Interregional Migration and Housing Vacancy: Theory and Empirics

Interregional Migration and Housing Vacancy: Theory and Empirics

Jeffrey P. Cohen
Professor of Real Estate and Finance
University of Connecticut
Jeffrey.Cohen@business.uconn.edu

Cletus C. Coughlin
Emeritus Economist – Research Division
Federal Reserve Bank of St. Louis
coughlincc@charter.net

Jonas Crews
Senior Research Associate
Heartland Forward
crews@heartlandforward.org

Daniel Soques
Assistant Professor of Economics
University of North Carolina Wilmington
soquesd@uncw.edu

10/16/2020

Abstract. We examine homeowner vacancy rate interdependencies over time and space through the channel of migration. Our theoretical analysis extends the Wheaton (1990) search and matching model for housing by incorporating interregional spillovers due to some households' desires to migrate between regions and by allowing for regime-switching behavior. We explicitly test the hypothesis that migration drives vacancy rates by estimating a Markov-Switching Vector Autoregression (MS-VAR), which captures both the regime-dependent nature and persistence of the vacancy rate. Our empirical results suggest that fluctuations in the vacancy rate across Census regions are driven more by migration into a given region rather than migration outflows. Additionally, the MS-VAR setup allows us to capture asymmetric effects of these migration flows on regional vacancy rates. Asymmetric effects are evident for the South and Midwest regions, wherein migration flows have larger effects on vacancy rates during relatively high vacancy rate and high volatility regimes. Conversely, the regional vacancy rates for the Northeast and West respond similarly to migration, regardless of the vacancy regime.

Introduction

Housing and labor markets are closely linked, with migration often playing a key role. For example, sometimes moves to a new house are coupled with migration across regions within a country to take advantage of a job opportunity. Housing market shocks in some regions of the country, frequently in conjunction with employment shocks, can cause homeowners to choose to migrate to other regions. These shocks can also lead to vacant homes if not all homes are sold at the time of the out-migration. In times of high vacancies, lack of adequate in-migration may impact the vacancy rates in a region. Vacancies in a region can be correlated with vacancies in other regions when national housing cycles hit multiple regions simultaneously.

At a given point in time, a region might be experiencing one of multiple possible vacancy rate "regimes" (such as low, intermediate, or high vacancies). Vacancy rates are likely to vary across regions not only due to shocks, but also due to many other factors, such as housing construction regulations. These other factors affect a region's response to shocks leading to different probabilities of home sales across "regimes", varied probabilities of finding a new house, and varied "transition rates" (i.e., probabilities of the desire to move in different regimes). Underlying all of this are homeowner preferences to migrate, and the likelihood of finding a new home when and where they migrate (i.e., probability of selling an old house and finding a new one). Thus, homeowner actions are intimately tied to the vacancy rates of the region from which the homeowner is migrating.

We explore connections between regions by examining migration across regions, along with migration's relationship with a region's vacancy rate in its housing market. We also explore how vacancy rates across regions vary. Our primary goals are to highlight specific theoretical connections and to test hypotheses on the relationships between migration and housing vacancies, based on empirical differences across regions and regimes (i.e., vacancy "states") of the economy. This latter point highlights the importance of the Markov-switching models that we estimate. Our empirical results provide insights into the interconnectedness of regions and the workings of an important feature of the nexus between housing and migration.

Vacancies in the housing market (i.e., unoccupied housing units) are similar to unemployment in the labor market in that some level is desirable and expected in a well-functioning market. For example, homeowners may experience changes in their family or employment situations such that their existing house no longer meets their needs in terms of location and/or house characteristics, and, after a search, more appropriate housing is purchased. Thus, until the first house is sold, a given homeowner may own two houses, one of which is vacant. Vacancies also arise for other reasons. For example, newly constructed houses may be vacant for a period before occupancy.[1] Vacancy rates are important considerations to understand

---

[1] Two other reasons might also be noted. An individual might own two houses, one of which is occupied most of the year and the other that is used for vacations. The vacation house will likely be unoccupied for large portions of a year. Such vacancy can be viewed as intentional and does not suggest any housing market problem. What is termed

in particular during housing bubbles. Migration between regions is another important reason for vacancies, where it may take time for households to find a buyer for their previous home.

In our analysis, we first propose a generalization of a theoretical model that implies a possible relationship between vacancy rates in a region and migration from another region into that region. We also test the hypothesis that vacancy rates depend on the state of the housing market and migration. For example, we estimate separate vacancy rate models for what we term a low-vacancy regime and for a high-vacancy regime, with the high-vacancy regime being relatively higher than the low-vacancy regime. There is neither an absolute minimum difference between these regimes, nor is there an absolute maximum for a low-vacancy regime or an absolute minimum for a high-vacancy regime. In some cases, three regimes are identified: low, intermediate, and high. Upon entering a specific regime, the regime is highly persistent. Numerous circumstances can cause vacancies to rise and lead to a high-vacancy regime. National and local recessions as well as geographic shifts in demand due to migration can generate rising vacancies and lead to what we characterize as a high-vacancy regime.

Recent history provides a stark example of a recession that propelled a sharp upward movement in vacancy rates. As shown in Figure 1, during the housing crisis associated with the Great Recession, homeowner vacancy rates in the United States reached levels far greater than at any time since measurement began in the mid-1950s. Rates reached 2.9% during the housing crisis. As far back as records have been kept, rates had always been below 2.0%. As of Q3/2019, the national vacancy rate was 1.4%.

From our regional perspective, homeowner vacancy rates differ across space and, similar to the United States as a whole, change over time.[2] Numerous factors are likely to drive such results. These factors include the cost of holding vacant units, search costs and the matching process, expectations about future housing prices, demand for specific housing characteristics, the quantity and quality of intermediaries, the specific characteristics of the existing housing stock, transaction costs, land use regulations, and credit market imperfections.[3] In addition, migration-induced regional vacancy rate changes can be affected differentially by national economic shocks and by regional shocks. These differences across space and changes over time

long-term vacancy (i.e., non-seasonal housing units that have been vacant for an unusually long period of time) suggests the possibility of some fundamental problem, such as a declining neighborhood. See Molloy (2016) and Harrison and Immergluck (2020) for recent analyses of long-term vacancies. Due to the level of aggregation of our analysis, we do not address this important issue.

[2] While our research examines owner-occupied housing, a similar analysis could be done for rental property. An early study is Gabriel and Nothaft (1988). From the mid-1950s to the present, the correlation between homeowner and rental vacancy rates is 0.77. These rates move in opposite directions prior to the Great Recession, a period worthy of a separate study, as well as for a much more complicated analysis that considers the interaction between rental and homeownership vacancies.

[3] See Fritzsche and Vandrei (2014) for a discussion of the theoretical causes of vacancies and a summary of empirical findings. A related paper by Cheshire, Hilber, and Koster (2015) finds that regulations restricting new house construction increases rather than decreases vacancy rates. Regulations lead to higher prices providing incentives for occupying houses, but also impede the matching process. Empirically, for a sample of local housing markets in England, this latter effect, which tends to increase vacancies, dominates the former effect.

create the possibility of different regional vacancy rates. For example, as of Q3/2019, the vacancy rate was 1.4% in the Northeast Census region, 1.3% in the Midwest, 1.7% in the South, and 1.2% in the West.

We address the spatial variation theoretically by adapting and extending the Wheaton (1990) search and matching model for housing to include a direct connection between interregional housing market differences via migration primarily motivated by job opportunities. Changing jobs is a key factor in migration within the United States. For example, Ihrke (2014) found that with-in county moves are associated with housing-related issues, while between county moves are associated with job-related issues. Molloy, Smith, and Wozniak (2011) report that approximately 1.5% of the population moves between two of the four Census regions annually.

Empirically, we first document key features of the vacancy rates for the United States and the four Census regions via simple univariate Markov-switching models.[4] In the context of housing vacancy, this is the first known research that uses Markov-switching models. These models are important because they allow us to deal with the large changes (or regime shifts) in vacancy rates during major economic shocks and recoveries. At both the national and aggregate level, we find the vacancy rate is well-described by models with regime-switching and autoregressive behavior (i.e., persistence).

These common features of the vacancy rate time series motivate our choice to analyze migration spillovers in a Markov-switching Vector Autoregression (MS-VAR) model. The MS-VAR models allow us to capture both regime-switching and persistence, and coincides with our regional theoretical model. These multivariate models also allow us to test the level of interdependence of the Census regions and how that interdependence varies across regimes, by incorporating other regions' vacancy rates into the empirical models. Our research appears to be the first attempt to empirically identify and explain interregional vacancy relationships. Broadly, we find that fluctuations in a given region's vacancy rate are driven by households' desire to migrate *into* that region. Our empirical results suggest there is no consistent effect across regions on the vacancy rate due to migration *out* of that region. In addition, the MS-VAR allows us to capture if these migration spillovers depend upon the state (i.e., regime) of the regional housing market. By using state-dependent generalized impulse response functions, we document that two regions, the South and Midwest, tend to have larger spillover effects of migration flows during time periods characterized by a high average vacancy rate and high vacancy volatility. However, asymmetric effects are not common across all regions. In particular, migration effects are relatively similar across the business cycle for both the Northeast and West regions.

The rest of the paper proceeds as follows. We begin with a review of the literature focused on vacancies and migration. Next, we develop an extension of the Wheaton (1990)

---

[4] Another focus would be on MSAs rather than Census regions. Our framework allows for such an exercise, but we lack sufficient data for estimation.

model. This extension provides the theoretical foundation for our use of MS-VAR models. We then outline the data and estimation procedure. Next, we describe the key empirical features of the vacancy rate. These key features motivate the next section, which outlines the estimated migration spillover effects from the MS-VAR models. The last section concludes.

Migration and Housing Vacancy: Background Literatures

Changes in internal migration have garnered much research attention in recent years. Molloy, Smith, and Wozniak (2011) highlight that migration rates between Census regions, states, and MSAs have trended downward since the early 1980s. As discussed by Frey (2019), this trend has continued during recent years. While migration rates have declined, the population changing residences is still large. Molloy, Smith, and Wozniak (2017) argue that a decline in job changing has caused migration to decline; however, they are unable to find specific reasons for the decline in job changing. Bayoumi and Barkema (2019) argue that increasing inequality has contributed to declining long-distance migration. Migration from less to more prosperous MSAs has been slowed by increasing housing costs dominating increasing earnings, while migration from more to less prosperous MSAs has been slowed by decreasing earnings dominating decreasing housing costs.

While the demand for housing can decline when there is out-migration from a region, and demand can increase when there is in-migration, there are undoubtedly frictions that prevent a complete and instantaneous adjustment.[5] Molloy (2016) provides a number of arguments and references that would lead one to expect that price would not fall immediately. Goodman and Ittner (1992) argue that owners tend to overestimate the value of their property and Genesove and Mayer (2001) find a reluctance to sell property for less than property owners judge as its worth.[6] Especially when demand declines, property owners might not recognize the decline in the value of their property. Anenberg (2016) and Guren (2018) argue that owners set their asking prices based on the transactions prices of comparable properties sold recently. Thus, given a decrease in demand, some owners might have unrealistic expectations concerning the values of their property. Finally, if an owner is offered less than the mortgage amount, a common occurrence during the recent housing crisis, then sales become quite complicated. To complete the sale, either the lender must forgive the difference between the mortgage amount and the transaction price or the seller must make up the difference.

---

[5] The possibility of vacant housing is illustrated by Figure A1 in Appendix A via a simple supply and demand framework.

[6] The loss aversion argument of Genesove and Mayer (2001) illustrates how psychological concepts from behavioral economics can affect vacancy rates, a topic discussed in Fritzsche and Vandrei (2014). Stein (1995) offers another argument related to Genesove and Mayer (2001). In a declining house price environment, potential sellers are adversely affected by the resulting decline in wealth and liquidity. This produces a reluctance to sell because ever-lower prices decrease the potential seller's options for relocating as their capability of making a given down-payment is reduced.

In light of the preceding frictions, assume that price does not adjust fully and instantaneously to a decline in demand. Given this partial adjustment, the vacancy rate may be higher than its previous rate. Turning to the case of an increase in demand, one can also argue that price will not adjust completely and instantaneously. If so, then the vacancy rate will decline below its previous rate.[7] The effect of vacancies on housing prices has been a topic of increased attention due to the housing bubble; for instance, see Whitaker and Fitzpatrick (2013).

The next question is what happens when more time is allowed for adjustments in the housing market due to a homeowner's desire to migrate. Because a supply curve for housing for periods longer than the short run likely has a positive slope and can shift, the quantity of housing units can adjust upward via new construction and downward by depreciation/teardowns. On the demand side, whether the shock is temporary or permanent is of utmost importance. If the shock is temporary, because of a one-time change in migration, for instance, then one should expect price and vacancy to return to their original values. On the other hand, if the shock is permanent, then price and quantity will adjust further and their effects on the vacancy rate are uncertain without more detailed information on various quantitative relationships as well as the cause of the shock.

To elaborate on the preceding point, assume there is a permanent decline in demand due to outmigration from a region. As a result, demand for all housing units falls, as does demand for all occupied units. Similar to the case with a fixed housing supply, such a decline puts downward pressure on housing prices. If price adjusts completely and instantaneously, the new equilibrium quantities of total and occupied housing would decline. The number of vacancies would be unchanged and the vacancy rate would increase. However, as discussed previously, the actual shifts of occupied housing and total housing need not be parallel and equidistant. It is possible that in this new case the decline in occupied housing is relatively less (more) than the decline in total housing demand. This creates the possibility that the new vacancy rate is lower (higher) than the original vacancy rate. What remains certain is that the quantity of housing supplied is now less, the quantity of occupied housing is less, and the price of housing is less.

The foundations for the preceding supply and demand discussion are a housing search and matching model. Wheaton (1990) provided a one-region search and matching model that yields a vacancy rate. The literature on this topic has advanced in various other directions. For a recent literature review see Han and Strange (2015). See Williams (1995) for a continuous time version of Wheaton's model. The role and impact of bargaining in the housing search process is examined by Ihlanfeldt and Mayock (2012) and by Merlo and Ortalo-Magné (2004). Piazzesi, Schneider and Stroebel (2015) extend the housing market matching literature by allowing for multiple market segments and heterogeneous searchers.

---

[7] See Figure A1 in Appendix A for an exposition of the analysis described here.

7

Prior to our extending (or generalizing) the Wheaton (1990) model, we highlight the basics of the search and matching process with his model. Wheaton (1990) assumed that there are two types of households (e.g., families and singles) and two types of housing units (e.g., large and small). A given household is viewed as "matched" when a family is in a large unit and a single is in a small unit and "mismatched" when a family is in a small unit or a single is in a large unit. A matched household becomes mismatched when a single becomes a family, or a family becomes a single person.[8] A household moves from mismatched to matched by finding and purchasing the other appropriate unit. Then the previously occupied house is put up for sale. An additional simplification in Wheaton's model is that while households can change between types, the aggregate distribution of households by type is stable.

At any time, a given household is in one of three occupancy categories: 1) HM_i indicates a matched household (i.e., the household occupies an appropriate housing unit); 2) HD_i indicates a matched household with a house to sell (i.e., the household owns two units - one being appropriate and not for sale and the other inappropriate for them and for sale); and 3) HS_i indicates a mismatched household looking for an appropriate unit. With a fixed number of households and units of each type, households change categories in the model according to the following differential equations:

(1) $\dot{HS}_i = -m_i HS_i - \beta_i HS_i + \beta_j HM_j$

(2) $\dot{HD}_i = -q_i HD_i + m_i HS_i + \beta_j HD_j - \beta_i HD_i$

(3) $\dot{HM}_i = -\dot{HS}_i - \dot{HD}_i$, i = 1,2, j ≠ i.

β_i is the transition rate from preferring type i to preferring type j. Wheaton (1990) defines a mismatched household as one that cannot find an appropriate house instantaneously, so this leads to a search process that produces matches. Wheaton (1990) assumes the arrival of matches in the entire market occurs with a Poisson process, with m_i the match rate. Sales of vacant houses also are assumed to occur with a Poisson process, q_i. The sales of vacant houses equal the flow of house purchases, so that:

(4) $q_i V_i = m_i HS_i$

Equation (1) is the time rate of change of becoming mismatched. The first term is those who have become newly matched thus reducing the rate of becoming mismatched, the second term is those exiting from i (mismatched and looking) to j, and the third term captures the newly

---

[8] Obviously, this is a major simplification. Households tend to move when a job change creates a large increase in commuting distance or the household experiences some other change in income or family size that makes a current house inadequate. Ihrke (2014) found that with-in county moves are associated with housing-related issues, while between-county moves are associated with job-related issues.

mismatched who are moving from j to i. Equation (2) is the time rate of change of those households being matched with a house to sell. The first term captures the sale of vacant houses of type i, thus reducing the rate of being matched with a house to sell. The second term captures those households who have become matched and now have a house to sell. The third term captures households who are newly matched with a type i house to sell. The fourth term captures those households who have changed from i into j. Equation (3) is time rate of change of those becoming matched, which is simply the difference between the negative of those households becoming mismatched and looking and those who have become matched with a house to sell.[9]

An Extended Search and Matching Model

We extend and adapt Wheaton's (1990) model to allow for multiple regions and migration between these regions.[10] One change is to incorporate an interactive aspect in which those attempting to buy and sell vacant homes of a given type (i.e., in one region) take the actions of all others (i.e., in other regions) as given. To accomplish this we modify the previous definitions to focus on two regions (or one region, i, and the rest of the country, j), rather than on households in one region. A second adaptation is allowing for multiple vacancy regimes, denoted by subscripts on the model structural parameters.

---

[9] Wheaton (1990) further simplifies the analysis by assuming that the two types of households are identical in number and behavior. With $\beta_1 = \beta_2$, $V_1 = V_2$, $H_1 = H_2$, $m_1 = m_2$, then $HS_1 = HS_2$, $HM_1 = HM_2$, and $HD_1 = HD_2$. The effect of this simplification is to reduce the system of six differential equations to the following two equations:

(i) $\dot{HS}_i = -HS(2\beta + m) + \beta H - \beta HD$

(ii) $\dot{HD}_i = mHS(1 - HD/V)$

Equations (9) and (10) allow the determination of HS and HD. The resulting steady state is characterized by:

(iii) $HS = (\beta(H - V))/(2\beta + m)$

and

(iv) $HD = V$.

Equation (iii) captures mismatched households looking to buy a house, while equation (iv) captures matched households who have a (vacant) house to sell. In our adaptation, we retain the generalization that there can be two types, i and j, in our model and analysis.

[10] For an alternative housing market model using a search-and-matching approach, see Lisi (2015). This model, similar to our model, highlights the existence of vacancies; however, in contrast to our model, it does not address interregional effects.

We also consider a different notion of "mismatch" than in Wheaton (1990). Rather than mismatches due to owning a house that is inappropriately sized for a given household (i.e., a single individual owns a large house or a family owns a small house), in our context mismatches are associated with a desire to migrate to one region, but owning a home in the region of origin. Thus, a geographic mismatch underlies the housing mismatch. As a result, $HS_{i,t}, HD_{i,t}$, and $HM_{i,t}$ are respectively those who live in region j and want to migrate to region i at a point in time; those who happily live in region i at a point in time, but still have a vacant home in region j; and those who happily live in region i at a point in time. Note that in other parts of this paper, we refer to HS_j, which denotes people wanting to move out of region i and into region j; and similarly, HM_j as individuals living happily in region j who do not desire to migrate to region i. Also, $q_{s,i}$ is the probability of sale for a vacant region i house in regime s, $m_{s,i}$ is the probability of a mismatched household finding a region i home to purchase in regime s, and $\beta_{s,i}$ is the transition rate from preferring region i to preferring region j in regime s. Given the focus on "mismatched" households, together with the desire of these people to become matched, the mechanism by which vacancies in a particular region are determined is through migration between regions. This is demonstrated in more detail below.

We have added time subscripts (denoted as t) to our variables and regime (or state) subscripts (denoted as s) to the parameters in equations (3) and (4). We additionally assume that all homes are owned by the market participants we explicitly consider in the model, which leads to $V_{i,t} = HD_{j,t}$ because the only vacant homes in region i would be owned by those who have migrated to region j but still own a home in region i. The additions and assumption yield:

(3') $\left(HM_{i,t} - HM_{i,t-1}\right) = -\left(HS_{i,t} - HS_{i,t-1}\right) - (V_{i,t} - V_{i,t-1})$

(4') $q_{s,i}V_{i,t} = m_{s,i}HS_{i,t}$

Considering the activities of buyers and sellers of vacant homes in region i, we use equations (1) and (2) from above (i.e., Wheaton (1990)), slightly adjusted by adding time (t) and regime or state (s) subscripts and acknowledging that $m_{s,i} HS_{i,t} = q_{s,i} V_{i,t} = q_{s,i} HD_{j,t}$:[11]

(5) $\dot{HS_i} \sim HS_{i,t} - HS_{i,t-1} = -q_{s,i} HD_{j,t} - \beta_{s,i} HS_{i,t} + \beta_{s,j} HM_{j,t}$

(6) $\dot{V_i} \sim V_{i,t} - V_{i,t-1} = -m_{s,i} HS_{i,t} + m_{s,j} HS_{j,t} + \beta_{s,i} HD_{i,t} - \beta_{s,j} HD_{j,t}$.

---

[11] Wheaton (1990) drops the subscripts to indicate that the economy (i.e., one region) is in steady state. We are retaining the subscripts here to allow for multiple regions, each of which is taking the other regions' variables as exogenous. Without keeping these subscripts, it would be as if we were saying there was one large homogeneous region in the country, which is not consistent with the regional problem we are studying. In other words, these subscripts are a driver for our hypothesis that vacancies differ and are interdependent across regions.

These equations are different from equations (1) and (2) above (i.e., as in Wheaton (1990)) in the following ways. First, we assume a non-steady state scenario. Second, $-q_{s,i}\,HD_{j,t}$ (the number of homeowners who migrated to region j but still own a home in a region, i, in a given regime, s, at a particular time, t) is on the right side of (5) instead of the scenario in equation (1) above where $-m_i\,HS_i$ is on the right side. Equation (6) is different from Wheaton (1990) (equation (2) above) since we are considering actions of housing market participants in region i, and $HD_{j,t}$ households (i.e., $V_{i,t}$) are the sellers in region i; we also have $-m_{s,i}\,HS_{i,t}$ on the right side of (6), as opposed to having $-q_{s,i}\,HD_{j,t}$ in equation (2) above (the analogous equation of Wheaton (1990)).

Solving (5) for $HS_{i,t-1}$, substituting this equation for $HS_{i,t-1}$ into (6), along with using the fact that $m_{s,i}\,HS_{i,t} = q_{s,i}\,V_{i,t} = q_{s,i}\,HD_{j,t}$, and then solving for $V_{i,t}$, yields:

$$(7) \quad V_{i,t} = \left[\frac{\beta_{s,i} - 1 + \beta_{s,j}(\beta_{s,i} - 1) + m_{s,i}q_{s,i}}{(\beta_{s,i} - q_{s,i} - 1)}\right]V_{i,t-1} + \left[\frac{m_{s,j}(\beta_{s,i} - 1)}{(\beta_{s,i} - q_{s,i} - 1)}\right]HS_{j,t-1}$$
$$+ \left[\frac{\beta_{s,i}(\beta_{s,i} - 1)}{(\beta_{s,i} - q_{s,i} - 1)}\right]V_{j,t-1} + \left[\frac{-m_{s,i}\beta_{s,j}}{(\beta_{s,i} - q_{s,i} - 1)}\right]HM_{j,t-1}$$

Because the parameters in the brackets are transition rates, they are each expected to vary between 0 and 1. Given this, the anticipated signs of the terms in brackets implied by the theory are as follows. The term preceding $V_{i,t-1}$ is indeterminate, while the terms preceding $V_{j,t-1}$, $HM_{j,t-1}$, and $HS_{j,t-1}$ are expected to be positive.[12] We take this equation to the data, generalized to incorporate more than one j region, and estimate VAR and MS-VAR models to examine how well our theory holds up empirically. Our empirical models fit our theoretical model well by allowing us to simultaneously estimate equation (7) for multiple regions, with the designation of region i and regions j alternating across equations. To completely internalize the empirical dynamics, we add separate equations for $HM_{j,t}$, $HS_{j,t}$, and $V_{j,t}$, each as a function of ($V_{i,t-1}$, $HS_{j,t-1}$, $V_{j,t-1}$, $HM_{j,t-1}$).[13]

Underlying the parameters ($m_{s,i}, m_{s,j}, \beta_{s,j}, q_{s,i}, q_{s,j}, \beta_{s,i}$) are economic and other factors affecting the supply and demand of housing in each region. These parameters can vary across

---

[12] For each term in brackets, the denominators are identical and are a negative value. The numerator for the first term in brackets is indeterminant because (dropping subscripts) mq is positive, but the other terms sum to a negative number. Empirically, we find the first term to be positive. The other 3 numerators are negative, so each of the remaining terms in brackets is positive.

[13] Note that we only include the structural equation for the vacancy rate since our research question focuses on how this variable responds to migration flows (similarly, our empirical results for the MS-VAR only show the response of the vacancy rate to shocks in migration flows). The full set of structural equations is available from the authors upon request.

regimes. While there may be one, two, or more regimes, we allow the data to help us determine how many regimes there are for each region.

Assuming multiple regimes, the parameters $(m_{s,i}, m_{s,j}, \beta_{s,j}, q_{s,i}, q_{s,j}, \beta_{s,i})$ all have regime subscripts, s, where for the two-regime model, $s = \{1, 2\}$ (or equivalently, in a two-regime model for our application, $s = \{\text{low vacancy, high vacancy}\}$). A key point is that there is a continuous equation constituting $V_{i,t}$ for the high vacancy regime, and a continuous equation constituting $V_{i,t}$ in a low vacancy regime. In other words, there are effectively jump discontinuities at the times of regime switches, with a binary regime indicator, but $V_{i,t}$ is continuous throughout a particular regime until the next switch. As a concrete example, suppose region i is in a low-vacancy regime (i.e., $s = 1$) from time t to time $t + 10$, before a switch to a high vacancy regime (i.e., $s = 2$) at time $t + 11$. From time t to time $t + 10$, $V_{i,t}$ will be characterized by a continuous, linear equation between it and $V_{j,t}$. Then, from $t + 11$ until the next regime switch, there will be a separate, regime 2-related continuous, linear equation describing the relationship between $V_{i,t}$ and $V_{j,t}$.

It is our hypothesis that major shifts in the U.S. economy, and thus the national housing market, also cause shifts in the migration of households across regional housing markets. For that reason, we attempt to fit switching models to the data to capture any shifts in the relationship between regional housing markets. Shocks are not restricted to the national economy as regional shocks are possible as well.

The relationships between $V_{i,t}$ and $HS_{j,t}$, $V_{j,t}$, and $HM_{j,t}$ can vary in different vacancy regimes, due to the regime subscripts underlying the parameters $(m_{s,i}, m_{s,j}, \beta_{s,j}, q_{s,i}, q_{s,j}, \beta_{s,i})$. We aim to determine the sign and magnitude of these relationships, which can be different in each vacancy regime. We accomplish this by estimating multiple regime, reduced form MS-VAR models.

A Markov-switching Model for Vacancy Rates

While we initially estimate constant-only and own-lag-only Markov-switching models to motivate the usage of regime switching, our model of interest is focused on the four U.S. Census regions. It is a reduced-form VAR model centered on a four-region generalization of equation (7). The preferred parametrization of this model would involve twelve equations, one for each region's V, HS, and HM variables as functions of the eleven other V, HS, and HM variables. This would allow a complete internalization of the interregional dynamics implied by equation (7). To illustrate, the equation for the South region's vacancy rate would be the following:

(8) $V_{S,t} = C_{s,t} + a_{s,t} V_S^k + b_{s,t} HS_S^k + c_{s,t} HM_S^k + d_{s,t} V_M^k + e_{s,t} V_N^k + f_{s,t} V_W^k + g_{s,t} HS_M^k + h_{s,t} HS_N^k + i_{s,t} HS_W^k + j_{s,t} HM_M^k + k_{s,t} HM_N^k + l_{s,t} HM_W^k + \varepsilon_{S,t},$

where subscripts S, M, N, and W indicate South, Midwest, Northeast, and West, respectively. For compact notation, we define V_S^k and $a_{s,t}$ as the 1st through kth lags of $V_{S,t}$, and the corresponding coefficient parameter vector of length k. The other variable and parameter vectors are defined similarly. While equation (7) implies that only one lag be included, we allow for the possibility of $k > 1$, as dictated by estimation results, to deal with the possibility of housing market frictions resulting in additional lags being necessary.

Unfortunately, data limitations do not provide sufficient degrees of freedom to estimate all parameters. We are able to estimate models involving individual region vacancy measures, but with HS and HM aggregated to three regions at a time. We use the three-region-aggregated measures to estimate four different models, each producing an equation as close as possible to the preferred generalization of equation (7). As an example, the model oriented toward the South is parametrized as follows:

$$(9a)\ V_{S,t} = C_{s,t} + a_{s,t}\ V_S^k + b_{s,t}\ HS_{MNW}^k + c_{s,t}\ V_M^k + d_{s,t}\ V_N^k + e_{s,t}\ V_W^k + f_{s,t} HM_{MNW}^k + \varepsilon_{S,t}$$

$$(9b)\ V_{M,t} = C_{s,t} + g_{s,t}\ V_S^k + h_{s,t}\ HS_{MNW}^k + i_{s,t}\ V_M^k + j_{s,t}\ V_N^k + k_{s,t}\ V_W^k + l_{s,t} HM_{MNW}^k + \varepsilon_{M,t}$$

$$(9c)\ V_{N,t} = C_{s,t} + m_{s,t}\ V_S^k + n_{s,t}\ HS_{MNW}^k + o_{s,t}\ V_M^k + p_{s,t}\ V_N^k + q_{s,t}\ V_W^k + r_{s,t} HM_{MNW}^k + \varepsilon_{N,t}$$

$$(9d)\ V_{W,t} = C_{s,t} + s_{s,t}\ V_S^k + t_{s,t}\ HS_{MNW}^k + u_{s,t}\ V_M^k + v_{s,t}\ V_N^k + w_{s,t}\ V_W^k + x_{s,t} HM_{MNW}^k + \varepsilon_{W,t}$$

$$(9e) HS_{MNW,t} = C_{s,t} + y_{s,t}\ V_S^k + z_{s,t}\ HS_{MNW}^k + \alpha_{s,t}\ V_M^k + \beta_{s,t}\ V_N^k + \gamma_{s,t}\ V_W^k + \delta_{s,t} HM_{MNW}^k + \varepsilon_{HS,t}$$

$$(9f) HM_{MNW,t} = C_{s,t} + \vartheta_{s,t}\ V_S^k + \mu_{s,t}\ HS_{MNW}^k + \pi_{s,t}\ V_M^k + \rho_{s,t}\ V_N^k + \tau_{s,t}\ V_W^k + \varphi_{s,t} HM_{MNW}^k + \varepsilon_{HM,t},$$

where the subscript MNW represents the aggregated region encompassing the Midwest, Northeast, and West, or the portion of the U.S. excluding the South region. Variants of equations (9a) – (9f) oriented toward another region would thus incorporate HS and HM values for the portion of the U.S. excluding that region. Due to data limitations, equation (9a) is as close as we can get to the fully generalized equation for $V_{S,t}$, (equation (8)). In correspondence with equation (7), with the South region as region i, HS_{MNW} and HM_{MNW} are the appropriate aggregated HS_j and HM_j. While each multiple-equation model estimated will only have one equation of significant interest to us, it is important to note that central to the estimation of any switching model is the estimation of vectors of probabilities for being in a given regime across the sample period. These regime probability vectors would likely be very different if we attempted to estimate only each model's equation of interest in a single-equation model, given that key model dynamics would not be internal to the estimation.

In terms of our theoretical model in (7), the parameters in (9a), in a one-lag model, correspond to those in (7) as follows:

$$a_{s,t} \approx \left[\frac{\beta_{s,S} - 1 + \beta_{s,MNW}(\beta_{s,S} - 1) + m_{s,S}q_{s,S}}{(\beta_{s,S} - q_{s,S} - 1)} \right]$$

$$b_{s,t} \approx \left[\frac{m_{s,MNW}(\beta_{s,S} - 1)}{(\beta_{s,S} - q_{s,S} - 1)} \right]$$

$$c_{s,t} \approx \left[\frac{\beta_{s,S_M}(\beta_{s,S_M} - 1)}{(\beta_{s,S_M} - q_{s,S_M} - 1)} \right]$$

$$d_{s,t} \approx \left[\frac{\beta_{s,S_N}(\beta_{s,S_N} - 1)}{(\beta_{s,S_N} - q_{s,S_N} - 1)} \right]$$

$$e_{s,t} \approx \left[\frac{\beta_{s,S_W}(\beta_{s,S_W} - 1)}{(\beta_{s,S_W} - q_{s,S_W} - 1)} \right]$$

$$f_{s,t} \approx \left[\frac{-m_{s,S}\beta_{s,MNW}}{(\beta_{s,S} - q_{s,S} - 1)} \right],$$

where, as examples of notation, $\beta_{s,S}$ is the transition rate from preferring the South to preferring the rest of the country, and β_{s,S_M} is the transition rate from preferring the South to preferring the Midwest. Based on our previous discussion, the sign of $a_{s,t}$ is indeterminate, while the signs of $b_{s,t}$, $c_{s,t}$, $d_{s,t}$, $e_{s,t}$ and $f_{s,t}$ are positive.

Data and Estimation Procedure

We will first consider variations to the model outlined in equations (9a) – (9f) which do not consider regional spillovers, but rather looks at the univariate time-series dynamics of vacancies at the national level. To estimate this national model, we only need data on national vacancies. We opt to use the vacancy rate rather than the raw number of vacancies due to the measure's tractability and common usage as a housing market indicator. The vacancy rate data comes from the U.S. Census Bureau and is at a quarterly frequency from 1956:Q1 through 2019:Q3. Note that we use the homeowner vacancy rate, which is the proportion of homeowner inventory that is vacant for sale. To eliminate seasonal trends, we seasonally adjust the vacancy rate data using the standard X-13 ARIMA method publicly available from the Census.

To estimate the regional variants of the model outlined in equations (9a) – (9f), we need regional data on: (i) vacancies, $V_{i,t}$, (ii) households who would like to migrate out of region i,

$HS_{j,t}$, and (iii) households who would not like to migrate into region i, $HM_{j,t}$. Due to data limitations at the MSA level, as well as the large numbers of parameters that would need to be estimated at the MSA level, we choose to focus on the four regions of the United States outlined by the Census: South, Northeast, Midwest, and West. As with the national data, we use the homeowner vacancy rate for each region in place of the actual number of vacancies. This regional vacancy rate data is from the Census for 1956:Q1 through 2019:Q3; however, we must again seasonally-adjust the data using the standard X-13 method.

Direct measures of the migration variables $HS_{j,t}$ and $HM_{j,t}$ are not readily available and must be constructed. Recall that $HS_{j,t}$ measures the number of home-owning households (unhappily) living outside of region j that would like to migrate to region j at time period t. To begin constructing $HS_{j,t}$, we use the number of individuals that actually leave region i at time $t + 1$, MO_{t+1}, to capture those people *desiring* to leave region i at time t. To generate the number of households that migrated out, we divide this measure of outmigration in number of individuals by the average number of individuals per household at time t, $F_{i,t}$. Finally, we multiply this measure of outmigration in number of households by the homeownership rate in region i at time t, $H_{i,t}$, to get the number of home-owning households that actually migrated out at $t + 1$. Explicitly, we use this measure to proxy the number of households wanting to leave region i at time t by:

$$HS_{j,t} = \frac{MO_{i,t+1}}{F_{i,t}} \times H_{i,t} \; .$$

Note that the last part of this calculation implicitly assumes that preferences for homeownership are reflected in the homeownership rate in the region the household left, which is necessary since measures of the percentage of migrating households who are homeowners are not directly available.

Regional data on outmigration (MO_i), average household size (F_i), and the homeownership rate (H_i) are all available from the Census. Some of the data is only available at a yearly frequency. To retain information in the quarterly data, we interpolate any annual data evenly across its respective four quarters. Since $HS_{j,t}$ is a count variable rather than a rate, we detrend the data in addition to seasonally adjusting it. We also normalize $HS_{j,t}$ to be mean zero and unit variance to ease interpretation of the impulse response functions outputted from the MS-VAR.

To measure $HM_{j,t}$, we need to capture the number of households who are happily living in region j at time t and do not desire to migrate out. We begin with the total number of households in region j who owned a home at time t, $HN_{j,t}$. We then subtract from $HN_{j,t}$ the number of households living in region j with a vacant home in region i ($HD_{j,t}$) and the number of households living in region j who want to live in region i ($HS_{i,t}$). To proxy for $HD_{j,t}$, we use

the outmigration of individuals from region i at time t ($MO_{i,t}$) divided by average household size in region i at time $t-1$ ($F_{i,t-1}$) and multiplied by the homeownership rate in region i at time t, ($H_{i,t-1}$). Intuitively, this proxy measures the number of home-owning households who migrated into region j the previous period and thus are relatively more likely to have a vacant home in region i. Since $HS_{i,t}$ captures households across three regions (i.e., households outside region i who wish to move to region i), we need to adapt the proxy outlined in the previous paragraph used for $HS_{j,t}$. For similar reasons as before, we use the number of individuals that migrated into region i at time $t+1$, $MI_{i,t+1}$ to capture those individuals that wished to leave region j at time t. We convert this to number of households by dividing $MI_{i,t+1}$ by $F_{j,t}$, which is the owner-occupied-household-weighted average number of individuals per household across the other three Census regions (i.e., all other regions besides region i). Finally, we multiply this number of outmigration households by $H_{j,t}$, the owner-occupied-household-weighted average homeownership rate across the three Census regions other than region i. Formally, the constructed proxy for the number of households happily living in region j at time t is given by:

$$HM_{j,t} = HN_{j,t} - \frac{MO_{i,t}}{F_{i,t-1}} \times H_{i,t-1} - \frac{MI_{i,t+1}}{F_{j,t}} \times H_{j,t}.$$

Similar to the data for $HS_{j,t}$, the underlying data for $HM_{j,t}$ is available from the Census and we detrend, seasonally-adjust, and normalize to obtain the finalized proxy.

We estimate each model using Bayesian methods. In particular, we approximate each model's posterior distribution via the Markov-chain Monte Carlo (MCMC) technique of Gibbs sampling which draws each set of parameters' conditional distribution given the other sets. This approach is a standard method for estimating highly parameterized nonlinear models as shown by Fruhwirth-Schnatter (2006) and Sims and Zha (2006). Appendix B outlines the technical details of the estimation.

Before moving to our empirical results, we should note that the seasonally-adjusted vacancy rate (both for the United States and for the individual regions) fails to reject the null of nonstationarity according to a standard Augmented Dickey-Fuller (ADF) test. Including a nonstationary series in our model could lead to spurious results and conclusions, especially as it relates to lags of the endogenous variable. However, Nelson et al. (2001) shows that standard ADF tests have low power when the data is truly generated from a Markov-switching process. Therefore, to check the vacancy rate for stationarity, we estimate a Markov-switching ADF (MS-ADF) model as in Hall et al. (1999), Kanas and Genius (2005), and Kanas (2006). After controlling for regime switching in the intercept, the MS-ADF model for each vacancy rate allows us to reject the null of nonstationarity at the 5% level.

Univariate Switching Regression Results

National Models

For illustrative purposes and for completeness, we begin with a special variant of equation (9a) for the U.S. as a whole (i.e., assuming there is only one region). In this section, we will consider a host of alternative univariate specifications for the U.S. vacancy rate. Before examining the estimation results for each alternative model, let's re-examine the raw data for the vacancy rate for the United States from 1956:Q1 through 2019:Q3. Figure 1 shows that the U.S. vacancy rate has normally been less than 2.0 percent. Only the period associated with the housing crisis exhibited rates in excess of 2.0 percent. Prior to the Great Recession, no period, even recessions, exhibited a national vacancy rate greater than 2.0 percent.

We consider four alternative Markov-switching specifications for the U.S. vacancy rate: (i) MS2 – a model with a two-regime Markov-switching intercept and variance, (ii) MS3 – a model similar to MS2 but with three regimes, (iii) MS2AR – a two-regime model with a switching intercept, variance, and autoregressive (AR) term on the first lag of the vacancy rate, and (iii) MS3AR – a model similar to MS2AR but with three regimes. By comparing the fit of these four models, we can deduce whether the vacancy rate tends to switch between two or three regimes and whether an autoregressive component is necessary to control for persistence. It is useful to consider these alternative specifications in the framework outlined by equation (9a). The simple MS2 and MS3 models corresponds to equation (9a) while imposing $a_{s,t}$, $b_{s,t}$, $c_{s,t}, d_{s,t}$, $e_{s,t}$ and $f_{s,t}$ to equal vectors of zeros. The MS2AR and MS3AR models are similarly structured, but allow the lagged vacancy rate in the one region (i.e., the U.S.) to affect the current vacancy rate. This is equivalent to assuming the parameters other than $a_{s,t}$ in (9a) are all set to zero.

Table 1 displays the estimation results for each univariate model of the U.S. vacancy rate. For each model, we show the average vacancy rate in each regime (along with its associated posterior standard deviation), the variance in each regime, the persistence probability and expected duration of each regime, and standard model fit statistics.[14] Throughout the paper, regime 1 is the low vacancy regime and regime 2 is the relatively higher vacancy regime. For the three-regime models, regime 1 has the lowest average vacancy rate, regime 3 has the highest, and regime 2 is between the two extremes.

The first column of Table 1 outlines the simplest model, MS2. Under this specification, the low-vacancy rate estimate for the United States is 1.43 percent, while the high-vacancy rate is 2.33 percent. Additionally, the 95% highest posterior density intervals for the intercepts do not overlap, allowing us to conclude that the regimes are well-identified and the average vacancy

---

[14] We consider three model fit statistics from the literature: Akaike Information Criteria (AIC), Bayesian Information Criterion (BIC), and Deviance Information Criterion (DIC). The first two are standard in both Frequentist and Bayesian approaches, while DIC is applicable when using Bayesian estimation methods.

rate is substantially higher in regime 2 relative to regime 1. Regime 1 can also be characterized as the relatively low volatility regime as the variance of the vacancy rate is considerably lower than under regime 2.[15]

As can be seen by the estimates of p_{11} and p_{21}, both regimes are highly persistent. If the United States is currently in regime 1, the low-vacancy regime, the probability of remaining in regime 1 is 99 percent. Similarly, if the United States is currently in regime 2, the high-vacancy regime, the probability of remaining in regime 2 is slightly lower at 97 percent. An alternative way to look at the transition probabilities is through the expected duration of a given regime. This calculation provides us with the answer to the intuitive question of, "how long will this regime last" (on average). The low vacancy regime, regime 1, has an average phase of 143 quarters, while the high vacancy regime lasts for 30 quarters on average.

Figure 2(a) shows the actual vacancy rates compared with the predicted vacancy rate from the MS2 model. Since the vacancy rate is only a function of the regime variable, the predicted vacancy rate changes only when there is an estimated change in regime. Figure 2(b) shows the posterior probability of when the United States was in either the low vacancy regime or high vacancy regime. These estimated probabilities suggest that the United States has been in a low vacancy regime for the entirety of the sample besides the Great Recession/Financial Crisis. Intuitively, the MS2 model wants to fit the data during this period since the vacancy rate is considerably higher than other periods in the time sample.

The regime timing results from the MS2 model beg the question if two regimes are sufficient to capture the switching dynamics of the aggregate vacancy rate. In the second column of Table 1, we present the posterior estimates for the MS3 model. Based on the standard model fit statistics, the MS3 model fits the data better than the MS2 model implying that three regimes are necessary to describe the U.S. vacancy rate. Figure 3(a) shows the actual and predicted vacancy rate based on the MS3 model and reinforces the better fit of the model relative to the MS2 model. Regime 1 has the lowest mean vacancy rate (1.16), regime 3 has the highest (2.47), while regime 2 is the intermediate case (1.64). Regime 3 also has a substantially higher variance than regime 1 or 2, implying it matches well with the high vacancy regime from the MS2 model as both are characterized by a higher degree of volatility. The expected duration of regime 1 and regime 2 are similar (31 quarters and 35 quarters, respectively), while the high vacancy regime has a considerably shorter duration (18 quarters).

Figure 3(b) displays the estimated timing of each regime across the entire sample. The United States is in regime 1, the low vacancy regime, during the late 1950s, the late 1960s through the 1970s, and in the very recent past (2019). Regime 2, the intermediate regime, dominated the early 1960s, the 1980s through 2005, and late 2010s. Similar to the high vacancy

---

[15] The finding of a regime with a higher mean and volatility of the vacancy rate is somewhat similar to the finding of Kim and Nelson (1999) as it relates to the business cycle. They find that during unstable times (i.e., recessions) the mean GDP growth rate is lower and the volatility is higher.

regime in the MS2 model, regime 3 from the MS3 model is only active during the extended Great Recession period. Note that, besides the Great Recession, there is no apparent comovement of any of the three regimes with the aggregate business cycle as reflected by the recession dating provided by the National Bureau of Economic Research (NBER).

The MS2 and MS3 models previously discussed do not consider that there may be an AR component to the vacancy rate. The third column of Table 3 outlines the estimation results for the MS2AR model that allows for regime switches, while also including a single AR term (i.e., the first lag of the U.S. vacancy rate US_{t-1}). The MS2AR model improves the in-sample fit compared to the MS3 model, and turns out to be the best description of the data for all four univariate models considered. The inclusion of the AR term is justified by the relatively high coefficient on the first lag of the vacancy rate across both regimes. Regime 1 is characterized by a relatively higher AR coefficient than regime 2, but we do not find a substantial difference across regimes for this parameter as their posterior coverage intervals overlap. Despite the estimated difference in the constant parameter across regimes, the two regimes have similar steady-state vacancy rates. However, regime 2 does have a notably higher variance than regime 1.

Figure 4(a) shows the predicted U.S. vacancy rate from the MS2AR model and the improved fit of the actual vacancy rate is stark. The estimated regime probabilities in Figure 4(b) however show that the regimes are relatively less identified compared to the MS2 or MS3 models. This result implies that perhaps the MS3 model is better for regime identification, whereas the MS2AR model is preferred for overall fit.

The last column of Table 1 and Figure 5 present the estimation results for the MS3AR model. This model does not outperform the less sophisticated MS2AR by the standard fit statistics; however, it does better than either the MS2 or MS3 models and reinforces the importance of including the lagged term. Similar to the MS2AR model, the estimated regime timing from the MS3AR model are less intuitive compared to those from the MS3 model. Therefore, we conclude that the U.S. vacancy rate is overall best described by a model with two regimes and an AR component but the simpler MS3 model offers useful insights into regime timing.

Region-Specific Models

This section considers the same four univariate regime-switching models from the previous section, but applies them to the four U.S. Census regions: South, Northeast, Midwest, and West. Tables 2 through 5 display the parameter estimates for each model applied to the four regions. Generally, the results for the regional vacancy rates qualitatively mimic the results for the aggregate vacancy rate. The model with two regimes and an AR component (the MS2AR model) fits the data the best for all four regions. The MS3AR is second best for each region, but the regime timing from the MS3 model is more intuitive, in our opinion. Thus, we focus our

discussion in this section on the MS3 model specification for ease of interpretation. The dominance of the MS2AR model motivates the MS-VAR models considered in the next section.

The second column of Table 2 contains the parameter estimates for the MS3 model applied to the South vacancy rate. Compared to the U.S. vacancy regimes, the South regimes have higher average vacancy rates. In addition, the South experiences a higher degree of volatility as the variance parameter is larger across all three regimes compared to the U.S. The South vacancy rate tends to switch marginally more often between regimes, as indicated by the shorter expected duration across regimes. Figure 6(a) shows that the estimated regime timing matches quite well with the estimated regimes for the United States. The only marked difference is the South's transient shift to the low vacancy regime in the early 1990s, whereas the United States vacancy rate stays in the intermediate regime.

In the second column of Table 3, we provide the MS3 parameter estimates for the Northeast vacancy rate. In contrast to the results for the South, the Northeast has a lower mean vacancy rate across all three regimes compared to the aggregate vacancy rate. Additionally, the low and high regimes for the Northeast have relatively lower variance than the corresponding U.S. regimes; the intermediate regimes have similar volatilities than the corresponding U.S. regimes, though the Northeast has a higher intermediate regime volatility. The expected duration of each regime for the Northeast is longer than the durations for the U.S. regimes. This result is reinforced by looking at the estimated regime timing for the Northeast in Figure 6(b). The Northeast vacancy rate experiences only three switches. Starting in the low regime, the vacancy rate switches to the intermediate regime in the mid-1980s. Similar to the aggregate vacancy rate, it switches to the high regime just prior to the Great Recession. The Northeast vacancy rate stays in the high regime relatively longer and only recently switches back to the intermediate regime. The Northeast is the only vacancy rate that is not currently estimated to be in the low regime based on the MS3 model. However, there has been a recent uptick in the probability of the low regime that could continue into the future.

The MS3 estimation results for the Midwest vacancy rate are given in the second column of Table 4. The mean vacancy rate and variance parameters across regimes are similar to those for the entire United States. The notable exception is that the intermediate regime, regime 2, for the Midwest has a moderately higher volatility (0.25) than the United States as a whole (0.19). The expected duration for the low regime is considerably higher for the Midwest (45 quarters) compared to U.S. aggregate low regime (31 quarters). Conversely, the expected duration for the Midwest's intermediate regime is substantially lower (17 quarters) to that of the nation's (35 quarters). The estimated timing of regimes for the Midwest is presented in Figure 6(c). In general, these regimes match the timing for the U.S. regimes with some minor differences. The Midwest does not have a short-lived intermediate regime in the early 1960s as the nation's vacancy rate does. However, the Midwest vacancy rate does switch back to a low regime in the late 1980s until 2000. Similar to the U.S. vacancy rate, the Great Recession period is the only

high vacancy rate period for the Midwest, and in the wake switches to an intermediate regime and most recently the onset of a low regime.

Finally, the second column of Table 5 provides the parameter estimates for the MS3 model applied to the West vacancy rate. Similar to the Midwest region, the West vacancy rate parameters largely match the nations with a markedly high variance for regime 2. The West experiences a higher number of switches across the time sample relative to the nation as a whole, as indicated by the lower expected duration of each regime. In terms of regime timing, Figure 6(d) shows that the West's vacancy rate follows a similar pattern to the overall United States. However, the intermediate regime in the early 1960s is identified to a lesser extent in the West compared to the nation as it places a considerable probability on being in either the intermediate or high regime. A similar phenomenon occurs in the mid-2000s as the model struggles to identify whether the West was in an intermediate or low regime prior to switching to the high regime of the Great Recession. Compared to the other regions and the United States, the West's vacancy rate has been in the current low regime for considerably longer since the estimated start date is approximately 2015.

These univariate regional models suggest that, although there is some heterogeneity across regional vacancy rates, there are notable similarities in regime timing, mean vacancy rates, and volatility across regions. These findings suggest it would be worthwhile to investigate if there are spillovers across regions similar to the theoretical structure outlined in equations (9a) – (9f). We investigate this idea in a multivariate setting in the next section.

Multivariate Switching Regression Results

The results from the previous section suggests that the vacancy rate (both at the national and regional levels) is best described by a Markov-switching process with two regimes and an autoregressive component. Additionally, there appears to be some broad similarities in vacancy dynamics across regions and at the national level. These findings motivate our choice to analyze the interregional effects of migration in a two-regime Markov-switching Vector Autoregression (MS-VAR) as in as in equations (9a), (9b), (9c), (9d), (9e) and (9f). This model setup allows for similar dynamics as in the univariate MS2AR but also captures spillover effects across the regions. The estimation details for the MS-VAR are presented in Appendix B. Note that due to data limitations for some of the migration variables, the data for the MS-VAR estimation is limited to the slightly shorter time frame of 1964:Q1 through 2017:Q2.

Figures 7-10 show the predicted vacancy rate for each region and their respective timing of low and high regimes. For each region, the predicted vacancy rates follow the actual vacancy rate closely across time illustrating the benefit of adding data in the cross-section and potential spillovers of housing markets across regions. With regards to regime timing, the South region experiences the high regime throughout the entirety of the sample besides the late 1970s.

Similarly, the Northeast region mostly experiences a high vacancy rate regime besides the early 2000s and some one-quarter blips of a low vacancy regime. The Midwest vacancy rate is in the high regime for the majority of the sample as well, with exceptions for the low regimes in the late 1970s, mid to late 1980s, and late 1990s. Lastly, the West region is in a high vacancy rate regime for the entire time series besides one-quarter low regimes throughout the sample.

The regime timing for the MS-VAR for each region differs substantially from the regime timing presented in the previous section for a number of reasons. First, the previous section considered a model with three regimes whereas the MS-VAR only has two regimes. It is possible to add a third regime, but this often over-fits the data (as with the MS3AR). Second, the MS-VAR allows all variables in the model to switch thus the switching process is trying to best fit the dynamics of a number of variables in the system rather than just a single region's vacancy rate.

Impact of Interregional Migration Effects

The multivariate setup of the MS-VAR allows us to determine the effect of one variable on another in the system. Presenting the estimated parameters of the entire system is often cumbersome and unintuitive. Instead, these dynamic effects are often illustrated with impulse response functions (IRF) that show the response of each variable in the system to a (structural) shock. For a standard VAR model, computation of the orthogonalized impulse response function is relatively straightforward. However, in a nonlinear framework such as our MS-VAR, it is necessary to calculate generalized impulse response functions (GIRF) as first considered by Koop et al. (1996) to capture the regime switching dynamics in the wake of a given shock. Intuitively, the GIRF captures the fact that the system's dynamics are not constant but switch between two discrete regimes across time (based on the estimated Markov process). Some previous studies such as Ehrmann et al. (2003) estimate "regime-dependent" IRFs that assume that the regime does not change after the shock. These regime-dependent IRFs may under or overestimate the dynamic effects of a shock because they fail to account for the changing dynamics of the system. For estimation details of GIRFs, see Koop et al. (1996) and Alessandri and Mumtaz (2019).

In this section, we focus our attention on the two migration variables, HS_j and HM_j, and how they affect the region-specific vacancy rates. Figure 11 displays the estimated dynamic response of each of the four regional vacancy rates to a shock to one of the two migration variables. For example, the first panel (i.e., top left) of Figure 11 shows the response of the South vacancy rate across future horizons $V_{S,t+h}$ to a one standard deviation shock to HS_j (the number of people who are unhappily living in the South and would like to migrate out) at time $t = 0$. The second panel (i.e., top right) of Figure 11 shows the response of the South vacancy rate $V_{S,t+h}$ to a one standard deviation shock to HM_j (the number of people who are happily living outside the South and do not prefer to migrate in) at time $t = 0$. The dashed lines show

22

the 68% highest posterior density interval, as is standard from Sims and Zha (2006). When this interval does not include 0, we conclude that the effect is substantially different from 0.

We find that an increase in the number of households wanting to migrate out of the South (i.e., a shock to HS_j) does not substantially affect the South vacancy rate. Intuitively, an increase in the number of households that are happily living outside the South (i.e., an increase in HM_j) leads to an increase in the South vacancy rate for an extended period of time.

The second row of Figure 11 presents the response of the Northeast vacancy rate to the two migration variables. The Northeast vacancy rate responds negatively to an increase in the number of households wanting to migrate out of the Northeast, whereas the vacancy rate responds positively to an increase in the number of households happily living outside the Northeast.

We do find a significant positive response of the Midwest vacancy rate to an increase in HS_j (the number of households living unhappily in the Midwest and want to migrate out). Similar to the South and Northeast vacancy rates, an increase in households happily living outside the Midwest causes the Midwest vacancy rate to rise.

A shock to the number of households inside the West wanting to migrate out of the region does not significantly change the vacancy rate. Interestingly, an increase in the number of households happily living outside the West causes an initial decrease in the West vacancy rate in the short-run. After a year and half, this effect turns positive and the result matches that of the other three regions.

In sum, there are two key takeaways from the estimated effects of the migration variables on regional vacancy rates. First, an increase in the number of households wishing to migrate into a given region does not have a consistent effect on that region's vacancy rate. Theory suggests the effect of HS_j (the number of households wishing to migrate out of region i) on region i's vacancy rate should be positive. However, our empirical results match the theoretical implication for only the Midwest for a short period of time, whereas for two of the other three regions we find no substantial effect of HS_j, and a significant negative effect for the Northeast. Perhaps an omitted factor, such as the state of the economy or long-term structural changes, may be driving a friction between the expected response of the vacancy rate to an increase in people wanting to leave a given region. Finally, an increase in the number of households happily living outside region i (HM_j) tends to increase region i's vacancy rate with the exception of the West, which experiences a short-run drop (but ultimately an increase in the medium- and longer-run). Overall, while HM_j is quite consistent in terms of its effects, HS_j seems not to be.

State-Dependent Effects of Interregional Migration

   The previous section analyzed the estimated effects of the migration variables on vacancy rates across regions. These estimated effects were calculated based on all time periods, both low and high vacancy regime periods. The primary difference between a standard VAR and the MS-VAR considered in this section is that the system's dynamics change across time. One implication of this assumption is that the MS-VAR has the ability to capture if the effect of one variable on another differs across regimes. For example, Auerbach and Gorodnichenko (2012) use a regime-switching VAR to find that fiscal multipliers on output are larger during recessions than expansions. In our application, the MS-VAR allows us to answer the interesting question of whether the effects of migration differ between relatively low- and high-vacancy states.

   Figure 12 shows the GIRFs of the regional vacancy rates to the two migration variables conditional on the initial regime. For example, the blue line in the first panel of Figure 12 shows the response of the South vacancy rate to a one standard deviation shock to HS_j (number of households wanting to migrate out of the South) at time $t = 0$ in the relatively low vacancy regime. In the same panel, the red line shows the response of the South vacancy rate to the same shock but if the regime at time $t = 0$ is the high vacancy regime. Because these estimated responses depend upon the initial regime, these GIRFs are described as *state-dependent*. Across both regimes, the South's vacancy rate is not significantly changed by an increase in the number of households unhappily living in the South wanting to migrate out. However, the vacancy rate in the South does respond asymmetrically across regimes to an increase in the number of households living outside the South wanting to migrate in. In the high-vacancy regime, the South's vacancy rate responds as expected: increasing in response to a shock to HM_j. But in the low vacancy regime, the same shock does not significantly alter the vacancy rate though the point estimate of the response is positive.

   The second row of Figure 12 shows the state-dependent responses of the Northeast vacancy rate to the migration shocks. The results for both regimes match the generalized responses for the Northeast presented in figure 11. An increase in households wanting to leave the Northeast drives the Northeast vacancy rate significantly down across both regimes. An increase in external households wanting to move-in drives the vacancy rate up in both regimes. Despite marginal differences in size of the effect, the shape and significance of the state-dependent responses are consistent across regimes.

   In the third row of Figure 12, we present the response of the Midwest vacancy rate to a shock to the migration variables. When the number of households wanting to leave the Midwest increases, the vacancy rate in the Midwest has the expected positive response in the high vacancy regime. However, in the low-vacancy regime, the point estimate of the response is estimated to be negative, but this effect is not significantly different from zero. In response to an increase in the number of households wanting to migrate into the Midwest, the vacancy rate eventually

increases across both regimes but there is evidence of short-run asymmetry across regimes as the vacancy rate initially decreases in the low-vacancy regime.

Finally, the fourth row of Figure 12 shows the response of the West vacancy rate to a shock to HS_j and HM_j. Across both shocks, the response of the vacancy rate is virtually identical conditional on being in either of the two regimes. This result is most likely explained by the fact that the low vacancy periods are relatively short-lived (i.e., one-quarter) based on the estimated timing in Figure 10(b). Therefore, the initial regime is relatively uninformative for how a given shock affects the vacancy rate; the model tends to switch to (or stay in) the high vacancy regime quickly so the dynamic effects are indistinguishable.

Altogether, the state-dependent effects show that two regions (the South and Midwest) have asymmetric responses of their respective vacancy rate to migration shocks. In both cases, the vacancy rate tends to be more responsive when the vacancy rate is in its high-mean/high-volatility regime. Conversely, the other two regions (the Northeast and the West) have similar responses across regimes, and therefore their housing-migration dynamics are relatively constant across the time series considered. We discuss potential rationale for some of these differences below.

Conclusion

This paper considers the hypothesis that there are linkages between regional single-family residential vacancy rates and migration flows between regions. We first extend the Wheaton (1990) theoretical model of search and matching for housing to allow for interactions across regions due to migration, regional housing spillovers, and multiple vacancy rate regimes. While our theoretical model implies that there could be potential interdependencies across regions, our empirical analysis enables us to test the hypothesis that there are actual interdependencies linking regions through migration flows.

More specifically, our theoretical model identifies potential interdependencies between regions that may arise due to the desire of homeowners to migrate to a different region. Our theoretical model implies that the two migration variables, HS_j and HM_j, should positively impact the future vacancy rate in region i. These relationships are intuitive because they imply testable hypotheses that more people wanting to move out of region i and into region j (i.e., higher HS_j) should lead to higher vacancy rates in region i; and more people happily living in regions other than region i (i.e., region j) who do not want to migrate into region i (i.e., HM_j) will drive up future vacancy rates in region i. Largely we find that regional vacancy rate movements are driven by HM_j, i.e., those households which are "happily" living outside of region i who have no desire to migrate into region i. However, our empirical results find no consistent effect on regional vacancy rates from changes in HS_j, households which are

"unhappily" living in region i wishing to migrate out (and into region j). The MS-VAR also allows us to capture if these migration effects are asymmetric across vacancy rate regimes. For the South and Midwest, our results suggest that migration flows influence regional vacancy rates more during high average vacancy rate and high volatility periods. But migration effects are relatively consistent across the vacancy-rate cycle for both the Northeast and West regions.

Empirically, on our path to testing the above hypotheses, we estimated an array of simple univariate Markov-switching models both at the national and regional levels. We found that a Markov-switching model with three regimes gave the most intuitive identification of high-, low-, and intermediate-vacancy rate periods. However, the model which best fit the vacancy rate data was one which included both two regimes and an autoregressive term to control for persistence.

These empirical features of the vacancy rate time series provided the motivation for us to test for migration spillovers in a MS-VAR. This multivariate specification allowed us to capture all of the features outlined by our theoretical model as well as the empirical dynamics of regional vacancy rates. The MS-VAR models tend to fit the data quite well, suggesting that a given region's vacancy rate is well-informed by both migration flows and other regions' vacancy rates.

These findings are intriguing. Change in the number of people living outside of the South who do not desire to migrate into the South, is the primary driver of vacancy rates in the South during high vacancy regimes. Similar reasoning holds for the Midwest. But we do not see this phenomenon in the Northeast and the West; vacancy rates responses to migration in these regions are symmetric across vacancy rate regimes.

To explain the intuition behind these factors, the inequality argument proposed by Bayoumi and Barkema (2019) may provide some insights. First note that the West and Northeast – with many high-tech centers such as Boston, New York, Washington DC, San Francisco, San Jose, and Los Angeles - tend to be the two most prosperous regions. On the other hand, the South and Midwest tend to be less prosperous, given the average incomes in many cities in those two regions. Bayoumi and Barkema (2019) find that migration from more to less prosperous areas has been hampered by the fact that earnings in the less prosperous areas have fallen by more than housing costs in those areas. In our context, individuals who are happy living in the more prosperous regions (Northeast and West) who do not desire to migrate to the South or Midwest may be driving higher vacancy rates in the South and Midwest during high vacancy rate regimes. This is particularly relevant for the South and Midwest since they are two regions with decreasing earnings generally dominating decreasing housing costs. Individuals living in the Northeast or West are generally earning more than their counterparts in the South or Midwest, and therefore their incentives to migrate to the South or Midwest may be relatively low. This may exacerbate the vacancy rate increases in the South and Midwest during times of high vacancy rates in those regions – as there are relatively few potential in-migrants who are able to sop up the growing excess supply of housing.

References

Alessandri, Piergiorgio, and Mumtaz, Haroon. 2019. "Financial Regimes and Uncertainty Shocks," *Journal of Monetary Economics*, 101, pp. 31-46.

Anenberg, Elliot. 2016. "Information Frictions and Housing Market Dynamics," *International Economic Review*, 57(4), pp. 1449-1479.

Auerbach, Alan J and Gorodnichenko, Yuriy. 2012. "Measuring the Output Responses to Fiscal Policy," *American Economic Journal: Economic Policy* 4(2), pp. 1–27.

Bayoumi, Tamim and Barkema, Jelle. 2019. "Stranded! How Rising Inequality Suppressed US Migration and Hurt Those 'Left Behind,'" IMF Working Paper WP/19/122, June.

Billio, Monica; Casarin, Roberto; Ravazzolo, Francesco; and van Dijk, Herman K. 2016. "Interconnections between Eurozone and US Booms and Busts Using a Bayesian Panel Markov-Switching VAR Model," *Journal of Applied Econometrics*, 31(7), pp. 1352-1370.

Blake, Andrew and Mumtaz, Haroon. 2015. "Applied Bayesian Econometrics for Central Bankers," *Handbooks*, Centre for Central Banking Studies, Bank of England, number 36.

Cermeño, Rodolfo. 2002. "Growth Convergence Clubs: Evidence from Markov-Switching Models Using Panel Data," 10th International Conference on Panel Data Discussion Paper D5-3, July.

Cheshire, Paul; Hilber, Christian; and Koster, Hans. 2015. "Regulating Housing Vacancies Away? The Paradoxical Effects of Mismatch," Spatial Economics Research Centre Discussion Paper 181, July.

Ehrmann, Michael; Ellison, Martin; and Valla, Natacha. 2003. "Regime-Dependent Impulse Response Functions in a Markov-Switching Vector Autoregression Model," *Economics Letters*, 78(3), pp. 295-299.

Fritzsche, Carolin and Vandrei, Lars. 2014. "Theoretical Causes for Vacancies in the Housing Market," *Credit and Capital Markets*, 47(3), pp. 465-483.

Frey, William H. 2019. "For the First Time on Record, Fewer than 10% of Americans Moved in a Year," *Brookings – the Avenue*, November 22.

Frühwirth-Schnatter, Sylvia. 2006. *Finite Mixture and Markov Switching Models*, Springer Series in Statistics, New York: Springer-Verlag.

Gabriel, Stuart A. and Nothaft, Frank E. 1988. "Rental Housing Markets and the Natural Vacancy Rate," *Real Estate Economics*, 16(4), pp. 419-429.

Genesove, David and Mayer, Christopher. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), pp. 1233-1260.

Goodman, Jr., John L. and Ittner, John B. 1992. "The Accuracy of Home Owners' Estimates of House Value," *Journal of Housing Economics* 2(4), pp. 339-357.

Guren, Adam. 2018. "House Price Momentum and Strategic Complementarity," *Journal of Political Economy*, 126(3), pp. 1172-1218.

Hall, Stephen G.; Psaradakis, Zacharias; and Sola, Martin. 1999. "Detecting Periodically Collapsing Bubbles: A Markov-Switching Unit Root Test," *Journal of Applied Econometrics* 14(2), pp. 143–154.

Hamilton, James D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), pp. 357-384.

Hamilton, James D. 1994. *Time-Series Analysis*, Princeton, NJ: Princeton University Press.

Hamilton, James D. 2010. "Regime Switching Models," in *Macroeconometrics and Time Series Analysis*, ed. by Durlauf Steven N. and Blume, Lawrence E., The New Palgrave Economics Collection. Palgrave Macmillan, London, pp. 202-209.

Hamilton, James D. and Owyang, Michael T. 2012. "The Propagation of Regional Recessions," *Review of Economics and Statistics*, 94(4), pp. 935-947.

Han, Lu and Strange, William C. 2015. "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage," in *Handbook of Regional and Urban Economics, Volume 5B,* ed. by Duranton, Gilles; Henderson, Vernon: and Strange, William, North Holland, pp. 813-886.

Harrison, Austin and Immergluck, Dan. 2020. "The Battle of the Belts: Comparing Housing Vacancy in Larger Metros in the Sun Belt and the Rust Belt since the Mortgage Crisis, 2012 to 2019," working paper, July.

Ihlanfeldt, Keith and Mayock, Tom. 2012. "Information, Search, and House Prices: Revisited," *Journal of Real Estate Finance and Economics*, 44(1), pp. 90-115.

Ihle, Rico; von Cramon-Taubadel, Stephan; and Zorya, Sergiy. 2009. "Markov-Switching Estimation of Spatial Maize Price Transmission Processes between Tanzania and Kenya," *American Journal of Agricultural Economics*, 91(5), pp. 1432-1439.

Ihrke, David. 2014. "Reason for Moving: 2012 to 2013," *Current Population Reports, P20-574*, Washington, DC: US Census Bureau.

Kanas, Angelos And Genius, Margarita. 2005. "Regime (Non) Stationarity in the US/UK Real Exchange Rate," *Economics Letters*, 87(3), pp. 407–413.

Kim, Chang-Jin and Nelson, Charles R. 1999. *State-Space Models with Regime Switching*. MIT Press, Cambridge, Massachusetts

Koop, Gary; Pesaran, M. Hashem; and Potter, Simon M. 1996. "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74(1), pp. 119-147.

Lisi, Gaetano. 2015. "Real Estate Macroeconomics and the Four-Quadrant Model: DiPasquale-Wheaton-Colwell Meet Mortensen-Pissarides," *Journal of Real Estate Practice and Education*, 18(1), pp. 87-106.

Merlo, Antonio and Ortalo-Magné, François. 2004. "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 56(2), pp. 192-216.

Molloy, Raven. 2016. "Long-Term Vacant Housing in the United States," *Regional Science and Urban Economics*, 59, pp. 118-129.

Molloy, Raven; Smith, Christopher L; and Wozniak, Abigail, 2011. "Internal Migration in the United States," *Journal of Economic Perspectives*, 25(3), pp. 173-196.

Molloy, Raven; Smith, Christopher L; and Wozniak, Abigail, 2017. "Job Changing and the Decline in Long Distance Migration in the United States," *Demography*, 54(2), 631-653.

Nelson, Charles R.; Piger, Jeremy; and Zivot, Eric. 2001. "Markov Regime Switching and Unit-Root Tests," *Journal of Business & Economic Statistics*, 19(4), pp. 404–415.

Piazzesi, Monika; Schneider, Martin; and Stroebel, Johannes. 2015. "Segmented Housing Search," NBER Working Paper w20823.

Piger, Jeremy. 2009. "Econometrics – Models of Regime Changes," in *Encyclopedia of Complexity and System Science*, ed. by Meyers, Robert A., Springer, New York, 2744-2757.

Sims, Christopher A. and Zha, Tao. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 96(1), pp. 54–81.

Stein, Jeremy C. 1995. "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Constraints," *Quarterly Journal of Economics*, 110(2), pp. 379-406.

Wheaton, William C. 1990. "Vacancy, Search, and Prices in a Housing Market Matching Model," *Journal of Political Economy*, 98(6), pp. 1270-1292.

Whitaker, Stephan and Fitzpatrick IV, Thomas J. 2013. "Deconstructing Distressed-Property Spillovers: The Effects of Vacant, Tax-Delinquent, and Foreclosed Properties in Housing Submarkets," *Journal of Housing Economics*, 22(2), pp. 79-91.

Williams, Joseph T. 1995. "Pricing Real Assets with Costly Search," *Review of Financial Studies,* 8(1), pp. 55-90.

Table 1

United States Univariate Markov-Switching Models

	MS2	MS3	MS2AR	MS3AR
Regime 1				
US_{t-1}	-	-	0.97	0.93
	-	-	0.02	0.03
Constant	1.43	1.16	0.05	0.09
	0.02	0.03	0.03	0.04
Sigma	0.28	0.22	0.13	0.13
Regime 2				
US_{t-1}	-	-	0.87	0.90
	-	-	0.12	0.07
Constant	2.33	1.64	0.23	0.25
	0.09	0.02	0.20	0.18
Sigma	0.43	0.19	0.31	0.28
Regime 3				
US_{t-1}	-	-	-	0.53
	-	-	-	0.29
Constant	-	2.47	-	0.83
	-	0.10	-	0.51
Sigma	-	0.39	-	0.29
p11	0.99	0.97	0.96	0.99
p21	0.97	0.97	0.73	0.92
p33	-	0.94	-	0.91
D1	143.26	31.00	26.32	102.24
D2	29.66	35.31	3.67	12.32
D3	-	17.84	-	11.11
Observations	258.00	258.00	257.00	257.00
Log-Likelihood	192.14	299.11	457.12	451.64
AIC	-376.28	-586.21	-902.25	-885.28
BIC	-362.06	-564.89	-880.95	-853.34
DIC	-374.38	-586.24	-893.88	-747.53

Table 2

South Univariate Markov-Switching Models

	MS2	MS3	MS2AR	MS3AR
Regime 1				
US_{t-1}	-	-	0.96	0.96
	-	-	0.02	0.02
Constant	1.45	1.39	0.07	0.06
	0.04	0.05	0.04	0.04
Sigma	0.30	0.28	0.16	0.17
Regime 2				
US_{t-1}	-	-	0.59	0.68
	-	-	0.25	0.21
Constant	2.21	2.00	0.71	0.50
	0.04	0.04	0.43	0.36
Sigma	0.39	0.22	0.37	0.33
Regime 3				
US_{t-1}	-	-	-	0.61
	-	-	-	0.21
Constant	-	2.78	-	0.76
	-	0.08	-	0.41
Sigma	-	0.42	-	0.34
p11	0.98	0.97	0.97	0.98
p21	0.98	0.96	0.82	0.89
p33	-	0.94	-	0.88
	-	-	-	-
D1	43.03	29.98	38.72	62.44
D2	42.01	25.43	5.67	9.39
D3	-	15.73	-	8.28
Observations	258.00	258.00	257.00	257.00
Log-Likelihood	153.45	230.52	359.34	356.44
AIC	-298.90	-449.04	-706.67	-694.88
BIC	-284.69	-427.72	-685.38	-662.94
DIC	-296.68	-450.34	-708.08	-647.98

Table 3

Northeast Univariate Markov-Switching Models

	MS2	MS3	MS2AR	MS3AR
Regime 1				
US_{t-1}	-	-	0.94	0.94
	-	-	0.03	0.03
Constant	0.93	0.89	0.07	0.07
	0.03	0.02	0.03	0.04
Sigma	0.22	0.21	0.15	0.15
Regime 2				
US_{t-1}	-	-	0.61	0.72
	-	-	0.23	0.19
Constant	1.64	1.40	0.67	0.49
	0.04	0.03	0.40	0.35
Sigma	0.30	0.26	0.36	0.33
Regime 3				
US_{t-1}	-	-	-	0.51
	-	-	-	0.25
Constant	-	1.87	-	0.78
	-	0.05	-	0.39
Sigma	-	0.29	-	0.34
p11	0.98	0.98	0.98	0.97
p21	0.98	0.97	0.85	0.87
p33	-	0.96	-	0.83
	-	-	-	-
D1	63.47	64.36	46.74	38.82
D2	53.11	38.03	6.74	7.86
D3	-	23.80	-	5.88
Observations	258.00	258.00	257.00	257.00
Log-Likelihood	247.52	282.91	376.21	374.20
AIC	-487.04	-553.82	-740.41	-730.40
BIC	-472.83	-532.50	-719.12	-698.46
DIC	-484.83	-552.46	-744.48	-726.97

Table 4

Midwest Univariate Markov-Switching Models

	MS2	*MS3*	*MS2AR*	*MS3AR*
Regime 1				
US_{t-1}	-	-	0.96	0.95
	-	-	0.02	0.04
Constant	1.21	1.12	0.06	0.07
	0.02	0.03	0.03	0.05
Sigma	0.26	0.23	0.17	0.16
Regime 2				
US_{t-1}	-	-	0.71	0.78
	-	-	0.24	0.20
Constant	2.20	1.62	0.43	0.33
	0.07	0.06	0.30	0.27
Sigma	0.47	0.25	0.37	0.35
Regime 3				
US_{t-1}	-	-	-	0.78
	-	-	-	0.18
Constant	-	2.52	-	0.35
	-	0.09	-	0.27
Sigma	-	0.38	-	0.35
p11	0.99	0.98	0.97	0.94
p21	0.98	0.94	0.76	0.74
p33	-	0.95	-	0.71
D1	134.38	44.56	29.01	15.55
D2	42.99	17.29	4.18	3.85
D3	-	19.17	-	3.50
Observations	258.00	258.00	257.00	257.00
Log-Likelihood	190.44	248.71	350.44	346.51
AIC	-372.87	-485.42	-688.89	-675.02
BIC	-358.66	-464.10	-667.59	-643.08
DIC	-370.48	-484.32	-692.87	-678.99

Table 5

West Univariate Markov-Switching Models

	MS2	MS3	MS2AR	MS3AR
Regime 1				
US_{t-1}	-	-	0.92	0.92
	-	-	0.03	0.04
Constant	1.44	1.26	0.13	0.13
	0.03	0.03	0.05	0.06
Sigma	0.29	0.25	0.20	0.19
Regime 2				
US_{t-1}	-	-	0.37	0.51
	-	-	0.29	0.30
Constant	2.15	1.71	1.00	0.75
	0.07	0.03	0.47	0.46
Sigma	0.41	0.25	0.44	0.40
Regime 3				
US_{t-1}	-	-	-	0.49
	-	-	-	0.31
Constant	-	2.44	-	0.76
	-	0.09	-	0.45
Sigma	-	0.42	-	0.39
p11	0.98	0.96	0.96	0.94
p21	0.95	0.96	0.73	0.73
p33	-	0.92	-	0.75
	-	-	-	-
D1	49.44	25.70	25.13	17.12
D2	18.35	25.73	3.69	3.72
D3	-	11.83	-	4.07
Observations	258.00	258.00	257.00	257.00
Log-Likelihood	170.37	225.60	285.75	283.77
AIC	-332.74	-439.21	-559.50	-549.54
BIC	-318.52	-417.89	-538.20	-517.59
DIC	-330.04	-438.23	-561.56	-555.11

Figure 1

U.S. Homeowner Vacancy Rate

Source: Census Bureau/Haver Analytics
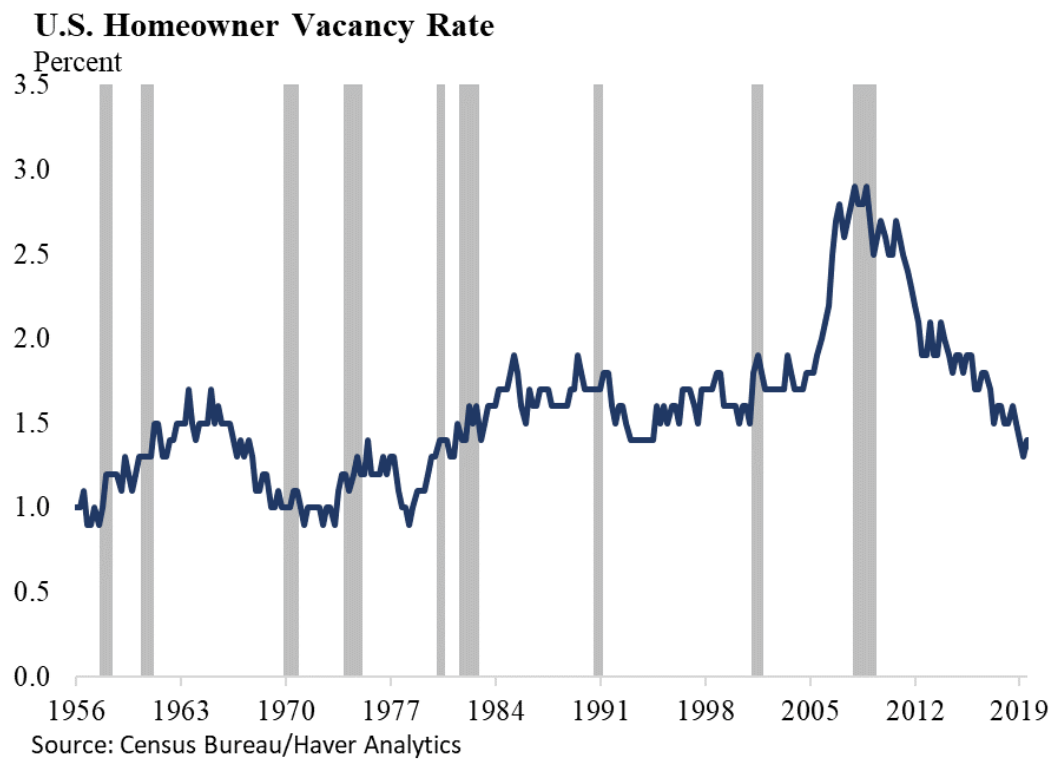
Figure 2

United States MS2 Model

(a) Predicted Vacancy Rate
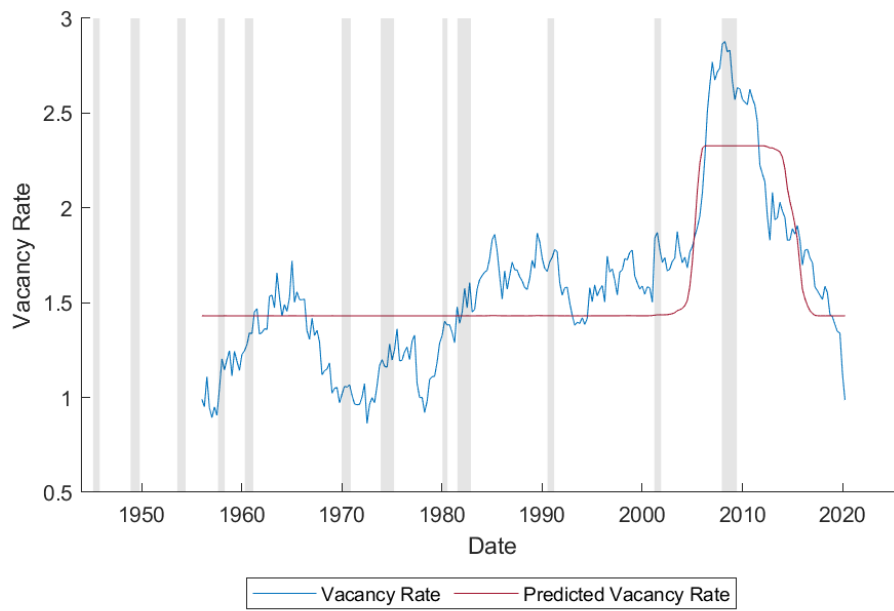
(b) Regime Probabilities
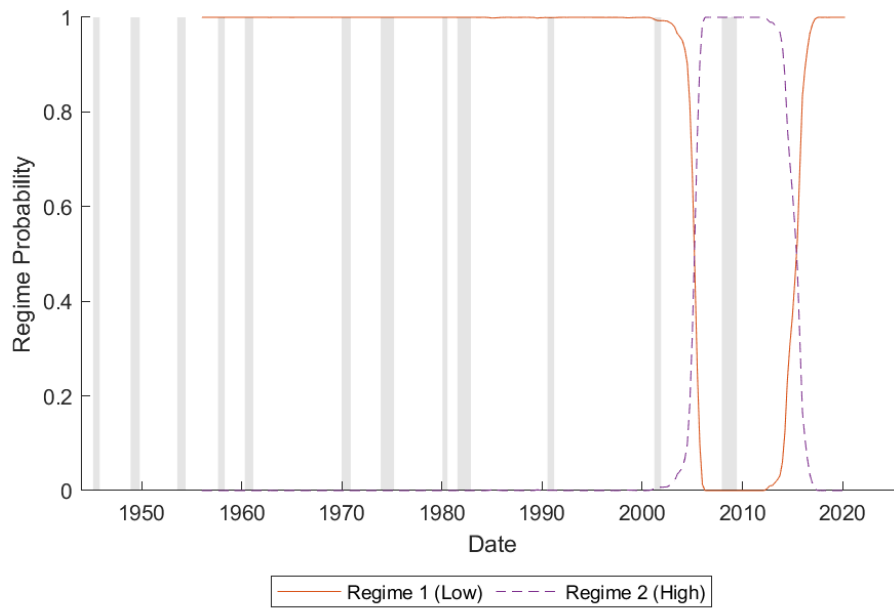
Figure 3

United States MS3 Model

(a) Predicted Vacancy Rate
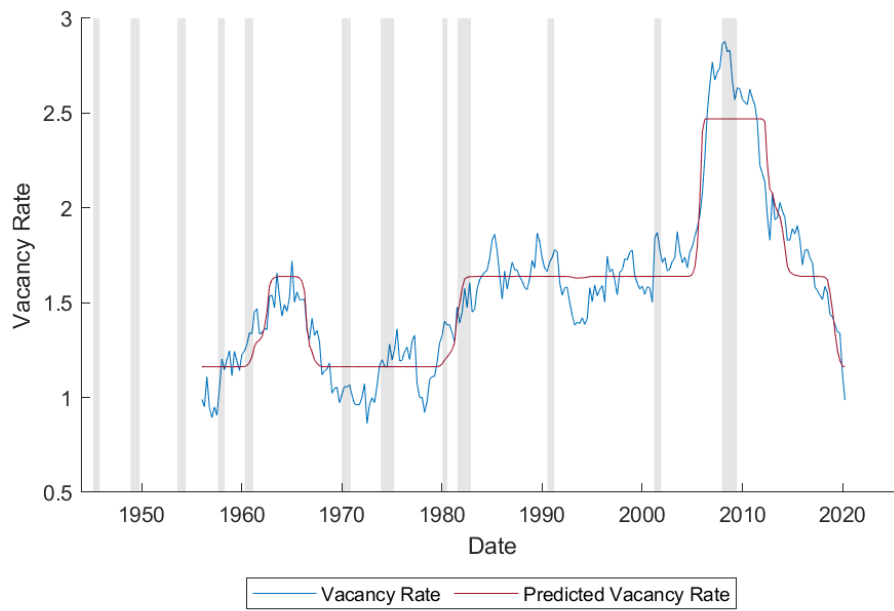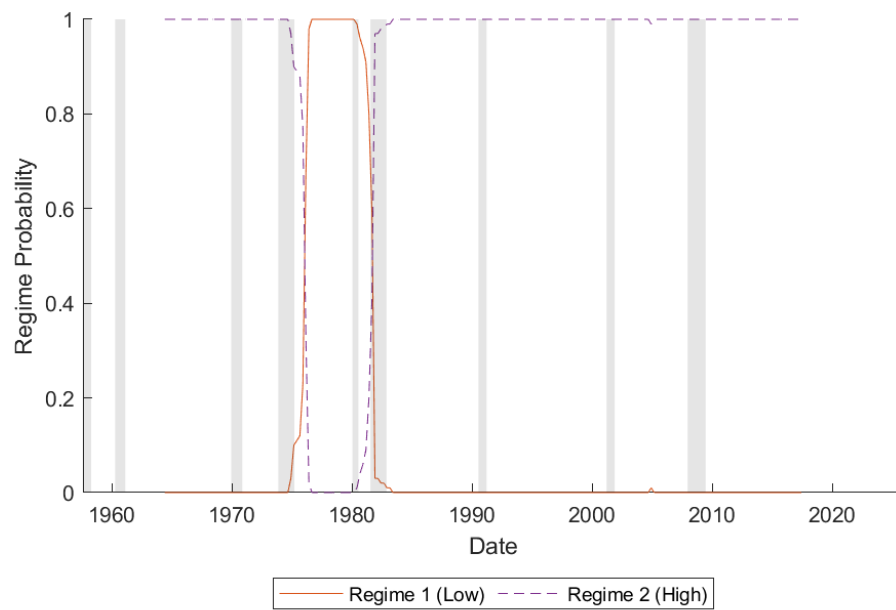

(b) Regime Probabilities
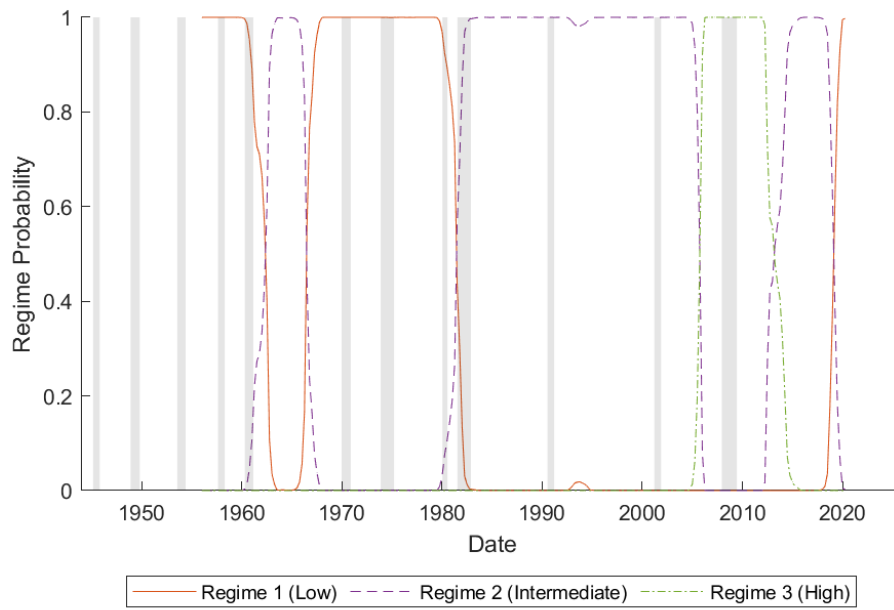
Figure 4

United States MS2AR Model

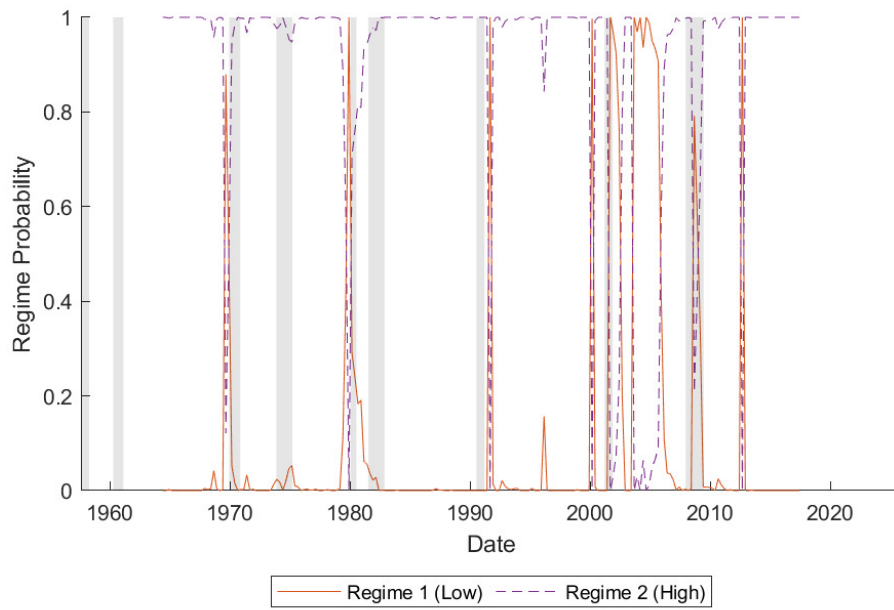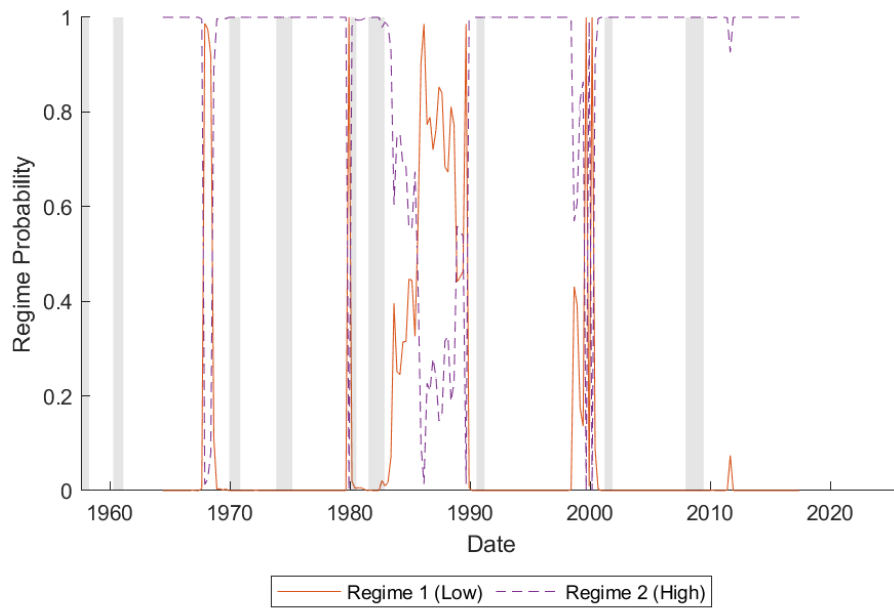*(a) Predicted Vacancy Rate*
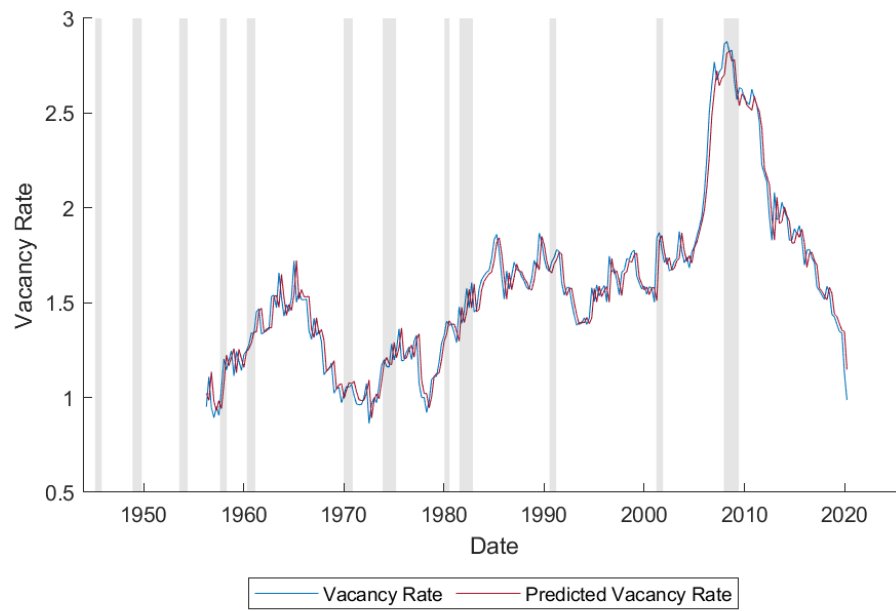
(b) Regime Probabilities
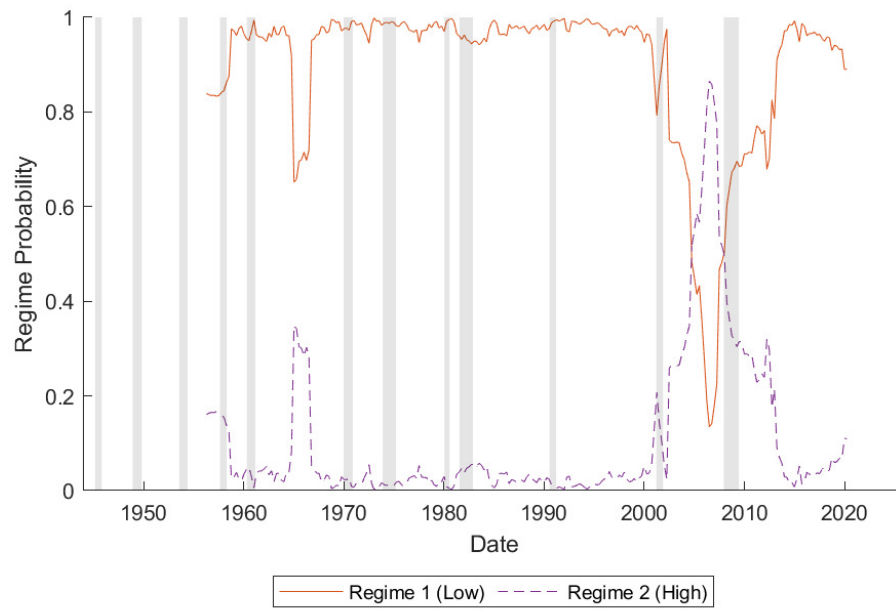
Figure 5

United States MS3AR Model

(a) Predicted Vacancy Rate
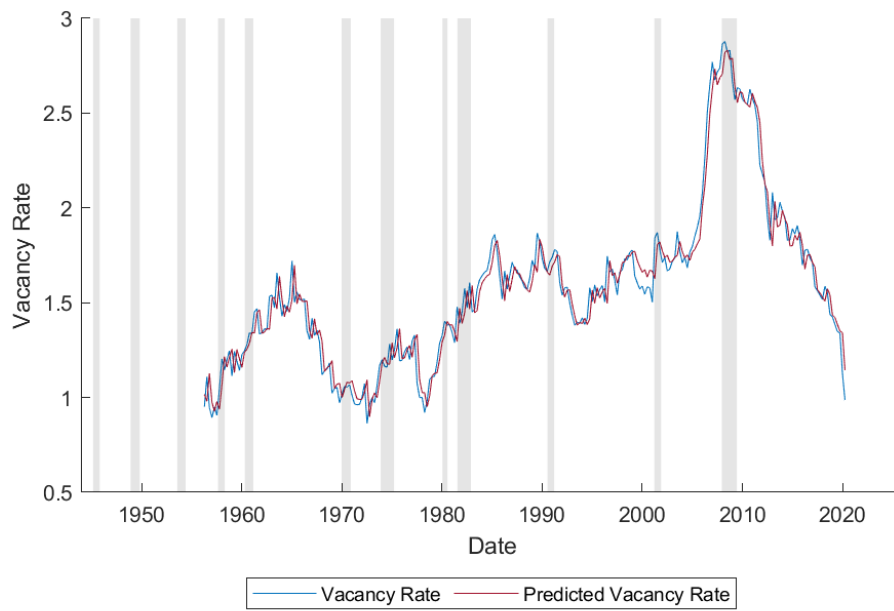
(b) Regime Probabilities
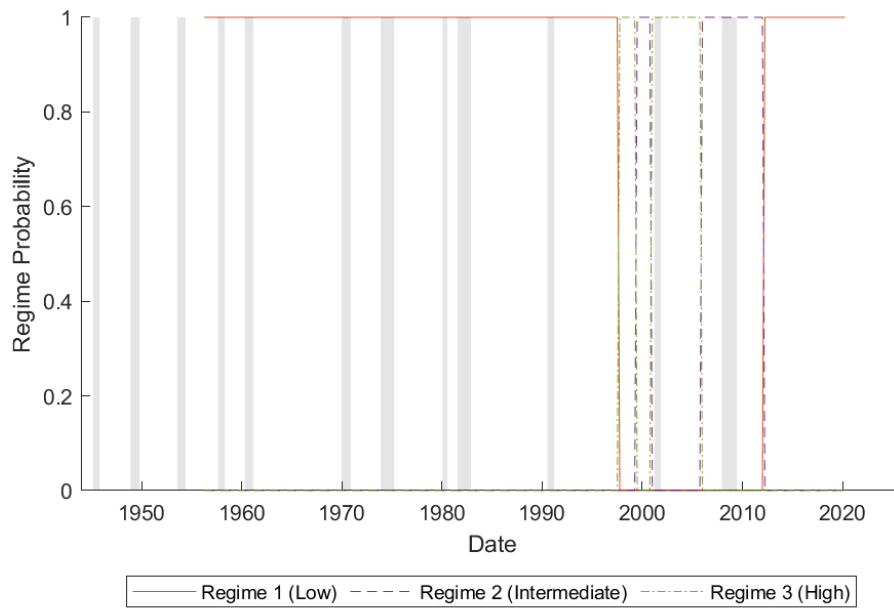
Figure 6

Regional MS3 Models: Regime Probabilities

(a) South

(b) Northeast

(c) Midwest

(b) West

Regime 1 (Low)　Regime 2 (Intermediate)　Regime 3 (High)
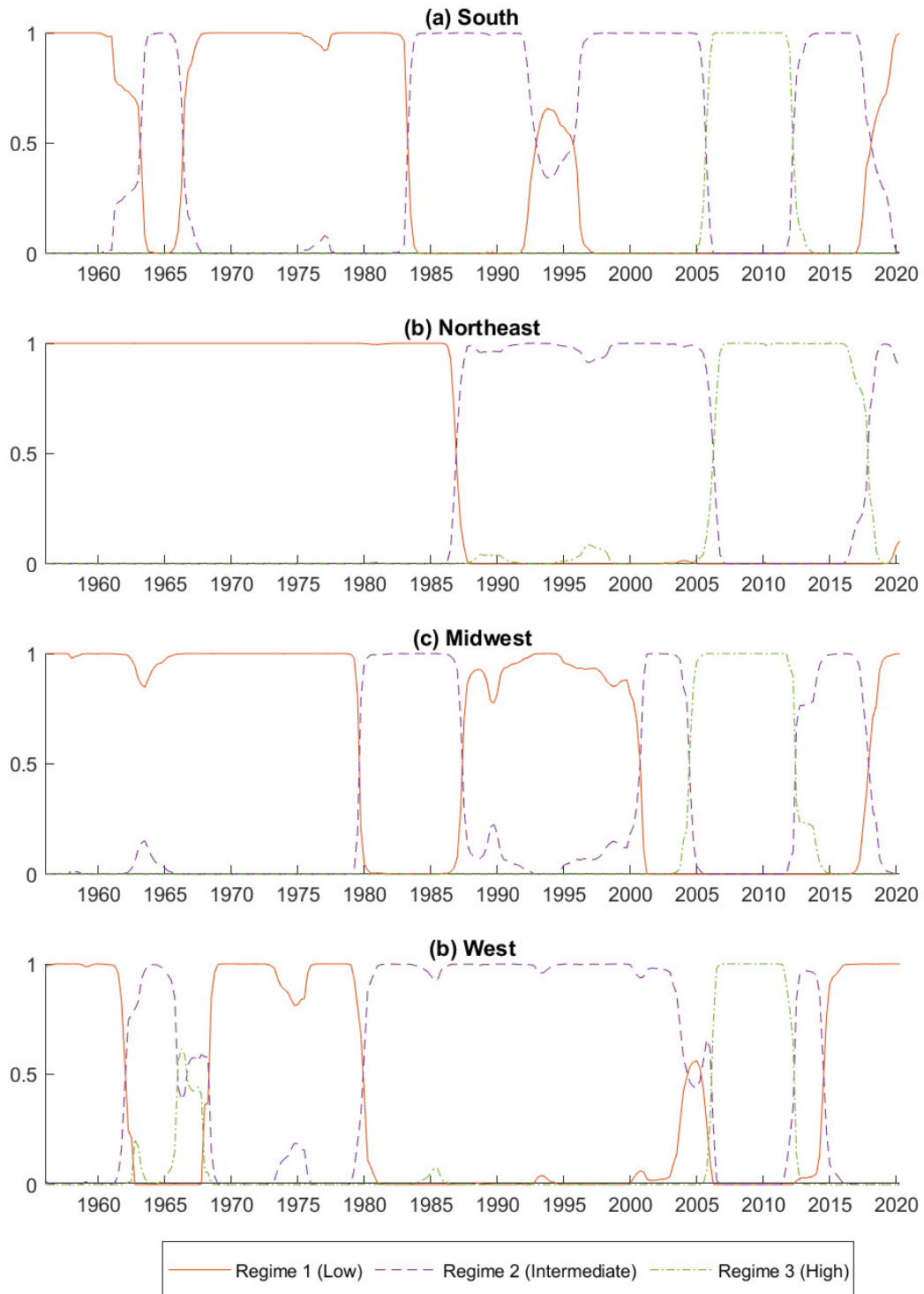
Figure 7

South MS-VAR Model

(a) Predicted Vacancy Rate
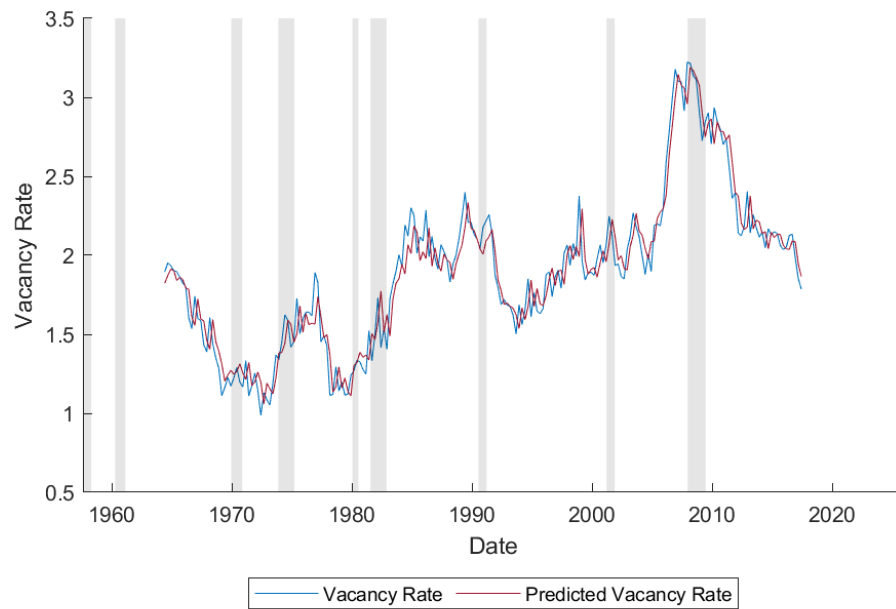
(b) Regime Probabilities

Figure 8

Northeast MS-VAR Model

(a) Predicted Vacancy Rate
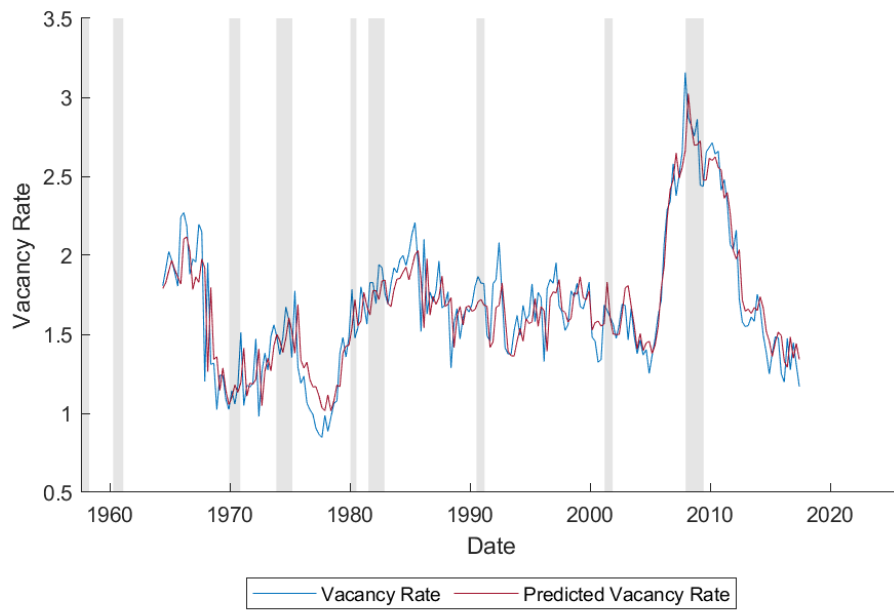
(b) Regime Probabilities

Figure 9

Midwest MS-VAR Model

(a) Predicted Vacancy Rate
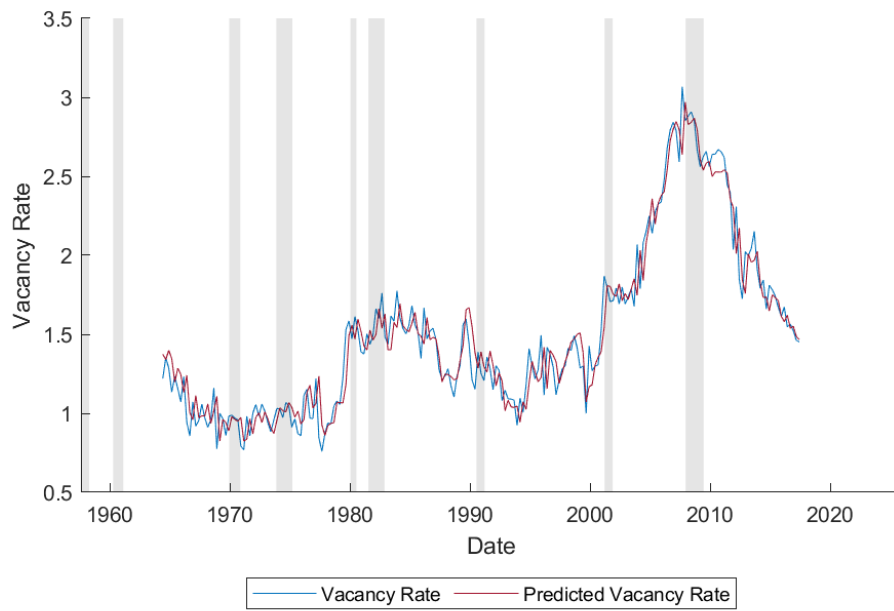
(b) Regime Probabilities

Figure 10

West MS-VAR Model

(a) Predicted Vacancy Rate
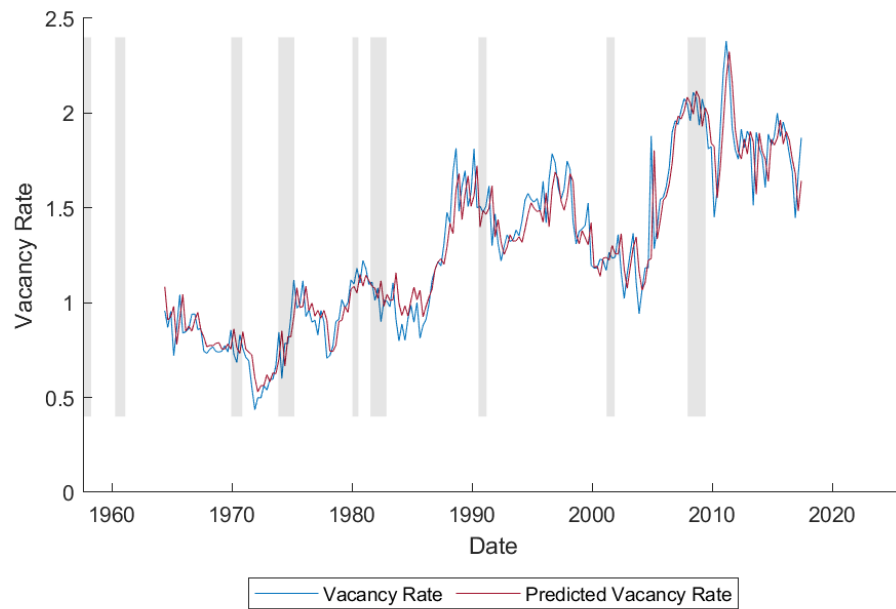
(b) Regime Probabilities
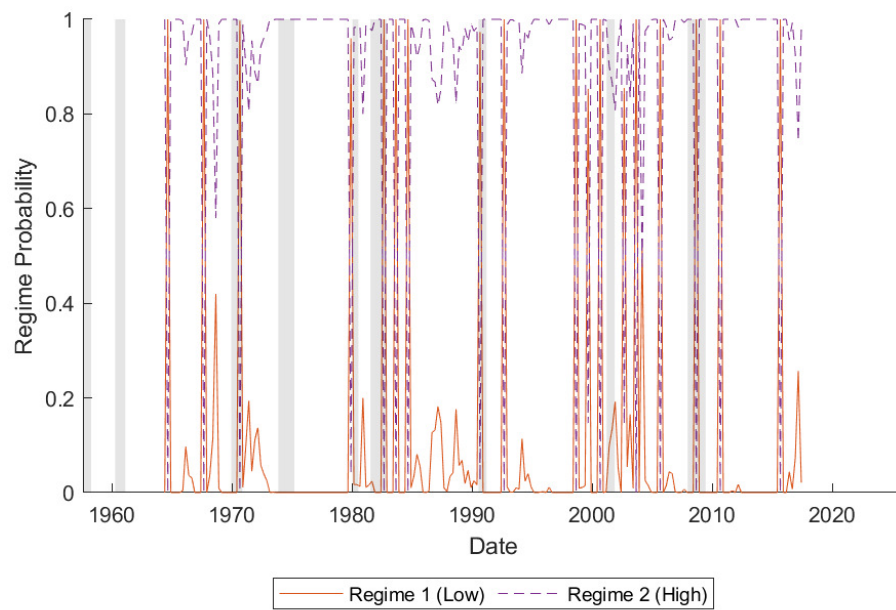
Figure 11

MS-VAR Generalized Impulse Response Functions
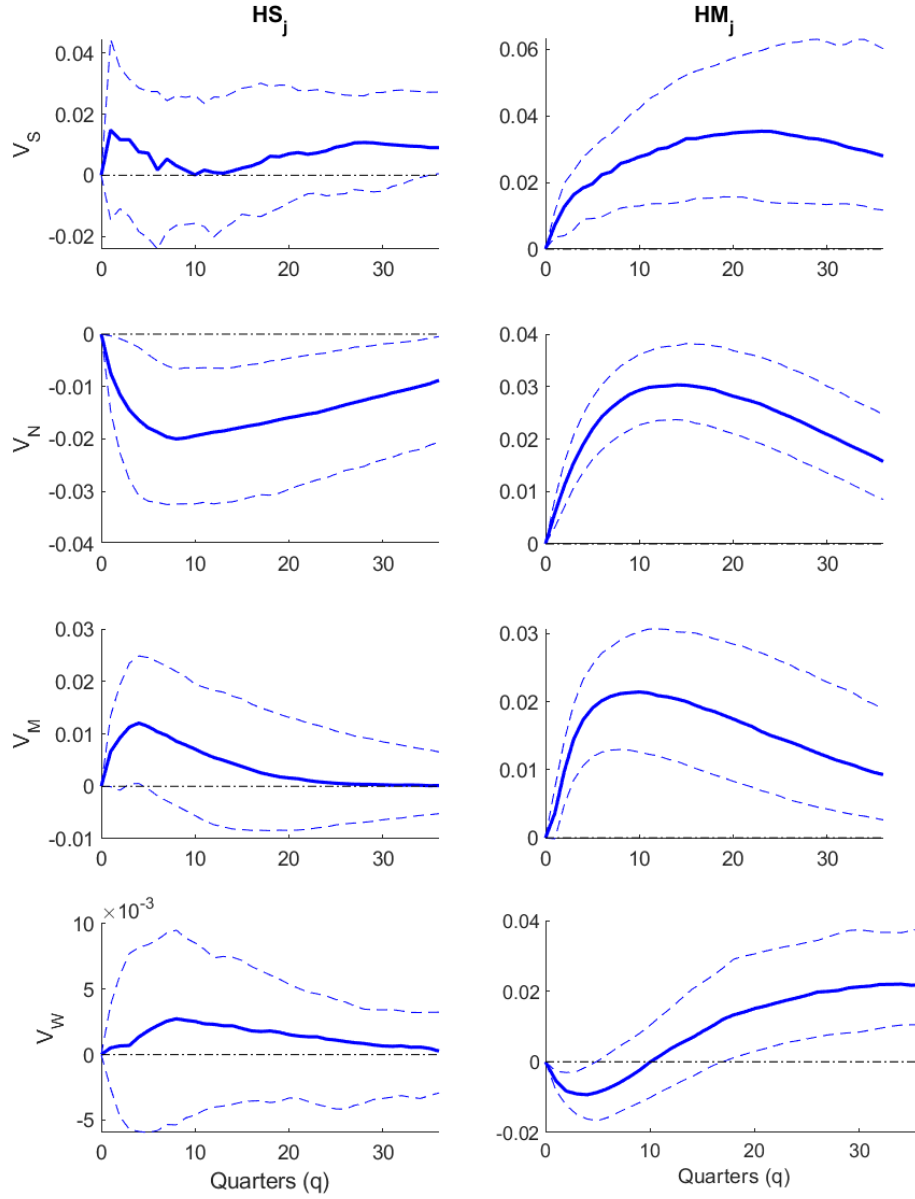
Note: The figure displays the effects of a one standard deviation shock to HS_j and HM_j on each region's vacancy rate, V_i. HS_j in this setting represents those living in the region i wishing to migrate outside of region i, while HM_j represents those happily living outside of region i. The median change in the vacancy rate is plotted in each of these figures from $t = 0$ to $t = 40$ along with dashed lines indicating the 68% highest posterior density interval.

Figure 12

MS-VAR State-Dependent Generalized Impulse Response Functions
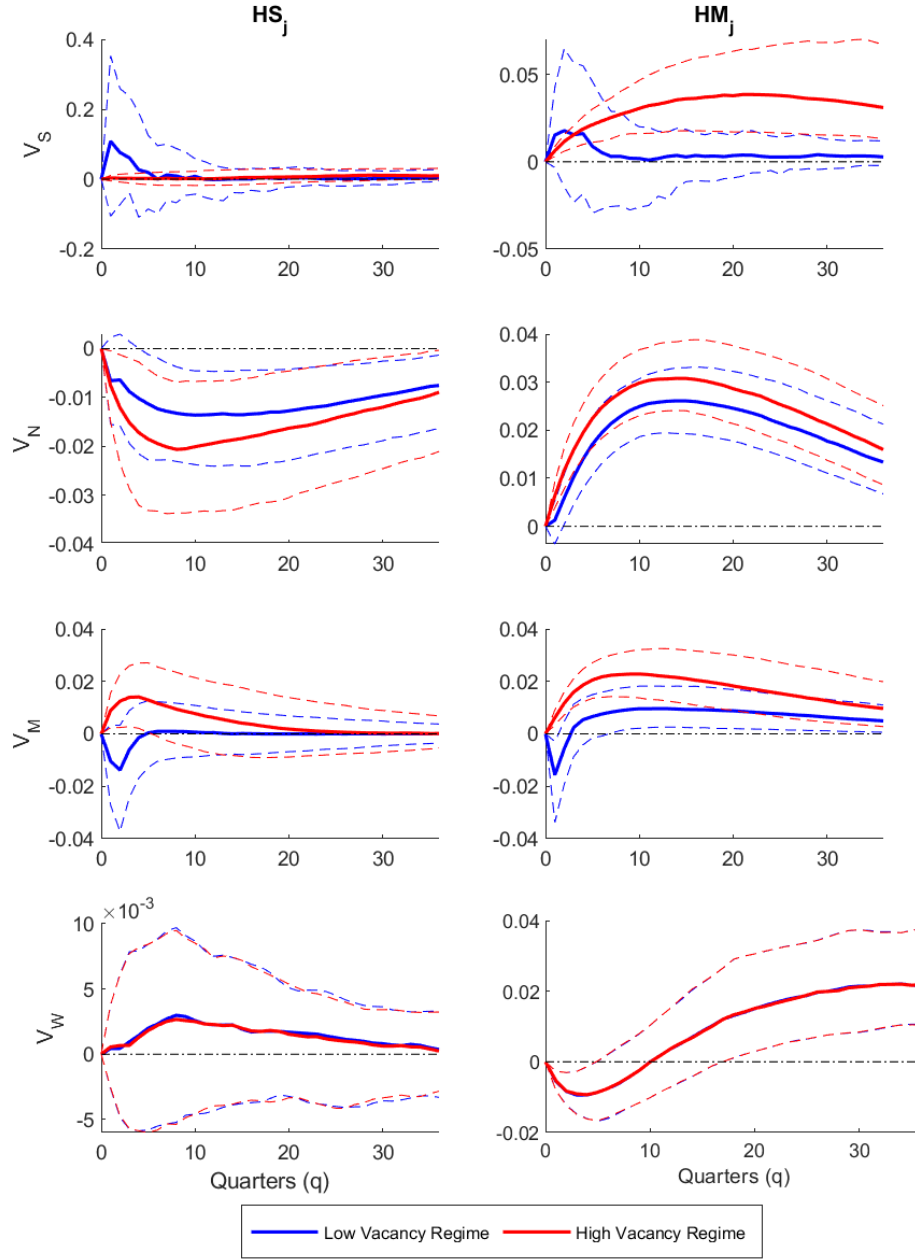
Note: The figure displays the effects of a one standard deviation shock to HS_j and HM_j on each region's vacancy rate, V_i, conditional on the initial regime at time $t = 0$. HS_j in this setting represents those living in the region i wishing to migrate outside of region i, while HM_j represents those happily living outside of region i. The median change in the vacancy rate is plotted in each of these figures from $t = 0$ to $t = 40$ along with dashed lines indicating the 68% highest posterior density interval.

Appendix A:

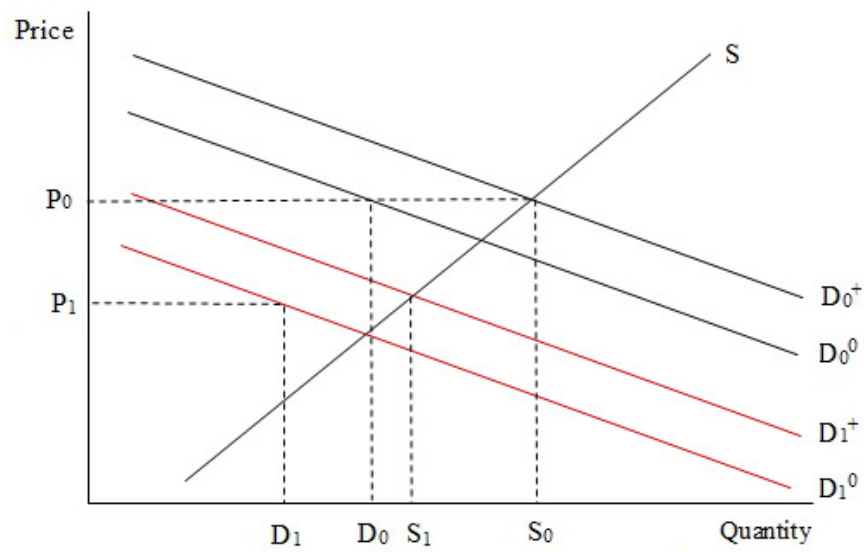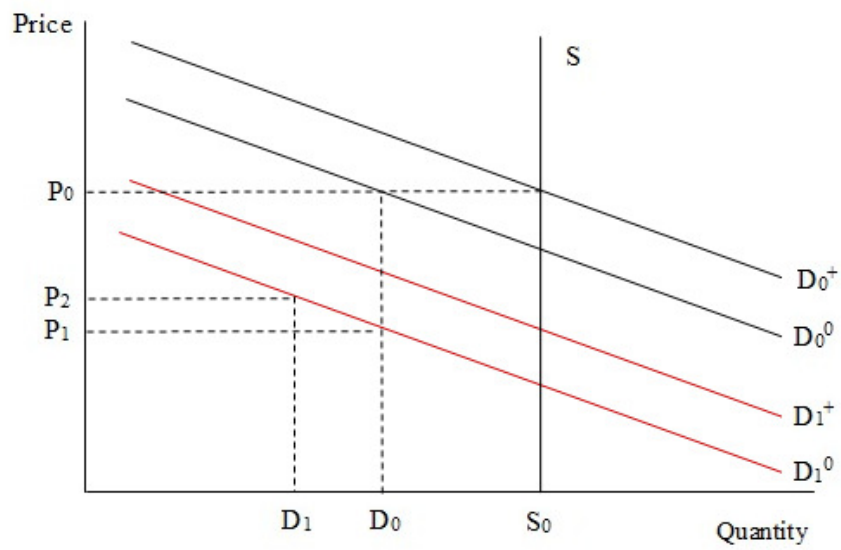*Vacant Housing in a Simple Supply and Demand Framework*

Figure A1 displays the possibility of vacant housing via a simple supply and demand framework.

As shown in the top panel, let the supply of housing units be fixed at S_o. With a demand for occupied housing units, D_o^o, and a total demand for housing units, D_o^+, then the equilibrium price would be P_o, the equilibrium quantity of occupied housing units would be D_o, and the equilibrium quantity of all housing units would be the same as the quantity supplied of all housing units (i.e., S_o). Associated with this equilibrium is a vacancy rate (VR) equal to $((S_o - D_o)/S_o) \times 100$.

Next, suppose demand for housing declines to D_1^+ for all units and D_1^o for occupied units. Such a decrease in demand can occur due to out-migration from a region. Note that the shifts are parallel and equidistant. This is a strong assumption that causes vacancies and the associated vacancy rate to remain unchanged. If the demand for occupied housing shifts relatively more (less) than the demand for total housing, then vacancies and the vacancy rate increase (decrease) relative to the initial equilibrium. Such a decline puts downward pressure on housing prices. If price adjusts completely and instantaneously, then the new equilibrium price would be P_1 and the equilibrium quantities of occupied and total housing would be unchanged. Vacancies and the associated vacancy rate would also remain unchanged.

Figure A1

The Housing Market and Vacancy Rates

Price

S

P_0

D_0^+

D_0^0

P_2
P_1

D_1^+

D_1^0

D_1 D_0 S_0 Quantity

Price

S

P_0

D_0^+

P_1

D_0^0

D_1^+

D_1^0

D_1 D_0 S_1 S_0 Quantity

48

Appendix B:

Discussion of Basic Markov-Switching Models

Some time series variables exhibit changes in behavior from one stretch of time to the next. For an example of such a variable, consider a city that has persistent periods of two types of easily distinguished weather - comfortable temperatures and heat waves. Daily construction activity, C, in a non-heat wave period at time t is described as $C_t = \beta_0 + \alpha_0'(Z_t)$, where β_0 and α_0 are respectively a constant and a coefficient column vector, both being specific to comfortable days. Z_t is a column vector of explanatory variables at time t. Meanwhile, construction activity during heat waves could be described as $C_t = \beta_1 + \alpha_1'(Z_t)$. It is reasonable to believe that, ceteris paribus, the former equation will predict much higher construction activity than the latter. Further, assuming we have all the data for independent variables and daily construction activity, estimation of these equations is relatively straightforward because we can distinguish whether past days were in a heat wave, and then estimate each equation using appropriate data.

Unfortunately, it is frequently difficult to distinguish between variables' different behaviors, which can be thought of as belonging to different states or regimes. Consider U.S. labor productivity growth, which likely goes through periods of low growth and high growth. We do not have any binary indicators that tell us when some shock has, for example, shifted labor productivity to a low-growth regime. Such shocks are random and difficult to spot even after they have occurred. Markov-switching models, first discussed in Hamilton (1989) and further detailed in Hamilton (1994), attempt to deal with this problem of having a variable characterized by different regimes that have no simple way of being identified. That is, Markov-switching models attempt to identify unobservable states of the world, and describe how variables behave in each state. "Markov" refers to the fact that the models use Markov chains to characterize the unobserved regimes.

A general Markov-switching model, with dependent variable at time t, y_t, and column vector of explanatory variables, possibly including lags of y, at time t, X_t, is:

$$y_t = \beta_{s_t} + \alpha'_{s_t}(X_t) + \varepsilon_t,$$

where s_t, derived through estimation of a Markov chain, indicates the regime of y at time t, and $\varepsilon_t \sim$ i.i.d. $N(0, \sigma^2_{s_t})$.[16] Thus, the variance of the error term, the value of the constant, β, and the coefficient column vector, α, can depend on s_t. It is not required that all parameters depend on the regime, but at least one must or (13) would become a simple regression. Note that if X contains no variables, (13) is reduced to a constant-only model where the constant for each regime will simply be the regime's average value for y.

Markov-switching models have been used in a variety of applications. Billio et al. (2016) and Hamilton and Owyang (2012) use the models to compare business cycles across countries and U.S. states, respectively. Cermeño (2002) uses a simple Markov-switching model to

---

[16] The inclusion of autoregressive variables tends to produce smoother regime changes. Models without autoregressive variables tend to generate abrupt switches.

characterize low-growth and high-growth regimes for per capita output of U.S. states and several countries. Ihle, von Cramon-Taubadel, and Zorya (2009) use a Markov-switching model to characterize the transmission of maize prices between two African countries. For surveys of Markov-swtiching models, see Hamilton (2010) and Piger (2009).

Estimation Details

The univariate and multivariate Markov-switching models were estimated using the Bayesian method of Gibbs sampling. The Gibbs sampler is a Markov-chain Monte Carlo (MCMC) technique which constructs the full joint posterior distribution of the model by drawing the parameters and latent regime in blocks from their conditional posterior distributions. In the following paragraphs, we outline the sampler for the most general models from which the approaches for the simpler specifications can be deduced.

For the univariate model, there are four distinct blocks: (i) the mean growth parameters C_{S_t} and coefficient on the lagged vacancy rate a_{S_t}, (ii) the error variance $\sigma_{S_t}^2$ where $\varepsilon_t \sim N\left(0, \sigma_{S_t}^2\right)$, (iii) the transition probabilities p_{ji} in the transition matrix P, and (iv) the latent regime time series $S = \{S_1, \ldots, S_T\}$. In order to identify the regimes and to avoid label-switching across Gibbs sampling iterations, we impose that the mean vacancy rate parameter is larger in regime 3 than regime 2 and larger in regime 2 than regime 1 (i.e., $C_1 < C_2 < C_3$). Additionally to ensure stationarity, we reject draws of a_{S_t} which imply a nonstationary process for V (i.e., each draw must satisfy $|a_{S_t}| < 1$. We use standard priors and Gibbs sampling steps from the literature as outlined by Kim and Nelson (1999).

For the multivariate model, there are again four distinct blocks: (i) the entire set of coefficients which we label the matrix B_{S_t}, (ii) the variance-covariance matrix Σ_{S_t} of the error terms, (iii) the transition matrix P, and (iv) the regime series S. Similar to the univariate model, we redraw any draw of B_{S_t} for which the steady-state vacancy rate of the region of interest is not higher in regime 2 (again to avoid label-switching across iterations). We also reject any draw of B_{S_t} which violates stationarity. The priors and sampling steps are standard as outlined by Blake and Mumtaz (2017). The full set of priors and initial conditions are available from the authors upon request.

Additional Results: Unconditional Generalized Impulse Responses

The main text only displays the impulse response functions for the two migration variables, HS_j and HM_j. However, the spillovers of housing dynamics across regions could be captured by the response of one region's vacancy rate to a shock in another region's vacancy rate. In what follows, we present the full set of generalized impulse responses for each region's MS-VAR model and discuss only the vacancy rate dynamics between regions since the migration effects are condensed and discussed in the main text.

Figure B1 displays the GIRF for the MS-VAR model applied to the South region. The first panel of Figure B1 shows the response of the South vacancy rate across future horizons $V_{S,t+h}$ to a one standard deviation shock to the South vacancy rate at time $t = 0$. The second panel of Figure B1 shows the response of the South vacancy rate $V_{S,t+h}$ to a one standard

deviation shock to the Northeast vacancy rate at time $t = 0$. In a similar fashion, the remaining four panels display the dynamic response of the South vacancy rate to a one standard deviation shock to the remaining endogenous variables. As expected, the South vacancy rate is positively affected by its own shock in a considerable way. A shock to the Northeast vacancy rate puts downward pressure on the South vacancy rate after 8 quarters, whereas shocks to the Midwest or West vacancy rate both increase the South vacancy rate immediately.

The corresponding GIRFs for the Northeast vacancy rate are presented in Figure B2. The Northeast vacancy rate increase in response to a one standard deviation shock to the South, Northeast, Midwest, or (to a lesser extent) West vacancy rates.

The responses of the Midwest vacancy rate presented in Figure B3 are qualitatively similar to those for the South. A positive shock to the South, Midwest, or West vacancy rate also increases the Midwest vacancy rate. Yet again, a shock to the Northeast vacancy rate decreases the Midwest vacancy rate.

Finally, Figure B4 displays the responses of the West vacancy rate to the various shocks in the model. The results for the West mostly match the findings for the South and Midwest. The West vacancy rate responds positively to a shock to the South, Midwest, and West vacancy rates; but a shock to the Northeast vacancy rates drives the West vacancy rate down.

We summarize the results from the GIRFs in table A1. A positive sign "+" indicates a significant positive response of the respective vacancy rate (columns) to each shock (row). Conversely, a negative sign "−" indicates a significant negative response and a "x" reflects no substantive relationship. The table illustrates a number of two key takeaways. First, the South, Midwest, and West all have a positive feedback to each other. Second, a shock to the Northeast actually decreases all other regions vacancy rates besides its own.

Additional Results: State-Dependent Generalized Impulse Responses

Figure B5 shows the GIRF of the South vacancy rate to various shocks in the model conditional on the initial regime. For example, the blue line in the first panel of Figure B5 shows the response of the South vacancy rate to its own shock if the regime at time $t = 0$ in the relatively low vacancy regime. In the same panel, the red line shows the response of the South vacancy rate to the same shock but if the regime at time $t = 0$ is the high vacancy regime. In both regimes, the South vacancy rate responds positively but the effect is considerably more significant in the high vacancy regime. The response of the South vacancy rate in a high vacancy regime to the other shocks qualitatively matches the unconditional GIRF presented in Figure B1. However, in the low vacancy regime, the only significant effect on the South vacancy rate comes from its own shock. This finding could imply that this time period was relatively isolating for the South and so the interregional effects are small, or it is possible that the response in the low regime is not well identified since there is only one period (the late 1970s) where the South is in in the low regime.

The state-dependent responses of the Northeast vacancy rate are presented in Figure B6. The Northeast vacancy rate responds qualitatively the same across the two regimes and six different shocks. The notable exception is that the effect of an increase in the West vacancy rate is only significant in the high vacancy regime and not the low vacancy regime. Responses to the other shocks mainly differ in timing of significance but the shape is consistent across regimes.

The Midwest vacancy rate responses across the two regimes are shown in Figure B7. Similar to the Northeast, the results for each regime for the Midwest match the unconditional GIRF results presented in Figure B3. The primary difference is that the response of the Midwest vacancy rate is significantly positive to the number of households outside the Midwest that would like to move there in the high vacancy regime, but the effect is insignificant in the low regime.

Lastly, Figure B8 displays the state-dependent GIRFs of the West vacancy rate to the six shocks. As mentioned in the main text, the responses are virtually identical across regimes, possibly because of the estimated short durations of the low vacancy regime for the West.

In summary, we find that the interregional effects of migration are broadly similar for three of the four regions across the two regimes. Similar to the unconditional GIRFs presented in the previous section, there tends to be a high degree of spillovers for these three regions regardless of whether a given region was in a low or high vacancy regime. However, the South briefly experienced some insulation from these spillover effects in the late 1970s. In the more recent past, all four regions tend to be connected to each other in a positive fashion except for the Northeast which tends to cause other regions' vacancy rates to fall.

Table B1

Summary of Interregional Effects from MS-VAR Model

		Response Variable			
		V_S	V_{NE}	V_M	V_W
Shocked Variable	V_S	+	+	+	+
	V_{NE}	-	+	-	-
	V_M	+	+	+	+
	V_W	+	+	+	+
	HS_j	x	-	+	x
	HM_j	+	+	+	-/+

Figure B1

South MS-VAR Generalized Impulse Response Function
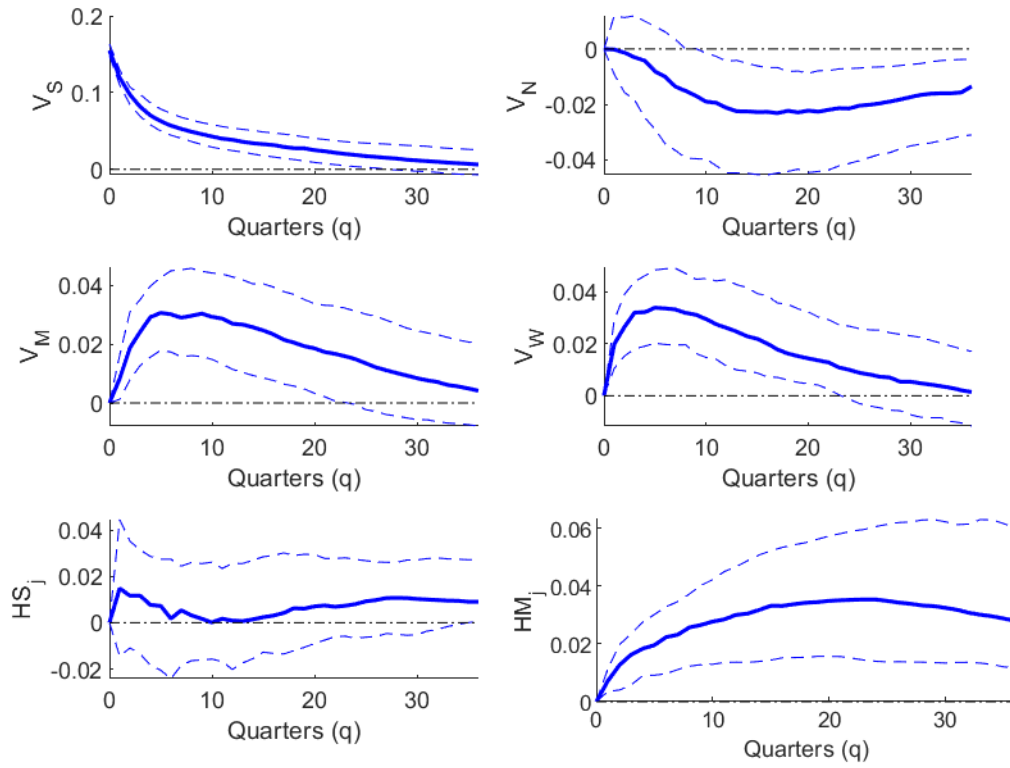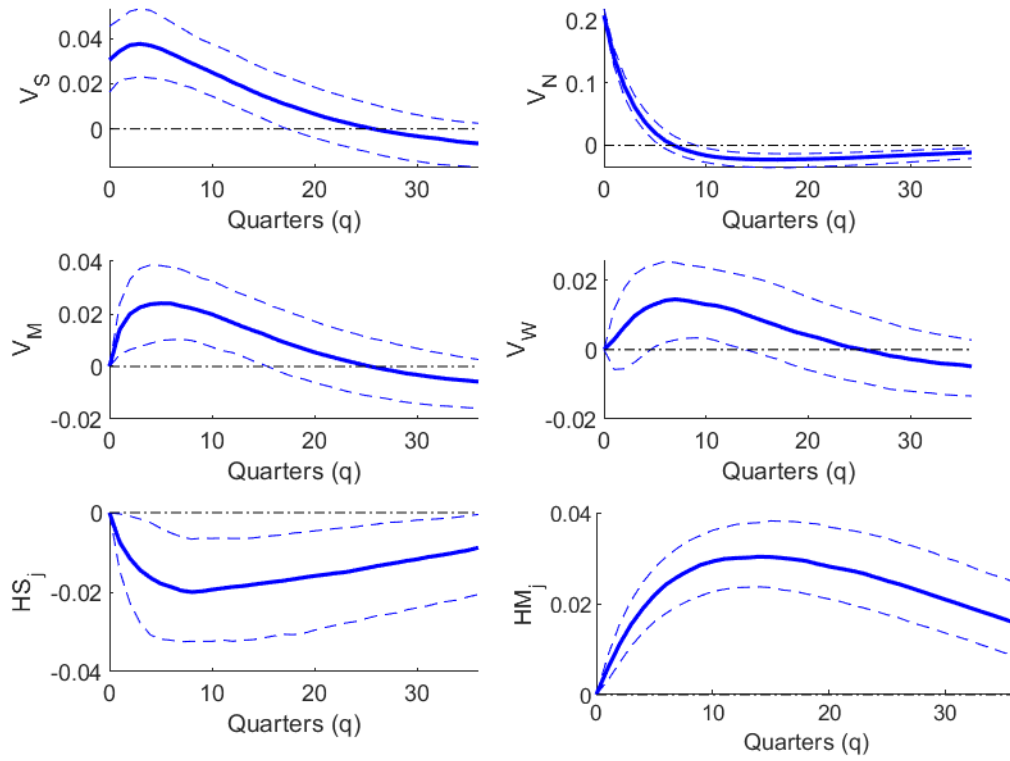
Figure B2

Northeast MS-VAR Generalized Impulse Response Function

Figure B3

Midwest MS-VAR Generalized Impulse Response Function
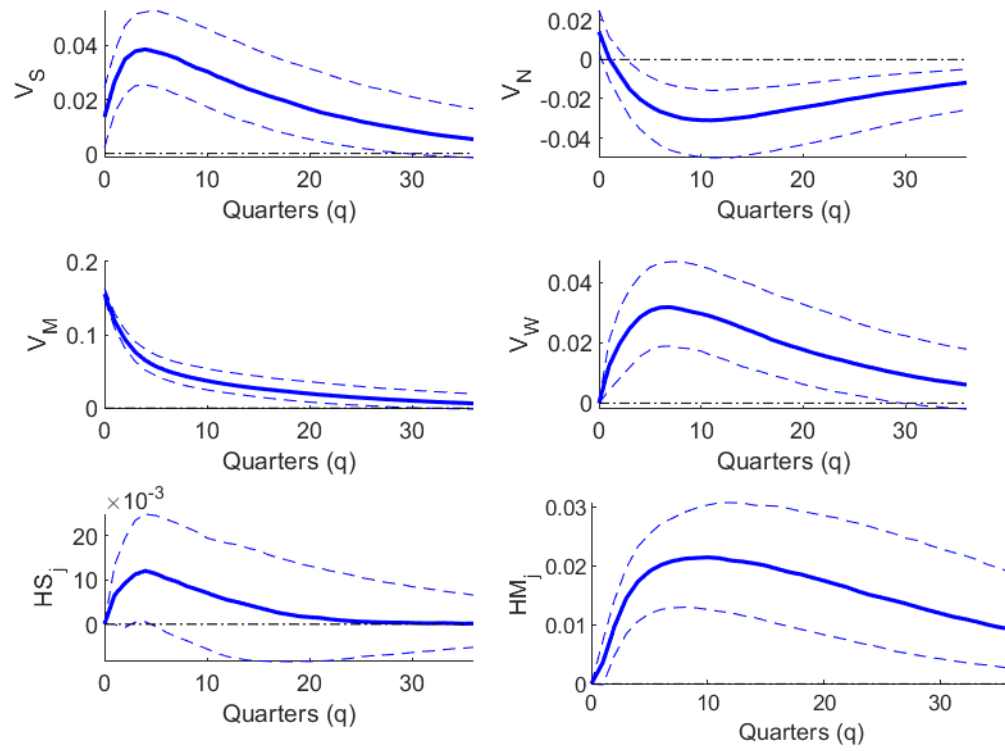
Figure B4

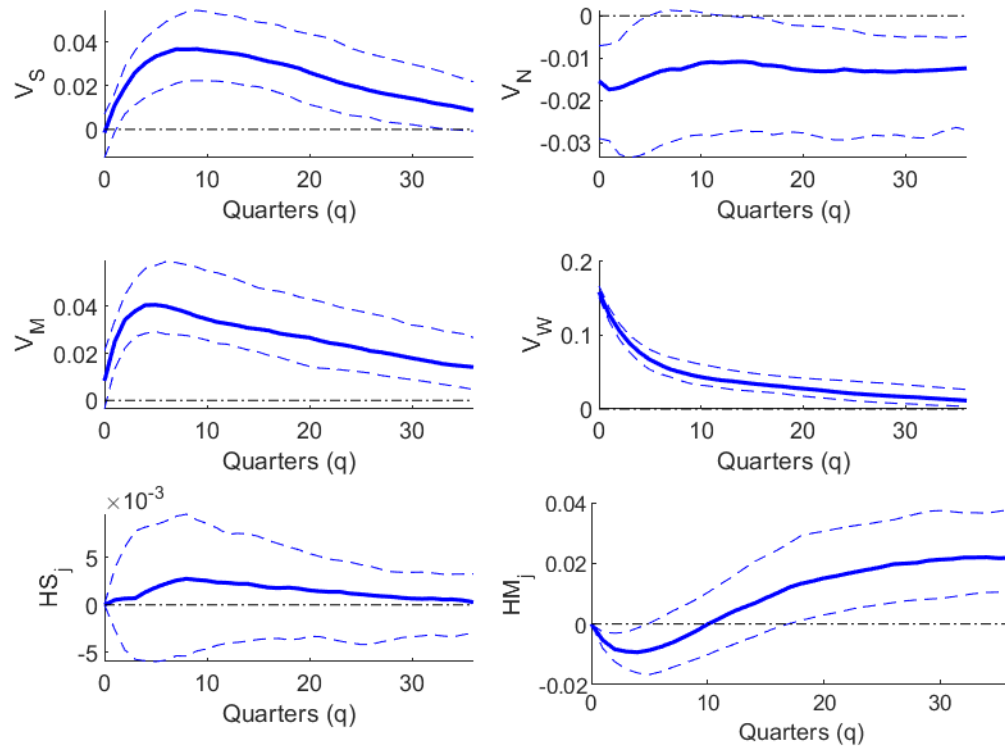**West MS-VAR Generalized Impulse Response Function**

Figure B5

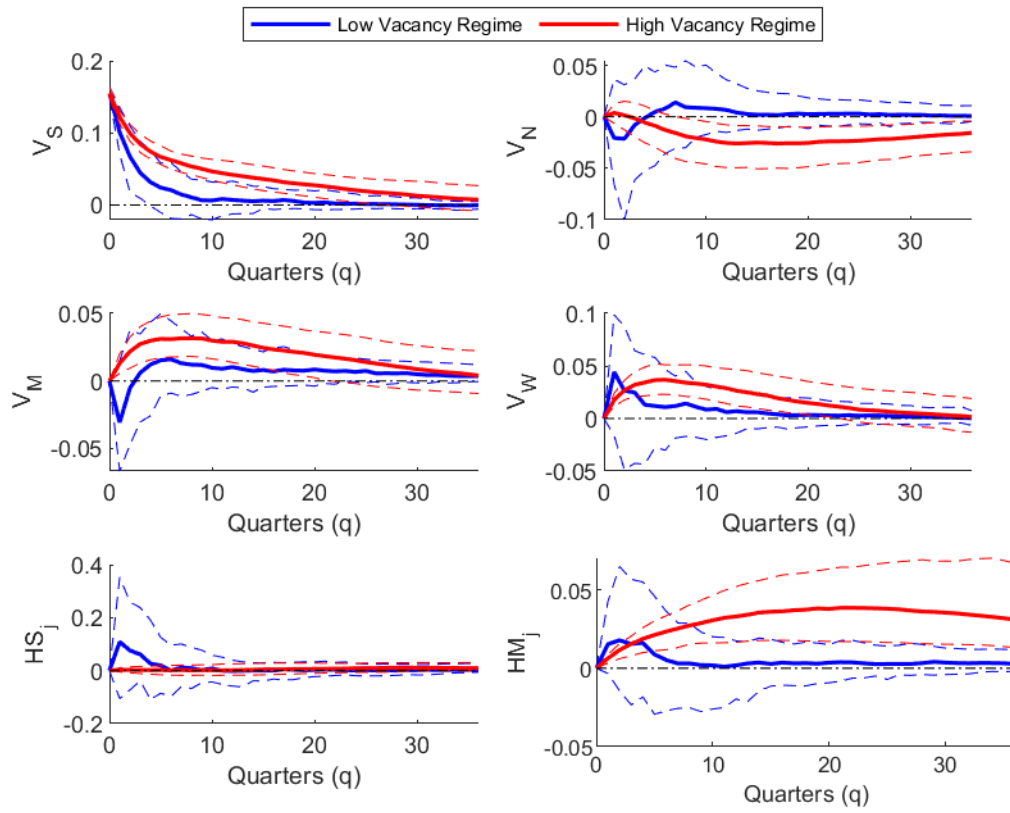**South MS-VAR State-Dependent Generalized Impulse Response Function**

Figure B6

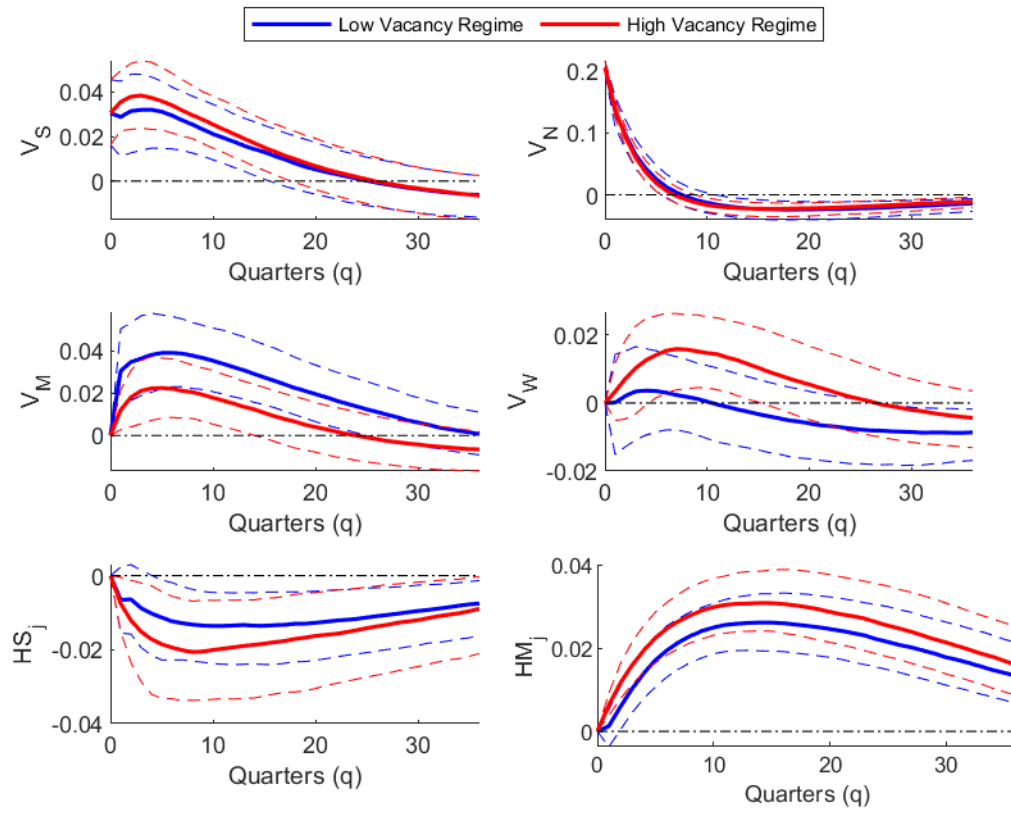**Northeast MS-VAR State-Dependent Generalized Impulse Response Function**

Figure B7

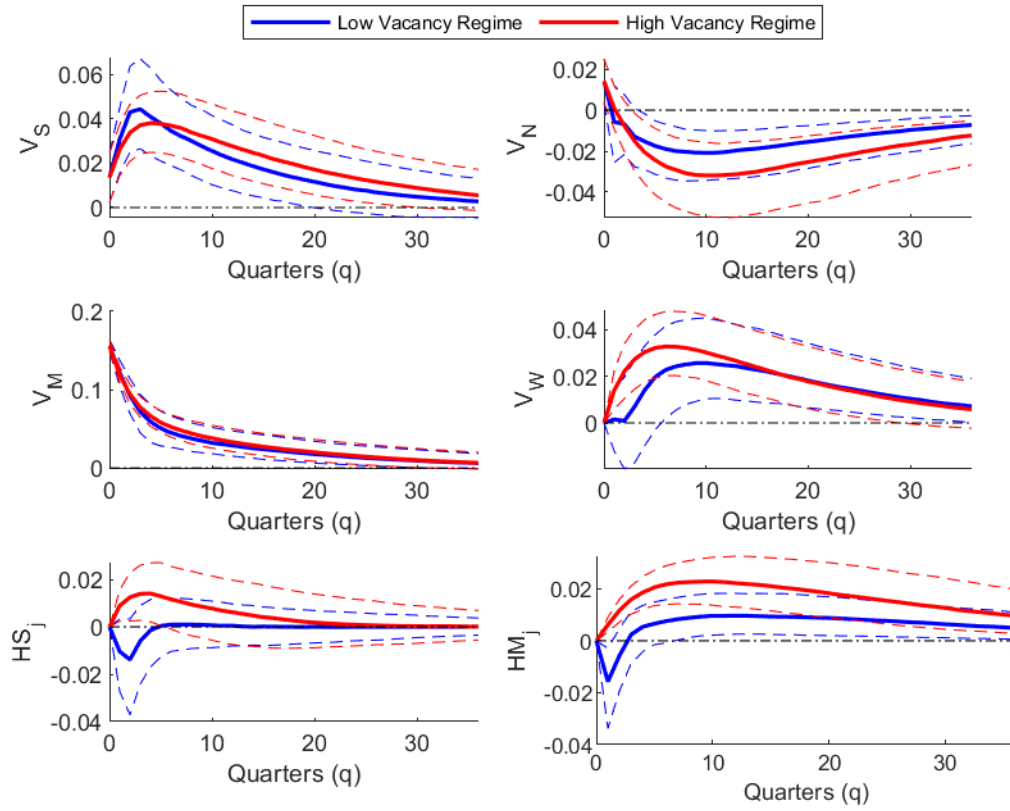**Midwest MS-VAR State-Dependent Generalized Impulse Response Function**

Figure B8

West MS-VAR State-Dependent Generalized Impulse Response Function