

Anatomy of the Beginning of the Housing Boom:  
U.S. Neighborhoods and Metropolitan Areas, 1993-2009

Fernando Ferreira and Joseph Gyourko  
The Wharton School, University of Pennsylvania  
& NBER

August 18, 2011

**Online Appendix**

## Appendix #1: Data Coverage

### Starting Dates for Each MSA in the Final Sample

start	msa_name	start	msa_name
1993q1	Providence-New Bedford-Fall River, RI-MA	1995q3	Nashville-Davidson--Murfreesboro, TN
1993q1	Sacramento--Arden-Arcade--Roseville, CA	1996q1	Flagstaff, AZ
1993q1	San Jose-Sunnyvale-Santa Clara, CA	1996q1	Kingston, NY
1993q1	Reno-Sparks, NV	1996q1	New York-Northern New Jersey-Long Island, NY-NJ-PA2/
1993q1	Portland-Vancouver-Beaverton, OR-WA	1996q1	Deltona-Daytona Beach-Ormond Beach, FL
1993q1	Olympia, WA	1996q1	Ocala, FL
1993q1	Pittsfield, MA	1996q1	Gainesville, FL
1993q1	Springfield, MA	1996q1	Port St. Lucie-Fort Pierce, FL
1993q1	Visalia-Porterville, CA	1996q1	Cape Coral-Fort Myers, FL
1993q1	Riverside-San Bernardino-Ontario, CA	1996q1	Knoxville, TN
1993q1	Tucson, AZ	1996q1	Yuma, AZ
1993q1	Oxnard-Thousand Oaks-Ventura, CA	1996q2	Panama City-Lynn Haven, FL
1993q1	Redding, CA	1996q2	Fort Walton Beach-Crestview-Destin, FL
1993q1	Modesto, CA	1996q3	Salem, OR
1993q1	Phoenix-Mesa-Scottsdale, AZ	1997q1	Barnstable Town, MA
1993q1	Merced, CA	1997q1	Erie, PA
1993q1	Hartford-West Hartford-East Hartford, CT	1997q1	Allentown-Bethlehem-Easton, PA-NJ
1993q1	Stockton, CA	1997q1	Palm Bay-Melbourne-Titusville, FL
1993q1	Madera, CA	1997q1	Sarasota-Bradenton-Venice, FL
1993q1	Bridgeport-Stamford-Norwalk, CT	1997q1	Tampa-St. Petersburg-Clearwater, FL
1993q1	Las Vegas-Paradise, NV	1997q1	Tallahassee, FL
1993q1	Fresno, CA	1997q1	Vero Beach, FL
1993q1	Seattle-Tacoma-Bellevue, WA	1997q1	Orlando, FL
1993q1	Napa, CA	1997q2	Baltimore-Towson, MD
1993q1	Hanford-Corcoran, CA	1997q2	Columbus, OH
1993q1	New Haven-Milford, CT	1997q2	Akron, OH
1993q1	Salinas, CA	1997q2	Lakeland-Winter Haven, FL
1993q1	Worcester, MA	1997q3	Jacksonville, FL
1993q1	Boston-Cambridge-Quincy, MA-NH	1997q3	Yakima, WA
1993q1	Bakersfield, CA	1997q3	Pensacola-Ferry Pass-Brent, FL
1993q1	Los Angeles-Long Beach-Santa Ana, CA	1997q3	Washington-Arlington-Alexandria, DC-VA-MD
1993q1	Norwich-New London, CT	1997q3	Cincinnati-Middletown, OH-KY-IN
1993q1	Vallejo-Fairfield, CA	1997q4	Springfield, OH
1993q1	Santa Rosa-Petaluma, CA	1998q1	Lincoln, NE
1993q1	San Francisco-Oakland-Fremont, CA	1998q1	Cleveland-Elyria-Mentor, OH
1993q2	Yuba City-Marysville, CA	1998q1	Chicago-Naperville-Joliet, IL-IN-WI
1993q3	Chico, CA	1998q1	Honolulu, HI
1994q1	Bremerton-Silverdale, WA	1998q1	Fort Collins-Loveland, CO
1994q1	San Diego-Carlsbad-San Marcos, CA	1998q1	Denver-Aurora, CO
1995q1	Corvallis, OR	1998q1	Dayton, OH
1995q1	Spokane, WA	1998q1	Detroit-Warren-Livonia, MI
1995q1	Eugene-Springfield, OR	1998q1	Colorado Springs, CO
1995q1	Medford, OR	1998q2	Oklahoma City, OK
1995q1	Bellingham, WA	1998q2	Tulsa, OK
1995q1	Carson City, NV	1998q2	Grand Junction, CO
1995q1	Mount Vernon-Anacortes, WA	1998q4	Richmond, VA
1995q1	Prescott, AZ	1998q4	Memphis, TN-MS-AR

## Appendix #2: Summary Statistics on Key Housing Characteristics

	MSAs	Neighborhoods	Neighborhoods, >10 transactions
	(1)	(2)	(3)
Sale Price	255,409 (94,028)	256,759 (160,139)	251,082 (152,062)
Number of Bedrooms	3.2 (0.2)	3.2 (0.4)	3.3 (0.4)
Number of Bathrooms	2.3 (0.3)	2.4 (7.8)	2.4 (8.5)
Square Footage	1,856 (144)	1,893 (3,839)	1,961 (4,691)
Age of House	30 (12)	29 (21)	26 (21)
Mean Number of Transactions	249,585 (333,188)	1,724 (2,001)	2,260 (2,810)

Notes: First column presents weighted averages and standard deviations (in parenthesis) for all MSAs in our final sample. Weights are based on number of transactions. Column 2 shows summary statistics by tract groups, while Column 3 presents descriptives for a subsample of tracts with more than 10 transactions in every half-year period.

### **Appendix #3: Hedonic Regression Specifications**

The hedonic regression in Equation (1) contains a number of categorical variables created to control for differences in housing quality. Separate vectors were created for the number of bedrooms (Bed), the number of bathrooms (Bath) and the age of the home (Age).

In the case of bedrooms, ten dichotomous dummies were used to control for the number of bedrooms ranging from less than 1 (which includes 0 and 0.5 bedrooms in the raw data) to a top code of 9 for homes with nine or more bedrooms. In this case, each dummy represented a unit increase in the number of bedrooms (e.g. there are dichotomous dummies created for homes with <1, 1, 2, 3, 4, 5, 6, 7, 8, and 9+ bedrooms).

In the case of bathrooms, we included controls for homes with fewer than 1 bathroom (again, 0 or more typically, 0.5 bathrooms), a top code for units with seven or more bathrooms, dummies for each half unit increase from 1 through 5, and then controls for each unit increase until seven. More specifically, the twelve categories were: <1, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 6, and 7+.

There are nine categories of the Age vector from Equation (1). They range from newly built homes with an age of zero to homes at least 40 years old. The specific age categories are as follows: 0, 1, 2-5, 6-9, 10-14, 15-19, 20-29, 30-39, and 40+ years old.

The other quality control in the hedonic estimation, the square footage of the living space in the home, is continuous in nature and was entered in quadratic form as noted in Equation (1).

**Appendix #4: Breakpoint Estimates Summary**  
MSAs

	estimated coefficient	t-stat	R <sup>2</sup>	number of quarters	number of MSAs
all MSAs	0.14	7.34	0.63	31	94
by year:					
1997	0.12	8.43	0.65	39	5
1998	0.12	9.10	0.66	38	8
1999	0.13	7.61	0.64	33	8
2000	0.13	7.84	0.63	36	4
2001	0.13	7.96	0.63	36	3
2002	0.17	8.21	0.64	36	11
2003	0.16	7.62	0.61	34	12
2004	0.19	9.25	0.68	35	20
2005	0.18	8.15	0.62	36	8
not stat. significant	-0.01	-1.74	0.27	14	6
not enough data	0.08	4.18	0.70	7	9

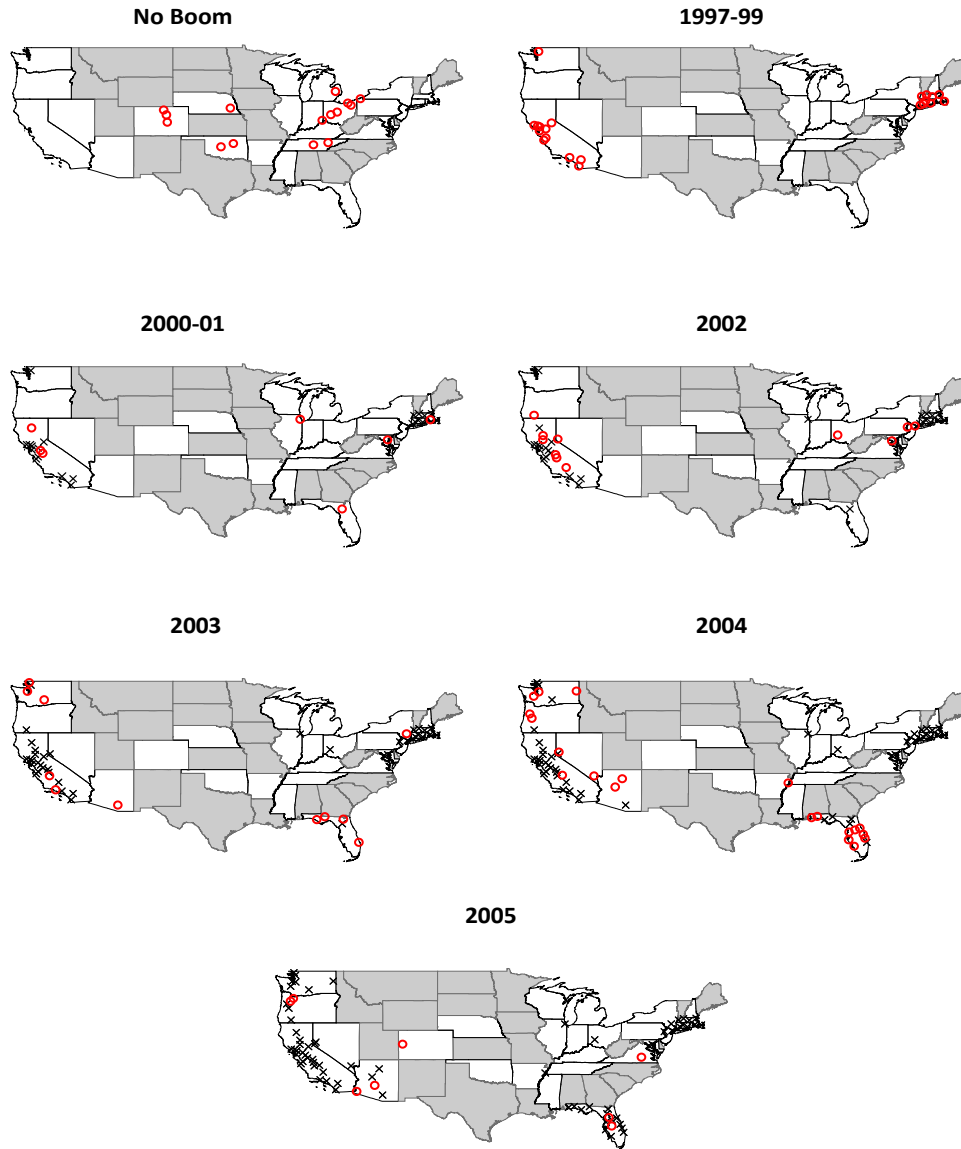
Neighborhoods

	estimated coefficient	t-stat	R <sup>2</sup>	number of half-years	number of neighborhoods
all tract groups	0.17	3.59	0.49	14	7335
by year:					
1994	0.29	2.91	0.36	16	19
1995	0.23	3.45	0.41	18	61
1996	0.18	3.66	0.44	18	116
1997	0.18	3.98	0.50	16	353
1998	0.18	4.50	0.53	17	656
1999	0.20	4.40	0.54	16	540
2000	0.21	4.56	0.56	16	495
2001	0.19	4.24	0.55	15	377
2002	0.20	4.18	0.52	16	463
2003	0.22	4.25	0.50	17	503
2004	0.24	4.67	0.52	18	837
2005	0.25	5.59	0.56	20	680
2006	0.21	2.91	0.32	20	29
2007	0.27	2.98	0.25	27	4
2008	0.12	2.75	0.29	20	4
not stat. significant	0.06	0.43	0.28	8	1728
not enough data	0.21	5.07	0.78	4	470

Notes: Both panels show summary stats of the break point estimation for MSAs and tract groups. The first column shows averages of the estimated coefficients  $d$ 's from equation (4), the second column show the average t-stat, the 3<sup>rd</sup> column shows average  $R^2$ 's, the fourth column shows the average number of periods used in the estimation, and the last column shows the total number of MSAs or tract groups.

## Appendix #5: Geographic Heterogeneity in the Starting Points of Housing Booms: MSAs

### Timing of Breakpoints by Metro Area



Note: Each red circle denotes a metropolitan area that is new to the time frame noted just above each map. Each black 'x' represents a metropolitan area from all previous maps. Shaded states are not represented in our sample. See the discussion in the text for more details.

The first map in Appendix #5 shows that the 15 MSAs that never had a meaningful boom in price growth are all located in the interior of the country. These markets are not shown in any subsequent map, each of which plots the geographic distribution and spreading of initial booms.

The first housing booms according to our metric occurred in the 3<sup>rd</sup> and 4<sup>th</sup> quarter of 1997 in two California markets (Los Angeles and Napa) and three New England regions (Springfield, MA, New Haven, CT, and Stamford, CT). The second map in the figure shows the location of all the markets that boomed between 1997 and 1999. This group includes other markets, both big and small, also in northern New England and coastal California, as well as the first market in Washington state. An interesting pattern emerges after that initial set of booms: from coastal California, booms proliferate in the west and north directions, while from northern New England new booms occur in the east and south directions.

For example, the third map in the figure adds in the seven metropolitan areas that first boomed at some point in 2000 or 2001. In addition to three smaller interior markets in California (Modesto, Merced, and Redding) and a couple of east coast markets (Providence-New Bedford-Fall River, RI-MA and Baltimore-Towson, MD), this time span sees the first major Midwestern market (Chicago-Naperville-Joliet, IL-IN-WI) and the first Florida market (Gainesville) experience their major jumps in price growth.

Calendar year 2002 sees the beginning of the bigger wave of housing booms. The fourth map in the figure shows this group of 11 to be a fairly disparate group. There are a number of smaller California markets that start booming (Yuba City-Marysville, Chico, Bakersfield, Madera, and Fresno), but we see other places in different western states boom, too. They include the first market in Nevada (Carson City), as well as one in Oregon (Medford). On the east coast, the major metropolitan areas of New York-Northern New Jersey-Long Island, NY-NJ, and Washington, DC, also experienced their global breakpoints, in addition to the smaller NJ-PA market of Allentown-Bethlehem-Easton.

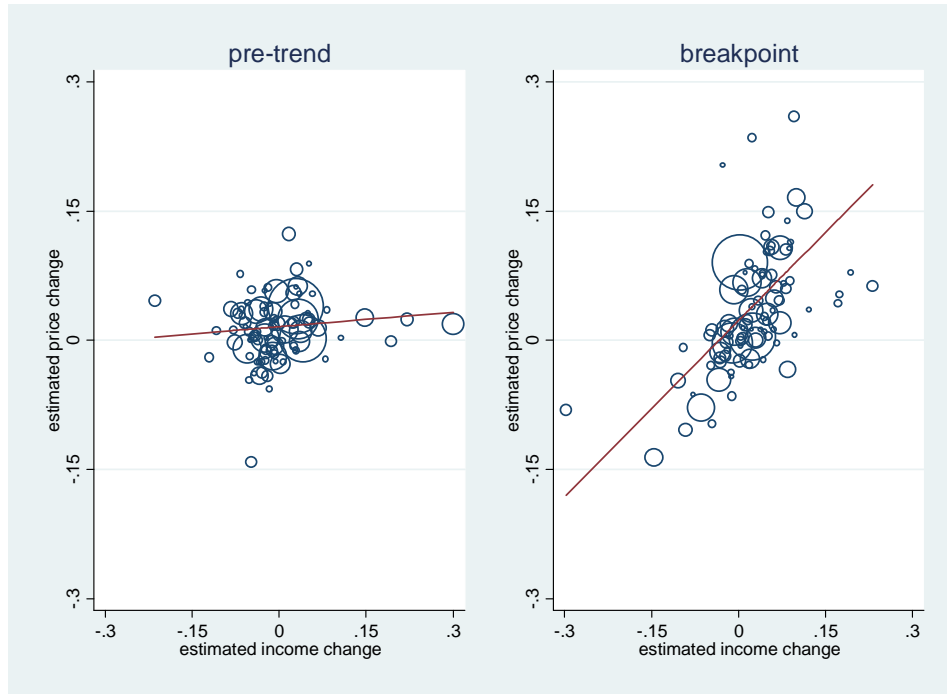
Calendar year 2003 then sees another twelve markets start to boom. Markets in the so-called 'sand states' are prevalent in this group. It includes three more Florida markets, along with the first Arizona metropolitan area (Tucson). Honolulu, three Washington state metros, and two more California markets also boom in 2003, as the geographic extent of the boom widens across the western states.

The largest number of metropolitan areas (20) boomed in 2004. This group also has a high concentration in the sand states. There are nine such metropolitan areas in Florida alone, including Orlando and Tampa. In Nevada, the Las Vegas-Paradise and Reno-Sparks metros experienced a boom. Other Arizona markets also begin experiencing a boom this year (Flagstaff and Prescott), although the Phoenix-Mesa-Scottsdale area does not do so until the beginning of 2005. This time period also sees a further widening of boom markets in the Pacific Northwest, including the large metros of Seattle-Tacoma-Bellevue, WA, and Portland-Vancouver-Beaverton, OR-WA.



## Appendix #6: Price and Income Correlations at Breakpoint

### a) Price and income OLS correlations, MSAs



Notes: We first estimate a regression of price on quarter and MSA fixed effects using the complete data set, and then use the residuals to measure the magnitude of the price changes around the breakpoint for each. A similar procedure is used for income. The figures above plot the MSA-level estimated changes, weighted by MSA population. The red line shows the estimate coefficient from OLS regressions that use the plotted data.

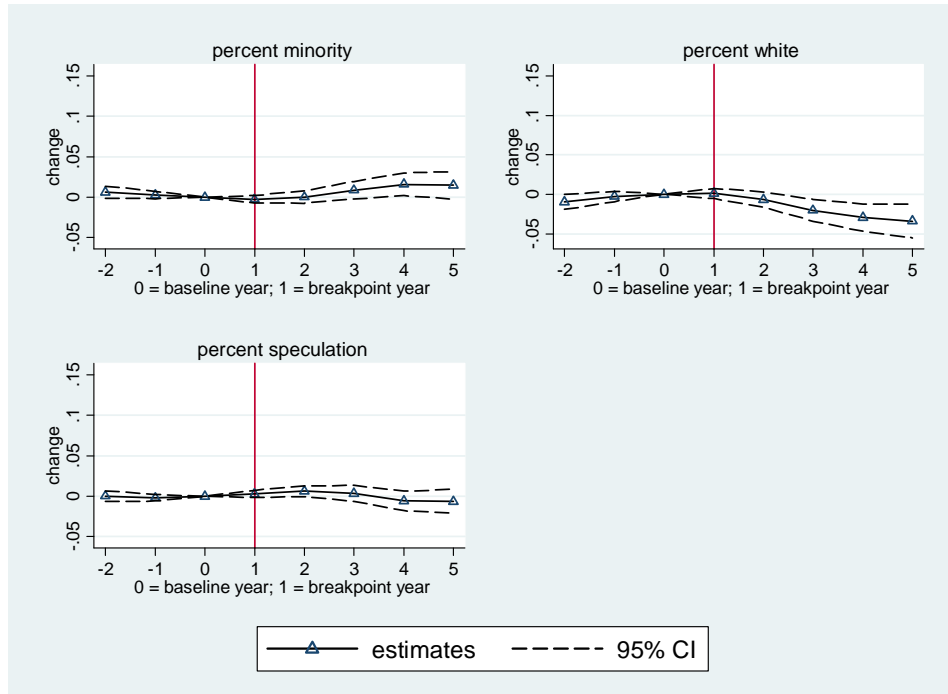
### b) Price and income OLS and IV correlations, MSAs

dependent variable: log price	pre-trend	breakpoint		
	OLS (1)	OLS (2)	OLS (3)	IV (4)
log income	0.06	<b>0.68</b>	<b>0.63</b>	<b>0.84</b>
	0.05	<b>(0.09)</b>	<b>(0.09)</b>	<b>(0.26)</b>
MSA and quarter effects	Y	Y	Y	Y
covariates	N	N	Y	Y
Observations	94	94	89	86
R-squared	0.02	0.37	0.48	0.46

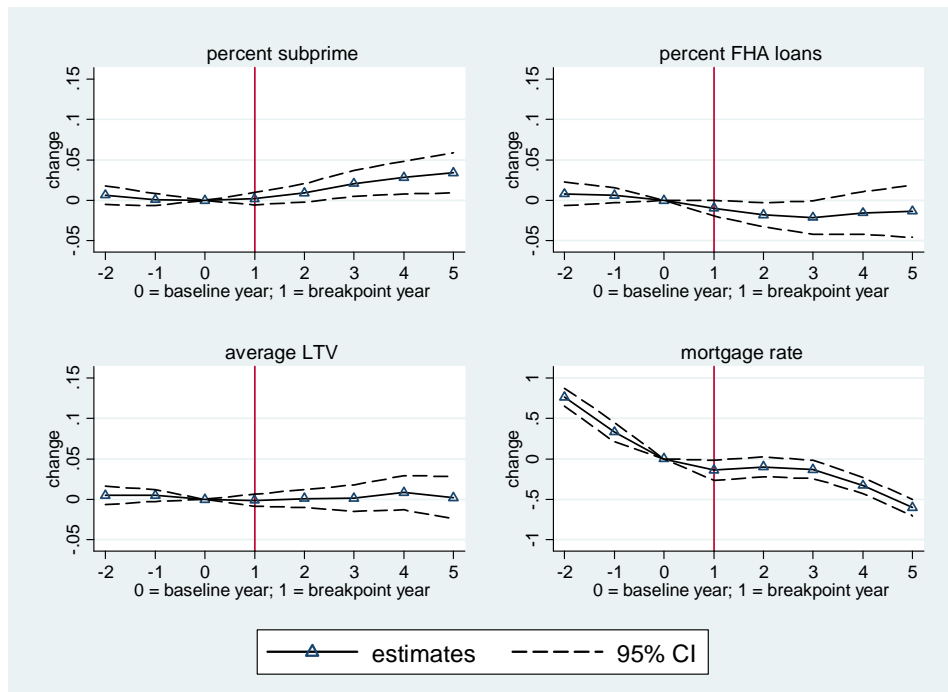
Notes: We first estimate a regression of price on quarter and MSA fixed effects using the complete data set and save the residuals. A similar procedure is used for income. The table above shows separate OLS and IV regressions of the residual price on the residual income, for the pre-trend and breakpoint periods. Covariates include percent minority, migration flow, percent speculators, average LTV, fraction subprime, and fraction FHA loans. Per capita income is used as instrumental variable for log homebuyer income in column IV. Bold coefficients are significant at 5% level.

## Appendix #7: Demand shifters and robustness tests around the breakpoint, MSA level

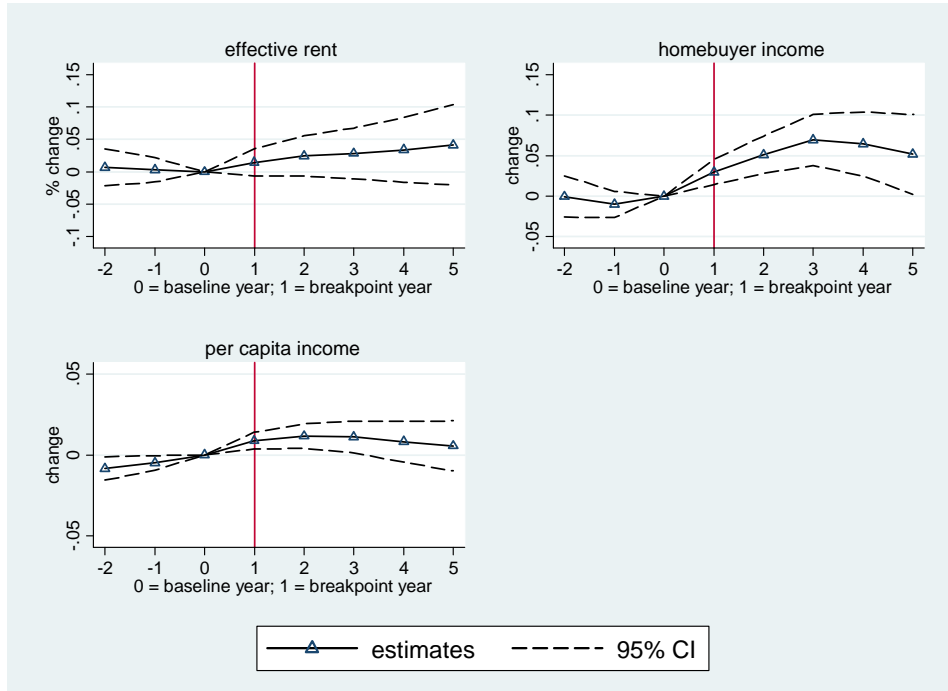
### a) Other buyer characteristics



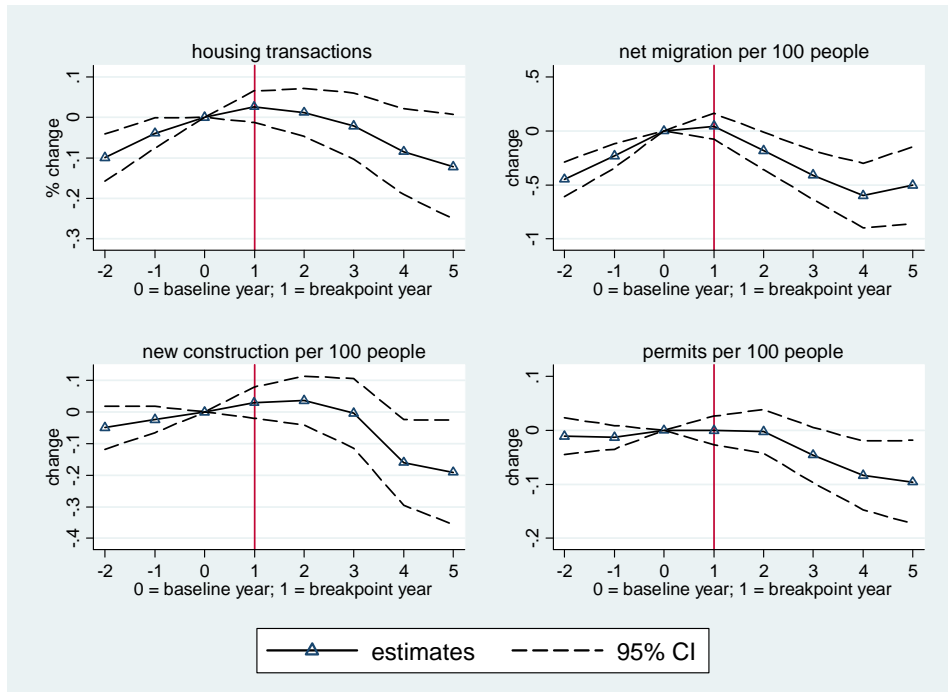
### b) Credit markets



### c) Rents and income



### d) Quantities, and new supply



e) Other robustness

