THE GENDER GAP IN HOUSING RETURNS*

PAUL GOLDSMITH-PINKHAM

KELLY SHUE

June 2020

Abstract

Using detailed data on housing transactions across the United States since 1991, we find that single men earn 1.5 percentage points higher returns per year on housing relative to single women. Approximately 45% of the gap can be explained by differences in the location and timing of transactions. The remaining gap is due to a 2% gender difference in execution prices at purchase and sale, arising from differences in the choice of initial list prices and negotiated discounts. Overall, the gender gap in housing explains 30% of the gender gap in wealth accumulation at retirement for the median household.

*Respective contact information and affiliations are: paul.goldsmith-pinkham@yale.edu, Yale School of Management; kelly.shue@yale.edu, Yale School of Management and NBER. We thank James Choi, Tim Landvoigt (discussant), Gonzalo Maturana (discussant), Barry Nalebuff, Chris Palmer (discussant), Paola Sapienza (discussant), Matt Spiegel, and seminar audiences at Berkeley, Colorado Finance Summit, Five Star Conference, Florida International University, FMA, Georgia State University, Johns Hopkins University, NBER Behavioral Economics, Stanford, University of Michigan, University of Delaware, University of Pittsburgh, University of Toronto, Yale, and Yale-RFS Conference on Real and Private-Value Assets for helpful comments. We thank Gen Li, Huijun Sun, and Kaushik Vasudevan for excellent research assistance. This research was funded in part by the International Center for Finance at Yale SOM.

I. Introduction

Housing accounts for the majority of most American households' wealth, with Americans investing more in the housing market than in the stock market.[1] Housing also differs from other common forms of household savings, such as bank deposits, bonds, and stocks, in that it is an illiquid and heterogeneous asset with prices determined through bilateral negotiation. Motivated by the existing research showing gender differences in negotiation, financial investment strategies, and preferences for risk and competition (e.g., Sunden and Surette, 1998; Ayres, 1990; Babcock and Laschever, 2009; Sapienza et al., 2009; Bertrand, 2011; Reuben et al., 2015; Niederle and Vesterlund, 2007; Gneezy et al., 2003), we investigate how men and women differ in their financial returns on housing investment.

We use detailed data from CoreLogic covering over 50 million housing transactions and matched property listings across the US from 1991 to 2017. For approximately 9 million transactions for which we can identify homeowner gender, and the initial purchase and eventual sale prices, we compute the homeowner's annualized realized return.

We find that single men earn 1.5 percentage points higher unlevered annualized returns relative to single women. Approximately 45% of this gender gap in raw housing returns can explained by market timing, i.e., the choice of where and when to buy, and when to sell. Women earn lower returns on housing partly because they tend to buy in locations when aggregate house prices are high and sell when they are low. However, a large gender gap persists even after accounting for market timing. Among men and women who buy and sell in the same zip code and year-month, women still earn 0.8 percentage points lower unlevered annualized returns.

We also find that couples underperform single women on a raw unadjusted basis, but the low returns for couples are primarily due to market timing. After adjusting for the location and timing of purchases and sales, couples significantly outperform single women and underperform single men. Couples also display unique dynamics, in that they underperform single men at purchase, but outperform single men at sale.

Most U.S. home buyers purchase housing using mortgage debt with loan-to-value ratios of 80 percent or higher, and have not paid down a large fraction of the principal at the time of sale. Therefore, the real return earned is typically a levered return. We estimate that, if homeowners use leverage via a standard 30-year fixed rate mortgage with a 20% downpayment and no refinancing, the gender gap

[1] We plot the share of net worth invested in housing and the stock market for each percentile in the wealth distribution in Appendix Figure A1 using data from the Survey of Consumer Finances.

would be significantly larger. Assuming this degree of leverage, men outperform women by almost 8 percentage points per year. This growth in the gender gap arises because leverage amplifies raw return differences.

The gender gap in housing returns varies with the business cycle, but exists in all years within our sample and has not declined significantly in recent years (similar to the persistence of the unexplained gender wage gap in the most recent two decades (Goldin, 2014; Blau and Kahn, 2017)). Zip codes with lower average education and income, greater average age, and higher fraction single female are associated with larger gender gaps. However, the gender gap remains large in regions with high average education, income, and house price levels. The gender gap is largest in the right tail of the return distribution, although it also present at the median, and does not significantly reverse in the left tail.

In more detailed analysis, we explore the sources of the gender gap in housing returns. Controlling for market timing, we show that the remaining gender gap arises primarily because of a gap in execution prices at the points of purchase and sale. What happens between purchase and sale, e.g., potential mechanisms in which men invest more in upgrades and maintenance or buying riskier properties that naturally earn higher returns, appear to be less important drivers of the gender gap.

We begin by examining data on repeat sales of the same property. We further compare men and women who transact in the same zip code and year-month. This allows us to examine the remaining gender gap after adjusting for differences in market timing. Holding the property fixed, and comparing men and women to transact in the same zip-year-month, women buy the same property for 1-2% more than men and sell for 2-3% less. This gender gap in transaction prices arises from differences in the choice of list price and negotiated discount relative to the list price. Again using repeat sales data that allow us to hold the property fixed, we find that female buyers purchase properties that are listed at higher relative prices. Despite buying at higher prices, female sellers also choose to list for lower prices when selling. In addition, women negotiate worse discounts relative to the list price.

This gender gap in prices varies with the match between the gender of the buyer and seller in each transaction. Again exploiting repeat sales, we find that the highest transaction prices are associated with male sellers and female buyers, and the lowest transaction prices are associated with female sellers and male buyers. Female sellers and male buyers are associated with the largest negotiated discounts relative to the list price, while male sellers and female buyers are associated with the smallest discounts.

These findings relating to list prices and negotiated discounts imply that women experience worse execution prices on real estate transactions at the points of purchase and sale. Differences in execution prices should matter much more for the annualized returns of short-term investors than for the returns of long-term buy-and-hold investors. Consistent with this insight, we find that the magnitude of the gender gap in annualized returns decays toward zero with holding length. Homeowners with shorter tenure in their properties "trade" assets more often, so variation in execution prices matter more for their annualized returns.

We also explore several other potential mechanisms which may drive the gender gap in housing returns. First, men may select properties with characteristics naturally associated with higher returns. In particular, men may purchase riskier properties, such that their higher return represents compensation for the additional risk. Second, men may invest more in housing maintenance and upgrades, such that their real investment return is lower than implied by analysis using only the purchase and sale prices. Third, women may be older, have more children, or have lower education and income, and these demographic factors may contribute to the gender gap in housing returns. Fourth, men may employ better real estate agents when engaging in housing transactions.

We find that differences in preferences for types of homes do not substantially affect the gender gap in housing returns. Men and women indeed differ in their choice of home type (e.g., new construction and number of bedrooms), but controlling for these features does not significantly affect the gender gap in housing returns. Then, using detailed demographics data from the American Housing Survey, we replicate our main results showing a gender gap in housing returns. We find that homeowner age, education, income, and number of children significantly predict returns. However, the unexplained gender gap remains large after controlling for these variables. We also find no gender differences in reported home maintenance investment in the American Housing Survey data. In addition to routine maintenance, men may invest more in non-routine upgrades or renovations. Using CoreLogic listings data, we find that men are more likely to list homes that have been upgraded or renovated, but the difference in upgrade rates is small. The gender gap in housing returns remains large in a restricted sample for which the house listing does not mention any upgrades or renovations. Further, we find that the gender gap in returns persists after controlling for listing agent fixed effects, implying that women do not hire systematically worse real estate agents on average.

We also use the empirical relation between annualized returns and holding length to assess the importance of the property risk or maintenance explanations. Gender differences in routine maintenance or property risk imply that the the gender gap in annualized housing returns should remain positive as holding length increases. For example, if men purchase riskier properties that warrant an extra 1% return each year, the gender gap in annualized returns should asymptote toward 1% as holding period increases. Instead, we observe a gender gap in annualized housing returns that decays toward zero with holding length. This pattern is more consistent with a gap in returns that arises from gender differences in execution prices.

Finally, we examine how the gender gap varies with market tightness. In tight markets, sale volume is high relative to the number of outstanding house listings. As markets tighten, the gender gap in returns, transaction prices, and negotiated discounts all shrink toward zero. This pattern suggests that bilateral negotiation may be an important driver of the *average* gender gap in housing returns. As markets tighten, bilateral negotiations are typically replaced with quasi-auctions, in which multiple interested buyers bid for homes. The reduction in the gender gap with market tightness is also inconsistent with an explanation in which men buy riskier properties or invest more in home maintenance and upgrades, because risk and home improvements should, if anything, be more efficiently priced in liquid markets. Similarly, the reduction in the gender gap with market tightness is inconsistent with women earning lower housing returns because they derive greater utility from housing and are therefore willing to pay more for the same house. If so, women should also submit higher bids for homes, leading to lower returns in tight housing markets.

Our analysis is related to existing research examining gender differences in stock market participation, portfolio allocation between stocks and bonds, and investment performance (e.g. Sunden and Surette, 1998; Bajtelsmit and VanDerhei, 1997; Hinz et al., 1997; Barber and Odean, 2001). We believe it is equally or more important to study gender differences in housing investment, given that housing represents a much larger proportion of the typical household's savings portfolio. Housing also differs from other common forms of household savings because prices are often determined through bilateral negotiation. In contrast, we would expect men and women to earn the same return on an investment in a S&P500 index fund (holding timing constant), even if one group were more financially sophisticated or derived greater personal utility from owning the asset.

Our findings that women negotiate worse discounts relative to the list price suggests that gender differences in negotiation contribute to the gap in housings returns. The fact that women choose to list homes at lower prices also suggests that gender differences in "first offers" within a negotiation framework may play an important role. These patterns are consistent with a large literature documenting gender differences in the ability, style, and willingness to negotiate. This literature has shown that women have more negative outcomes when negotiating in laboratory settings, as well as in labor market and automobile market settings (e.g., Ayres, 1990; Ayres and Siegelman, 1995; Castillo et al., 2013; Leibbrandt and List, 2014; List, 2004; Morton et al., 2003; Reuben et al., 2015).

Housing negotiations differ from standard labor market or automobile negotiations in that buyers and sellers are often advised by real estate agents, who may vary in skill and may not perfectly share the interests of their clients.[2] We find that the gender gap remains similar after controlling for listing agent fixed effects, implying that the gender gap is not driven by female sellers systematically choosing worse agents. However, it remains possible that the same agent gives different advice to men and women, or that men and women differ in whether they follow the (good or bad) advice of their agents. Our observational data unfortunately does not allow us to examine the details of discussions between clients and agents, so we leave further study of the role of agents to future work.

The format of housing negotiations resembles that of semi-anonymous experimental ultimatum games. In housing negotiations, buyers and sellers typically submit written bids through agents rather than negotiating face to face, although buyers and sellers are usually aware of basic facts regarding each other's identities, such as gender. Solnick (2001) studies a related ultimatum game between two players in which Player 1 chooses how to split $10 between him or herself and Player 2. Player 2 sees the offer and can choose to accept the offer or reject, in which case both players receive zero. The players are randomly matched and are aware of each other's genders, but remain otherwise anonymous and do not interact face to face. Solnick (2001) finds that gender significantly affects game payoffs. Player 1s, especially female Player 1s, on average offer more money to Player 2 if Player 2 is known to be male. Male Player 2s choose higher minimum acceptable offers, especially if Player 1 is known to be female. These experimental results are consistent with an important idea that may apply to housing negotiations: both men and women expect women to be more willing to share the surplus from negotiations, and are more willing to walk away from aggressive offers if those offers are proposed by women.

We caution that our results do not necessarily imply that women make mistakes in housing negotiations or have lower negotiation ability. Exley et al. (2016) show that women can experience more negative outcomes by "leaning-in" and negotiating aggressively, and Ayres and Siegelman (1995) find

[2] Some labor market negotiations with documented gender pay gaps also involve agents, such as those in which talent agents represent professional actors or athletes. Our housing setting is also unique because personal interactions usually end after housing transactions are completed, so the gender gap is unlikely to be explained by women placing greater value on continued relationships, a factor that might impact labor market negotiations (Babcock and Laschever, 2009).

that female buyers receive worse counteroffers during automobile negotiations even when they follow identical scripts used by male buyers. Moreover, women may equal men in negotiation ability, but conduct the housing search differently because they derive greater utility from a fast, low-risk, or non-confrontational negotiation process. It is not the goal of this paper to disentangle negotiation ability from other preferences that could affect negotiated outcomes.

Regardless of the exact channel, our analysis shows that gender differences in housing returns are economically large. A simple calculation implies that the housing-related gender gap in dollars is approximately $1,600 per year for the median single female homeowner. This is approximately half as large as the unexplained gender pay gap, which has been the subject of numerous academic studies and policy debates (see e.g., Blau and Kahn (2017); Card et al. (2015)). Because housing wealth is the dominant form of savings for most households, our findings also offer insight into variation in wealth accumulation (e.g., Ruel and Hauser (2013)). We estimate that the gender gap in housing returns can explain approximately 30% of the overall gender gap in wealth accumulation at retirement.

Finally, our research is related to contemporaneous work by Andersen et al. (2018) (AMNV), which examines gender differences in negotiations for real estate transactions in Denmark. Our analysis differs from and complements AMNV in several ways. First, AMNV (along with related earlier research by Harding et al. (2003)[3]) is focused on demographic variation in bargaining. In contrast, we focus on the total gender gap in housing returns from the perspective of wealth accumulation. We are interested in factors beyond negotiation; for instance, we show that market timing can explain 45% of the gender gap in realized returns, and that the gender gap in housing is an important contributor to the gender gap in wealth accumulation at retirement. Second, we exploit additional features of the US data to explore the role of listing agents, variation by market tightness, and dependence on investment holding length, which help to reject potential alternative explanations. Third, we differ from AMNV in our usage of data on both list prices and final transaction prices, which allows us to examine negotiated discounts and how they vary with the match between the genders of buyers and sellers. Fourth, we find large gender differences in transaction prices using repeat sales data in the US, while AMNV find smaller and insignificant differences using repeat sales data in Denmark. These results complement each other, and suggest that the gender gap may vary by country and culture.

[3] Harding et al. (2003) uses data from the American Housing Survey and structural modeling methods, instead of a repeat sales approach, to estimate how bargaining power varies with demographics. Our evidence of inequality in housing markets is related to findings by Bayer et al. (2017), who use panel data covering four U.S. cities to show that Black and Hispanic buyers pay a 3% premium when purchasing homes. We unfortunately cannot examine racial gaps in our main data sample because we do not observe detailed demographic information for buyers and sellers, and instead infer gender from names.

II. Data and Measurement

In this section, we describe the data sources for our analysis, the construction of key measures, including identification of gender, and summarize the overall data set.

A. CoreLogic Deeds and Listings Data

Our main housing transaction data comes county deeds records gathered by CoreLogic. We restrict our analysis to arms-length transactions (sales between two unaffiliated parties) and exclude non-transaction deed events such as mortgage refinancings. Each observation reflects a housing transaction, containing information on the date of the transaction, the sale price, the exact address of the property, and the names of both sides of the transaction. This last set of data fields allows us to partially identify the gender of the participant, as well as the number of participants on each side of the transaction (discussed further below).

To supplement the transactions data with time-varying measures of the properties' characteristics, we link the deeds dataset by property location to a dataset of property listings also constructed by CoreLogic.[4] These data come from Multiple Listing Service (MLS) systems operated by regional real estate boards. Each listing includes a large number of fields describing the property and the status of the listing. These include when the property is listed, the list price, listing agent identification number, and the listed property features such as the number of bedrooms and bathrooms and age of the structure. If the listing sells, we observe the close date and sale price.

When linking the deeds and listings data, we match properties based on unique geographic location identifiers, which are the parcel identifier and county. This provides a many-to-many match between deed transactions and listings. We then match a deed transaction to a listing by: 1) matching the closest pair where the deed transaction date is within a year after the listing sale date, and 2) restricting that the difference in the log sale price in the deeds data and the log sale price in the listings data is less than one.

B. Measurement of Gender and Family Structure

We identify the gender and family structure (single or couple) of the buyers and seller on each transaction using reported names on the deed. For each deed record reflecting an arms-length transaction, CoreLogic reports the full name of the first and second buyers on a deed and the full names of the first and second sellers. We identify two pieces of information from these name fields: first, we parse the

[4]Properties are uniquely identified via parcel number (assigned by county deeds offices) and county code.

fields to identify how many parties exist on each side of the transaction, since in some cases, couples are transcribed as "John and Mary Smith" in one field, rather than being split across fields as "John Smith" and "Mary Smith." Second, we use the first names to probabilistically assign a gender to each party in the transaction. We follow Chari and Goldsmith-Pinkham (2017) and use data from Tang et al. (2011) to measure the probability that a given name is male or female (based on self-reported data to Facebook). Then, for all names with associated gender probabilities greater than 95%, we assign either male or female gender. For those who do not match, or whose probabilities are less than 95%, we treat as unknown genders.

Identification of the number of parties, and their respective genders, allows us to group each side of the transaction into four "gender groups": single male, single female, couples (2 individuals with both genders identified), and other, where other is the residual category and will include single individuals without gender identified, couples where only one gender is identified, couples where neither gender is identified, and institutions. For each transaction, this grouping is done for both the buyer and seller side.

In a small set of counties during the early years of our sample period, we observe an unusually high percentage of single male buyers and sellers, and an unusually low percentage of couple buyers and sellers. Since housing records are maintained at the county level, we believe these county-years represent cases in which only one name, usually the male name, was recorded for transactions actually conducted by couples. To address this measurement issue, we impose that, in a given county-year, we are able to identify the gender and family structure (single or couple) for at least five percent of deeds, and that within identified deeds, at least 20% are identified as couples, and at least five percent as single women. Otherwise, all observations in the county year are reclassified as the "other" category due to inability to credibly identify gender and family structure. The procedure reclassifies approximately four hundred thousand observations. In Appendix Table A1, we show that our main results remain similar if we do not apply this screen.

C. Measurement Error

Our measures of gender and family structure may be subject to three types of measurement error. First, we fail to identify gender for some individuals, and they will be relegated into the "Other" category. This is more likely with non-Anglo-Saxon names where the gender is less predictable based on name. Second, we may miscategorize some single women as single men and vice versa. Given our cutoff for gender categorization is 95% or greater certainty based on names, we are less concerned about this type of error, but it remains possible. Finally, some single men and women identified in our data may actually correspond to couples who choose (or follow local convention) in recording only a single name in a real estate transaction. As noted in the previous section, we filter out county-years which appear to follow a convention in which transactions involving couples only list one person's name (usually the husband's name). These cases usually occur in the early years of our data sample. However, it remains possible that our filter does not identify all such cases.

We address measurement error in three ways. First, we examine supplementary data from the American Housing Survey (AHS), a representative panel dataset of US homes. Homeowner gender and marital status in the AHS are self reported and much less likely to suffer from measurement error. We are able to replicate our baseline finding of a larger than 1 percentage point gender gap in unlevered housing returns. These supplementary results show that our main conclusions are unlikely to be driven by errors in our gender and family structure identification algorithm.

Second, we show via a simple extension of Aigner (1973) that, under reasonable assumptions, measurement error in our gender and family structure measure would cause us to underestimate the extent to which single women earn lower returns on housing relative to single men. Full details concerning assumptions and derivations are presented in Appendix B.

Third, we report the gender gap in realized returns for each year within our sample period in Figure 5. In recent years, there is likely to be less measurement error in which true couples are misclassified as single men or women because only one name is reported on deed records (see Appendix Figure A3 which shows that the ratio of identified single male buyers to identified couple buyers falls in the 1990s and then stabilizes in the post-2000 period). We continue to find economically large and statistically significant gender gaps in housing returns in recent sample years.

Finally, we note that, while a strength of our data is the large sample coverage, we face the limitation that we can only analyze differences in housing returns among the women, men, and couples who appear in our data. In particular, we measure a gender gap among men and women who *choose* to own housing. Our estimates are not meant to represent potential housing returns for the universe of men and women. Despite these limitations, we believe our estimates are informative of the sources of the gender gap among the large set of Americans who choose to be homeowners. Our data limitations are similar to the limitations faced by the large literature examining the gender pay gap, which likewise can only be measured among the set of men and women who choose to work.

D. Measuring Unlevered Returns and Levered Returns

In our full dataset of transactions, we are able to identify consecutive arms-length market transactions for each property. We focus on realized returns, earned in the period from purchase to sale. Using these consecutive transactions, we can identify the unlevered annualized return for property i in sale year s: $r_{is} = \left(\frac{P_{is}}{P_{ib}}\right)^{\frac{1}{(s-b)}} - 1$, where P_{ib} is the purchase price, i.e., the previous market transaction price on the property in year b. Because we observe exact dates for transactions, we allow years b and s to be non-integers to better measure the exact holding length of each property. We require a minimum holding length of three months to be included in the returns sample, and also present results separately by holding length buckets. To ensure that we correctly measure r_{is} for single male, single female, and couples, we focus on the subsample of returns that has three additional restrictions: (1) we have identified the gender and family structure in both periods b and s, (2) the gender and family structure of the buyer in period b corresponds to the gender and family structure in period s, and (3) the names of the buyers in period b is sufficiently close to the names of the sellers period s by string matching distance.[5] This final sample is used for our analysis of housing returns. These filters substantially restrict our analysis sample, since we need to observe multiple transactions and correctly identify gender and family structure, but ensures that we are not incorrectly measuring returns. Our final returns sample contains 8.9 million observations.

Restriction (2) also ensures that we can focus on the housing returns of men and women who were identified as single at both purchase and sale. Observations in which a house is purchased by a couple and sold by a single person (which can occur after death or divorce), or purchased by a single person and sold by a couple (which can occur after marriage), are not used when comparing returns across single men, women, and couples.

In reality, the majority of homeowners in the United States buy their homes using debt, with leverage of five-to-one or higher. Moreover, this leverage tends to persist over a long period of time, due to long duration mortgages whose fixed amortization schedules pay mainly interest upfront. Therefore, the real return earned is typically a levered return. Ideally, given the mortgage type, term, interest rate and downpayment, we could identify the exact levered return earned by each homeowner. However, many of these fields are missing from the data. Therefore, we compute hypothetical levered returns if the homeowner used the most common type of mortgage in the data (a 30-year fixed rate loan with an initial loan-to-value (LTV) ratio of 80% with the market interest rate in the year-quarter of the initial

[5] We use a tolerance of 0.7 for the string matching function *matchit* in Stata.

purchase) and does not refinance. We also present summary statistics for the average level of leverage used by single men, women, and couples over time, as well as the rates of missing data within each group. Finally, in supplementary results, we calculate levered returns using actual LTV ratios for a more recent sample period which has fewer missing data issues.

We measure the initial LTV as the ratio of the initial mortgage amount divided by the purchase price. In many cases, the mortgage amount is missing. In recent years, we believe missing mortgage amounts represent homes purchased with cash. In 2017, for example, 28.8% of all housing transactions in our data have missing mortgage amounts, which approximately matches external estimates of the fraction of all cash purchases in the 2000s.[6] However, in the 1990s, the share of missing mortgage amounts in our data is substantially higher, and suggests that there may also be missing data issues.

To address these data issues, we provide a simple benchmark in which we convert unlevered housing returns into levered returns assuming all purchases were done with an 80% LTV mortgage. This holds fixed any potential leverage differences across transactions, and instead converts the unlevered returns into a measure that captures the modal degree of leverage that homeowners face. Based on the purchase value and assuming 80% LTV, we calculate the initial downpayment D_{ib} and initial mortgage amount Mortgage_{ib} used to buy the home. We then calculate the average interest rate in the year-quarter of initial purchase by taking the 30-year fixed rate mortgage rate from Freddie Mac. Using this interest rate ρ_{ib}, we calculate the relative principal pay down at every monthly duration horizon, assuming no refinancing, and use this to identify the share of remaining mortgage principal outstanding when the house is sold in period s (Mortgage_{is}). This allows us to calculate the total cash out payment for the home at the time of sale: $\text{Equity}_{is} = max\{P_{is} - \text{Mortgage}_{is}, 0\}$.[7] We then approximate the time b net present value of equity as the sum of the downpayment plus the discounted value of principal paydown payments: $\text{Equity}_{ib} \approx D_{ib} + \sum_{\tau=b}^{s} W_{i\tau}/(1+\rho_{ib})^{\tau-b}$. As a result, our levered annualized return is $r_{is}^{\text{lev}} = \left(\frac{\text{Equity}_{is}}{\text{Equity}_{ib}}\right)^{\frac{1}{(s-b)}} - 1$. Note that in the case of a full cash purchase, $r_{it}^{\text{lev}} = r_{it}$.

In additional robustness results reported in the Appendix, we use the exact initial mortgage amount in the data to calculate the LTV ratio at the time of purchase, and use this to calculate levered returns. Due to missing data issues in the early years of our sample, we limit this analysis to purchases in 2000 and after. We continue to make the same interest rate and maturity structure assumptions as described above, as many of these fields are missing in the data even in recent years.

[6]The Wall Street Journal, available https://www.wsj.com/articles/want-that-house-youd-better-pay-in-cash-1512469800, estimates 20-30% of home purchases in the 2000s were all cash transactions.

[7]By using the max operator, we are implicitly assuming that homeowners cannot lose more than 100% of their original equity. Removing the max operator yields similar results.

It is important to note that these calculations estimate *realized* levered and unlevered returns conditional on sale. These returns do not represent returns in a counterfactual world in which all homeowners are forced to sell at a fixed horizon. Instead, they measure realized returns for those homeowners who choose to sell.

E. Description of Data

In Panels A and C of Table 1, we report the average sale transaction price broken across buyer and seller gender groups (single male, single female, couple, and other) and pooled into overall averages. We are able to credibly identify the buyer gender and family structure according to the criteria described in Section B for approximately 62% of the sample, or around 32 million transactions. We are able to credibly identify the seller gender and family structure according for approximately 40% of the sample, or around 21 million transactions. Among identified gender groups, couples have the largest share of transactions, followed by single men and then single women. We refrain from interpreting gender differences in raw sale prices in the summary statistics table, because men and women may purchase properties with different average quality. Our subsequent analysis will focus on housing returns or exploit repeat sales, which hold the property constant.

In Panels B and D, we report the averages for our sample of sales transactions that are linked to listings data, which has more limited coverage for the early years of our sample period. The linked sample is substantially smaller and covers roughly 20 million transactions. Approximately 68% and 50% of this sample's buyer and seller gender groups can be credibly identified, respectively. Unsurprisingly, this sample has higher prices, since the listings data covers a later sample period.

We measure discounts off the list price as (list price - transaction price)/list price $\times$ 100. A larger discount benefits buyers and hurts sellers in terms of returns on housing investment. Among the identified gender groups, single male buyers receive the largest purchase discounts, followed by couple buyers, and then single female buyers. Single female sellers give the largest sale discounts, followed by single male sellers, followed by couple sellers. The three groups have similar average days on market, equal to the number of days between the earliest available listing associated with a transaction and the sale date.

Finally, in Panel E, we report the averages for our sample where we observe both the purchase and the sale of a property, and we are able to confirm that the seller is the same as the buyer in the previous transaction. This restricted sample is smaller at roughly 8.9 million transactions. In this sample, we exclude the other group. We find that, on average, single men have the highest annualized unlevered return, at 8.40 percent, followed by single women at 6.85 percent, and couples at 6.46 percent. Single men hold their properties for 0.4 years less time than single women, and 0.6 years less than couples.

III. Empirical Results

This section describes our regression methodology and summarizes our main results measuring the difference in returns between single men, single women, and couples. We then assess the various channels that can explain this difference in returns.

A. Estimation Approach

Our main analysis takes two forms. Both approaches use a simple linear regression framework to account for potential differences across gender groups. The first is an analysis of the unlevered and levered annualized returns:

$$r_{is} = \text{Single Female}_{is}\beta_1 + \text{Couple}_{is}\beta_2 + X_{is}\tau + \epsilon_{is}, \tag{1}$$

where Single Female_{is} is an indicator for a single female seller in period s and Couple_{is} is an indicator for a couple seller in period s. We estimate this regression using the returns sample, which only includes single male, single female, and couple sellers. As a result, β_1 and β_2 capture the relative effect when compared to Single Male_{is}, the omitted category. X_{is} represents control variables. Without any control variables, the coefficients on Single Female_{is} and Couple_{is} measure the average raw return difference for each group relative to Single Male_{is}. We also present specifications in which X_{is} represents control variables for location and time. For instance, if X_{is} includes five-digit zip code interacted with sale-year-month fixed effects, then the main coefficients measure the average difference in returns among people who sold in the same zip code and time period.

Our second set of analyses, focusing on the channels affecting housing return, is similar but uses alternative outcome measures, such as the $\log(\text{Sale Price}_{it})$:

$$Y_{it} = \text{Single Female}_{it}\beta_1 + \text{Couple}_{it}\beta_2 + \text{Other}_{it}\beta_3 + X_{it}\tau + \epsilon_{it}. \tag{2}$$

Since these outcomes are not measured using within-home changes in price, we additionally include a property fixed effect in X_{it} to capture unobserved quality in the property that may be correlated with gender or family structure. To better estimate this property fixed effect, we include transactions that are not included in our returns data sample, and additionally control for the Other_{it} indicator.

B. Baseline Results

We begin by showing how housing returns differ by the gender group of the homeowner. We use observations at the sale transaction level. The sample is restricted to observations for which the gender of all sellers can be identified, and for which we can match the identity of the seller at the time of sale to the identity of the buyer at the time of initial purchase. The results are shown graphically in Figure 1 with detailed regression results reported in Table 2.

In column 1 of Table 2 Panel A, we find that single women earn 1.5 percentage points lower unlevered annualized returns than single men (the omitted category). We then explore how much of this overall gender gap can be explained by market timing, i.e., gender differences in the choice of when and where to buy, and when to sell. As we move from column 1 to column 5, we introduce more detailed control variables for market timing, including zip-year-month fixed effects for the initial purchase transaction and zip-year-month fixed effects for the sale transaction. We also control for the interaction between year-month of purchase and year-month of sale fixed effects, which subsume the control variable for holding length. We find that 45% the raw gender gap in returns in column 1 can be explained by more detailed control variables for market timing. In other words, women earn lower returns partly because they are more likely to buy in locations when aggregate house prices are high and sell when they are low. However, a large gender gap persists even after introducing detailed controls for market timing. Among men and women who buy and sell in the same zip code and year-months, women still earn 0.8 percentage points lower unlevered annualized returns on housing.

We also find that couples underperform single women in terms of raw returns but outperform single women and underperform single men after adjusting for market timing. Column 1 shows that the raw return gap in annualized unlevered returns between couples and single men is 1.9 percentage points. However, the relative returns for couples are very sensitive to the inclusion of controls for market timing. Moving from column 1 to column 5, we find that 79% of the return gap between couples and single men can be attributed to market timing. Among households that buy and sell in the same location and time period, couples outperform single women by 0.4 percentage points and underperform single men by 0.4 percentage points. These results indicate that couples earn lower returns primarily due to poor market timing, but outperform single women holding the location and transaction period fixed. These findings are consistent with the idea that couples (and possibly single women) face more binding constraints in the timing of real estate transactions due to child care and the school calendar system. In later analysis, we examine the behavior of each group separately at purchase and sale, and find that couples underperform when purchasing, but outperform even single men when selling.

Because most home buyers purchase housing using loan-to-value ratios of 80 percent or higher, and have not paid down a large fraction of the principal at the time of sale, the real return earned is typically a levered return. Appendix Figure A4 plots the average loan-to-value (LTV) at the time of initial purchase for each gender group over time. In Panel A, we report the LTV for the sample with mortgage amount data, while in Panel B we report the share of transactions with missing mortgage data information. Single men and women have higher average LTV than couples across all years conditional on non-missing mortgage data, but also have higher rates of missing mortgage data. Because missing mortgage data can represent cash purchases or true missing data, we do not draw strong conclusions about differences in leverage across groups.

In Panel B of Table 2, we assess variation in housing returns if homeowners used a standard mortgage contract: a 30-year fixed rate mortgage with a 20% downpayment. In general, allowing for leverage leads to a much larger gender gap in housing returns because leverage amplifies raw returns, and therefore amplifies differences in raw returns between groups. For each return observation, we compute the annualized levered return following the procedure described in Section II.D. Regression coefficients measure the the expected gender gap if households used the most common form of leverage during our sample period. We find the women underperform men by 7.9 percentage points per year in terms of these levered returns. Approximately half of the gap in levered returns can be attributed to market timing. However, women continue to underperform men by 3.7 percentage points per year after controlling flexibly for the location and timing of purchases and sales. We again find that couples underperform in terms of raw levered returns, but earn levered returns in the intermediate range after controlling for market timing: among households that buy and sell in the same zip code and year-months, couples outperform women by 2.2 percentage points and underperform men by 1.5 percentage points.

In addition to examining the gender gap in mean returns, we also compare the distribution of returns across groups. Panels A and B of Figure 2 show unlevered and levered annualized returns at various percentiles of the return distribution for each gender group. These figures report raw returns (without any adjustments for market timing). This set of figures reveal that the gender gap exists in all parts of the return distribution except for the left tail where women and men fare equally poorly. However, the gender gap is larger at the 90th percentile of the returns distribution than at the median. Couples underperform at most points of the raw return distribution, but as shown in Figure 1, couples earn returns between that of single men and women after adjustments for market timing.

In Panel A of Figure 3, we plot the density of unlevered annualized returns. The figure again shows that men weakly outperform women at all parts of the return distribution, with the largest differences in the right tail. In Panel B, we zoom in to the return near zero, where all groups have distributions with missing mass just to the left of zero. This dip in the distribution is consistent with loss aversion and the disposition effect (see e.g., Genesove and Mayer, 2001; Shefrin and Statman, 1985), in which people are reluctant to sell at less than their initial purchase price. Finally, Figures 2 and 3 show that men do not have worse left tail outcomes than women in terms of realized returns.[8] Therefore, compensation for greater downside risk in realized returns is unlikely to explain the higher average returns for men in our sample. However, we caution that we don't observe other adverse outcomes such as personal bankruptcy costs that may differ by gender.

In supplementary analysis reported in Appendix Table A2, we repeat the estimation procedure in Table 2, but focus on purchases in the year 2000 and afterwards, where there is less missing mortgage data. To calculate levered returns, we use the initial mortgage amount reported in the data (instead of assuming an initial LTV of 80%) and assume a cash purchase when the mortgage amount is missing. In Panel A, we see that restricting the sample to a more recent time period yields similar estimates for for the relative performance of single women, single men, and couples in terms of *unlevered* returns. In Panel B, we see very similar results in *levered* returns for single women relative to single men, but greater underperformance for couples. This is due to the fact that couples choose mortgages with lower initial LTVs (as seen in Appendix Figure A4). However, the lower return for couples comes with a benefit, as they are exposed to less risk due to their lower leverage. The same is not true for single women, who have similar average LTV compared to single men.

C. Heterogeneity

In Table 3, we explore how the average gender gap in housing returns within a zip code varies with zip-level demographics from the 2010 American Community Survey. We measure the gender gap in each zip code as the average difference between male and female returns and present the average

[8]Using a Mann-Whitney-Wilcoxon test, we formally reject the hypothesis that the returns for single men and single women are drawn from the same distribution, with a z-score of -71.292 and p-value of 0.000.

gender gap across quartiles of various zip-level demographic characteristics. We find that the magnitude of the gender gap decreases with education, and increases with age, fraction black, and fraction single female. We also find that the gender gap remains large even in zip codes in the top quartile of education, income, and house prices (measured relative to the state-year-month average).

Figure 4 shows the magnitude of the average difference in unlevered returns between single men and women across each state in our sample.[9] The gender gap is positive in almost all states within our sample with good data coverage. We believe that variation in the gender gap across states could be partly caused by differences in data quality and estimation error across states. As noted earlier in Section II.B, some single men and women identified in our data may actually correspond to couples who choose (or follow local convention) to record only a single name in a real estate transaction. If the degree of estimation error also varies across states, that could contribute to variation in the estimated gender gap across states (see Appendix B).

Figure 5 and Appendix Figure A5 show how the average and median realized returns varied over time and across gender groups. Realized returns on housing are positive in all years of our sample, but display significant business cycle variation, with the highest returns in the run-up to the housing market crash in 2006. This is consistent with recent findings in Sakong (2019) showing a relation between cyclical housing transactions and wealth inequality. The magnitude of the gender gap in returns appears to increase with average returns, although the gender gap remains large in magnitude in recent years and does not exhibit a strong secular decline over time.

Appendix Figure A3 shows the composition of transactions by gender group over time. Changing composition, combined with business cycle variation in average returns, implies that gender differences in market timing can play a role in the overall gender gap, but the overall shift in composition is relatively small.

D. The Gender Gap in Execution Prices

So far, we have shown that the gender gap in housing returns can be partly explained by gender differences in market timing. In this section, we explore gender variation in transaction prices, list prices, and transaction discounts, among buyers and sellers who transact in the same zip code and time period. We also discuss why the impact of a gender gap in execution prices on annualized returns depends on the holding length.

The unlevered annualized return on housing depends mechanically on the ratio of the sale price to

[9] We exclude states where the number of observations is less than 500.

the initial purchase price, annualized to account for holding length. To assess gender variation in each transaction price, we exploit repeat sales data and control for zip-year-month fixed effects to account for time trends within a zip code. Each observation in this analysis is a transaction. To better estimate property fixed effects, we do not restrict the sample to buyers or sellers with identified genders and matched names across sales and initial purchase. All observations corresponding to non-identified parties are included and coded as the "other" category. Thus, our sample size expands to over 50 million observations.

The results in Table 4 show that women purchase homes at approximately 1-2% higher prices than men, holding the property fixed and adjusting for local time trends in prices. Women also sell the same property for 2-3% less than men. Couples underperform single women in terms of purchasing at higher prices, but also outperform even single men in terms of selling at higher prices.

We can also examine how transaction prices vary with the match between categories of sellers and buyers. In Figure 6, we plot the coefficients from a regression of log transaction price on the interaction of seller gender and buyer gender, controlling for property fixed effects and zipcode by sale-year-month fixed effects. The base category is male buyers and male sellers, and each estimate should be interpreted as the relative price compared to that group. Among the four possible matches between male and female sellers and buyers, the highest transaction prices occur when there is a male seller and female buyer, and the lowest transaction prices occur when there is a female seller and male buyer. These results relating to the interaction between buyer and seller gender are consistent with field evidence from Hernandez-Arenaz and Iriberri (2018). Using data from a televised game show, Hernandez-Arenaz and Iriberri find that bargaining matches composed of a woman and man are the most favorable to men and least favorable to women.

While these results suggest that there may be strategic reasons to match with female buyers and sellers, we do not find strong evidence of unusual matching patterns. In Appendix Table A3, we examine how sellers and buyers of different genders and family structure match. We find that sellers and buyers of the same "type" (single male, single female, and couple) tend to match with themselves slightly more than would be expected under random matching, but otherwise do not find large deviations from what would be expected under random matching.

The final sale price for each transaction depends on the price at which the property is listed and the negotiated discount relative to the list price. To examine gender variation in list prices and negotiated discounts, we restrict the sample from Table 4 to observations that can be matched to MLS data on home listings, leading to approximately 20 million observations, 10 million of which correspond to repeat sales. Using repeat sales, and holding the property fixed and adjusting for local time trends in listed prices, we find in Table 5 that women choose to purchase the same property when it is listed for approximately 2% higher than when it is purchased by men. Even though women purchase homes at relatively higher prices, female sellers also choose to list the same property for 2% less than when it is listed by men. Couples again show differential behavior at purchase and sale. Couples purchase homes that are listed for 1-2% higher than when they are purchased by single men. However, couples choose list prices at sale that are only 0.4% lower than single men, after adjusting for zip-year-month fixed effects.

Next, we examine how negotiated discounts vary by gender in Table 6. We measure purchase and sale discounts as the percentage discount relative to the list price, (list price - transaction price)/list price $\times$ 100, so a larger purchase discount contributes to a higher return on housing investment and a larger sale discount contributes to a lower return on housing investment.[10] We find that female buyers purchase homes at a 0.26 percentage point lower discount relative to men. Thus, female buyers receive lower discounts on home purchases even though they also choose to purchase properties with relatively high list prices. Further, female sellers agree to 0.09 percentage points greater discounts at sale. Thus, female sellers agree to larger discounts off the list price even though they already choose to list at lower prices than male sellers.

The discounts negotiated by couple buyers and sellers also display interesting patterns. Couple buyers lie between single men and women in terms of purchase discounts. However, couples negotiate more advantageous discounts than even single men when selling properties. In combination with the earlier results on list prices, we find that couple sellers list the same property at approximately equal prices set by single men, but are less willing to agree to discounts off the chosen list price.

In Appendix Figure A8, we plot the full purchase and sale discount distributions. There is a large mass of discounts at exactly zero, with female buyers bunching the most at zero discount. When we condition on positive discounts, we see that for purchase discounts, female buyers have more mass on the left than male buyers (so female buyers receive lower discounts). This behavior flips for sale discounts (so female sellers agree to larger discounts).

We can also examine how transaction discounts vary with the match between categories of sellers

[10] To adjust for outliers, we censor discounts that are outside of the [-10,25%] range, implying a sale price that is less than 25% of the list price, or 10% higher than the list price. The endpoints of the range are chosen to approximately match the top and bottom 1% of the data.

and buyers. In Figure 7, we plot the results from a regression of discounts on the interaction of seller gender and buyer gender, including zipcode by year-month fixed effects. The base category is male buyers and male sellers, and each estimate should be interpreted as the relative discount compared to that group. Among the four possible matches between male and female sellers and buyers, the largest discount occurs when there is a female seller and male buyer, and the smallest discount occurs when there is a male seller and female buyer.

Given the gender differences in list prices and discounts, one may wonder whether female sellers benefit from less aggressive pricing with faster transaction times. For approximately 10 million observations, we observe the number of days on market between initial listing and sale resolution. Because some houses are listed multiple times if they fail to sell, we use the earliest available listing date and list price for each transaction. In Table 7, we find that women sell homes with 2.7% shorter transaction periods relative to men (the difference falls to 1% if we control for property fixed effects). This difference, while statistically significant, is economically small: the average transaction period is 42 days, and single women transact one day faster. Column 3 shows that the gender gap in sale prices remains large after controlling for the days on market. In Appendix Figure A9, we plot the full distribution of days on market for each gender group, and do not see substantial differences in the right tail of the distribution, implying that male sellers are not more likely to experience unusually long transaction times.

D.1 Variation by Holding Length

The gender gap in list prices and transaction discounts together imply that single women experience worse execution prices on real estate transactions at the points of purchase and sale. Differences in execution prices should matter much more for the annualized returns of short term investors than for the returns of long term buy-and-hold investors. In a simple model in which women buy properties for δ fraction more and sell δ fraction less then men and hold for t years, we expect that women will earn $100 \times 2\delta/t$ percentage points lower annualized unlevered returns than men.[11] In other words, the impact of a gender gap in execution prices on the gender gap in annualized returns should asymptote toward zero with holding length.

[11] Let P_0 represent the market price of the property at the time of purchase. Suppose that the market value of the property grows by a fraction r each year. Suppose men buy and sell at the market price, so their annualized return equals r regardless of the holding period. Suppose that women buy properties for a fraction δ more and sell for δ less than the market price and hold for a period of t years. We can solve for the annualized return for women r_F such that $[(1+r_F)^t = (1-\delta) * P_0 * (1+r)^t]/[(1+\delta) * P_0]$. Solving for r_F after applying the approximation that $log(1+x) \approx x$ for x close to zero implies that $r_F = r - 2\delta/t$.

Panel A of Figure 8 shows that the gender gap in annualized returns converges toward zero as holding length increases. Homeowners with longer tenure in their properties "trade" assets less often, so any advantage or disadvantage in execution prices will matter less for their annualized returns.[12]

In Panels B and C of Figure 8, we examine how the gender gap in execution prices vary with holding length. While there is some variation, the gender gap in execution prices does not asymptote toward zero as holding length increases. Women have significantly worse execution prices at purchase and sale for all holding length buckets in our sample. However, worse execution prices at the points of purchase and sale matter less for annualized returns on investment as holding period increases.

In other words, we find a large gender gap in execution prices even for properties held over many years. However, the gender gap in *annualized* returns asymptotes to zero with increased holding length. This occurs because the fixed dollar disadvantage at purchase and sale is divided over a greater number of holding years when computing annualized returns.

These empirical patterns also help to reject a potential mechanism involving "house flipping." House flippers have a strategy of buying, rehabbing, and quickly reselling properties for a profit. Single men may be more likely to pursue flipping strategies than single women, and that may contribute to their higher returns (Appendix Figure A7 shows that single men are slightly more likely than single women and couples to have short holding length). However, we continue to find a large female disadvantage in transaction prices for homes held for eight or more years, which are less likely to be flipped properties.

In our baseline analysis, we equally weighted each completed housing transaction when estimating the gender gap in housing returns. We can instead weight each transaction by holding length, so that each year in which a property is held receives equal weight. This alternative weighting scheme would place, e.g., six times the weight on a observation for a home held for six years relative to an observation for a home held for one year. We present results using this alternative weighting scheme for unlevered and levered returns in Panels A and B of Appendix Table A4, respectively. We find that the gender gap shrinks significantly: single women earn 0.4 percentage points lower unlevered returns and 1.4 percentage points lower levered returns relative to single men. The smaller estimated gender gap is consistent with the fact that gender differences in execution prices should matter less for annualized returns for investments with longer holding periods, and this alternative specification

[12] The *level* of annualized returns also declines with holding length for all groups, as shown in Appendix Figure A6. This pattern may occur due to selection, in which homeowners sell early only if they expect a high sale price that can cover transaction costs. However, the level of returns asymptotes toward 5% per year rather than zero, and cannot directly explain why the *gender gap* in returns asymptotes toward zero.

increases the weights on observations with longer holding periods.

E. Other Potential Mechanisms

In this section, we explore several other potential mechanisms which may help explain the gender gap in housing returns. First, men may select properties with characteristics associated with higher returns. In particular, men may buy riskier homes, such that their higher return represents compensation for the additional risk. Second, men may invest more in housing maintenance and upgrades, such that their real investment return is lower than implied by analysis using only the sale price and purchase price. Third, women may be older, have more children, or have lower education and income, and these demographic factors may drive the gender gap in housing returns. Fourth, men may earn higher returns because they choose better or more effective real estate agents.

E.1 Property characteristics, risk, maintenance, and upgrades

In Appendix Table A5, we find that gender is predictive of the types of properties held, i.e., age of house, square footage, and whether it was new construction at the time of purchase. Table 8 shows that, while some of these characteristics are predictive of housing returns, controlling for detailed home characteristics does not have a large impact on the estimated magnitude of the gender gap in returns (as evidenced by the small difference in coefficients on single female between columns 2 and 3). This analysis shows that women do not, on average, sort toward a set of housing characteristics that are associated with lower returns.

Next, we explore whether men earn higher returns because they invest more in housing maintenance and upgrades. In particular, men may be more likely to purchase fixer uppers, which would explain why men buy at low prices and sell at high prices, holding the property fixed. Given that maintenance and upgrades are costly (Harding et al., 2007), men's real investment return may be lower than implied by analysis using only the sale price and purchase price. We measure a renovated home using the description of the listing in the CoreLogic listings data, matching on any string that matches one of the following: renovation, remodel, new, update or restore.[13] In Appendix Table A6, we find that men are more likely to list homes that have been upgraded or renovated, but the difference in upgrade rates across genders is small. Table 8 shows that the gender gap remains large after controlling for whether the house has been upgraded or renovated, and Appendix Table A6 shows

[13]Specifically, we use a regular expression to match the following: "renov | remodel | new | update | restore".

that the gender gap in housing returns remains large in a restricted sample for which the house listing does not mention any synonyms for upgrades.

Aside from upgrades and renovations that are noted in property listings, men may also invest more in routine maintenance. In column 1 of Table 9, we use data on self-reported annual home maintenance costs (scaled by the price of the home) from the American Housing Survey and find insignificant and close-to-zero gender differences in maintenance investment.

So far, we have shown that the gender gap remains large for homes that have not experienced major upgrades/renovations, and men and women invest similar dollar amounts into home maintenance. However, it remains possible that men invest more personal time and effort into home maintenance. To further assess the importance of gender differences in home maintenance, we return to the empirical relation between holding length and the gender gap in returns. Examining this relation will also help to rule out a risk-based story in which men purchase riskier properties.

If men invest more in maintenance each year or purchase riskier properties, then the gender gap in annualized housing returns should not decay toward zero as holding period increases. For example, if men purchase riskier properties that warrant an extra 1% return each year as compensation for risk, then the gender gap in annualized returns should asymptote toward 1% as holding period increases. We instead observe a gender gap in housing returns in Figure 8 that decays toward zero with holding length. This pattern is more consistent with a gap in returns that arises from gender differences in execution prices at the points of purchase and sale (as discussed previously, differences in execution prices of δ and a holding length of t predicts that the gender gap equals $2\delta/t$, which approximately matches the shape of the decay in the data).

E.2 Demographic Variation using the American Housing Survey

We use detailed demographic data from the American Housing Survey to explore whether demographic factors correlated with gender may help explain the gender gap in housing returns. Women outlive men on average, and older individuals may earn worse returns on housing.[14] Women may also have more children, lower education, and lower income, factors that may impact housing returns.

Table 9 presents regression analysis using the American Housing Survey (AHS) over the 2001-2013 sample period. In column 3, we are able to replicate our baseline results of a gender gap in housing

[14] After the death of a spouse, widows or widowers may sell homes at a discount for a variety of reasons. Such cases are excluded from our returns analysis because we require that the homeowner be single at the time of home purchase as well as the time of home sale in order to be classified as single male or single female. We also restrict our sample to arms-length transactions, which excludes transfers to family members.

returns using this alternate data source. This shows that our baseline results are unlikely to be due to measurement error in the identification of gender groups, since these fields are self-reported rather than imputed in the AHS. Interestingly, the gender gap in realized returns in column 3 exceeds the gender gap in self-reported estimated returns in column 2, which uses self-reported estimates of the current market value of the property relative to the purchase price to calculate returns.[15] This comparison suggests that women underestimate their investment return disadvantage in housing markets.[16]

In column 4, we show that number of children, ethnicity, and income all significantly predict realized unlevered housing returns. A greater number of other adults in the home (which may represent care given to elderly relatives) and lower education also predict lower returns, although the coefficients are not statistically significant. However, the gender gap remains large at 1 percentage point per year after controlling for all of these demographic variables.

In column 5, we explore how having children may impact the housing returns of single men, women, and couples.[17] The omitted category in this regression is single male without any children. For all groups, we find that having children is associated with lower returns on housing (having at least one child lowers the annualized return for single men and women by approximately 1 percentage point, and the returns of couples by 0.5 percentage points). However, single women without any children still earn 1 percentage points lower returns relative to single men without children, after controlling for education, income, and ethnicity.

E.3 Real Estate Agents

The gender gap in housing returns could also arise because men employ better or more effective real estate agents. Housing transactions are typically intermediated by agents who advise their clients on the choice of listing prices, offers, and counter offers in real estate negotiations. Using merged MLS data, we observe unique codes for each listing agent representing the seller.[18] In Table 10, we

[15]The estimated unlevered return and the real unlevered return are calculated in slightly different samples. For the real unlevered return, we follow Harding et al. (2007) and require that a transaction happens between two adjacent surveys, and that both the purchase and sale of the property are arms-length transactions. Since not all houses transact in this way and we have a shorter sample period, we are left with a much smaller sample. For the estimated return, we observe this value in almost all cases, and thus have a much larger sample to draw on.

[16]Bordalo et al. (2019) show that women exhibit lower performance when they face negative setting-specific stereotypes. For example, women may perform poorly in automobile negotiations because the automobiles setting is stereotypically dominated by men. It is less obvious that housing is a male-dominated setting, given that women invest a significantly greater percentage of their wealth into housing than men. The results in column 2 of Table 9 also suggest that women may be unaware of the full extent to which they underperform relative to men when engaging in housing transactions.

[17]This test is motivated by evidence in Kleven et al. (2019) showing that women's earnings and labor market outcomes are disproportionately affected by childbirth.

[18]We unfortunately do not observe codes for buyer agents, so our analysis is limited to tests in which we control for sellers' listing agent fixed effects.

re-estimate the gender gap in returns, discounts, list prices, and transaction prices, after controlling for listing agent fixed effects. The fixed effects control for the fixed impact of each agent on prices, discounts, etc. across all transactions in our data. In other words, we can control for the time-invariant impact of agent gender, talent, effort, etc., on housing transactions. We continue to find similarly sized gender gaps in returns, discounts, list prices, and transaction prices after controlling for agent fixed effects. Thus, the gender gap cannot be explained by the possibility that single women match with worse seller agents.

An important caveat to this analysis is that we can only rule out the influence of agent effects that are *fixed* across clients. The impact of agents may depend on the interaction between agents and the gender of their clients. For example, the same real estate agent may give different advice to male and female clients. Men and women may also differ in whether they follow the (good or bad) advice of their agents. Our observational data unfortunately does not allow us to examine the details of discussions between clients and agents, so we leave further study of the role of agents to future work.

E.4 Variation by Market Tightness

Finally, we explore how the gender gap in housing markets varies with market tightness in Table 11. To measure market tightness, we use our listings data to construct a county-by-month measure of the number of sales in a given month scaled by the total number of listings. As markets tighten, it becomes easier to sell a given property and the housing market becomes more liquid. We replicate our analyses from Tables 2, 4 and 6, interacting the gender group indicators with market tightness. We find that the gender gaps in returns, prices, and discounts all shrink substantially with market tightness. The gender gap in returns would approximately approach zero when market tightness is around 1. There are similar offsetting interaction coefficients for transaction prices and discounts.

This pattern suggests that bilateral negotiation may be an important driver of the *average* gender gap in housing returns. As markets tighten, bilateral negotiation is often replaced with quasi-auctions, in which multiple interested buyers simultaneously bid for homes. The reduction in the gender gap with market tightness is also inconsistent with an explanation in which men buy riskier properties or invest more in home maintenance and upgrades. Differences in maintenance or risk should lead to real differences in home value appreciation that would, if anything, be more efficiently priced in liquid markets. Similarly, variation by market tightness is inconsistent with the idea that women earn lower housing returns because they derive greater utility from housing and are therefore willing to pay more for the same house. If so, women should also submit higher bids for homes, leading to lower returns in tight housing markets.

IV. Discussion and Implications

Overall, we believe the gender gap in housing returns arises because of (1) differences in market timing and (2) differences in negotiated execution prices, conditional on transacting in the same zip code and time period. What happens in between purchase and sale, e.g., potential mechanisms in which men invest more in upgrades and maintenance or buying riskier properties with naturally higher returns, appear to be less important explanations.

We also believe that the root causes of gender differences in market timing and negotiated outcomes may be complex and are deserving of further research. For example, women may engage in less advantageous market timing because they face timing constraints, rather than misunderstanding market conditions. Women may have worse negotiated outcomes for a variety of reasons, including expectations regarding how women should behave (e.g., Egan et al. (2017)), preferences and beliefs regarding negotiation strategy which may be context dependent (e.g., Bordalo et al. (2019)), or willingness and ability to search. Unobserved demographic differences across gender may also matter, although we do find that the gender gap remains after controlling for income, education, and other observed demographic characteristics.

Regardless of the exact channel, our analysis shows that gender differences in housing returns are economically large. In our main analysis, we presented percentage point gender gaps in housing returns, transaction prices, and negotiated discounts. We can also think about the gender gap in terms of dollars lost per year for the median single female homeowner.

We assume that a woman holds a home worth $200,000 (the median home value from the 2013-2017 American Community Survey five-year estimates) for the median holding period of five years, and experiences a 2% execution price disadvantage at the points of purchase and sale. These assumptions imply that women lose $1,600 each year as a homeowner relative to single men. We can compare this magnitude to the dollar loss from the gender pay gap. Blau and Kahn (2017) estimate a gender pay gap of 8% among similar men and women in terms of education, occupation, and other observables. Based on the median wage for single men of $35,000 per year, the gender pay gap implies that women with similar observables earn $2,800 less per year. Thus, the gender gap in housing markets is approximately half as large as the gender pay gap, which has been the subject of numerous academic studies and policy debates.

Because housing wealth is the dominant form of savings for most households, our findings also offer insight into variation in wealth accumulation (e.g., Ruel and Hauser (2013)). Appendix Figures A1 and A2 report equity in housing as a percentage of total net worth and the gender gap in wealth accumulation at retirement, respectively. We estimate the impact of gender differences in the housing market on wealth accumulation under various assumptions. The details of this calculation are reported in the Appendix. We estimate that the gender gap in housing returns can explain approximately 30% of the overall gender gap in wealth accumulation at retirement.

It is important to note that these calculations of the gender gap in dollars use estimated differences in execution prices; these calculations do not require any assumptions regarding the degree of leverage taken by female and male homeowners. Even if all homeowners took on zero debt, we would still conclude that the gender gap in housing markets is an important contributor to the gender gap in wealth accumulation.

Moreover, these calculations may represent conservative estimates of the gender gap in dollars. The calculations use average differences in execution prices among men and women who transact in the same zip code and year-month. The calculations do not take into account the significant gender differences in market timing. The gender gap in dollars would be larger if we allow women to be more likely to buy when aggregate prices are high and sell when they are low.

Finally, we emphasize that women may derive greater utility from home ownership despite earning lower financial returns. Therefore, we do not draw conclusions regarding gender gaps in welfare. However, we also believe that a simple explanation in which women value housing more than men is unable to fully match the empirical patterns documented in this paper. In particular, it does not explain why women buy homes for more, but also sell for less. It does not explain why women negotiate worse discounts when facing male counterparties. It also does not explain why the gender gap approaches zero in tight housing markets. If women attach higher valuations to homes, they should submit higher bids when multiple potential buyers submit competing bids, leading to lower returns for women in tight housing markets.

V. Conclusion

We uncover a large gender gap in the returns to housing investment in recent decades in the US. This gender gap is an important contributor to gender differences in wealth accumulation, given that housing wealth represents the dominant form of savings for most US households. Using detailed data on housing transactions across the US, we find that single men earn 1.5 percentage points higher unlevered returns per year on housing investment relative to single women. However, the real return earned by most households is a levered return. The gender gap in raw returns grows significantly larger after adjusting for mortgage borrowing. Assuming the modal 30-year fixed rate mortgage with a 20% downpayment, men earn approximately 7.9 percentage points higher levered returns per year relative to women. Using data on repeat sales, we show that women buy the same property for approximately 2% more and sell for 2% less, after controlling for market timing. The gender gap in housing returns arises because of gender differences in the location and timing of transactions, choice of initial list price, and negotiated discount relative to the list price. While the gender gap varies with demographic characteristics, it remains substantial in regions with high average education, income, and house price levels. It also has not displayed a secular decline over time.

In addition to the gender gap resulting from market timing, we find that women experience significantly more negative negotiation-related outcomes in housing markets. Women negotiate smaller discounts relative to the list price when buying and offer larger discounts when selling. Holding the property fixed, the highest transaction prices occur in cases with a male seller and female buyer, and the lowest transaction prices occur in cases with a female seller and male buyer. However, these results do not necessarily imply that women make mistakes in housing negotiations. In particular, recent research by Exley et al. (2016) suggests that women can sometimes experience even more negative outcomes by "leaning-in" and negotiating more aggressively. Given the importance of housing investment for household savings, we believe that further exploration of factors that determine the gender gap in housing is an important direction for future research.

References

Aigner, Dennis J, "Regression with a binary independent variable subject to errors of observation," *Journal of Econometrics*, 1973, *1* (1), 49–59.

Andersen, Steffen, Julie Marx, Kasper Meisner Nielsen, and Lise Vesterlund, "Gender differences in negotiation: Evidence from real estate transactions," *Working Paper*, 2018.

Ayres, Ian, "Fair driving: Gender and race discrimination in retail car negotiations," *Harvard Law Review*, 1990, *104*, 817.

_ **and Peter Siegelman**, "Race and gender discrimination in bargaining for a new car," *American Economic Review*, 1995, pp. 304–321.

Babcock, Linda and Sara Laschever, *Women don't ask: Negotiation and the gender divide*, Princeton University Press, 2009.

Bajtelsmit, Vickie L and Jack L VanDerhei, "Risk aversion and pension investment choices," *Positioning Pensions for the Twenty-First Century*, 1997, *45*, 66.

Barber, Brad M and Terrance Odean, "Boys will be boys: Gender, overconfidence, and common stock investment," *The Quarterly Journal of Economics*, 2001, *116* (1), 261–292.

Bayer, Patrick, Marcus D Casey, Fernando Ferreira, and Robert McMillan, "Estimating racial price differentials in the housing market," *Journal of Urban Economics*, 2017.

Bertrand, Marianne, "New perspectives on gender," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1543–1590.

Blau, Francine D and Lawrence M Kahn, "The gender wage gap: Extent, trends, and explanations," *Journal of Economic Literature*, 2017, *55* (3), 789–865.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, "Beliefs about gender," *American Economic Review*, 2019, *109* (3), 739–73.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly Journal of Economics*, 2015, *131* (2), 633–686.

Castillo, Marco, Ragan Petrie, Maximo Torero, and Lise Vesterlund, "Gender differences in bargaining outcomes: A field experiment on discrimination," *Journal of Public Economics*, 2013, *99*, 35–48.

Chari, Anusha and Paul Goldsmith-Pinkham, "Gender representation in economics across topics and time: Evidence from the NBER summer institute," Technical Report, National Bureau of Economic Research 2017.

Egan, Mark L, Gregor Matvos, and Amit Seru, "When Harry fired Sally: The double standard in punishing misconduct," Technical Report, National Bureau of Economic Research 2017.

Exley, Christine Linman, Muriel Niederle, and Lise Vesterlund, "Knowing when to ask: The cost of leaning-in," *Working Paper*, 2016.

Genesove, David and Christopher Mayer, "Loss aversion and seller behavior: Evidence from the housing market," *The Quarterly Journal of Economics*, 2001, *116* (4), 1233–1260.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini, "Performance in competitive environments: Gender differences," *The Quarterly Journal of Economics*, 2003, *118* (3), 1049–1074.

Goldin, Claudia, "A grand gender convergence: Its last chapter," *American Economic Review*, 2014, *104* (4), 1091–1119.

Harding, John P, Stuart S Rosenthal, and CF Sirmans, "Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model," *Journal of Urban Economics*, 2007, *61* (2), 193–217.

_ , _ , and Clemon F Sirmans, "Estimating bargaining power in the market for existing homes," *Review of Economics and Statistics*, 2003, *85* (1), 178–188.

Hernandez-Arenaz, Iñigo and Nagore Iriberri, "Women ask for less (only from men): Evidence from bargaining in the field," *Journal of Economic Behavior & Organization*, 2018, *152*, 192–214.

Hinz, Richard P, David D McCarthy, and John A Turner, "Are women conservative investors? Gender differences in participant-directed pension investments," *Positioning Pensions for the Twenty-First Century*, 1997, *91*, 103.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019, *11* (4), 181–209.

Leibbrandt, Andreas and John A List, "Do women avoid salary negotiations? Evidence from a large-scale natural field experiment," *Management Science*, 2014, *61* (9), 2016–2024.

List, John A, "The nature and extent of discrimination in the marketplace: Evidence from the field," *Quarterly Journal of Economics*, 2004, *119* (1), 49–89.

Morton, Fiona Scott, Florian Zettelmeyer, and Jorge Silva-Risso, "Consumer information and discrimination: Does the internet affect the pricing of new cars to women and minorities?," *Quantitative Marketing and Economics*, 2003, *1* (1), 65–92.

Niederle, Muriel and Lise Vesterlund, "Do women shy away from competition? Do men compete too much?," *The Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales, "Taste for competition and the gender gap among young business professionals," Technical Report, National Bureau of Economic Research 2015.

Ruel, Erin and Robert M Hauser, "Explaining the gender wealth gap," *Demography*, 2013, *50* (4), 1155–1176.

Sakong, Jung, "Cyclical Housing Transactions and Wealth Inequality," *Working Paper*, 2019.

Sapienza, Paola, Luigi Zingales, and Dario Maestripieri, "Gender differences in financial risk aversion and career choices are affected by testosterone," *Proceedings of the National Academy of Sciences*, 2009, *106* (36), 15268–15273.

Shefrin, Hersh and Meir Statman, "The disposition to sell winners too early and ride losers too long: Theory and evidence," *Journal of Finance*, 1985, *40* (3), 777–790.

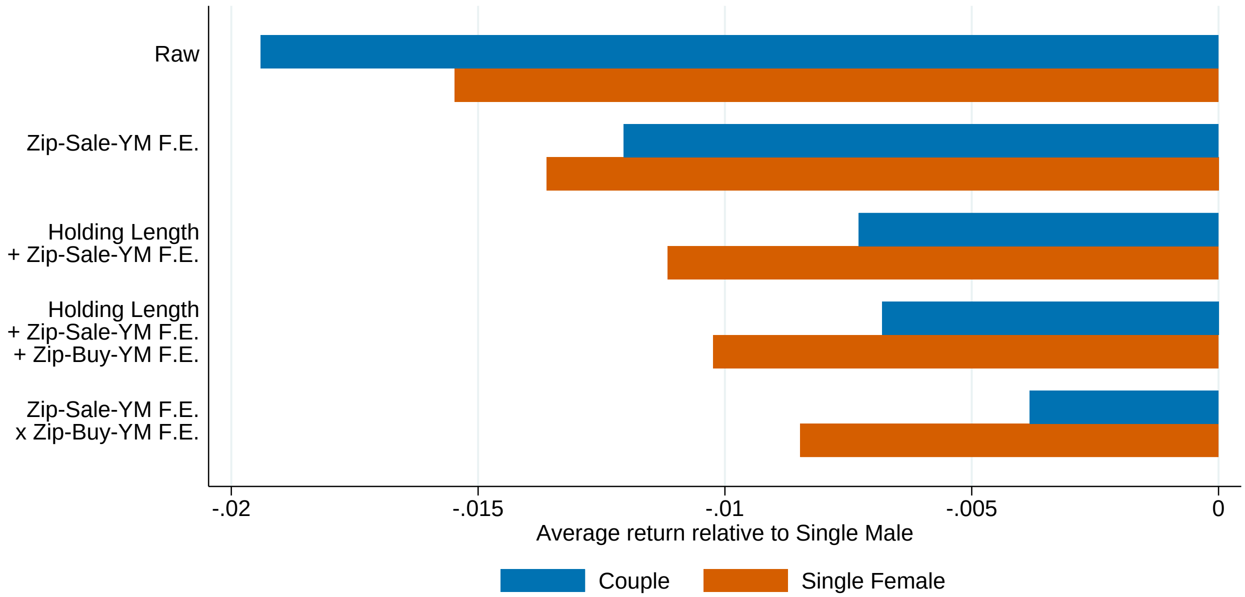
Solnick, Sara J, "Gender differences in the ultimatum game," *Economic Inquiry*, 2001, *39* (2), 189–200.

Sunden, Annika E and Brian J Surette, "Gender differences in the allocation of assets in retirement savings plans," *American Economic Review*, 1998, *88* (2), 207–211.

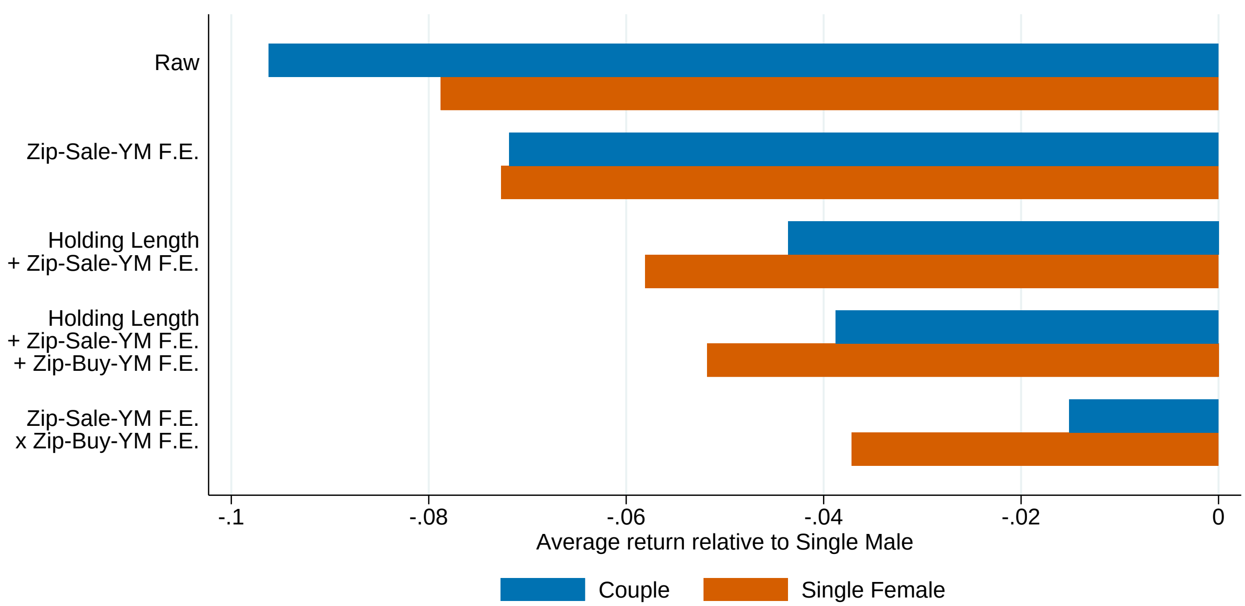
Tang, Cong, Keith Ross, Nitesh Saxena, and Ruichuan Chen, "What's in a name: A study of names, gender inference, and gender behavior in facebook," in "International Conference on Database Systems for Advanced Applications" Springer 2011, pp. 344–356.

Figure 1: Annualized returns by gender group across different specifications

Panel A: Unlevered Returns



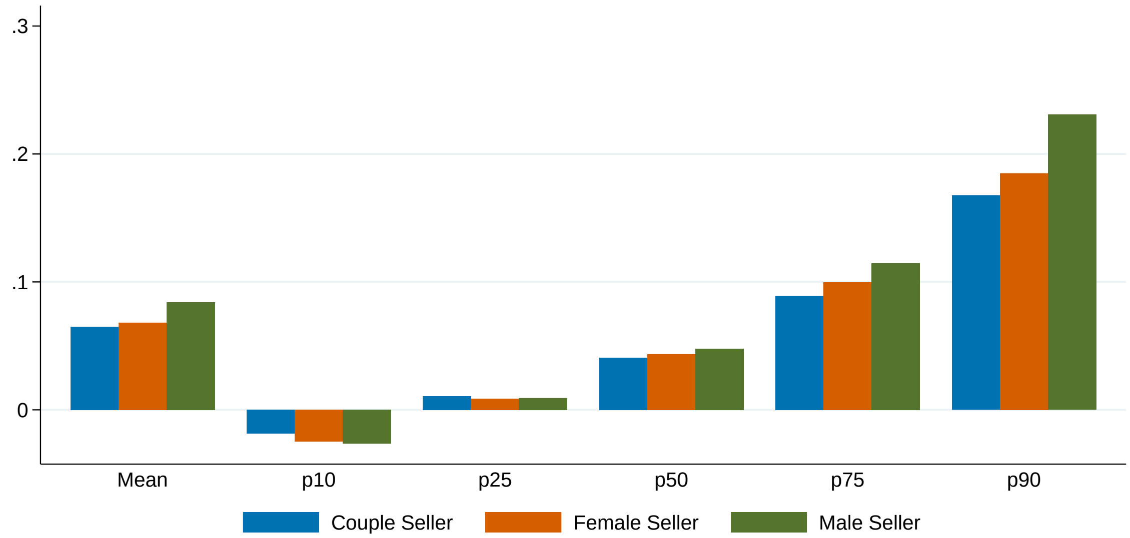
Panel B: Levered (80%) Returns



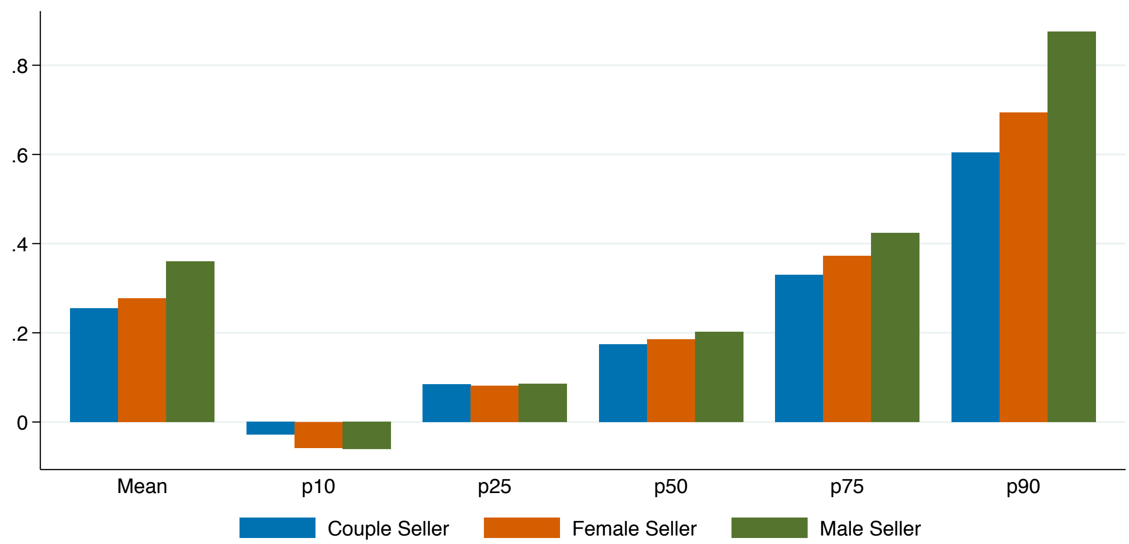
Note: This figure plots the difference in annualized returns between single women and couples, relative to single men, conditional on the set of control variables listed on the y-axis. Panels A and B correspond to the coefficient estimates in Panels A and B of Table 2, respectively. The levered returns are calculated following the formula in Section II.D, assuming an initial LTV of 80%. See Section II.B for more details on the definition of gender groups (single female, single male, and couple).

Figure 2: Distribution of annualized returns by gender group

Panel A: Unlevered Returns

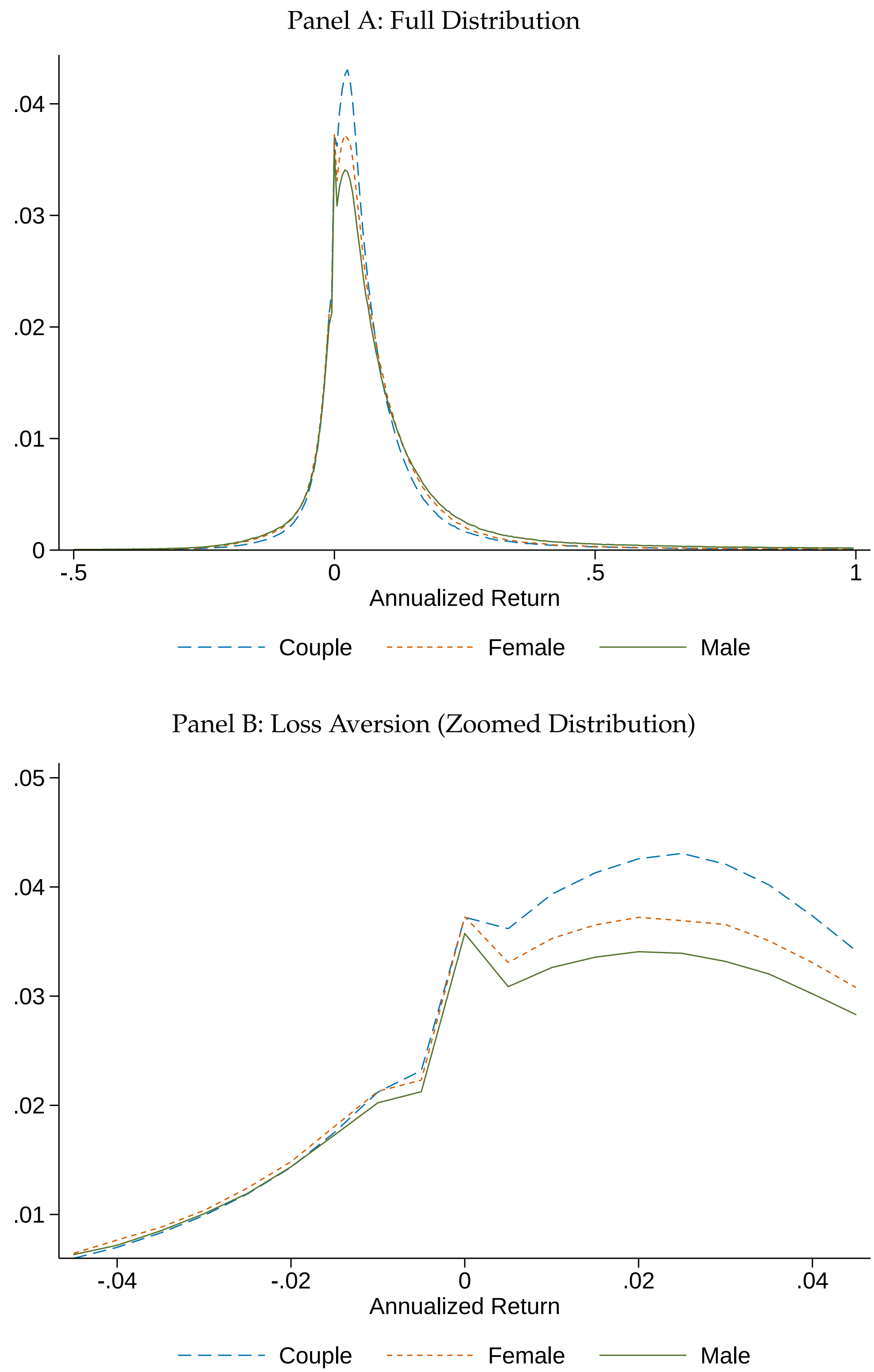


Panel B: Levered (80%) Returns



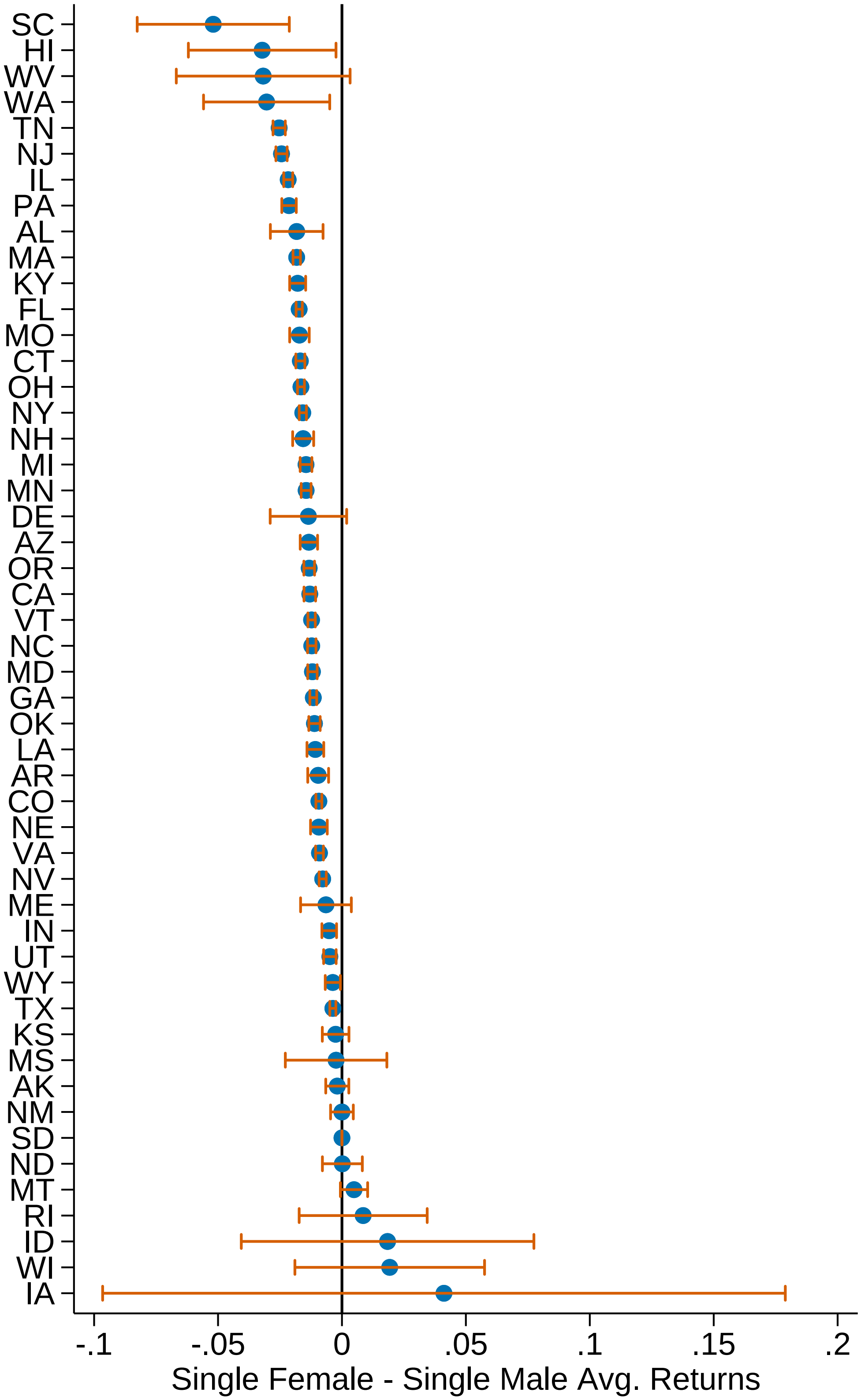
Note: This figure plots summary statistics for the annualized returns for housing transactions by three gender groups: couples, single women, and single men. The levered returns are calculated following the formula is Section D, assuming an initial LTV of 80%.

Figure 3: Density of unlevered annualized returns by gender group



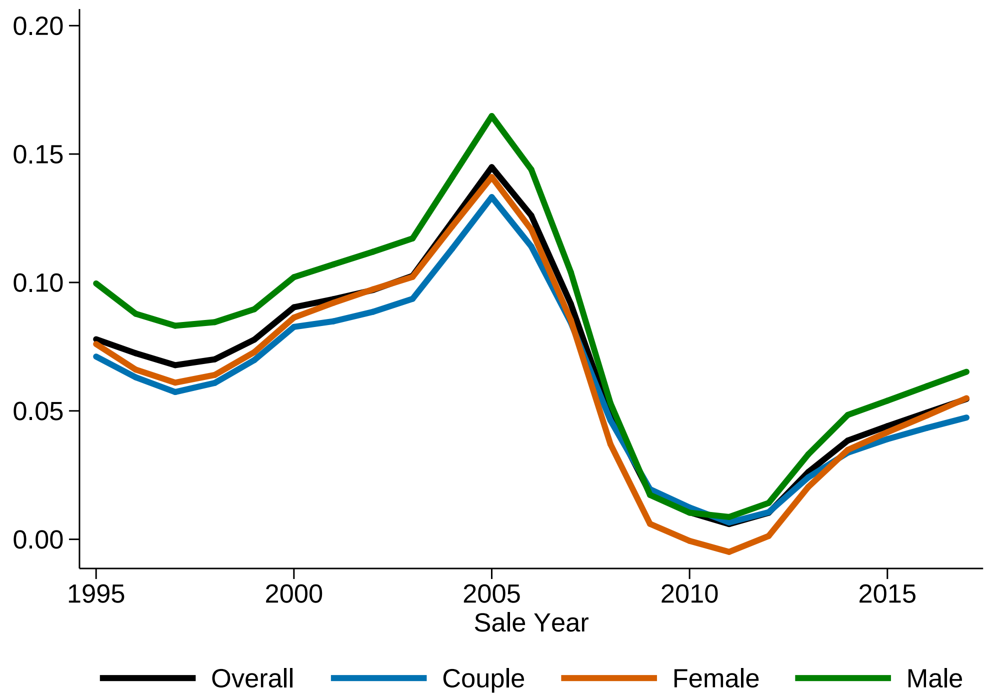
Note: This figure plots the density of the annualized unlevered returns for housing transactions by three gender groups: couples, single women, and single men. Returns are truncated at -50% and +100% in Panel A. Returns are truncated at -4% and +5% in Panel B.

Figure 4: Gender gap in unlevered returns across states



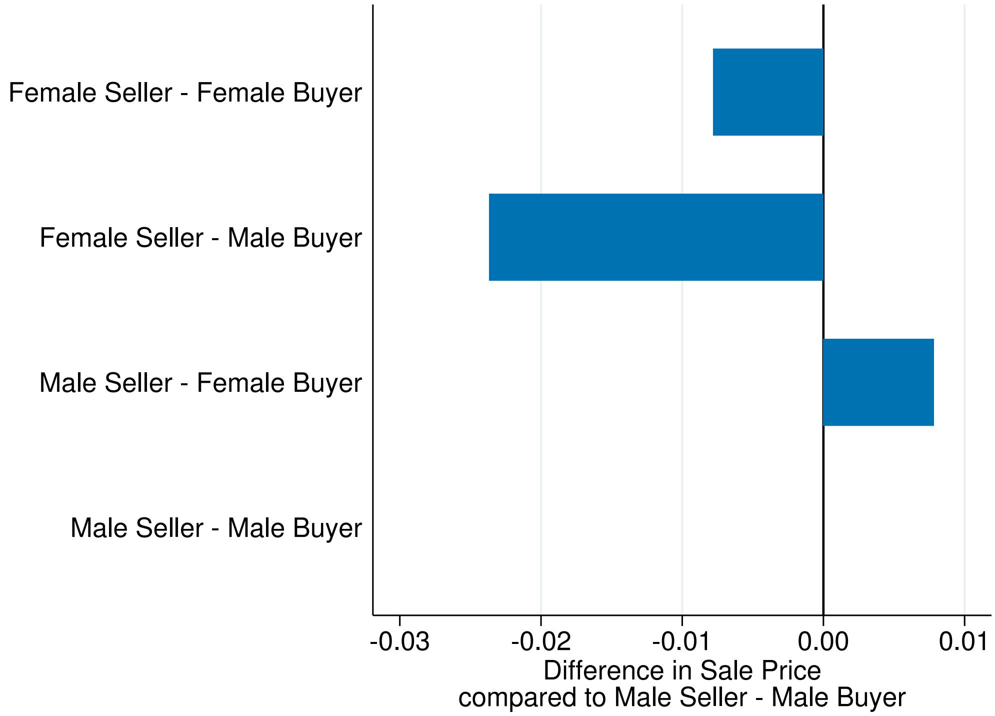
Note: This figure plots the average difference in unlevered annualized returns between single men and women across states. The points represent the estimated difference in realized returns, while the bars represent the 95% confidence interval. Standard errors are clustered at the zip-code level.

Figure 5: Average unlevered returns over time by gender group



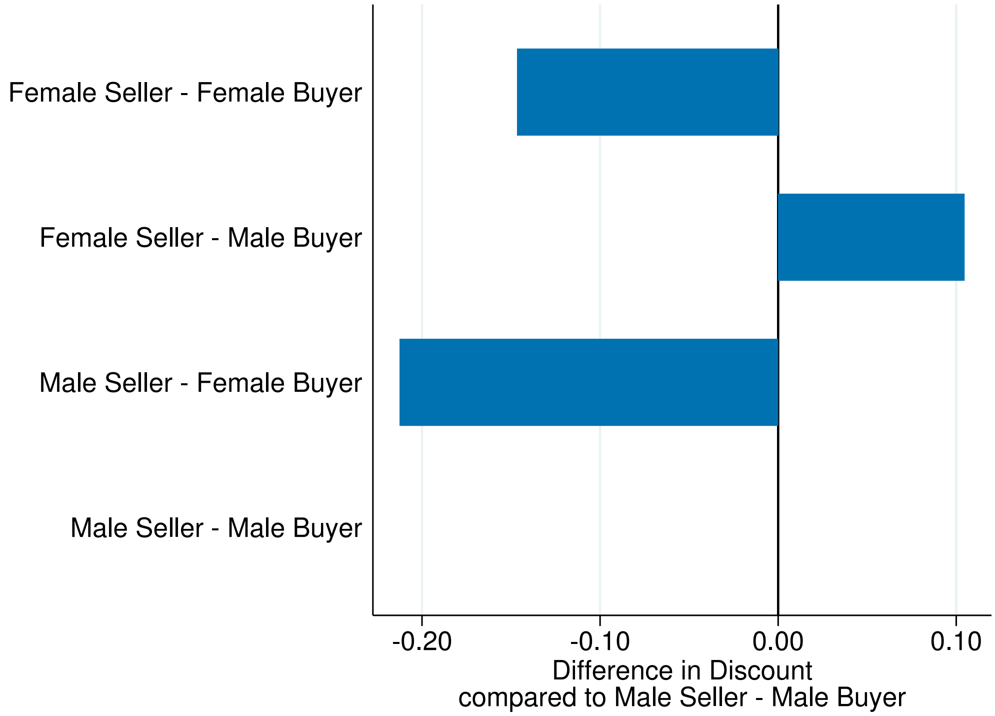
Note: This figure plots the average unlevered annualized return for couples, single women, and single men by sale year. As our sample begins in 1991, we begin this figure in 1995 to allow for sufficient data to avoid truncation.

Figure 6: Transaction price by seller-buyer gender pairing



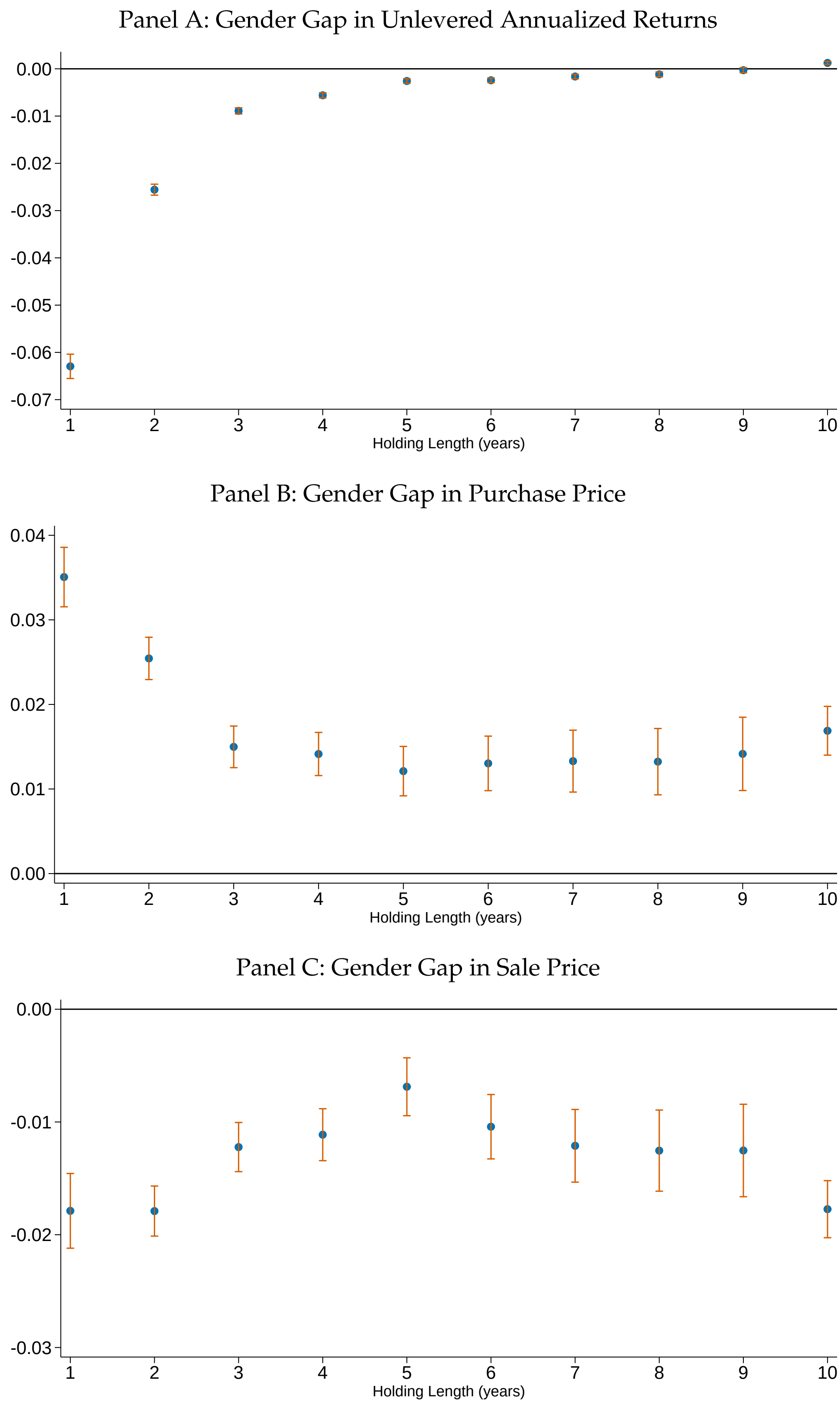
Note: This figure plots the average difference in log transaction prices for each possible seller-buyer gender pair, relative to transactions involving single male sellers and single male buyers. These estimates come from a regression of the form in Table 4 column 4, but allowing for the buyer and seller gender group indicators to interact. We plot only the coefficients representing single male or female buyers and sellers, with male seller-male buyer as the omitted base coefficient.

Figure 7: Discounts by seller-buyer gender pairing



Note: This figure plots the average difference in discounts for each possible seller-buyer gender pair, relative to discounts for transactions involving single male sellers and single male buyers. We measure discounts as (list price - transaction price)/list price $\times$ 100, so a larger discount contributes to a higher return on housing investment for buyers and a lower return for sellers. These estimates come from a regression of the form in Table 6 column 4, but allowing for the buyer and seller gender group indicators to interact. We plot only the coefficients representing single male or female buyers and sellers, with male seller-male buyer as the omitted base category.

Figure 8: Differences between single men and women by holding period



Note: This figure plots the average difference in outcomes between single men and women by holding length for the property. We exclude holding periods longer than 11 years. Panel A plots the female minus male gender gap in returns controlling for zip-sale-year-month fixed effects. Panels B and C plot the female minus male gender gap in log purchase and sale prices, respectively, controlling for property fixed effects and zip-year-month fixed effects. Vertical bars represent 95% confidence intervals.

Table 1: Summary statistics

	Buyer Gender Group				
	Single Male	Single Female	Couple	Other	Overall
Panel A: Full Sample					
Log(Sale Price)	11.9482	11.8821	12.1105	12.1842	12.0704
Sample Size	11,561,755	7,614,231	13,424,894	20,282,986	52,883,866
Panel B: Listing Sample					
Log(Sale Price)	12.0567	11.9758	12.2847	12.0711	12.1102
Log(List Price)	12.0739	11.9918	12.3067	12.0958	12.1308
Purchase Discount (p.p.)	2.7826	2.5422	2.7200	3.1715	2.8415
Log(Days on Market)	3.7355	3.7024	3.7856	3.7377	3.7444
Sample Size	4,751,834	3,240,324	5,236,683	5,711,909	18,940,750

	Seller Gender Group				
	Single Male	Single Female	Couple	Other	Overall
Panel C: Full Sample					
Log(Sale Price)	12.0264	11.9572	12.1414	12.0748	12.0704
Sample Size	6,092,999	5,110,755	9,960,567	31,719,545	52,883,866
Panel D: Listing Sample					
Log(Sale Price)	12.1459	12.0678	12.2737	12.0332	12.1102
Log(List Price)	12.1651	12.0914	12.2919	12.0546	12.1308
Sale Discount (p.p.)	2.7882	2.9663	2.5274	2.9744	2.8415
Log(Days on Market)	3.7176	3.6967	3.7018	3.7840	3.7444
Sample Size	2,484,028	2,423,657	4,554,616	9,478,449	18,940,750
Panel E: Returns Sample					
Log(Sale Price)	12.1699	12.0844	12.3354	-	12.2292
Annualized Unlevered Returns	0.0840	0.0685	0.0646	-	0.0712
Annualized Levered (80%) Returns	0.3024	0.2237	0.2062	-	0.2389
Holding Length (Years)	5.3427	5.7665	5.9889	-	5.7456
Log(Purchase Price)	11.9294	11.8477	12.0807	-	11.9828
Sample Size	2,666,894	2,021,915	4,244,322	-	8,933,131

Note: This table reports summary statistics for the samples used in the analysis, split by four gender groups (single male, single female, couple, and other), and also pooled. Panels A and B are split by buyer gender group, and Panels C, D, and E are split by seller gender group. Each cell is the overall mean within the relevant the sample group. Panels A and C represent the full sample of all arms-length sales transactions reported in the data. Panels B and D represent the sample of sales transactions successfully matched to listings data. Panel E represents the sample of sales transactions where we successfully match the identified gender and family structure in the sale transaction with that in the previous purchase transaction, and also match the names of the sellers with the names of the buyers in the previous transaction. See Section II.B for more details on how we identify gender groups (single female, single male, couple, and other). See Section II.A for more details on how we match the data.

Table 2: Housing returns: market timing

Panel A: Unlevered Returns

	Unlevered Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.015***	-0.014***	-0.011***	-0.010***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Couple	-0.019***	-0.012***	-0.007***	-0.007***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Length			-0.006***	-0.001	
			(0.000)	(0.001)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.004	0.352	0.378	0.534	0.592
Observations	8,933,131	8,933,131	8,933,131	8,933,131	8,933,131

Panel B: Levered (80%) Returns

	Levered (80%) Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.079***	-0.073***	-0.058***	-0.052***	-0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Couple	-0.096***	-0.072***	-0.044***	-0.039***	-0.015***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Holding Length			-0.037***	-0.026***	
			(0.000)	(0.006)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.004	0.292	0.326	0.481	0.628
Observations	8,933,131	8,933,131	8,933,131	8,933,131	8,933,131

Note: This table shows the difference in annualized returns for single women and couples relative to single men (the omitted category). The levered returns are calculated following the formula in Section II.D, assuming an initial LTV of 80%. Columns 2 through 5 introduce additional control variables for the location and timing of housing transactions. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3: Heterogeneity by zip-level demographics: quartile averages

Male - Female Unlevered Ann Return	Quartile 1	Quartile 2	Quartile 3	Quartile 4
	(1)	(2)	(3)	(4)
Frac Black	0.0137	0.0140	0.0143	0.0194
Frac HS Education or Less	0.0089	0.0146	0.0196	0.0240
Frac 60+	0.0118	0.0154	0.0182	0.0184
Frac Single Female	0.0115	0.0135	0.0150	0.0222
Median Family Income	0.0234	0.0177	0.0145	0.0109
House Price	0.0191	0.0165	0.0139	0.0118

Note: This table presents the average male minus female gender gap in unlevered annualized housing returns across quartiles of various zip-level demographic characteristics from the 2010 American Community Survey. The gender gap within each zip code is measured as the simple average across all transactions. Zip-codes within each quartile are equally weighted to compute the quartile average.

Table 4: Transaction price

	Log(Purchase Price)		Log(Sale Price)	
	(1)	(2)	(3)	(4)
Single Female	0.013***	0.018***	-0.029***	-0.026***
	(0.001)	(0.001)	(0.001)	(0.001)
Couple	0.029***	0.032***	0.014***	0.016***
	(0.002)	(0.001)	(0.001)	(0.001)
Other	0.091***	0.022***	-0.051***	-0.051***
	(0.005)	(0.003)	(0.002)	(0.001)
Property FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	No	Yes	No
Zip-Year-Month FE	No	Yes	No	Yes
R-squared	0.794	0.886	0.793	0.886
Observations	52,883,866	52,883,866	52,883,866	52,883,866

Note: This table examines variation in transaction prices across gender groups. Gender groups in columns 1 and 2 refer to buyers and gender groups in columns 3 and 4 refer to sellers. We use repeat sales data that allows us to control for property fixed effects. We also control for zip-year-month fixed effects to account for time trends within a zip code. Each observation is a transaction. To better estimate property fixed effects, we do not restrict the sample to buyers or sellers with identified genders and matched names across sales and initial purchase. All observations corresponding to non-identified single women, single men, and couples are included and coded as the "other" category. Single males are the omitted category. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 5: List price

	Log(Purchase List Price)		Log(Sale List Price)	
	(1)	(2)	(3)	(4)
Single Female	0.021***	0.022***	-0.021***	-0.018***
	(0.001)	(0.001)	(0.001)	(0.000)
Couple	0.023***	0.015***	-0.008***	-0.004***
	(0.001)	(0.000)	(0.001)	(0.001)
Other	-0.022***	-0.027***	-0.099***	-0.080***
	(0.002)	(0.001)	(0.002)	(0.002)
Property FE	Yes	Yes	Yes	Yes
Year-Month FE	Yes	No	Yes	No
Zip-Year-Month FE	No	Yes	No	Yes
R-squared	0.916	0.961	0.918	0.962
Observations	10,255,902	10,255,902	10,255,902	10,255,902

Note: This table examines variation in list prices across gender groups. Gender groups in columns 1 and 2 refer to buyers and gender groups in columns 3 and 4 refer to sellers. We use repeat sales data that allows us to control for property fixed effects. We also control for zip-year-month fixed effects to account for time trends within a zip code. Each observation is a listing matched to a sales transaction and the sample is restricted to properties for which we observe at least two transactions. To better estimate property fixed effects, we do not restrict the sample to buyers or sellers with identified genders and matched names across sales and initial purchase. All observations corresponding to non-identified single women, single men, and couples are included and coded as the "other" category. Single males are the omitted category. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6: Discount relative to the list price

	Purchase Discount		Sale Discount	
	(1)	(2)	(3)	(4)
Single Female	-0.276***	-0.256***	0.167***	0.092***
	(0.007)	(0.005)	(0.007)	(0.005)
Couple	-0.117***	-0.084***	-0.280***	-0.228***
	(0.012)	(0.005)	(0.012)	(0.005)
Zip-Year-Month FE	No	Yes	No	Yes
R-squared	0.001	0.222	0.002	0.257
Observations	13,690,139	13,690,139	9,658,833	9,658,833

Note: This table examines how negotiated transaction discounts vary by gender group. Gender groups in columns 1 and 2 refer to buyers and gender groups in columns 3 and 4 refer to sellers. We measure purchase and sale discounts as (list price - transaction price)/list price $\times$ 100, so a larger purchase discount contributes to a higher return on housing investment and a larger sale discount contributes to a lower return on housing investment. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 7: Days on market

	Sale Log(Days on Mkt)		Log(Sale Price)
	(1)	(2)	(3)
Single Female	-0.027***	-0.009***	-0.024***
	(0.002)	(0.002)	(0.001)
Couple	-0.048***	-0.049***	-0.006***
	(0.002)	(0.002)	(0.001)
Other	-0.002	-0.049***	-0.087***
	(0.003)	(0.003)	(0.002)
Sale Log(Days on Mkt)			0.002***
			(0.000)
Property FE	No	Yes	Yes
Zip-Year-Month FE	Yes	Yes	Yes
R-squared	0.283	0.659	0.937
Observations	10,412,065	10,412,065	10,412,065

Note: Columns 1 and 2 examine how days on market, measured as the number of days between the earliest available listing associated with a transaction and the sale date, varies with seller gender group. Column 3 examines variation in sale prices by gender group after controlling for the number of days on market. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 8: Controlling for timing and property characteristics

	Unlevered Ann Return		
	(1)	(2)	(3)
Single Female	-0.012***	-0.007***	-0.006***
	(0.000)	(0.000)	(0.000)
Couple	-0.015***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)
Log(Age of Unit)			0.006***
			(0.000)
Garage			-0.006***
			(0.000)
Pool			-0.002***
			(0.000)
Cooling			-0.002***
			(0.000)
Fireplace			-0.004***
			(0.000)
Basement			0.001**
			(0.000)
Waterfront			0.003***
			(0.001)
Bathrooms			0.001
			(0.001)
Log(Sq Ft)			0.003***
			(0.001)
Bedrooms			0.001***
			(0.000)
Upgraded			0.008***
			(0.000)
New Construction			-0.004***
			(0.000)
Property Type FE	No	No	Yes
Zip-SaleYM FE	No	Yes	Yes
Zip-BuyYM FE	No	Yes	Yes
SaleYM FE x BuyYM FE	No	Yes	Yes
R-squared	0.003	0.646	0.649
Observations	3,008,113	3,008,113	3,008,113

Note: This table examines how the gender gap in housing returns varies with additional control variables for market timing and property and listing agent characteristics. The sample is restricted to observations for which we have matched listings data on property characteristics. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 9: Maintenance and housing returns in the American Housing Survey

	Maint/price	Est Unlevered Return	Real Unlevered Return		
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.000233 (0.000159)	-0.00595** (0.00241)	-0.0120** (0.00509)	-0.0102** (0.00511)	
Couple	0.000372*** (0.000133)	-0.00784*** (0.00213)	-0.00634 (0.00506)	-0.00377 (0.00479)	
Age of House	0.0000975*** (0.00000210)	0.000367*** (0.0000319)	0.000226*** (0.0000477)	0.000226*** (0.0000476)	0.000224*** (0.0000476)
Age of Householder	0.0000469*** (0.00000322)	-0.000617*** (0.0000529)	-0.0000800 (0.0000734)	-0.000140 (0.0000935)	-0.000146 (0.0000974)
Single Female with Child					-0.0193** (0.00776)
Single Female w/o Child					-0.0103* (0.00576)
Male with Child					-0.0124 (0.00859)
Couple with Child					-0.0105 (0.00643)
Couple w/o Child					-0.00565 (0.00567)
Number of Children				-0.00293** (0.00123)	
Number of Adults				-0.00231 (0.00180)	-0.00197 (0.00178)
Some College				0.00145 (0.00368)	0.00156 (0.00366)
College Degree				0.00101 (0.00320)	0.000987 (0.00321)
Graduate Degree				-0.000806 (0.00365)	-0.000769 (0.00365)
Black				-0.00689 (0.00627)	-0.00737 (0.00631)
American Indian				-0.0114 (0.00941)	-0.0114 (0.00971)
Asian				-0.0117* (0.00695)	-0.0120* (0.00698)
Other Race				-0.00701 (0.0143)	-0.00782 (0.0146)
Log Family Income				0.00239* (0.00128)	0.00246* (0.00128)
MSA x Survey or Sale Year FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.063	0.023	0.294	0.298	0.298
Observations	124,505	135,669	3,716	3,678	3,678

Note: This table uses data from the American Housing Survey. In column 1, the dependent variable is reported annual maintenance scaled by home purchase price. In column 2, estimated unlevered return is the annualized unlevered return, calculated using the homeowner's self reported estimate of current home value relative to purchase price. In columns 3-5, real unlevered return is the annualized unlevered return, calculated using the actual purchase price and sale price. Standard errors are double clustered by household and survey year. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 10: Controlling for listing agents

	Unlevered Ann Return	Purchase Discount	Sale Discount	Log(Purchase List Price)	Log(Sale List Price)	Log(Purchase Price)	Log(Sale Price)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Single Female	-0.012***	-0.233***	0.068***	0.013***	-0.013***	0.016***	-0.024***
	(0.000)	(0.005)	(0.004)	(0.000)	(0.000)	(0.001)	(0.001)
Couple	-0.010***	-0.042***	-0.175***	0.010***	0.004***	0.025***	0.014***
	(0.000)	(0.004)	(0.004)	(0.000)	(0.000)	(0.001)	(0.001)
Other		0.431***	0.246***	-0.017***	-0.030***	0.017***	-0.042***
		(0.006)	(0.006)	(0.000)	(0.001)	(0.003)	(0.001)
Property FE	No	No	No	Yes	Yes	Yes	Yes
Zip-Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Listing Agent FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.454	0.286	0.285	0.975	0.975	0.887	0.888
Observations	5,377,609	19,941,435	19,941,435	10,210,719	10,210,719	28,673,949	28,673,949

Note: This table re-estimates the main regressions from Tables 2, 4, 5, and 6, controlling for listing agent fixed effects. The sample is restricted to observations with matched listing agent data. Agents are identified using their listing agent code on a listing matched to a transaction. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 11: Variation by market tightness

	Unlevered Ann Return	Purchase Discount	Sale Discount	Log(Purchase Price)	Log(Sale Price)
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.017***	-0.284***	0.039***	0.024***	-0.029***
	(0.000)	(0.019)	(0.014)	(0.001)	(0.001)
Couple	-0.012***	-0.240***	-0.012	0.013***	0.017***
	(0.000)	(0.018)	(0.014)	(0.001)	(0.001)
Other		0.014	0.389***	0.032***	-0.060***
		(0.016)	(0.016)	(0.002)	(0.002)
Singe Female X Tightness	0.020***	0.284***	-0.487***	-0.039***	0.019***
	(0.002)	(0.095)	(0.066)	(0.004)	(0.004)
Couple X Tightness	-0.001	0.413***	-0.026	-0.004	-0.024***
	(0.002)	(0.091)	(0.065)	(0.006)	(0.005)
Other X Tightness		0.068	0.135*	-0.006	0.059***
		(0.077)	(0.074)	(0.008)	(0.009)
Property FE	No	No	No	Yes	Yes
Zip-Year-Month FE	Yes	Yes	Yes	Yes	Yes
R-squared	0.354	0.207	0.208	0.886	0.886
Observations	8,265,449	19,845,356	19,845,356	46,602,251	46,602,251

Note: This table re-estimates the main regressions from Tables 2, 4, and 6, interacting the gender group indicators with a measure of market tightness. Market tightness is defined as the number of sales in a given county-month, scaled by the outstanding number of listings currently for sale in that county-month. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Appendix for

The Gender Gap in Housing Returns

Paul Goldsmith-Pinkham Kelly Shue

A. Impact on Wealth Accumulation

In this section, we estimate that gender differences in the housing market can explain up to 30% of the gender gap in wealth accumulation at retirement.[19]

Before proceeding to the calculations, we note that wealth is measured at the household level and shared by married couples. Therefore, we estimate wealth accumulation by women and men who remain single until retirement, which represent approximately 10 and 13 percent of the US population, respectively.[20] Our estimations may still be informative when considering the welfare of ever-married individuals because we offer insight into each party's financial outside options if they were to remain single, which may affect bargaining power within married households. The limitations of our calculations are similar to those faced when considering how the gender pay gap contributes to differences in wealth accumulation, which likewise can only be estimated for single households but may matter for bargaining power and welfare within married households.

We assume that men and women buy their first house at age 30. They remain homeowners until age 65 when they retire and sell their home. House value grows at 2.5% annually. People pay off their mortgages in full by retirement and remain single from age 30 to 65. People move to new homes four times, at ages 35, 40, 45, and 54 (this timing is designed so that people move less frequently as they age). Men and women invest the same initial amount in housing at age 30 and do not move additional income into housing other than to cover transaction costs, which we assume to be the same for men and women. Men trade at fair market prices whereas women lose 2% of housing wealth at each house purchase and sale.

Under these assumptions, women accumulate 18.4% lower housing wealth relative to men by age 65. Intuitively, the difference arises because the initial male investment in housing grows at a rate of 2.5% annually. The initial female investment in housing grows at the same rate, but women experience a loss of 2% of housing wealth at each purchase and sale.

The gender gap in median net worth at retirement for single homeowners is approximately 40%,[21] and a single woman with the median level of wealth has 70% of her net worth invested in housing (see Appendix Figures A1 and A2). This implies that approximately 32.3% $(= (70\% \times 18.4\%)/40\%)$ of the gap in wealth accumulation at retirement can be explained by the gender gap in housing markets. Note that this calculation does not make any assumptions regarding the degree of leverage or mortgage debt taken by men and women, other than to assume that mortgages are paid in full by retirement.

We can modify our baseline calculation by changing the number of housing transactions. Women are less disadvantaged if they engage in fewer transactions. We assume that people move three times instead of four, at the ages of 35, 42, and 52, between the initial purchase at age 30 and final sale at age 65. Under these modified assumptions, the gender gap in housing wealth at retirement is 15.0%, implying that 26.3% of the gap in total wealth accumulation can be explained by the gender gap in housing markets.

We can also modify our baseline calculation by allowing people to invest more in housing with each move. This is equivalent to allowing people to contribute to their housing portfolio over time by moving to more expensive homes. We assume that men invest an additional 15% of the market value of their current house into their next house each time they move. We assume women move at the same time and invest the same additional dollar amount as men each time they move. Under these modified assumptions, the gender gap in housing wealth at retirement is 14.3%, implying that 25.1% of the gap in total wealth accumulation can be explained by the gender gap in housing markets.

These estimates may be conservative because they rely on an estimate of the gender gap in exe-

[19]We thank Tim Landvoigt for recommending this exercise and for detailed suggestions regarding execution.

[20]The numbers come from the 2018 American Community Survey, and represent the percentage of never-married women and men aged 55 to 64.

[21]Our estimate of the gender gap in wealth accumulation is similar to the estimated gap of 36% for never-married men and women at age 65 presented in Ruel and Hauser (2013).

cution prices at purchase and sale, which comes from regression results in Table 4, which control for property fixed effects and zip-year-month fixed effects. These regressions estimate the gender gap in execution prices among men and women who choose to transact in the same zip-year-month. Our calculation does not take into account that women also earn lower returns due to market timing, in that they are more likely to buy when aggregate prices are high and sell when they are low.

B. Measurement Error

In this section, we discuss the potential bias in our estimates that may arise due to measurement error in our gender group variables (single male, single female, couple, and other).

The result below is a simple extension of Aigner (1973). We show that measurement error would cause us to underestimate the extent to which single women earn lower returns on housing relative to single men under the following assumptions: (1) measurement error in our measured gender groups (single male, single female, couple, and other) is random (not correlated with true gender groups), (2) observations observed empirically as couples are not true single men or women (because we identify an observation as a couple only if we observe two separate full names with identified genders, so there are unlikely to be false positives in the identification of couples), (3) the rates of false identification of true single men as single women and true single women as single men are low (based on the fact that our gender identification algorithm uses a 95% confidence cutoff), (4) a true couple is more likely to be mistakenly identified as a single man than as a single woman (because of pre-existing norms described in Section II.B), and (5) couples earn weakly lower returns on housing relative to single men (based on supplementary results from the American Housing Survey).

With these assumptions at hand, we show below that our estimate of the gender return gap between single women and single men will be smaller than the true gender gap due to measurement error.

Result We focus on the simple returns regression, without controlling for location or timing, for clarity of exposition. The true returns regression is

$$r_i = \beta_0 + C_i\beta_c + F_i\beta_f + \epsilon_i, \tag{A1}$$

where M_i, F_i, and C_i denote the true gender group status on a transaction i, corresponding to single male, single female, and couple.

Let m_i, f_i, c_i denote our empirically measured gender group, with o_i denoting the "other" category for which we cannot identify the gender group.

Note that the measurement error for each term is:

$$c_i = C_i + u_{c,i}, \tag{A2}$$
$$f_i = F_i + u_{f,i}, \tag{A3}$$
$$m_i = M_i + u_{m,i}, \tag{A4}$$

where $u_{\cdot,i} \in \{-1, 0, 1\}$ such that m_i, f_i, c_i, and $o_i \in \{-1, 0, 1\}$ and $m_i + f_i + c_i + o_i = 1$. We assume that the vector of u are i.i.d. random draws (with probabilities outlined below).

We can empirically estimate the following regression:

$$r_i = \beta_0 + c_i\beta_c + f_i\beta_f - u_{c,i}\beta_c - u_{f,i}\beta_f + \epsilon_i, \tag{A5}$$

or in compressed notation:

$$r_i = X_i\beta - u_{c,i}\beta_c - u_{f,i}\beta_f + \epsilon_i, \tag{A6}$$

where $X_i = (1\ c_i\ f_i)$, and $\beta = (\beta_0, \beta_c, \beta_f)'$.

Note that we condition on $o_i = 0$ in our regressions. We revisit this below.

Now consider the estimator of β from Equation A6:

$$\hat{\beta} = \beta - (X_n'X_n)^{-1}X_n'(u_{c,n}\beta_c + u_{f,n}\beta_f) + (X_n'X_n)^{-1}X_n'\epsilon_n, \tag{A7}$$

where X_n is an $n \times 3$ stacked matrix of X_i, and $u_{c,n}$ is an $n \times 1$ stacked vector of $u_{c,i}$ (and the same for $u_{f,n}$ and ϵ_n). For simplicity of notation, let the observed number of single female and single males be $n_f = an$, and $n_c = bn$, respectively, where $a, b \in (0,1)$.

We can derive the following:

$$(X_n'X_n) = \begin{bmatrix} n & n_c & n_f \\ n_c & n_c & 0 \\ n_f & 0 & n_f \end{bmatrix}, \tag{A8}$$

$$(X_n'X_n)^{-1} = \begin{bmatrix} \frac{1}{n(1-a-b)} & \frac{-1}{n(1-a-b)} & \frac{-1}{n(1-a-b)} \\ \frac{-1}{n(1-a-b)} & \frac{n(1-a)}{bn(n(1-a-b))} & \frac{1}{n(1-a-b)} \\ \frac{-1}{n(1-a-b)} & \frac{1}{n(1-a-b)} & \frac{n(1-b)}{an(n(1-a-b))} \end{bmatrix}, \tag{A9}$$

$$X_n'u_{c,n} = \begin{bmatrix} \sum_{i=1}^n u_{c,i} \\ \sum_{i=1}^n c_i u_{c,i} \\ \sum_{i=1}^n f_i u_{c,i} \end{bmatrix}, \tag{A10}$$

$$X_n'u_{f,n} = \begin{bmatrix} \sum_{i=1}^n u_{f,i} \\ \sum_{i=1}^n c_i u_{f,i} \\ \sum_{i=1}^n f_i u_{f,i} \end{bmatrix}. \tag{A11}$$

Hence, the bias for β_f will be

$$\hat{\beta}_f - \beta_f = \beta_c \frac{\sum_{i=1}^n u_{c,i}}{n(1-a-b)} - \beta_c \frac{\sum_{i=1}^n c_i u_{c,i}}{n(1-a-b)} - \beta_c \frac{n(1-b)\sum_{i=1}^n f_i u_{c,i}}{an(n(1-a-b))} \tag{A12}$$

$$+\beta_f \frac{\sum_{i=1}^n u_{f,i}}{n(1-a-b)} - \beta_f \frac{\sum_{i=1}^n c_i u_{f,i}}{n(1-a-b)} - \beta_f \frac{n(1-b)\sum_{i=1}^n f_i u_{f,i}}{an(n(1-a-b))} + o_p(1). \tag{A13}$$

If we take limits, and assume all three groups will be non-trivial in size, then we get:

$$\hat{\beta}_f - \beta_f = \frac{\beta_c E(u_{c,i})}{1-a-b} - \frac{\beta_c E(c_i u_{c,i})}{1-a-b} - \frac{(1-b)\beta_c E(f_i u_{c,i})}{a(1-a-b)} \tag{A14}$$

$$+ \frac{\beta_f E(u_{f,i})}{1-a-b} - \frac{\beta_f E(c_i u_{f,i})}{1-a-b} - \frac{(1-b)\beta_f E(f_i u_{f,i})}{a(1-a-b)}. \tag{A15}$$

Now consider the joint distribution of our observed data and the true data:

$$Pr(X_i, x_i) = \begin{array}{c|ccc} & C & F & M \\ c & (1-\lambda_c-\kappa_C-\eta_C)\pi_C & 0 & 0 \\ f & \kappa_C\pi_C & \kappa_F\pi_F & \kappa_M\pi_M \\ m & \eta_C\pi_C & \eta_F\pi_F & \eta_M\pi_M \\ o & \lambda_c\pi_C & \lambda_F\pi_F & \lambda_M\pi_M, \end{array} \tag{A16}$$

where $\kappa_M = Pr(f_i = 1|M_i = 1)$, $\pi_M = Pr(M_i = 1)$, $\eta_M = Pr(m_i = 1|M_i = 1)$ and $\lambda_M = Pr(o_i = 1|M_i = 1)$. $\kappa_C, \pi_C, \eta_C, \lambda_C, \kappa_F, \pi_F, \eta_F$, and λ_F are defined similarly. Note that we assume that observed couples do not correspond to true single male or female. Since we condition on $o = 0$, we can recondition this distribution, by scaling the terms by $Pr(o = 0) = 1 - (\lambda_c \pi_C + \lambda_F \pi_F + \lambda_M \pi_M) = \alpha$.

Finally, we use this joint distribution to consider the joint probability of our measured variables and error terms (recall that $Pr(u_{c,i} = -1) = Pr(C_i = 1, c_i = 0)$, $Pr(u_{c,i} = 0) = Pr(C_i = 1, c_i = 1) + Pr(C_i = 0, c_i = 0)$ and $Pr(u_{c,i} = 1) = Pr(C_i = 0, c_i = 1)$, conditional on $o_i = 0$):[22]

$u_{f,i}$ \ f_i	0	1	$f(u_{f,i})$
-1	$\eta_F \pi_F / \alpha$	0	$\eta_F \pi_F / \alpha$
0	$((1 - \lambda_c - \kappa_C)\pi_C + \eta_M \pi_M)/\alpha$	$\kappa_F \pi_F / \alpha$	$((1 - \lambda_c - \kappa_C)\pi_C + \eta_M \pi_M + \kappa_F \pi_F)/\alpha$
1	0	$(\kappa_M \pi_M + \kappa_C \pi_C)/\alpha$	$(\kappa_M \pi_M + \kappa_C \pi_C)/\alpha$
$f(f_i)$	$((1 - \lambda_c - \kappa_C)\pi_C + \eta_M \pi_M + \eta_F \pi_F)/\alpha$	$(\kappa_F \pi_F + \eta_M \pi_M + \kappa_C \pi_C)/\alpha$	1

$u_{f,i}$ \ c_i	0	1	$f(u_{f,i})$
-1	$\eta_F \pi_F / \alpha$	0	$\eta_F \pi_F / \alpha$
0	$(\kappa_F \pi_F + \eta_M \pi_M + \eta_C \pi_C)/\alpha$	$(1 - \lambda_c - \kappa_C - \eta_C)\pi_C/\alpha$	$((1 - \lambda_c - \kappa_C)\pi_C + \eta_M \pi_M + \kappa_F \pi_F)/\alpha$
1	$(\kappa_M \pi_M + \kappa_C \pi_C)/\alpha$	0	$(\kappa_M \pi_M + \kappa_C \pi_C)/\alpha$
$f(c_i)$	$\vdots$	$(1 - \lambda_c - \kappa_C - \eta_C)\pi_C/\alpha$	1

$u_{c,i}$ \ f_i	0	1	$f(u_{c,i})$
-1	$\eta_C \pi_C / \alpha$	$\kappa_C \pi_C / \alpha$	$(\eta_C + \kappa_C)\pi_C/\alpha$
0	$(\eta_F \pi_F + \eta_M \pi_M + (1 - \lambda_C - \kappa_C - \eta_C)\pi_C)/\alpha$	$(\kappa_F \pi_F + \kappa_M \pi_M)/\alpha$	$\dots$
1	0	0	0
$f(f_i)$	$(\eta_F \pi_F + (1 - \lambda_c)\pi_C + \kappa_M \pi_M)/\alpha$	$(\kappa_F \pi_F + \eta_M \pi_M)/\alpha$	1

$u_{c,i}$ \ c_i	0	1	$f(u_{c,i})$
-1	$(\eta_C + \kappa_C)\pi_C/\alpha$	0	$(\eta_C + \kappa_C)\pi_C/\alpha$
0	$((\kappa_M + \eta_M)\pi_M + (\kappa_F + \eta_F)\pi_F)/\alpha$	$(1 - \lambda_c - \kappa_C - \eta_C)\pi_C/\alpha$	$\dots$
1	0	0	0
$f(c_i)$	$(\kappa_F \pi_F + \kappa_M \pi_M + \eta_M \pi_M + \eta_F \pi_F)/\alpha$	$(1 - \lambda_c)\pi_C/\alpha$	1

Solving for expectations yields:

$$E(u_{c,i}) = -(\eta_C + \kappa_C)\pi_C/\alpha, \tag{A17}$$
$$E(u_{f,i}) = (\kappa_M \pi_M + \kappa_C \pi_C - \eta_F \pi_F)/\alpha, \tag{A18}$$
$$E(c_i u_{c,i}) = 0, \tag{A19}$$
$$E(f_i u_{c,i}) = -\kappa_C \pi_C / \alpha, \tag{A20}$$
$$E(c_i u_{f,i}) = 0, \tag{A21}$$
$$E(f_i u_{f,i}) = (\kappa_M \pi_M + \kappa_C \pi_C)/\alpha. \tag{A22}$$

Now we consider our bias term for β_f again:

[22] For space reasons, we omit certain cells, but these are simply the sum of the terms in each column or row.

$$\hat{\beta}_f = \beta_f - \frac{\beta_c(\eta_C + \kappa_C)\pi_C/\alpha}{1-a-b} + \frac{(1-b)\beta_c\kappa_C\pi_C/\alpha}{a(1-a-b)} \tag{A23}$$

$$+ \frac{\beta_f(\kappa_M\pi_M + \kappa_C\pi_C - \eta_F\pi_F)/\alpha}{1-a-b} - \frac{(1-b)\beta_f(\kappa_M\pi_M + \kappa_C\pi_C)/\alpha}{a(1-a-b)} \tag{A24}$$

$$= \beta_f - \beta_c\frac{a\eta_C\pi_C - (1-b-a)\kappa_C\pi_C}{\alpha a(1-a-b)} \tag{A25}$$

$$- \beta_f\frac{a\eta_F\pi_F + (1-b-a)(\kappa_M\pi_M + \kappa_C\pi_C)}{\alpha a(1-a-b)} \tag{A26}$$

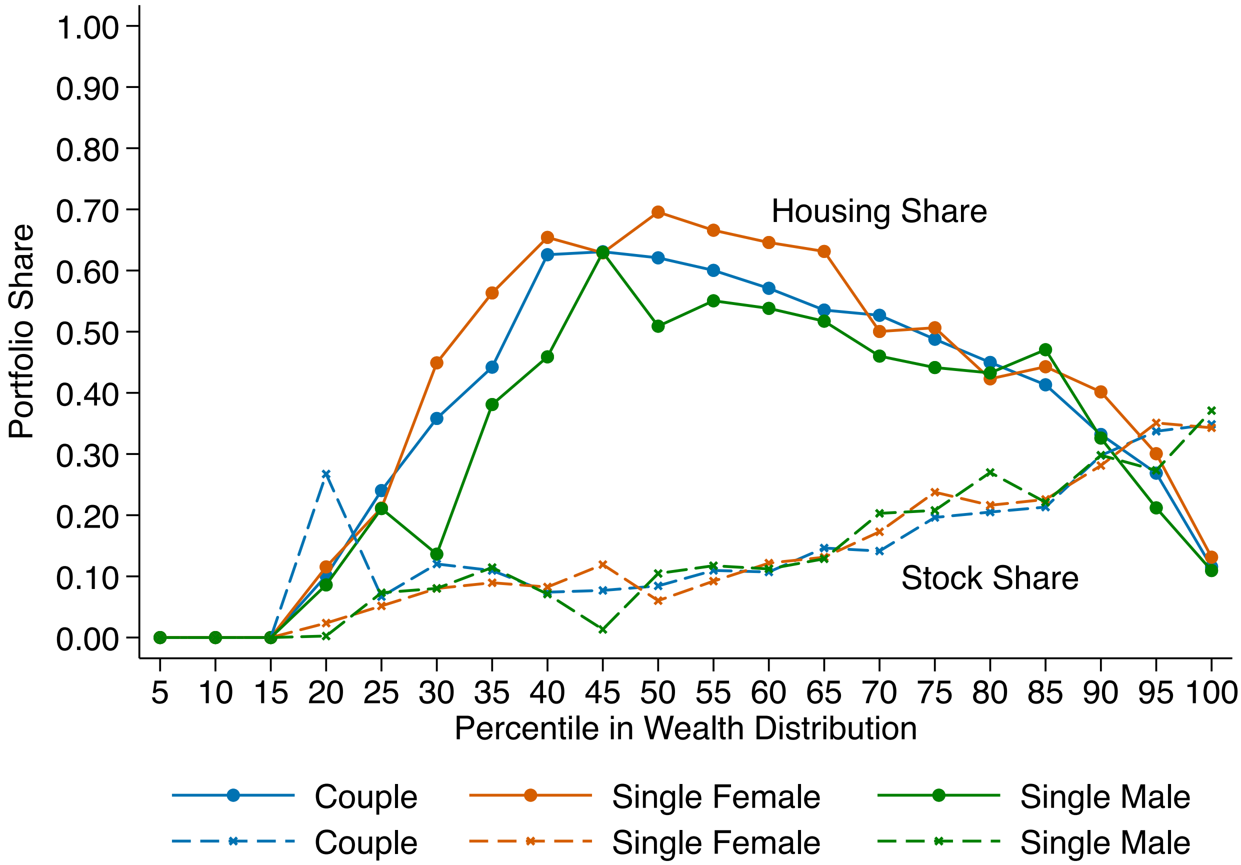
$$= \beta_f\left(1 - \frac{a\eta_F\pi_F + (1-b-a)(\kappa_M\pi_M + \kappa_C\pi_C)}{\alpha a(1-a-b)}\right) \tag{A27}$$

$$- \beta_c\frac{a\eta_C\pi_C - (1-b-a)\kappa_C\pi_C}{\alpha a(1-a-b)}. \tag{A28}$$

Our measure of the gender gap in returns, $\hat{\beta}_f$, can be biased due to two factors expressed in lines A27 and A28. The first represents attenuation towards zero as in Aigner (1973) if we assume that $\frac{a\eta_F\pi_F+(1-b-a)(\kappa_M\pi_M+\kappa_C\pi_C)}{\alpha a(1-a-b)} \in [0,1]$. We believe this assumption is reasonable because of the following: $\eta_F\pi_F$ is the probability that the true gender group is single female and is mistakenly identified as single male. $\kappa_M\pi_M$ is the probability that the true gender group is single male and is mistakenly identified as single female. Both terms are weakly positive and likely to be close to zero because we identify first names as male or female only if our matching algorithm output exceeds a 95% confidence level. $\kappa_C\pi_C$ is the probability that the true gender group is couple and is mistakenly identified as single female. This term is weakly positive and also likely to be close to zero. Empirical patterns suggest that some couples list only the male name on deeds records. However, it is much more rare for a true couple to list only the female name on deeds records (see discussion in Section II.B).

The second bias factor in line A28 represents upward bias if we assume that $-\beta_c\frac{a\eta_C\pi_C-(1-b-a)\kappa_C\pi_C}{\alpha a(1-a-b)} > 0$. We believe this assumption is reasonable because of the following: First note that $a\eta_C\pi_C - (1-b-a)\kappa_C\pi_C > 0$ if $a/(1-b-a) > \kappa_C/\eta_C$. $a/(1-b-a)$ is the ratio of the observed share of single women to the share of single men, and is equal to approximately 0.7 in our data (see Appendix Figure A3). κ_C/η_C is the ratio of the probability that a true couple is identified as a single female to the probability that a true couple is identified as a single male. As discussed in Section II.B, it is much more common for a couple to list only the male name than to list only the female name, so this fraction is likely to be close to zero and therefore less than 0.7. Finally, β_C is the return earned by couples relative to single men. β_C is likely to be weakly less than zero because we estimate that couples slightly underperform single men using supplementary data from the American Housing Survey (AHS). In the AHS, marital status and gender are measured with minimal error because they are reported by survey respondents rather than inferred from names.

Figure A1: Share of net worth invested in housing versus stocks, Survey of Consumer Finances

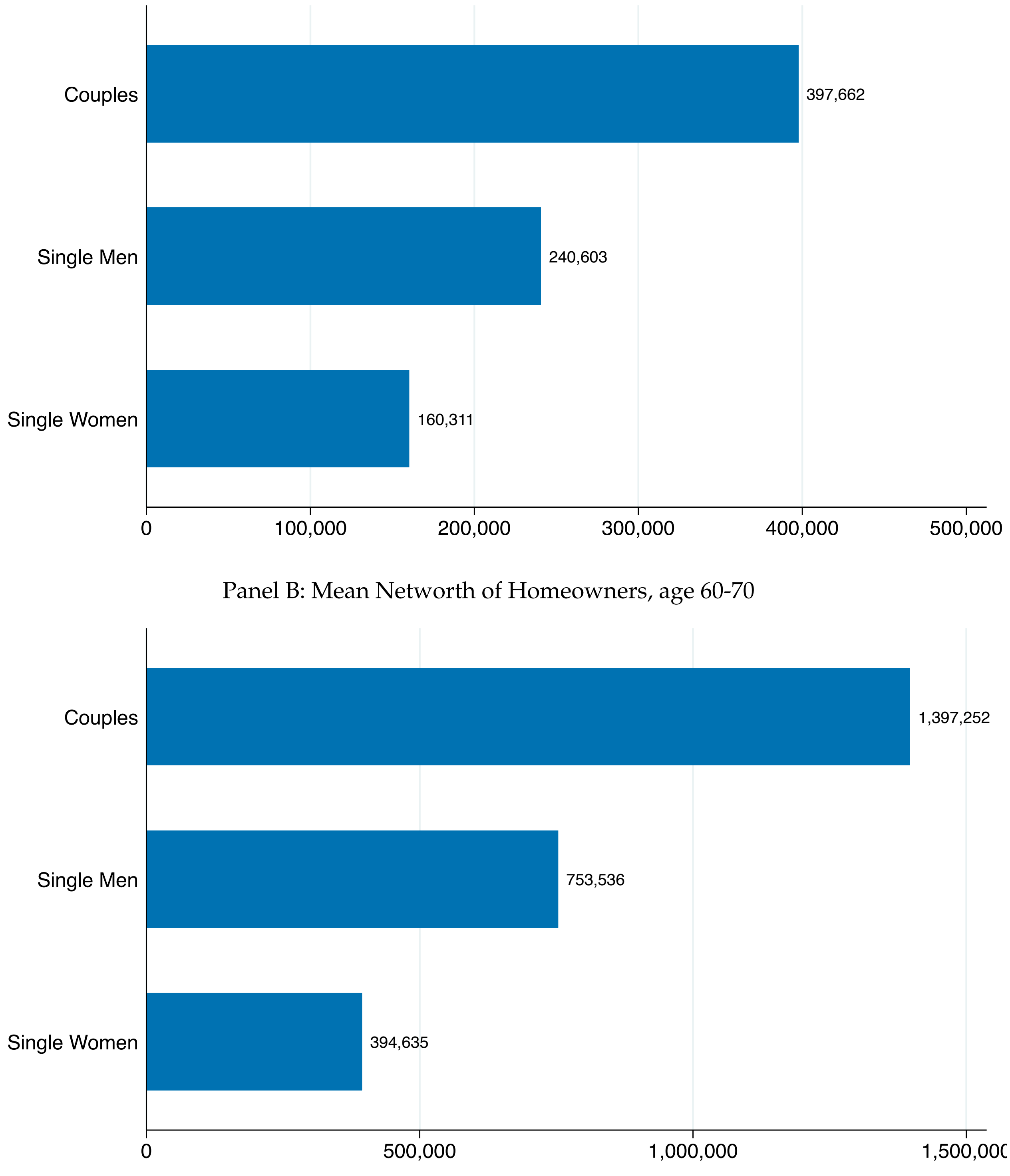


Note: This figure plots the average share of net worth invested in housing and financial equities (stocks) across the wealth distribution, for individuals between the ages of 60 and 70, split by single women, single men, and couples. Single women and single men are defined by gender of head of household, and not living with a partner. Couples are defined as those living with partner or married. Housing share is total housing equity divided by net worth. Stock share is the share of wealth in all equity investments (including retirement accounts, IRA/Keogh accounts, directly held pooled investment funds held by household, directly held stocks held by household, account-type pension plans from the head of household and spouse's current jobs, and trusts investments) divided by net worth. Both share variables are set to zero if the numerator is negative or zero, or net worth is negative or zero. We pool across all years in the Survey of Consumer Finance (1989-2016), and all variables are measured in 2016 dollars.

Figure A2: Wealth at retirement, Survey of Consumer Finances

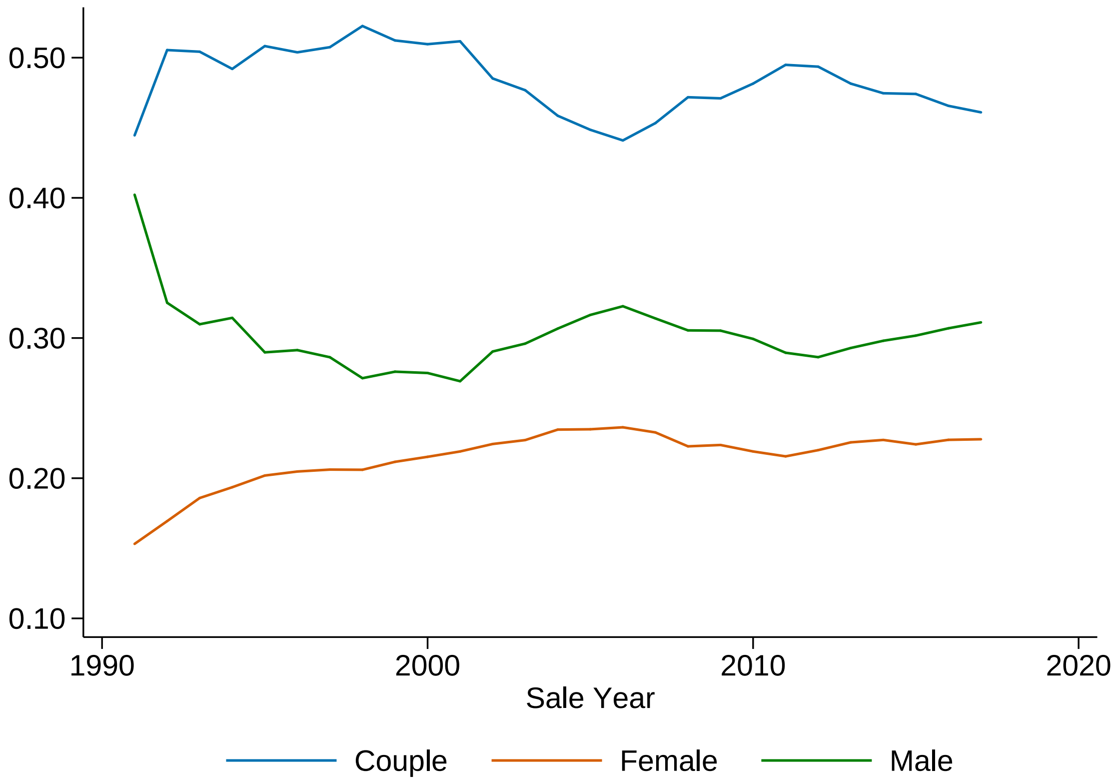
Panel A: Median Networth of Homeowners, age 60-70

Panel B: Mean Networth of Homeowners, age 60-70



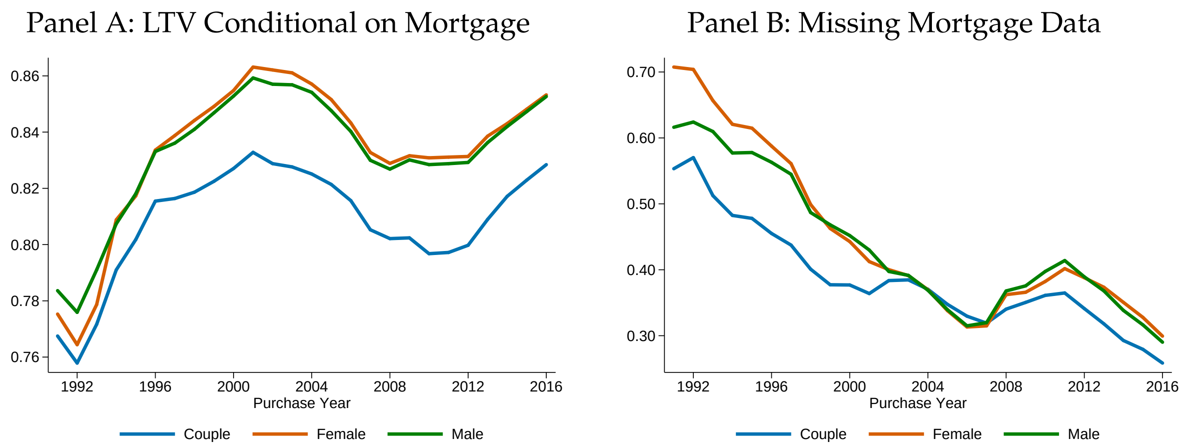
Note: This figure plots the mean and median net worth across couples, single males, and single females between the ages of 60 and 70. Age is defined by the head of the household, as reported in the survey. Single women and single men are defined by gender of head of household, and not living with partner. Couples are defined as those living with partner or married. We pool across all years in the Survey of Consumer Finance (1989-2016), and all variables are measured in 2016 dollars.

Figure A3: Composition of transactions by gender group over time



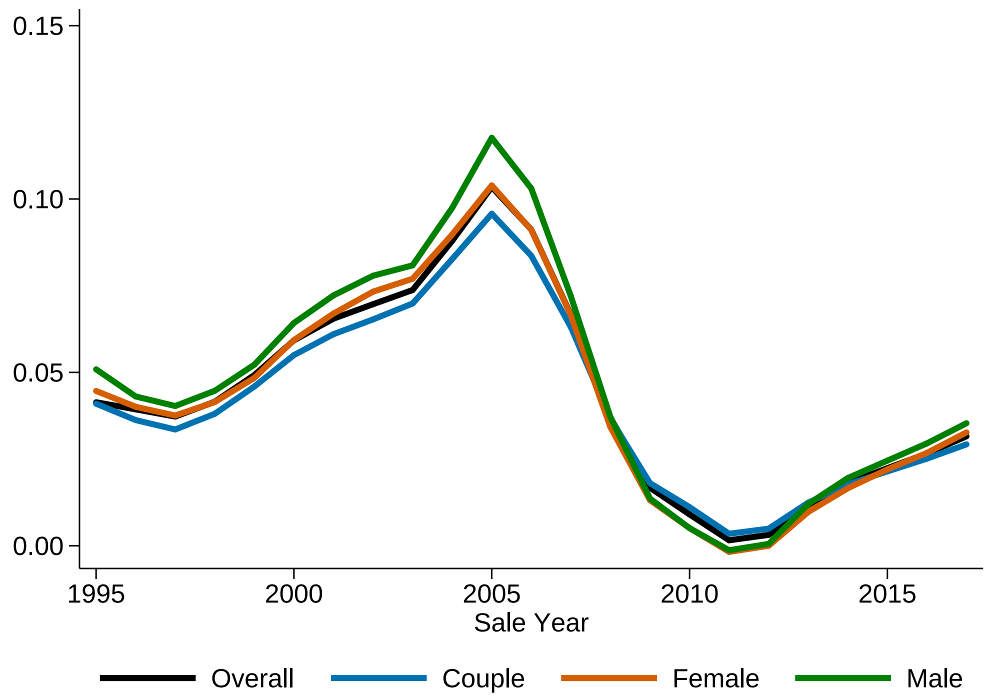
Note: This figure plots the relative composition of sale transactions across couples, single males, and single females within the sample of transactions used for returns estimation.

Figure A4: Original LTV over time by gender group



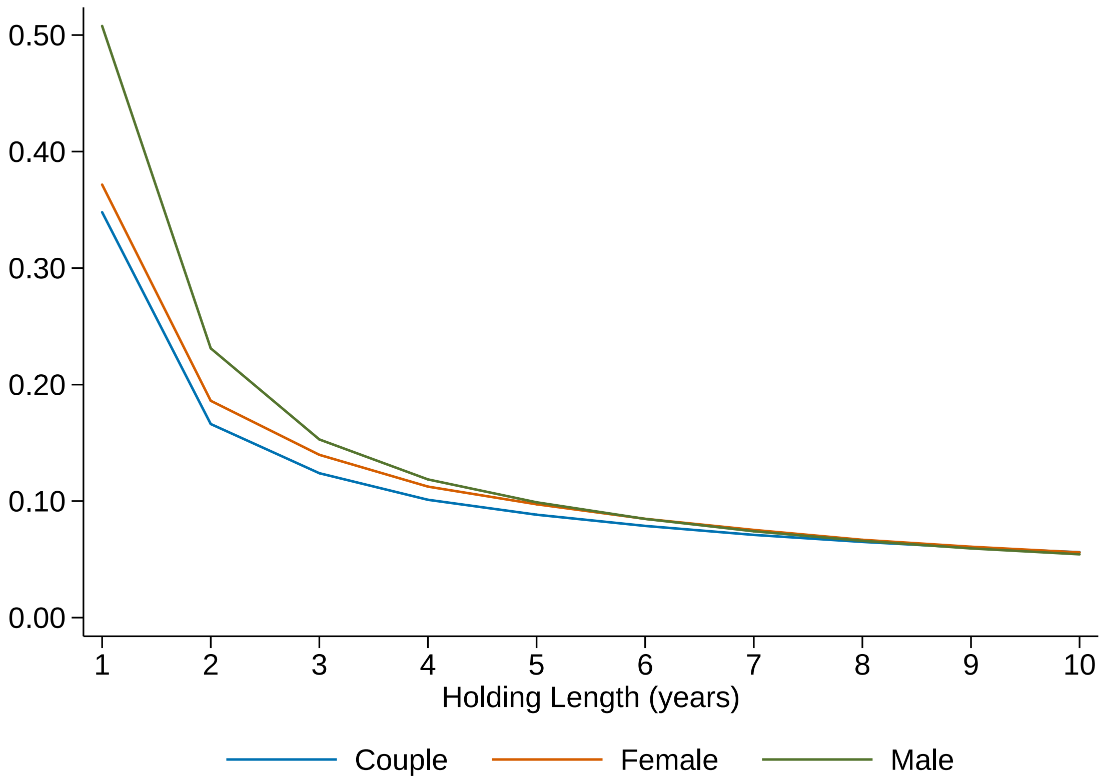
Note: This figure presents average loan-to-value (LTV) at the time of purchase for couples, single women, and single men. In Panel A, we plot the average LTV at time of purchase, conditional on having on mortgage data. In Panel B, we plot the share of transactions with missing mortgage data. This combines two forces: full cash transactions and observations with missing mortgage amount data.

Figure A5: Median unlevered returns over time by gender group



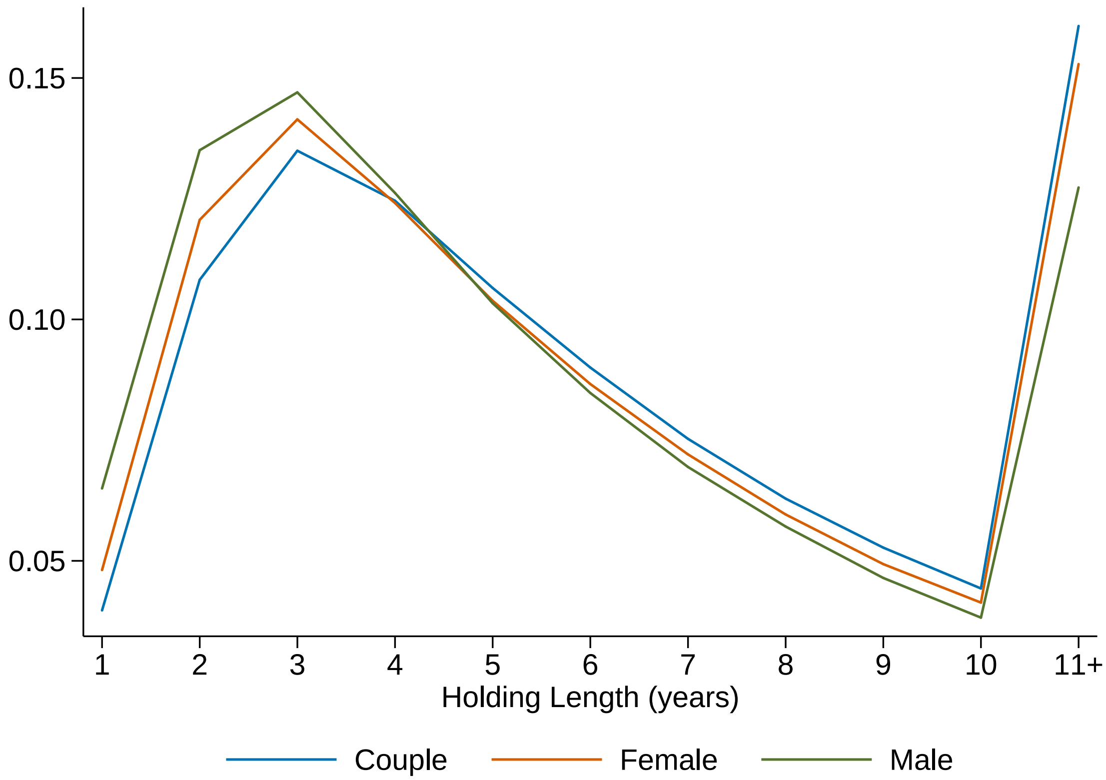
Note: This figure plots the median unlevered annualized return for couples, single women, and single men by sale year. As our sample begins in 1991, we begin this figure in 1995 to allow for sufficient data to avoid truncation. See Section II.B for more details on the definition of gender and family structure.

Figure A6: Unlevered returns by holding period



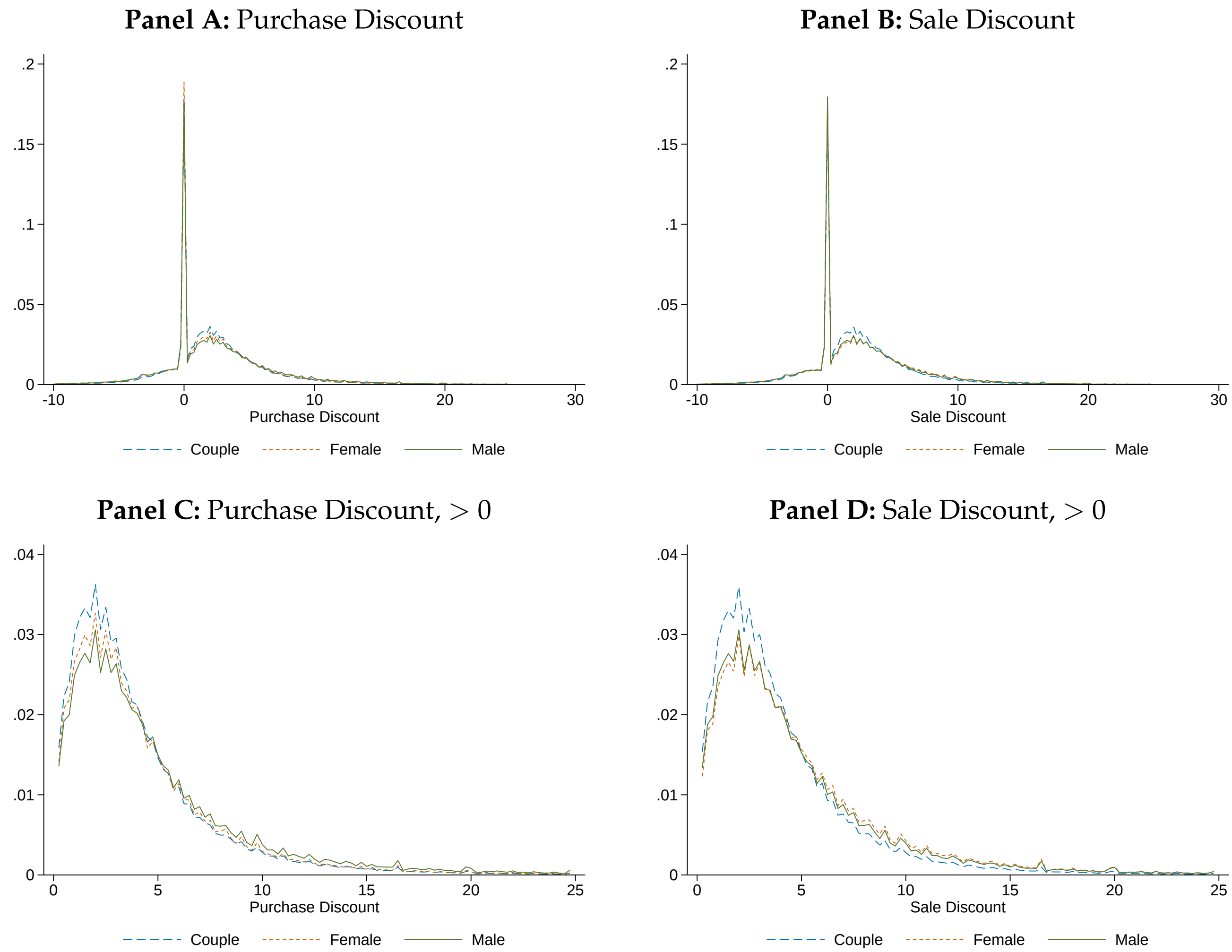
Note: This figure plots the average unlevered annualized returns for couples, single women, and single men by holding length. We exclude holding periods longer than 11 years.

Figure A7: Transaction share by holding length



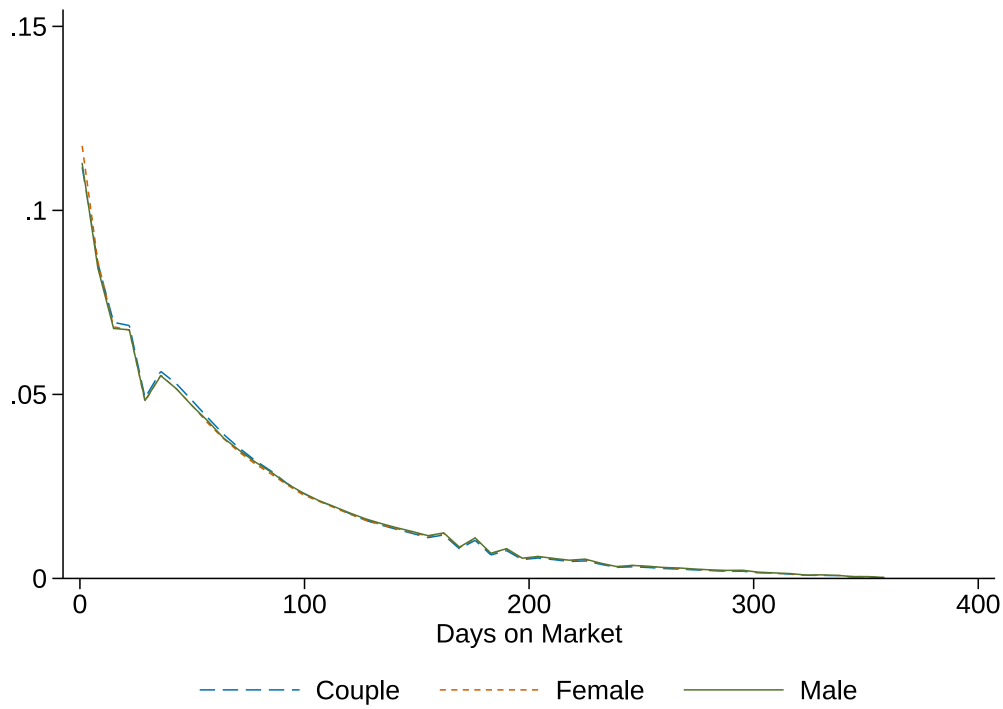
Note: This figure plots the distribution of transactions for couples, single women, and single men by holding lengths within our returns sample. We restrict the sample to properties with a minimum holding period of 3 months.

Figure A8: Sale and purchase discount distributions



Note: This figure plots the distribution of purchase and sale discounts for couples, single women and single men. We measure discounts as (list price - transaction price)/list price × 100, so a larger purchase discount contributes to a higher return on housing investment and a larger sale discount contributes to a lower return on housing investment. In Panels A and B, we plot the full distributions for all three groups. In Panels C and D, we restrict the distribution to values greater than zero to exclude the spike at 0.

Figure A9: Days on market distribution



Note: This figure plots the distribution of days on market for listings sold by couples, single women, and single men. Days on market equals the number of days between the earliest available listing associated with a transaction and the sale date.

Table A1: Housing returns: market timing with less conservative screens

Panel A: Unlevered Returns

	Unlevered Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.016***	-0.013***	-0.011***	-0.010***	-0.009***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Couple	-0.020***	-0.012***	-0.007***	-0.007***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Length			-0.006***	-0.001	
			(0.000)	(0.001)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.005	0.354	0.379	0.534	0.592
Observations	9,351,419	9,351,419	9,351,419	9,351,419	9,351,419

Panel B: Levered (80%) Returns

	Levered (80%) Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.078***	-0.071***	-0.057***	-0.051***	-0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Couple	-0.099***	-0.071***	-0.043***	-0.038***	-0.015***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Holding Length			-0.037***	-0.028***	
			(0.000)	(0.006)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.004	0.295	0.330	0.482	0.628
Observations	9,351,419	9,351,419	9,351,419	9,351,419	9,351,419

Note: This table re-estimates Table 2, using less conservative screens to identify the gender groups single male, single female, and couple of the owners. These less conservative screens do not require county-years to satisfy certain cut-offs of single female owner share and couple share (see Section II.B). Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A2: Housing returns with estimated leverage

Panel A: Unlevered Returns

	Unlevered Ann Return (Post-2000)				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.019***	-0.015***	-0.013***	-0.011***	-0.009***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Couple	-0.024***	-0.013***	-0.008***	-0.007***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Length			-0.010***	-0.005***	
			(0.000)	(0.002)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.005	0.414	0.442	0.589	0.636
Observations	5,359,865	5,359,865	5,359,865	5,359,865	5,359,865

Panel B: Levered (Estimated LTV) Returns

	Levered Ann Return (Estimated LTV)				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.095***	-0.082***	-0.065***	-0.060***	-0.035***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Couple	-0.188***	-0.116***	-0.079***	-0.073***	-0.040***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Holding Length			-0.075***	-0.084***	
			(0.001)	(0.028)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.002	0.245	0.254	0.380	0.428
Observations	5,359,865	5,359,865	5,359,865	5,359,865	5,359,865

Note: This table re-estimates Table 2, using the initial loan-to-value estimates from the underlying Corelogic data in Panel B. When mortgages are missing in the underlying data, we assume that it is a cash purchase (LTV of 0%). We restrict the sample in both panels to observations with purchase year in 2000 or later, where there is less missing mortgage data (see Panel B of Appendix Figure A4). Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A3: Match rates

Panel A: Overall Match Rates

	Buyer Gender			
Seller Gender	Single Male	Single Female	Couple	Overall
Single Male	0.1385	0.0868	0.1010	0.3262
	[0.1207]	[0.0830]	[0.1225]	
Single Female	0.0936	0.0748	0.0752	0.2437
	[0.0901]	[0.0620]	[0.0915]	
Couple	0.1378	0.0930	0.1993	0.4301
	[0.1591]	[0.1095]	[0.1615]	
Overall	0.3700	0.2546	0.3755	1

Panel B: Zip-Year-Quarter Match Rates

	Buyer Gender			
Seller Gender	Single Male	Single Female	Couple	Overall
Single Male	0.1503	0.0997	0.1141	0.3372
	[0.1400]	[0.0936]	[0.1204]	
Single Female	0.1032	0.0869	0.0866	0.2535
	[0.0990]	[0.0736]	[0.0940]	
Couple	0.1533	0.1088	0.2225	0.4642
	[0.1596]	[0.1171]	[0.2000]	
Overall	0.3799	0.2696	0.3958	1

Note: This table presents the joint probability of a seller of a given gender group matching with a buyer of a given gender group. The first number in each cell is the empirical match rate as seen in the data. The number in brackets is the theoretical number if match rates were random (using the product of the two marginal empirical rates). Non-categorized observations (the other category) are excluded from the matching exercise. Panel A pools the full sample, while Panel B calculates the actual and random match rates at the zip-year-quarter level, and then reports the unweighted average across zip-year-quarters.

Table A4: Housing returns: weighted by holding length

Panel A: Unlevered Returns

	Unlevered Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.005***	-0.005***	-0.004***	-0.005***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Couple	-0.004***	-0.001***	0.000***	-0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Length			-0.002***	0.003***	
			(0.000)	(0.001)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.001	0.383	0.388	0.559	0.601
Observations	8,933,131	8,933,131	8,933,131	8,933,131	8,933,131

Panel B: Levered (80%) Returns

	Unlevered Ann Return				
	(1)	(2)	(3)	(4)	(5)
Single Female	-0.017***	-0.017***	-0.014***	-0.017***	-0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Couple	-0.007***	-0.002***	0.004***	-0.001***	0.006***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Holding Length			-0.007***	0.001	
			(0.000)	(0.003)	
Zip-SaleYM FE	No	Yes	Yes	Yes	Yes
Zip-BuyYM FE	No	No	No	Yes	Yes
SaleYM FE x BuyYM FE	No	No	No	No	Yes
R-squared	0.000	0.321	0.326	0.509	0.618
Observations	8,933,131	8,933,131	8,933,131	8,933,131	8,933,131

Note: This table re-estimates Table 2, weighting each observation by holding length. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A5: Selection of property characteristics

	New Construction	Log(House Age)	Log(Sq Ft)
	(1)	(2)	(3)
Single Female	0.003***	-0.021***	-0.066***
	(0.000)	(0.003)	(0.001)
Couple	0.033***	-0.137***	0.143***
	(0.001)	(0.004)	(0.002)
Zip-Year-Month FE	Yes	Yes	Yes
R-squared	0.271	0.513	0.445
Observations	8,933,131	2,131,508	1,937,762

Note: This table examines gender differences in preferences for property characteristics. Standard errors are clustered by zipcode. The sample is restricted to observations with matched listings data covering the relevant housing characteristic. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A6: Upgrades

	Had Upgrades	Returns (Not Upgraded)	Returns (Upgraded)
	(1)	(2)	(3)
Single Female	-0.010***	-0.010***	-0.017***
	(0.001)	(0.000)	(0.000)
Couple	0.000	-0.009***	-0.020***
	(0.001)	(0.000)	(0.001)
Zip-Year-Month FE	Yes	Yes	Yes
R-squared	0.296	0.415	0.387
Observations	3,432,795	2,330,733	1,102,062

Note: Properties are considered to be upgraded if the listing text contain synonyms for upgrades, renovations, new features, or expansions. Column 1 measures relative upgrade rates across gender groups. Columns 2 and 3 estimate the gender gap in unlevered annualized returns for subsamples of the data that have not been upgraded or have been upgraded, respectively. The sample is restricted to observations that are matched to listings with non-missing property descriptions. Standard errors are clustered by zipcode. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.