NBER WORKING PAPER SERIES

MISINFORMED SPECULATORS AND MISPRICING IN THE HOUSING MARKET

Alex Chinco
Christopher Mayer

Working Paper 19817
http://www.nber.org/papers/w19817

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2014

Vivek Sampathkumar, Daniel Hubbard, Laura Vincent, and James Witkin provided dedicated research assistance. We thank Viral Acharya, Xavier Gabaix, Ed Glaeser, Robin Greenwood, Gur Huberman, Aurel Hizmo, Andrew Paciorek (discussant), Tomasz Piskorski, Alexi Savov, Chester Spatt (discussant), David Sraer (discussant), Jeremy Stein, Johannes Stroebel (discussant), Stijn Van Nieuwerburgh, Reed Walker, and Jeff Wurgler as well as seminar participants at Harvard Business School, NYU Stern, Columbia Business School, the NBER Behavioral Finance/Housing Bubbles joint session, the NBER Conference Project on Housing and the Financial Crisis, the Federal Reserve Bank of Cleveland, AEA/ASSA, and Riksbank for extremely helpful comments and suggestions. We are grateful for Dataquick, 1010data, and an anonymous firm for providing data and assistance that were crucial for this project. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Misinformed Speculators and Mispricing in the Housing Market
Alex Chinco and Christopher Mayer
NBER Working Paper No. 19817
January 2014
JEL No. G14,R3

ABSTRACT

This paper uses transactions-level deeds records to examine how out-of-town second house buyers contributed to mispricing in the housing market. We document that out-of-town second house buyers behaved like misinformed speculators and drove up both house price and implied-to-actual rent ratio (IAR) appreciation rates in cities like Phoenix, Las Vegas, and Miami in the mid 2000s. Our analysis has 3 parts. First, we give evidence that out-of-town second house buyers behaved like misinformed speculators. Compared to local second house buyers, out- of-town second house buyers had worse exit timing (i.e., were likely misinformed) and were also less able to consume the dividend from their purchase (i.e., were likely speculators). Second, we show that increases in out-of-town second house buyer demand predict increases in future house price appreciation rates and IAR appreciation rates. A 10%pt increase in the fraction of sales made to out-of-town second house buyers is associated with a 6%pt increase in house price appreciation rates and a 9%pt increase in IAR appreciation rates over the course of the next year in that city. Third, we address the issue of reverse causality using a novel econometric strategy. The key insight is that an increase in the fundamental value of owning a second house in Phoenix is a common shock to the investment opportunity set of all potential second house buyers. If changes to fundamentals were driving both price dynamics as well as out-of-town second house buyer demand, we would expect to see large jumps in house price and IAR appreciation rates preceded by increases in out-of-town second house buyer demand from across the country. The data do not display this symmetric response, and are thus inconsistent with reverse causality. We conclude by discussing both the economic magnitudes of out-of-town second house buyer flows and the broader applicability of our econometric approach.

Alex Chinco
University of Illinois
College of Business
1206 S 6th St
Rm: Wohlers 406
Champaign, IL 61820
AlexChinco@gmail.com

Christopher Mayer
Columbia Business School
3022 Broadway, Uris Hall #805
New York, NY 10027
and NBER
cm310@columbia.edu

1. Introduction

Researchers have long puzzled over the inefficiency of housing prices. Starting with the seminal paper by Case and Shiller (1989), analysts have documented the predictability of home prices: positive short-run serial correlation and negative serial correlation in the long-run. For a broader discussion of the literature on mispricing in the residential housing market see Mayer (2010). More recently, Glaeser, Gottlieb, and Gyourko (2010) have pointed out that fundamental factors like declining interest rates, income growth, and increased credit access cannot explain the price run-ups in many markets over the previous decade. An additional challenge to models based only on fundamentals is the fact that the largest price increases occurred in housing markets like Las Vegas and Phoenix where land supply is quite elastic. Thus researchers have increasingly turned to examining the role of speculation as an additional factor in explaining the strong increase in housing prices in certain markets.

This paper uses transactions-level deeds records to document that out-of-town second house buyers behaved like misinformed speculators, driving up both house price and implied-to-actual rent ratio (IAR) appreciation rates in cities like Phoenix, Las Vegas, and Miami in the mid 2000s. In the process we employ a novel econometric approach to address concerns about reverse causality. Our data list both the property address as well as the tax bill mailing address for every single family house purchase in 21 cities during the period from January 2000 to December 2007. We use this information to classify each single family house purchase as made by either an owner occupant, a local second house buyer, or an out-of-town second house buyer. A local second house buyer is someone who buys a second house in Phoenix and also lives in Phoenix. An out-of-town second house buyer is someone who buys a second house in Phoenix but lives year-round in San Francisco.

Our analysis has 3 parts. First, in Section 3 we give evidence that out-of-town second house buyers behaved like misinformed speculators. We start by showing that out-of-town second house buyers had worse exit timing than local second house buyers—i.e., were likely misinformed. Out-of-town second house buyers realized roughly 6% per year lower capital gains on their purchases than local second house buyers who bought at the exact same time in the exact same city as illustrated

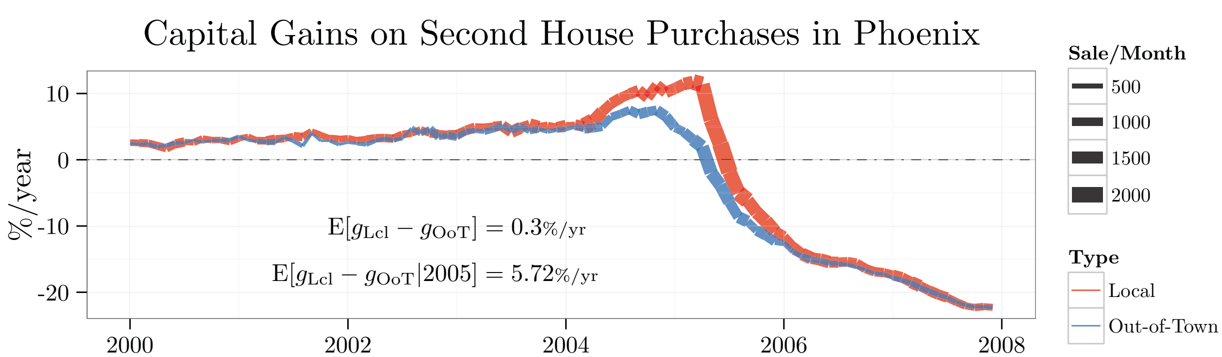


Figure 1. *Line height: Average capital gains on second house purchases made in Phoenix each month split by buyer type (local vs. out-of-town) in units of percent per year. Line width: Number of purchases made each month in Phoenix by each buyer type. Sample: January 2000 to December 2008. Reads: "While local second house buyers who bought in March 2005 earned nearly an 11% per year capital gain on their purchase, out-of-town second house buyers who bought in the same month earned only a 5% per year capital gain on these purchases."*

in Figure 1. Then, we show that out-of-town second house buyers were less able to consume the housing dividend from their purchase due to their distance from the house. A typical property manager charges a fee of 1 month's rent plus an additional 8% of the annual rent each year to lease a house and manage relations with the tenant. As a result, these buyers must have relied more on price appreciation in their ex ante return calculations—i.e., were likely speculators. Finally, we show that out-of-town second house buyers don't appear to have fundamentally different preferences than local second house buyers. These buyers weren't just the super rich since the median value of their primary residences was actually lower than that of the median sale price in many cities. Moreover, at peak the typical out-of-town second house buyer in places like Los Angeles and Jacksonville owned 2 investment properties making diversification an unlikely motivation.

Second, in Section 4 we show that increases in out-of-town second house buyer demand predict increases in future house price and IAR appreciation rates. A 10%pt increase in the fraction of sales made to out-of-town second house buyers is associated with a 6%pt increase in house price appreciation rates and a 9%pt increase in IAR appreciation rates over the next year in that city as depicted in Figure 2. By contrast, increases in local second house buyer demand are not associated with increases in either house price or IAR appreciation rates. Of course, in order to make statements about mispricing, we must pick a pricing model. In this paper we subscribe to the user-cost model and use deviations of city-wide log IAR ratios from zero as a proxy for mispricing. Concretely, the log of the IAR ratio as calculated as in Himmelberg, Mayer, and Sinai (2005) represents the excess return over the apartment rental rate of borrowing money, buying a house, living in it for 1 year, and then selling it in a year's time. The higher the log IAR ratio is, the cheaper it is to rent rather than buy, and the higher the level of mispricing on the rent-vs-own margin. One of the main criticisms of this model is that expected house price appreciation rates are an *input*. To mitigate this concern we show that our empirical results do not depend on the gritty details of how this expectation is calculated. In summary, Sections 3 and 4 show that out-of-town second house buyers behaved like misinformed speculators in the mid 2000s, and increases in demand from these traders are associated with higher house price and IAR appreciation rates in subsequent months.

Section 5 houses the third and final component of our analysis. In this section we ask the natural follow up question: "Is this just reverse causality? Is it just that changes in fundamentals are both attracting out-of-town second house buyers and also driving up prices?" No. Our results suggest this is not the case. The key insight is that an increase in the fundamental value of owning a second house in Phoenix represents a common shock to the investment opportunity set of all potential second house buyers. Thus, if changes to fundamentals were driving both price dynamics as well as out-of-town second house buyer demand, we would expect to see large jumps in house price and IAR appreciation rates preceded by increases in out-of-town second house buyer demand from across the country. The data do not display this symmetric response. As a result, they are inconsistent with an explanation based on reverse causality. To test this prediction econometrically, we condition our regressions on city size. For instance, we find that a 10%pt increase in demand for second houses in Phoenix by buyers living in Los Angeles predicts a bigger increase in Phoenix house price and IAR appreciation rates than a 10%pt increase in demand by buyers living in Milwaukee. We include month and ordered city pair fixed effects in our specification to account for variation in

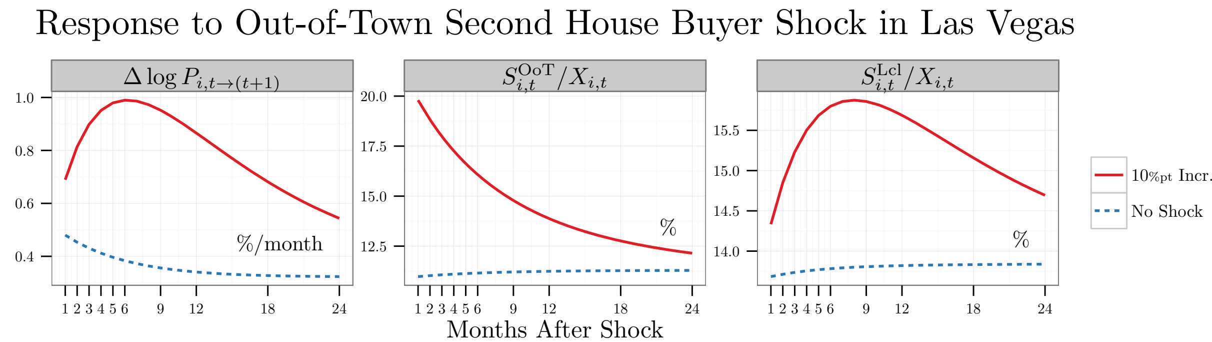


Figure 2. *Response of house price appreciation rates, out-of-town second house buyer demand, and local second house buyer demand both to (red, solid) a* $10\%_{\mathrm{pt}}$ *increase in the fraction of sales made to out-of-town second house buyers in Las Vegas as well as to (blue, dashed) no shock. We compute the figure using the panel VAR in Equation* (9) *whose point estimates calculated over the time period from January 2000 to December 2007 are housed in Table 8. Reads: "In the* 12 *months following a* $10\%_{\mathrm{pt}}$ *increase in the fraction of purchases made by out-of-town second house buyers in Las Vegas house price appreciation rates will rise by roughly* $6\%_{\mathrm{pt}}$*."*

macroeconomic conditions and the fact that potential second house buyers in San Francisco are more likely to buy a in Las Vegas than Cleveland.

A common first response to this specification is: "Of course! The $10\%_{\mathrm{pt}}$ increase in demand from the bigger city should have a larger price impact. Isn't this tautological?" No. It is not, though the logic is subtle. Imagine two groups of potential second house buyers living in Los Angeles and Milwaukee respectively that are each thinking about buying into Phoenix, and suppose that there are 10 times as many potential buyers living in Los Angeles, 10000, as there are living in Milwaukee, 1000. First, consider the null hypothesis, h_0, that *a common shock to the fundamental value of owning a second house in Phoenix is both attracting out-of-town second house buyers and driving up prices*. In this world, potential second house buyers in both Los Angeles and Milwaukee see the same shock (the shock to fundamentals), and will always increase their demand by $10\%_{\mathrm{pt}}$ at the exact same time. As a result, Phoenix will always see aggregate out-of-town second house buyer demand shocks of either 1100 buyers or none at all. Here is the punchline. Under the null hypothesis, the relative sizes of Los Angeles and Milwaukee are irrelevant. It doesn't matter how out-of-town second house buyers are distributed across the country because they are responding to a common Phoenix-specific shock. Now, consider the alternative hypothesis, h_A, that *shocks to fundamentals in Phoenix are not the only thing attracting out-of-town second house buyers*. For instance, the Milwaukee Journal Sentinel might have run a glowing review of Phoenix as a winter getaway, or Home and Garden Television might have played a Phoenix-based episode of Flip That House in Los Angeles cable market. In this world, potential second house buyers in Los Angeles and Milwaukee will not in general see the same shocks. As a result, Phoenix may see a 1000 buyer demand shock from Los Angeles in one month and then a 100 buyer demand shock from Milwaukee in the next. It's only in this alternative world that the $10\%_{\mathrm{pt}}$ increase in demand from the bigger city has a larger price impact. It's only in this alternative world that the geographic distribution of potential second house buyers matters.

We conclude in Section 6 by discussing both the economic magnitudes of out-of-town second house buyer capital flows as well as the broader applicability of our econometric approach. We add up the sales prices of the houses purchased by these buyers in each city and compare this value to the city's GDP. This calculation builds on the analogy between purchases by out-of-town second house buying and foreign direct investment. The results are large. In 2004 out-of-town second house buyers spent in excess of 5% of the GDP of Las Vegas which is on par with the size of foreign direct investment in many developing countries à la Javorcik (2004). Our econometric approach allows us to identify price movements that can't be solely explained by shocks to fundamentals. Importantly, it completely sidesteps the thorny problem of instrumenting for misinformed speculator demand across assets. It applies wherever there is market segmentation and data on flows between segments.

This paper primarily builds on 3 strands of the finance literature. First, there is a recent string of real estate finance papers studying second house buyers. Bayer, Geissler, and Roberts (2011) analyze the behavior of second house buyers in Los Angeles who buy and sell small numbers of houses and try unsuccessfully to time the market. These authors find that speculator trading behavior is strongly associated with neighborhood price instability. Haughwout, Lee, Tracy, and van der Klaauw (2011) examine credit report data and show that mortgages on second houses represented nearly half of all mortgages originated in the 4 states with the highest price appreciation rates at the peak of the market. Li and Gao (2012) present theoretical results that second house buyers can fuel a boom as well as empirical evidence showing that second house buyers are both more likely to be present in cities with high house price appreciation and also more likely to subsequently default at higher rates. The results in these papers are complimentary to ours in that all of these papers document the large growth in second house purchases in the highest appreciating cities; however, none of the existing papers is able to directly address the issue of reverse causality. This argument involving reverse causality is commonly referred to as the Friedman Critique and dates to Friedman (1953). Note that buying a second house is just one way of speculating in the housing market. e.g., Choi, Hong, and Scheinkman (2013) show that owner occupants likely over-invested in home improvements.

Second, our work extends the analysis above to differentiate between local second house buyers and out-of-town second house buyers. Importantly, we find that only the purchases of the latter group lead to mispricing. Thus, this paper connects to the existing work on the relationship between distance and investment performance. Within the context of the real estate market Kurlat and Stroebel (2014) give a theoretical model of markets with many heterogeneous assets and differentially informed agents. However, this effect applies more broadly. e.g., Cohen, Frazzini, and Malloy (2010) study effect of a sell side equity analyst's social network on her ability to gather information. They find that analysts who are more distant from a firm in terms of having gone to school with fewer of the senior management team underperform their peers. In a similar vein, Hong, Kubik, and Stein (2005) gives evidence that mutual fund managers are more likely to buy a particular stock if other managers in the same city are buying it as well.

Third and finally, all of these papers are couched in a broader behavioral finance literature on price formation and noise traders. Beginning with the work of Black (1986) and De Long, Shleifer, Summers, and Waldmann (1991), many authors have conjectured that the trading behavior of misinformed speculators can cause prices to deviate from their fundamentals. According to these models other traders may not be able to restore equilibrium because of limits to arbitrage such

Deeds Data Structure

	Property Address	Tax Bill Address	Price	Date
1	1 Telegraph Hill Blvd, SF	1 Telegraph Hill Blvd, SF	$151	04/15/2002
2	200 Fremont St, LV	888 W Bonneville Ave, LV	$154	10/20/2003
3	200 Fremont St, LV	709 N La Brea Ave, LA	$300	05/01/2006

Table 1. *Fictitious observations illustrating the structure of the deeds data. The columns display the reported property address, tax bill mailing address, sales price, and sales date. Row 1 represents a purchase by an owner occupant, row 2 represents a purchase by a local second house buyer and row 3 represents a purchase by an out-of-town second house buyer. Reads: "A house at 200 Fremont St. in Las Vegas sold for $300k to someone living at 709 N La Brea Ave. in Los Angeles in May 2006 would be classified as an out-of-town second house purchase."*

as capital constraints, informational frictions, or a limited supply of tradable shares. See Shleifer and Vishny (1997), Scheinkman and Xiong (2003), and Ofek and Richardson (2003) respectively for examples. The interaction of traders who are not fully rational and constrained arbitrageurs can explain common asset pricing puzzles such as momentum—e.g., see Hong and Stein (1999). The residential housing market is a natural laboratory for studying behavioral finance questions about the motivations of traders as the data have additional variables that extend the reach most financial datasets. e.g., in the stock market traders are generally anonymous. By comparison, housing markets offer a wealth of information about traders via deeds records, mortgage records, tax rolls, etc...

2. Data Description

To perform our analysis, we use data drawn from three main sources: county deeds records obtained from Dataquick and an anonymous data provider, HPI data from Zillow, Inc. and IAR data computed according to the procedure developed in Himmelberg, Mayer, and Sinai (2005). Subsections 2.1, 2.2 and 2.3 describe each of these data sources and present summary statistics. Once cleaned, our data represents 21 MSAs indexed by $i = 1, 2, \ldots, I$ over the time period $t = 1, 2, \ldots, T$ with $t = 1$ denoting January 2000 and $t = T$ denoting December 2007.

2.1. Transaction Level Deeds Records.

We use transactions-level deeds records to classify single family house purchases as made by either owner occupants, local second house buyers, or out-of-town second house buyers. A deed is a written legal instrument that passes the rights to a particular property (in our case a single family house) from one owner to the next. The deeds records are public in most states due to information disclosure acts and are maintained by the local county. Deeds records document any time a property is sold or a new mortgage is taken out by an owner using the property as collateral. Together, these data contain a complete sales history of any parcel of land. Below, we define variables denoting the number of sales in an MSA in a given month.

Definition 1 (Sales)**.** *Define $X_{i,t}$ as the annualized number of single family houses sold in MSA i at month t in units of houses per year.*

One advantage of using the US residential real estate market to study asset pricing questions is that we can obtain information on all buyers and sellers via county deeds records. Namely, for each

Out-of-Town Second House Purchases as a % of Monthly Sales

	Mean	Sd	Min	Q25	Q50	Q75	Max
Baltimore	4.76	1.94	2.34	3.19	4.34	6.32	9.69
Charlotte	3.33	2.32	0.497	1.39	2.38	5.88	7.80
Cincinnati	6.27	1.18	2.78	5.82	6.29	6.84	9.46
Cleveland	5.37	0.942	3.03	4.87	5.38	5.98	7.49
Denver	2.20	1.28	0.676	1.08	1.70	3.29	5.54
Jacksonville	5.92	2.62	2.20	3.82	4.97	7.51	12.3
Las Vegas	11.0	3.83	4.67	7.03	12.0	14.5	17.1
Los Angeles	1.15	0.437	0.224	0.877	1.13	1.40	2.19
Miami	4.59	1.52	2.02	3.23	4.41	5.88	7.39
Milwaukee	1.28	0.599	0.193	0.857	1.19	1.63	2.94
Minneapolis	1.38	0.813	0.177	0.708	1.32	2.00	3.39
Orlando	9.86	3.41	3.16	7.48	9.99	12.4	15.7
Philadelphia	2.63	1.29	0.757	1.46	2.57	3.56	5.58
Phoenix	7.67	2.95	3.58	5.52	6.73	9.40	15.5
Riverside	8.33	1.30	5.62	7.41	8.23	9.41	11.4
Sacramento	6.49	0.951	4.31	5.88	6.56	7.37	8.28
San Diego	3.07	1.46	1.38	1.87	2.57	4.02	7.51
San Francisco	2.33	0.393	1.60	2.06	2.31	2.53	3.65
San Jose	1.86	0.451	0.693	1.56	1.83	2.13	3.06
Tampa	7.74	2.49	3.66	5.71	7.34	9.86	12.5
Washington	1.35	0.426	0.591	1.05	1.24	1.63	2.52
Mean	4.57	1.50					

Table 2. *Percent of single family house purchases made by out-of-town second house buyers in each MSA i in each month t over the time interval from January 2000 to December 2007. The shaded region in each sparkline plot covers the interquartile range for each MSA and is not a constant scale. Reads: "Over the entire sample period,* 5.92% *of all sales were made to out-of-town second house buyers each month in Jacksonville, FL; however, the fraction of sales made to out-of-town second house buyers was more than double this number at its peak in early 2006."*

property transaction in our database, we observe not only an address for the property itself but also a mailing address where the county sends the tax bill for the property. Below, we define variables denoting the identity of various type of house purchasers in an MSA in a given month.

Definition 2 (Second House Purchases)**.** *Define $S_{i\to j,t}$ as the annualized number of single family houses sales in MSA j at month t where:*

(1) The mailing address of the tax bill and the property address recorded in the deeds records do not match, and

(2) The mailing address is located in an MSA i.

$S_{i\to j,t}$ has units of houses per year.

Definition 3 (Out-of-Town Second House Purchases)**.** *Define $S^{\text{OoT}}_{j,t} = \sum_{j\neq i} S_{i\to j,t}$ as the annualized number of second house purchases in MSA j at month t where the mailing address is located in an MSA i with $j \neq i$. $S^{\text{OoT}}_{i,t}$ has units of houses per year.*

Definition 4 (Local Second House Purchases)**.** *Define $S^{\text{Lcl}}_{j,t} = S_{j\to j,t}$ as the annualized number of*

Local Second House Purchases as a % of Monthly Sales

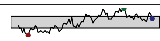
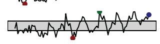
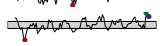
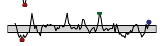
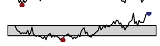
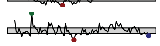
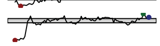
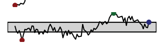
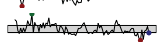
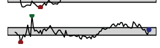
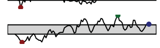
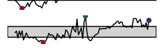
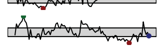
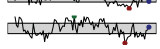
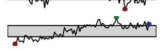
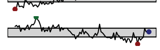
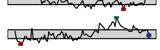
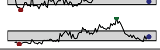
		Mean	Sd	Min	Q25	Q50	Q75	Max
Baltimore		13.1	3.14	7.12	10.3	13.6	15.5	18.9
Charlotte		9.49	1.46	6.72	8.38	9.38	10.5	12.6
Cincinnati		12.3	1.77	6.87	11.1	12.2	13.6	16.7
Cleveland		10.5	1.64	6.84	9.19	10.4	11.4	15.2
Denver		9.94	2.42	6.41	7.99	9.17	11.7	16.2
Jacksonville		17.0	1.73	13.1	15.8	16.9	18.0	24
Las Vegas		12.8	3.36	7.30	10.3	12.8	14.7	19.7
Los Angeles		10.3	2.29	2.95	9.92	10.8	11.6	13.5
Miami		14.6	1.95	10.8	13.2	14.3	16.1	18.8
Milwaukee		10.1	1.96	5.46	8.63	9.82	11.6	16.3
Minneapolis		13.3	4.32	5.82	9.04	14.1	16.5	22.6
Orlando		15.9	2.16	10.9	14.1	15.9	17.3	22.7
Philadelphia		16.0	3.08	9.99	13.8	16.4	18.1	22.6
Phoenix		16.2	2.60	11.9	13.9	16.1	18.1	22.1
Riverside		10.4	1.08	8.38	9.58	10.2	11.2	13.3
Sacramento		11.6	1.37	8.65	10.6	11.5	12.7	14.1
San Diego		12.7	2.41	7.78	10.5	13.5	14.6	17.7
San Francisco		9.97	1.51	6.70	8.66	10.0	11.1	14.2
San Jose		8.05	1.92	5.03	6.78	7.68	8.87	15.2
Tampa		17.7	2.55	12.9	15.4	17.3	19.3	24.9
Washington		8.98	1.78	6.52	7.51	8.76	10.2	13.9
Mean		13.8	2.32					

Table 3. *Percent of single family house purchases made by local second house buyers in each MSA i in each month t over the time interval from January 2000 to December 2007. The shaded region in each sparkline plot covers the interquartile range for each MSA and is not a constant scale. Reads: "Over the entire sample period,* 16.0% *of all sales were made to local second house buyers each month in Philadelphia, PA. At peak, this number climbed to* 22.6%*."*

second house purchases in MSA j at month t where both the mailing and property addresses are located in MSA j. $S_{j,t}^{\mathrm{Lcl}}$ has units of houses per year.

Table 1 gives an example of an owner occupant, a local second house buyer, and an out-of-town second house buyer in our data. In the mid 2000s, the number of purchases by out-of-town second house buyers in MSAs like Las Vegas, Miami, and Phoenix grew appreciably relative to their level at the beginning and end of our sample period. Table 2 which gives summary statistics for the number of out-of-town second house purchases in each MSA i as a fraction of the total number of sales in MSA i each month. At peak, out-of-town second house purchases always represented a minority of house purchases. In the most extreme market, Las Vegas, out-of-town second house buyers purchased 17% of all sales in 2004, up from roughly 7% in the early 2000s. Many of these MSA specific sparkline plots display a similar hump-shaped pattern in the number of out-of-town second house purchases as a fraction of sales. A key insight for our analysis is that the scale of the patterns differs by orders of magnitude. For example, while both Miami and Milwaukee show similar percent change rises in the fraction of all houses bought by out-of-town second house buyers from 2002 to 2006, at the peak of the housing boom Miami had around 3 times as large a fraction

House Price Appreciation Rates in % per Year

		Mean	Sd	Min	Q25	Q50	Q75	Max
Baltimore		6.49	9.31	−12.3	−1.77	10.1	12.0	21.4
Charlotte		0.951	3.43	−6.52	−1.56	1.30	3.87	6.39
Cincinnati		0.048	2.39	−6.32	−1.05	0.796	1.70	3.23
Cleveland		−1.87	4.23	−11.2	−4.98	−0.322	1.09	4.48
Denver		−0.309	4.40	−9.61	−2.63	−0.237	0.976	11.3
Jacksonville		4.59	9.94	−19.3	2.26	8.20	10.2	17.0
Las Vegas		3.69	18.0	−34.5	−3.66	5.10	7.83	44.3
Los Angeles		6.21	14.9	−28.6	−4.49	10.9	16.6	27.6
Miami		6.48	16.7	−31.1	−1.66	12.9	14.9	27.3
Milwaukee		1.49	5.53	−9.36	−2.20	1.15	4.58	14.2
Minneapolis		1.86	7.00	−13.2	−2.99	4.43	7.07	10.5
Orlando		5.25	15.4	−28.7	0.897	7.25	12.6	32.3
Philadelphia		5.62	6.69	−8.54	0.499	7.70	11.4	12.9
Phoenix		3.53	16.3	−25.4	−7.37	3.67	7.28	39.4
Riverside		6.49	19.1	−40.2	−2.73	11.6	16.8	33.2
Sacramento		5.06	15.7	−26.8	−12.0	13.4	17.0	22.7
San Diego		4.47	14.7	−25.2	−7.63	9.04	15.7	27.9
San Francisco		2.29	12.0	−25.2	−5.46	3.96	11.3	21.9
San Jose		1.33	11.0	−19.2	−5.65	0.327	9.23	25.8
Tampa		4.66	14.3	−26.5	−2.03	9.44	11.5	23.9
Washington		6.43	12.8	−21.7	−6.27	11.1	15.9	20.7
Mean		3.56	11.14					

Table 4. *House price appreciation rates in each MSA i from month t to month $t+12$ in units of* %/yr *over the time interval from January 2000 to December 2007. The shaded region in the sparkline plots covers the interquartile range for each MSA and is not a constant scale. Reads: "Over the entire sample period, house prices in Las Vegas grew by an average of* 3.69% *per year; however, this growth rate skyrocketed to* 44.3% *per year in 2004."*

of purchases made by out-of-town second house buyers as Milwaukee.

Research on the role of second house buyers in the housing market typically treats local and out-of-town second house buyers in the same way. However, as demonstrated in Table 3, purchases by local second house buyers exhibit a very different time series pattern than that of out-of-town second house buyers. The overall share of purchases by local second house buyers varies much less across markets compared to the variability in house price appreciation. As well, in most cases (Las Vegas is an appreciable exception), the share of local second house buyers does not exhibit a hump with a peak at or near the peak of house prices.

2.2. **House Price Appreciation Rate.** We use house price appreciation rate data to measure the price impact of out-of-town second house buyer demand shocks. We obtain monthly house price index (HPI) data from Zillow, Inc. at the MSA level. Zillow data are available for a larger number of locations than the S&P/Case and Shiller index data and use a methodology that potentially makes the index less sensitive to changes in the mix of properties that sell at a given point in time. The Zillow indexes behave quite similarly to the S&P/Case and Shiller indexes during the boom, but show less of a sharp decline in 2007.

IAR Appreciation Rates in % per Year

		Mean	Sd	Min	Q25	Q50	Q75	Max
Baltimore		2.04	9.75	−18.5	−4.14	2.39	9.36	24.8
Charlotte		−1.00	8.53	−16.4	−6.42	−1.83	4.06	21.9
Cincinnati		−1.72	5.31	−12.6	−5.26	−1.55	1.76	14.9
Cleveland		−3.41	5.70	−18.2	−6.04	−2.63	−0.545	13.1
Denver		−2.42	7.51	−17.5	−7.18	−1.91	2.36	20.4
Jacksonville		0.792	11.8	−25.9	−4.13	0.494	6.11	31.4
Las Vegas		0.640	16.6	−39.4	−5.42	0.678	6.85	36.9
Los Angeles		0.210	14.0	−39.6	−4.44	3.12	8.47	24.7
Miami		2.72	14.9	−33.1	−0.856	4.66	12.1	24.7
Milwaukee		0.047	6.67	−15.1	−3.43	0.548	4.54	14.3
Minneapolis		−0.095	7.37	−19.6	−2.46	0.872	4.76	13.5
Orlando		1.84	14.7	−32.5	−4.01	1.66	9.34	32.6
Philadelphia		2.37	7.11	−14.5	−0.852	2.36	6.65	17.8
Phoenix		0.205	16.6	−32.3	−8.55	−1.65	7.01	36.6
Riverside		0.369	16.4	−46.3	−3.20	3.69	10.7	27.1
Sacramento		0.938	13.0	−35.2	−4.28	4.78	9.30	20.1
San Diego		−1.38	12.8	−37.2	−4.87	1.64	5.91	25.0
San Francisco		−0.385	13.8	−41.1	−3.47	3.68	7.85	18.2
San Jose		0.20	13.2	−36.1	−2.39	1.81	7.42	23.9
Tampa		1.57	13.5	−30.9	−2.26	3.61	8.49	27.8
Washington		1.64	11.6	−29.1	−3.04	4.59	10.5	18.6
Mean		0.246	11.5					

Table 5. *IAR appreciation rate in each MSA i from month t to month $t+12$ in units of* %/yr *over the time interval from January 2000 to December 2007. The shaded region in the sparkline plots covers the interquartile range for each MSA and is not a constant scale. Reads: "Over the entire sample period, implied-to-actual rent ratios in Phoenix grew by an average of* 0.205% *per year; however, this growth rate jumped to* 36.6% *per year in 2004."*

Definition 5 (House Price Appreciation Rate)**.** *Define* $\Delta \log P_{i,t\to(t+\tau)} = \log P_{i,t+\tau} - \log P_{i,t}$ *as the house price appreciation rate in MSA i at month t in units of* $1/\tau\text{mo}$*, where $P_{i,t}$ is the HPI index level normalized to be unity in a base year.*

Table 4 gives summary statistics for the house price appreciation rate in units of percent per year. A number of the markets saw annual house price appreciation rates above 20% per year, with house price appreciation rates exceeding 35% per year in Las Vegas and Phoenix near the peak of their booms. What's more, the sparkline plots show that the timing of these peaks varied substantially from MSA to MSA with the house price appreciation rate peak in Las Vegas arriving more than a year prior to the peak in Phoenix. As documented in Ferreira and Gyourko (2011) the recent boom began at different times in different MSAs, and house prices exhibited different appreciation rates across these markets. Even the start dates of the subsequent decline in prices differed by a year or more.

2.3. **Implied-to-Actual Rent Ratio Appreciation Rate.** In order to make statements about mispricing, we must pick a pricing model. We use IAR appreciation rate data suggested by the user cost model to measure the impact of out-of-town second house buyer demand shocks on mispricing.

Input Variables to User-Cost Model

Variable	Source	Description
ρ_t	CRSP	Risk-free rate computed as annualized 10yr T-Bill.
$\omega_{i,t}$	Emrath (2002)	Property tax rate.
μ_t	Federal Reserve Bank of St. Louis	Mortgage interest rate.
$\kappa_{i,t}$	NBER	Federal marginal tax rate.
δ	Harding, Miceli, and Sirmans (2000)	Housing capital depreciation rate.
$\mathrm{E}\left[\Delta \log P_{i,t+12}\right]$	Ferreira and Gyourko (2011), the US Census, and the Livingston Survey	Expected house price appreciation rate equals historical long-term real growth rates by MSA plus expected inflation.
π	Flavin and Yamashita (2002)	Risk premium associated with owning a house.

Table 6. *Data sources and short descriptions of the input variables used to compute the user cost of housing in Himmelberg, Mayer, and Sinai (2005). All variables have units of* $1/\mathrm{yr}$ *except for the federal marginal tax rate,* $\kappa_{i,t}$*, which is a fraction. All variables reflect rates over the time interval* t *to* $t+12$ *and are known at time* t*.*

Beginning with Poterba (1984), many authors have priced residential real estate by comparing the price of a house to the present value of its stream of rental payments, taking into account the favorable tax treatment given to owner occupied properties and mortgage interest payments. This pricing strategy is similar to the dividend discount model for the stock market. We refer to models that price housing along this margin as user cost models.

Unlike the stock market where analysts have actual dividends and share prices, in the housing market it is quite unusual to have matched data on the sale price and rental rate over the next year for a particular house. Himmelberg, Mayer, and Sinai (2005) suggest a methodology that allows us to create an index of mispricing by comparing the ratio of the imputed rent level to the actual rent level, where the imputed rent is calculated by multiplying the user cost times the price of an owner occupied house. We use the user cost of housing data from Himmelberg, Mayer, and Sinai (2005) updated through December 2007. Table 6 gives the data sources and a set of short descriptions for the input variables used to compute the user cost of housing in Equation (1).

Definition 6 (User Cost of Housing)**.** *Define* $U_{i,t\to(t+12)}$ *as the user cost of housing in MSA* i *in month* t *which reflects the fraction of the price of a house that an owner must pay in order to live in that house over the next year from time* t *to time* $t+12$*:*

$$U_{i,t\to(t+12)} = \rho_t + \omega_{i,t} - \kappa_{i,t} \cdot \{\mu_t + \omega_{i,t}\} + \delta - \mathrm{E}\left[\Delta \log P_{i,t\to(t+12)}\right] + \pi \tag{1}$$

where the user cost of housing has units of $1/\mathrm{yr}$*.*

In the standard user cost model, the price of a house in an MSA i at month t multiplied by the prevailing user cost of housing should equal the rental rate over the next year, or $P_{i,t} \cdot U_{i,t\to(t+12)} = R_{i,t\to(t+12)}$. REIS collects monthly estimates of the annualized rent for a 2-bedroom apartment.

Definition 7 (Apartment Rental Rate Index)**.** *Define* $R_{i,t\to(t+12)}$ *as the apartment rental rate index* *in MSA i at month t which reflects the annual rent payment required to live in 2-bedroom apartment in MSA i from month t to $t+12$ in units of* $1/\mathrm{yr}$.

The log IAR can be thought of as the excess return over the apartment rental rate of a trading strategy whereby an agent borrows money at rate ρ_t per year to buy a house, lives in the house for a year while paying a constant proportion of the house value in depreciation costs δ per year and earning the tax shield $\kappa_{i,t}$ on his debt payments of $(\mu_t + \omega_{i,t})$ per year and then sells the house after one year getting capital gains at the expected price appreciation rate of $\mathrm{E}\left[\Delta \log P_{i,t\to(t+12)}\right]$ per year while enduring a constant risk premium of π per year. Himmelberg, Mayer, and Sinai (2005), do not allow the risk premium or leverage to change over time. Thus the computation can be thought of as a long-run measure of the relative price of owning versus renting, abstracting from important short-run considerations like easy and cheap leverage in the mid 2000s and time varying risk premia. When the IAR in a given metropolitan area exceeds unity, owning a house is more expensive than renting relative to the average value over the sample period.

Definition 8 (Implied-to-Actual Rent Ratio (IAR) Appreciation Rate)**.** *$Z_{i,t}$ denotes the IAR in MSA i at month t reflecting the ratio of the cost to a potential owner of borrowing money, purchasing a house and then selling it in 1 year to the cost at which he can rent a comparable property for the same amount of time:*

$$\begin{aligned} Z_{i,t} &= \frac{1}{\bar{Z}_i} \cdot \left(\frac{P_{i,t} \cdot U_{i,t\to(t+12)}}{R_{i,t\to(t+12)}} \right) \\ \bar{Z}_i &= \frac{1}{T} \cdot \sum_{t=1}^{T} \left(\frac{P_{i,t} \cdot U_{i,t\to(t+12)}}{R_{i,t\to(t+12)}} \right) \end{aligned} \tag{2}$$

The IAR is scaled to equal 1 relative to the average value of the ratio from January 1980 to December 2007. $\Delta \log Z_{i,t\to(t+\tau)} = \log Z_{i,t+\tau} - \log Z_{i,t}$ denotes the IAR appreciation rate in units of $1/\tau\mathrm{mo}$.

The IAR is computed using HPI data from both the Federal Housing Finance Administration and Zillow since the Zillow house price indexes are not available prior to 1996. Table 5 gives summary statistics for the annual IAR appreciation rates. This measure of mispricing varies substantially across markets such as Phoenix and Denver, respectively. At the peak in Phoenix, a tenant renting an apartment for \$1000 per month would have to pay \$1658 per month in mortgage payments and other costs in order to buy an equivalent house and live in it from January 2004 to December 2004. By comparison, in Denver, this ratio was 1.267 between 2004 and 2006, so a tenant would have paid about \$1267 per month to purchase a house that rented for \$1000 per month and live in it from January 2004 to December 2004. While houses in Denver were still priced at a small premium relative to renting at the peak of the national boom, they appeared much less overpriced than houses in Phoenix at the same time.

Researchers have critiqued the user cost approach in a number of ways. For example, Glaeser and Gyourko (2007) point out that very few single family houses are rented, so any rental index is not assured to match up with the price index. Also, the user cost model as estimated above is inherently static, so it cannot easily incorporate time varying factors like risk premia, the expected growth rates of house prices, mean-reverting interest rates, credit constraints, and mobility. Glaeser, Gottlieb, and Gyourko (2010) gives a model that attempts to correct the simple user cost model for some of these time-varying features. Mayer (2010) provides a discussion of the pros and cons of the user cost model as well as a collection of alternative measures of mispricing in the housing market.

Nevertheless, a simple analysis of the user cost model suggests it is well-suited for the purposes of our paper in that it allows us to estimate a single index value that proxies for mispricing that accounts for the fact that dividend streams earned by house buyers systematically vary due to factors such as tax treatment and prevailing interest rates. Comparing house prices to variables like employment and income has no firm theoretical prediction and fails to account for changes in economic fundamentals like interest rates and variable land supply across locations. Comparing house prices to construction costs only works in markets where land has very low value and thus is in abundant supply relative to demand. Even in locations with low land prices, house prices should still equal the present value of rents. By analogy, note that in equity markets there are circumstances where shares of the same stock confer different dividend streams and a user cost-like model is warranted. i.e., shares with voting rights command a higher price to dividend ratio than shares of the same stock without voting rights as documented in Zingales (1995).

Finally, Table 2 in Hubbard and Mayer (2009) estimates a log-linearized version of the user cost model in levels:

$$\log P_{i,t} = \alpha_i + \kappa_t + \beta \cdot \log R_{i,t\to(t+12)} + \gamma \cdot \log U_{i,t\to(t+12)} + \varepsilon_{i,t} \tag{3}$$

over the time interval from January 1980 to December 2007 with both MSA and year fixed effects. The authors find coefficients of $\gamma = 0.93$ and $\beta = -0.75$, which are very close to the values of 1.0 and -1.0 respectively as predicted by the static user cost model. Thus, even though it has many imperfections, the user cost appears to provide a simple benchmark for what housing prices might be in a long-term equilibrium. In all of the specifications below, we repeat our analysis with both house price and IAR appreciation rates and report both sets of coefficients. The findings are quite similar for both measures. As well, all of our results involving IAR appreciation rates are robust to computing this measure with a variety of different assumptions about the expected future house price appreciation rate.

3. Misinformed Speculators

This section gives evidence that out-of-town second house buyers behaved like misinformed speculators. In Subsection 3.1, we show that out-of-town second house buyers are less informed about local market conditions relative to local second house buyers and owner occupants. Out-of-town second house buyers earned lower capital gains on their purchases in MSAs such as Las Vegas, Phoenix, and Miami relative to local second house buyers who were better able to time the market. Of course, returns are composed of both capital gains and dividends. In Subsection 3.2, we then argue that out-of-town second house buyers are less able to consume the dividends generated by their housing purchase. Finally, in Subsection 3.3 we address concerns that out-of-town second house buyers might have different motivations for making their purchase than local second house buyers or owner occupants. For instance, they might be extremely wealthy or have a demand for portfolio diversification. We show that these concerns don't appear to be consistent with our data.

Out-of-town second house buyers look very much like average home buyers except for the fact that they've decided to treat housing as a financial asset and make a speculative bet.

In summary, we view local second house buyers as agents who are often engaging in a similar trade as out-of-town second house buyers, but who are better informed about future house price appreciation rates. To be clear, there are a variety of differences between the two groups. For instance, out-of-town second house buyers are more likely to live in their second house part time, while local second house buyers are more likely to rent out their house as a source of income. However, after acknowledging these differences, the fact remains that both groups of traders are less able to consume the dividend stream from their second house purchase relative to owner occupants and thus both groups are more reliant on capital gains to earn positive returns on their investments. Thus, we think of local second house buyers as a somewhat comparable "control group" of traders who are more informed than out-of-town second house buyers.

3.1. **Informational Disadvantage.** Out-of-town second house buyers are uninformed relative to local second house buyers or owner occupants. By definition, out-of-town second house buyers live farther away from the houses they have purchased than local second house buyers and owner occupants. Thus, these traders don't "know the neighborhood" as well as local buyers. In addition, out-of-town second house buyers face a difficult principal agent problem when dealing with local real estate agents who are paid on commission. Levitt and Syverson (2008) find that real estate agents have substantial discretion in the timing and pricing of house sales with brokers receiving about 3.7% more than other local owner occupants when selling their own houses. Out-of-town second house buyers with higher monitoring costs likely face an even larger distortion.

As more direct evidence, we show that out-of-town second house buyers are less successful in timing their exit from the market when compared to local second house buyers. Figure 3 shows the average realized capital gains on single family house purchases made by local and out-of-town second house buyers in MSA i in each month t in units of percent per year. We compute this capital gain by taking the weighted average of the annualized house price appreciation rates earned by all second house buyers who purchased a property in MSA i in month t and then resold it in month $t + \tau$ for $\tau \in [1, \bar{\tau}]$, where $\bar{\tau}$ represents the number of months between December 2007 and t where our data are right censored. We assign observations that are right censored the house price appreciation rate from t to the end date $(t + \bar{\tau})$ = December 2008. The width of the out-of-town second house buyers line is scaled to represent the number of out-of-town second house purchases in MSA i in month t.

In key markets such as Las Vegas, Phoenix, Miami and Tampa, out-of-town second house buyers earned lower capital gains on their investments relative to local second house buyers. For instance, out-of-town second houes buyers purchasing in Las Vegas in March 2004 earned an 8% per year capital gain on average; whereas, local second house buyers purchasing in the same month earned a 17% per year capital gain on average. In addition, the average capital gain on out-of-town second house buyer purchases decreased from 8% per year to −15% per year as the number of out-of-town second house purchases as a percent of all sales rose from 5% in March 2004 to 13% in January 2007. While out-of-town second house buyers realized 3% per year lower capital gains than local second house buyers in Las Vegas during the entire sample period, this gap is largest for buyers

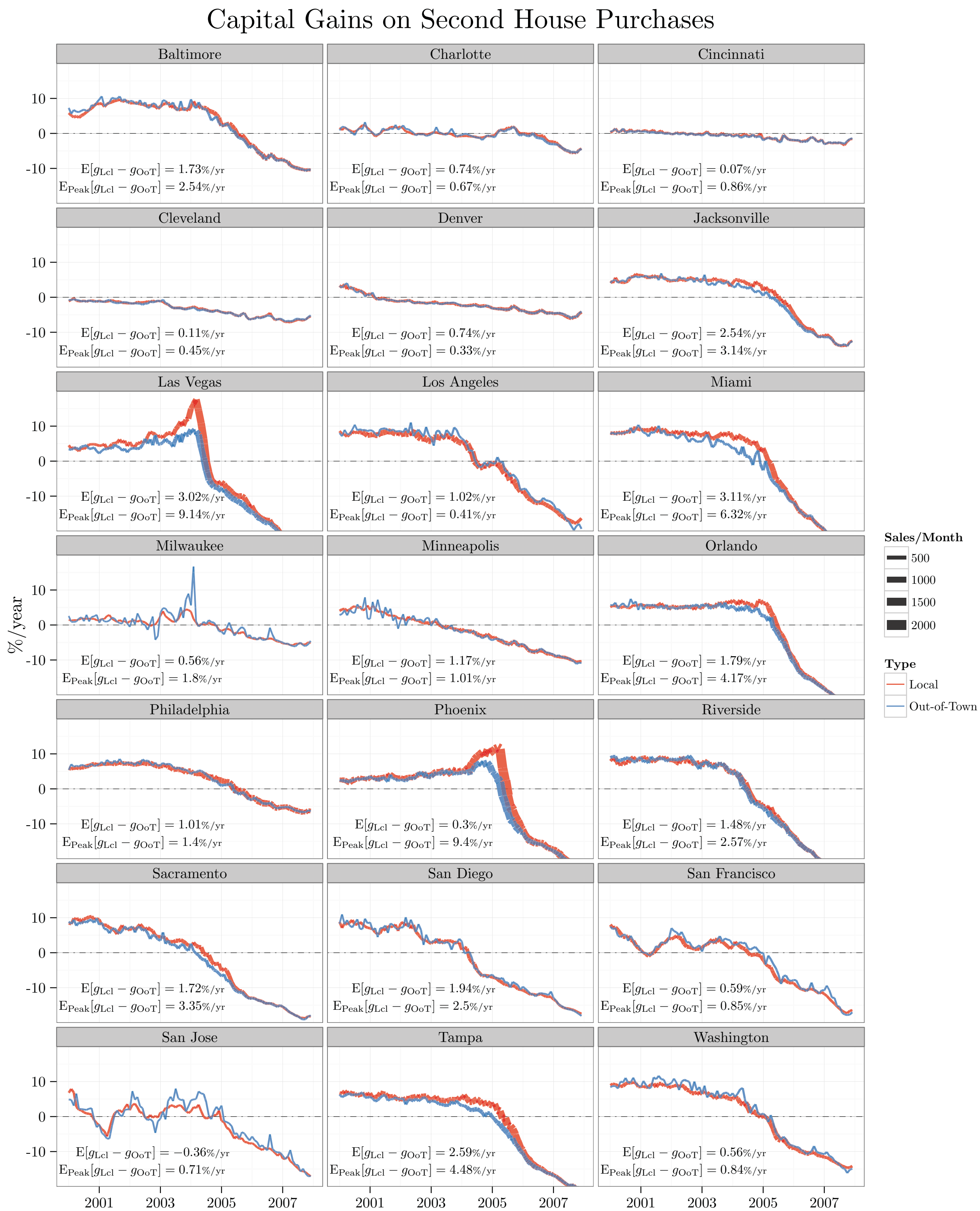


Figure 3. *The capital gain on single family house purchases made by local and out-of-town second house buyers from January 2000 to December 2008 using ZIP code by month level house price index data from Zillow. The width of each line is scaled by the number of purchases by each buyer type.* g_{OoT} *and* g_{Local} *are the realized capital gains for out-of-town and local second house buyers in units of* %/yr. *Reads: "Out-of-town second house buyers purchasing in Las Vegas in March 2004 realized an* 8% *per year capital gain on average; whereas, local second house buyers realized a* 17% *per year capital gain on average. The average capital gain earned by out-of-town second house buyers decreased from* 8% *per year to* −15% *per year as the number of out-of-town second house purchases as a percent of all sales rose from* 5% *in March 2004 to* 13% *in January 2007."*

who bought near the peak of the housing boom in Las Vegas. These patterns exist only for "boom" markets and are either absent or reversed in other markets such as San Francisco or Cleveland which traditionally have either very cyclical or very flat house price appreciation rates.

Since both out-of-town and local second house buyers bought their houses at the same time in Figure 3, the differences in capital gains earned by each group of traders must stem from differences in exit timing. Put differently, the figure suggests that local second house buyers in markets such as Las Vegas, Phoenix, Miami and Tampa were better able to time the market downturn than out-of-town second house buyers. To quantify this intuition, we estimate the regression specification in Equation (4) which captures the extent to which out-of-town and local second house buyers were able to recognize the most appropriate time to sell their house prior to the crash. In particular, we estimate the probability that a second house buyer "flips" their house within 6 months as a function of (a) the buyer type, (b) whether house prices have hit their peak, (c) the extent to which house prices are rising or falling in the upcoming year, and (d) the interaction of these terms:

$$\begin{aligned} F_{n,i,t-6} = \alpha_i + \widehat{\alpha}_i \cdot 1_n^{\mathrm{OoT}} + \kappa_t + \widehat{\kappa}_t \cdot 1_n^{\mathrm{OoT}} + \widehat{\kappa} \cdot 1_n^{\mathrm{OoT}} \\ + \beta \cdot \Delta \log P_{i,t\to(t+12)} + \widehat{\beta} \cdot \Delta \log P_{i,t\to(t+12)} \cdot 1_n^{\mathrm{OoT}} \\ + \gamma \cdot 1_{i,t}^{\mathrm{PostPeak}} + \widehat{\gamma} \cdot 1_{i,t}^{\mathrm{PostPeak}} \cdot 1_n^{\mathrm{OoT}} \\ + \delta \cdot \Delta \log P_{i,t\to(t+12)} \cdot 1_{i,t}^{\mathrm{PostPeak}} + \widehat{\delta} \cdot \Delta \log P_{i,t\to(t+12)} \cdot 1_{i,t}^{\mathrm{PostPeak}} \cdot 1_n^{\mathrm{OoT}} \\ + \varepsilon_{n,i,t} \end{aligned} \tag{4}$$

If local second house buyers are better informed about future house price appreciation rates, then this knowledge should be revealed in their resale timing. These buyers should be more likely to exit the each market immediately before the house price appreciation rate begins to collapse. Naïvely, we might expect that more informed traders would always flip at a higher rate over the interval $(t-6) \to t$ when house price appreciation rates are lower over the interval from $t \to (t+12)$. However, quickly reselling a house is difficult when house prices are collapsing. Thus, this naïve estimate of a β% response to a 1%pt per year increase in the house price appreciation rate in MSA i from $t \to (t+12)$ is a weighted average of the decline in the flipping rate in order to earn the capital gains and the increase in the flipping rate due to market liquidity. To disentangle these two offsetting effects, we interact the house price appreciation rate in MSA i from month $t \to (t+12)$ with a dummy variable $1_{i,t}^{\mathrm{PostPeak}} \in \{0,1\}$ which is 1 if the house price appreciation rate in MSA i peaked in months $(t-6) \to t$ and house price appreciation rates in MSA i reached at least 20% per year to ensure we are not identifying small local peaks, but rather the culmination of a large increase in prices.

Table 7 displays the estimated regression coefficients from Equation (4). In all of our regression specifications with both time and group fixed effects, we report unclustered standard errors as well as standard errors clustered at along both the time and group dimensions. Reporting each of these three values allows both verifies the robustness of the coefficient estimates and also allows readers to diagnose potential problems with the specification as suggested in Petersen (2009). First, we see that out-of-town second house buyers are 5% less likely than local second house buyers to resell their house within 6 months over the entire sample. Next, we find that while local second house buyers are

Second House Buyer Market Timing

Dep. Var.: House resells within 6 Months				
	Estimate	Std. Error		
Out-of-town Second House Buyer	−0.050	0.015	0.019	0.015
Future House Price Apprec. Rate	0.123	0.014	0.031	0.032
Post Peak Resale	0.043	0.010	0.020	0.019
Post Peak × Future Apprec. Rate	−0.150	0.046	0.055	0.041
Out-of-town × Future Apprec. Rate	−0.131	0.019	0.018	0.025
Out-of-town × Post Peak	−0.031	0.015	0.020	0.021
Out-of-town × Post Peak × Future Apprec. Rate	0.112	0.066	0.071	0.090
	Clustering	$\emptyset$	t	i
	N	1390118		
	R^2	0.083		

Table 7. *Estimated coefficients and standard errors from Equation (4). Resale within 6 months is defined as one if a house purchase in month $t-6$ in MSA i resells during the interval $(t-6, t]$. Future house price appreciation rate is the house price appreciation rate in MSA i over the interval from $t \to (t+12)$ in units of percent per year. Post peak is a dummy variable which is 1 if the house price appreciation rate in MSA i peaked in months $(t-6, t]$ and MSA i's house price appreciation rate peak reached 20% per year or more. The regression uses monthly data from July 2000 to June 2008 on all house sales to local and out-of-town second house purchases the 21 MSAs weighted by the number of second house purchases in each MSA in each month. Fixed effect estimates of α_i, $\hat{\alpha}_i$, κ_t and $\hat{\kappa}_t$ are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time and clustering across MSAs respectively.*

4.3% more likely to flip their house purchase within the 6 months immediately following the peak in local house price appreciation rates, out-of-town second house buyers are only $4.3 - 3.1 = 1.2\%$ more likely to flip their house purchase during this key interval. What's more, a t-test reveals that the point estimate for out-of-town second house buyers is not statistically different from zero, suggesting that the likelihood of flipping is nearly unchanged for out-of-town second house buyers immediately after a house price peak. Finally, while local second house buyers are more likely to flip a purchase when prices are rising rapidly during the subsequent 12 months, this effect disappears immediately following the peak in house price appreciation rates further suggesting that local second house buyers are strategically changing their behavior in order to time the market. On the other hand, while out-of-town second house buyers tend to flip houses more often when house prices are declining during the entire sample, this effect disappears immediately following the peak in house price appreciation rates. Taken together, the evidence presented in Table 7 suggests that out-of-town second house buyers are not using insights about future house price appreciation rates to strategically exit their investments in the local housing market to the same extent as local speculators.

3.2. **Dividend Consumption.** Returns are composed of both capital gains and dividends. We now argue that out-of-town second house buyers are less able to consume the dividends generated by their housing purchase. There are a number of reasons someone might want to buy a second house in a different city: a buyer might want to live in the house for part of the year, rent the property out as an additional source of income, or renovate the house and sell it for a profit at sometime in the future. According to a 2005 survey conducted by the National Association of

Realtors,[1] 31% of second house buyers in 2004 planned to use their house as a vacation home, while the remaining 69% of second house buyers planned to use their house as a rental property. In each of these instances, an out-of-town second house buyer gets lower dividends from the purchased house than a local second house buyer or an owner occupant.

We first examine out-of-town second house buyers who use the house only for weekends, holidays, or vacations. Part time residents can only consume the dividend (e.g., live in the house) for the portion of each year that they live in the house and thus get lower use than an owner occupant. Next, consider out-of-town second house buyers who wish to rent out their purchase. Out-of-town second house buyers face potentially higher costs of property maintenance, renovation, and rental management. It is costly and difficult to supervise contractors or maintenance people from far away. As a proxy for the full opportunity cost, note that a typical property manager charges a fee of one months rent plus an additional 8% of the annual rent each year to lease a house and manage relations with the tenant. Direct costs to maintain and pay for repairs to appliances and the house itself are extra. To put these figures in perspective, imagine you decided to purchase a \$200k house in Las Vegas (the median price in 2004) with a 20% downpayment (the median downpayment for out-of-town second house buyers in 2004) in order to rent it out for \$1000 per month (the median rental price for 2 bedroom houses in Las Vegas in 2004). Over the course of the next 2 years the monthly property management fees would amount to 7% of your initial \$40k investment. This is a non-trivial drag on your returns, and even assumes that you would be able to immediately rent out your second house purchase and keep a tenant for all 24 months. To make matters worse, any second house buyer (both local and out-of-town) wishing to rent out their property faces the prospect of higher physical depreciation costs as rental tenants may treat the house relatively poorly as compared to owner occupants as noted in Harding, Miceli, and Sirmans (2000). Finally, out-of-town second house buyers who plan on renovating a house and selling it for a profit do not live in the property and are thus almost entirely motivated by future capital gains as pointed out in Bayer, Geissler, and Roberts (2011).

3.3. **Preference Heterogeneity.** Even if out-of-town second house buyers are both less informed and less able to consume the dividend from their purchase than local second house buyers, they would not be misinformed speculators if they had completely different motivations for making their purchases. For example, a recent New York Times article on the marriage market in China[2] reported that “70% of single women said they would tie the knot only with a prospective husband who owned a home.” Obviously, we wouldn’t want to classify Chinese bachelors as misinformed speculators if we found that they were willing to overpay for housing and tended not to live in their purchase until married. There doesn’t appear to be any such explanation four out-of-town second house purchases in the US residential housing market. An out-of-town second house buyer in our data looks very much like an average home buyer except for the fact that she’s decided to treat housing as a financial asset and make a speculative bet.

In this subsection, we consider 3 different motivations. First, you might be concerned that the bulk of out-of-town second house buyers that we study are simply rich occupants in coastal cities.

[1] Investment and Vacation Home Buyers Survey. *National Association of Realtors*, Mar. 6 2005.

[2] Andrew Jacobs. For many Chinese men, no deed means no dates. *The New York Times*, Apr. 14 2011.

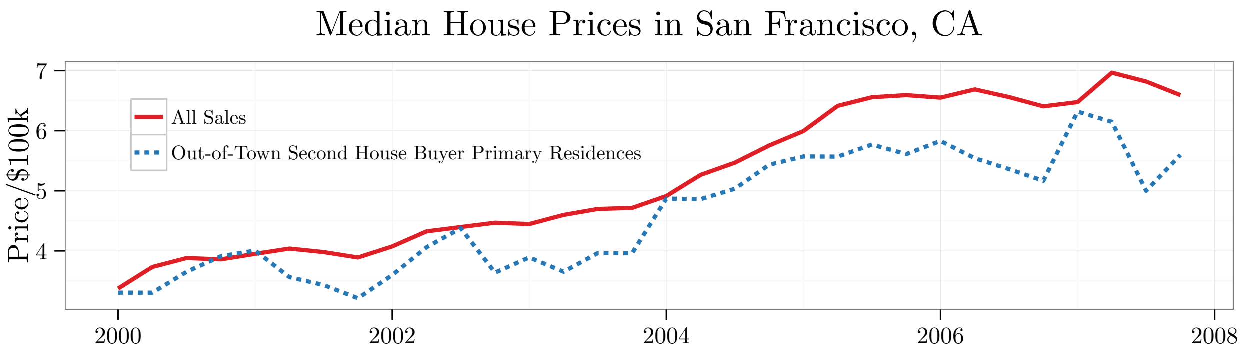


Figure 4. *Median primary residence house price for the populations of out-of-town second house buyers and of all buyers living in San Francisco in units of* \$100k *over the time period from January 2000 to December 2007. The price of the primary residences of out-of-town second house buyers living in San Francisco is computed by scaling up the most recent sale price by the* Zillow *ZIP code level price index. Reads: "In January 2005, the median value of all single family houses purchased in San Francisco was* \$600k. *By contrast, the median value of the primary residences of out-of-town second house buyers who live in San Francisco and bought a second house in another MSA in January 2005 was only* \$555k.*"*

Perhaps, they are simply getting lots of utility from owning a vacation house in the Phoenix or a weekend getaway in Miami. The data do not appear to be consistent with this hypothesis. For example, Figure 4 shows that the typical out-of-town second house buyer is not a very rich household. We examine the price of the house that is the primary residence of out-of-town second house buyers in the highest income cities including San Francisco, San Jose, and New York. In January 2005, the median value of all single family houses purchased in San Francisco was \$600k. By contrast, the median value of primary residences of out-of-town second house buyers who live in San Francisco and bought a second house in another MSA in January 2005 was only \$555k. While the value of their primary residence is not a complete characterization of out-of-town second house buyers' wealth, this evidence suggests that the super-rich are not the only traders buying second houses in other cities.

Second, perhaps out-of-town second house buyers are just soon-to-be retirees looking to buy a house to retire in? For instance, a 2004 Money Magazine article[3] writes that "every 8 seconds, a boomer turns 50. . . so if you're thinking about retiring in a new home—whether it's your only base or one of two—now's the time to find it. 'There's a concern that the prices of properties most desired by boomers may get out of reach,' says David Hehman, CEO of EscapeHomes.com, a San Francisco realty firm specializing in second homes. . . Our advice: Buy now, retire later." Do we really want to call Nana a misinformed speculator? In a word: Yes. There is no rule saying that misinformed speculators can't sport bifocals and wear scratchy sweaters. If you replace the word "properties" with "tech stocks" the article may well have been written 4 years earlier. Baby boomers that bought a second house in Phoenix or Las Vegas with the goal of waiting a few years and then retiring there made a calculated bet on future house prices and certainly didn't consume the full dividend from their housing purchase while still employed.

Third and finally, perhaps out-of-town second house buyers are interested in the diversification

[3]Marion Asnes. Buy Now, Retire Later. *Money Magazine*, Jul. 1 2004.

benefits of owning a second house in a market where returns are less correlated with other assets in the portfolio. i.e., Lustig and Van Nieuwerburgh (2005) give empirical evidence of a housing capital risk premia due to the covariance of its returns with the returns to the household's human capital. Nevertheless, diversification can't be the main motivation for out-of-town second house buyers in our sample because out-of-town second house buyers tended to own several houses in MSAs with highly correlated price dynamics. We calculate that the typical out-of-town second house buyer who owned an investment property in Los Angeles in January 2006 owned 2.1 houses in addition to his primary residence—i.e., his primary residence, the second house in Los Angeles that he bought in January 2006, and 1 additional house that he bought in previous months.

It's even hard to see how out-of-town second house buyers who only bought 1 additional house could have been motivated primarily by diversification benefits. e.g., imagine you are a home owner living in Los Angeles, and you are considering investing \$40k as a downpayment on a second house in Phoenix. If you don't make this purchase, you will realize utility over the next year of:

$$U_{\text{Don't Invest}} = \mathrm{E}\left[\, \Delta P_{\text{LA}} + D_{\text{LA}} \,\right] - \frac{\gamma}{2} \cdot \mathrm{Var}\left[\, \Delta P_{\text{LA}} + D_{\text{LA}} \,\right] \tag{5}$$

where ΔP_{LA} denotes the dollar house price appreciation in Los Angeles, and D_{LA} denotes the housing dividend paid out by your primary residence in Los Angeles. If you decide to purchase the second house in Phoenix, you will realize a utility over the next year of:

$$\begin{aligned} U_{\text{Invest}} = \mathrm{E}\left[\, (\Delta P_{\text{LA}} + D_{\text{LA}}) + (\{\Delta P_{\text{Phx}} - \kappa\} + \{D_{\text{Phx}} - \lambda\}) \,\right] - \$40\text{k} \\ - \frac{\gamma}{2} \cdot \mathrm{Var}\left[\, (\Delta P_{\text{LA}} + D_{\text{LA}}) + (\{\Delta P_{\text{Phx}} - \kappa\} + \{D_{\text{Phx}} - \lambda\}) \,\right] \end{aligned} \tag{6}$$

where ΔP_{Phx} denotes the dollar house price appreciation in Phoenix, κ denotes the wedge between the house price appreciation rates earned by out-of-town and local buyers, D_{Phx} denotes the housing dividend paid out by the second house in Phoenix, and λ denotes the dividend loss due to being an out-of-town buyer. Subsection 3.1 suggests that on a \$200k house the dollar house price appreciation wedge would be about $\kappa = 0.09 \times \$200\text{k} = \$18\text{k}$ at peak. Subsection 3.2 suggests that the dividend drag of being an out-of-town owner would be around $\lambda = \$2\text{k}$ in the first year of ownership. Thus, in order for the second house purchase in Phoenix to make sense for you, it has to be the case that:

$$\$20\text{k} < -\, \gamma \cdot \rho_{\text{LA,Phx}} \cdot \sigma_{\text{LA}} \cdot \sigma_{\text{Phx}} \tag{7}$$

after setting $\$40\text{k} = \mu_{\text{Phx}} + D_{\text{Phx}} - \frac{\gamma}{2} \cdot \sigma^2_{\text{Phx}}$ since there is an active residential housing market in Phoenix. On a \$500k primary residence in Los Angeles, the yearly standard deviation in dollar price appreciation over the time period from 1996 to 2011 is $\sigma_{\text{LA}} = \$60\text{k}$. On a \$200k second house in Phoenix, the yearly standard deviation in dollar price appreciation over the same time period is $\sigma_{\text{Phx}} = \$30\text{k}$. Thus, for a risk aversion coefficient $\gamma = 2$, there just needs to be a really weak negative correlation of -0.01 between dollar house price appreciation in Los Angeles and Phoenix for the second house in Phoenix to be a good investment. Here is the punchline: the actual correlation is extremely positive at $\rho_{\text{LA,Phx}} = 0.81$! Buying a second house in Phoenix just isn't a good hedge against house price fluctuations in Los Angeles.

Diversification benefits surely play some role in determining macro-level house price dynamics. However, the previous 2 paragraphs suggest they are unlikely to be key factors governing the choices of out-of-town second house buyers. This can't be what's driving out-of-town second house buyer demand.

4. Predictive Regressions

In this section, we show that increases in out-of-town second house buyer demand predict increases in future house price and IAR appreciation rates; whereas, increases in local second house buyer demand are not associated with increases in either of these measures. Subsection 4.1 details the house price appreciation rate results, and Subsection 4.2 outlines the IAR appreciation rate results. Of course, in order to make statements about mispricing, we must pick a pricing model. In this paper we subscribe to the user-cost model and use deviations of city-wide log IAR ratios from zero as a proxy for mispricing. The higher the log IAR ratio is, the cheaper it is to rent rather than buy, and the higher the level of mispricing on the rent-vs-own margin. One of the main criticisms of this model is that expected house price appreciation rates are an *input*. To mitigate this concern we show that our empirical results do not depend on the gritty details of how this expectation is calculated by rerunning the resulting using a moving average calculation.

4.1. **House Price Appreciation Rate Regressions.** We estimate a panel VAR characterizing the relationship between the house price appreciation rate in an MSA i from month t to $t+1$ and the numbers of local and out-of-town second house purchases as a percent of sales in MSA i in month t using a panel data set indexed by MSA and month. The state vector $\mathbf{Y}_{i,t}$ contains the house price appreciation rate from month $(t-1) \to t$ in MSA i as well as the fraction of all house purchases in MSA i in month t that were made by out-of-town and local second house buyers respectively:

$$\mathbf{Y}_{i,t} = \begin{bmatrix} \Delta \log P_{i,(t-1)\to t} & \frac{S^{\mathrm{OoT}}_{i,t}}{X_{i,t}} & \frac{S^{\mathrm{Lcl}}_{i,t}}{X_{i,t}} \end{bmatrix}^{\top} \tag{8}$$

The omitted category is the fraction of sales made by owner occupants. Using this state vector, we study the regression specified in Equation (9) below:

$$\mathcal{E}_{i,t} = (\mathbf{I} - \mathbf{\Theta}\mathcal{L}_1)\,(\mathbf{Y}_{i,t} - \mathbf{A}_i - \mathbf{K}_t) \tag{9}$$

In this representation, $\mathbf{I}$ denotes a 3×3 identity matrix, $\mathbf{\Theta}$ denotes the 3×3 transition matrix, $\mathcal{L}_1$ denotes the 1 month lag operator, $\mathbf{A}_i$ and $\mathbf{K}_t$ denote 3×1 vectors of MSA and month specific fixed effects and $\mathcal{E}_{i,t}$ denotes a 3×1 vector of error terms.

We report the point estimates and standard errors for the elements of the $\mathbf{\Theta}$ transition matrix in Table 8. Panel (a) of Table 8 reveals that a 10%pt increase in the number of out-of-town second house purchases as a fraction of all purchases in an MSA i in month t is associated with a 0.21%pt increase in the house price appreciation rate in the subsequent month. To get a better sense of the size of this relationship at the yearly horizon in the presence of the other variables, we compute the cumulative change in the house price appreciation in Las Vegas in response to a 10%pt increase in the fraction of purchases made by out-of-town second house buyers as was experienced during the mid 2000s. We plot this impulse response function in Figure 2 above and calculate that this one time 10%pt increase corresponds to a 6%pt per year increase in the annual house price appreciation rate or roughly 1/5th the realized 30%pt jump in the annual house price appreciation rate in Las Vegas during the boom.

House Price Appreciation Rate Panel VAR

(a) Dep. Var.: House Price Appreciation Rate

	Estimate	Std. Error		
Lagged Price Apprec. Rate	0.864	0.012	0.022	0.025
Lagged Percent Out-of-Town	0.021	0.007	0.009	0.009
Lagged Percent Local	−0.007	0.004	0.005	0.006
	Clustering	$\emptyset$	t	i

(b) Dep. Var.: Out-of-Town Second House Buyer Percent

	Estimate	Std. Error		
Lagged Price Apprec. Rate	0.080	0.019	0.025	0.023
Lagged Percent Out-of-Town	0.882	0.011	0.017	0.014
Lagged Percent Local	0.012	0.006	0.006	0.011
	Clustering	$\emptyset$	t	i

(c) Dep. Var.: Local Second House Buyer Percent

	Estimate	Std. Error		
Lagged Price Apprec. Rate	0.137	0.038	0.046	0.024
Lagged Percent Out-of-Town	0.065	0.021	0.025	0.021
Lagged Percent Local	0.824	0.013	0.013	0.025
	Clustering	$\emptyset$	t	i

Table 8. *Parameter values and standard errors of the transition matrix* $\boldsymbol{\Theta}$ *specified in Equation* (9) *estimated using three panel regressions on monthly data for the* 21 *MSAs from February 2000 to December 2007 using* $N = 1995$ *observations. Fixed effect estimates of* $\mathbf{A}_i$ *and* $\mathbf{K}_t$ *are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across MSAs. Reads: "A* $10\%_{\mathrm{pt}}$ *increase in the number of out-of-town second house purchases as a percent of all purchases in Los Angeles in April is associated with a* $0.21\%_{\mathrm{pt}}$ *increase in the house price appreciation rate in Los Angeles in May."*

This estimate is likely biased downward for 2 reasons. First, the impulse response calculation only considers a one time shock whose effect peaks out at around the 6 month horizon as Figure 2 shows. However, MSAs such as Las Vegas realized sustained increases in the fraction of all sales made by out-of-town second house buyers that lasted for upwards of 2 years. Second, group fixed effects tend to explain absorb too much variation in panel VARs with short time series as suggested in Nickell (1981). For intuition, recall that in finite samples principle component analysis over-estimates the size of the first principle component and leaves too little variation to be explained by subsequent factors. Similarly, by picking the group fixed effects that best explain the average level of the group, panel VARs on data with a short time series dimension tend to assign too much of the variation across groups to the group fixed effects $\mathbf{A}_i$ and leave too little to be explained by the transition matrix $\boldsymbol{\Theta}$.

Note that while an increase in the fraction of purchase made by out-of-town second house buyers predicts and increase in house price appreciation rates over the next year, an increase in the fraction of purchases made by local second house buyers has a *negative* but statistically insignificant effect. Thus, it is clear that not all second house buyers have the same price impact in this market. Some speculators were created more equal than others.

IAR Appreciation Rate Panel VAR

(a) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.438	0.021	0.076	0.092
Lagged Percent Out-of-Town	0.073	0.014	0.015	0.029
Lagged Percent Local	0.004	0.008	0.010	0.012
	Clustering	$\emptyset$	t	i

(b) Dep. Var.: Out-of-Town Second House Buyer Percent

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.046	0.016	0.018	0.017
Lagged Percent Out-of-Town	0.888	0.011	0.017	0.014
Lagged Percent Local	0.012	0.006	0.005	0.011
	Clustering	$\emptyset$	t	i

(c) Dep. Var.: Local Second House Buyer Percent

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.088	0.031	0.033	0.026
Lagged Percent Out-of-Town	0.074	0.021	0.024	0.023
Lagged Percent Local	0.824	0.013	0.013	0.025
	Clustering	$\emptyset$	t	i

Table 9. *Parameter values and standard errors of the transition matrix* $\mathbf{\Theta}$ *specified in Equation* (9) *estimated using three panel regressions on monthly data for the* 21 *MSAs from February 2000 to December 2007 using* $N = 1995$ *observations but with IAR rather than house price appreciation rates. Fixed effect estimates of* $\mathbf{A}_i$ *and* $\mathbf{K}_t$ *are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across MSAs. Reads: "A* 10%pt *increase in the number of out-of-town second house purchases as a fraction of all purchases in Los Angeles in April is associated with a* 0.73%pt *increase in the IAR price appreciation rate in Los Angeles in May."*

4.2. **IAR Appreciation Rate Regressions.** The results in the previous subsection suggest that an increase in the fraction of purchases made by out-of-town second house buyers in a given month predicts an increase in the house price appreciation rate in the subsequent month. However, large price movements do not necessarily indicate mispricing; instead, these movements in price could be due to fluctuations in housing market fundamentals. In order to address this concern, we augment our analysis with a similarly specified panel VAR regression using the monthly IAR appreciation rate rather than the monthly house price appreciation rate:

$$\mathbf{Y}_{i,t} = \begin{bmatrix} \Delta \log Z_{i,(t-1)\to t} & \frac{S_{i,t}^{\mathrm{OoT}}}{X_{i,t}} & \frac{S_{i,t}^{\mathrm{Lcl}}}{X_{i,t}} \end{bmatrix}^{\top} \tag{10}$$

and report these estimates in Table 9.

Comparing Panel (a) in Tables 8 and 9 reveals that the predictive power of an increase in the percent of all purchases made by out-of-town second house buyers is even larger when examining IAR appreciation rates rather than house price appreciation rates. e.g., a 10%pt increase in the number of out-of-town second house purchases as a percent of all purchases in an MSA i in month t is associated with a 0.73%pt increase in the IAR appreciation rate in the subsequent month. This evidence suggests that out-of-town second house buyer demand shocks appreciably distort the own

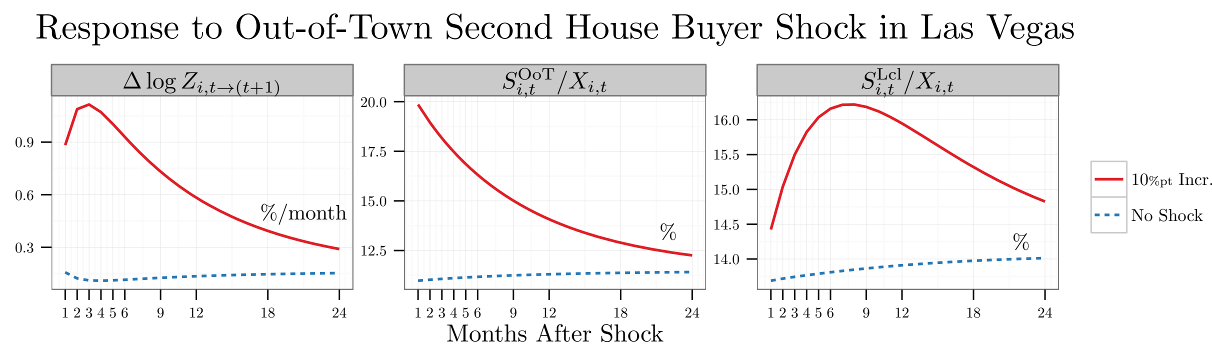


Figure 5. *Response of IAR appreciation rates, out-of-town second house buyer demand, and local second house buyer demand both to (red, solid) a* 10%pt *increase in the fraction of sales made to out-of-town second house buyers in Las Vegas as well as to (blue, dashed) no shock. We compute the figure using the panel VAR in Equation* (9) *whose point estimates calculated over the time period from January 2000 to December 2007 are housed in Table 8. Reads: "In the* 12 *months following a* 10%pt *increase in the percent of purchases made by out-of-town second house buyers in Las Vegas, IAR appreciation rates will rise by roughly* 9%pt*."*

vs. rent calculus of people living in the target MSA. As before, we find very little evidence that an increase in purchases by local second house buyers affects mispricing. The predictive power of an increase in the fraction of purchases made by these local speculators is a tightly estimated zero in Table 9.

Computing the impulse response to the same 10%pt increase in the percent of purchases made by out-of-town second house buyers on IAR rather than house price appreciation rates in Las Vegas as shown in Figure 5 reveals that this shock again explains around 1/5th the increase in mispricing in Las Vegas. This estimate remains relatively unchanged even though the point estimates on the lagged out-of-town second house buyer percent in Table 9 are much larger than those in 9 because IAR appreciation rates are substantially less predictable than the house price appreciation rates. The autoregressive coefficient falls from 0.864 in Table 8 to 0.438 in Table 9, while the adjusted R^2 drops from 0.75 to only 0.26. As further evidence, note that the peak effect of the 10%pt shock in Figure 5 occurs at 3 months rather than at 6 months in the house price appreciation rate calculation in Figure 2. Switching from house price appreciation rates to IAR appreciation rates leaves the point estimates in Panels (b) and (c) of Table 9 nearly unchanged as compared to Table 8 which gives a check on the stability of the estimation procedure. Moreover, Table 10 shows that our results are not dependent upon the precise way that the expected price appreciation is computed in the IAR appreciation rate calculation. It shows that a 10%pt increase in the number of out-of-town second house purchases as a fraction of all purchases in Los Angeles in April is associated with a 0.33%pt increase in the IAR appreciation rate in Los Angeles in May even when using 5 year moving average for the expected price appreciation in the user cost calculation.

5. Reverse Causality

The previous sections give evidence that out-of-town second house buyers behaved like misinformed speculators in the mid 2000s, and increases in demand from these traders are *associated*

IAR Appreciation Rate Panel VAR Using Moving Average for House Price Appreciation Rates

(a) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.626	0.032	0.032	0.047
Lagged Percent Out-of-Town	0.033	0.010	0.010	0.013
Lagged Percent Local	0.002	0.007	0.007	0.007
	Clustering	$\emptyset$	t	i

(b) Dep. Var.: Out-of-Town Second House Buyer Percent

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.087	0.027	0.030	0.027
Lagged Percent Out-of-Town	0.887	0.016	0.017	0.014
Lagged Percent Local	0.011	0.006	0.005	0.011
	Clustering	$\emptyset$	t	i

(c) Dep. Var.: Local Second House Buyer Percent

	Estimate	Std. Error		
Lagged IAR Apprec. Rate	0.120	0.053	0.052	0.039
Lagged Percent Out-of-Town	0.076	0.016	0.024	0.022
Lagged Percent Local	0.824	0.012	0.013	0.025
	Clustering	$\emptyset$	t	i

Table 10. *Parameter values and standard errors of the transition matrix* $\mathbf{\Theta}$ *specified in Equation* (9) *estimated using three panel regressions on monthly data for the* 21 *MSAs from February 2000 to December 2007 using* $N = 1995$ *observations but with IAR appreciation rates calculated using* 5 *year moving average price appreciation rates. Fixed effect estimates of* $\mathbf{A}_i$ *and* $\mathbf{K}_t$ *are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across MSAs. Reads: "A* $10\%_{\mathrm{pt}}$ *increase in the number of out-of-town second house purchases as a fraction of all purchases in Los Angeles in April is associated with a* $0.33\%_{\mathrm{pt}}$ *increase in the IAR appreciation rate in Los Angeles in May even when using* 5 *year moving average for the expected price appreciation in the user cost calculation."*

with higher house price and IAR appreciation rates in subsequent months. This section investigates the obvious follow up question: "Are these associations causal?" Using a novel econometric approach, we show that out-of-town second house buyer demand cannot be entirely explained by changes in fundamentals. Subsection 5.1 outlines our identification strategy. The crucial insight is that an increase in the fundamental value of owning a second house in Phoenix represents a common shock to the investment opportunity set of all potential second house buyers. Thus, if changes to fundamentals were driving both price dynamics as well as out-of-town second house buyer demand, we would expect to see large jumps in house price and IAR appreciation rates preceded by increases in out-of-town second house buyer demand from across the country. Subsection 5.2 then shows that the data do not display this symmetric response. As a result, they are inconsistent with an explanation based on reverse causality. To test this prediction econometrically, we condition our regressions on city size. For instance, we find that a $10\%_{\mathrm{pt}}$ increase in demand for second houses in Phoenix by buyers living in Los Angeles predicts a bigger increase in Phoenix house price and IAR appreciation rates than a $10\%_{\mathrm{pt}}$ increase in demand by buyers living in Milwaukee. We include month and ordered city pair fixed effects in our specification to account for variation in macroeconomic conditions and the fact that potential second house buyers in San Francisco are more likely to buy a in Las Vegas than Cleveland. Finally, in Subsection 5.3, we consider additional specifications to investigate the robustness of the main results.

5.1. **Empirical Predictions.** Consider a short example to cement the intuition behind our identification strategy. Appendix A gives a fully fledged economic model. Consider 2 groups of people living in Los Angeles and Milwaukee respectively that are each thinking about buying a second house in Phoenix. For the purposes of this example, suppose that there are roughly 10 times as many potential second house buyers living in Los Angeles, 10000, as there are living in Milwaukee, 1000. First, consider the null hypothesis, h_0, that *a common shock to the fundamental value of owning a second house in Phoenix is both attracting out-of-town second house buyers and driving up prices.* In this world, potential second house buyers in both Los Angeles and Milwaukee see the same shock (the shock to fundamentals), and will always increase their demand by 10%pt at the exact same time. As a result, Phoenix will always see aggregate out-of-town second house buyer demand shocks of either 1100 buyers or none at all. Here is the key observation. Under the null hypothesis, the relative sizes of Los Angeles and Milwaukee are irrelevant. Because out-of-town second house buyers are responding to a common Phoenix-specific shock it doesn't matter how they are distributed across the country. Now, consider the alternative hypothesis, h_A, that *shocks to fundamentals in Phoenix are not the only thing attracting out-of-town second house buyers.* For instance, the Milwaukee Journal Sentinel might have run a glowing review of Phoenix as a winter getaway, or Home and Garden Television might have played a Phoenix-based episode of Flip That House in Los Angeles cable market. In this world, potential second house buyers in Los Angeles and Milwaukee will not in general see the same shocks. As a result, Phoenix may see a 1000 buyer demand shock from Los Angeles in one month and then a 100 buyer demand shock from Milwaukee in the next. It's only in this alternative world that the 10%pt increase in demand from the bigger city has a larger price impact. It's only in this alternative world that the geographic distribution of potential second house buyers matters.

To implement this identification strategy, we need to calculate both the number of potential out-of-town second house buyers living in each MSA i in every month t as well as the time varying demand per out-of-town second house buyer in MSA i for second houses in MSA j at month t.

Definition 9 (Number of Out-of-Town Second House Buyers)**.** *Let Q_i denote the number of potential out-of-town second house buyers in MSA i measured as the average annualized number of second house purchases made by buyers living in MSA i each month over the period from January 2000 to December 2007 so that $T = 96$:*

$$Q_i = \frac{1}{96} \cdot \sum_{t=1}^{96} \left(\sum_{i \neq j} S_{i \to j,t} \right) \tag{11}$$

Definition 10 (Out-of-Town Second House Buyer Share)**.** *Let $\theta_{i \to j,t}$ denote demand for houses in*

MSA j at time t by buyers in MSA i as a fraction of the number of second house buyers in MSA i:

$$\theta_{i\to j,t} = \frac{S_{i\to j,t}}{Q_i} \tag{12}$$

where $\theta_{i\to j,t}$ has units of houses per trader.

Using these variables, we estimate Equation (13), which studies the relationship between the house price appreciation rate from time t to time $t+1$ and the proportion of second house buyers in each MSA i that purchase an out-of-town second house in MSA j at time t represented by the coefficient γ on the variable $\theta_{i\to j,t}$:

$$\Delta \log P_{j,t\to(t+1)} = \beta \cdot \Delta \log P_{j,(t-1)\to t} + \gamma \cdot \theta_{i\to j,t} + \alpha_{i\to j} + \kappa_t + \varepsilon_{i\to j,t}, \qquad i \neq j \tag{13}$$

The ordered MSA pair dummy variables control for two key effects as displayed in Equation (14) below:

$$\alpha_{i\to j} = \bar{\alpha}_j - \gamma \cdot \mathrm{E}[\theta_{i\to j,t}] \tag{14}$$

First, each $\alpha_{i\to j}$ accounts for the mean house price appreciation rate $\bar{\alpha}_j$ in each MSA j over the sample period (or the mean IAR appreciation rate over the same time period). Second, each $\alpha_{i\to j}$ adjusts the predicted house price appreciation rate (or IAR appreciation rate) in MSA j for the average rate at which second house buyers living in MSA i purchase second houses in MSA j. For instance, $\gamma\cdot\mathrm{E}[\theta_{(\mathrm{SFO},j),t}]$ differentially controls for the tendency of potential out-of-town second house buyers living in San Francisco to purchase more houses in Phoenix than in Milwaukee:

$$\mathrm{E}[\theta_{(\mathrm{SFO,PHX}),t}] \neq \mathrm{E}[\theta_{(\mathrm{PHX,SFO}),t}] \neq \mathrm{E}[\theta_{(\mathrm{SFO,MIL}),t}] \tag{15}$$

We estimate all regression equations in this section using a panel dataset at a monthly frequency from February 2000 to December 2007 on the $21 \times 20 = 420$ ordered MSA pairs with all $i = j$ pairs removed. Observations from January 2000 are removed due to the missing 1 month lagged values yielding a balanced panel of 39900 observations. We also consider the specification outlined in Equation (13) replacing the house price appreciation rate in MSA j from time t to time $t+1$ with the IAR appreciation rate from time t to $t+1$. Consistent with the results in Section 4, we expect to estimate a positive γ for both specifications indicating that, for instance, IAR appreciation rates rise by $\gamma\%_{\mathrm{pt}}$ per month in MSA j when the proportion of second house buyers in MSA $i \neq j$ that invest in MSA j increases by $1\%_{\mathrm{pt}}$.

Next, we augment this baseline specification in order to investigate the null hypothesis that second house buyers in all MSAs $i \in \{I \setminus j\}$ proportionally increase their demand for houses in MSA j after appropriate controls. We do this by including an interaction between the number of second house buyers in MSA i and the proportion of these speculators buying houses in MSA j. Specifically, we define the three indicator variables below which divide the set of 21 MSAs in our sample into terciles based on the number of second house buyers where 1_i^{Sml} denotes one of the 7 MSAs with the fewest potential out-of-town second house buyers, 1_i^{Med} denotes the next 7 MSAs which have a moderate number of potential out-of-town second house buyers, and 1_i^{Lrg} denotes the 7 MSAs with the most potential out-of-town second house buyers. We then estimate the regression specification

House Price Appreciation Rate Reverse Causality Specification

(a) Dep. Var.: House Price Appreciation Rate

	Estimate	Std. Error		
Lagged House Price Appreciation Rate	0.853	0.003	0.021	0.005
Out-of-Town Second House Buyer Share	0.213	0.024	0.059	0.035
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.753		

(b) Dep. Var.: House Price Appreciation Rate

	Estimate	Std. Error		
Lagged House Price Appreciation Rate	0.869	0.003	0.021	0.006
Out-of-Town Second House Buyer Share	0.052	0.038	0.052	0.040
Medium MSA × Out-of-Town Buyer Share	0.253	0.055	0.058	0.057
Large MSA × Out-of-Town Buyer Share	0.318	0.055	0.072	0.090
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.755		

Table 11. *Panel (a): Coefficient estimates from Equation (13). Panel (b): Coefficient estimates from Equation (16). All regressions use monthly data from February 2000 to December 2007 on the* 420 *ordered MSA pairs with all $i = j$ pairs removed. Fixed effect estimates of $\alpha_{i\to j}$ and κ_t are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across ordered MSA pairs. Reads: "A* 10%pt *increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a* 2.13%pt *increase in monthly house price appreciation rates in Phoenix in February 2005. This impact jumps to* 3.70%pt *if the out-of-town second house buyer choice shock happens in Los Angeles (a large MSA) and falls to* 0.52%pt *if the out-of-town second house buyer choice shock happens in Milwaukee (a small MSA)."*

in Equation (16) below where δ_{Med} and δ_{Lrg} have units of houses per person per month:

$$\begin{aligned}\Delta \log P_{j,t\to(t+1)} &= \alpha_{i\to j} + \kappa_t + \beta \cdot \Delta \log P_{j,(t-1)\to t} + \gamma \cdot \theta_{i\to j,t} \\ &\quad + \delta_{\text{Med}} \cdot 1_i^{\text{Med}} \cdot \theta_{i\to j,t} + \delta_{\text{Lrg}} \cdot 1_i^{\text{Lrg}} \cdot \theta_{i\to j,t} + \varepsilon_{i\to j,t} \qquad i \neq j \end{aligned} \tag{16}$$

using both house price appreciation rates and IAR appreciation rates.

What will the coefficients from Equation (16) tell us? First, if the null hypothesis h_0 that a common shock to the fundamental value of owning a second house in Phoenix is both attracting out-of-town second house buyers and driving up prices is true, then we should find $\delta_{\text{Med}} = \delta_{\text{Lrg}} = 0$. i.e., a 1%pt increase in the demand per potential buyer living in Los Angeles (a large market) for second houses in Phoenix should be equally predictive of an increase in house price appreciation rates in Phoenix as a 1%pt increase in the demand per potential buyer living in Milwaukee (a small market). Next, if the alternative hypothesis h_A that shocks to fundamentals in Phoenix are not the only thing attracting out-of-town second house buyers is true, then we should find $\delta_{\text{Lrg}} > \delta_{\text{Med}} > 0$. i.e., that the geographic distribution of potential second house buyers matters. House price and IAR appreciation rates are higher in Phoenix following a demand shock from Los Angeles as opposed to a demand shock from Milwaukee.

IAR Appreciation Rate Reverse Causality Specification

(a) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Appreciation Rate	0.505	0.005	0.073	0.020
Out-of-Town Second House Buyer Share	0.769	0.049	0.115	0.090
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.254		

(b) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Appreciation Rate	0.503	0.005	0.073	0.020
Out-of-Town Second House Buyer Share	0.356	0.079	0.088	0.100
Medium MSA $\times$ Out-of-Town Buyer Share	0.493	0.115	0.117	0.173
Large MSA $\times$ Out-of-Town Buyer Share	0.821	0.115	0.169	0.243
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.255		

Table 12. *Panel* (a)*: Coefficient estimates from Equation* (13) *using IAR appreciation rates rather than price appreciation rates as the dependent variable. Panel* (b)*: Coefficient estimates from Equation* (16) *using IAR appreciation rates rather than price appreciation rates as the dependent variable. All regressions use monthly data from February 2000 to December 2007 on the* 420 *ordered MSA pairs with all* $i = j$ *pairs removed. Fixed effect estimates of* $\alpha_{i \to j}$ *and* κ_t *are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across ordered MSA pairs. Reads: "A* 10%pt *increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a* 7.69%pt *increase in monthly IAR appreciation rates in Phoenix in February 2005. This impact jumps to* 11.77%pt *if the out-of-town second house buyer choice shock happens in Los Angeles (a large MSA) and falls to* 3.56%pt *if the out-of-town second house buyer choice shock happens in Milwaukee (a small MSA)."*

5.2. **Estimation Results.** The evidence presented in Tables 11 and 12 strongly rejects the null hypothesis, h_0, that common shocks to the fundamental value of owning a second house in, say, Phoenix are both attracting out-of-town second house buyers and driving up prices. The tables show that the geographic distribution of potential second house buyers matters in both a statistically measurable as well as an economically meaningful way. Consistent with our alternative hypothesis, h_A, unobserved shocks to fundamentals cannot fully explain the behavior of out-of-town second house buyers.

Panel (a) in both Table 11 and Table 12 reports the estimated coefficients and standard errors from Equation (13) using both price and IAR appreciation rates as the dependent variable and indicates that an increase in the fraction of potential out-of-town second house buyers investing in an MSA has a positive and statistically significant effect. The point estimate for in Table 11 implies that a 10%pt increase in the fraction of out-of-town second house buyers purchasing a house in, say, Las Vegas predicts a 2.13%pt increase in the monthly house price appreciation rate in Las Vegas. Similarly, in Table 12, a 10%pt increase in the fraction of out-of-town second house buyers purchasing a house in Las Vegas predicts a 7.69%pt increase in the IAR appreciation rate in the subsequent month. i.e., when more out-of-town second house buyers pile into a city, renting begins to look like a more and more attractive option in the subsequent months.

Next, looking at Panel (b) in both Tables 11 and 12 we see that the coefficient δ_{Lrg} in both Equation (16) is statistically different from zero. The impact on house price appreciation rates of a 10%pt increase in the fraction of potential out-of-town second house buyers living in a large MSA (e.g., Los Angeles) and investing in, say, Phoenix is almost twice as large as that of a 10%pt increase in the fraction of potential out-of-town second house buyers living in a small MSA (e.g., Milwaukee). As well, the ordering of the interaction terms is consistent with the alternative hypothesis that demand from distant speculators causes house price and IAR appreciation rates to increase. In all specifications $\delta_{\text{Lrg}} \geq \delta_{\text{Med}} \geq 0$. Panel ($b$) of Table 11 reads that while a 10%pt increase in January 2004 in the fraction of potential out-of-town second house buyers living in Milwaukee and investing in Phoenix predicts an 0.52%pt increase in the monthly house price appreciation rate in February 2004, that same 10%pt increase in the fraction of potential out-of-town second house buyers living in los Angeles and investing in Phoenix predicts a $0.52 + 3.18 = 3.70$%pt increase in the monthly house price appreciation rate in Phoenix in February 2004.

5.3. **Robustness Checks.** In this subsection, we discuss possible ways that the regression specified in Equation (16) might lead to spurious conclusions and then describe how we address these issues. First, while the baseline specification assumes that the house price and IAR appreciation rates in all MSAs realize a common shock κ_t in each month, macroeconomic forces during our sample period likely affected potential out-of-town second house buyers living in different MSAs in different ways. For example, potential buyers living in New York might always have more accurate beliefs about the fundamental value of owning a second house than potential buyers living in Milwaukee. In Table 13 we re-run the specifications using home MSA by month rather than simply month fixed effects to account for this concern as shown below:

$$\begin{aligned}\Delta \log P_{j,t\to(t+1)} = \alpha_j + \kappa_{i,t} + \beta \cdot \Delta \log P_{j,(t-1)\to t} + \gamma \cdot \theta_{i\to j,t} & \\ + \ \delta_{\text{Med}} \cdot 1_i^{\text{Med}} \cdot \theta_{i\to j,t} + \delta_{\text{Lrg}} \cdot 1_i^{\text{Lrg}} \cdot \theta_{i\to j,t} + \varepsilon_{i\to j,t} & \quad i \neq j \end{aligned} \tag{17}$$

where the $\kappa_{i,t}$ terms capture the time varying effect of shocks to different out-of-town second house buyer home MSAs. We find that the previous results are qualitatively unaffected by including home MSA by month fixed effect. e.g., after controlling for the average level of out-of-town second house buying by traders living in Los Angeles and Milwaukee each month, it is still true that a 10%pt increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a 1.47%pt increase in monthly house price appreciation rates in February if the buyer choice shock occurs in Los Angeles (a big MSA) but only a 0.28%pt increase if the buyer choice shock happens in Milwaukee (a small MSA).

This result means that out-of-town second house buyer demand is a fundamentally networked phenomenon. Any explanation must suggest particular *pairs* of cities at particular times. Any story that is specific to a particular city or a special date just doesn't cut it. e.g., in order to confound our results, potential second house buyers in Los Angeles must get differentially better information about the fundamental value of housing in Phoenix than potential second house buyers living in Milwaukee in an extremely precise way: (a) traders living in Los Angeles would need to receive extremely good information about buying a second house in Phoenix only during the period from

(i,t) and j Fixed Effects Reverse Causality Specification

(a) Dep. Var.: House Price Appreciation Rate

	Estimate	Std. Error		
Lagged House Price Appreciation Rate	0.855	0.003	0.025	0.004
Out-of-Town Second House Buyer Share	0.028	0.031	0.021	0.033
Medium MSA × Out-of-Town Buyer Share	0.098	0.042	0.038	0.046
Large MSA × Out-of-Town Buyer Share	0.119	0.040	0.051	0.043
	Clustering	$\emptyset$	(i,t)	j
	N		39900	
	R^2		0.747	

(b) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Appreciation Rate	0.457	0.004	0.096	0.018
Out-of-Town Second House Buyer Share	0.242	0.065	0.064	0.064
Medium MSA × Out-of-Town Buyer Share	0.250	0.089	0.096	0.097
Large MSA × Out-of-Town Buyer Share	0.250	0.084	0.153	0.082
	Clustering	$\emptyset$	(i,t)	j
	N		39900	
	R^2		0.227	

Table 13. *Panel (a): Coefficient estimates from Equation (16) with (i,t) and $i \to j$ fixed effects. Panel (b): Coefficient estimates from Equation (16) with (i,t) and $i \to j$ fixed effects using IAR appreciation rates rather than price appreciation rates as the dependent variable. All regressions use monthly data from February 2000 to December 2007 on the 420 ordered MSA pairs with all $i = j$ pairs removed. Fixed effect estimates of $\alpha_{i \to j}$ and $\kappa_{i,t}$ are omitted for clarity. Standard errors are estimated three different ways to account for clustering over home MSA by time or across ordered MSA pairs. Reads: "After controlling for the average level of out-of-town second house buying by traders living in Los Angeles and Milwaukee each month, it is still true that a 10%pt increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a 1.47%pt increase in monthly house price appreciation rates in February if the buyer choice shock occurs in Los Angeles (a big MSA) but only a 0.28%pt increase if the buyer choice shock happens in Milwaukee (a small MSA)."*

2004 through 2006 when Phoenix realized its highest house price appreciation rates, and (b) this information could only have applied to Phoenix and not to other cities that traders in Los Angeles might have invested in. Thus, while a time-varying ordered city pair correlation between both house price and IAR appreciation rates and out-of-town second house buyer shares would confound our results, it is difficult to think of such an explanation.

Second, perhaps the size of potential out-of-town second house buyers' belief distortions are random variables that move over time. In such a world, covariance between the home MSA size and the size of potential out-of-town second house buyers' belief distortion may bias our results. To address this concern, in Table 14 we again re-run our analysis, only this time we instead compute the number of potential out-of-town second house buyers living in each MSA using the ranking in 2000. e.g., let $\widehat{Q}_i$ denote the number of potential out-of-town second house buyers living in MSA i

Year 2000 Ranking Period Reverse Causality Specification

(a) Dep. Var.: House Price Appreciation Rate

	Estimate	Std. Error		
Lagged House Price Appreciation Rate	0.869	0.003	0.021	0.006
Out-of-Town Second House Buyer Share	0.069	0.015	0.021	0.025
Medium MSA × Out-of-Town Buyer Share	−0.012	0.020	0.013	0.032
Large MSA × Out-of-Town Buyer Share	0.142	0.028	0.036	0.056
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.754		

(b) Dep. Var.: IAR Appreciation Rate

	Estimate	Std. Error		
Lagged IAR Appreciation Rate	0.504	0.005	0.073	0.020
Out-of-Town Second House Buyer Share	0.205	0.033	0.036	0.035
Medium MSA × Out-of-Town Buyer Share	0.019	0.042	0.028	0.069
Large MSA × Out-of-Town Buyer Share	0.436	0.059	0.082	0.149
	Clustering	$\emptyset$	t	$i \to j$
	N	39900		
	R^2	0.255		

Table 14. *Panel (a): Coefficient estimates from Equation (16) with Q_i estimated over the period from January 2000 to December 2000. Panel (b): Coefficient estimates from Equation (16) with Q_i estimated over the period from January 2000 to December 2000 using IAR appreciation rates rather than price appreciation rates as the dependent variable. All regressions use monthly data from February 2000 to December 2007 on the* 420 *ordered MSA pairs with all $i = j$ pairs removed. Fixed effect estimates of $\alpha_{i \to j}$ and κ_t are omitted for clarity. Standard errors are estimated three different ways to account for clustering over time or across ordered MSA pairs. Reads: "When ranking MSAs according to their potential out-of-town second house buyer populations in 2000, a* 10%pt *increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a* 2.11%pt *increase in monthly house price appreciation rates in February if the buyer choice shock occurs in Los Angeles (a big MSA) but only a* 0.69%pt *increase if the buyer choice shock happens in Milwaukee (a small MSA)."*

similarly defined but measured over the period from January 2000 to December 2000 so that $T = 12$:

$$\widehat{Q}_i = \frac{1}{12} \cdot \sum_{t=1}^{12} \left(\sum_{i \neq j} S_{i \to j,t} \right) \tag{18}$$

By contrast, the first definition of the number of potential second house buyers in each MSA i represents the sample average over the entire period from January 2000 to December 2007. Since this variable is computed using the entire time series, it is potentially simultaneously determined with investment opportunities in the largest markets for out-of-town second house buyers that appear attractive later in the sample period. e.g., some out-of-town second house buyers might only have entered the housing market because MSAs like Las Vegas and Phoenix appeared to have had great investment opportunities. This observation motivates the use of the second definition that includes only data from the year 2000 which predates the rapid rise in house price appreciation rates in all MSAs and minimizes the possibility for correlation between home MSA size and the level of belief distortion.

The results in Table 14 are a bit less robust than those in Tables 11 and 12 with the interaction terms having smaller coefficients, but present a consistent story. In all cases, the coefficients on the interaction of out-of-town second house buyer share and large MSAs is statistically different from zero no matter which clustering of standard errors we use. When ranking MSAs according to their potential out-of-town second house buyer populations in 2000, a 10%pt increase in the fraction of out-of-town second house buyers buying into Phoenix in January 2005 is associated with a 2.11%pt increase in monthly house price appreciation rates in February if the buyer choice shock occurs in Los Angeles (a big MSA) but only a 0.69%pt increase if the buyer choice shock happens in Milwaukee (a small MSA). In Table 14, the coefficient on the interaction with medium size cities is negative, but is not statistically different from zero. Finally, we observe that the empirical results are strongest in Panel (b) where we use the IAR appreciation rate as the dependent variable. To the extent that the IAR appreciation rate proxies for mispricing, these results present a consistent picture that out-of-town second house buyers contribute to mispricing.

6. Conclusion

We use transactions-level deeds records to show that out-of-town second house buyers behaved like misinformed speculators and drove up both house price and IAR appreciation rates. First, we give evidence that out-of-town second house buyers behaved like misinformed speculators relative to local second house buyers. Out-of-town buyers had poorer exit timing and were also less capable of consuming the dividend from their purchase. Second, we demonstrate that increases in out-of-town second house buyer demand predict increases in future house price and IAR appreciation rates. Specifically, a 10%pt increase in the fraction of sales made to out-of-town second house buyers is associated with a 6% increase in house price appreciation rates and a 9% increase in IAR appreciation rates. Third and finally, we confront the issue of reverse causality using a new econometric approach. Our central insight is that an increase in the fundamental value of owning a second house in Phoenix is a common shock to the investment opportunity set of potential second house buyers living everywhere. Thus, if it were the case that changes to fundamentals were driving both price dynamics as well as out-of-town second house buyer demand, then we should see large jumps in house price and IAR appreciation rates preceded by increases in out-of-town second house buyer demand from everywhere. The data do not display this symmetric response.

We conclude by discussing both the economic magnitudes of out-of-town second house buyer capital flows and the broader applicability of our econometric approach. First, to get a sense of the magnitudes, we examine how total out-of-town second house buyer purchases compare to the size of the local economy. Figure 7 plots the sum of the sales prices on out-of-town second house buyer purchases as a percent of MSA-specific gross domestic product from 2000 to 2007. The MSA-specific GDP data comes from the Bureau of Economic Analysis. These calculations treat all purchases as being net new capital coming from outside the MSA, whether financed by debt or equity since the majority of debt financing came via selling residential mortgage-backed securities. This figure shows that the sum of the sales prices in Las Vegas exceeded 5% of Las Vegas's GDP in 2004. Thus, demand shocks from out-of-town second house buyers appear to be quite substantial when compared

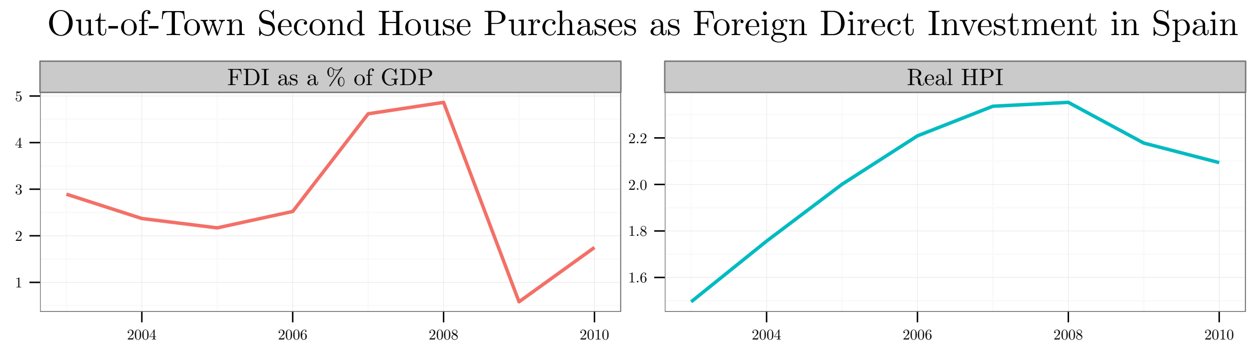


Figure 6. *Left Panel: Net foreign direct investment (henceforth, FDI) in Spain from the World Bank as a percent of Spain's GDP from 2003 to 2010. Reads: "Net FDI inflows into Spain amounted to a little less than* 5% *of Spain's GDP in 2008." Right Panel: Real HPI index level in Spain over this same time period. Reads: "The real HPI index level rose by just over* 230% *from a base of* 1 *in 2000."*

to the aggregate economic output of many MSA-level economies. As points of comparison, Barro and Ursúa (2008) define a 10% drop in the GDP of a country as an economic disaster, and Javorcik (2004) examines firm-level data in Lithuania and finds that foreign direct investment from the US on the order of 3.4% of the Lithuanian GDP in 2000 leads to substantial spillover effects in its real economy. Building on this analogy to FDI, we see an opportunity in future work to study the impact of these spillovers from the residential housing market on local economies.

Second, we emphasize that our econometric approach allows us to identify price movements that can't be solely explained by shocks to fundamentals. To our knowledge, this approach which completely sidesteps the thorny problem of instrumenting for misinformed speculator demand across assets is new to the finance and economics literature. The identification strategy applies wherever there is a) market segmentation and b) data on flows between segments. We conjecture that distant demand driven mispricings may not be a phenomenon confined to the US residential real estate market. For instance, Office for National Statistics (2007) found that 1.8Mil households in England owned a second home and, among these properties, 87k were in Spain and being used as part time residences during the peak of the Spanish housing boom. To give some idea of the scale of this investment expenditure by overseas second home buyers in Spain, in Figure 6 we plot the net foreign direct investment (henceforth, FDI) in Spain as a percent of Spain's GDP from 2003 to 2010 using data from the World Bank alongside the real HPI level in Spain over this same time period. We find that FDI as a percent of GDP spikes to just under 5% in 2008, a similar percentage to the total of outside purchases of homes in Las Vegas at peak, and that the timing of this spike corresponds to the peak of the HPI level. Data do not show a similar peak in FDI in other southern European counties. Similar events unfolded in the US commercial real estate market in the late 1980s when a 1986 tax code change made purchases of commercial real estate less attractive for US companies and invited a host of foreign investors from countries like Japan to large scale purchases of commercial office buildings. See Sagalyn (1999), which discuss the purchase of Rockefeller Center by Mitsubishi Trust, Co. for more than $1Bil in the late 1980. Thus, distant speculators may be an important class of traders playing a role in mispricing more generally and an interesting topic of future research.

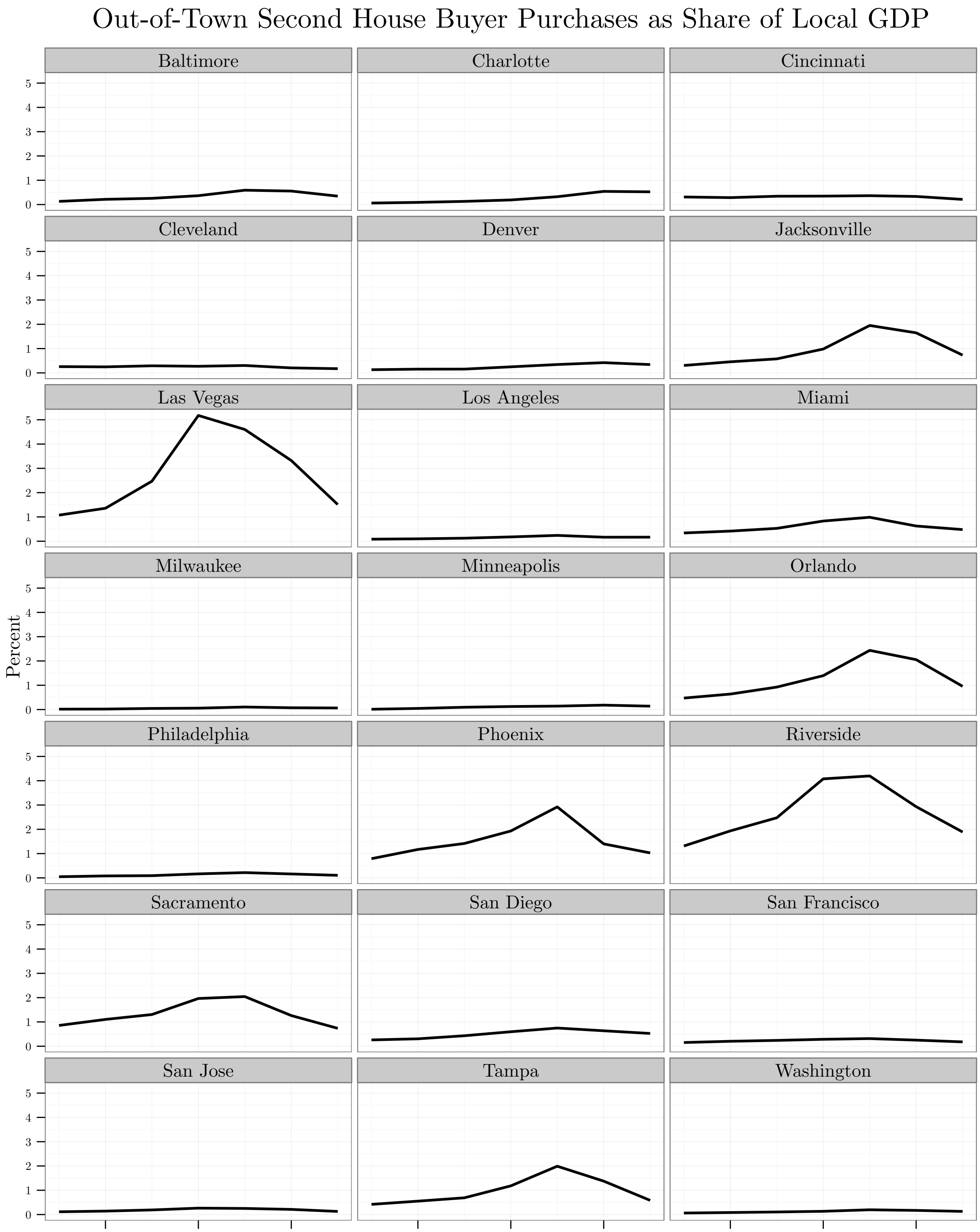


Figure 7. *Sum of the sales prices of single family houses sold to distant speculators as a percent of total GDP in each MSA from 2000 to 2007. We compute MSA-specific GDP using data from the BEA as the product of the per capita income in each MSA times the population. Reads: "The sum of the sales prices in Las Vegas exceeded* 5% *of the GDP for the entire MSA in 2004."*

References

Barro, R. and J. Ursúa (2008). Macroeconomic crises since 1870. *Brookings Papers on Economic Activity*.

Bayer, P., C. Geissler, and J. Roberts (2011). Speculators and middlemen: The role of flippers in the housing market. *National Bureau of Economic Research Working Paper Series*.

Black, F. (1986). Noise. *The Journal of Finance 41*.

Case, K. and R. Shiller (1989). The efficiency of the market for Single-Family homes. *The American Economic Review 79*.

Choi, H., H. Hong, and J. Scheinkman (2013). Speculating on home improvements. *Journal of Financial Economics forthcoming*.

Cohen, L., A. Frazzini, and C. Malloy (2010). Sell side school ties. *Journal of Finance 65*.

De Long, B., A. Shleifer, L. Summers, and R. Waldmann (1991). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance 45*.

Emrath, P. (2002). Property taxes in the 2000 census. *Housing Economics 12*.

Ferreira, F. and J. Gyourko (2011). Anatomy of the beginning of the housing boom: U.S. neighborhoods and metropolitan areas, 1993-2009. *National Bureau of Economic Research Working Paper Series*.

Flavin, M. and T. Yamashita (2002). Owner-Occupied housing and the composition of the household portfolio. *The American Economic Review 92*.

Friedman, M. (1953). *Essays in Positive Economics*. University of Chicago Press.

Glaeser, E., J. Gottlieb, and J. Gyourko (2010). Can cheap credit explain the housing boom? *National Bureau of Economic Research Working Paper Series*.

Glaeser, E. and J. Gyourko (2007). Arbitrage in housing markets. *National Bureau of Economic Research Working Paper Series*.

Harding, J., T. Miceli, and C. Sirmans (2000). Do owners take better care of their housing than renters? *Real Estate Economics 28*.

Haughwout, A., D. Lee, J. Tracy, and W. van der Klaauw (2011). Real estate investors, the leverage cycle, and the housing market crisis. *Federal Reserve Bank of New York Staff Reports 514*.

Himmelberg, C., C. Mayer, and T. Sinai (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *The Journal of Economic Perspectives 19*.

Hong, H., J. Kubik, and J. Stein (2005). Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. *The Journal of Finance 60*.

Hong, H. and J. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance 54*.

Hubbard, G. and C. Mayer (2009). The mortgage market meltdown and house prices. *The B.E. Journal of Economic Analysis and Policy 9*.

Javorcik, B. (2004). Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages. *American Economic Review 94*.

Kurlat, P. and J. Stroebel (2014). Testing for information asymmetries in real estate markets. *SSRN Working Paper*.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica 53*.

Levitt, S. and C. Syverson (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics 90*.

Li, W. and Z. Gao (2012). Real estate investors and the boom and bust of the us housing market. *SSRN*.

Lustig, H. and S. Van Nieuwerburgh (2005). Housing collateral, consumption insurance and risk premia: an empirical perspective. *The Journal of Finance 60*.

Mayer, C. (2010). Housing bubbles: A survey. *Annual Review of Economics 3*.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica 49*.

Ofek, E. and M. Richardson (2003). Dotcom mania: The rise and fall of internet stock prices. *The Journal of Finance 58*.

Office for National Statistics (2007). Community housing survey.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies 22*.

Poterba, J. (1984). Tax subsidies to Owner-Occupied housing: An Asset-Market approach. *The Quarterly Journal of Economics 99*.

Sagalyn, L. (1999). *Cases in Real Estate Finance and Investment Strategy*. Urban Land Institute.

Scheinkman, J. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy 111*.

Shleifer, A. and R. Vishny (1997). The limits of arbitrage. *The Journal of Finance 52*.

Zingales, L. (1995). What determines the value of corporate votes. *The Quarterly Journal of Economics 110*.

Appendix A. A Simple Model of Speculation

In this section we develop a simple noisy rational expectations model of the US residential housing market to clarify our identification strategy and better interpret our empirical results. We begin in Subsection A.1 by outlining the basic economic framework. Then, in Subsection A.2 we study the pricing implications in two alternative regimes. The first regime admits only fully informed traders while the second allows for misinformed traders as well. In Subsection A.3, we interpet our econometric analysis through the lens of the model.

A.1. **Economic Framework.** Consider a static housing market with $I \geq 1$ MSAs. The price of a house in MSA i is P_i and the true value of a house in MSA i is V_i where both P_i and V_i are measured in dollars per house. We model the true value of housing in each MSA i as an iid random variable drawn from a normal distribution $V_i \overset{\text{iid}}{\sim} \mathrm{N}(\mu_v, \sigma_v^2)$. There are Q_i traders in each MSA i indexed by $q = 1, 2, \ldots, Q_i$. Let $\vartheta_{q,i\to j}$ denote the number of houses in MSA j demanded by the q^{th} trader in MSA i and let $S_{i\to j}$ denote the total number of houses in MSA j demanded by traders in MSA i. We denote the average demand for houses in MSA j by traders living in MSA i as $\theta_{i\to j} = {}^1\!/_{Q_i} \cdot \sum_{q=1}^{Q_i} \vartheta_{q,i\to j}$ and can interpret this quantity as the probability that a randomly selected trader in MSA i buys a house in MSA j. Total demand for housing in MSA j, denoted X_j, is defined as the sum of the housing demand from each MSA i plus an MSA specific demand shock ε_j:

$$X_j = \sum_{i=1}^{I} S_{i\to j} + \varepsilon_j = \sum_{i=1}^{I} \left(\sum_{q=1}^{Q_i} \vartheta_{q,i\to j} \right) + \varepsilon_j = \sum_{i=1}^{I} \left(Q_i \cdot \theta_{i\to j} \right) + \varepsilon_j \tag{19}$$

where ε_j is an iid draw from a normal distribution $\varepsilon_j \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ and X_j has units of houses.

There is a collection of market makers who operate under perfect competition. These agents only observe the aggregate demand X_j in each MSA and as a result of perfect competition set the price level equal to the expected value of housing in MSA j given the realized aggregate demand:

$$P_j = \mathrm{E}\left[V_j | X_j\right] = \alpha + \beta \cdot X_j \tag{20}$$

The coefficient β can be interpreted as the dollar change in the price of housing in MSA j when traders demand one additional unit of housing in MSA j. Market makers might be developers or property managers who either build new housing units to match demand or reclaim unused housing units by turning them into rental properties or razing them to build office or industrial space.

Traders in each MSA i know the true value of housing in every other MSA j. For instance, in this view of the world a trader living in San Francisco that purchases a second house in Las Vegas knows the true value of housing in Las Vegas. The competitive market makers assume that traders use a linear demand rule given by:

$$\vartheta_{q,i\to j} = \gamma_{q,i\to j} + \delta_{q,i\to j} \cdot V_j \tag{21}$$

This is the standard ansatz for Kyle (1985) type models and can easily be verified in equilibrium. The coefficient $\gamma_{q,i\to j}$ has units of houses per trader and the coefficient $\delta_{q,i\to j}$ has units of houses per trader dollar. Each individual trader optimizes their value function $W_{q,i}$ by choosing how many houses to buy in each MSA j:

$$\begin{aligned} W_{q,i} &= \sum_{j=1}^{I} W_{q,i\to j} \\ W_{q,i\to j} &= \max_{\vartheta_{q,i\to j}} \mathrm{E}\left[(V_j - P_j)\cdot \vartheta_{q,i\to j} | V_j\right] \end{aligned} \tag{22}$$

Definition 11 (Equilibrium)**.** *An equilibrium consists of price parameters (α^*, β^*) and demand parameters $\{(\gamma^*_{q,i\to j}, \delta^*_{q,i\to j})\}$ for each trader over every ordered MSA pair such that:*

1. *(1) Given market makers follow the pricing rule in Equation (20), the housing demand schedule $\{\vartheta_{q,i\to j}\}_{i,j\in I}$ dictated by the demand rule parameters $\{(\gamma^*_{q,i\to j}, \delta^*_{q,i\to j})\}_{i,j\in I}$ solves each trader's optimization problem in Equation (22).*
2. *(2) Given all traders follow the demand rules specified in Equation (21), the price parameters (α^*, β^*) satisfy the expectations equality in Equation (20).*

A.2. **Equilibrium Housing Prices.** First, we solve for the equilibrium in this economy when all traders are fully informed. This equilibrium is identical to the standard Kyle (1985) equilibrium in all aspects except for the fact that each trader represents only $1/\sum_{i'=1}^{I} Q_{i'}$ of the total market demand. Thus parameters defining the number of houses demanded per trader $\theta_{i\to j}$ as well as the price impact of each trader's demand decisions $(\gamma_{i\to j}, \delta_{i\to j})$ are both deflated by a factor of $1/\sum_{i'=1}^{I} Q_{i'}$.

Proposition 1 (Fully Informed Equilibrium)**.** *When traders in all markets have correct beliefs about the true value of housing V_j in MSA j, traders in all MSAs demand the same number of houses in MSA j:*

$$\bar{\theta}_j = \theta_{1\to j} = \theta_{2\to j} = \cdots = \theta_{I\to j} \tag{23}$$

Proof. Substituting both the functional form for the housing price in MSA j from Equation (20) and the functional form for the aggregate demand in MSA j from Equation (19) into the objective function for an individual trader q from MSA i yields an expression:

$$\begin{aligned} W_{q,i\to j} &= \max_{\vartheta_{q,i\to j}} \mathrm{E}\left[(V_j - \alpha - \beta\cdot X_j)\cdot \vartheta_{q,i\to j} | V_j\right] \\ &= \max_{\vartheta_{q,i\to j}} \mathrm{E}\left[\left(V_j - \alpha - \beta\cdot \sum_{i'=1}^{I}\left(\sum_{q'=1}^{Q_{i'}} \vartheta_{q',i'\to j}\right) - \beta\cdot\varepsilon_j\right)\cdot \vartheta_{q,i\to j}\middle| V_j\right] \end{aligned} \tag{24}$$

Taking the derivative of this optimization program with respect to trader q's demand gives the first order condition:

$$0 = \mathrm{E}\left[\left(V_j - \alpha - \beta\cdot \sum_{i'=1}^{I}\left(\sum_{q'=1}^{Q_{i'}} \vartheta_{q',i'\to j}\right) - \beta\cdot\varepsilon_j\right) - 2\cdot\beta\cdot \vartheta_{q,i\to j}\middle| V_j\right] \tag{25}$$

where we assume $Q_i \approx Q_i - 1$ for simplicity. Evaluating the conditional expectation operator yields:

$$0 = V_j - \alpha - \beta\cdot \sum_{i'=1}^{I}\left(\sum_{q=1}^{Q_{i'}} \vartheta_{q,i'\to j}\right) - 2\cdot\beta\cdot \vartheta_{q,i\to j} \tag{26}$$

We then solve for $\vartheta_{q,i\to j}$ to derive the expression below:

$$\vartheta_{q,i\to j} = -\frac{\alpha + \beta \cdot \sum_{i'=1}^{I} \left(\sum_{q'=1}^{Q_{i'}} \vartheta_{q',i'\to j} \right)}{2 \cdot \beta} + \left(\frac{1}{2 \cdot \beta} \right) \cdot V_j \tag{27}$$

This expression would be identical for any trader q living in MSA $i \in I$ implying that $\theta_{i\to j} = \theta_{i'\to j}$ for all $i, i' \in \{1, 2, \ldots, I\}$. □

The key implication of this framework is that, in a world where all traders are fully informed, the proportion of traders from MSA i investing in MSA j is the same for each $i = 1, 2, \ldots, I$. i.e., variation in the housing demand in MSA j per person in MSA i is proportional to variation in the value of housing in MSA j as fluctuations in V_j represent a common shock. While full information is perhaps the most natural benchmark, note that the symmetry in Proposition 1 still holds if traders are not fully informed but instead similarly misinformed. For instance, if potential second house buyers in every MSA all over-valued housing in Phoenix by 10%, then traders in all MSAs would still demand the same number of houses in Phoenix—this common demand per trader would just be too high.

Next, we solve for an equilibrium when traders in some MSA i are misinformed about the value of housing in MSA j. Specifically, suppose that traders in MSA i believe that the value of a house in MSA j is $\tilde{V}_j = V_j + \eta$ dollars with $\eta > 0$ rather than the true value of V_j dollars assuming that traders in MSA i think that all other traders share the same beliefs. Let $\tilde{P}_j^{(i)}$ denote the price of housing in MSA j when traders from MSA i have overconfident beliefs about V_j.

Proposition 2 (Price Distortion)**.** *Suppose that misinformed traders in MSA i believe that the value of housing in MSA j is $\tilde{V}_j = V_j + \eta$ with $\eta > 0$. Then the price of a house in MSA j will be distorted by an amount proportional to the number of traders in MSA i:*

$$\tilde{P}_j^{(i)} - P_j = \left(\frac{Q_i}{\sum_{i'=1}^{I} Q_{i'}} \right) \cdot \frac{\eta}{2} \tag{28}$$

Proof. If the market makers do not realize that traders may be overconfident or uninformed, they will adopt the same pricing rule as in Proposition 1. What's more, both traders with correct beliefs in MSAs $i' \neq i$ and traders with overconfident beliefs in MSA i think that all other agents share their beliefs so that they anticipate a price in MSA j of:

$$\mathrm{E}\left[P_j | \mathrm{MSA}\right] = \begin{cases} \alpha^* + \beta^* \cdot \sum_{i'=1}^{I} Q_{i'} \cdot \left(\bar{\gamma}^* + \bar{\delta}^* \cdot V_j \right) & \text{if MSA} \neq i \\ \alpha^* + \beta^* \cdot \sum_{i'=1}^{I} Q_{i'} \cdot \left(\bar{\gamma}^* + \bar{\delta}^* \cdot \{V_j + \eta\} \right) & \text{if MSA} = i \end{cases} \tag{29}$$

However, the realized total demand in MSA j given that traders in MSA i have inflated beliefs, $\tilde{X}_j^{(i)}$, will be given by:

$$\begin{aligned} \tilde{X}_j^{(i)} &= \sum_{i' \neq i} Q_{i'} \cdot \left(\bar{\gamma}^* + \bar{\delta}^* \cdot V_j \right) + Q_i \cdot \left(\bar{\gamma}^* + \bar{\delta}^* \cdot \{V_j + \eta\} \right) \\ &= \sum_{i'=1}^{I} Q_{i'} \cdot \left(\bar{\gamma}^* + \bar{\delta}^* \cdot V_j \right) + Q_i \cdot \bar{\delta}^* \cdot \eta \end{aligned} \tag{30}$$

Thus, the difference between the price levels in MSA j in the fully informed regime and the regime with misinformed speculators will be given by $\tilde{P}_j^{(i)} - P_j = Q_i \cdot \beta^* \cdot \bar{\delta}^* \cdot \eta$. Substituting in the functional forms for the equilibrium coefficients β^* and $\bar{\delta}^*$ from Proposition 1 yields the desired result. □

This proposition is easiest to interpret via a short numerical example. Suppose that there are 55×10^6 traders split across 10 MSAs with the largest MSA i' containing 10×10^6 traders and the smallest MSA i'' containing only 1×10^6 traders. Then, the price increase in MSA j when traders from MSA i' or i'' alternately believe that housing values in MSA j are $\tilde{V}_j = V_j + \$5000$ are:

$$\tilde{P}_j^{(\text{MSA})} - P_j = \begin{cases} \left(\frac{10\times10^6}{55\times10^6}\right) \cdot \frac{\$5000}{2} = \$454.55 & \text{if MSA} = i' \\ \left(\frac{1\times10^6}{55\times10^6}\right) \cdot \frac{\$5000}{2} = \$45.45 & \text{if MSA} = i'' \end{cases} \tag{31}$$

In other words, when misinformed traders from a larger market attempt to purchase investment properties, they have a bigger impact on prices than misinformed traders from a smaller market.

A.3. **Empirical Strategy.** The goal of this simple model is to provide a scaffolding within which to better understand the empirical strategy we employ. First, we identify a group of misinformed speculators. Within the model, this task corresponds to identifying a group of traders who are likely to have misinformed beliefs about future price levels, i.e. an $\eta > 0$. In Section 3 we give evidence that out-of-town second house buyers satisfy this criteria. Second, we show that an increase in demand from out-of-town second house buyers predicts increases in house price and IAR appreciation rates. Within the model, this task is tantamount to checking if housing appears overpriced—i.e., that $P_j/\mathrm{E}[P_j] > 1$ or $\log P_j - \log \mathrm{E}[P_j] > 0$ after taking logs—when out-of-town second house buyers have above average demand. While the model is cast in levels, in the empirical implementation we study $\log P_{j,t} - \log P_{j,t-1}$ in place of $\log P_j - \log \mathrm{E}[P_j]$ under the assumption that $\mathrm{E}[P_j] = P_{j,t-1}$.

Finally, we address the issue of reverse causality. Within the model, this task corresponds to checking whether or not high realized prices in MSA j are due to high realized housing values V_j or to some group of traders in MSA i having misinformed beliefs $\eta > 0$. We exploit the natural geographic segmentation in the housing market to address this challenge. Proposition 1 demonstrates that if an increase in the price of housing in MSA j is due to an unobserved (from the point of view of an econometrician) increase in house values, then out-of-town second house buyers from each other MSA should increase their demand for housing in MSA j in equal proportions. In Section 5 we test for this symmetry and show it to be violated. From this evidence, we conclude that out-of-town second house buyers are not simply responding to unobserved information when making their purchases. In Proposition 2 we show that if out-of-town second house buyers from MSA i have a belief distortion η about the value of housing in MSA j, then the size of the resulting price distortion should be proportional to the share of traders residing in MSA i. We find exactly this pattern in the data; the correlation between the house price and IAR appreciation rates and the share of out-of-town second house buyers going from MSA i to MSA j is bigger when the total number of out-of-town second house buyers living in MSA i is larger. We interpret these results as evidence that MSA specific variation in out-of-town second house buyer beliefs about MSA j (perhaps due to local news sources or word of mouth) is contributing to the realized price distortion.