buyers have significantly higher levels of income and wealth than black and Hispanic buyers, we would generally expect whites to be systematically more likely to buy improved or especially well-maintained houses. If this were indeed the case, the estimated γ would *understate* the actual premiums paid by minorities (blacks and Hispanics) relative to whites. We address this potential concern by using the detailed assessor records associated with each housing transaction in the data set.¹³

Prior Evidence on Price Differentials in the Housing Market

To estimate price differentials in the housing market, one would ideally like to compare the price that buyers of different races would pay for identical properties. The previous literature has pursued this strategy by comparing the prices of homes purchased in the same neighborhood at roughly the same time. The pioneering studies that first developed and implemented this research design (King and Mieskowski, 1973; Kain and Quigley, 1975; Yinger, 1978) estimated that minorities paid a premium for comparable housing. In contrast, subsequent analyses based on larger samples by Follain and Malpezzi (1981), Chambers (1992) and Kiel and Zabel (1996) found statistically significant *discounts* for black buyers relative to whites. These studies were still constrained by the size and nature of the data sets available at the time, limiting the researchers' ability to control for unobserved differences in house quality within neighborhoods or neighborhood attributes. Given higher levels of income and wealth, it would not be surprising if white buyers generally purchased higher quality housing within the same neighborhood,

¹³ In particular, the records for most homes include an assessor flag that indicates whether there has been a major renovation to the home and when such a renovation occurred. Dropping these houses from the sample has a negligible impact on the parameter estimates, indicating that there is little confounding within-neighborhood correlation between race and housing improvements/maintenance.

potentially leading to a spurious finding that white buyers pay a ‘premium’ for housing. By including house fixed effects in equation (1), we can address this limitation directly.

In this sense, our approach is related to specifications reported in Myers (2004) that include house fixed effects estimated with data from three waves of the American Housing Survey, supplemented with special neighbors samples. Myers finds evidence that blacks pay premiums for their homes of around 10 percent relative to whites, although the estimates are significant only at the 10 percent level.¹⁴ Another recent study, by Ihlanfeldt and Mayock (2009), examines price differentials by race in a specification that controls for detailed structural characteristics and block-group fixed effects for data on 20 metropolitan areas in Florida. As in much of the prior literature, Ihlanfeldt and Mayock find evidence of price discounts for nonwhite buyers, and argue that correlation between race and unobserved property characteristics may bias the estimated race coefficients downwards.¹⁵ It is precisely this potential bias that the inclusion of house fixed effects is designed to eliminate.

Relative to Myers (2004) and Ihlanfeldt and Mayock (2009), our analysis offers two main advantages. The first is that we have a complete census of housing transactions in major metropolitan areas over a long period of time, including many houses that sell to buyers of different races. Such ‘switches’ in the race of the buyer are needed to identify racial differences in pricing when house fixed effects are included in the analysis. We effectively have multiple orders of magnitude more data, and this allows us to estimate racial differences far more precisely and characterize how they vary across sellers, neighborhoods, and market conditions. The second main advantage of our research design is the inclusion of neighborhood-by-time fixed effects that effectively control for time-varying factors in each neighborhood that might influence prices. Again, it is the sheer size and scope of our data set that allows us to control for neighborhood dynamics in this especially flexible way.

IV. Main Results

We now present the main results of our analysis. We begin by estimating limited versions of equation (1) that leave out the house and/or neighborhood-by-time fixed effects. Comparing our final model (1) with these incomplete specifications provides a clear demonstration of the importance of the additional features in estimating racial differences in purchase prices for

¹⁴ Myers’ sample consists of just under 22,000 observations. She notes that a larger sample would increase the precision of fixed effects estimates reported in the paper, as more changes in the race of home owners would then be likely.

¹⁵ Building on a model of bargaining and an estimation approach put forward in Harding, Rosenthal, and Sirmans (2003), they ultimately argue that their results provide evidence of discrimination.

comparable housing. We then add controls that proxy for other aspects of the buyer's financial position – income, down-payment, and lender fixed effects – to examine whether the racial differences estimated in our baseline specification might be driven by these financial considerations, which are correlated with race, rather than the buyer's race/ethnicity itself.

Baseline Results

Table 2 reports baseline results for minority-white price differentials from the estimation of equation (1). The table reports results for each metropolitan area organized into two panels: Panel A presents the black coefficient and Panel B presents the Hispanic coefficient for each specification. In each case, the coefficient measures the difference in transaction price relative to a white buyer. Each panel reports results for four specifications. The first controls for neither house nor neighbor-by-time fixed effects. Not surprisingly, the parameter estimates in this case are uniformly negative and large in magnitude as black and Hispanic buyers purchase substantially less expensive properties than white buyers in each metropolitan area.

The second column of each panel reports parameter estimates for a specification that includes neighborhood-by-time fixed effects but does not include house fixed effects. This specification compares more closely to the specifications typically reported in the existing literature, which generally include controls for neighborhood-level attributes but not house fixed effects. And as in the existing literature, the parameter estimates reported in column (2) remain negative in almost every sample. The estimated black-white and Hispanic-white price differentials for the full sample in this specification are -0.044 and -0.061, respectively, giving no indication that black or Hispanic buyers may pay more for *comparable* housing than their white counterparts.

The specification reported in the third column of Table 2 includes house fixed effects, while the one reported in the fourth column controls simultaneously for both house and neighborhood-by-time fixed effects. Focusing on the parameter estimates reported in column (4) of each panel, it is immediately obvious that the inclusion of house fixed effects flips the sign on the estimated minority-white price differentials from negative to positive in every sample for both black and Hispanic buyers. The estimated price differentials for black and Hispanic buyers are 0.020 and 0.021, respectively, implying that black and Hispanic buyers typically pay about 2 percent more for comparable housing than whites.

There is considerable variation in the estimated price differentials across cities. The estimated black-white price differential is highest in Cook County, which includes the central city of Chicago (4.7 percent), and the Baltimore area (2.9 percent), while estimates range between

0.8-1.9 percent in the other six areas. The estimated Hispanic-white price differential, on the other hand, is in the 2.0-3.0 percent range for much of Chicago, Maryland, and San Francisco, contrasting with the estimates for the Los Angeles area, which are little more than 1 percent.

Controlling for Other Buyer Attributes

While the results presented in Table 2 provide strong statistical evidence that black and Hispanic buyers pay more for comparable housing than whites, it is not necessarily the case that these estimated differentials are motivated by racial considerations. Instead, race may simply be correlated with other buyer attributes that affect the home sales process. To assess this possibility, Table 3 presents results from additional specifications that control for various aspects of the buyer's financial position, including the income, down-payment percentage, and a set of lender fixed effects. As previously mentioned, the buyer's financial position might be correlated with sales price for a number of reasons including the ability to secure mortgage financing and search costs. Because some lenders specialize in high-priced or subprime loans, the inclusion of lender fixed effects proxies to some extent for the buyer's credit worthiness.

The first column of Table 3 repeats the final specification from Table 2, which controls for house and neighborhood-by-time fixed effects. The specification reported in column (2) adds controls for measures of the buyer's income and down-payment percentage. The coefficients for both the black-white and Hispanic-white differentials increase in a relatively uniform way for each city sub-sample with the inclusion of these controls. As shown in the bottom row, both the black-white and Hispanic-white gaps increase from approximately 2.0 to 2.8 percent for the full sample. Column (3) reports results for a specification that adds second-order polynomial terms in buyer attributes to the specification reported in column (2). The addition of these controls again results in larger estimates of the black- and Hispanic-white price differentials in almost all city sub-samples, increasing the estimate of the black differential to 3.3 percent and the Hispanic differential to 2.9 percent. Column (4) reports results for a specification that adds lender fixed effects to the specification reported in column (2). The inclusion of these lender controls has a negligible effect on the estimates in each sub-sample, and the 'all-city' estimated price premiums for blacks and Hispanics are almost identical and very precisely-estimated: 2.8 percent for blacks and 2.7 percent for Hispanics. Taken as a whole, the results presented in Table 3 provide clear

and consistent evidence that the correlation of race and financial considerations leads to an under- rather than over-estimate of the magnitude of racial price differentials.¹⁶

V. Heterogeneity Across Sellers and Neighborhoods

The results presented in Tables 2 and 3 establish premiums in sales prices for comparable housing on the order of 3 percent for both black and Hispanic buyers relative to whites. We now present a series of results that characterize how these premiums vary across sellers and neighborhoods. The main goal of this portion of our analysis is to examine whether racial animosity or prejudice can explain the observed price premiums.

Seller's Race

We begin by decomposing the estimated premiums by the race of both the buyer and seller.¹⁷ To this end, we expand equation (1) to include a full set of interactions between buyer and seller race:

$$(3) \ln(p_{ijt}) = (\text{race_seller} \times \text{race_buyer})_{it} \lambda + X_{ijt} \beta + \mu_j + \theta_{nt} + \varepsilon_{ijt},$$

reporting all of the estimated premiums, λ , relative to white-to-white transactions. Again, for expositional purposes, it is helpful to consider a differenced version of equation (3):

$$(4) \ln(p_{ijt}) - \ln(p_{i'jt'}) = (\text{race_s} \times \text{race_b}_{it} - \text{race_s} \times \text{race_b}_{i't'}) \lambda + (X_{ijt} - X_{i'jt'}) \beta + (\theta_{nt} - \theta_{nt'}) + \omega_{ijt}$$

Because, as noted, *seller's race* is unobserved for the first transaction for each house in our repeat sales sample, we consider three alternative approaches to estimating equation (4). All three approaches yield very similar point estimates for the key parameters of interest, λ and are reported below or in appendices. Our first approach is to restrict the sample to observations from homes that sell at least three times, dropping the first transaction for each of these houses (for which seller's race is unknown) from the analysis. As one might expect, this approach greatly

¹⁶ A potential explanation for this finding is that individuals with higher levels of income and wealth have higher search costs due to a higher value of time. As a result of such a mechanism, conditioning on income and wealth, which is highly correlated with race, would lead to larger estimates of racial price premiums.

¹⁷ This test bears some relation to the test for racial profiling proposed by Anwar and Fang (2006) and Close and Mason (2007). In these settings, researchers compare the search rates of drivers of different races by officers of different races to study whether racial differences in search are motivated by prejudice versus differential search costs.

diminishes the number of observations in the sample (i.e., it entirely eliminates the use of houses that sell twice and cuts in half the number of observations in equation (4) for houses that sell three times). As a result, the neighborhood-by-time effects, θ , are not very precisely estimated, leading to less precise estimates of our main coefficients of interest, λ , although these are still statistically significant in many cases.

In order to estimate the neighborhood-by-time effects more precisely, our two alternative approaches make use of the full repeat sales sample to identify the pattern of price appreciation at the neighborhood level, while making sure that the use of observations which are missing seller's race do not contaminate the estimates of λ . Our second approach is thus to create a new category for seller's race that equals one if it is unobserved. We then simply include interactions of this unobserved seller race category with the appropriate buyer race for the first transaction that we observe for each house. We treat the parameters on interactions that involve sellers of unknown race and buyers of each race as nuisance parameters. The advantage of this approach over the first is that it allows all of the consecutive pairs of transactions to help identify the neighborhood-by-time fixed effects. For our third and final approach, we estimate equation (4) for all consecutive pairs of transactions for each house j in the repeat-sales sample. But for the first pair of transactions for house j , we simply set the buyer-seller race difference (i.e., the first term on the right hand side of equation (4)) to zero. In this way, houses that sell only twice, for example, are helpful in getting a much more precise estimate of price appreciation patterns by neighborhood, but do not contribute directly to the identification of λ .

The results from this final and preferred approach – Approach 3 – are shown in Table 4. (The results from the first two approaches are reported in Appendix Tables 2 and 3 respectively and, as mentioned above, yield qualitatively very similar point estimates across all four metropolitan areas.) The panels of Table 4 report results for each buyer-seller race combination relative to a white-to-white transaction separately for each of the four metropolitan samples. For each, the buyer's race is indicated in the row heading and the seller's race is indicated in the column heading. Thus, comparing results across rows shows how the sales price varies with buyer's race for sellers of the race/ethnicity shown in the column heading.

Comparing the row average price differentials for buyers of different races is, in essence, the focus of Tables 2 and 3. In contrast, our main focus in examining the results presented in Table 4 is instead to compare across columns, which shows how the sales price for each type of buyer varies with the seller's race. If racial animosity or prejudice is an important factor driving seller behavior, we would expect sellers to systematically favor buyers of their own race or ethnicity. Instead, in the Chicago, Los Angeles, and San Francisco sub-samples, the estimated

sales price premiums are almost identical, irrespective of seller's race. For black buyers in Chicago, for example, the estimated premium falls within a narrow range between 5.2 and 5.3 percent for white, black, and Hispanic sellers, respectively. If anything, for black and Hispanic buyers in these cities, the estimated premiums are slightly larger for black and Hispanic sellers relative to their white counterparts. For white buyers in these cities, the estimated premiums for black and Hispanic sellers are all within 1.0 percentage points of the baseline estimate for white-to-white transactions. The results for Maryland differ from those for the other city samples in that black and Hispanic sellers in Maryland earn a premium of 5-8 percent relative to their white counterparts but again, the size of this premium does not vary with the race of the buyer. Taken as a whole, the results presented in Table 4 provide strong evidence that the price differentials associated with buyer's race reported in Tables 2 and 3 do not vary systematically with the seller's race.

Another interesting way to view the results presented in Table 4 is to consider what they tell us about the relative market returns among sellers of each race. In Chicago, Los Angeles, and San Francisco, sellers of each race receive prices that are within about a single percent of one another when they sell to white buyers, and a similar pattern holds for sales to Hispanic and black buyers. In the Maryland area, black and Hispanic sellers command a premium of about 5 percent when they sell to buyers of each race. Averaging across the four metropolitan areas as a whole, these results imply that there is very little difference in the prices that households of each race receive when they sell their homes – if anything, black and Hispanic sellers appear to do slightly better than white sellers.¹⁸

Neighborhood Initial Demographic Composition

Table 5 reports results that explore the heterogeneity in estimate racial price differentials across another important dimension of heterogeneity – the existing composition of the neighborhood (Census tract). In particular, the table reports estimates for black and Hispanic buyers in each city separately for (i) all Census tracts, (ii) tracts that were more than 50 percent non-Hispanic white in 2000, and (iii) tracts with 80 percent non-Hispanic white or greater in 2000. For black buyers, the parameter estimates are fairly stable across these three specifications

¹⁸ We return to this result in the concluding section below, as it points to a role for experience in the real-estate bargaining process in explaining the premiums that black and Hispanic buyers pay on average.

in the four city subsamples, implying that the price premium paid by black buyers does not vary much with the neighborhood setting.¹⁹

A different pattern emerges for Hispanic buyers. In particular, the estimated price premium for Hispanic buyers is uniformly higher in neighborhoods with a lower fraction of white residents, falling from 1.6-3.4 percent in all tracts to 0.3-1.1 percent in neighborhoods that were more than 80 percent white in 2000 in the four city subsamples. In fact, the point estimate for the Hispanic price premium is statistically insignificant in all of the cities except Chicago in neighborhoods that were initially at least 80 percent white. Thus, the price premium essentially goes away for Hispanic buyers when they purchase houses in predominantly white neighborhoods, while a premium remains for black buyers in these neighborhoods.²⁰ One possible explanation for the fact that the Hispanic premium diminishes in predominantly white neighborhoods is that the Hispanic buyers in these neighborhoods may be systematically different from those that purchase homes in more racially diverse neighborhoods. Buyers in the latter set of neighborhoods, for example, may include a higher percentage of recent immigrants, an attribute that we cannot observe directly in the data.

VI. Do We Detect Discrimination?

The magnitude and remarkably robust nature of the premiums paid by blacks and Hispanics provide strong evidence that these households pay more than their white counterparts when purchasing comparable homes. The lack of any same-race preference on the part of sellers, however, makes it less clear whether such differentials actually represent the consequences of discriminatory behavior or are an artifact of some other process related to home-buying.

Largely uncontroversial, however, is the notion that discriminatory behavior is ongoing in housing markets. Audit studies provide a particularly powerful and complementary strategy for detecting signs of discriminatory behavior in housing markets. Generally speaking, these studies aim to test for discrimination by sending individuals of different races or ethnicities, matched as well as possible on other characteristics, to inquire about housing units either for sale or rent.

¹⁹ It is worth noting that the estimates for the Chicago area imply that blacks pay a greater premium for housing in neighborhoods in which the initial percentage of white buyers is less than 50 percent. This result may be capturing in part the activity of white buyers moving into gentrifying neighborhoods.

²⁰ We also tested whether racial price differentials are larger in 'hot' market conditions, in which a seller is more likely to receive multiple offers simultaneously and might, more generally, be willing to decline offers knowing that subsequent offers will arrive relatively quickly. In this way, hot market conditions might make it easier for a seller to avoid transacting with a buyer of a certain race, should that be his or her preference. Estimated racial premiums do not vary much with market conditions, although some estimates are noisy. These results are available upon request.

Using this approach, Yinger (1996) and Ondrich *et al.* (2003) find direct evidence of the discriminatory treatment of minority buyers and renters along a number of dimensions. In particular, they find evidence of statistical discrimination likely stemming from uncertainty about black potential buyers ability to put forward successful bids. For example, they find evidence that agents do not increase their effort in response to higher sales prices and substantial steering of blacks to homes with fewer features than they request. These findings suggest that agents may be skeptical of black ability to purchase more expensive homes. If such beliefs are pervasive, one can imagine a situation where conditional on choosing a particular house, blacks may need to submit higher bids than observationally equivalent whites to be taken seriously.

While audit studies provide a powerful way to identify the pervasiveness of exclusion and steering in the market, they yield little evidence about transactions that are actually consummated.²¹ This is important both because audit studies may miss aspects of discrimination that occur through price negotiation/determination and because it is difficult to gauge the ultimate impact of the observed exclusion or steering on housing outcomes (Goldberg, 1996; Yinger, 1998). Conversely, a limitation of studies that focus on the pricing of observed transactions (like our own) is that they miss exactly the kinds of effects that steering and exclusion have on the choice of homes captured by audit studies. In this sense, the two approaches for detecting signs of discriminatory behavior in consumer markets neatly complement each other.

Another possible explanation for the robust premiums we find is that nonwhite buyers pay higher prices because of differences in bargaining power or search costs. If a substantial portion of sellers, for example, were unwilling to sell to buyers of a particular race, the effective cost of continuing to search for a suitable home would be higher for these buyers. Faced with a lower value of continuing to search, buyers of this race might settle for a higher transaction price in order to secure a property of their liking. The results examining how the premiums vary across seller race, however, do not point in an obvious way to this potential explanation, as nonwhite buyers do not appear to get a ‘better’ deal through transacting with someone of their own race.²²

²¹ Another limitation of audit studies is that they require all participants to behave identically, despite the fact that participants are likely aware of the goals of the study. To address this concern, a new study by Hanson and Hawley (2011) uses email communication to respond to online apartment listings, signaling race through the use of names that are predominantly used by either white or black parents. Another interesting recent paper by Ewens, Tomin and Wang (2012) uses a similar approach to distinguish between prejudice and statistical discrimination as explanations for callback behavior. Their results point to statistical discrimination rather than prejudice as the most likely explanation for landlord behavior; this is broadly consistent with our findings.

VII. Summary and Conclusion

The main goals of this paper have been to measure the extent of price differentials in the housing market on the basis of buyer's race and to gauge the explanatory power of a set of alternative explanations for any estimated price differences. In this concluding section of the paper, we summarize our main findings briefly and discuss their implications for the evolution of racial differences in wealth and home ownership and the persistence of residential segregation.

Combining a novel data set with a fixed effects approach that allowed us to control for unobserved differences in housing and neighborhood quality to a much greater extent than the prior literature, our analysis documented meaningful, statistically significant premiums in the prices paid by black and Hispanic buyers in each of the four large metropolitan areas that we studied – a result which sharply contrasts with the best-known results in the recent literature. The unique features of our data set also allowed us to examine whether racial differences in income, wealth, or access to credit might be responsible for these premiums. Here, the results indicated that the correlation of race with these factors works in the opposite direction: the estimated racial premiums conditional on flexible functions of income and down-payment percentage, as well as lender fixed effects are actually positive and larger in size.

Having established the presence of robust differences in the prices paid by buyers of each race, we then took up the question of whether systematic cross-racial bias by sellers could explain these differences. In particular, we examined whether sellers are willing to accept lower prices from buyers of the same race, finding that the premiums paid by black and Hispanic buyers in fact vary very little with the race of the seller. While these results provide evidence against racial bias on the part of sellers as the primary explanation for the observed racial price premiums, it is important to note that we cannot rule out that animosity or prejudice may lead to the exclusion of minority buyers from purchasing certain properties in the first place, as suggested by previous audit studies.

Our analysis also revealed that sellers of each race do roughly as well on average when they sell their homes, no matter what the race of the buyer is. Thus, any disadvantage that black and Hispanic buyers face when purchasing homes disappears when it comes time to sell. While certainly not conclusive, this pattern suggests that the relative inexperience of black and Hispanic buyers, due to the historically lower rates of home ownership, may contribute to the higher prices that they initially pay upon entering the market. With the experience gained from having participated in the real estate bargaining process at least once, black and Hispanic sellers do as well as their white and Asian counterparts. What this means for subsequent purchasing experiences of more seasoned buyers is unclear, however, as sellers of each race may continue to

statistically discriminate against experienced black and Hispanic buyers because race/ethnicity is highly correlated with experience, which may be difficult to observe.

Of course, relative inexperience with the bargaining process is not the only possible basis for statistical discrimination on the part of sellers. Another potential explanation, put forward by Yinger (1995, 1998), is that black and Hispanic buyers face higher search costs in the housing market. Such higher search costs might stem from a variety of fundamental sources or because of the expectation of discriminatory treatment in the market. Knowing that black and Hispanic buyers typically face higher costs of continuing to search, sellers might statistically discriminate by holding out for a higher price (i.e., use a higher reservation price) in negotiations.

No matter what the ultimate reason for the price premium, our results imply that systematic, robust racial differences in the price paid to buy a home – on the order of 3 percent on average in multiple major US markets – persist to the present day, long after many of the most overt forms of institutional discrimination have been eliminated. Considering the average purchase price paid by a black homebuyer in our sample is \$177,000, this translates to an average premium of about \$5,000 per transaction, a substantial amount given the average income of black households in these cities.

These price differentials are likely to have important implications for a range of other social and economic outcomes, including the evolution of racial differences in wealth, home ownership rates, and location decisions. Faced with what amounts to a substantial transaction tax with each home purchase, one would naturally expect home ownership rates to be lower and the benefits of ownership for wealth accumulation to be systematically diminished for minority households. Moreover, to the extent that these differentials represent price discrimination, the added cost may alter black and Hispanic household location decisions (i.e., leading to the choice of neighborhoods with more rental properties or lower prices). Hence, existing residential segregation may be reinforced, which has further important consequences for educational and labor market outcomes.²³

²³ See Yinger (1995) and Cutler and Glaeser (1997) for a discussion of the consequences of segregation for education and labor market outcomes. It is particularly noteworthy that the estimated premiums are largest for black households in Chicago and Maryland – metropolitan areas that remain largely segregated along racial lines to this day.

References

Anwar, S. and H. Fang (2006), "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence," *American Economic Review*, **96**(1): 127-151.

Arrow, K.J. (1998), "What has Economics to Say About Racial Discrimination?" *Journal of Economic Perspectives*, **12**(2): 91-100.

Ayres, I. (1991), "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations," *Harvard Law Review*, **104**: 817-872.

Ayres, I. (2003), Pervasive Prejudice?: Unconventional Evidence of Race and Gender Discrimination. Chicago: University of Chicago Press

Ayres, I and P. Siegelman (1995), "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, **85**: 304-321.

Chambers, D.N. (1992), "The Racial House Price Differential and Racially Transitional Neighborhoods," *Journal of Urban Economics*, **32**: 214-232

Charles, K. and J. Guryan (2011), "Studying Discrimination: Fundamental Challenges and Recent Progress," NBER Working Paper #17156.

Charles, K. and E. Hurst (2002), "The Transition to Home Ownership and the Black/White Wealth Gap," *Review of Economics and Statistics*, **84**(2): 281-297

Close, B. and P. Mason (2007), "Searching for Efficient Enforcement: Officer Characteristics and Racially-Biased Policing," *Review of Law and Economics*, **3**(2).

Cornelissen, T. (2008), "The Stata command felsdreg to fit a linear model with two high dimensional fixed effects," *The Stata Journal*, **8**(2): 170-189.

Cutler, D.M., and E. Glaeser (1997), "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, **112**(3): 827-872.

Cutler, D.M., E. Glaeser and J. Vigdor (1999), "The Rise and Decline of the American Ghetto," *Journal of Political Economy*, **107**(3): 455-506.

Ewens, M., B. Tomlin, and C. Wang (2012), "Statistical Discrimination or Prejudice? A Large Sample Field Experiment," mimeo, Carnegie Mellon University.

Follain, J.R. and S. Malpezzi. (1981), "Another Look at Racial Differences in Housing Prices," *Urban Studies*, **18**: 195-203.

Goldberg, P.K. (1996), "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, **104**(3): 622-654.

Hansen, A. and Z. Hawley (2011), "Do Landlords Discriminate in the Rental Housing Market? Evidence from an Internet Field Experiment in U.S. Cities," *Journal of Urban Economics*, forthcoming.

Harding, J.P., S. Rosenthal, and D.F. Sirmans (2003), "Establishing Bargaining Power in the Market for Existing Homes," *Review of Economics and Statistics*, **85(1)**: 178 – 188.

Hirsch, A. (1998), *Making the Second Ghetto: Race and Housing in Chicago 1940-1970*, University of Chicago Press, Chicago, IL.

Ihlanfeldt, K. and T. Mayock (2009), "Price Discrimination in the Housing Market," *Journal of Urban Economics*, **66**: 125 – 140.

Kain, J.F. and J.M. Quigley (1975), *Housing Markets and Racial Discrimination: A Microeconomic Analysis*, NBER, Cambridge, MA.

Kiel, K.A. and J. Zabel (1996), "House Price Differentials in US Cities: Household and Neighborhood Racial Effects," *Journal of Housing Economics*, **5**: 171-189.

King, A. and P. Mieskowski (1973), "Racial Discrimination, Segregation, and the Price of Housing," *Journal of Political Economy*, **81**: 590-606.

Munnell, A.H, G.M.B. Tootell, L.E. Browne, and J. McEneaney (1996) "Mortgage Lending in Boston: Interpreting HMDA Data," *American Economic Review*, **86(1)**: 25-53.

Myers, C.K. (2004), "Discrimination and Neighborhood Effects: Understanding Racial Differentials in US Housing Prices," *Journal of Urban Economics*, **56**: 279-306.

Ondrich, J., S. Ross and J. Yinger (2003), "Now You See It, Now You Don't: Why Do Agents Withhold Available Houses from Black Customers?" *Review of Economics and Statistics*, **85(4)**: 854-873.

Yinger, J. (1978), "The Black-White Price Differential in Housing: Some Further Evidence," *Land Economics*, **54**: 187-206.

Yinger, J. (1986), "Measuring Discrimination with Fair Housing Audits: Caught in the Act," *American Economic Review*, **76**: 881-893.

Yinger, J. (1995), *Closed Doors, Opportunities Lost: The Continuing Cost of Housing Discrimination*, Russell Sage Foundation, New York.

Yinger, J. (1998), "Evidence on Discrimination in Consumer Markets," *Journal of Economic Perspectives*, **12(2)**: 23-40.