Economics Bulletin

Volume 39, Issue 2

I am Jane. Do I pay more in the housing market?

Myongjin Kim
University of Oklahoma

Brent Norwood
University of Oklahoma

Sean O'Connor
University of Oklahoma

Leilei Shen
Kansas State University

Abstract

Do women pay more than men in the housing market? We utilize repeat-sales housing data from ZTRAX to examine if gender gaps exist in house purchase prices and loan-to-price ratios. We find that female homebuyers pay a 2% premium on average. In addition, female homebuyers' loan to price ratio is 3 percentage points lower than that of male buyers. We also show that female buyers pay less when the seller is female than when the seller is male. However, the gender price differentials and loan-to-price differentials are disappearing in more recent years. One possible explanation for the disappearing gender price differentials in house prices and loan-to-price ratios is the shrinking of the gender wage gap in recent years.

We would like to thank the seminar participants at The University of Oklahoma and Kansas State University.
Citation: Myongjin Kim and Brent Norwood and Sean O'Connor and Leilei Shen, (2019) "I am Jane. Do I pay more in the housing market?", *Economics Bulletin*, Volume 39, Issue 2, pages 1612-1620
Contact: Myongjin Kim - mjkim@ou.edu, Brent Norwood - bnorwood09@ou.edu, Sean O'Connor - sean.oc@ou.edu, Leilei Shen - lshen@ksu.edu.
Submitted: April 09, 2019. **Published:** June 23, 2019.

Economics Bulletin

Submission Number: EB-19-00346

I am Jane. Do I pay more in the housing market?

Myongjin Kim
University of Oklahoma

Brent Norwood
University of Oklahoma

Sean O'connor
University of Oklahoma

Leilei Shen
Kansas State University

Abstract

Do women pay more than men in the housing market? We utilize repeat-sales housing data from ZTRAX to examine if gender gaps exist in house purchase prices and loan-to-price ratios. We find that female homebuyers pay a 2% premium on average. In addition, female homebuyers' loan to price ratio is 3 percentage points lower than that of male buyers. We also show that female buyers pay less when the seller is female than when the seller is male. However, the gender price differentials and loan-to-price differentials are disappearing in more recent years. One possible explanation for the disappearing gender price differentials in house prices and loan-to-price ratios is the shrinking of the gender wage gap in recent years.

Submitted: April 09, 2019.

1 Introduction

"Owning a home is part of that (American) dream," said George W. Bush. Is there a price differential between men and women for the similar dream? After all, a recent study from New York's Department of Consumer Affairs demonstrated that women, on average, pay 7 percent more than men for similar products across 35 products categories that ranged from toys and accessories to personal care products (Bessendorf 2015). Since an individual's house is considered as the average person's greatest source of wealth and investment and its unobserved quality can be controlled by using the repeat-sales approach, in this paper, we utilize transactional data from Zillow, a leading online platform for real estate properties, to examine if gender gaps exist in house purchase prices and loan to price ratios.

Using transactional data based on repeat-sales to control for the time-invariant unobserved heterogeneities in house qualities, we find that women pay more in house prices in the past but the difference between the prices paid by women and men is disappearing. In addition, we find that female homebuyers' loan to price ratio is lower than that of male buyers, and this gender gap in loan to price ratio is also disappearing in recent years.

There are two alternative explanations of our findings of gender price differentials in the housing market. One is that women are statistically-discriminated[1] against by sellers: women are less desirable buyers because sellers may find women to be less attractive including, but not limited to, the following reasons. It's harder for women to get loans (Ongena and Popov 2016) and when women do get loans they have to pay higher interest rates (Alesina, Lotti and Mistrulli 2013, and Cheng, Lin and Liu 2011). This explanation is consistent with our findings that in the past women's loan to price ratio is lower than men's. In other words, women may have to pay more in down payment than men to get loans for houses with the same price.

Another explanation is that women are taste-discriminated against by sellers: women are less desirable buyers because they are women, period. We investigate this possibility by identifying the seller's gender and examine the price differentials between each seller-gender and buyer-gender tuple. We find that women are not paying more than men do when sellers are women, but women pay 2.7% more than men do when sellers are men. This suggests that the estimated premium is partly driven by taste-based discrimination because if the estimated premium was entirely due to statistical-based discrimination, women would be paying more when sellers were women as well. However, according to Becker (1961), if women are discriminated against and pay higher prices, why don't other sellers make extra profits by not discriminating? This then should put a cap on the gender-based price differential. This could explain why we find that on average women pay 2% more than men.

We also find that gender gaps in house prices and loan-to-price ratios are disappearing in recent years. One possible explanation for this phenomenon is the closing of the gender wage gap in recent years. This has been documented in the recent literature. Chen (2018) finds that both women and men's wages are declining in recent years but men's wages decline faster, thus closing up the gender gap by 2 percentage points. Autor, Dorn and Hanson (2018) also find that among manufacturing workers, men's wages decline more than women's due to import competition.

[1]The theory of statistical discrimination was pioneered by Arrow (1973).

The literature on discrimination in the real estate market focuses primarily on race. King and Mieszkowski (1973) find that for the same rental units, African American households have to higher markups relative to white households. Ihlanfeldt and Mayock (2009) and Bayer, Casey, Ferreira, McMillan (2017) find that black and Hispanic buyers pay more in house prices. Others studies in the past find that whites pay more (Follain and Malpezzi 1981, Chambers 1992, and Kiel and Zabel 1996). This is due to lack of access to repeat-sales data for which properly measuring housing quality based on observables is hard.

The rest of the paper is organized as follows. Section II describes the data. Section III describes our identification strategy and estimating equations. Section IV presents our estimates, and Section V offers concluding remarks.

2 Data

We utilize the transactional data ZTRAX (Zillow Transaction and Assessment Database) to examine if gender gaps exist in house purchase prices. ZTRAX contains digitized county records of all information relevant to the sale of a house, including the price, names of the buyers, home characteristics, and loan details from 1985 to 2017.

The process of cleaning the data involves removing any transactions that are not deed transfers. Many observations that are not deed transfers occur below $10,000, suggesting personal or inter-family transfers as in the case of inheritances. We have therefore restricted observations to those above $10,000.

In order to identify the gender of buyers, we follow the strategy used in Bertrand and Mullainathan (2004). Bertrand and Mullainathan (2004) use hospital data for babies born in Massachusetts between 1974 and 1979 to determine the most common names by gender. We use the list of female names identified in Bertrand and Mullainathan (2004) and thus limit the transactional data to the state of Massachusetts. We also complement this list using the data from Social Security.[2] The Social Security database identifies the frequency of a given name and whether the sex at birth is male or female. Using data from 1912 to 2016, we calculate the percentage of times a name is female or male. For example, for the given name "Abigail", 99.8% of the occurrences is female, i.e. 99.8% of the U.S. births with the given name "Abigail" is recorded with the sex female. We identify the gender of the individual to be female if at least 90% of the occurrences of this name is female, and identify the given name as male if less than 10% of the occurrences of this name is female in the Social Security data. Our empirical results that follow are robust to different cutoffs and are available upon request.[3]

The ZTRAX database contains information on multiple buyers if the deed is registered with multiple names. In the case of multiple buyers, if one of the buyers is female, then the gender is recorded as female for this transaction. Note that in the analysis that follows,

[2]This is the National Data on the relative frequency of given names in the population of U.S. births where the individual has a Social Security Number. Names with fewer than 5 occurrences in any geographic area are dropped.

[3]Results are similar using 95% and 5% cutoffs, and 80% and 20% cutoffs. However, our measure of gender may still be subject to measurement errors. Wooldridge (2012) suggests that measurement error in the independent variable leads to attenuation bias. This implies the coefficient for Gender would be biased towards zero in the presence of measurement errors. This suggests that the coefficient reported for Gender can be interpreted as a lower bound to the true gender-price differential.

the sample of single-name buyers and the sample of multiple-names buyers are mutually exclusive of each other.

Each buyer is assigned a buyer ID in the ZTRAX database. Therefore, a buyer may show up multiple times if she purchases multiple homes in our sample period. We let experience equals to one at time t if the buyer has purchased a house in periods before t, and zero otherwise.[4]

Table 1 presents the descriptive statistics for the main variables used in our analysis. The average house price among single buyers is \$185,350 and the median house price is \$176,310. Among multiple buyers, the mean is \$287,794 and the median is \$299,539. Among single buyers, women account for 31% of the transactions and 40% of the buyers have bought houses before. Among multiple buyers, at least one of the buyers is a woman in 75% of the transactions and at least one of the buyers is an experienced buyer in 39% of the transactions.

Table 1 Summary Statistics

	Mean	Median	Std Dev	Min	Max	Obs
	Panel A. Single buyer					
house price (log)	12.13	12.08	0.58	10.46	13.76	47,051
loan ratio	0.81	0.8	0.16	0	1	47,051
gender	0.31	0	0.46	0	1	47,051
experienced	0.40	0	0.49	0	1	47,051
	Panel B. Multiple buyers					
house price (log)	12.57	12.61	0.54	10.46	13.76	54,375
loan ratio	0.79	0.8	0.16	0	1	54,375
gender	0.75	1	0.43	0	1	54,375
experienced	0.39	0	0.49	0	1	54,375

3 Empirical Strategy

In this section, we describe our empirical strategy to examine if gender gaps exist in the house purchase price and the loan to price ratio as follows:

$$y_{ihnt} = \beta_1 Gender_i + \beta_2 Exp_{it} + \alpha_h + \alpha_{nt} + \varepsilon_{ihnt}, \tag{1}$$

where y_{ihnt} is the purchase price or the loan to price ratio for house h in the neighborhood with zip code n purchased by individual i at time t. In the baseline regression, we control for the house fixed effects, α_h, which absorb will all time-invariant observable and unobservable house characteristics. Note that by controlling for house fixed effects, we are essentially using the repeat-sales in the sample. We also control for demand and cost shifters that may affect the dependent variable by including neighborhood-by-time fixed effects α_{nt}, where time is at the year-month level. Also note that inflation is captured by α_{nt}. All regressions are

[4]We do not observe a name change for each buyer ID and therefore we are less concerned that buyers endogenously manipulate their names to pay higher or lower house prices.

clustered at the census-block level.[5]

The variable of interest is *Gender*. This variable is equal to one if the buyer is female, and zero otherwise. In the case of multiple buyers, if one of the buyers is female, then *Gender* is equal to one. *Exp* is equal to one if individual i has purchased houses before period t and zero otherwise. Note that because we are interested in the gender gap in the house purchase price and the loan to price ratio, we cannot include individual buyer fixed effects because this will absorb the variable of interest *Gender*.[6]

4 Results

We begin our analysis by first reporting the fixed effects estimates of equation (1) for house purchase prices, which are presented in columns 1 to 4 in Table 2. Column 1 suggests that among single buyers, i.e. only one name is recorded on the deed transfer, women on average pay 2.3% more in house prices than men. This effect is statistically significant. In terms of magnitudes, women pay $4,263 more than men for a house with the average price. The median earnings of women is $41,977 and the median earnings of men is $52,146 in 2017 (Census Bureau). This effect accounts for 10% of the median woman's earnings and is therefore economically significant.

Column 2 allows gender gaps in house purchase prices to be heterogeneous across decades. We create decade dummies for 1987–2000, 2001–2010, and 2011-2017 and interact these with *Gender*. We find that although gender gaps still exist before in the previous decades, in the most recent decade, women are no longer paying more for house prices than men. Columns 3 and 4 repeat the analysis for multiple buyers. Among multiple buyers, when one of the buyers is a woman, the buyers tend to pay more in house prices than buyers who are all men. In terms of magnitudes, buyers with at least one woman pay $3,166 more than buyers who are all men for a house with the average price. Similar to to the single buyers, among the multiple buyers, the gender gap in house purchase price is also disappearing in more recent years. Note that our results are robust to different decade cutoffs.[7]

Having established that women pay more in house purchase prices than men using repeat-sales, one possible explanation for the gender gap in house purchase prices is that women are tasted-discriminated against by sellers. Women could be less desirable buyers because it's harder for them to take out loans, this is especially the case for single buyers who are women. Next, we check if gender gaps exist in the amount of loan taken out to purchase the house. We use the loan to price ratio as the dependent variable, and estimate equation (1) using a

[5]Note that there are 230 census blocks and 507 zip codes in the data.

[6]One concern about the model captured by equation (1) is that homes purchased by men and women may have experienced different rates of appreciation within the same neighborhood due to time-varying house characteristics unobserved by econometricians. If one gender tends to buy houses with higher appreciation rate, this would bias the Gender variable. However, due to the gender-wage gap (established in the labor literature), we would expect that men would be more likely to outbid women on houses with higher appreciation rates. This implies that our estimate of β_1 is more likely to be a lower bound of the gender-price differential. Another concern is that there is selection to buy based on income. We do not have buyers' income information, however, we can apply the logic before and we would expect men are more likely to outbid women on the same house given the gender-wage gap. As a result, our estimate of β_1 provides a lower bound to the true gender-price differential.

[7]Results are similar when we use 4 dummies, 1987-1990, 1991-2000, 2001-2010, and 2011-2017. These are available upon request.

Table 2 Disappearing Gender Gaps in House Prices

	(1) Single-name	(2) Single-name	(3) Multiple-names	(4) Multiple-names
Gender	0.023***		0.011**	
	(0.005)		(0.004)	
Gender ×		0.025***		0.018***
1987-2000		(0.007)		(0.007)
Gender ×		0.030***		0.008
2001-2010		(0.007)		(0.005)
Gender ×		-0.009		0.004
2011-2017		(0.012)		(0.010)
experienced	-0.021***	-0.021***	-0.007***	-0.007***
	(0.005)	(0.005)	(0.003)	(0.003)
Observations	47,051	47,051	54,375	54,375
House FE	YES	YES	YES	YES
Zip code-year-month FE	YES	YES	YES	YES
Adjusted R-squared	0.901	0.901	0.940	0.940
R-squared	0.976	0.976	0.987	0.987

NOTE: The dependent variable is the log price. Robust standard errors clustered at the cencus-block levels are reported in the parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

linear probability model. Results are reported in columns 1 to 4 in Table 3. The estimates from column 1 suggests that among single buyers, women's loan to price ratio is lower than men's. This implies that women take out less loans than men to purchase a house for which men and women pay the same price. In addition, results from column 2 suggests that the gender gap in loan to price ratio is disappearing in more recent years. It is worthwhile to point out that among multiple buyers, when one of the buyers is a woman, there is no gender gap in the loan-to-price ratio on average. More interestingly, in more recent years, when one of the buyers is a woman, the multiple buyers (potentially couples) take out more loans.

Our finding that women have to pay more in down payment than men to get loans for houses with the same price is consistent with the taste-based discrimination explanation. An alternative, not mutually exclusive, explanation is that in very stressful situations such as house-buying process, female sellers may be more friendly to female buyers than male sellers. We investigate this possibility by estimating the following equation:

$$y_{isht} = \mathbf{Gender}_{is}\Gamma + \beta_2 Exp_{it} + \alpha_h + \alpha_t + \varepsilon_{isrt}, \tag{2}$$

where $\mathbf{Gender}_{is}$ is a vector containing the gender dummies for buyer i and seller s. We use {seller male; buyer male} as the base category, and include the following dummies: {seller female; buyer female}, {seller female; buyer male}, {seller male; buyer female} in equation

Table 3 Disappearing Gender Gaps in Loan Ratios

	(1) Single-name	(2) Single-name	(3) Multiple-names	(4) Multiple-names
Gender	-0.030***		-0.003	
	(0.005)		(0.005)	
Gender ×		-0.032***		-0.006
1987-2000		(0.006)		(0.008)
Gender ×		-0.031***		-0.009
2001-2010		(0.005)		(0.006)
Gender ×		-0.018		0.019**
2011-2017		(0.014)		(0.009)
experienced	-0.011***	-0.011***	-0.006*	-0.006*
	(0.004)	(0.004)	(0.003)	(0.003)
Observations	47,051	47,051	54,375	54,375
House FE	YES	YES	YES	YES
Zip code-year-month FE	YES	YES	YES	YES
Adjusted R-squared	0.188	0.188	0.215	0.215
R-squared	0.801	0.801	0.832	0.832

NOTE: The dependent variable is the loan ratio (loan / price). Robust standard errors clustered at the cencus-block levels are reported in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

(2). Note that in our data, seller name is not recorded in most transactions, therefore our sample size is greatly reduced when we estimate equation (2). After controlling for house fixed effects (i.e. still using repeat-sales approach), we can only control for year-month fixed effects.[8] Standard errors are robust and clustered at the census block level.

We first report the results estimated from equation (1) in Table 4 columns 1 and 2 to confirm that gender gaps exist but are disappearing in house purchase prices using this reduced sample. Next, we report the estimates from equation (2) for single buyers in column 3. Results suggest that compared to when both buyers and sellers are male, female buyers pay 2.7% more in house purchase prices than male buyers. More interestingly, when sellers are female, female buyers pay 4.1% less in house purchase prices, and male buyers pay 4.9% less in house purchase prices. This pattern is similar for multiple buyers in column 6, but the magnitude is smaller. The results in Table 4 suggest that female sellers are more friendly than male sellers to female buyers, but female sellers also tend to accept lower prices from buyers regardless of buyers' gender compared to male sellers.

5 Conclusion

Using repeat-sales data from a online real estate database, we find that women pay more in house prices and take out less loans. However, this difference in gender is disappearing in

[8]There are not enough observations to control for neighborhood-by-time fixed effects.

Table 4 Heterogeneity in gender price differentials based on seller's gender

	(1) Single-name	(2) Single-name	(3) Single-name	(4) Multiple-names	(5) Multiple-names	(6) Multiple-names
Gender	0.029***			0.014***		
	(0.011)			(0.004)		
Gender ×		0.028			0.037***	
1987-2000		(0.020)			(0.008)	
Gender ×		0.042***			0.008	
2001-2010		(0.015)			(0.005)	
Gender ×		-0.002			-0.004	
2010-2017		(0.033)			(0.010)	
Seller female ×			-0.041*			-0.016***
Buyer female			(0.023)			(0.005)
Seller male ×			0.027**			0.015***
Buyer female			(0.013)			(0.004)
Seller female ×			-0.049***			-0.023***
Buyer male			(0.017)			(0.006)
experienced	-0.008	-0.008	-0.008	-0.009***	-0.009***	-0.008***
	(0.012)	(0.012)	(0.012)	(0.003)	(0.003)	(0.003)
Observations	2,574	2,574	2,574	17,108	17,108	17,108
House FE	YES	YES	YES	YES	YES	YES
Year-month FE	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.889	0.888	0.890	0.942	0.942	0.942
R-squared	0.956	0.957	0.957	0.971	0.971	0.971

NOTE: The dependent variable is the log price. Robust standard errors clustered at the census-block levels are reported in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

more recent years.

This study has limitations, for example, we cannot control for time-variant heterogeneities in house qualities. House characteristics in our data set such as square footage of the house, number of bedrooms, number of bathrooms, and square footage of the yard are mostly time-invariant and therefore are absorbed by house fixed effects. We also cannot directly control for time-variant unobserved heterogeneities in house qualities, however, some of these unobserved heterogeneities may be captured by the neighborhood-year-month fixed effects. In addition, we have not controlled for other individual demographic characteristics that may affect the house purchase price and loan to price ratio, such as a person's income, race, occupation, and education. Next, we plan to use this data for all 50 U.S. states and merge with the census data using a matching algorithm based on a person's name, sex, and zip code to obtain information on other individual characteristics to help explain why the gender gaps in house purchase prices are disappearing in more recent years.

Using loan contracts from Italy, Alesina, Lotti and Mistrulli (2013) find that women pay more for credit than men and this is not driven by lack of credit history or women using different banks than men. We also want to merge our data with the U.S. Loan Application Registrar (LAR) to obtain information on the terms of the loan and answer questions related to gender gaps in down payment and mortgage interest rates.

References

[1] Alesina, A., F. Lotti, and P.M. Mistrulli. (2013) "Do Women Pay More for Credit? Evidence from Italy" *Journal of the European Economic Association* **11**, 45-66.

[2] Arrow, K. J. (1973) "The Theory of Discrimination" in *Discrimination in Labor Markets* by O. Ashenfelter and A. Rees, Eds., Princeton, NJ: Princeton University Press.

[3] Becker, G. (1961) *The Economics of Discrimination*. University of Chicago Press.

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

[5] Chen, H. (2018) "Why are U.S. Women Decreasing Their Labor Force Participation If Their Wages are Rising?" *Economic Inquiry* **56**(4), 2010-2026.

[6] Bayer, P., M. Casey, F. Ferreira, and R. McMillan. (2017) "Racial and Ethnic Price Differentials in the Housing Market" *Journal of Urban Economics* **102**, 91-105.

[7] Bessendorf, A. (2015) "From Cradle to Cane: The Cost of Being a Female Consumer" New York City Department of Consumer Affairs.

[8] Cheng, P., Z. Lin, and Y. Liu. (2011) "Do Women Pay More for Mortgages?" *The Journal of Real Estate Finance and Economics* **43**(4), 423-440.

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

[10] Ihlanfeldt, K. and T. Mayock. (2009) "Price Discrimination in the Housing Market" *Journal of Urban Economics* **66**(2), 125-140.

[11] 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.

[12] King, AT. and P. Mieszkowski. (1973). "Racial Discrimination, Segregation, and the Price of Housing" *Journal of the Political Economy* **81**(3), 590-606.

[13] Ongena, S. and A. Popov. (2016) "Gender Bias and Credit Access" *Journal of Money, Credit, and Banking* **48**(8), 1691-1724.

[14] Zillow. (2017) "ZTRAX: Zillow Transaction and Assessment Database, 2017-Q4" Zillow Group, Inc.