Clustered housing cycles☆,☆☆

Rubén Hernández-Murillo[a], Michael T. Owyang[b,*], Margarita Rubio[c]

[a] External Outreach & Regional Analytics, Federal Reserve Bank of Cleveland, United States
[b] Research Division, Federal Reserve Bank of St. Louis, St. Louis, MO, United States
[c] Department of Economics, University of Nottingham, United Kingdom

ARTICLE INFO

ABSTRACT

Using a panel of U.S. city-level building permits data, we estimate a Markov-switching model of housing cycles that allows cities to systematically deviate from the national housing cycle. These deviations occur for clusters of cities that experience simultaneous housing contractions. We find that cities do not form housing regions in the traditional geographic sense. Instead, similarities in factors affecting the demand for housing (such as population growth or availability of credit) appear to be more important determinants of cyclical co-movements than similarities in factors affecting the supply for land (such as the availability of developable land or the elasticity of land supply).

1. Introduction

Recent macroeconomic research has argued that housing market movements are the source of—rather than the consequence of—business cycle fluctuations.[1] For example, housing was the key instigating component of the recent financial crisis (see Bernanke (2008)) and Leamer (2007) argues that, at a national level, housing emphasis the business cycle.[2] While business cycles are typically measured at a national level, housing markets are generally believed to be highly localized.[3] At the subnational level, the relationship between housing and the business cycle is less clear. Indeed, Ghent and Owyang (2010) find that, while housing cycles exist at disaggregate

levels, the relationship between business cycles and housing appears to break down at subnational levels.[4] The disconnect generated between the two levels of granularity can present problems for macroeconomic analysis and policy making.[5]

What produces the disconnect between housing and business cycles at different levels of disaggregation? One hypothesis is that a national cycle exists across all housing markets, but this more pervasive cycle is lost in the heterogeneity once the data is disaggregated. In other words, housing cycles may have both a national and regional element (Del Negro and Otrok, 2007). Local deviations from the national cycle can be small timing differentials (i.e., cities are just out-of-sync enough that the average cycle does not match the national cycle) or major

[1] Housing wealth effects lead to a correlation between housing and consumption expenditures and are typically accompanied by changes in housing investment in the same direction. For example, Davis and Heathcote (2005) find that residential fixed investment leads non-housing investment and is more than twice as volatile.
[2] These empirical regularities have prompted macroeconomic researchers to consider the theoretical underpinnings of housing and the business cycle in general equilibrium models (see Iacoviello (2005), Iacoviello and Neri (2010)).
[3] A few papers have modeled subnational business cycles (e.g., Carlino and Sill, 2001 for regions, Owyang et al. (2005) for states, and Owyang et al. (2008) for cities).
[4] Several papers study housing at the subnational level with mixed results. Glaeser et al. (2011), Del Negro and Otrok (2007) show that the variation in house prices is primarily driven by subnational factors rather than national factors, while Moench and Ng (2011) find that national shocks have larger effects on the housing cycle than regional shocks. Stock and Watson (2010) show that although building permits co-move across states, housing markets can be uncorrelated across regions. Other studies find that housing markets have become more integrated in the last decade (see Cotter et al. (2012), Kallberg et al. (2014), Landier et al. (2013)).
[5] Ghent and Owyang (2010) imply implies that models need to account for regional differences in U.S. housing cycles, as regional differences may influence optimal policy conduct. Identifying regions with similar housing characteristics would help policymakers more accurately predict cross-regional effects, and identify regions to target with particular policies.

departures. In the former case, a pervasive national cycle could be detected once we account for the deviations. In the latter case, city-level cycles could be viewed as completely independent of the national cycle.

While much of the recent work has focused on house price dynamics, we are interested in housing as it relates to the business cycle rather than house price dynamics alone. Saiz (2010) argues that housing volumes may be a better indicator of business cycle dynamics than house prices. Indeed, comparing the data on housing permits and prices reveals important differences in how the two series are connected to the business cycle. Price growth slows during many of the previous downturns but only becomes negative during the 2007–2009 recession. Permits, on the other hand, exhibit clear negative growth rates in all NBER recessions, and therefore represent a more reliable indicator of business cycles. Thus, in this paper, we consider city-level construction permits to analyze housing cycles at the regional level.[6]

We adapt the model from Hamilton and Owyang (2012) (henceforth HO), who identify national and regional business cycle phases. The approach in HO is to allow the data to group states into regions that have the same business cycles. In our model, during the *national* phases, all of the cities will be in expansion or contraction, respectively. Furthermore, cities in the same region (or *cluster*) experience simultaneous idiosyncratic contractions, which could lead, lag, or occur separately from national contractions.

An advantage of the HO model is that it determines the key factors in the grouping of cities—i.e., it generates endogenous clusters or regions. Cities may cluster for economic (e.g., similar industrial composition), geographic (e.g., inability to increase the housing stock due to geographic constraints), or other inherent reasons (e.g., weather). We consider nine covariates in determining our clusters: housing density, population growth, the share of manufacturing employment, average winter temperature, the average unemployment rate, Saiz's (2010) index of undevelopable land and elasticity of land supply, as well as two financial variables reflecting the incidence of subprime mortgages across cities.

We find evidence of a national housing cycle that appears to be linked to the national business cycle. We also find 3 clusters of cities that experience their own idiosyncratic contractions. These contractions can occur before and lead into a national downturn, continue after a national downturn, or occur completely independently of a national downturn. In addition, our method allows us to determine some of the factors that affect cluster composition. These factors appear to be proxies of city-level housing demand characteristics rather than housing supply characteristics, factors that influence business cycle, or factors related to the similarity of financial conditions.[7]

Our finding that housing cycles may depend on local factors in addition to national factors has implications for policy implementation, as monetary policy may be transmitted in different ways throughout the set of clusters. In fact, Füss et al. (2012) find that specific demand and supply characteristics at the level of a Metropolitan Statistical Area (henceforth MSA), such as population growth and the elasticity of housing supply, are crucial links that transmit national monetary policy and sentiment into housing price inflation at the local level. Therefore, identifying regional clusters of housing cycles represents relevant information for the optimal conduct of monetary policy. We also find a cluster in which financial variables are more important, which

suggests that regional differences may influence the conduct of macroprudential policies as well. Our results open up a new line of research to explore whether fiscal policy or macroprudential policies should consider regional differences in attempting to balance out housing cycles across regions and evaluating the heterogeneity in the monetary transmission mechanism.

The rest of the paper is outlined as follows: Section 2 describes the data construction. Section 3 describes the model. Section 4 provides a brief overview of the estimation of the model. Section 5 presents the results and discusses the implications for the determination of the national cycle. Section 6 offers some conclusions.

2. The data

In the next subsections, we describe the collection and transformation of the data used to construct housing cycles and city-level covariates used to form the clusters.

2.1. City-level building permits

While city-level building permit data are available for various time periods for the majority of cities, MSA definitions have changed a number of times over the years. These definitional changes can present a problem for business cycle analysis, which requires longer time series to detect switching between phases. Using current MSA-level data for our analysis would limit the number of cities with available data and the length of the time series which only go back to the mid-1990s. Therefore, we construct MSA-level permits by aggregating county-level permit data using the counties included in the December 2009 MSA definitions.[8]

County-level building permits are released monthly as year-to-date levels that accumulate the monthly changes. Although differencing would yield a monthly flow of permits, the year-to-date data are constantly revised, confounding identification of an actual change versus a revision. In addition, the noise in the monthly data affects our filter's ability to detect switches in cycle phase. To mitigate these two problems, we use the band-pass-filtered, year-over-year growth rate of the (final release) monthly building permits series. This transformation also has the benefit of smoothing outliers that would not be considered business cycle phase changes but could be misidentified by our estimation algorithm.

Our sample consists of data from 1989:01 to 2012:11. We select 135 cities with population greater than 250,000 residents (based on 1990 populations computed with the 2009 MSA definition) for which all covariates and permits data are available. We limit our sample size to large cities because the data for the small cities are inherently noisier, and clustering in larger panels is more computationally-intensive and more likely to be imprecisely estimated.

Fig. 1 depicts the building permits growth series for a few of the cities in our sample.[9]

The national data are included for comparison; shaded areas represent NBER recession dates. The national data have a clear cyclical pattern that is roughly coincident with the timing of the NBER contractions. The U.S. permits data experience some fluctuations apart from the business cycle dates but the largest downturns begin just before an NBER-defined peak.[10] City-level experiences vary widely: Chicago, for example, behaves

---

[6] Alternative results using housing prices yield qualitatively similar conclusions. These results are available upon request.

[7] Traditionally, the main driver of house price co-movement is geography. Pollakowsky and Ray (1997), Can (1990), Ioannides and Zabel (2003) find positive feedback effects on house prices between contiguous regions and within neighborhoods. Stevenson (2004), Oikarinen (2006) show that co-movements or regional housing markets are driven by substitution effects, where changes in one region's prices generate a time-lag movement in another region's prices. Brady (2011), Holly et al. (2011) also explore the spatial and temporal diffusion of house price shocks in a dynamic system.

[8] December 2009 MSA definitions are available from the Census Bureau website http://www.census.gov/population/metro/files/lists/2009/List4.txt.

[9] The growth rates illustrated in Fig. 1 as well as the summary statistics reported in Table 1 refer to the rates of growth in the monthly permit series after the application of the band-pass filter. The application of the band-pass filter does not necessarily smooth the growth rates but reduces outliers in the growth series when there are jumps in the level series.

[10] As evidenced by the 2001 recession, permit growth does not always fall with the business cycle. Of course, we have only three national recession experiences in our sample so the results should be extrapolated with caution.
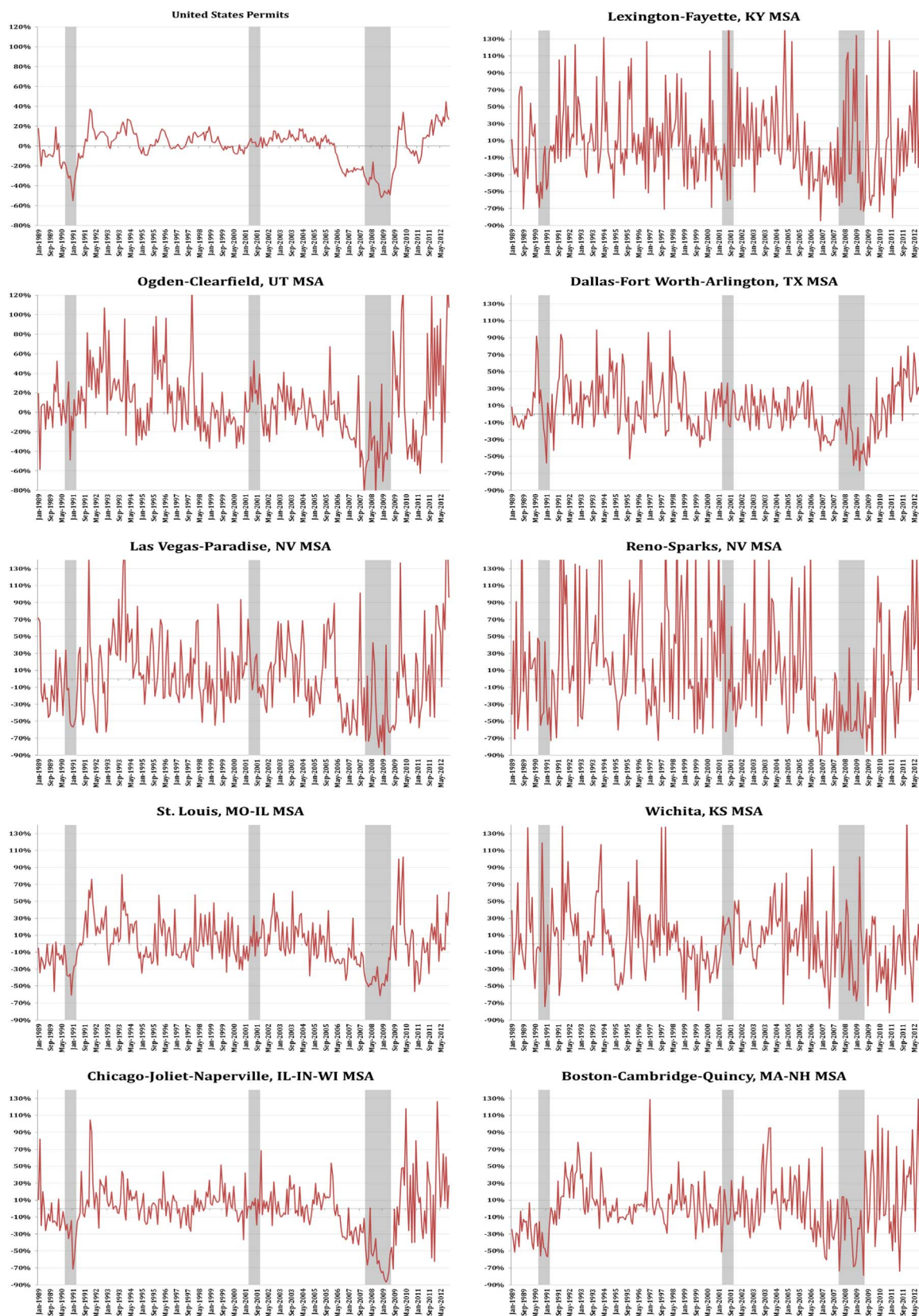
Fig. 1. Building Permits. A Few Cities and the Nation.

very similarly to the nation, while Los Angeles has the same broad features as the nation but has more small, non-national housing contractions. Reno-Sparks, NV experiences even more small, non-national housing contractions than LA, such that the series almost appears to

exhibit seasonality. Table 1 contains the full list of cities in our sample and summary statistics of the permits growth series. The table provides the Core Based Statistical Area (CBSA) code, the metropolitan area name, and the mean and standard deviation of the growth rates in permits.

Table 1
List of cities and permits growth rate statistics.

CBSA	MSA	Permits Growth Rate (%)	
		Mean	SD
10420	Akron, OH	−3.5	35.0
10580	Albany-Schenectady-Troy, NY	1.7	45.0
10740	Albuquerque, NM	4.9	42.9
10900	Allentown-Bethlehem-Easton, PA-NJ	1.1	36.9
11460	Ann Arbor, MI	0.7	75.6
11700	Asheville, NC	3.6	41.7
12060	Atlanta-Sandy Springs-Marietta, GA	2.4	35.7
12260	Augusta-Richmond County, GA-SC	5.2	40.6
12420	Austin-Round Rock-San Marcos, TX	23.7	67.3
12540	Bakersfield-Delano, CA	3.3	48.0
12580	Baltimore-Towson, MD	0.9	38.4
12940	Baton Rouge, LA	15.2	65.0
13140	Beaumont-Port Arthur, TX	27.2	88.9
13820	Birmingham-Hoover, AL	8.4	46.8
14260	Boise City-Nampa, ID	12.8	47.4
14460	Boston-Cambridge-Quincy, MA-NH	2.8	34.5
14860	Bridgeport-Stamford-Norwalk, CT	7.5	55.7
15180	Brownsville-Harlingen, TX	10.6	42.8
15940	Canton-Massillon, OH	2.5	41.5
15980	Cape Coral-Fort Myers, FL	7.0	47.6
16620	Charleston, WV	11.6	66.9
16700	Charleston-North Charleston-Summerville, SC	6.0	39.2
16740	Charlotte-Gastonia-Rock Hill, NC-SC	5.1	39.2
16860	Chattanooga, TN-GA	1.2	31.2
16980	Chicago-Joliet-Naperville, IL-IN-WI	−0.4	31.3
17140	Cincinnati-Middletown, OH-KY-IN	−1.9	27.1
17460	Cleveland-Elyria-Mentor, OH	−1.4	27.4
17820	Colorado Springs, CO	13.1	52.4
17900	Columbia, SC	6.3	38.8
17980	Columbus, GA-AL	3.9	51.6
18580	Corpus Christi, TX	14.0	62.7
19100	Dallas-Fort Worth-Arlington, TX	7.3	31.2
19340	Davenport-Moline-Rock Island, IA-IL	11.5	58.9
19380	Dayton, OH	5.3	56.0
19660	Deltona-Daytona Beach-Ormond Beach, FL	−0.7	40.0
19740	Denver-Aurora-Broomfield, CO	11.6	46.6
19780	Des Moines-West Des Moines, IA	15.7	63.1
19820	Detroit-Warren-Livonia, MI	1.7	38.6
20260	Duluth, MN-WI	8.1	53.3
20500	Durham-Chapel Hill, NC	12.1	60.5
21340	El Paso, TX	12.8	61.0
21500	Erie, PA	17.6	78.8
21660	Eugene-Springfield, OR	12.4	65.0
21780	Evansville, IN-KY	4.7	46.2
22180	Fayetteville, NC	18.3	68.5
23060	Fort Wayne, IN	1.9	37.6
23420	Fresno, CA	4.6	46.3
24340	Grand Rapids-Wyoming, MI	1.6	36.5
24660	Greensboro-High Point, NC	5.5	45.8
24860	Greenville-Mauldin-Easley, SC	5.9	37.8
25420	Harrisburg-Carlisle, PA	2.9	40.8
25540	Hartford-West Hartford-East Hartford, CT	−0.6	39.4
25860	Hickory-Lenoir-Morganton, NC	−0.5	38.3
26420	Houston-Sugar Land-Baytown, TX	12.6	36.9
26620	Huntsville, AL	6.7	52.4
26900	Indianapolis-Carmel, IN	2.4	31.1
27140	Jackson, MS	7.8	51.9
27260	Jacksonville, FL	5.1	42.1
28020	Kalamazoo-Portage, MI	9.3	62.0
28140	Kansas City, MO-KS	6.0	50.8
28660	Killeen-Temple-Fort Hood, TX	26.8	99.9
28700	Kingsport-Bristol-Bristol, TN-VA	4.4	41.7
28940	Knoxville, TN	−26.8	276.0
29460	Lakeland-Winter Haven, FL	7.5	51.7
29540	Lancaster, PA	4.6	48.8
29620	Lansing-East Lansing, MI	3.1	53.8
29820	Las Vegas-Paradise, NV	4.1	45.9
30460	Lexington-Fayette, KY	6.7	47.4
30780	Little Rock-North Little Rock-Conway, AR	13.3	56.2
31140	Louisville/Jefferson County, KY-IN	4.7	39.0
31540	Madison, WI	6.0	44.1

Table 1 (continued)

CBSA	MSA	Permits Growth Rate (%)	
		Mean	SD
31700	Manchester-Nashua, NH	0.1	44.5
32580	McAllen-Edinburg-Mission, TX	11.7	43.9
32820	Memphis, TN-MS-AR	7.8	55.0
33100	Miami-Fort Lauderdale-Pompano Beach, FL	1.1	40.5
33340	Milwaukee-Waukesha-West Allis, WI	−1.2	35.7
33460	Minneapolis-St. Paul-Bloomington, MN-WI	2.2	36.3
33660	Mobile, AL	15.4	65.6
33700	Modesto, CA	1.5	60.7
33860	Montgomery, AL	7.5	65.0
34980	Nashville-Davidson−Murfreesboro−Franklin, TN	6.7	42.6
35300	New Haven-Milford, CT	−1.9	44.6
35380	New Orleans-Metairie-Kenner, LA	22.1	97.4
35620	New York-Northern New Jersey-Long Island, NY-NJ-PA	3.4	33.2
35980	Norwich-New London, CT	0.3	49.1
36260	Ogden-Clearfield, UT	6.5	37.1
36420	Oklahoma City, OK	9.9	42.2
36540	Omaha-Council Bluffs, NE-IA	7.5	43.7
36740	Orlando-Kissimmee-Sanford, FL	5.4	43.6
37100	Oxnard-Thousand Oaks-Ventura, CA	7.5	69.6
37340	Palm Bay-Melbourne-Titusville, FL	0.0	39.4
37860	Pensacola-Ferry Pass-Brent, FL	6.7	49.8
37900	Peoria, IL	9.6	58.2
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	−1.1	28.3
38060	Phoenix-Mesa-Glendale, AZ	4.8	37.4
38300	Pittsburgh, PA	1.2	28.9
38860	Portland-South Portland-Biddeford, ME	1.0	34.6
38940	Port St. Lucie, FL	1.1	52.4
39300	Providence-New Bedford-Fall River, RI-MA	−4.5	29.1
39340	Provo-Orem, UT	17.9	53.4
39740	Reading, PA	1.2	50.4
39900	Reno-Sparks, NV	11.1	66.0
40060	Richmond, VA	−0.4	33.7
40140	Riverside-San Bernardino-Ontario, CA	−1.3	37.6
40380	Rochester, NY	−2.3	35.5
40420	Rockford, IL	1.5	49.3
41180	St. Louis, MO-IL	−0.6	27.7
41420	Salem, OR	10.2	71.2
41500	Salinas, CA	13.2	90.5
41700	San Antonio-New Braunfels, TX	12.9	46.9
41740	San Diego-Carlsbad-San Marcos, CA	7.2	59.2
41860	San Francisco-Oakland-Fremont, CA	7.3	50.9
41940	San Jose-Sunnyvale-Santa Clara, CA	39.2	156.2
42220	Santa Rosa-Petaluma, CA	3.9	63.5
42340	Savannah, GA	8.3	51.9
42540	Scranton−Wilkes-Barre, PA	0.9	43.1
42660	Seattle-Tacoma-Bellevue, WA	3.9	36.7
43780	South Bend-Mishawaka, IN-MI	6.8	63.4
44060	Spokane, WA	13.2	71.1
44140	Springfield, MA	−3.3	37.8
44180	Springfield, MO	7.9	47.8
44700	Stockton, CA	2.8	48.1
45060	Syracuse, NY	4.5	55.5
45300	Tampa-St. Petersburg-Clearwater, FL	4.8	42.1
45780	Toledo, OH	4.1	52.5
46060	Tucson, AZ	8.2	53.3
46140	Tulsa, OK	8.8	41.5
46700	Vallejo-Fairfield, CA	18.3	93.1
47260	Virginia Beach-Norfolk-Newport News, VA-NC	−4.1	55.3
47300	Visalia-Porterville, CA	1.9	38.2
47900	Washington-Arlington-Alexandria, DC-VA-MD-WV	2.9	34.9
48620	Wichita, KS	3.7	41.3
49340	Worcester, MA	3.4	44.9
49620	York-Hanover, PA	−0.1	39.7
49660	Youngstown-Warren-Boardman, OH-PA	0.2	42.3

2.2. Covariate data

In addition to the panel of permits data, we require a set of time-invariant covariates to parameterize the prior distribution that defines the clusters. Our choice of covariates is meant to represent factors that affect housing cycles: demand, supply, geography, economic, and financing conditions. Meen (1999) suggests that co-movements may be caused by factors that affect the demand or the supply of housing. In particular, Meen (1999) finds that co-movements are mainly caused by migration, equity transfers, spatial arbitrage, and local economic development.

One limitation of the algorithm described below is that the covariate data used to populate the prior must be time invariant as the model assumes time-invariant clusters. Thus, the cross-sectional covariates are all computed either as long-run averages or by taking a snapshot at some point in the sample.

We selected a small set of covariates and calculated long run averages, values at the beginning of our sample, or averages during the sample period, depending on data availability and relevance for our analysis. For example, we obtained long run historical averages of winter temperatures. We also calculated housing density and the share of manufacturing employment as of 1990. Population growth represents the historical annual average in the two decades prior to the beginning of our sample.[11] To capture economic conditions during the period of analysis, we computed the average unemployment rate from 1988 to 2012.[12] Similarly, we calculated financial variables representing the increase in availability of credit for the period of subprime credit expansion, from 2002 to 2005. Finally, we obtained variables representing housing supply conditions from Saiz (2010). Table 2 presents summary statistics of the cross-section of these covariates.

Average population growth over the period 1970 to 1990 proxies the average change in housing demand, as higher average population growth suggests a higher average demand for new housing. In addition, the city-level average unemployment rate represents long-run differences in local economic conditions that may also affect the demand for housing.[13] Average winter temperature may be another proxy for housing demand, especially for a higher income demographic. Housing density is a proxy for the supply of housing, and Saiz's (2010) index of undevelopable land reflects geographic constraints on the elasticity in the supply of housing.[14] We include the share of manufacturing employment because prior studies have shown that it is a determinant of business cycle similarity and may also influence

Table 2
Summary statistics of covariates.

Covariate	Mean	SD	Median	Min	Max
Housing units per sq. km. 1990	56.5	53.2	39.7	4.1	380.7
Avg. population growth (%) 1970–1990	1.5	1.3	1.2	−0.6	6.0
Manufacturing employment share (%) 1990	21.3	8.5	20.0	3.2	58.5
Average winter temperature (° Celsius)	4.3	6.9	3.1	−9.6	19.6
Average unemployment rate (%) 1988–2012	6.0	1.8	5.6	3.3	14.2
Saiz's (2010) index of undevelopable land (%)	27.1	21.0	21.9	0.9	79.6
Saiz's (2010) elasticity of land supply	2.1	1.0	1.8	0.7	5.5
Change in share of subprime mortgages 2002–2005	−1.5	8.3	−1.9	−36.8	44.1
Growth in subprime mortgages 2002–2005	11.5	12.0	11.6	−20.3	50.4

Notes:
MSA population, housing units, and land area were aggregated from county-level data from the Census Bureau to match 2009 MSA definitions available at http://www.census.gov/population/metro/files/lists/2009/List4.txt (Retrieved on 11 April 2013).
The manufacturing employment share was computed as the ratio of MSA manufacturing employment to total employment, aggregating county-level data from the 1990 County Business Patterns from the Census Bureau.
Average winter temperatures represent long-run typical temperatures obtained for each city from the Department of Energy.
Unemployment rates were computed aggregating the number of the unemployed and the labor force at the county level with data from the Bureau of Labor Statistics, and subsequently were averaged over the period 1988:1–2012:12.
Saiz's (2010) index represents the percent of area in each city that cannot be developed because of geographic constraints, and was graciously provided by the author.
The change in the proportion of subprime mortgage loans and the growth in mortgage loan balances were obtained from credit bureau data at the zip code level aggregated at the metropolitan area.

housing cycle similarity. Finally, the two measures of credit availability are the change in the proportion of subprime mortgage loans relative to total mortgage loan volume and the growth in the loan volume of subprime mortgages over the period 2002 to 2005.

3. The empirical model

The model is a first-order Markov-switching model in the mean growth rate of each city's building permits series. We allow for two regimes at the city level: an expansion regime with a higher average growth rate and a contraction regime with a lower average growth rate. In the most general framework, each city building permit series could have an independent, unobserved 2-state Markov-switching process. In the most degenerate case, each city would have the same business cycle, a national cycle. Each of these models yields a regime process that can be summarized in a single national Markov state variable. In the former case, the Markov variable has 2^N possible regimes; in the latter case, the aggregate variable is a 2-state variable.

We are interested in an intermediate model which simultaneously limits the possible regimes to a tractable number, estimates a national regime, and allows some heterogeneity across cities. This framework can be obtained by assuming that a national regime—subject to some restrictions—exists, but that departures from this national regime (i.e., idiosyncratic contractions) must be relatively pervasive (i.e., experienced by a group of cities). Thus, our model has both national cycles and periods during which groups of cities experience contractions by

---

[11] The average population growth is taken from the Census and represents the percent increase in the population of the counties in the 2009 definition of the MSA between 1970 and 1990. The share of manufacturing employment is computed from Census county-level data from County Business Patterns. Average winter temperature represents long-run typical temperatures obtained for each city. The unemployment rates were computed aggregating the number of the unemployed and the labor force at the county level with data from the Bureau of Labor Statistics, and subsequently averaging over the period 1988:1–2012:12. Housing density is total housing units in 1990 divided by the land area of the MSA. Loan volumes are aggregated from ZIP code level data on loan volumes from the Federal Reserve Bank of New York Equifax Consumer Credit Panel Data. *Subprime* loans are defined as those granted to individuals with an Equifax Risk Score of 660 or less. (This score is a internal generic risk score that ranges between 280 and 850.) Saiz's (2010) index of undevelopable land represents the proportion of area in each city that cannot be developed because of geographic constraints.

[12] Taking averages of the unemployment rate captures persistent differences in economic conditions across cities at the cost of ignoring trends during the period of analysis.

[13] Clayton et al. (2010) find evidence that housing demand, in particular, is the main determinant of housing cycles. They find that both house prices and trading volumes are significantly affected by changes in the labor market, which include changes in total non-agricultural employment, average household income, and the unemployment rate.

[14] Expansion of the population away from central urban areas into rural and remote areas (i.e., urban sprawl) is another factor identified in the literature as affecting house price co-movement. Rising incomes, growing population, and low commuting costs boost demand for space in distant locations where land is relatively cheap, causing urban expansion. See, for instance Brueckner (2000), Couch and Karecha (2006) and O'Sullivan (2009).

themselves. Let κ represent the number of groups of cities; the aggregate Markov variable has $K = \kappa + 2$ regimes (one for each idiosyncratic contraction, the national expansion, and the national contraction).[15]

3.1. Clustered Markov-switching

Formally, let $\mathbf{y}_t$ denote an $(N \times 1)$ vector of observed city-level building permit growth rates at date t and $\mathbf{Y}_t = (\mathbf{y}'_t, \mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_1)'$. Denote $\mathbf{S}_t$ as an $(N \times 1)$ vector of contraction indicators (so $S_{nt} = 1$ when city n is in contraction and $S_{nt} = 0$ when city n is in expansion). Suppose that

$$\mathbf{y}_t = \mu_0 + \mu_1 \odot \mathbf{S}_t + \varepsilon_t, \tag{1}$$

where the nth element of the $(N \times 1)$ vector $\mu_0 + \mu_1$ is the average building permit growth in city n during contraction, the nth element of the $(N \times 1)$ vector μ_0 is the average building permit growth in city n during expansion, and $\odot$ is the Hadamard (element-by-element) product. For identification, we assume that $\mu_{0n} > 0$ and $\mu_{1n} < 0$; that is, contractions are defined as strictly lower mean growth rates than expansions. Let $E(\varepsilon_t \varepsilon'_t) = \Sigma$, and we assume that the covariance matrix is diagonal with representative element σ_n^2. The diagonality restriction is made for parsimony and implies that the correlation across cities is driven primarily through simultaneity in their cycles.

We can summarize the individual S_{nt} with a scalar aggregate regime indicator Z_t that represents the time-t aggregate regime. Let $\mathbf{H}$ denote an $(N \times K)$ matrix whose elements are all zeros and ones and where K is the allowed number of possible aggregate permutations (regimes), including both idiosyncratic contractions and national expansion and contraction. The row n, column k element of $\mathbf{H}$ is 1 if city n is in a contraction when the aggregate regime is k. In a model in which all cities enter and exit contractions at the same time, $K = 2$, with the first column being all zeros and the second column is all ones. As an example, suppose that $K = 3$; this indicates three regimes: national-level expansion, national-level contraction, and one idiosyncratic contraction.

For exposition purposes only, consider the example of a cluster consisting of cities with manufacturing sectors larger than some predefined threshold. For $k = 1$, all cities are in expansion, by definition. For $k = 2$, all cities are in contraction, by definition. For $k = 3$, only the cities with manufacturing sectors above a certain threshold are in contraction; all other cities are in expansion. For purposes of this discussion, we refer to the regimes in which all cities move together as *national* regimes and refer to regimes in which some cities are in contraction but others are not as *idiosyncratic* contractions.[16]

The aggregate regime follows a polychotomous K-state Markov process with $(K \times K)$ transition kernel $\mathbf{P}$. In principle, we could model a world in which Z_t is allowed to transition to and from any aggregate regime. HO impose additional restrictions for identifying the clusters. They assume that the aggregate regime is free to transition to and from national-level expansions or contractions at any time. That is, if we label the first two regimes as national-level expansions or contractions, the first two rows and columns of $\mathbf{P}$ are unrestricted. However, HO impose the restriction that the aggregate regime cannot transition from one idiosyncratic cluster contraction to another. That is, if $K = 4$ and $Z_{t-1} = 4$, Z_t can take on only values of 1, 2, or 4. This aspect presents as zero restrictions on the transition kernel $\mathbf{P}$ and is described in more

detail in the estimation section below.

The model can alternatively be depicted as a mixture of distributions for the mean growth rate of permits. The distribution of the growth rate of permits conditional on being in (aggregate) regime k is

$$\mathbf{y}_t | z_t = k \sim N(\mathbf{m}_k, \Sigma),$$

where

$$\mathbf{m}_k = \mu_0 + \mu_1 \odot \mathbf{h}_k$$

for $\mathbf{h}_k$, the kth column of $\mathbf{H}$.[17]

It is important to note here that our setup allows cities to belong to only one cluster at each iteration of the sampler. At the end of the estimation, however, the results for each city will yield a posterior probability of belonging to each cluster. We then impose membership into a single cluster by setting a threshold for the posterior probabilities and assign each city to the unique cluster where the probability exceeds the threshold. The advantage of imposing membership in only one cluster is that these clusters can be interpreted as proper "regions". In our framework, the regions are not necessarily geographic but depend on the cyclical similarity of the member cities.

One might wonder what the advantage of using the clustered panel approach is compared with estimating each city separately. In the univariate Markov switching model, the posterior regime probability tends to identify a contraction whenever the growth rate begins to turn negative. Thus, the model can pick up very short-lived, city-specific negative growth periods. When the data are noisy, the univariate filter may identify a large number of turning-points as a city can be in contraction at any time, independent of the other cities. In the panel, contractions are only identified when a number of cities move together. Thus, our cluster recessions will tend to be more persistent than those identified by a collection of univariate Markov switching models. This feature is not imposed as a restriction, per se, but is a product of the multivariate filter.

3.2. Logistic clustering

One of the main features of the model is that it can be used to explain why cities' housing cycle experiences are correlated. To do this, we can model the cluster indicators h_{nk} as functions of a vector of fixed city-level covariates $\mathbf{x}_{nk}$ that influences whether city n is in a contraction when $Z_t = k$. Following Frühwirth-Schnatter and Kaufmann (2008) and HO, we assume that the probability that a city is in cluster k is defined by:

$$p(h_{nk} = 1) = \begin{cases} \frac{1}{1 + \sum_{k=2}^{\kappa} \exp(\mathbf{x}'_{nk}\beta_k)} & \text{for } k = 1 \\ \frac{\exp(\mathbf{x}'_{nk}\beta_k)}{1 + \sum_{k=2}^{\kappa} \exp(\mathbf{x}'_{nk}\beta_k)} & \text{otherwise} \end{cases} \tag{2}$$

for $n = 1, \ldots, N$; $k = 1, \ldots, \kappa$. Note that (2) resembles the probability that we would obtain from a multinomial logistic regression model.

For implementation, we consider $p(h_{nk})$ as the prior probability that city n belongs to cluster k conditional only on the observed covariates. The business cycle data in conjunction with the prior probability will determine the posterior probability through an application of Bayes' rule. As explained in HO, we can take the βs as population parameters even though $p(h_{nk})$ represents a form of the prior probability. We can estimate the βs as we would other model parameters and they will determine which and to what extent the covariates are important for clustering.

---

[15] The model is taken from Hamilton and Owyang (2012) and is also similar to Kaufmann (2010). These papers use clustering algorithms similar to Frühwirth-Schnatter and Kaufmann (2008) to reduce the dimension of the aggregate Markov-switching process. Other papers with multivariate Markov-switching models include Paap et al. (2009) and Leiva-Leon (2012).

[16] We do not refer to idiosyncratic *expansions* as these are simply idiosyncratic recessions for the complement set of states.

[17] Note that this specification imposes that the mean rate of growth during a contraction for each city is the same regardless of whether the cities fall into a national or idiosyncratic contraction. The model could be extended to allow for regime-specific growth rates, as in Eo and Kim (2016) or to differentiate between national contraction and idiosyncratic contraction growth rates.

4. Estimation

The model presented above is straightforward to estimate in a Bayesian environment. Given a prior, the joint posterior of the model parameters including the regimes can be generated by the Gibbs sampler (see Gelfand and Smith (1990), Casella and George (1992), Carter and Kohn (1994)). The Gibbs sampler iterates over draws from each parameter's conditional posterior distribution. After discarding a number of initial draws to achieve convergence, the remaining draws form the full joint posterior distribution of all of the model parameters.

Let Θ represent the full set of parameters. Then, Θ includes the regime growth rates, μ_0 and μ_1; the covariance matrix, Σ; the transition probabilities, $\mathbf{P}$; the time series of aggregate regimes, $\mathbf{Z}_T$; the matrix defining the clusters $\mathbf{H}$; and the logistic parameters, β, ψ, and ξ. The number of clusters, κ, is assumed to be fixed to estimate the other model parameters and is discussed further below. For now, we will assume that the number of clusters κ is determined exogenously and is suppressed in the notation. There are four blocks of parameters to be sampled: each city's parameter set, $\theta_n = (\mu_{0n}, \mu_{1n}, \sigma_n^2)$; the aggregate business cycle, $\mathbf{Z}_T$, and its associated transition matrix, $\mathbf{P}$; the matrix $\mathbf{H}$ determining the cluster membership; and the logistic parameters, β, ψ, and ξ.

4.1. Priors

The Bayesian environment requires a set of prior distributions for the model parameters. The distributional assumptions for the priors will, in turn, yield distributional assumptions on the posteriors. The city cycle parameters θ_n are assumed to have a normal-inverse Gamma prior distribution. The transition probabilities for the aggregate regime process are assumed to have a Dirichlet prior distribution given the fixed number of regimes. The cluster indicators have the logistic prior discussed above with population parameters β that are normal. Prior hyperparameters are shown in Table 3.

4.2. Posterior inference

As we noted above, the Gibbs sampler consists of iterative draws from the conditional distributions of the model parameters. In this subsection, we describe the draws; details for the sampler's posterior distributions can be found in HO.

Conditional on the other model parameters, the set of city-level parameters, θ, are conjugate normal-inverse Gamma and independent for all n. Thus, for each n, we first draw the μ_{0n} and μ_{1n} from a normal posterior distribution that depends in part on the regime-dependent conditional mean. We then draw the σ_n^{-2} from a Gamma posterior distribution. These draws can be made independently because the ε_{nt}s are assumed to be uncorrelated.

Conditional on the other model parameters, the posterior distribution of $\mathbf{Z}_T$ can be obtained from a multi-regime extension of the Hamilton (1989) filter, a discrete-state modification of the familiar

Kalman filter. For the sampler, we compute the posterior regime probabilities for each time period and draw each Z_t recursively, starting with Z_T. This draw is described in Kim and Nelson (1999). Once we have obtained the regimes, we can compute the posterior distributions for probabilities in the transition kernel. These are conjugate Dirichlet distributions (the multiple regime equivalent to the beta distribution) and depend on the observed number of transitions from one regime to another. Because we are restricting the number of transitions between idiosyncratic contractions to zero, the posterior distribution for these transition probabilities will also be identically zero.

The cluster indicators can be drawn by an application of Bayes' rule. We draw the value of the cluster membership indicator for each city-cluster combination conditional on the memberships of all of the other cities. The posterior probability is influenced by the logistic prior probability (i.e., the covariates and the estimated βs) and the similarities of city n 's housing cycle to the cycles of the cities in the cluster in question.

The logistic prior parameters are drawn in three steps. The marginal inclusion coefficient, β_{nk}, is drawn from a conjugate normal. The logistic prior requires two other parameters: a latent logistic variable, ξ_{nk}, whose sign is determined by the value of h_{nk} and the variance of this variable, λ_{nk}. We follow Holmes and Held (2006) who generate ξ_{nk} from a truncated logistic and then generate λ_{nk} conditional on ξ_{nk}.

4.3. Choosing the number of clusters

The model outlined in Section 3 is defined for a fixed number of clusters, κ. HO use techniques outlined in Chib (1995) and Chib and Jeliazkov (2001) to compute marginal likelihoods to determine κ. These methods use resampling techniques to compute the posterior ordinate – a component of the marginal likelihood – and are accurate

Table 3
Priors for Estimation.

Parameter	PriorDistribution	Hyperparameters	
$[\mu_{0n}, \mu_{1n}]'$	$N(\mathbf{m}, \sigma^2\mathbf{M})$	$\mathbf{m} = [2, -1]'$; $\mathbf{M} = \mathbf{I}_2$	$\forall\, n$
σ_n^{-2}	$\Gamma\left(\frac{\nu}{2}, \frac{\delta}{2}\right)$	$\nu = 2$; $\delta = 2$	$\forall\, n$
$\mathbf{P}$	$\mathbf{D}(\alpha)$	$\alpha_i = 0$	$\forall\, i$
β_k	$N(\mathbf{b}, \mathbf{B})$	$\mathbf{b} = \mathbf{0}_p$; $\mathbf{B} = \frac{1}{2}\mathbf{I}_p$	$\forall\, k$

Notes:
The table shows the prior distributions and their hyperparameters for all model parameters.
$D(\cdot)$ is the Dirichlet distribution.

Table 4
Smallest and largest permits growth rates: expansion .

Cities with smallest μ_0		
CBSA	MSA	μ_0
25860	Hickory-Lenoir-Morganton, NC	0.0402
39300	Providence-New Bedford-Fall River, RI-MA	0.0461
28940	Knoxville, TN	0.0565
33340	Milwaukee-Waukesha-West Allis, WI	0.0575
40380	Rochester, NY	0.0586
44140	Springfield, MA	0.0601
17140	Cincinnati-Middletown, OH-KY-IN	0.0715
17460	Cleveland-Elyria-Mentor, OH	0.0733
31700	Manchester-Nashua, NH	0.0734
10420	Akron, OH	0.0764

Cities with largest μ_0		
CBSA	MSA	μ_0
46700	Vallejo-Fairfield, CA	0.2665
14260	Boise City-Nampa, ID	0.2671
12940	Baton Rouge, LA	0.2703
41700	San Antonio-New Braunfels, TX	0.2704
21500	Erie, PA	0.2779
17820	Colorado Springs, CO	0.2803
13140	Beaumont-Port Arthur, TX	0.3203
12420	Austin-Round Rock-San Marcos, TX	0.3607
28660	Killeen-Temple-Fort Hood, TX	0.3813
41940	San Jose-Sunnyvale-Santa Clara, CA	0.4702

Notes:
The table shows the cities with the smallest (top panel) and largest (bottom panel) growth rates in permits during expansion (μ_0).
CBSA is the city's Core Based Statistical Area code.
Growth rates are the average year-over-year changes given in decimal points (0.01 is one percent).

Table 5
Smallest and largest permits growth rates: contraction.

Cities with smallest $\mu_0 + \mu_1$

CBSA	MSA	$\mu_0 + \mu_1$
38940	Port St. Lucie, FL	−0.3281
40140	Riverside-San Bernardino-Ontario, CA	−0.3229
33700	Modesto, CA	−0.3071
15980	Cape Coral-Fort Myers, FL	−0.3032
37100	Oxnard-Thousand Oaks-Ventura, CA	−0.2874
37340	Palm Bay-Melbourne-Titusville, FL	−0.2842
12060	Atlanta-Sandy Springs-Marietta, GA	−0.2668
35300	New Haven-Milford, CT	−0.2534
38060	Phoenix-Mesa-Glendale, AZ	−0.2503
33100	Miami-Fort Lauderdale-Pompano Beach, FL	−0.2415

Cities with largest $\mu_0 + \mu_1$

CBSA	MSA	$\mu_0 + \mu_1$
19340	Davenport-Moline-Rock Island, IA-IL	−0.0240
21340	El Paso, TX	−0.0128
21500	Erie, PA	−0.0080
30780	Little Rock-North Little Rock-Conway, AR	−0.0064
16620	Charleston, WV	−0.0046
39340	Provo-Orem, UT	0.0032
22180	Fayetteville, NC	0.0121
33660	Mobile, AL	0.0501
35380	New Orleans-Metairie-Kenner, LA	0.1049
13140	Beaumont-Port Arthur, TX	0.1490

Notes:
The table shows the cities with the smallest (top panel) nd largest (bottom panel) growth rates in permits during contraction ($\mu_0 + \mu_1$).
CBSA is the city's Core Based Statistical Area code.
Growth rates are the average year-over-year changes given in decimal points (0.01 is one percent).

Table 6
Smallest and largest permits growth rate: volatility.

Cities with smallest σ^2

CBSA	MSA	σ^2
41180	St. Louis, MO-IL	0.0840
17140	Cincinnati-Middletown, OH-KY-IN	0.0875
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.0910
17460	Cleveland-Elyria-Mentor, OH	0.0926
38300	Pittsburgh, PA	0.0927
39300	Providence-New Bedford-Fall River, RI-MA	0.0946
16980	Chicago-Joliet-Naperville, IL-IN-WI	0.0990
16860	Chattanooga, TN-GA	0.1069
33460	Minneapolis-St. Paul-Bloomington, MN-WI	0.1077
26900	Indianapolis-Carmel, IN	0.1121

Cities with largest σ^2

CBSA	MSA	σ^2
16620	Charleston, WV	0.4333
22180	Fayetteville, NC	0.4437
28660	Killeen-Temple-Fort Hood, TX	0.4671
21500	Erie, PA	0.4939
44060	Spokane, WA	0.4946
46700	Vallejo-Fairfield, CA	0.5942
13140	Beaumont-Port Arthur, TX	0.6681
41500	Salinas, CA	0.6875
41940	San Jose-Sunnyvale-Santa Clara, CA	1.0995
28940	Knoxville, TN	1.3928

Notes:
The table shows the cities with the smallest (top panel) and largest (bottom panel) residual volatility in permits growth rates (σ^2).
CBSA is the city's Core Based Statistical Area code.
Growth rates are the average year-over-year changes given in decimal points (0.01 is one percent).

when computed with a large number of post-convergence iterations. While simple to code, the methods of Chib (1995) and Chib and Jeliazkov (2001) use Monte Carlo integration to compute the posterior ordinate and are computationally intensive when the model has a large number of parameter blocks. The number of clusters is then chosen as the model yielding the highest marginal likelihood which must be computed for each possible value of κ. Because HO considered states, the number of clusters was assumed to be small and only a small number of marginal likelihoods were required to determine κ.

In the present analysis, we have a large number of cities, which would require a larger number of marginal likelihood computations. Because the methods used to compute the marginal likelihoods are time-consuming, we opted instead for computing a modified version of the BIC for each κ for a number of clusters between 3 and 9. Computing the BIC has been shown to approximate the marginal likelihood (e.g., Kass and Raftery, 1995; Raftery, 1995) and is faster to compute because it does not require resampling. The modification attaches a prior to the model dimension, putting more weight on parsimonious models. We thus compute the average of the BICs computed at each draw of the Gibbs sampler. We then select the number of clusters that minimizes this measure, which in our results corresponds to $\kappa = 3$ and $K = 5$.

5. Results

The results of the estimation are comprised of the growth rates and variances of permits for each city, the regime processes for the nation and the subgroupings, and the cluster compositions.
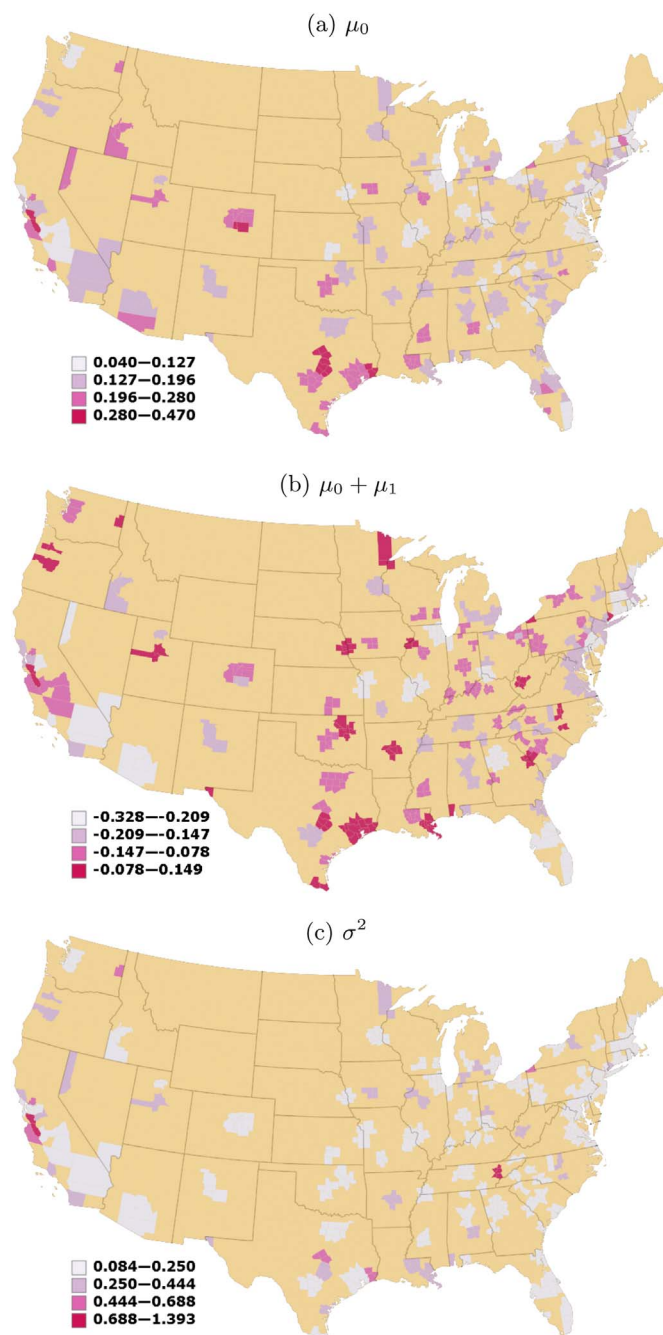
5.1. Growth rates

The model posits that the growth rate in city-level permits take on two average values over the housing cycle. During expansions, the mean growth rate is μ_0; during contractions, the mean growth rate falls to $\mu_0 + \mu_1$, since $\mu_1 < 0$. Table 4 lists the cities with the largest and smallest average expansion growth rates; Table 5 lists the cities with the largest and smallest average contraction growth rates, and Table 6 lists the cities with the largest and smallest variances. Fig. 2 maps the means of μ_0, $\mu_0 + \mu_1$, and σ^2 for each of the cities in our sample.[18]

While there are large differences in the means across cities, there are surprisingly few geographic patterns. A number of cities in Texas, for example, have high expansion growth rates but also have large negative contraction growth rates. However, these combinations are not limited to Texas—or even warm or southern cities. Some cities in upstate New York, for example, experience similarly large expansion growth rates. Cities that have large expansion rates appear to also have large contraction rates: the correlation between μ_0 and $\mu_0 + \mu_1$ is 24.8 percent.

Perhaps not surprisingly, cities with high growth rates also typically have a higher conditional variance; the correlation between μ_0 and σ^2 is 46.1 percent. Because the model restricts the cycle processes to achieve parsimony, city-level fluctuations will manifest in the residuals. This result then suggests that cities with a rapidly growing housing stock are likely to also experience large, cycle-independent fluctuations.

---

[18] Table 1 shows the mean growth rates and the city-level variances for all of the 135 cities in our sample and includes the mean and sample standard deviation of each city's raw building permits growth series.

(a) μ_0

☐	0.040—0.127
☐	0.127—0.196
☐	0.196—0.280
☐	0.280—0.470

(b) $\mu_0 + \mu_1$

☐	-0.328—-0.209
☐	-0.209—-0.147
☐	-0.147—-0.078
☐	-0.078—0.149

(c) σ^2

☐	0.084—0.250
☐	0.250—0.444
☐	0.444—0.688
☐	0.688—1.393

Fig. 2. Model estimates, $\kappa = 3$.

5.2. National housing cycles

One of the main features of the HO model is that it estimates a pervasive national cycle, which affects—by assumption—all of the cities simultaneously. Because of the restriction that the national cycle includes all cities, one believing in abundant city-level heterogeneity might imagine that there would not be much of a national cycle. On the other hand, one believing in pervasive cross-city linkages might imagine that idiosyncratic cycles would be less common.

Fig. 3 plots the posterior probabilities of a *national*-level housing contraction, $\Pr[Z_t = 2|\mathbf{Y}_T]$. The probabilities reflect the uncertainty around the polychotomous outcomes: a national housing expansion, a national contraction, or one of the κ cluster contractions. We note a few key features of the national housing contractions. First, the nation experiences two major housing downturns that coincide with business
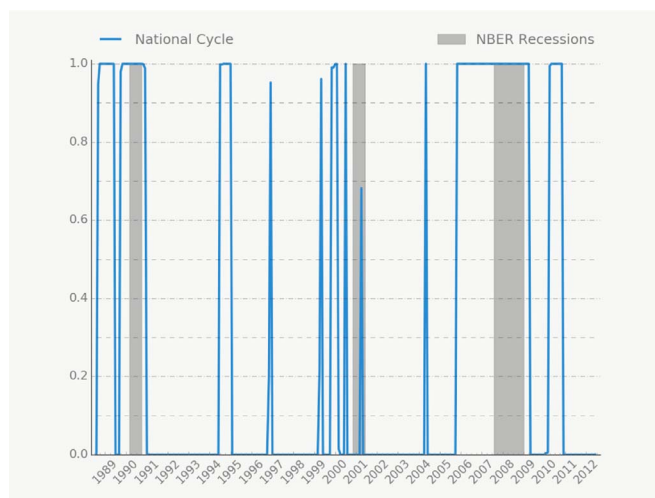
Fig. 3. Posterior probabilities of a national level housing recession, $\kappa = 3$.

cycle downturns. The first housing downturn is around the time of the 1991 NBER recession and the second is around the time of the 2007–2009 NBER recession; the housing contraction lasts a little longer than the business cycle contraction in both cases. This timing is not surprising for the 2007–2009 episode, as it was an economic downturn specifically associated with a decline in the housing market.

Second, national housing contractions are identified with little uncertainty—that is, there are very few periods for which the regime probability is between zero and one. Third, a few instances of national housing contractions are not associated with national recessions.

Finally, a long-lasting national housing contraction does not occur during the 2001 NBER recession in our sample (only short-lived episodes). This divergence in cycles is likely because there were localized housing downturns during these periods but they were not pervasive enough to include all of the cities in our sample. Thus, they can be characterized by cluster contractions instead of national contractions.[19]

5.3. City-level housing cycles

Before we can compare the city-level cycles, we must first determine the number of clusters preferred by the data. We treat the number of clusters as a model selection problem and compute a modified Bayesian information criterion (BIC) for various numbers of clusters. Our objective is to be as parsimonious as possible, so the parameter penalty in the BIC helps reduce the chance that the model becomes over-parameterized.

We choose the model with $\kappa = 3$. As our method generates posterior probabilities of belonging to each cluster, we classify a city falling in a cluster if it has higher than 50 percent posterior probability in that cluster. If a city belongs to no cluster by this metric, we place it in a separate group. The clusters so defined are mutually exclusive because the method generates cluster posterior probabilities that add up to one.

Once we have determined the number of city-level idiosyncratic cycles, we can examine how the aggregate regime process evolves. Recall that, while we have defined the *national cycles* as the regimes for which all cities are either in or out of housing contractions, that is, national cycles imply that all cities are either in the high-growth or in the low-growth state simultaneously, there will still be idiosyncratic contractions for some cities that are realized when the aggregate

---

[19] Indeed, this period is identified as a national housing contraction in models with a smaller the number of clusters. However, as we will see below, the data prefer a model with a larger number of idiosyncratic clusters.This finding is consistent with a large but not full set of cities experiencing a housing downturn in 2001.

Table 7
Transition probabilities.

		To Exp	To Rec	To C1	To C2	To C3
		$Z_{t+1} = 1$	$Z_{t+1} = 2$	$Z_{t+1} = 3$	$Z_{t+1} = 4$	$Z_{t+1} = 5$
From Exp	$Z_t = 1$	0.9303	0	0	0.0697	0
From Rec	$Z_t = 2$	0	0.8825	0.0331	0	0.0845
From C1	$Z_t = 3$	1	0	0	**0**	**0**
From C2	$Z_t = 4$	0	1	**0**	0	**0**
From C3	$Z_t = 5$	0.1851	0	**0**	**0**	0.8149

Notes:

The table shows the transition probabilities (**P**) for the aggregate state variable (Z_t) for $\kappa = 3$ clusters and $K = \kappa + 2 = 5$ states.

Exp and **Rec** signify national contractions.

C1 through **C3** signify the corresponding cluster contraction.

Bolded zeros represent ex ante identifying restrictions on the probabilities.

Z_t cannot transition from one idiosyncratic regime to another. In addition to the imposed restrictions, we find that a few other transition probabilities are estimated to be zero. For example, we find no transitions from $Z_t = 1$ to $Z_{t+1} = 2$. How can this occur? The data indicate that transitions from national expansions to national contractions can only occur through one of the cluster contractions—that is, a group of cities always begins contracting before it spreads to the entire country.

Fig. 4 plots the posterior probabilities of the idiosyncratic clusters, $\Pr[Z_t = k + 2|\Omega_T]$. We also plot the probability of the national contraction in each panel for reference. These figures illustrate that cities in Cluster 2 enter into a contraction ahead of other cities. Following a national contraction, cities in Cluster 1 come out of the contraction a period later than other cities, and cities in Cluster 3 take even longer to come out of a contraction. According to the estimated persistence
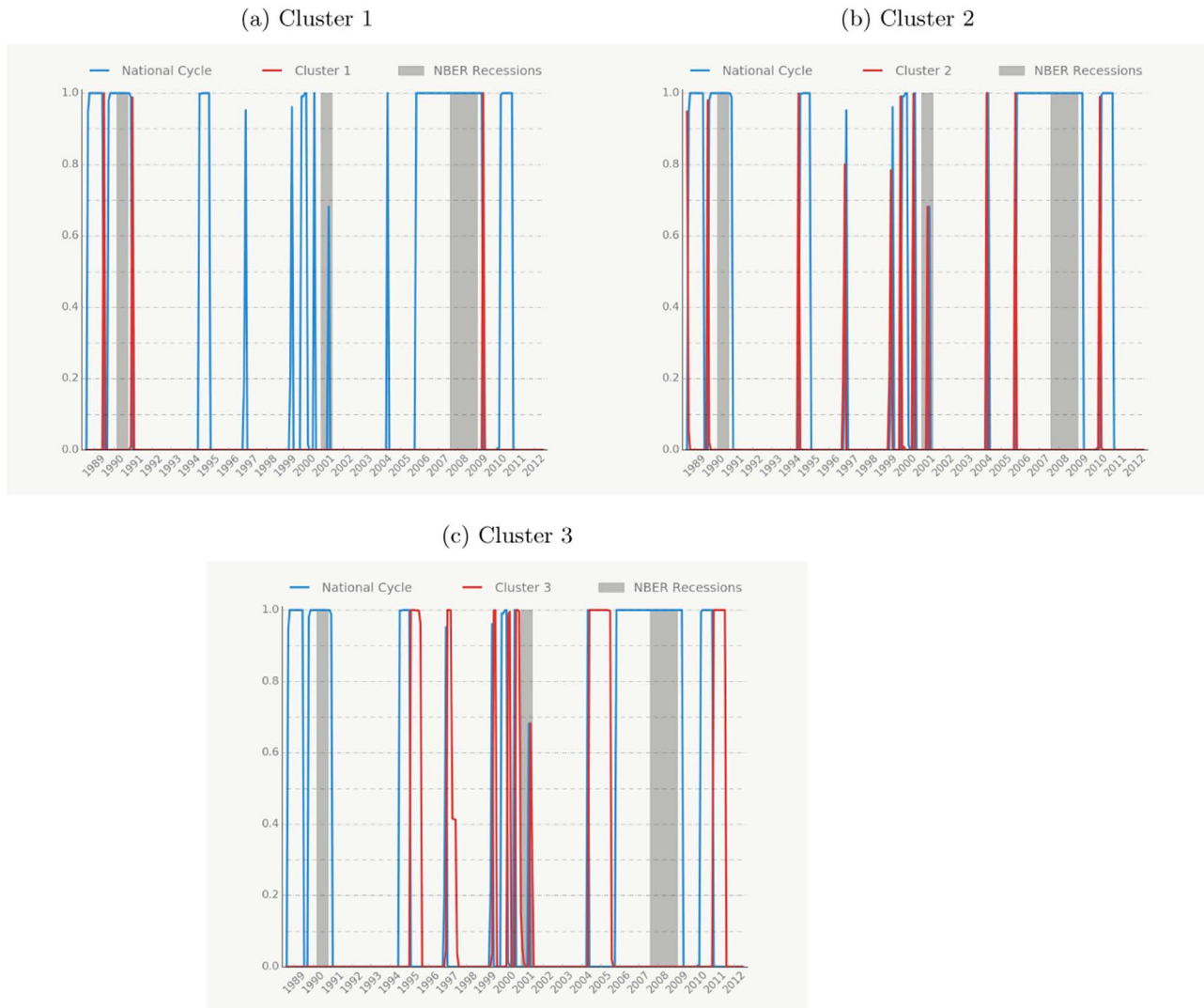
(a) Cluster 1

(b) Cluster 2

(c) Cluster 3

Fig. 4. Posterior probabilities. National cycle and idiosyncratic clusters, $\kappa = 3$.

regime is $Z_t = k \geq 3$. During a cluster-k contraction, cities in cluster k will be in the low-growth state while other cities are in the high-growth state. Recall also that we have imposed the identifying restriction that disables transitions between these idiosyncratic contractions. Thus, the (aggregate) economy must pass through either full expansion or full contraction before transitioning into another idiosyncratic contraction.

Table 7 shows the estimated transition probabilities for the aggregate regime, Z_t. Bold zeros reflect the imposed restriction that

probability, the latter group stays on average 5 periods, i.e. nearly half a year, longer in contraction than all other cities.

Fig. 5 shows the composition of the different clusters. In these maps, we display the cities that are members of each particular cluster. As with the growth rates, we find only loose geographic patterns in the behavior of the city housing cycles.

Cluster 1 contains cities all across the country without an obvious geographical pattern, although primarily in the Midwest, South, and
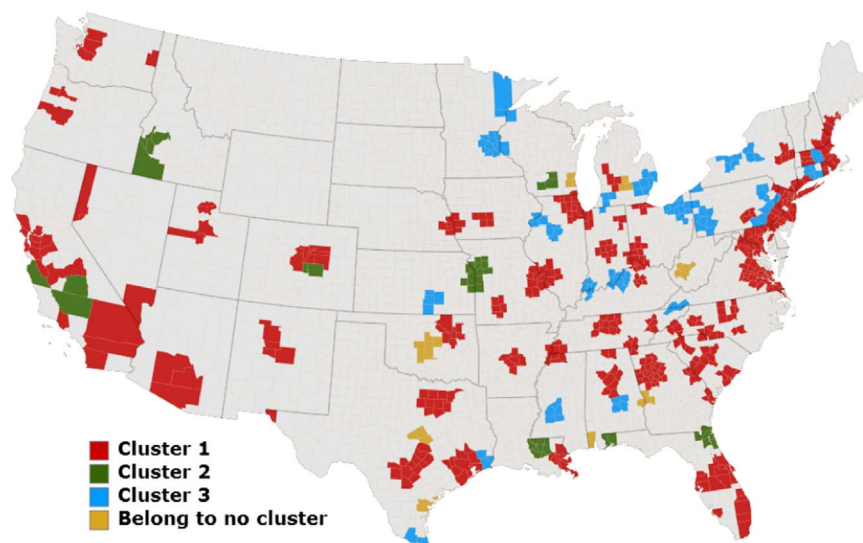
Fig. 5. City clusters.

the West and East coasts. The predominant feature of these cities' housing experiences is that they contracted in the early 1990s and late 2000s.

Cluster 2 is composed of cities that have idiosyncratic contractions but has a much smaller membership. Again, there is no real geographic pattern to membership. Neighboring cities do not appear to influence each other's membership to this cluster. The predominant feature of these cities' housing experiences is that they entered into contraction ahead of other cites. These results suggest that national downturns begin in a large number of cities and eventually grow to affect all cities.

Cities in Cluster 3 do, to some extent, exhibit a more geographical pattern with cities in the Midwest, the Northeast, and in the South appearing to be more likely to be included. Cities in this cluster remained in contraction following the national downturns of the mid 1990s. Cities in Cluster 3, if they stay longer in contraction than the other cities, they stay considerably longer. According to Fig. 4 panel (c), the contractions of cities in Cluster 3 lasted twice as long as for all other cities. Cities in Cluster 3 also remain in contraction after the final national contraction in our sample, which occurred around 2010.

In general, we conclude that geography and size (population) do not appear to be the only key determinants of cluster composition. The larger clusters might appear regionally connected but they are so large as to obscure any real patterns. In the next section, we formally investigate what factors might determine cluster membership.

5.4. What factors affect the co-movements

To help determine the factors that might make city-level housing cycles co-move, we included a set of city-level covariates in the prior for the clusters.[20]

These variables are intended to characterize differences and similarities in the supply and demand for housing across cities. Demand elements include population growth (which could reflect migration in or out of the city), average winter temperature (which may reflect location preference), and the average unemployment rate (which may proxy for economic conditions affecting household income). Housing density and the index of undevelopable land are both indicators of

Table 8
Marginal effects of the covariates.

Covariate	β_2	β_3
Constant	**−0.180**	**−0.223**
Housing units per sq. km. 1990	**−0.062**	−0.042
Avg. population growth (%) 1970–1990	−0.009	**−0.192**
Manufacturing employment share (%) 1990	−0.050	**0.103**
Average winter temperature (° Celsius)	−0.004	−0.027
Average unemployment rate (%) 1988–2012	0.001	0.057
Saiz's (2010) index of undevelopable land (%)	0.000	−0.003
Saiz's (2010) elasticity of land supply	0.003	−0.017
Change in share of subprime mortgages 2002–2005	−0.011	**0.092**
Growth in subprime mortgages 2002–2005	−0.020	0.030

Notes:
Coefficients other than the constant, *c*, have been scaled by $\Lambda(c)(1 - \Lambda(c))$, where $\Lambda(x) = (1 + \exp(-x))^{-1}$ is the logistic *cdf*. Because the covariates have been standardized to have mean zero and unit standard deviation, the scaled coefficients can be interpreted as marginal effects evaluated at the means, and represent the change in the likelihood of cluster membership in response to a one-standard-deviation from the mean in the variable of interest.
Coefficients in bold indicate that zero lies outside the interior 95 percent coverage interval.
In the multinomial logistic, coefficients β_1 have been normalized to 0.

housing supply and indicate how easily new housing can be constructed in the metro area.[21] We also include covariates related to subprime mortgages, as an indicator of local financial conditions and credit availability and the share of manufacturing employment as proxy for industrial composition in the city.

Table 8 shows the estimated values of the coefficients in the multinomial logistic prior, where bold indicates values for which zero is not contained in the 95-percent coverage interval of the posterior distribution of the βs. Because we estimate a multinomial logistic model, the parameters of Cluster 1 are normalized to 0 and the coefficients for Clusters 2 and 3 are interpreted relative to Cluster 1.

We find that housing demand factors play a more important role for determining membership to Cluster 3 relative to Cluster 1 (as indicated

[20] The prior can be thought of as sorting a city into or out of a cluster when the cyclical data (the city's cycle compared to the cluster's cycle) is inconclusive.The effect of the prior, in a sense, is to sort the city into the cluster if the city has similar characteristics to the other cities in that cluster. The prior is only important if the likelihood of the being in or being out of the cluster are about equal.

[21] We acknowledge that the categorization of the covariates in terms of housing demand and supply factors is, to some extent, arbitrary, however we believe it helps organize the discussion.

Table 9
Summary statistics by cluster.

Covariate	Cluster	Mean	SD	Median	Min	Max	N
Housing units per sq. km. 1990	1	61.9	60.4	42.1	5.4	380.7	84
	2	22.3	13.4	24.2	4.1	46.8	10
	3	55.9	37.5	52.6	5.9	165.4	33
	No cluster	43.8	43.7	31.0	13.8	148.6	8
Avg. population growth (%) 1970–1990	1	1.8	1.4	1.5	−0.1	6.0	84
	2	1.9	0.7	1.9	0.7	2.7	10
	3	0.6	0.9	0.4	−0.6	3.8	33
	No cluster	0.8	0.8	0.8	0.0	2.3	8
Manufacturing employment share (%) 1990	1	20.2	8.3	19.3	3.2	58.5	84
	2	14.4	4.4	15.0	6.9	19.8	10
	3	27.1	7.0	27.9	12.6	39.9	33
	No cluster	18.1	6.3	16.1	10.9	26.7	8
Average winter temperature (° Celsius)	1	5.5	6.5	5.3	−6.7	19.6	84
	2	6.3	7.7	8.1	−7.8	15.3	10
	3	0.4	6.1	−1.5	−9.6	16.9	33
	No cluster	5.1	7.7	6.3	−5.8	13.6	8
Average unemployment rate (%) 1988–2012	1	5.8	1.6	5.5	3.6	12.6	84
	2	7.2	3.4	5.6	3.3	13.8	10
	3	6.2	1.9	5.7	4.2	14.2	33
	No cluster	5.9	0.8	6.0	4.5	7.1	8
Saiz's (2010) index of undevelopable land (%)	1	29.3	22.8	23.2	1.0	79.6	84
	2	31.9	19.1	28.9	5.8	66.1	10
	3	20.6	14.3	17.9	0.9	55.3	33
	No cluster	25.0	24.8	18.1	2.5	71.7	8
Saiz's (2010) elasticity of land supply	1	1.9	1.0	1.6	0.7	5.4	84
	2	1.8	0.6	1.7	1.1	3.2	10
	3	2.4	1.1	2.2	0.9	5.5	33
	No cluster	2.6	1.2	2.3	1.0	4.6	8
Change in share of subprime mortgages 2002–2005	1	−1.6	10.5	−2.2	−36.8	44.1	84
	2	−3.1	1.6	−3.7	−5.0	−0.9	10
	3	−0.8	1.4	−1.0	−3.1	3.1	33
	No cluster	−1.3	0.9	−1.4	−2.6	0.4	8
Growth in subprime mortgages 2002–2005	1	9.1	11.5	10.0	−20.3	37.0	84
	2	9.1	9.9	9.5	−8.1	24.1	10
	3	17.8	12.8	14.5	−3.5	50.4	33
	No cluster	13.2	8.1	12.4	−1.6	23.0	8

Notes:
MSA population, housing units, and land area were aggregated from county-level data from the Census Bureau to match 2009 MSA definitions available at http://www.census.gov/population/metro/files/lists/2009/List4.txt (Retrieved on 11 April 2013).
The manufacturing employment share was computed as the ratio of MSA manufacturing employment to total employment, aggregating county-level data from the 1990 County Business Patterns from the Census Bureau.
Average winter temperatures represent long-run typical temperatures obtained for each city from the Department of Energy.
Unemployment rates were computed aggregating the number of the unemployed and the

labor force at the county level with data from the Bureau of Labor Statistics, and subsequently were averaged over the period 1988:1–2012:12.
Saiz's (2010) index represents the percent of area in each city that cannot be developed because of geographic constraints, and was graciously provided by the author.
The change in the proportion of subprime mortgage loans and the growth in mortgage loan balances were obtained from credit bureau data at the zip code level aggregated at the metropolitan area.

by the importance of population growth).

Saiz's indicators of land supply do not appear to be crucial factors in synchronizing cities' housing cycles, supporting the finding that geographic factors are not the main determinants of clustering. However, housing density seems to decrease the probability of belonging to Cluster 2 relative to Cluster 1.

These results do not suggest that supply factors are not important for housing markets in general. They may play a role in determining similarities in the trends in housing markets and may be more important for prices than for permits.[22]

Financial variables are important covariates for determining membership to Cluster 3. The change in the share of subprime mortgages has a positive and significant effect on the probability of being part of Cluster 3. Membership to Cluster 3, is additionally driven by economic factors. In particular, membership to Cluster 3 is positively influenced by a higher share of manufacturing employment.

The importance of economic factors, particularly for Cluster 3, is supported by traditional location-quotient and shift-share analyses that characterize the most important industries in subnational regions.[23] These traditional approaches show that cities in Cluster 1 are more specialized in the construction and leisure and hospitality sectors. For cities in Cluster 2, the most important sector is natural resources and mining. For cities in Cluster 3, manufacturing and education and health are the most important industries. (The location-quotient and shift-share calculations using employment data for 11 supersectors are available upon request.) The importance of the manufacturing sector is illustrated also in Table 9, which presents summary statistics of the covariates grouped by cluster.

We find that the determinants of housing cycle similarity differ substantially from the determinants of business cycle similarity (as in HO), which supports Ghent and Owyang's (2010) conclusion that, at a subnational level, housing cycles and business cycles are not as connected as they appear to be at the national level. We find that both housing demand- or supply-side factors are important determinants of similarities in the cyclical fluctuations in housing.

6. Conclusions

Despite the strong linkages often found between housing market cycles and business cycles at the national level, the relationship between housing and the business cycle is less clear when considering subnational data because housing markets are highly localized. In this paper, we examine the cross-city linkages in housing cycles by estimating a cluster Markov-switching model of building permits. We find that there does exist a national housing cycle which roughly coincides with the national business cycle as determined by the NBER. In addition, cities experience idiosyncratic housing contractions either through early entry into national contractions, prolonged exposure to national contractions, or purely idiosyncratic contractions.

We find that the presence of regional housing cycles may depend on

---

[22] In fact, the elasticity of land supply plays an important role in determining the clusters when we use house price growth as indicators of housing cycles.
[23] The location quotient is defined as the share of industry employment in a region relative to the share of industry employment in the nation. Shift-share analysis decomposes regional industry growth (e.g., in employment) into national trends, industry trends, and idiosyncratic regional factors. Both approaches are used to identify areas where a regional economy exhibits competitive advantages or disadvantages when compared with the nation. See, e.g. Coulson, 2010; Coulson et al., 2013.

local factors in addition to national factors. We estimate that idiosyncratic contractions occur in three clusters, for which membership is influenced by similarities in factors influencing both housing demand and supply or factors influencing the similarity of business cycles. For the most part, we also find that geography does not seem to be an important determinant of cluster membership.

Finally, regional differences in housing cycles can have implications for the transmission of monetary policy or for implementing macroprudential policies across regions in the U.S. The possibility of different reactions to a single national-level policy suggests that models should take heterogeneity into account and consider, for example, differences in housing demand, the size of the population, or different strengths of housing shocks across regions when designing policy. In light of our results, housing demand heterogeneity could be an important consideration when assessing the relative effectiveness of a policy across multiple regions. While monetary policy is designed to be determined at the national level, macroprudential policies or financial regulations could potentially address idiosyncratic housing cycles.

References

Bernanke, Ben S., 2008. Housing, Mortgage Markets, and Foreclosures, in Remarks delivered at The Federal Reserve System Conference on Housing and Mortgage Markets, Washington, D.C., December.

Brady, Ryan R., 2011. Measuring the diffusion of housing prices across space and over time. J. Appl. Econom. 26 (2), 213–231.

Brueckner, Jan K., 2000. Urban sprawl: diagnosis and remedies. Int. Reg. Sci. Rev. 23 (2), 160–171.

Can, Ayse S., 1990. The measurement of neighbourhood dynamics in urban house prices. Econ. Geogr. 66 (3), 254–272.

Carlino, Gerald, Sill, Keith, 2001. Regional income fluctuations: common trends and common cycles. Rev. Econ. Stat. 83 (3), 446–456.

Carter, Christopher K., Kohn, Robert, 1994. On Gibbs sampling for state space models. Biometrika 81 (3), 541–553.

Casella, George, George, Edward I., 1992. Explaining the Gibbs sampler. Am. Stat. 46 (3), 167–174.

Chib, Siddhartha, 1995. Marginal likelihood from the Gibbs output. J. Am. Stat. Assoc. 90 (432), 1313–1321.

Chib, Siddhartha, Jeliazkov, Ivan, 2001. Marginal likelihood from the Metropolis-Hastings output. J. Am. Stat. Assoc. 96 (453), 270–281.

Clayton, Jim, Miller, Norman G., Peng, Liang, 2010. Price-volume correlation in the housing market: causality and co-movements. J. Real Estate Financ. Econ. 40 (1), 14–40.

Cotter, John, Gabriel, Stuart, Roll, Richard, 2012. Can metropolitan housing risk be diversified? A cautionary tale from the recent boom and bust, Technical report, UCD Geary Institute.

Couch, Chris, Karecha, Jay, 2006. Controlling urban sprawl: some experiences from Liverpool. Cities 23 (5), 353–363.

Coulson, N. Edward, 2010. The Long Run Shift-Share: Uncovering the Sources of Sectoral Fluctuations in Metropolitan Areas, in Tim Bollerslev, Jeffrey Russell and Mark Watson, eds., Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, Oxford University Press, Oxford.

Coulson, N. Edward, Liu, Crocker H., Villupuram, Sriram V., 2013. Urban economic base as a catalyst for movements in real estate prices. Reg. Sci. Urban Econ. 43 (43), 1023–1040.

Davis, Morris A., Heathcote, Jonathan, 2005. Housing and the business cycle. Int. Econ. Rev. 46 (3), 751–784.

Del Negro, Marco, Otrok, Christopher, 2007. 99 Luftballons: monetary policy and the house price boom across U.S. States. J. Monet. Econ. 54 (7), 1962–1985.

Eo, Yunjong, Kim, Chang-Jin, 2016. Markov-switching models with evolving regime-specific parameters: are postwar booms and recessions all alike? Rev. Econ. Stat. 98 (5), 940–949.

Frühwirth-Schnatter, Sylvia, Kaufmann, Sylvia, 2008. Model-based clustering of multiple times series. J. Bus. Econ. Stat. 26 (1), 78–89.

Füss, Roland, Zhu, Bing, Zietz, Joachim, 2012. Why do local housing prices react so differently to a monetary stimulus? Technical report, University of St. Gallen.

Gelfand, Alan E., Smith, Adrian F.M., 1990. Sampling-based approaches to calculating marginal densities. J. Am. Stat. Assoc. 85 (410), 398–409.

Ghent, Andra C., Owyang, Michael T., 2010. Is housing the business cycle? Evidence from U.S. Cities. J. Urban Econ. 67 (3), 336–351.

Glaeser, Edward L., Gyourko, Joseph, Morales, Edward, Nathanson, Charles G., 2011. Housing Dynamics, May, unpublished Manuscript.

Hamilton, James D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.

Hamilton, James D., Owyang, Michael T., 2012. The propagation of regional recessions. Rev. Econ. Stat. 94 (4), 935–947.

Holly, Sean, Pesaran, Hashem M., Yamagata, Takashi, 2011. The temporal and spatial diffusion of house prices in the UK. J. Urban Econ. 69 (1), 2–23.

Holmes, Chris C., Held, Leonhard, 2006. Bayesian auxiliary variable models for binary and multinomial regression. Bayesian Anal. 1 (1), 145–168.

Iacoviello, Matteo, 2005. House prices, borrowing constraints, and monetary policy in the business cycle. Am. Econ. Rev. 95 (3), 739–764.

Iacoviello, Matteo, Neri, Stefano, 2010. Housing market spillovers: evidence from an estimated DSGE model. Am. Econ. J.: Macroecon. 2 (2), 125–164.

Ioannides, Yannis M., Zabel, Jeffrey E., 2003. Neighborhood effects and housing demand. J. Appl. Econom. 18, 563–584.

Kallberg, Jarl, Liu, Crocker H., Pasquariello, Paolo, 2014. On the price co-movement of U.S. residential real estate markets. Real Estate Econ. 42 (1), 71–108.

Kass, Robert E., Raftery, Adrian E., 1995. Bayes factors. J. Am. Stat. Assoc. 90 (430), 773–795.

Kaufmann, Sylvia, 2010. Dating and forecasting turning points by Bayesian clustering with dynamic structure: a suggestion with an application to Austrian data. J. Appl. Econom. 25 (2), 309–344.

Kim, Chang-Jin, Nelson, Charles R., 1999. State-space models with regime switching. The MIT Press, Cambridge, MA.

Landier, Agustin, Sraer, David, Thesmar, David, 2013. Banking deregulation and the rise in house price co-movement, Technical report.

Leamer, Edward E., 2007. Housing is the business cycle, Working Paper 13428, NBER, September.

Leiva-Leon, Danilo, 2012. Monitoring synchronization of regional recessions: a Markov-switching network approach, October, unpublished manuscript.

Meen, Geoffy, 1999. Regional house prices and the ripple effect: a new interpretation. Hous. Stud. 14 (6), 733–753.

Moench, Emanuel, Ng, Serena, 2011. A hierarchical factor analysis of U.S. housing market dynamics. Econom. J. 14 (1), C1–C24.

Oikarinen, Elias, 2006. Price linkages between stock, bond and housing markets-evidence from Finnish data, Technical Report 1004, Keskusteluaiheita, Discussion papers.

O'Sullivan, Arthur, 2009. Urban economics. McGraw Hill, New York, NY.

Owyang, Michael T., Piger, Jeremy M., Wall, Howard J., 2005. Business cycle phases in U.S. States. Rev. Econ. Stat. 87 (4), 604–616.

Owyang, Michael T., Piger, Jeremy M., Wall, Howard J., Wheeler, Christopher H., 2008. The economic performance of cities: a Markov-switching approach. J. Urban Econ. 64 (3), 538–550.

Paap, Richard, Segers, Rene, van Dijk, Dick, 2009. Do leading indicators lead peaks more than troughs? J. Bus. Econ. Stat. 27 (4), 528–543.

Pollakowsky, Henry O., Ray, Traci S., 1997. Housing price diffusion patterns at different aggregation levels: an examination of housing market efficiency. J. Hous. Res. 8 (1), 107–124.

Raftery, Adrian E., 1995. Bayesian model selection in social research. Sociol. Methodol. 25, 111–163.

Saiz, Albert, 2010. The geographic determinants of housing supply. Q. J. Econ. 125 (3), 1253–1296.

Stevenson, Simon, 2004. House price diffusion and inter-regional and cross-border house price dynamics. J. Prop. Res. 21, 301–320.

Stock, James H., Watson, Mark W., 2010. The Evolution of National and Regional Factors in U.S. Housing Construction, in Tim Bollerslev, Jeffrey Russell and Mark Watson, eds., Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, Oxford University Press, Oxford.