A Time-series Test to Identify Housing Bubbles

Diego Escobari

The University of Texas – Pan American
escobarida@utpa.edu

Damian Damianov

Durham University
damian.damianov@durham.ac.uk

Andres Bello

The University of Texas – Pan American
belloa@utpa.edu

SEA

November 2013

Outline

- Introduction
  - Motivation
  - Contribution
- Data
- Empirical Strategy
  - Identification of the Bubble
  - Testing Methodology
- Empirical Results
- Conclusion

Introduction

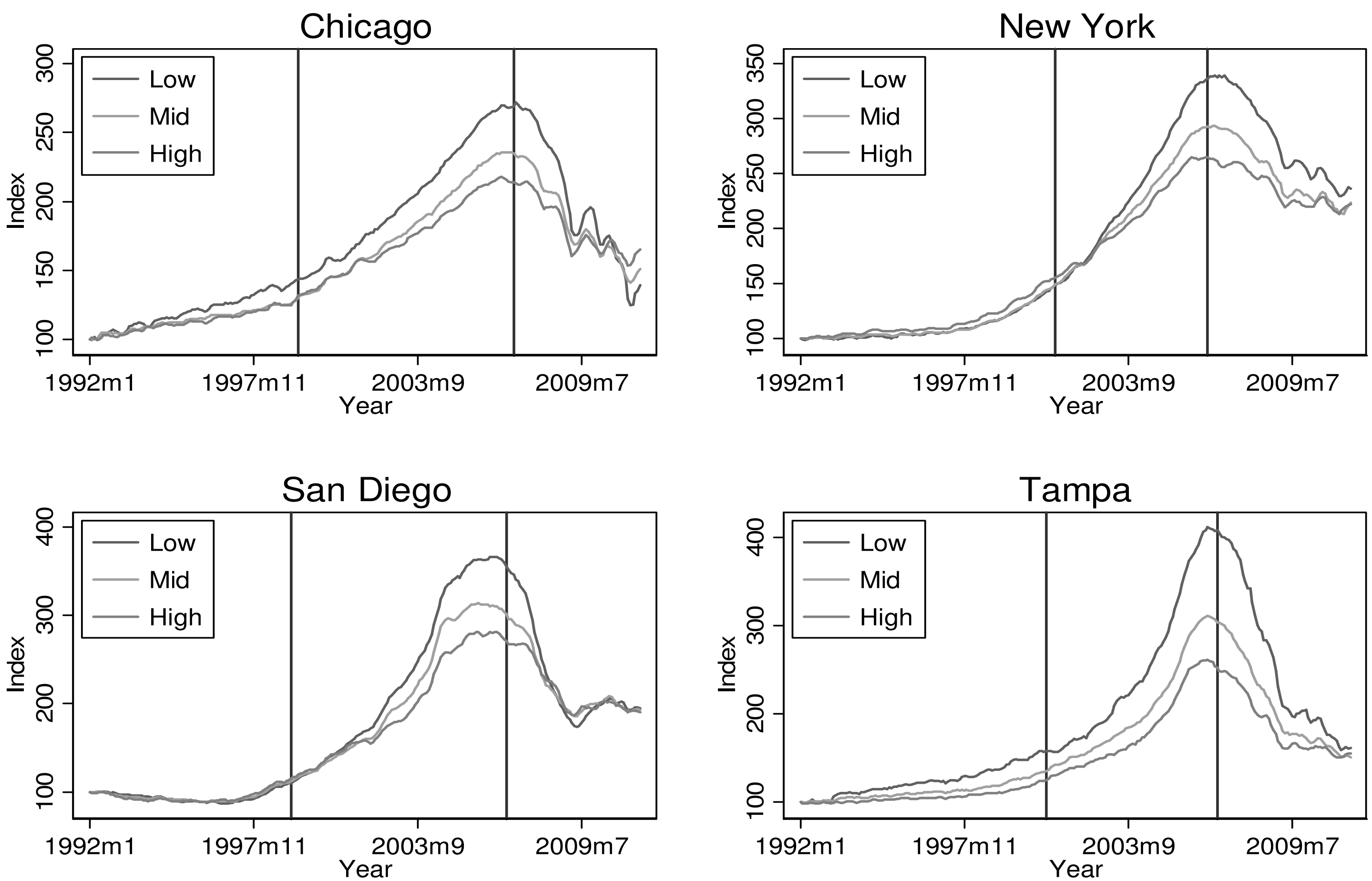
- Motivation
  - Bubbles: Unsustainable growth in asset prices that cannot be explained by "fundamental" factors.
  - Recent housing bubble was a key force behind the recent financial crisis.
  - Recent housing boom and bust was marked by large differences in appreciation/depreciation across price tiers.
- Contribution
  - New empirical time series test for the existence of housing bubbles.
  - The procedure endogenously determines the beginning and the end of the bubble.
  - Identifies the bubble without observing fundamentals.

Data

- Time series S&P Case-Shiller seasonally adjusted Tier Price Indices.
- Between January 1992 and August 2011 with 15 Metropolitan Statistical Areas (MSA):
  - Atlanta, Boston, Chicago, Denver, Los Angeles, Miami, New York, Minneapolis, Phoenix, Portland, San Diego, San Francisco, Seattle, Tampa, and Washington DC.
- For each MSA we have three indices, the Low-, Medium-, and High-Tier.
- The indices we employ are constructed using a three month moving average, where home sales pairs are aggregated in rolling three month periods.

Figure 1

Low, Mid and High Tiers Indexes, 1992-01 through 2011-08.



Identification Strategy

$$p_t = p_t^f + B_t$$

Difference between two price tiers:

$$y_t^{ij} \equiv p_{i,t} - p_{j,t} = \left(p_{i,t}^f + B_{i,t}\right) - \left(p_{j,t}^f + B_{j,t}\right)$$

Test if the difference is trend stationary with zero mean process:

$$\lim_{k\to\infty} E_t\left(p_{i,t+k} - p_{j,t+k} \middle| I_t\right) = \beta_0 + \beta_1 t$$

Combining the last two equations:

$$\lim_{k\to\infty} E_t\left(p_{i,t+k} - p_{j,t+k} \middle| I_t\right) = \lim_{k\to\infty} E_t\left(p_{i,t+k}^f - p_{j,t+k}^f \middle| I_t\right) + \lim_{k\to\infty} E_t\left(B_{i,t+k} - B_{j,t+k} \middle| I_t\right)$$

Testing Methodology

- Minimum LM unit root test proposed by Lee and Strazicich (2003).
- Data-generating process:

$$y_t = \delta' Z_t + e_t \qquad\qquad e_t = \beta e_{t-1} + \varepsilon_t$$

- Includes two changes in levels and trends:

$$Z_t = [1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]'$$

where $DT_{mt} = t - T_{Bm}$ for $t \geq T_{Bm} + 1$, $m = 1,2$

$DT_{mt} = 0$ otherwise.

Empirical Results

- With no breaks find very little evidence of stationarity and cannot identify the bubble.

Table 2

Differences across Tiers with No breaks. ADF and KPSS tests, 1992-01 through 2011-08.

	$y_t^{ML} \equiv p_{M,t} - p_{L,t}$		$y_t^{HL} \equiv p_{H,t} - p_{L,t}$		$y_t^{HM} \equiv p_{H,t} - p_{M,t}$	
	ADF	KPSS	ADF	KPSS	ADF	KPSS
Atlanta	-0.255	0.378[c]	-0.027	0.395[c]	-0.244	0.424[c]
Boston	-1.324	0.265[c]	-1.332	0.266[c]	-0.890	0.257[c]
Chicago	0.083	0.329[c]	-0.995	0.270[c]	-0.810	0.214[c]
Denver	-1.081	0.500[c]	-1.057	0.489[c]	-1.225	0.391[c]
Los Angeles	-2.353	0.171[a]	-2.598[a]	0.194[b]	-1.898	0.211[b]
Miami	-1.078	0.244[c]	-1.884	0.233[c]	-0.369	0.208[b]
New York	-2.093	0.231[c]	-1.930	0.240[c]	-1.965	0.241[c]
Minneapolis	-0.771	0.378[c]	-0.404	0.412[c]	0.042	0.456[c]
Phoenix	-1.736	0.316[c]	-1.592	0.315[c]	-1.204	0.272[c]
Portland	-0.421	0.466[c]	-0.463	0.411[c]	-0.843	0.215[b]
San Diego	-1.499	0.255[c]	-2.582[a]	0.247[c]	-1.813	0.214[b]
San Francisco	-1.381	0.268[c]	-1.792	0.268[c]	-1.368	0.254[c]
Seattle	0.361	0.392[c]	-0.452	0.262[c]	-1.225	0.157[a]
Tampa	-1.138	0.251[c]	-1.314	0.260[c]	-0.700	0.266[c]
Washington DC	-2.313	0.146[b]	-2.711[a]	0.165[b]	-1.989	0.179[b]

Table 3

Differences in Mid and Low ($y_t^{ML} \equiv p_{M,t} - p_{L,t}$) with Breaks, 1992-01 through 2011-08.

	$\hat{\phi}$	Test Statistic	$\hat{k}$	$\hat{T}_{B1}$	$\hat{T}_{B2}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Atlanta	-0.126	-3.037	12	2000-01	2009-02	0.42	0.90
Boston	-0.217	-4.637	12	1997-12	2005-04	0.31	0.70
Chicago	-0.286	-5.724[b]	7	1995-12	2008-04	0.21	0.86
Denver	-0.138	-4.592	12	1997-03	2003-05[d]	0.27	0.60
Los Angeles	-0.044	-4.237	12	1995-11	2004-02	0.20	0.64
Miami	-0.220	-6.254[b]	10	2006-03	2008-06	0.75	0.86
New York	-0.189	-4.977	9	2001-09	2005-10	0.51	0.72
Minneapolis	-0.165	-4.606	10	1999-10	2007-11	0.41	0.83
Phoenix	-0.204	-6.460[c]	8	2006-10	2008-03	0.78	0.85
Portland	-0.421	-6.636[c]	12	1998-06	2009-02	0.34	0.90
San Diego	-0.104	-5.360[a]	12	1998-03	2006-08	0.33	0.77
San Francisco	-0.080	-3.881	11	2004-03[d]	2009-01[d]	0.64	0.89
Seattle	-0.184	-4.540	11	1994-02	2008-04	0.11	0.86
Tampa	-0.308	-5.753[b]	7	2000-08	2006-05	0.45	0.75
Washington DC	-0.148	-5.057	10	2002-01	2007-10	0.53	0.83

Notes: $\hat{k}$ is the optimal lagged first-differenced terms, $\hat{T}_{Bm}$ for $m = 1, 2$ denotes the year and month of the estimated break points and $\hat{\lambda}_m = \hat{T}_{Bm}/T$ for $m = 1, 2$ denote the location of the breaks. [a], [b], and [c] denote significant at the 10%, 5%, and 1% levels, respectively. [d] denotes that the identified break point is not significant at the 10%.

Table 4

Differences in High and Low ($y_t^{HL} \equiv p_{H,t} - p_{L,t}$) with Breaks, 1992-01 through 2011-08.

	$\hat{\phi}$	Test Statistic	$\hat{k}$	$\hat{T}_{B1}$	$\hat{T}_{B2}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\%\Delta p_{H,t}$
Atlanta	-0.188	-4.049	12	2001-01	2009-02	0.47	0.90	
Boston	-0.122	-4.358	12	1997-12	2004-08[d]	0.31	0.66	96.6%
Chicago	-0.265	-4.600	11	1999-04	2006-09	0.38	0.77	62.5%
Denver	-0.146	-4.588	11	2002-09	2007-10	0.56	0.83	
Los Angeles	-0.032	-3.761	11	1995-11	2003-10	0.20	0.62	
Miami	-0.148	-4.717	11	2000-12	2007-09	0.47	0.82	79.4%
New York	-0.226	-5.155	9	2000-11	2006-02[d]	0.47	0.74	70.4%
Minneapolis	-0.142	-4.124	12	1999-10	2007-08	0.41	0.82	48.4%
Phoenix	-0.345	-7.211[c]	8	2007-04	2008-12[d]	0.80	0.89	
Portland	-0.284	-5.919[b]	12	1997-03[d]	2008-07	0.27	0.87	56.3%
San Diego	-0.063	-5.175	11	1999-01	2006-06	0.37	0.76	134.5%
San Francisco	-0.064	-3.096	10	2003-12	2008-08[d]	0.63	0.87	
Seattle	-0.132	-3.845	12	1997-06	2007-07[d]	0.29	0.81	119.2%
Tampa	-0.271	-5.516[a]	7	2000-07	2006-06	0.45	0.76	103.4%
Washington DC	-0.078	-3.640	10	2001-05	2007-07	0.49	0.81	61.8%

Notes: $\hat{k}$ is the optimal lagged first-differenced terms, $\hat{T}_{Bm}$ for $m = 1, 2$ denotes the year and month of the estimated break points and $\hat{\lambda}_m = \hat{T}_{Bm}/T$ for $m = 1, 2$ denote the location of the breaks. [a], [b], and [c] denote significant at the 10%, 5%, and 1% levels, respectively. [d] denotes that the identified break point is not significant at the 10%.

Empirical Results

- For example, from Table 4, in Chicago the boom of the housing bubble started in April 1999 and the bust was in September 2006.
- In Chicago the prices of the high tier homes went up by 62.5% between April, 1999 and September, 2006.
- Appreciation was most pronounced in San Diego with an increase of 134.5%, followed by Seattle and Tampa with 119.2% and 103.4%, respectively.
- The beginning of the bubbles that are statistically significant at a 10% level are all between June 1997 and May 2001, starting with Seattle and finishing with Washington DC.
- The statistically significant end-of-bubble dates are all between June, 2006 (San Diego and Tampa) and July 2008 (Portland).

Table 5

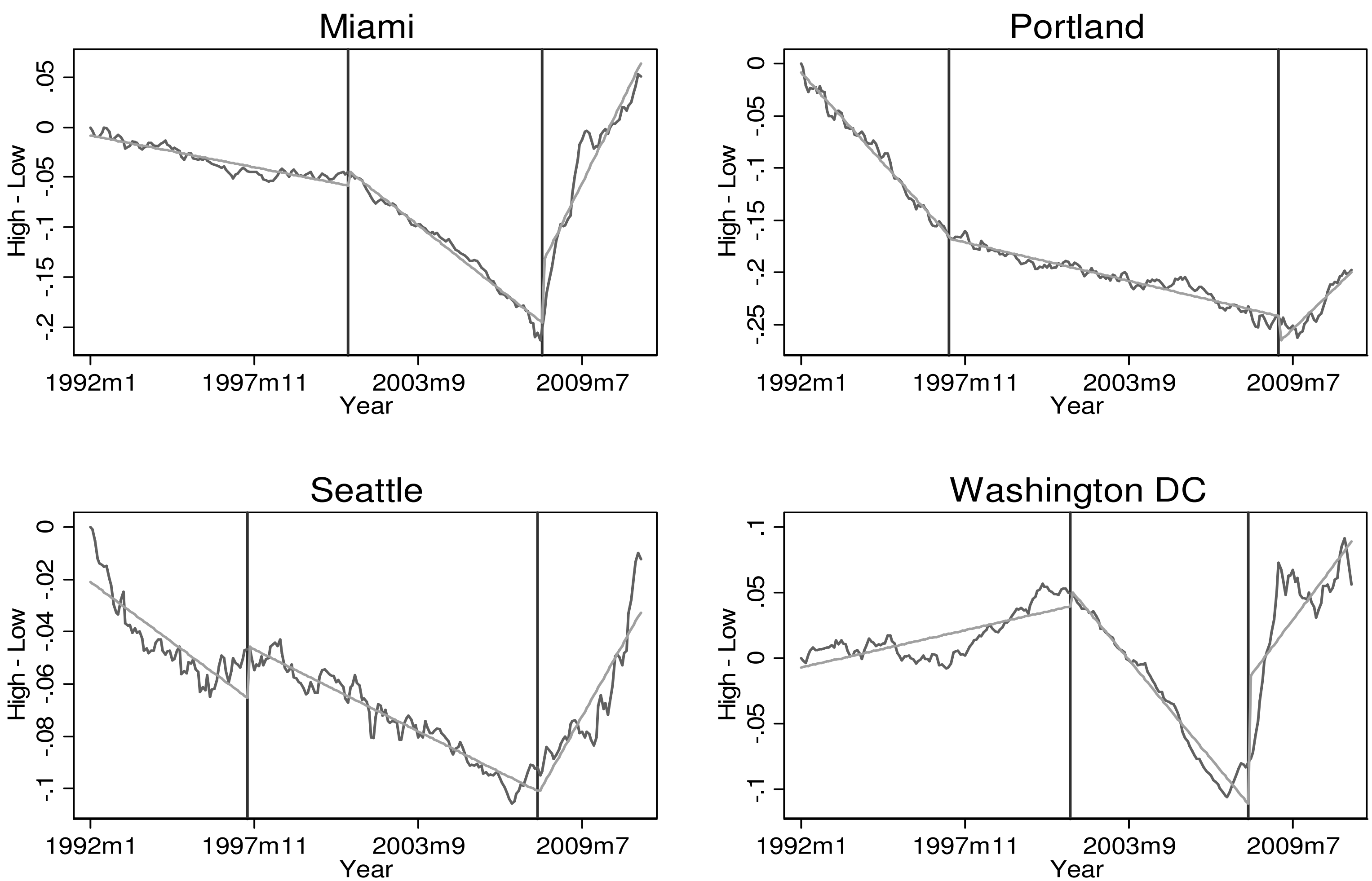
Differences in High and Mid ($y_t^{HM} \equiv p_{H,t} - p_{M,t}$) with Breaks, 1992m01 through 2011m08.

	$\hat{\phi}$	Test Statistic	$\hat{k}$	$\hat{T}_{B1}$	$\hat{T}_{B2}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Atlanta	-0.468	-7.105[c]	11	2001-06	2009-01	0.50	0.89
Boston	-0.118	-4.016	5	1998-10	2005-08	0.36	0.72
Chicago	-0.116	-3.253	12	2001-06	2008-07	0.50	0.87
Denver	-0.162	-4.516	5	2003-06	2008-03	0.60	0.85
Los Angeles	-0.078	-4.024	12	1998-05	2006-09	0.33	0.77
Miami	-0.140	-4.040	6	2000-10	2007-07	0.46	0.81
New York	-0.214	-5.258	12	2000-11	2006-12	0.47	0.78
Minneapolis	-0.256	-4.599	11	2000-11	2007-09	0.47	0.82
Phoenix	-0.270	-4.831	12	2004-09	2009-02	0.67	0.90
Portland	-0.240	-5.694[a]	12	1997-06	2008-11	0.29	0.89
San Diego	-0.094	-4.057	11	2001-01	2006-10	0.47	0.78
San Francisco	-0.151	-4.606	7	2000-09	2007-05	0.46	0.81
Seattle	-0.240	-4.991	12	1997-03	2008-06	0.27	0.86
Tampa	-0.304	-4.848	10	2001-12	2007-10	0.52	0.83
Washington DC	-0.056	-3.659	11	2002-09	2008-09[d]	0.56	0.88

Notes: $\hat{k}$ is the optimal lagged first-differenced terms, $\hat{T}_{Bm}$ for $m = 1, 2$ denotes the year and month of the estimated break points and $\hat{\lambda}_m = \hat{T}_{Bm}/T$ for $m = 1, 2$ denote the location of the breaks. [a], [b], and [c] denote significant at the 10%, 5%, and 1% levels, respectively. [d] denotes that the identified break point is not significant at the 10%.

Figure 2

Differences Between High and Low Tiers with Breaks, 1992-01 through 2011-08.



Conclusion

- The traditional approach to test for housing market bubbles needs market fundamentals.
- We exploit the property that low tier homes increase at a faster pace during the boom and depreciate more during the bust.
- Employ cointegration techniques that allow for structural breaks to estimate the dates of boom and bust.
- Misalignment in the appreciation rates of the home price tiered indices can be a symptom for a regime change in the borrowing and lending behavior.