ONLINE APPENDIX
Conflicts of Interest and Steering in Residential Brokerage

By PANLE JIA BARWICK, PARAG A. PATHAK, AND MAISY WONG

Table of Contents for the Online Appendix

APPENDIX: SAMPLE AND VARIABLE CONSTRUCTION

A1. Housing transactions

We begin with 722,925 non-rental listings for condominiums, single-family, and multi-family properties. We first drop 52,226 duplicate listings, 221 listings with list or sale prices that are below $10,000, and 5,546 listings with problematic listing office codes. We then keep listings whose status is cancelled, expired, sold, or withdrawn (this removes 4,721 listings) and drop 4,377 listings with missing market information. We lose 1512 listings with 0 commission rates, 540 listings with missing commission rates, and 307 listings with buying commission rates greater than 5 percent (which implies a total commission rate greater than 10 percent). This leaves us with a final sample of 653,475 listings and 421,329 sold listings. We have geocoded street addresses and property identifiers for 646,460 listings. We are able to identify 133,903 properties that have repeat listings (for a total of 344,832 listings) and 62,843 properties with repeat sales (for a total of 137,085 sales).

A2. Offices

Each office is identified by an office ID. Two big chains (Coldwell Banker and Dewolfe) merged in 2002. Some offices changed office IDs as a result of this merger but kept the same office location. We recognize them as the same office and assign them a unique office ID. In addition, offices that use the same office location (e.g., 1000 Mass Ave, Fl 2, Cambridge, 02138) during the same time period are recognized as the same office and assigned a unique office ID.

We identify 172 chains, representing 486,189 listings (74%) and 316,571 purchases (75%). We first identify offices that have multiple locations and offices that have at least 100 listings and purchases. Within this group, we group offices that have similar names as chains. For example, all offices that have "Century 21" in the name are categorized under the Century 21 chain.

Many agents and offices have only a few transactions in our sample. We determine which offices and agents are active according to the *average annual number of transactions*, which is the total number of transactions divided by the number of years an office or agent spans our data (calculated as the last year the office or agent is in our data minus the first year, plus one). We use this average to identify active and top offices and agents. Our analyses focus on offices with five or more average annual number of listings and agents with two or more average annual number of listings. They account for 95% and 92% of listings, respectively.

In our office-year analyses (Table 7 and Table C1), each office is assigned a primary market in each year. We define a primary market by ranking the total number of listings and purchases by an office in a market in a year, followed by the total value of transactions. Ties are broken by the alphabetical order of market names.

A3. Defining markets

We have a total of 87 markets. Outside of Boston, markets are defined by cities and towns. We combine small markets with a nearby continguous market that account for the most cross-market listings by brokerage offices in these small markets. The combined markets include Cohasset-Hull, Avon-Holbrook, Lynn-Nahant, Sherborn-Natick, Topsfield-Middleton, Lincoln-Wayland, Concord-Carlisle, Danvers-Wenham, Stow-Acton, Dover-Wellesley, Millis-Medfield, and Handon-Rockland. We split the city of Boston into 15 sub-markets according to a GIS shapefile of Boston neighborhoods defined by Zillow. These sub-markets include Dorchester, Allston-Brighton, Back Bay-Beacon Hill, Charlestown, East Boston, Fenway-Kenmore, Jamaica Plain, Roslindale, Roxbury, West Roxbury, South Boston, South End, Central, Hyde Park, and Mattapan. A few thousand listings with missing cities or GIS location are assigned to a market using a variable called *area* in the MLS dataset. We end up with 87 markets from 84 cities outside Boston, less 12 small cities plus 15 neighborhoods in Boston.

A4. Sale outcomes

A listing is sold if its reported status is sold or under agreement. There are 2,649 sold listings with missing sales prices. We replace these missing values with their listing price. Listings and sales prices are winsorized at the top 1 percent. For sold properties, the days on market is measured by the difference between the listing date and the sold date.

A5. Distance instruments

We have two distance instruments: distance to the nearest Coldwell Banker office in each year and distance to the nearest Century 21 office in each year. We geocode office locations to obtain latitudes and longitudes. Eighteen Coldwell offices and ten Century 21 offices have missing latitudes and longitudes. We winsorize distances at the top percentile and replace missing distances with the median distance. The IV coefficients are similar if missing distances are not replaced with the median.

A6. Seller fixed effects

We obtain seller names for sold listings from county deed records up to 2008. We merge MLS and deeds data using property address, sale date (within 28 days), and sale price (within $10,000). We are able to fill in seller names for listings that are not merged by tracing the chain of ownership. We assume that when a property is sold, the buyer in that transaction becomes the seller of subsequent MLS listings of the same property, until the next change in ownership. Likewise, the seller of a property remains the same through different listings until the property is sold.

A7. Cumulative days on market

We define *cumulative days on market* by combining unsold listings for the same property into the same *marketing history*. For example, if we see a listing for a property on January 1st 2001 that was withdrawn on June 30th 2001, but re-listed on December 1st 2001 and sold on February 1st 2002, we combine these two listings and calculate the *cumulative days on market* as the difference between the initial listing date and the final date when the property is off the market (the cumulative days on market is $365 + 31 = 396$ days in this example). To belong to the same marketing history, listing dates have to be less than one year apart. Using the same example, if the property was also listed on January 1st 1998 and was withdrawn on June 30th 1998, we do not combine this 1998 listing with the 2001 listing.

A8. Full list of controls for transaction-level analyses

Property controls

-Square footage in thousands of square feet (0 to 40+)

-10 dummies for number of bedrooms, including a dummy for missing values

-14 dummies for number of bathrooms in half bath increments

-9 dummies for number of other types of rooms

-9 dummies for groups of years (6-10 years, 11-25 years, and so on up to 151+ years, plus a dummy for missing age values. The omitted group is 0 to 5 years

-1 if property type is multifamily, 0 otherwise. The omitted group is condominiums

-1 if property type is singlefamily, 0 otherwise. The omitted group is condominiums

-Lot size in acres

-Master bathrooms: 1 if yes, 0 if no

-Finished basement is included in sqft estimation

-1(Beach front), 1(Water front)

-Availability of adult community

-Basement: 1 if yes, 0 if no

-4 dummies: 0, 1, 2 or 3 fireplaces, 99 (missing)

-Entry only: Listing agent's only service is to enter property info into MLS

-Lender owned

-Seller disclosure

-Short sale with lender approval required

-Sub-agency relationship offered

-9 dummies for types of listing agreement, including Exclusive Right to Sell with Named Exclusion, Exclusive Agency, Exclusive Right To Sell With Variable Rate of Commission, Exclusive Right To Sell With Dual Rate of Commission, Facilitation/Exclusive Right To Sell, Facilitation/Exclusive, Facilitation/Exclusive Right To Sell With Variable Rate of Commission, Missing information

-14 dummies for different types of showing methods

-Dummies for the following phrases: Needs Updating, Estate Sale, Foreclosure, Handyman, As-Is, Needs Tlc, Rehabber'S, Bank-Owned, Priced For A Quick Sale, Motivated, Potential, Youthful, Close, !, New, Spacious, Elegance, Beautiful, Appealing, Renovated, Remodeled, Vintage, State-Of-The-Art, Maintained, Wonderful, Brandnew, Fantastic, Charming, Stunning, Amazing, Granite, Immaculate, Breathtaking, Neighborhood, Spectacular, Landscaped, Art Glass, Builtin, Tasteful, Must See, Fabulous, Leaded, Delightful, Move-In, Gourmet, Copper, Corian, Custom, Unique, Maple, Newer, Hurry, Pride, Clean, Quiet, Dream, Block, Huge, Deck, Mint, Stately, Priced To Sell

Listing office controls

-One year lagged fraction of listings sold in a year by an office

-Ln(number of active agents in the Office+1), lagged by one year

-Lagged fraction of agents who are in the top 10 percentile of average annual listings and purchases

-Top 4 office in a market, by average annual number of listings

-1(office has at least 2 entry-only listings or share of exclusive right to sell listings is less than 50%, by office-market). Under exclusive right-to-sell contracts, the listing broker acts as the representative of the seller, and the seller agrees to pay a commission to the listing broker, regardless of whether the property is sold through the efforts of the listing broker.

Listing agent controls

-Whether among top decile of all agents, by average number of listings

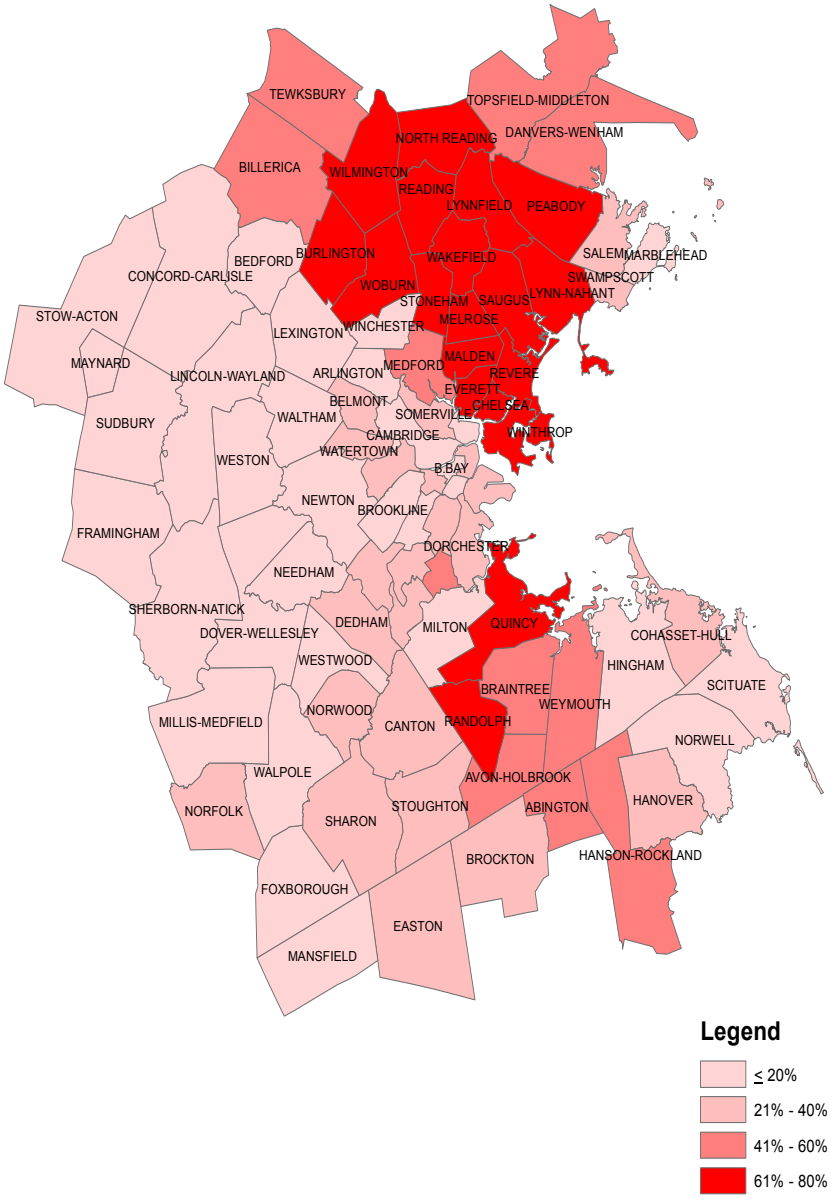
-Agent's average annual number of listings is at least the median amongst listing agents (the median is 2)

-Ln(Cumulative number of listings/purchases by a listing agent, up to the last year)

-Agent's experience in years

APPENDIX: FIGURES

Figure B1. : Percent of listings with low commission rates by market



Notes: Percent of listings in a market with commission rates below 2.5 percent.

APPENDIX: TABLES

C1. Growth paths for low and high commission firms

We refine the comparison in Figure 2 by controlling for firm attributes in the following regression:

$$1(TopRev_{lmt}) = \gamma frcRtL25_{lm,t-1} + X_{lm,t-1}\beta + \mu_{mt} + \varepsilon_{lmt}, \tag{C1}$$

where $1(TopRev_{lmt})$ is 1 if office l's listing commission revenue is in the top quartile in market m and year t, X represents office controls and μ represents market-year fixed effects. The key regressor is $frcRtL25_{lmt}$, the fraction of office l's listings that is below 2.5 percent in the most recent three years $t-2$ to t. Results using a one-year window instead of a three-year window are similar but noisier because some entrants have few listings in a year. The one-year lag of $frcRtL25_{lmt}$ alleviates concerns that it might be jointly determined with the dependent variable. A two-year lag of $frcRtL25_{lmt}$ leads to similar results. Firms' top-quartile status tends to be persistent over time, thus we control for a one-year lagged top status in X (except in the specification with office fixed effects to avoid biases due to the correlation between the residual and the lagged dependent variable). Results without the lagged status are more pronounced.

Table C1 reports estimates of γ for entrants (Panel A) and all offices (Panel B). Column 1 includes market-year fixed effects. Column 2 adds office quality, including the fraction of listings that are sold, average days on market for sold listings, fraction of agents who are the top ten percent highest performing agents, log of the number of active agents, age of the firm in years. Column 3 controls for the composition of an office's listings by adding the fraction of listings that are condominiums, the fraction that are single family, the square footage, number of bedrooms, number of bathrooms, age of the property, and list price, averaged among an office's listings at time t. This mitigates concerns that the weak performance of low commission entrants is driven by their tendency to list properties that deliver lower commission revenues. Column 4 adds office fixed effects.

Across the columns, low commission entrants are significantly less likely to be top-revenue firms, even after adjusting for observable differences among them. Our specification with the most saturated set of controls and office fixed effects suggests that an entrant that specializes in low commissions ($frcRtL25 = 1$) in the past is 12 percentage points (p.p.) less likely to report top-quartile revenues than an entrant that specializes in high commissions ($frcRtL25 = 0$). This effect is considerable given that the mean of the dependent variable is only 17%. When we repeat the analysis using all offices in Panel B, we continue to find much weaker performance for low commission offices. Although not shown, our results are robust to using different measures of dominance (the number of listings, the number of listings and purchases, etc.) and different sample cuts.

Table C1—: Effect of past commission policy on office success

Dependent variable:	**Whether top quartile in market-year**			
	(1)	(2)	(3)	(4)
Panel A: Entrants				
Low comm. offices	-0.08***	-0.06***	-0.05***	-0.12***
	(0.01)	(0.01)	(0.01)	(0.04)
N	6,294	6,294	6,294	6,294
R-squared	0.53	0.57	0.58	0.69
Panel B: All offices				
Low comm. offices	-0.08***	-0.05***	-0.05***	-0.10***
	(0.01)	(0.01)	(0.01)	(0.03)
N	13,255	13,255	13,255	13,255
R-squared	0.62	0.66	0.66	0.72
Market-year FE	Y	Y	Y	Y
Office controls	N	Y	Y	Y
Portfolio controls	N	N	Y	Y
Office FE	N	N	N	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

Notes: This table reports the effect of past commission rate policy on the probability of becoming a top quartile revenue firm, where *revenue* is total listing commission revenue, and *top quartile* is defined by market-year among all offices in that market. Firm *i*'s fraction of listings with a low commission rate at year *t*, *frcRtL25*, is the ratio of the total number of listings under 2.5% to total listings in year *t-2* to year *t*. We only keep firms whose average annual listing is at least five (these are active firms that represent 95% of listings) and firms with two or more firm-year observations. Panel A restricts to entrants, i.e., firms that first appear in our sample in 1999 or later. There are 902 market-year fixed effects in all columns and 1202 office fixed effects in the last column. Panel B uses all firms. There are 1131 market-year fixed effects and 1898 office fixed effects. Office controls and portfolio controls are lagged by a year. Standard errors are clustered at the office level.

C2. Robustness to heterogeneous samples

Table C2 shows that our estimates are stable across different samples. We repeat our main specification in column 6 of Table 3 for all three outcomes.

The results are similar for listings in Boston and listings outside Boston (columns 1 to 2), for condominiums, single-family houses and multi-family properties (columns 3 to 5). The last two columns divide the sample into high and low income markets using the median income in the city from the 2010 census (the results are similar if we use the median income in 2000 or the mean income in 2010).

Table C2—: Robustness checks across different samples for all sales outcomes

	Boston	**not Boston**	**Condos**	**Houses**	**Multifamily**	**High Income**	**Low Income**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Probability of sale							
Low commmission listings	-0.05***	-0.05***	-0.06***	-0.05***	-0.05***	-0.05***	-0.05***
	(0.01)	(0.004)	(0.01)	(0.005)	(0.01)	(0.01)	(0.004)
N	58474	286358	105306	191059	48467	113306	231526
R-squared	0.52	0.51	0.53	0.51	0.53	0.52	0.51
Panel B: Ln(Days on market)							
Low commmission listings	0.14***	0.11***	0.09***	0.13***	0.16***	0.13***	0.12***
	(0.05)	(0.02)	(0.03)	(0.02)	(0.05)	(0.03)	(0.02)
N	18443	118181	38979	82196	15449	48809	87815
R-squared	0.57	0.57	0.59	0.58	0.59	0.58	0.56
Panel C: Ln(Sale price)							
Low commmission listings	0.0005	0.0003	0.0009	-0.0009	0.001	-0.0005	0.0001
	(0.003)	(0.001)	(0.002)	(0.001)	(0.004)	(0.002)	(0.001)
N	18548	118537	39197	82356	15532	48910	88175
R-squared	0.99	0.99	1.00	0.99	0.98	1.00	0.99
Controls in Table 3 column 6	Y	Y	Y	Y	Y	Y	Y

* p<0.1, ** p<0.05, *** p<0.01

Notes: The effect of low commission rates on all three outcomes, by sub-samples. The sub-samples are: Boston only (column 1), outside Boston (column 2), condominiums (column 3), single-family (column 4), multi-family (column 5), high and low income markets (columns 6 and 7, respectively, where cities are split using median income in 2010 from the census). Each column repeats column 6 in Table 3.

C3. Robustness checks using different controls

Table C3 explores robustness to different types of controls. We first explore whether the results change when we use different geographic units to control for market conditions. Column 1 replicates column 6 in Table 3, column 2 uses zipcode-year fixed effects instead of market-year fixed effects, and column 3 uses tract-year fixed effects. In column 4, we add office fixed effects to our main specification.

Table C3—: Robustness checks using different controls

Dependent Variable:	**Probability of sale**			
	(1)	(2)	(3)	(4)
Low commission listings	-0.05***	-0.05***	-0.05***	-0.03***
	(0.003)	(0.003)	(0.004)	(0.004)
N	344832	344832	344832	326054
R-squared	0.51	0.52	0.53	0.54
Market-year FE	Y	N	N	Y
Zipcode-year FE	N	Y	N	N
Tract-year FE	N	N	Y	N
Office FE	N	N	N	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

Notes: Columns 1 to 3 replicate the main OLS specification in column 6 of Table 3 but with different set of controls for market conditions. Column 1 uses 1217 market-year fixed effects (same as Table 3), column 2 uses 3178 zipcode-year fixed effects, and column 3 uses 9030 tract-year fixed effects. Column 4 adds 2239 listing office fixed effects to our main OLS specification. This analysis only includes offices with average annual listings at or above 5 and drops 18,778 listings.

C4. Two-way clustering of standard errors

Table C4 shows that our main results are robust to a two-way clustering of standard errors by property and year (Cameron, Gelbach and Miller, 2011).

Table C4—: Robustness to two-way clustering of standard errors

Dependent variable:	**Pr(Sold)**	**Ln(Days on market)**	**Ln(Sale price)**
	(1)	(2)	(3)
Low commission listings	-0.05***	0.12***	0.0003
	(0.005)	(0.01)	(0.0008)
N	344832	136624	137085
Controls in Table 3 column 6	Y	Y	Y

* p¡0.1, ** p¡0.05, *** p¡0.01
Notes: Repeats column 6 of Table 3, but cluster standard errors by property and year.

C5. Robustness to right censoring for probability of sale

Here we address concerns that the probability of sale regression is affected by right censoring in the *sold* dummy (some listings in 2011 are sold after our sample period ends). We repeat our probability of sale analysis using whether a listing is sold within 30, 60, 90, and 180 days of the listing date as alternative dependent variables. We also experiment with dropping properties that are listed after 2009. Our conclusions are similar in all cases.

Table C5—: Probability of sale within 30, 60, 90, 180 days

Dependent variable:	**Sold within:**			
	30 Days	60 Days	90 Days	180 Days
	(1)	(2)	(3)	(4)
Low commission listings	-0.03***	-0.05***	-0.06***	-0.06***
	(0.003)	(0.003)	(0.003)	(0.003)
N	344832	344832	344832	344832
R-squared	0.51	0.52	0.53	0.52

* p¡0.1, ** p¡0.05, *** p¡0.01
Notes: The controls are the same as in column 6 of Panel A of Table 3. The dependent variable for each column is whether the listing is sold within 30 days, 60 days, 90 days, and 180 days, respectively.

C6. Robustness to probit for probability of sale

Next, Table C6 shows that our probability of sale results are robust to using probit instead of OLS. Our probit analysis resembles the OLS analysis in Panel A of Table 3, except we do not include property fixed effects and use all listings instead of repeat listings only. Our STATA program of probit with property fixed effects does not converge despite numerous attempts. Our most saturated probit specification in column 5 controls for market-year and month fixed effects, as well as the full set of 148 property controls, seller patience, office and agent controls.

Table C6—: Effect of low commission on probability of sale using probit

Dependent variable:	**Probability of sale**				
	(1)	(2)	(3)	(4)	(5)
Low commission listings	-0.09***	-0.07***	-0.07***	-0.05***	-0.05***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
N	653475	653475	653475	653475	653475
Market-year FE, month FE	Y	Y	Y	Y	Y
Property controls	N	Y	Y	Y	Y
Seller patience	N	N	Y	Y	Y
Office controls	N	N	N	Y	Y
Agent controls	N	N	N	N	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

C7. Selection correction for sold listings

Table C7a and Table C7b repeat the analyses for the effects on days on market and sale price for sold listings, using selection correction methods to address the concern these two outcomes are unobserved for properties that do not sell.

Panel A implements the Heckman (1979) selection correction method. We first estimate a probit model with the sold dummy as the dependent variable and the full sample of 653,475 listings. Our controls for the probit estimation include market-year and month fixed effects, the full set of 148 property controls, seller patience, office and agent controls. We do not include property fixed effects. We then construct the inverse Mills ratio using our probit estimation and include it as a control in our sale price and days on market regressions.

Panel B controls for the selection bias non-parametrically using fixed effects to relax the distributional assumption that the error terms in the outcome and selection equations are jointly Normally distributed. We first estimate the same probit model and predict the probability of sale. We then create dummies for each decile of the predicted probability of sale and include these decile fixed effects in our outcome regressions. In both cases, the results are similar to those in Table 3.

Table C7a—: Selection correction for effects on days on market

Dependent Variable:	**Ln(Days on Market)**					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Inverse Mills Ratio					
Low commission listings	0.10***	0.10***	0.12***	0.12***	0.12***	0.12***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
N	419116	419116	136624	136624	136624	136624
R-squared	0.11	0.14	0.56	0.56	0.57	0.57
	Panel B: Decile bins for selection probability					
Low commission listings	0.10***	0.10***	0.12***	0.12***	0.12***	0.12***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
N	419116	419116	136624	136624	136624	136624
R-squared	0.12	0.14	0.56	0.57	0.57	0.57
Market-year FE, month FE	Y	Y	Y	Y	Y	Y
Property controls	N	Y	Y	Y	Y	Y
Property FE	N	N	Y	Y	Y	Y
Seller patience	N	N	N	Y	Y	Y
Office controls	N	N	N	N	Y	Y
Agent controls	N	N	N	N	N	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

Table C7b—: Selection correction for effects on sale price

Dependent Variable:	**Ln(Sale Price)**					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Inverse Mills Ratio					
Low commission listings	0.02***	-0.01***	0.003*	-0.0005	0.0003	0.0003
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
N	421329	421329	137085	137085	137085	137085
R-squared	0.53	0.88	0.98	0.99	0.99	0.99
	Panel B: Decile bins for selection probability					
Low commission listings	0.02***	-0.01***	0.004**	-0.0004	0.0003	0.0003
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
N	421329	421329	137085	137085	137085	137085
R-squared	0.53	0.88	0.98	0.99	0.99	0.99
Market-year FE, month FE	Y	Y	Y	Y	Y	Y
Property controls	N	Y	Y	Y	Y	Y
Property FE	N	N	Y	Y	Y	Y
Seller patience	N	N	N	Y	Y	Y
Office controls	N	N	N	N	Y	Y
Agent controls	N	N	N	N	N	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

C8. Regressions with seller fixed effects

Next, Table C8 repeats the seller fixed effect regressions, but drops properties listed after 2008. Our county deeds data with seller names end in 2008. The analysis in the paper with seller fixed effects (Table 5) includes listings from 2008 to 2011 for which we could trace the seller names. To address the concern that we might mismatch sellers to listings after 2008, we repeat the seller fixed effect analysis using only listings between 1998 and 2008. The results are similar to those reported in Table 5.

Table C8—: Seller fixed effect regressions, 1998-2008

Dependent variable:	**Probability of sale**	
	(1)	(2)
Specification:	Seller name	No common names
Low commission listings	-0.07***	-0.07***
	(0.02)	(0.02)
N	30597	29333
R-squared	0.48	0.49
Controls in Table 5 column 3	Y	Y

* p¡0.1, ** p¡0.05, *** p¡0.01

C9. The effect of commission rates on cumulative days on market

Table C9 investigates the distributional effect of low commission rate on cumulative days on market. The dependent variable is the cumulative days on market between the first listing date and the sold date. The key regressor is whether the initial listing for the entire marketing history is strictly below 2.5 percent. Column 1 replicates column 6 in Panel B of Table 3 using the cumulative days on market. Columns 2 to 7 report quantile regressions for the 25th, 50th, 75th, 90th, 95th, and 99th percentiles, respectively. The controls for the quantile regression are similar to those in column 6 of Table 3, except we use market fixed effects plus year fixed effects instead of market by year fixed effects and we drop property fixed effects.

Table C9—: The effect of commission rate on cumulative days on market

Dependent variable:	**Cumulative days on market**						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Initially low commission	20.22***	2.71***	6.68***	13.21***	23.92***	29.78***	38.42***
	(1.58)	(0.19)	(0.36)	(0.72)	(1.33)	(1.96)	(5.00)
N	137081	417887	417887	417887	417887	417887	417887
Statistic	Mean	25th	50th	75th	90th	95th	99th

* p¡0.1, ** p¡0.05, *** p¡0.01

Notes: This table reports results on *cumulative days on market* discussed in Section V. Column 1 replicates our main OLS specification. We drop 4 repeat sales with outliers (days on market below -90 or above 1500 days). Standard errors are clustered at the property level. The quantile regressions in columns 2 to 7 use all sold listings (not just repeat sales), but drops the 9 outliers and 3,433 sales with missing property identifiers (we need property identifiers to cumulate days on market for each property).

APPENDIX: REDUCED NUMBER OF BUYERS

This section explains the back-of-the-envelope calculation discussed in Section IV.B. We use the estimate in column 5 of Table 7 (-0.04) to calculate the reduction in the number of potential buyers visiting a low commission property as a result of large offices steering buyers to high commission properties. The six dominant chains account for 54% of buyers. The average market share for offices affiliated with these chains is 17%, which is 2.8 times bigger than that for non top-chain offices (6%). At an elasticity of 0.04, this translates to a 6 p.p. reduction (54%*2.8*.04) in the number of potential buyers visiting low commission properties.

According to NAR (2014), a listing is visited by on average ten potential buyers. We make the simplifying assumptions that the matching event between a potential buyer and a seller is i.i.d. across individuals, and that the successful match rate is identical across properties and individuals. Suppose the probability that a listing matches with a potential buyer is x, then the probability that a listing is sold is 1 minus the probability that all of the ten potential matches fail, which is $1-(1-x)^{10}$. On average, 64.7% of all listings are sold, implying x is 9.9%.

From a base of ten potential buyers, a 6 p.p. reduction in the number of buyers lowers the likelihood of being sold to 62.5% (which is $1-(1-9.9\%)^{9.4}$). This accounts for about 40% of the 5 p.p. reduction in the sale probability that is documented in the paper. The magnitude is similar when the number of potential buyers is assumed to vary between five and twenty.

*

REFERENCES

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47: 153–161.
NAR

NAR. 2014. "2014 Profile of Home Buyers and Sellers: Highlights." *Accessed online, available at: http://www.realtor.org/reports/highlights-from-the-2014-profile-of-home-buyers-and-sellers*.