What Happens in Vegas Doesn't Always Stay in Vegas: The Dynamics of House Prices and Foreclosure Rates Across Space and Time *

Hua Kiefer†[1], Leonard C. Kiefer‡[2], and Diana Wei§[3]

[1]Federal Deposit Insurance Corporation
[2]Freddie Mac
[3]Office of the Comptroller of the Currency

May 17, 2019

Abstract

This paper identifies novel instruments for house prices and foreclosure rates, and use Dynamic Spatial Simultaneous Equation system (DSSES) to investigate the causal impact of each variables. Our results show that there is an economically significant impact of foreclosure rates on house prices and vice versa. Shocks to the foreclosure rate in one state not only impact house prices in that state, but also the foreclosure rate and house prices in nearby states, and ripple across the country. When it comes to the housing market, what happens in Vegas doesn't always stay in Vegas. We estimate that a one standard deviation foreclosure shock leads to a 2 percent decline in real house prices over the long run. These results provide evidence that could be useful for policymakers evaluating the effectiveness of foreclosure mitigation programs at a national level.

JEL Classification: C33, C36, R31

Keywords: Simultaneous Equation, Dynamic Panel, Spatial Spillover, House Price, Foreclosure

*The views expressed in this paper are those of the authors alone and do not reflect those of The Federal Deposit Insurance Corporation, the Office of the Comptroller of the Currency, and Freddie Mac.

†Email: HKiefer@fdic.gov

‡Email: Leonard_Kiefer@freddiemac.com

§Email: Diana.Wei@occ.treas.gov

1 Introduction

In the aftermath of the global financial crisis, the U.S. housing market experienced its worst years since the Great Depression. State average home prices declined by as much as 30 percent year-over-year and foreclosure start rates rose from an average under 0.5 percentage points prior to the crisis to a high of over 3 percentage points in Nevada. The U.S. government took unprecedented steps to stabilize the housing market and the macroeconomy. Besides the quantitative easing, Congress passed the Housing and Economic Recovery Act of 2008 (HERA), and Obama Administration launched the Home Affordable Refinance Program (HARP) and Home Affordable Modification Program (HAMP).[1]

These government programs were built based on the belief that subsidizing the housing market would help stabilize the U.S. housing market at large and that reducing foreclosures will help to stabilize house prices. However, empirical evidence to date on this question is limited with respect to the magnitude of the dynamic relationship between house prices and foreclosures. Although it is widely accepted in the literature that foreclosures influence house prices mostly through two channels: the disamenity effect of foreclosed properties, and the fire sale-induced supply effect, few studies focus on quantifying the aggregate effect of foreclosures on house prices in a macro setting.

We find two relevant studies examining the impact of foreclosures on house prices at the state level. However, the findings from these two studies differ in the estimate magnitude of foreclosure rates on house prices. Mian et al., 2015 (referred to as MST thereafter) find foreclosures lead to a large decline in house prices during the foreclosure peak following the recent crisis. Specifically, a 4.3 percentage points of increase in foreclosure rate (i.e., a one standard deviation of foreclosures per homeowner during the period of 2008-2009) leads to a 8% to 12% relative drop in house price growth over a nine-quarters time horizon between 2007Q4 and 2010Q1. Calomiris et al., 2013 (referred to as CLM thereafter) find a much smaller impact of foreclosures on house prices based on their PVAR results using data from 1981 through 2009: a foreclosure shock that results in a two-year increase in foreclosure rate of 4.3 percentage points leads to a nine-quarter cumulative decline of house prices of 2.7%. CLM compare the long-term (over 6 years) impact of prices on

[1] The Federal Reserve purchased approximately $1.8 trillion of longer-term agency Mortgage Backed Securities and agency debt to help lower mortgage interest rates. Congress passed HERA to appropriate considerable funds to help stabilize the US housing market including substantial tax credits to first time home buyers which were later extended to include all home buyers. HERA also initiated the Neighborhood Stabilization Program (NSP) to stabilize communities that suffered from foreclosures and abandonment, which was later renewed by American Recovery and Reinvestment Act of 2009 (ARRA). In early 2009 the Obama Administration launched the Making Home Affordable program that was aimed to avoid foreclosures by having Freddie Mac and Fannie Mae refinance or modify mortgages so that homeowners can avoid delinquency and foreclosure - HARP and HAMP. Later, the Federal Housing Finance Agency (FHFA), which oversees Freddie Mac and Fannie Mae, extended the deadline of HARP to the year end of 2018, and HAMP to the end of 2016 with a similar program (Flex Modification) following on the heels of the expiring HAMP.

foreclosures against that of foreclosures on prices and find that price shocks are 79% larger than foreclosure shocks. Based on the dominant effect coming from price shocks instead foreclosure shocks, CLM also suggest that the entire effect of foreclosures on house prices can be explained by the firesale-induced increase of supply. We note that although the models in both MST and CLM allow for the interactions between house prices and foreclosure, they isolate each state and restrict interactions within states. However, if house prices or foreclosures are spatially correlated, such restrictions ignore the amplification due to spillover effects and therefore underestimate the impact of foreclosure rates on house and vice versa.

In this paper, we provide causal evidence on house price and foreclosure under a dynamic panel framework - the Dynamic Spatial Simultaneous Equations System (DSSES). Our structural approach allows for spatial interaction and joint movement of house prices and foreclosures. The reinforcement between the spatial spillover mechanism and the joint movement behavior provides an amplification channel for the transmission of a shock through both the time and space dimensions.

The ideal instruments that help to disentangle the causal effect between house prices and foreclosure are those correlated with only one but not both of them. In this paper, we introduce two novel instruments meeting such criteria: the adjustable rate mortgages (ARMs) reset rate and the change in the natural population growth rate.

The ARM reset rate is defined as the proportion of mortgages within a state that that experience a payment shock as adjustable rate mortgages reach the expiration of their introductory rate and encounter an upward interest rate reset.[2] This indicator serves as an instrument for the foreclosure rate. Higher monthly payments resulting from an ARM interest rate can lead to more foreclosures. Financial institution regulators' interagency policies, such as the OCC Bulletin 2007-14,[3] and former Chair of the Federal Deposit Insurance Corporation, Sheila Bair's Congressional testimony [4] concern the negative impact of ARM reset and encouraged mortgage services to take advantage various government programs and work constructively with borrowers subject to ARM reset risk. The high correlation of foreclosures and payment shock suggests ARM reset as a natural instrument for foreclosures in studying the causal effect of foreclosures and house prices. A reset

[2] We found two empirical papers utilizing a similar concept in identifying the linkage between foreclosures and house prices. Both Gupta, 2018 and Makridis and Ohlrogge, 2018 approach the causal effect measurement issue from a micro standpoint by exploiting the loan-level interest rate variation due to an ARM reset in their two stage estimations. Our work differs from these two studies in the way that our ARM reset instrument measures the proportion of adjustable rate mortgages within a state that experience an upward change of their payment due to ARM reset.

[3] OCC Bulletin 2007-14, Statement on Working with Mortgage Borrowers, https://www2.occ.gov/news-issuances/bulletins/2007/bulletin-2007-14.html .

[4] Statement of Sheila C. Bair, Chairman, Federal Deposit Insurance Corporation on Accelerating Loan Modifications, Improving Foreclosure Prevention and Enhancing Enforcement before the Financial Services Committee; U.S. House Of Representatives; 2128 Rayburn House Office Building December 6, 2007, https://www.fdic.gov/news/news/speeches/archives/2007/chairman/spdec0607.html .

of an ARM is determined by its contractual terms at the time of origination. For example, the rate of a 3/27 ARM will be reset 3 years since its origination, and a 2/28 ARM will be reset 2 years since its origination. Also, whether the subsequent payment of an ARM after a reset will go up or down is determined by the previous introductory rate, the prevailing market rate which the mortgage is indexed to, and the margin specified in the mortgage contract. It is hard to expect an ARM reset is correlated with current house price shocks, since the tenor of introductory rate is usually selected by the borrower based on their forecast of their income change. It is also hard to expect the subsequent payment after a reset will go up or down at the initiation of the contract, unless we believe mortgage issuers can perfectly foresee mortgage rate paths.

To instrument houses prices, we use an indicator of housing demand: the change in the natural population growth rate defined as the quarterly change in births minus deaths divided by the population in the state. This indicator captures housing demand shocks through population growth that is less likely to be correlated with house price shocks than the pure population growth rate. The population growth rate in a state is affected by migration patterns, which themselves partially reflect economic trends, especially job growth. While fertility and mortality are themselves affected by economic conditions the impact is much smaller and with a delay thus we argue below that the natural population growth is a good instrument for housing demand.

There are three sources of endogeneity in our DSSES: 1). the endogenous own time-lagged effect after Helmert's transformation;[5] 2). the endogenous joint movement of house price and foreclosure; and 3). the endogenous spillover effect of home prices and foreclosures respectively. To address these multiple sources of endogeneity residing in our model, we adapt the finite moments instrument variable (FMIV) method designed for a single equation setting in Lee and Yu, 2014 to our simultaneous equations system setup and apply Yang and Lee, 2018's three stage least squares (3SLS) estimator[6] to handle the additional complication from our simultaneous equation system. Each of our structural equations in the simultaneous equations system satisfies the sufficient and necessary rank and order conditions specified in Yang and Lee, 2018's Proposition 1,[7] suggesting

[5]To take care of location and time fixed-effects in a dynamic panel model - the dependent variable is a function of its own time-lagged term, we adopt Helmert's transformation for all the variables in both sides of our equation. The transformed own time-lagged term becomes correlated with the residual and thus can no longer be treated as a pre-determined variable. See Section 3 for more details.

[6]The consistency of the quasi maximum likelihood method in Yang and Lee, 2018 relies on the assumption of large T to avoid handling the initial observation problem, and its asymptotic distribution depends on the growth rate of N and T (i.e., $(\mathsf{N}-1)/\mathsf{T}^3 \to 0$). Because the T in our empirical analysis is small relatively to N, we choose the 3SLS over the QML for this paper.

[7]Yang and Lee, 2018 articulate the sufficient and necessary conditions for identifying the coefficients of each structural equation in their Proposition 1. The rank condition requires the rank of a matrix representing all exclusive coefficient restrictions for the given equation to be equal to the total number of equations in the system minus 1. And the order condition requires that the number of all the excluded parameters of the given equation is no less than the total number of equations in the system minus 1. It is worth noting, though spatial lag and FOD-transformed

all the structural parameters in our system are identifiable. Our estimated results also meet their stable condition,[8] indicating that our model is stable in both space and time dimensions. This finite moments IV-based estimator allows us to arrive at a consistent estimator in the presence of multiple sources of endogeneity.

Our estimated result of a statistically significant and economically meaningful foreclosure externality on house price when the multi-dimensional endogeneity issues are not adequately controlled. Shocks to the foreclosure rate in one state not only impacts house prices in that state, but also the foreclosure rate and house prices in nearby states. When it comes to the housing market, what happens in Vegas doesn't always stay in Vegas. Our DSSES model estimation results show that a one standard deviation of foreclosure shock leads to a short-run real house price decline of 1.6 percent and a 2 percent decline in real house prices over the long run. A one standard deviation shock to real house prices lowers the foreclosure rate 13 percent in the short run. We also find significant spatial spillovers in both house prices and foreclosure rates across states. For example, four quarters after a one standard deviation shock to Nevada's foreclosure rate, real house prices in California experience a cumulative decline of 1 percent.

2 Literature Review

This paper relates to several streams of literature. First, we contribute to the literature of spatial spillovers. Neighborhood spatial spillovers have been explored extensively in the recent urban and real estate economics literature. Nevertheless, most of this spillover studies focus on house price (see, among others, Se Can and Megbolugbe, 1997; Basu and Thibodeau, 1998; Pace et al., 1998; LeSage and Pace, 2004; Clauretie and Daneshvary, 2009; Kiefer, 2011). These studies suggest the existence of spatial interdependence of house prices in a neighborhood. Anselin, 2008; Fingleton and Le Gallo, 2008 provide their justifications for the spatially correlated house prices, and refer to an omission of spatially autocorrelated regressors, and displaced demand and supply effects. Omitted variable issue is almost inevitable in modeling house price due to the uniqueness of location, which is an important determinant of property value. If there exist some common characteristics in the neighborhood that influence housing value, but omitted from the model specification (e.g., accessibility of shopping centers/parks, exposure to traffic noise, major employment center, etc.), prices of nearby houses tend to serve as proxies of these omitted neighborhood effects. The recognition of the proxy effect

own time lag are endogenous from a statistical perspective, it is appropriate to regard both of them as "exogenous" variables for the purpose of counting excluded variables in a structural model. Therefore, extending a first order time lag equation as specified in Yang and Lee, 2018 to a more general specification as our DSSES with p time lags ($p \geqslant 2$) will not change our identification conclusion.

[8] In order to have an nonexplosive system either in time and space, the aggregated effects of spatial spillovers, serial correlations, and cross effects cannot be too large. More specifically, the sufficient (not necessary) condition defined by Yang and Lee, 2018 requires the total effect to be less than 1 for row-normalized spatial weight matrices.

has another piece of empirical evidence - residential appraisals often rely on sales comparables to determine the value of a property in a housing transaction. The widely used Sales Comparison Approach (SCA) method is an appraisal procedure that is essentially a weighted average of sales comparables in the vicinity. The argument of displaced demand and supply effects is also intuitive. High prices discourage demand. If the price in a local area is high, the quantity demanded there should decrease, and thus demand from this location will be displaced into a nearby location suggesting a positive relationship between the local demand and nearby prices. The same argument holds for the supply function: a high price in a local area will attract the supply from nearby places suggesting a negative relationship between the local supply and the nearby prices. In equilibrium where demand equals supply, the displaced demand and supply effects translate into a positive spatial relationship of nearby house prices.

Studies on spillovers of mortgage defaults and foreclosures are rather limited. The existing works in this line of research show that the default risk of a borrower is affected by the household/loan characteristics of surrounding properties (see, among others, Goodstein et al., 2011; Agarwal et al., 2012; Zhu and Pace, 2014). However, the models considered in the aforementioned works have not yet established a direct link connecting default decisions of neighboring homeowners to fully capture the spillover effect of mortgage defaults, and thus underestimate the impact. Towe and Lawley, 2013 establish the connection between a homeowner's default decision and her observations of neighbors' time-lagged default decisions and explain it as a result of social interaction behaviors. Chomsisengphet et al., 2018 further this line of work by providing empirical evidence of spatial spillovers in homeowners' mortgage default decisions in forms of both time-lagged and contemporaneous effects.

It is worth mentioning that the spillover literature discussed above emphasizes the spatial neighborhood effect is a local phenomenon. The neighborhood of the spatial interactions among house prices or foreclosures is usually defined as a planned community within a city (e.g, Clauretie and Daneshvary, 2009), a county (e.g., Se Can and Megbolugbe, 1997; Towe and Lawley, 2013; Chomsisengphet et al., 2018; LeSage and Pace, 2004; Kiefer, 2011) or a Metropolitan Statistical Area (e.g., Basu and Thibodeau, 1998). The "local" characteristic of price or foreclosure spillovers makes sense because housing properties are closely tied to their locations. However, we are interested in testing whether this type of local neighborhood effects appear in our aggregated state level data as a collective outcome.

However, it is interesting to examine whether these local spillovers will be filtered out entirely when data is aggregated to a higher geographic level. In other words, are these spatial spillover effects still observable at the state level? Intuitively, the within-unit joint movement of house prices and foreclosures reinforces each other (i.e., unit i's house price move together with unit i's foreclosure) and helps fuel the ripple effects of house prices and foreclosures, and could result in a nontrivial impact received by a spatial unit far from the origin of a shock. Meanwhile, the spatial spillover transmission mechanism provides an amplification channel for the price response to a foreclosure shock within the same spatial unit such as a state.

By introducing the micro-based spillover literature to the limited studies on the quantification of the causal effect between foreclosures and house prices at the state level, our results shed light on the extent to which spatial spillovers impact the joint movement of house prices and foreclosures and potentially motivate the wide range of federal programs targeting stressed borrowers.

In addition, our work contributes to a growing literature on the causal linkage between house prices and foreclosures. Accurately measuring the price impact of foreclosures is not a simple task due to the reverse causality issue - house price can affect foreclosure, and it is likely true vice versa. Strong instruments as we propose in this paper - ARMs reset rate and change in the natural population growth rate - are useful to overcome this identification difficulty.

The causality direction can move from house prices to foreclosures. Foster and Van Order, 1984's option-based model suggests that a put option (i.e., mortgage default) is in the money when house prices fall. Falling house prices shrink homeowners' equity. For homeowners with thin equity to start with (i.e., low or no down payment mortgages), a sharp price drop like what we have seen in the recent crisis can easily push borrowers underwater (e.g., negative equity), which is a necessary (but not sufficient) condition for mortgage default. While literature focusing on the default impact of declining house prices is still immature (especially in terms of accurately quantifying the price externality), the existing studies have suggested a significant role played by house price changes on homeowners' default decisions (e.g., among others, Bajari et al., 2008; Foote et al., 2008; Guiso et al., 2013).

The opposite causality direction is equally plausible. Foreclosed properties are usually sold with a foreclosure discount accounting for either the below average physical condition, or the stigma effect - simply because these properties have been involved in foreclosure proceedings. The literature has suggested a foreclosure discount rate of 20 percent or more (e.g., among others, Clauretie and Daneshvary, 2009; Carroll et al., 1997; Harding et al., 2009; Campbell et al., 2011). Besides the self-discount of foreclosed properties themselves, negative price impacts of these distressed properties are found in forms of externality: through either a disamenity channel - deferred maintenance or attracting crime due to vacancy; or a supply channel - foreclosed properties add to the local house inventory available for sale. Both effects put downward pressure on nearby house prices.

A large literature focuses on disentangling these two types of underlying mechanisms as well as examining the magnitude of the foreclosure externality on house prices. Recent studies have particularly emphasized the importance of controlling for the reverse causality in order to accurately measure the price impact of foreclosures. For example, to take care of the simultaneity issue, Harding et al., 2009 adopt repeat sales approach, Campbell et al., 2011, and Hartley, 2014 use a difference in difference identification strategy in their hedonic estimations, Mian et al., 2015 employ an instrument variable capturing the differences in state foreclosure laws (i.e., judicial vs.

non-judicial), and Gerardi et al., 2015 include triple-interaction fixed effects (i.e., times of initial and subsequent sales and geographic location) in their repeat sales specification. The consensus in the literature suggests that the foreclosure spillovers on nearby house prices are less than 2 percent after properly controlling for the simultaneity issue in the estimation procedure.[9]

These empirical studies have so far focused on "controlling" for the reverse causality of house price on foreclosure by the inclusion and exclusion of variables representing housing quality and housing supply. Though the potential of a simultaneous move of house price and foreclosure is widely recognized, few studies have focused explicitly on modeling the co-movement pattern. The exception is CLM. In their study, the simultaneous relationship between house prices and foreclosures is examined and the magnitudes of price externality and foreclosure externality are compared. They employ a 5-equation PVAR consisting of price appreciation and foreclosure rate in addition to three macroeconomic indicators - growth rates of employment, permits, and home sales. Their findings suggest that the causality indeed exists in both directions. However, the cumulative impact (over six years) that prices have on foreclosures is 79% larger than the impact of foreclosures on prices. As a result, the strong connection between house price and foreclosure mainly reflect the house price impact on foreclosure activity rather than the other way around. By allowing for spatial autocorrelation and contemporaneous interactions between house price and foreclosure, our empirical results of long run analysis suggest the cumulative response to a standardized shock is only 36% larger for house prices on foreclosure than for foreclosure on house prices, which stands in contrast to CLM's claim of 79%. A shock to the foreclosure equation in our estimated DSSES that increases the foreclosure rate 1 standard deviation after eight quarters decreases real house prices 7.9 percent over that same period, which is in line with MST's findings of 8% to 12% over 9 quarters in response to a similar shock.

The rest of the paper is organized as follows. Section 3 presents two alternative econometric specifications, the PVAR and the DSSES, and their estimation methodologies. Section 4 describes the data and summary statistics and discusses our instrumental variables. Section 5 discusses our empirical results. Section 6 concludes.

3 Econometric Model

In this section, we start with a Panel Vector Autoregression (PVAR) specification commonly used in dynamic panel studies. Then we present a dynamic spatial simultaneous equations system (DSSES). Our DSSES differs from the conventional panel VAR approaches of house price and foreclosure panel data modeling in two ways: the simultaneous equations system setup allows for a *simultaneous-cross effect* between house price and foreclosure; the spatial lags in the system intro-

[9] Due to its relatively small effect, we leave out the cross spillovers between house prices and foreclosures across spatial units from our already complicated model specification.

duce a *contemporaneous-spillover effect* of house price and foreclosure. The DSSES specification emphasizes the amplification mechanism arising from the intertwined cross-effect dynamics and spatial spillovers of house prices and foreclosures.

3.1 Panel VAR

As the way Canova and Ciccarelli, 2013 put it: "PVARs have the same structure as VAR models, in the sense that all variables are assumed to be endogenous and interdependent, but a cross sectional dimension is added to the representation". Our PVAR system consists of m equations with house prices and foreclosures as two of the m endogenous variables. The model also includes the individual time lags and location fixed effects, as well as the interaction of the endogenous variables but with time lags - the *time-lagged-cross effects*. Let n denote the total number of spatial units, and T denote the total number time periods, the PVAR system at time period $t(\forall t = 1, 2, \cdots, T-p)$ can be written as

$$Y^*_{n2}(t) = \sum_{j=1}^{p} Y^*_{n2}(t-j)P_j + d' \otimes l_n + C + U^*_{n2}(t), \tag{1a}$$

and the lth equation $(\forall l = 1, 2, \cdots, m)$ in the system is expressed as,

$$y^*_{l,nm}(t) = \sum_{j=1}^{p} Y^*_{nm}(t-j)\rho_{j,\cdot l} + d_l \otimes l_n + c_{\cdot l} + u^*_{l,nm}(t), \tag{1b}$$

The dependent variable, $Y^*_{nm}(t) = [y^*_{1,nm}(t), y^*_{2,nm}(t), \cdots, y^*_{m,nm}(t)]$, with each column representing one endogenous variable (e.g., house prices, foreclosures, etc.), is an $n \times m$ matrix. Similarly, $Y^*_{nm}(t-j) = [y^*_{1,nm}(t-j), y^*_{2,nm}(t-j), \cdots, y^*_{m,nm}(t-j)]$ is an $n \times m$ matrix representing the time-lagged dependent variables to the jth order. $U^*_{nm}(t) = [u^*_{1,nm}(t), u^*_{2,nm}(t), \cdots, u^*_{m,nm}(t)]$ is the disturbance term. We assume the errors are i.i.d. across space and time.[10] P_j is a $m \times m$ matrix with the diagonal elements capturing the own time-lagged effect from j periods ago, and the off-diagonal elements denoting the cross time-lagged effect from j periods ago. We use $\rho_{j,\cdot l}$ to denote the lth column of P_j. C and d are, respectively, an $n \times m$ matrix of location fixed effects,[11] and a m-dimensional column vector of intercepts, while l_n is an $n \times 1$ vector of ones and $\otimes$ denotes the kronecker product. $c_{\cdot l}$ represents the lth column of C and d_l is the lth element of d.

It is well known that, in a dynamic panel, the fixed effects estimator is not consistent because they are correlated with the regressors due to lags of the dependent variables. We apply the forward orthogonal difference (FOD) transformation (i.e., Helmert's transformation) to each variable input of Equation (1) to remove the fixed effects.[12] The FOD transformation eliminates both the intercept

[10] PVAR estimation usually doesn't require zero correlations across equations.

[11] Note, to avoid perfect collinearity, we impose a normalization condition of $\sum_{i=1}^{n} c_{1,i} = 0$.

[12] See Appendix(A) for details of the operation of FOD.

and location fixed effects from Equation (1) and reduced the total observation from $mn(T-p)$ to $mn(T-p-1)$, let $\mathcal{T} = T-p-1$ to simplify the notation, we now have

$$Y_{nm,\mathcal{T}} = \sum_{j=1}^{p} Y_{nm,\mathcal{T}}^{(-j)} P_j + U_{nm,\mathcal{T}}, \tag{2}$$

where the superscript $(-j)$ of $Y_{nm,\mathcal{T}}^{(-j)}$ indicates the value of the variable is lagged by j periods. To consistently estimate Equation (2), restrictions are typically imposed on the coefficient matrices P_js to make the variance of $Y_{nm,\mathcal{T}}$ bounded and to make sure that P_js exists.

3.2 Dynamic Spatial Simultaneous Equations System

The previously described PVAR specification takes into account the interactions of house prices and foreclosures, however, it imposes the cross effect happens with time lags. There are reasons to believe, however, that this assumptions may be unrealistic. For example, if foreclosure rates respond quickly to price changes, our house price measure and foreclosure measure may be simultaneously determined. To allow for such possibilities, we need a system allowing for simultaneously determined dependent variables - *simultaneous-cross effect*. We introduce a dynamic spatial simultaneous equations system (DSSES) in this section to take care of not only the simultaneous-cross effect but also the *contemporaneous-spillover effect* to fully account for the interactive dynamics of house prices and foreclosures and their spatial transmission mechanism. More specifically, we adopt the format of the widely used spatial autoregressive (SAR) model from Cliff and Ord, 1973 in each time period and for the house price and foreclosure equations to capture the contemporaneous-spillover effect. And these two panel SAR models are then built into a simultaneous equations system to allow for the simultaneous-cross effect. The FOD transformed transformed functions consist of both types of simultaneity and have the form of

$$y_{1,i}(t) = -\gamma_{12} y_{2,i}(t) + \psi_{11} W_n Y_{1,n2}(t) + \sum_{j=1}^{p} \rho_{j,11} y_{1,i}(t-j) + \sum_{j=1}^{p} \rho_{j,12} y_{2,i}(t-j) + x'_{1,i}(t)\pi_{\cdot 1} + u_{1,i}(t), \tag{3a}$$

and

$$y_{2,i}(t) = -\gamma_{21} y_{1,i}(t) + \psi_{22} W_n Y_{2,n2}(t) + \sum_{j=1}^{p} \rho_{j,22} y_{2,i}(t-j) + \sum_{j=1}^{p} \rho_{j,21} y_{1,i}(t-j) + x'_{2,i}(t)\pi_{\cdot 2} + u_{2,i}(t), \tag{3b}$$

for $t = 1, 2, \cdots, \mathcal{T}$ and $i = 1, 2, \cdots, n$. In Equation (3a) (i.e., the house price equation), $y_{1,i}(t)$ is the dependent variables, $y_{1,i}(t-j)$ is the own time lag to the jth order with a scalar coefficient $\rho_{j,11}$ denoting the own time-lagged effect, $y_{2,i}(t-j)$ is the cross time lag to the jth order with $\rho_{j,12}$ denoting the cross time-lagged effect, $y_{2,i}(t)$ represents the simultaneous-cross effect - the effect from the dependent variable of Equation (3b), -γ_{12} and ψ_{11} denote the scalar coefficients, and $x_{1,i}(t)$ is the k_1- dimensional vector of control variables with corresponding parameter vector $\pi_{\cdot 1}$.[13] The error terms across equations, $u_{1,i}(t)$ and $u_{2,i}(t)$ are no longer uncorrelated due to the simultaneous-cross effect. We assume that the disturbances within each equation are still i.i.d., and the cross-equation disturbances follow a conventional correlation structure as

$$E(u_{m,i}(t)u_{l,j}(s)) = \{ \begin{array}{c} 0 \text{ if } i \neq j \text{ or } t \neq s \\ \sigma_{ml} \text{ if } i = j \text{ and } t = s \end{array} .$$

Alternatively, after stacking observations over space is ($\forall i = 1, \cdots, n$), we can write both the house price and foreclosure equations into a system as

$$Y_{n2}(t)\Gamma = W_n Y_{n2}(t)\Psi + \sum_{j=1}^{p} Y_{n2}(t-j)P_j + X_n(t)\Pi + U_{n2}(t), \tag{4a}$$

and the lth equation ($\forall l = 1, 2$) in the system is expressed as,

$$y_{l,n2}(t) = -Y_{n2}(t)\gamma_{\cdot l} + W_n Y_{n2}(t)\psi_{\cdot l} + \sum_{j=1}^{p} Y_{n2}(t-j)\rho_{j,\cdot l} + X_{l,n}(t)\pi_{\cdot l} + u_{l,n2}(t), \tag{4b}$$

for $t = 1, 2, ..., \mathcal{T}$. The dependent variable, $Y_{n2}(t) = [y_{1,n2}(t), y_{2,n2}(t)]$, with the first column, $y_{1,n2}(t) = [y_{1,1}(t), \cdots, y_{1,n}(t)]'$ representing the FOD-transformed dependent variable in the house price equation, and the second column, $y_{2,n2}(t) = [y_{2,1}(t), \cdots, y_{2,n}(t)]'$ representing the FOD-transformed dependent variable in the foreclosure equation, is an $n \times 2$ matrix. Γ is a 2×2 matrix with ones at the main diagonal, and its off-diagonal elements capture the simultaneous-cross effect (in a negative term). W_n is a time-invariant $n \times n$ weight matrix of known constants,[14] whose ijth entry is w_{ij}, and Ψ is the corresponding spatial autoregressive coefficient matrix with zeros off-diagonal elements. $X_n(t) = X_{1,n}(t) \cup X_{2,n}(t)$, is the FOD-transformed exogenous variable with an dimension of $n \times k$ (with x variables appearing in both equations counted for once only to avoid perfect multicollinearity, so $k \leqslant (k_1 + k_2)$), and $U_{n2}(t) = [u_{1,n2}(t), u_{2,n2}(t)]$ is the FOD-transformed disturbance term with a covariance matrix of $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$. Π is a $k \times 2$ coefficient matrix for exogenous regressors, and P_j is a 2×2 matrix for $j = 1, ...p$, with the diagonal elements capturing the own time-lagged effects, and off-diagonal elements capturing the time-lagged cross effects. If $j \geqslant 2$, Equation (4b) is a high order dynamic model. We use -$\gamma_{\cdot l}$, $\psi_{\cdot l}$, $\rho_{j,\cdot l}$, and $\pi_{\cdot l}$ to denote the lth column of the corresponding parameter matrices as before.

[13] Instead of a total number of m equations in the PVAR system, the DSSES consists of two equations only: the house price equation and the foreclosure equation. It also includes xs as exogenous variables.

[14] We are assuming that the system only involves one weights matrix. This assumption is made for ease of presentation, but also seems to be the typical specification in applied work. Our results can be generalized in a straight forward way to the case in which each spatially lagged variable depends upon a weights matrix which is unique to that variable.

Equation (4a) falls under the general form of spatial dynamic panel simultaneous equations models described in Yang and Lee, 2018.[15] Besides the previously discussed endogeneity arriving from the FOD-transformed own time lag (i.e., represented by the nonzero diagonal elements of P_j) as in the PVAR, there are two new sources of endogeneity: the simultaneous-cross effect represented by the nonzero off-diagonal elements in Γ, and the contemporaneous-spillover effect represented by the nonzero diagonal elements of Ψ(i.e., often referred to as spatial lag in the spatial literature). Valid IVs for he FOD-transformed own time lag, $y_{l,n2}(t-j)$, suggested in the dynamic panel literature include exogenous variables from all the time periods (i.e., $X_n(t)$ $\forall t = 1, 2, \cdots, \mathcal{T}$), and own untransformed terms from the current time period and all the previous time periods (i.e., $y^*_{l,i}(t-j)$, $y^*_{l,i}(t-j-1), \cdots$).[16] The number of feasible IVs is a function of t, and therefore the dimension of corresponding IV matrix increases with t. A large number of IVs constructed from all available time lags is beneficial, in principle, in terms of improving the asymptotic efficiency of the IV estimator. However, when $\mathcal{T}$ gets large, the many IVs issue occurs: many IVs decrease the variances of the IV estimator, but increase its bias. Considering the bias and variance trade-offs and the benefit of maintaining a constant number of IVs across ts for an easy construction of the IV matrix, we adopt Lee and Yu, 2014's finite moments instrument variable (FMIV) approach [17] by limiting the number of time lags on the exogenous variables as well as the own untransformed term.

The non-spatial simultaneous equations system literature suggests the use of exogenous variables excluded from the lth equation and the exogenous variables included in the lth equation as instruments to take care of the endogeneneity arising from the simultaneous movement of multiple equations. Meanwhile, common IVs suggested in the spatial literature for dealing with spatial lags call for first-order and higher-order spatially lagged exogenous variables. We follow Yang and Lee, 2018's IV strategy and construct our IV matrix as

$$\mathcal{G}_n(t) = \begin{bmatrix} Y^*_{n2}(t-p) & W_n Y^*_{n2}(t-p) & W_n^2 Y^*_{n2}(t-p) & X_n(t) & W_n X_n(t) & W_n^2 X_n(t) \end{bmatrix} \tag{5}$$

where the exogenous variables (i.e., $X_n(t)$) and predetermined variables (i.e., $Y^*_{n2}(t-p)$) are raised to first- and second- orders (i.e., pre-multiplied by W_n and W_n^2) to take care of the contemporaneous-

[15]Yang and Lee, 2018's general form allows for a four-ways channel of spatial spillovers: contemporaneous and with a time lag; within-equation and cross-equations. Our DSSES specification leaves out the time-lagged and cross-equation spillovers terms.

[16]Among others, see Arellano and Bond, 1991; Elhorst, 2010; Alvarez and Arellano, 2003; Kelejian and Prucha, 2004; Revelli, 2001 for examples.

[17]Lee and Yu, 2014 first show that the 2SLS estimator based on FMIV approach is consistent and asymptotically normal in Theorem 1 (though less efficient than the GMM estimator proposed in their paper). The set of IVs suggested by Lee and Yu, 2014 consists of both linear and quadratic moments. Due to the complication in our main model specification - the DSSES is a equations system but not a single equation, the quadratic moments become less straight forward. We therefore adopt only the linear moments from Lee and Yu, 2014 instead of pursuing their best IV estimator.

spillover. The IV matrix, $\mathcal{G}_n(t)$, is not equation specific, because $X_n(t) = X_{1,n}(t) \cup X_{2,n}(t)$ represents the complete set of exogenous variables of the system, and $Y^*_{n2}(t-p)$ reflects all the predetermined variables of the system. After stacking observations from ts ($\forall t = 1, 2, \cdots, \mathcal{T}$), the IV can be written into a matrix as

$$\mathcal{G}_{n,\mathcal{T}} = \begin{bmatrix} Y^*_{n2}(1-p) & W_n Y^*_{n2}(1-p) & W_n^2 Y^*_{n2}(1-p) \\ \vdots & \vdots & \vdots \\ Y^*_{n2}(\mathcal{T}-p) & W_n Y^*_{n2}(\mathcal{T}-p) & W_n^2 Y^*_{n2}(\mathcal{T}-p) \end{bmatrix},$$
$$\begin{matrix} X_n(1) & W_n X_n(1) & W_n^2 X_n(1) \\ \vdots & \vdots & \vdots \\ X_n(\mathcal{T}) & W_n X_n(\mathcal{T}) & W_n^2 X_n(\mathcal{T}) \end{matrix}\Bigg]$$

with a dimension of $n\mathcal{T} \times (6 + 6k)$.

We stack observations from all ts ($\forall t = 1, \cdots, \mathcal{T}$) for the lth equation , and let $\mathcal{W}_{n2,\mathcal{T}} = I_2 \otimes I_{\mathcal{T}} \otimes W_n$ for conciseness, Equation (4b) becomes

$$\begin{aligned} y_{l,n2,\mathcal{T}} = {} & -Y_{n2,\mathcal{T}}\gamma_{\cdot l} + \mathcal{W}_{n2,\mathcal{T}} Y_{n2,\mathcal{T}}\psi_{\cdot l} + X_{l,n,\mathcal{T}}\pi_{\cdot l} + \sum_{j=1}^{p} Y^{(-j)}_{n2,\mathcal{T}}\rho_{j,\cdot l} + u_{l,n2,\mathcal{T}}, \\ & \forall l = 1, 2, \end{aligned} \tag{6}$$

where $Y_{n2,\mathcal{T}}$, $\mathcal{W}_{n2,\mathcal{T}} Y_{n2,\mathcal{T}}$, and $Y^{(-j)}_{n2,\mathcal{T}}$ ($\forall j = 1, \cdots, p$) are all endogenous. Let

$$\begin{aligned} Z_{l,n,\mathcal{T}} &= \begin{bmatrix} Y_{n2,\mathcal{T}} & \mathcal{W}_{n2,\mathcal{T}} Y_{n2,\mathcal{T}} & X_{l,n,\mathcal{T}} & Y^{(-1)}_{n2,\mathcal{T}} & \cdots & Y^{(-p)}_{n2,\mathcal{T}} \end{bmatrix}, \text{and} \\ \theta_l &= \begin{bmatrix} -\gamma'_{\cdot l} & \psi'_{\cdot l} & \pi'_{\cdot l} & \rho'_{1,\cdot l} & \cdots & \rho'_{p,\cdot l} \end{bmatrix}', \end{aligned}$$

Equation (6) can be simplified to

$$\begin{aligned} y_{l,n2,\mathcal{T}} &= Z_{l,n,\mathcal{T}}\theta_l + u_{l,n2,\mathcal{T}}, \\ \forall l &= 1, 2. \end{aligned} \tag{7}$$

We first estimate Equation (7) separately for each equation. The 2SLS estimator of the lth equation has the form of

$$\widehat{\theta}^{\mathcal{G}}_{l,2sls} = (Z'_{l,n,\mathcal{T}} P^{\mathcal{G}} Z_{l,n,\mathcal{T}})^{-1} Z'_{l,n,\mathcal{T}} P^{\mathcal{G}} y_{l,n2,\mathcal{T}}$$

where $P^{\mathcal{G}} = \mathcal{G}_{n,\mathcal{T}}(\mathcal{G}'_{n,\mathcal{T}}\mathcal{G}_{n,\mathcal{T}})^{-1}\mathcal{G}'_{n,\mathcal{T}}$ is the projection matrix of the instrument matrix $\mathcal{G}_{n,\mathcal{T}}$. The 2SLS estimator's asymptotic distribution follows

$$\sqrt{n\mathcal{T}}(\widehat{\theta}^{\mathcal{G}}_{l,2sls} - \theta_l) \xrightarrow{d} N\Big(0, \operatorname*{plim}_{n\to\infty} \Big[\frac{\sigma_l}{n\mathcal{T}}(Z'_{l,n,\mathcal{T}} P^{\mathcal{G}} Z_{l,n,\mathcal{T}})\Big]^{-1}\Big),$$

with large n and $\mathcal{T}$.

To further extend the 2SLS estimator to a 3SLS estimator suggested by Yang and Lee, 2018 for an improved estimation efficiency, we stack both equations from Equation (6) and write the system into a vectorized form as

$$y_{n2,\mathcal{T}} = Z_{n,\mathcal{T}}\theta + u_{n2,\mathcal{T}}, \tag{8}$$

where $y_{n2,\mathcal{T}} = \begin{bmatrix} y'_{1,n2,\mathcal{T}} & y'_{2,n2,\mathcal{T}} \end{bmatrix}'$, $Z_{n,\mathcal{T}} = \mathrm{diag}[\ Z_{1,n,\mathcal{T}},\ \ Z_{2,n,\mathcal{T}}\ ] = \begin{bmatrix} Z_{1,n,\mathcal{T}} & 0 \\ 0 & Z_{2,n,\mathcal{T}} \end{bmatrix}$, $\theta = \begin{bmatrix} \theta'_1 & \theta'_2 \end{bmatrix}'$, and $u_{n2,\mathcal{T}} = \begin{bmatrix} u'_{1,n2,\mathcal{T}} & u'_{2,n2,\mathcal{T}} \end{bmatrix}'$. The variance-covariance matrix of the residuals in Equation (8) is of the form, $\Sigma \otimes I_{n,\mathcal{T}}$, with $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$, and $I_{n,\mathcal{T}}$ denoting an $n\mathcal{T} \times n\mathcal{T}$ identify matrix. The 3SLS estimator of the entire equation system is

$$\widehat{\theta}^{\mathcal{G}}_{3sls} = [\widehat{Z}'_{n,\mathcal{T}}(\widehat{\Sigma}^{-1} \otimes I_{n,\mathcal{T}})Z_{n,\mathcal{T}}]^{-1}\widehat{Z}'_{n,\mathcal{T}}(\widehat{\Sigma}^{-1} \otimes I_{n,\mathcal{T}})y_{n2,\mathcal{T}}$$

with $\widehat{Z}_{n,\mathcal{T}} = \mathrm{diag}[\ P^{\mathcal{G}}Z_{1,n,\mathcal{T}},\ \ P^{\mathcal{G}}Z_{2,n,\mathcal{T}}\ ] = \begin{bmatrix} P^{\mathcal{G}}Z_{1,n,\mathcal{T}} & 0 \\ 0 & P^{\mathcal{G}}Z_{2,n,\mathcal{T}} \end{bmatrix}$, and the variance-covariance components of $\widehat{\Sigma} = \begin{bmatrix} \widehat{\sigma}_{11} & \widehat{\sigma}_{12} \\ \widehat{\sigma}_{21} & \widehat{\sigma}_{22} \end{bmatrix}$ can be estimated using a first stage estimator, $\widehat{\theta}^{\mathcal{G}}_{l,2sls}$.[18] Yang and Lee, 2018 note the asymptotic distribution of this estimator is

$$\sqrt{n\mathcal{T}}(\widehat{\theta}^{\mathcal{G}}_{3sls} - \theta) \xrightarrow{d} N\Big(0, \operatorname*{plim}_{n\to\infty} \frac{1}{n\mathcal{T}}[\widehat{Z}'_{n,\mathcal{T}}(\Sigma^{-1} \otimes I_{n,\mathcal{T}})\widehat{Z}_{n,\mathcal{T}}]^{-1})\Big),$$

with large n and $\mathcal{T}$.

3.3 Spillover Enhanced Cross Effect

In our DSSES specification, contemporaneous-spillover effects serve as an amplification channel enhancing the interactive feedback between house prices and foreclosures. To understand how these spillover effects matter in terms of amplifying the transmission of a structural shock, we compute the impulse response to a structural shock.

To evaluate our DSSES specification in its original non-transformed format, we first revert the FOD-transformation, Equation (4a) now becomes

$$Y^*_{n2}(t)\Gamma = W_n Y^*_{n2}(t)\Psi + Y^*_{n2}(t-1)P + X^*_n(t)\Pi + d' \otimes l_n + C + U^*_{n2}(t),$$

where the input variables with superscript $*$ indicate they are in the original form without the FOD transformation. For simplicity, let $R^*(t) = X^*_n(t)\Pi + d' \otimes l_n + C$ denote the sum of all the exogenous terms. After moving all the endogenous terms to the left-hand-side of the functions, the system is expressed as

$$(I_n + \gamma_{12}I_n - \psi_{11}W_n)y^*_{1,n2}(t) = \rho_{11}y^*_{1,n2}(t-1) + R^*_{\cdot 1}(t) + u^*_{1,n2}(t), \tag{9a}$$

[18] As suggested by Kelejian and Prucha, 2004, the first stage estimator, $\widehat{\theta}^{\mathcal{G}}_{l,2sls}$, can be used to calculate the residuals in $\widehat{u}_{l,n2,\mathcal{T}}$, which in turn are used to form estimates of the elements in Σ as $\widehat{\sigma}_{ml} = \frac{1}{n\mathcal{T}}\widehat{u}'_{m,n2,\mathcal{T}}\widehat{u}_{l,n2,\mathcal{T}} (\forall m, l = 1, 2)$.

and

$$(I_n + \gamma_{21} I_n - \psi_{22} W_n) y^*_{2,n2}(t) = \rho_{22} y^*_{2,n2}(t-1) + R^*_{\cdot 2}(t) + u^*_{2,n2}(t), \tag{9b}$$

with $R^*_{\cdot l}(t)$ and $u^*_{l,n2}(t)$ denoting the lth column of $R^*(t)$ and $U^*_{n2}(t)$ respectively $\forall l = 1, 2$. We then stack Equation (9a) and Equation (9b) into a $2n$- dimensional vector form as

$$\Phi y^*_{n2}(t) = \mathcal{P} y^*_{n2}(t-1) + r^*(t) + u^*_{n2}(t), \tag{10}$$

where $\Phi = \begin{bmatrix} I_n - \psi_{11} W_n & \gamma_{12} I_n \\ \gamma_{21} I_n & I_n - \psi_{22} W_n \end{bmatrix}$, $\mathcal{P} = \begin{bmatrix} \rho_{11} I_n & 0_n \\ 0_n & \rho_{22} I_n \end{bmatrix}$, $y^*_{n2}(t) = [y^*_{1,n2}{}'(t), y^*_{2,n2}{}'(t)]'$, $y^*_{n2}(t-1) = [y^*_{1,n2}{}'(t-1), y^*_{2,n2}{}'(t-1)]'$, $r^*(t) = [R'_{\cdot 1}{}^*(t), R'_{\cdot 2}{}^*(t)]'$, and $u^*_{n2}(t) = [u^*_{1,n2}{}'(t), u^*_{2,n2}{}'(t)]'$.

We can invert Equation (10) to express $y^*_{n2}(t)$ as a function of its history, the exogenous variables and structural innovations as

$$y^*_{n2}(t) = \Phi^{-1} \mathcal{P} y^*_{n2}(t-1) + \Phi^{-1} r^*(t) + \Phi^{-1} u^*_{n2}(t), \tag{11}$$

Equation (11) can then be used to calculate the impulse response function to a structural innovation, $u^*_{n2}(t)$.

4 Data and Summary Statistics

We use data for the lower 48 contiguous United States, omitting the District of Columbia.

4.1 House Price and Foreclosure

We use data measured at a quarterly frequency. For house prices we are interested in the real (inflation-adjusted) quarterly log difference in house prices. We use the FHFA seasonally adjusted all-transactions price index to measure nominal prices and deflate it by the Consumer Price Index for all items less shelter as published by the U.S. Bureau of Labor Statistics.

For foreclosure rate we use the natural log of the state level foreclosure start rate as estimated by the Mortgage Bankers Association's National Delinquency Survey. Our foreclosure rate is the percent of all active loans that start a foreclosure in a quarter.

4.2 Neighborhood Specification

We construct the weight matrix of Equation (3), W_n, as an adjacency matrix by setting the weight element, w_{ij}, equal to $1/c_i$ if i and j border to each other, where c_i is the number of i's adjacent neighbors; and zero otherwise. Our adjacency weight matrix thus describes the proximity of between our observation units.

To test the robustness of our model, we also estimate Equation (3) using an alternative specifications of the weight matrix. The alternative weight matrices and their results are described in next section.

4.3 Instrumental Variables

To distinguish the impact of foreclosures on house prices from that of house prices on foreclosures in the DSSES specification, our 3SLS estimator calls for valid IVs for dealing with the endogeneity arising from FOD-transformed own time-lags, contemporaneous-spillover effect, and simultaneous-cross effect in each of the lth equation. As described in Equation (5), the building block, $X_n(t)$, in the IV matrix consists of two types of exogenous variables: the variables that are excluded from the lth equation and those that are included in the lth equation. The main challenge comes from identifying the appropriate variables meeting such exclusion criteria. In other words, we need economic variables that are correlated with house prices but not correlated with foreclosures, and vice versa.

4.3.1 ARM Resets as an Instrument for Foreclosures

To quantify the causal effect of foreclosures on house prices, we propose a novel instrument by leveraging the space and time variations of the number of adjustable rate mortgages (ARMs) that are hit by their contractual interest rate reset clock. Specifically, we calculate the percentage of ARMs encountering upward rate reset over the total number of active loans in a given month for a given state and use it as the exclusion restriction for identifying the house price equation.

ARMs were quite popular during the housing boom years of 2003-2006 as they offered borrowers low initial payments. ARMs often charge a low introductory rate (i.e., teaser rate) that helps entice borrowers and increase the marketability of ARMs over fixed rate mortgages (FRMs). ARM products typically involve two phases: the introductory period in which the interest rate is fixed, and followed by a second phase in which the rate is periodically moved to reflect prevailing market rates. The contract of an ARM features several key variables. It defines the length of the introductory period, in which the interest rate is fixed; a selected market index (e.g., LIBOR, TBill, Prime rate, etc.) which is used to reset rates in the second phase; the margin, or the spread between the index rate and the reset rates; floors and caps which determine the maximum and minimum amount the rate may move either for one reset period or for the life of the loan; and the frequency of rate adjustment which is usually one year.

The introductory rate is only temporary and most time ranges between 3 and 10 years. After the expiration of the introductory period and moving onto the second phase of an ARM, the borrower is confronted with the fluctuation of interest rate and its frequent reset. ARM indeed carries a financial risk for the borrowers. Due to this risk, rational borrowers tend to convert their ARMs to FRMs or refinance into new ARMs, prior to the rate reset date providing the prevailing market rate is favorable. Often time, ARM borrowers who are unable to refinance prior to the expiration date are those experiencing rising market rates, or a weak equity position, or both.[19] These ARM

[19]One could argue that the borrower might have experienced declining market rates, and therefore anticipated the

borrowers are more likely to encounter an increased difficulty to repay their mortgage if the rate resets up. The more the number of borrowers who are stuck with their ARM and experiencing the rate reset, the more the foreclosures. Thus a positive correlation is expected between the percentage of ARMs that are reset up and the foreclosure rate.

A hot housing market might attract more ARM borrowers due to the low initial payments. It is reasonable to believe that the number of ARMs issued might be correlated with house price at the loan origination time period. However, we do not expect the number of ARMs that are reset in a future time period have much to do the future house price directly, unless it is through the foreclosure channel.

We focus on the loans experiencing a rate increase during their initial rate reset using Black Knight's McDash 1^{st} lien data. We derive two indicators for our ARM reset calculation. First, we create a variable capturing the date when a reset hits. An ARM reset is flagged at the introductory expiration date or when the first principle and insurant (P&I) payment amount changes, whichever comes first.

Then, we compare the scheduled P&I payment from the current month with that of the previous month to identify whether the rate increases at the reset day.[20] The McDash database contains historical monthly loan-level information for more than 180 million mortgages; between 2005 and 2018 the McDash data has covered between 52 and 70 percent of the U.S. mortgage market. When aggregated across states and time, the average reset rate is 0.19% and the average reset rate with higher payment is 0.11%. The number stays roughly the same across states. For example, California has an average re-set rate of 0.22% and 0.13% are paying higher payment upon reset. Texas has an average reset rate of 0.18% and 0.09% are paying higher payment upon re-set. But these rates fluctuate significantly across time.

4.3.2 Natural Population Growth as an Instrument for House Prices

To quantify the causal effect of house prices on foreclosures, we use the quarterly change in the growth rate of natural population (i.e., Δ(births - deaths)/population) as our instrument. Popula-

rate will further decline at her ARM reset date. So she would not mind moving into the second phase of her ARM instead of re-financing out of it. However, if the market rate is trending down, from a theoretical perspective, the borrower can always better off by re-financing into a new ARM with a more favorable term instead of waiting for the expiration of her existing ARM. Though from an empirical perspective, due to frictions (e.g., re-finance cost), this scenario is possible, but not in a large scale - especially when the market rate has dropped significantly over a long enough period (e.g., the post-crisis time), borrowers qualified for re-financed loans would have done so already

[20]Because our ARM reset measure is used to proximate the number of borrowers who are likely to experience a payment shock and therefore fall behind their payment schedule afterwards, we are mainly interested in tracking ARM borrowers who have not yet defaulted prior to the rate reset as our base population. To alleviate the potential data pollution from those defaulted borrowers, we drop out loans from our tracking population when they become delinquent(i.e., 90DPD) at the time of the delinquency. Loans ever modified are excluded from our data sample if the time of modification occurs prior to the reset date.

tion growth reflects housing demand and is an important variable in many models of house prices. When population growth increases, household formation rates tend to rise, driving up housing demand. In markets with elastic housing supply over time the impact of increased population will be mitigated by expansion of the housing supply. However, many markets in the U.S. have inelastic housing supply. Moreover, housing is long-lasting and thus inelastic with respect to negative shocks. The housing supply just doesn't shrink when housing demand contracts. Rather, vacancy rates tend to rise and house price growth suffers.

State population growth by itself however, is not a suitable candidate instrument due to migration across states. Areas experiencing robust job growth tend to higher population growth rates as a booming economy attracts workers. For example, during the energy boom from 2007-2014 North Dakota experience a large influx of workers and the resident population increased sharply. This migration flow is likely correlated with the same shocks that drive foreclosure rates, so we need to find a way to control for demand shocks that are largely uncorrelated with economic conditions.

The quarterly change in the natural population growth rate is an instrument we can use. The natural population growth rate is defined as births minus deaths divided by population. While there may be some correlation between economic conditions and fertility/mortality the impact is much smaller, likely with considerable lags thus most of the cross-state variation in natural population rates is likely uncorrelated with state economic conditions.

4.4 Summary Statistics

Our estimation window covers 2005Q1-2018Q1, a period of 13.25 years (53 quarters). In addition to our instruments described above, we also include additional controls to account for economic and general housing market conditions. These controls include nonfarm payroll employment, per capita income, and single-family housing permits. For each of these controls we take the quarterly log difference in the variable. Summary statistics and variable definitions for our variables are displayed in Table (1).

Table 1: Summary Statistics (2005Q1-2018Q1)

Variable	**Mean**	**Std**	**Min**	**Max**	**Obs**
dlemp_lag1	0.002	0.006	-0.066	0.03	2544
dlperm_lag1	-0.012	0.176	-2.466	2.641	2544
dlpinc_lag1	0.003	0.013	-0.095	0.12	2544
dlrhpi	0	0.019	-0.109	0.092	2544
dnpopg	0	0.001	-0.007	0.007	2544
lfcl	-0.646	0.592	-2.303	1.324	2544
log_arm	-7.236	0.842	-9.261	-4.125	2544

Notations are providede below.

dlemp :	log difference in nonfarm payroll employment
dlperm:	log difference in single-family housing permits
dlpinc:	log difference in per capita income
dlemp_lag1:	1 quarter lag in dlemp
dlperm_lag1:	1 quarter lag in dlperm
dlpinc_lag1:	1 quarter lag in dlpinc˙lag1
dlrhpi:	log difference in real house price index
dnpopg:	quarterly difference in the natural population growth rate (births- deaths)/population
lfcl:	log of foreclosure start rate
log_arm:	log of arm reset share

5 Estimation Results

In this section, we estimate the causal effect between house prices and foreclosures for the PVAR and the DSSES. Based on the estimates, we compute the short-run and long-run cross effects and discuss the different outcomes from these two model specifications.

5.1 PVAR

Our main result involves the dynamic system of spatial simultaneous equations systems (DSSES) described in Section (3.2). However, before we proceed to the more complicated dynamics embedded in the DSSES framework it is useful to consider the results of PVAR, discussed in Section (3.1). The PVAR approach provides a useful benchmark and also allows us to compare our results directly to previous literature, particularly CLM who use a PVAR.

The PVAR treats all the variables as endogenous and estimates a reduced form equation. Then, identification of structural innovations is achieved through some strategy. The most common approach, and the one used by CLM is a recursive identification scheme. The recursive identification scheme requires that the variables only respond contemporaneously to innovations in variables ordered ahead of that equation. Thus, the first variable in the system is assumed to respond to only itself. The second variable responds to the first and itself, and so on, until the last variable in the system responds to innovations in all other equations. Table (10) in Appendix (B) presents the estimation results for a PVAR(12).

The coefficients of the PVAR are difficult to interpret so it is useful to consider summary statistics. One useful summary is the forecast error variance decomposition. The forecast error variance decomposition provides an estimate of the proportion of variation at a given horizon that is attributable to innovations in one of the variables.

The first panel provides estimates for the proportion of forecast errors for employment attributable to various innovations at horizons of 4, 8 and 24 quarters. For the employment variable, over 80% of the variation 4 quarters is due to innovations to the employment equation. The foreclosure rate only contributes a small amount, slightly over 1%. At 24 quarters innovations to the foreclosure equation contribute less that 5 percent of the variation. Note that for each of the first 3 variables (employment, per capita income and permits) innovations to the foreclosure equation contribute less than 5 percent even out to 24 quarters.

The fourth panel, which shows the house price response to shocks is of key interest. This tells us what proportion of real house price variation can be attributed to innovations in various variables. We find that innovations to the foreclosure equation account for between 16.6 and 22.8 percent of the variation in house prices. This is notably higher than the results in CLM, which find that foreclosure innovations only explain about 5 percent of the variation. Our PVAR includes different variables, but that is not the main reason for the divergence. If we include the same variables as

Table 2: Forecast Error Variance Decomposition

FEVD: Employment Response to Shocks (2005-2018)

horizon	**dlemp**	**dlpinc**	**dlperm**	**dlrhpi**	**lfcl**
4	0.808	0.013	0.087	0.081	0.011
8	0.521	0.022	0.186	0.235	0.036
24	0.488	0.032	0.219	0.213	0.048

FEVD: Per capita income Response to Shocks (2005-2018)

horizon	**dlemp**	**dlpinc**	**dlperm**	**dlrhpi**	**lfcl**
4	0.053	0.916	0.012	0.016	0.003
8	0.102	0.836	0.013	0.042	0.008
24	0.106	0.74	0.05	0.066	0.038

FEVD: Single-Family Permits Response to Shocks (2005-2018)

horizon	**dlemp**	**dlpinc**	**dlperm**	**dlrhpi**	**lfcl**
4	0.055	0.007	0.893	0.044	0.002
8	0.054	0.009	0.864	0.051	0.021
24	0.057	0.019	0.829	0.066	0.03

FEVD: House Price Response to Shocks (2005-2018)

horizon	**dlemp**	**dlpinc**	**dlperm**	**dlrhpi**	**lfcl**
4	0.04	0.039	0.038	0.718	0.166
8	0.051	0.034	0.046	0.651	0.218
24	0.084	0.043	0.175	0.471	0.228

FEVD: Foreclosure Response to Shocks (2005-2018)

horizon	**dlemp**	**dlpinc**	**dlperm**	**dlrhpi**	**lfcl**
4	0.004	0.003	0.031	0.029	0.933
8	0.007	0.004	0.132	0.091	0.766
24	0.007	0.012	0.307	0.073	0.601

in CLM we still would see a divergence between our results and CLM's.

The main reason we see a divergence between our result and CLM is due to the sample selection period. CLM estimate their PVAR on data from 1981 to 2009. This period excludes much of the long recovery in housing markets following the Great Recession. Our sample covering 2005Q1-2018Q1, while shorter, extends beyond the Great Recession and into the recent housing recovery. During our sample period, the foreclosure rate was much higher and more volatile. Our sample, allows us to trace out the dynamic linkages between house prices and foreclosure rates during periods when foreclosure start rates spiked and subsequently fell.

5.2 DSSES

The PVAR does not fully exploit the information contained in the state data. In particular, it does not take advantage of possible spatial autocorrelation. It seems unlikely that the innovations in one state would not have any impact on the outcomes in a neighboring state. To help control for possible spatial autocorrelation and to allow for contemporaneous interactions between foreclosure starts and house prices, we estimate the DSSES model described in Equation (3). We use the instrumental variables described in Section (4.3) to identify the contemporaneous impact of a foreclosure innovation on house prices and vice versa.

We chose to limit our estimation to a single time lag (i.e., $p = 1$) for each equation. While the estimation is straightforward to include additional lags into the specification, the results become more difficult to interpret. The inclusion of a single time lag (own and cross equation) along with a spatial lag provides a sufficiently rich model of the state house price and foreclosure rate interactions. For an easy cross reference, we rewrite Equation (3) for $p = 1$ (to simplify the notations, we leave out the subscript of p from the coefficients of ρs and replace p by 1 for the time index of time-lagged ys) below

$$y_{1,i}(t) = -\gamma_{12}y_{2,i}(t)+\psi_{11}W_nY_{1,n2}(t)+\rho_{11}y_{1,i}(t-1)+\rho_{12}y_{2,i}(t-1)+x'_{1,i}(t)\pi_{\cdot 1}+u_{1,i}(t), \quad (12a)$$

and

$$y_{2,i}(t) = -\gamma_{21}y_{1,i}(t)+\psi_{22}W_nY_{2,n2}(t)+1\rho_{22}y_{2,i}(t-1)+\rho_{21}y_{1,i}(t-1)+x'_{2,i}(t)\pi_{\cdot 2}+u_{2,i}(t), \quad (12b)$$

for $t = 1, 2, \cdots, \mathcal{T}$ and $i = 1, 2, \cdots, n$. We treat employment, per capita income and permits as predetermined variables and include them with a time lag of a single quarter. Table (3) presents our main results corresponding to Equation (12).

Our DSSES includes two variables the real house price growth rate (HPA) and the log foreclosure rate (FCL). Let us consider the HPA equation first before turning to the FCL equation.

Table 3: DSSES Estimation - State Adjacency Weights

	Beta 3SLS	**Std Error**	**t value**	**Pvalue**
HPA: FCLonHPA ($-\gamma_{12}$)	-0.054	0.005	-11.194	0
HPA: Spatial lag (ψ_{11})	0.444	0.038	11.831	0
HPA: owntime_lag1 (ρ_{11})	0.228	0.046	4.934	0
HPA: cross FCL_lag1 (ρ_{12})	0.05	0.005	10.467	0
HPA: dnpopg ($\pi_{.1}$)	0.352	0.194	1.818	0.035
HPA: dlemp_lag1 ($\pi_{.1}$)	0.459	0.057	8.063	0
HPA: dlpinc_lag1 ($\pi_{.1}$)	-0.085	0.021	-4.133	0
HPA: dlperm_lag1 ($\pi_{.1}$)	-0.002	0.001	-1.755	0.04
FCL: HPAonFCL ($-\gamma_{21}$)	-6.684	0.73	-9.161	0
FCL: Spatial lag (ψ_{22})	-0.044	0.028	-1.539	0.062
FCL: crossHPA_lag1 (ρ_{21})	1.212	0.742	1.634	0.051
FCL: owntime_lag1 (ρ_{22})	0.932	0.035	27.002	0
FCL: log_arm ($\pi_{.2}$)	0.01	0.005	1.774	0.038
FCL: dlemp_lag1 ($\pi_{.2}$)	2.699	1.03	2.62	0.004
FCL: dlpinc_lag1 ($\pi_{.2}$)	-0.689	0.326	-2.11	0.017
FCL: dlperm_lag1 ($\pi_{.2}$)	-0.076	0.02	-3.9	0

HPA refers to dlrhpi equation

FCL refers to lfcl equation

5.2.1 House Price Equation

Our results indicate that a 1 percent increase in the foreclosure rate ceteris paribus reduces real house prices by 5.4 basis points. The effect is dampened by the positive coefficient on the cross lag which nearly equals the contemporaneous coefficient. The reflects the fact that over the long run foreclosure innovations ought to have limited impact on the level of real house prices.

We find a significant positive coefficient on the spatial lag of 0.444. This implies that a 1 percent increase in the neighboring states' house prices leads to a 0.444 percent increase in a state's house prices. We also see a own time lag of 0.228, implying house price innovations are persistent. The IV for house prices, $\mathtt{dnpopg}$ has a small, but statistically significant coefficient of 0.352. This implies that a 1 percentage point increase in the state's natural population growth rate (birth rate minus death rate) leads to a 0.352 percent increase in real house prices.

The control variables dlemp_lag1, dlpinc_lag1, and dlperm_lag1 show up as significant drivers of house prices though their signs are difficult to interpret in isolation due to considerable collinearity between the three variables. For example, employment growth shows up with a positive coefficient while per capita income has a negative coefficient. However, employment is unlikely to increase without an impact on per capita income. When we estimate the model with only one of the predetermined variables the signs are as expected (positive), but our ability to identify the contemporaneous impact of foreclosure shocks is reduced.

5.2.2 Foreclosure Equation

We now move on to the foreclosure equation. We find a large and significant negative coefficient on HPA, indicating an increase in house prices lower the foreclosure start rate as economic theory would suggest.

The spatial lag coefficient for the foreclosure equation shows up with a small, but statistically significant sign. The interpretation of this negative spatial lag coefficient is not straightforward, but it is important to remember that we have estimated a complex dynamic equation system. Results below will provide additional insight (e.g., see Figure (**??**) for FCL response to FCL shock).

There is a lot of persistence in the foreclosure equation with an owntime_lag1 of 0.932. This implies that four quarters following a shock over 3/4 of the effect ($0.932^4 = 0.7545$) remains. The cross lag of HPA on FCL has opposite sign to the contemporaneous impact of HPA on FCL indicating some dampening over time.

Our instrument log_arm has a statistically significant impact on the foreclosure rate. A one percent increase in the share of mortgage loans that experience a payment shock increases the foreclosure rate by 1 basis point. That effect may seem economically small, but the standard deviation of our log_arm indicator is 0.842, indicating a typical shock to log˙arm raises the foreclosure rate by 0.842 percent.

As in the HPA equation interpreting the control variables individually is difficult, but we see statistically significant impacts from each control.

5.2.3 Short Run and Long Run Anlaysis

Even though the DSSES results are easier to inspect than the PVAR(12) coefficients, it is still somewhat difficult to build intuition for such a complex model by only considering coefficients. The tables below shew some additional light on the results.

First let us consider the short-run and long-run effect. We first can solve Equation (11) to consider the short-run impact of a shock. We have 48 states in our estimation sample, and the impact for each state differs slightly based on the number of neighbors, we can compute a summary which is the average of the state responses. A 1 standard deviation innovation to the HPA equation (=1.08%) results in a 1.98 percent increase in house prices and a 13.1 percent decline in the foreclosure rate. A standard deviation (=16.5%) to the log foreclosure start rate leads to a 1.62 percent decline in house prices and a 27.1 percent increase in the foreclosure rate.

There there is no long-run impact to the foreclosure rate. The model is stationary. However because we specified house prices in log differences, we can compute the cumulative long-run impact to the level of house prices. A one standard deviation shock to house prices leads to a cumulative increase in house prices of 2.6%, while a 1 standard deviation shock to the foreclosure rate leads to a 2% decline in house prices.

Using the estimated coefficients we can also compute the impulse response to a structural innovation to house prices or foreclosure rates. Note that this is only a partial response as we have not estimated the full dynamic relationship between our endogenous variables (HPA and FCL) and our predetermined variables.

Because our interest is in foreclosure innovations and because our Forecast Error Variance Decompositions above indicates that foreclosure innovations only contribute a small amount to variations in our predetermined variables (dlemp, dlpinc, and dlperm) this partial impulse response captures most of the dynamics that a fully specified model is likely to generate.

Let's consider first the response of HPA to a 1 standard deviation structural shock to one state. For illustration purposes we choose Nevada and present the results for nearby states, AZ, CA, ID, OR, and UT. Due to the spatial lag in our model, shocks to NV generated responses in its neighbors (the results extend to all states, but die off quickly with distance). Table (4) shows the cumulative response of house prices 4, 8, and 24 quarters following a structural shock to NV FCL.

This shows that after 4 quarters a 1 standard deviation shock the NV foreclosure equation results in a cumulative decline of 2.5% to NV house prices. Neighboring CA also experiences a decline, but of a more modest amount (1.1%). After 24 quarters the cumulative decline is smaller in absolute value as prices recover. And in nearby states the cumulative response is almost zero

Table 4: Cumulative Real House Price Response to a standardized NV Foreclosure Shock

horizon	AZ	CA	ID	NV	OR	UT
4	-0.0075	-0.0109	-0.007	-0.0252	-0.0097	-0.0063
8	-0.0053	-0.0076	-0.0049	-0.0229	-0.0068	-0.0044
24	-0.0014	-0.0021	-0.0013	-0.0195	-0.0019	-0.0011

(0.21% for CA).

We can also compute the HPA response of each of the six states to a NV HPA shock in Table (5).

Table 5: Cumulative Real House Price Response to a a standardized NV Real House Price Shock

horizon	AZ	CA	ID	NV	OR	UT
4	0.0107	0.0156	0.0099	0.0321	0.0139	0.0089
8	0.0094	0.0137	0.0087	0.0307	0.0121	0.0078
24	0.0056	0.0085	0.0049	0.0265	0.0072	0.0045

Four quarters following a 1 sd shock to NV HPA, NV HPA is up 3.2%. 24 quarters later, NV house prices are up 2.65%. Neighboring CA is up 1.6% four quarters later and 0.85% 24 quarters later.

We can also consider the FCL response of each of the six states to a NV FCL shock and a NV HPA shock. For these we do not compute the cumulative response but the response in the log of foreclosure starts 4, 8, and 24 quarters following a shock.

Table 6: Log Foreclosure Response to a a standardized NV Foreclosure Shock

horizon	AZ	CA	ID	NV	OR	UT
1	0.0198	0.0301	0.0175	0.2731	0.0256	0.016
4	0.0244	0.0339	0.0235	0.244	0.032	0.0205
8	-0.0012	-0.0037	-0.0005	0.1699	-0.0013	-0.0012
24	-0.0129	-0.0214	-0.0117	0.0561	-0.0167	-0.0115

Following a foreclosure shock, NV FCL rate is up 0.27%. Nearby CA is up 3% as spatial spillovers are positive (recall that the system accounts for impact on both FCL and HPA). But after 24 quarters NV foreclosure rates are still up 5.6%, but now the level of neighboring states have experienced a modest decline (2.1% for CA). This could be do to displaced demand. Households who were foreclosed in NV may move to nearby states, bolstering housing markets and leading to (very modest) declines in the FCL rate in those states.

Table 7: Log Foreclosure Response to a a standardized NV Real House Price Shock

horizon	AZ	CA	ID	NV	OR	UT
1	-0.0275	-0.0422	-0.0242	-0.1329	-0.0356	-0.0223
4	-0.043	-0.062	-0.0403	-0.1409	-0.0559	-0.0361
8	-0.0177	-0.0256	-0.0164	-0.0918	-0.0228	-0.0148
24	0.0049	0.0077	0.0045	-0.0229	0.0064	0.0042

Our results show that the cumulative response of the level of real house prices to a one standard deviation shock to foreclosure rates (see Table (4)) is similar to the response in the log level of the foreclosure rate in response to a one standard deviation house price shock (see Table (7)). After six years, a one standard deviation shock to the Nevada foreclosures lowers real house prices in Nevada by 1.95 percent, while six years after a one standard deviation shock to Nevada's house prices foreclosure rates have declined by 2.29 percent. Thus, the cumulative response to a standardized shock is only 36 percent larger for house prices on foreclosure than for foreclosure on house prices. This result stands in contrast to CLM (Figure 3) who find that the standardized foreclosure response to prices is 79% larger than the standardized price response to foreclosures.

It is also useful to compare our results in terms of the magnitude of responses. MST find that a one-standard deviation increases in foreclosures results in an 8% to 12% decline in house prices over 9 quarters. In contrast, CLM find a shock that results in a two-year increase in the foreclosure start rate of 4.3 percentage points results in a nine-quarter cumulative decline in house prices of 2.7%, and 6.8% over the long run. Tables 4-8 report the response of a standardized shock, and thus are not directly comparable to MST or CLM.

Instead, we can compute the size of a standardized foreclosure shock that is sufficient to increase a state's own foreclosure rate by 1 standard deviation (as measured in data). The standard deviation of the log foreclosure rate in our sample (see Table (1)) is 0.59 percent. A standardized shock to Nevada's foreclosure rate increases Nevada's foreclosure rate 0.1699 after 8 quarters (see Table (6)). Thus, we need a shock that is 3.47 standard deviations (0.59/0.1699) to generate a one-standard deviation increase in Nevada's foreclosure rate. Multiplying our result in Table 4 by 3.47 indicates that a 1-standard deviation shock (comparable to the one considered in MST and CLM) decreases real house prices 7.9 percent after 8 quarters.

5.3 Alternative Weight Matrices

In this section we present results using alternative weight matrices. We consider two alternatives.

In the first, we group states based on the U.S. Census Bureau's division. States within the same division are all neighbors, while states in other divisions are not neighbors. For example, TX, LA, OK, and AR (members of the West South Central Division) are all neighbors, but NM is not a neighbor with TX because it is int he Mountain Division. Also note that we continue to exclude HI and AK from our analysis, so the Pacific division consists of only CA, OR, and WA.

In our second alternative weight matrix, we use state-to-state migration flows based on IRS data. We consider the number of exemptions that were filed in a particular state where the return had been filed in another state in the previous year. We base our weight matrix on the number of in-migrants (not net flows) from one state to another. To smooth out volatility we compute the annual average number of immigrants to each state from 1995-2005. The weights are based on the share of all in-migrants over that period which came from a particular state.

Because California has a large number of out-migrants, this weighting scheme effectively places more weight on CA.

Flow data is compiled from administrative records from IRS's Individual Master File which includes a record for every individual income tax return filed. The data is developed by matching the records of individual income tax returns filed in the "base year", using the social security number of the primary taxpayer with the tax return filed the following year.

When the SSN of the primary taxpayer on the return filed in the base year matches the SSN of the return filed in the following year the county residence was compared to determine if they were the same. If the county address matched, then the taxpayer was counted as a "non-migrant". If the county address did not match, then the taxpayer was considered an "out-migrant relative to the county address on the return filed in the base year and an "in-migrant" relative to the county address on the current year. Only returns for which the SSN reported on the return in the "base" year matched the SSN reported on the return in following year are included.

Table 8: DSSES Estimation- Division Weights

	Beta 3SLS	**Std Error**	**t value**	**Pvalue**
HPA: FCLonHPA ($-\gamma_{12}$)	-0.037	0.005	-7.547	0
HPA: Spatial lag (ψ_{11})	0.516	0.039	13.398	0
HPA: owntime_lag1 (ρ_{11})	0.206	0.05	4.165	0
HPA: cross FCL_lag1 (ρ_{12})	0.034	0.005	6.932	0
HPA: dnpopg ($\pi_{.1}$)	0.219	0.181	1.211	0.113
HPA: dlemp_lag1 ($\pi_{.1}$)	0.355	0.051	6.983	0
HPA: dlpinc_lag1 ($\pi_{.1}$)	-0.079	0.019	-4.196	0
HPA: dlperm_lag1 ($\pi_{.1}$)	-0.002	0.001	-1.539	0.062
FCL: HPAonFCL ($-\gamma_{21}$)	-5.422	0.842	-6.437	0
FCL: Spatial lag (ψ_{22})	-0.025	0.035	-0.712	0.238
FCL: crossHPA_lag1 (ρ_{21})	0.72	0.863	0.834	0.202
FCL: owntime_lag1 (ρ_{22})	0.902	0.04	22.295	0
FCL: log_arm ($\pi_{.2}$)	0.017	0.006	2.905	0.002
FCL: dlemp_lag1 ($\pi_{.2}$)	1.095	1.06	1.033	0.151
FCL: dlpinc_lag1 ($\pi_{.2}$)	-0.394	0.34	-1.16	0.123
FCL: dlperm_lag1 ($\pi_{.2}$)	-0.073	0.019	-3.763	0

HPA refers to dlrhpi equation

FCL refers to lfcl equation

Table 9: DSSES Estimation- IRS Migration Weights

	Beta 3SLS	**Std Error**	**t value**	**Pvalue**
HPA: FCLonHPA ($-\gamma_{12}$)	-0.055	0.006	-8.823	0
HPA: Spatial lag (ψ_{11})	0.245	0.028	8.676	0
HPA: owntime_lag1 (ρ_{11})	0.221	0.063	3.534	0
HPA: cross FCL_lag1 (ρ_{12})	0.05	0.006	8.167	0
HPA: dnpopg ($\pi_{.1}$)	0.754	0.237	3.184	0.001
HPA: dlemp_lag1 ($\pi_{.1}$)	0.527	0.065	8.099	0
HPA: dlpinc_lag1 ($\pi_{.1}$)	-0.134	0.023	-5.729	0
HPA: dlperm_lag1 ($\pi_{.1}$)	-0.003	0.002	-1.663	0.048
FCL: HPAonFCL ($-\gamma_{21}$)	-4.479	0.933	-4.799	0
FCL: Spatial lag (ψ_{22})	0.007	0.016	0.439	0.33
FCL: crossHPA_lag1 (ρ_{21})	1.121	0.85	1.319	0.094
FCL: owntime_lag1 (ρ_{22})	0.862	0.031	28.033	0
FCL: log_arm ($\pi_{.2}$)	0.03	0.006	5.079	0
FCL: dlemp_lag1 ($\pi_{.2}$)	-0.412	1.131	-0.364	0.358
FCL: dlpinc_lag1 ($\pi_{.2}$)	-0.271	0.359	-0.753	0.226
FCL: dlperm_lag1 ($\pi_{.2}$)	-0.06	0.02	-3.047	0.001

HPA refers to dlrhpi equation

FCL refers to lfcl equation

6 Conclusion

In this paper we studied the dynamic relationship of house prices and foreclosure rates across space and time using a panel of U.S. states. Our results show that there is an economically significant impact of house prices on foreclosure rates and foreclosure rates on house prices. Moreover, even at the state level neighborhood effects are important. Shocks to the foreclosure rate in one state not only impacts house prices in that state, but also the foreclosure rate and house prices in nearby states. When it comes to the housing market, what happens in Vegas doesn't always stay in Vegas. Our DSSES model estimation results show that a one standard deviation foreclosure shock leads to a short-run real house price decline of 1.6 percent and a 2 percent decline in real house prices over the long run. A one standard deviation shock to real house prices lowers the foreclosure rate 13 percent in the short run. We also find significant spatial spillovers in both house prices and foreclosure rates across states. For example, four quarters after a one standard deviation shock to Nevada's foreclosure rate, real house prices in California experience a cumulative decline of 1 percent.

This paper also introduces a novel estimation strategy for Dynamic Spatial Simultaneous Equations System (DSSES). The DSSES allows researchers to estimate a dynamic simultaneous equation system with simultaneous equations, time lags and spatial lags in a panel data setting. This estimation technique could be a useful approach for modeling dynamics across space and time in regional studies.

We also contributed to the growing literature interested in understanding the dynamics of house prices and foreclosure rates by identifying two potentially useful instruments for state panel models. The ARM payment shock could be used to identify the effects of foreclosure rates, while the natural rate of population growth could serve as a useful instrument for housing demand shocks that is less likely to be correlated with economic factors than pure population growth.

The fact that foreclosure rates have an economically meaningful impact on house prices at the state level could be useful information for policymakers evaluating the effectiveness of foreclosure mitigation programs. While the literature has established that spillovers of any individual property die off after a short distance, the aggregate effect of multiple foreclosures in an area have impacts not only on local housing markets, but ripple across space and time, magnifying their aggregate impact. Studies that omit these important effects—contemporaneous causality and spatial lags—are likely to underestimate that impact of foreclosure rates on house prices and thus understate the potential benefits of foreclosure mitigation activities.

Appendix A FOD Transformation

Arellano and Bover, 1995 show the FOD transformation preserves the i.i.d. feature of the original error terms. We follow Lee and Yu, 2014's notation and express the FOD operator as, $F_{T-p,T-p-1}$, as a $(T-p)\times(T-p-1)$ matrix consisting of a subset of the eigenvectors of $J_{T-p}=(I_{T-p}-\frac{1}{T-p}l_{T-p}l'_{T-p})$ (where I_{T-p} is a $(T-p)\times(T-p)$ identity matrix and l_{T-p} is a $(T-p)$-dimensional vector of ones) - the eigenvectors corresponding to the unit eigenvalues, i.e., $J_{T-p}F_{T-p,T-p-1}=F_{T-p,T-p-1}$, $F_{T-p,T-p-1}F'_{T-p,T-p-1}=J_{T-p}$, and $F'_{T-p,T-p-1}F_{T-p,T-p-1}=I_{T-p-1}$. To illustrate the idea of FOD transformation, it is convenient to express the input variables of Equation (1a) in their vectorized forms. We let

$$
\begin{aligned}
&\Big[\mathrm{vec}(Y_{nm}(1)),\mathrm{vec}(Y_{nm}(2)),\cdots,\mathrm{vec}(Y_{nm}(T-p-1))\Big]= \\
&\qquad\Big[\mathrm{vec}(Y^*_{nm}(1)),\mathrm{vec}(Y^*_{nm}(2)),\cdots,\mathrm{vec}(Y^*_{nm}(T-p))\Big]F_{T-p,T-p-1}, \\
&\Big[\mathrm{vec}(Y_{nm}(0)),\mathrm{vec}(Y_{nm}(1)),\cdots,\mathrm{vec}(Y_{nm}(T-p-2))\Big]= \\
&\qquad\Big[\mathrm{vec}(Y^*_{nm}(0)),\mathrm{vec}(Y^*_{nm}(1)),\cdots,\mathrm{vec}(Y^*_{nm}(T-p-1))\Big]F_{T-p,T-p-1},\text{ and}
\end{aligned}
$$

We can further show the FOD transformation at the individual observation level (e.g., for spatial unit i at time t in equation l) as

$$
y_{l,i}(t)=\left(\frac{T-p-t}{T-p-t+1}\right)^{\frac{1}{2}}\left[y^*_{l,i}(t)-\frac{1}{T-p-t}\sum_{h=t+1}^{T-p}y^*_{l,i}(h)\right], \tag{13a}
$$

$$
y_{l,i}(t-1)=\left(\frac{T-p-t}{T-p-t+1}\right)^{\frac{1}{2}}\left[y^*_{l,i}(t-1)-\frac{1}{T-p-t}\sum_{h=t}^{T-p-1}y^*_{l,i}(h)\right], \tag{13b}
$$

Similar definitions apply to the disturbances, intercepts and location fixed effects. Because $F'_{T-p,T-p-1}l_{T-p}=0$, both the intercept and location fixed effects are eliminated from Equation (1). Also, the FOD-transformed residuals, $u_{l,i}(t)$s, are still i.i.d. across is and ts with

$$
E(u_{m,i}(t)u_{l,j}(s))=\{\begin{array}{c} 0 \text{ if } m\neq l \text{ or } i\neq j \text{ or } t\neq s \\ \sigma_{ml} \text{ if } m=l \text{ and } i=j \text{ and } t=s \end{array},
$$

because $F'_{T-p,T-p-1}F_{T-p,T-p-1}=I_{T-p-1}$.

The FOD-transformed lth equation ($\forall l=1,2,\cdots,m$) at time t can be specified as

$$
y_{l,nm}(t)=\sum_{j=1}^{p}Y_{nm}(t-j)\rho_{j,\cdot l}+u_{l,nm}(t). \tag{14}
$$

Stacking observations from all ts, the lth equation becomes

$$
y_{l,nm,T-p-1}=\sum_{j=1}^{p}Y^{(-j)}_{nm,T-p-1}\rho_{j,\cdot l}+u_{l,nm,T-p-1}. \tag{15}
$$

The superscript $(-j)$ of $Y^{(-j)}_{nm,T-p-1}$ indicates the time lagged property of this variable. It is important to note that the total number of observations in the lth equation reduces from $n(T-p)$ to $n(T-p-1)$ after the FOD transformation. $y_{l,nm,T-p-1}$ and $u_{l,nm,T-p-1}$ are now $n(T-p-1)$ vectors. The dimensions of $Y^{(-j)}_{nm,T-p-1}$ is $n(T-p-1) \times m$.

From Equation (13b), it is obvious that after FOD transformation, $y_{l,i}(t-1)$, depends on observation not only at $t-1$, but also those in the future time periods (i.e., $t, t+1, \cdots, T-p-1$). Therefore, the transformed own time-lagged term, $y_{l,i}(t-1)$, is now correlated with the transformed error term, $u_{l,i}(t)$.

Appendix B PVAR Estimate

We chose 12 lags to match CLM. The PVAR was estimated using the R package panelvar Sigmund and Ferstl, 2017 which implements GMM estimator described in Section (3.1) and the references therein.

Table 10: Dynamic Panel VAR Estimation: One-step GMM

	dlemp	dlpinc	dlperm	dlrhpi	lfcl
lag1_dlemp	0.3155***	0.3998**	2.0836	0.1371	0.6284
	-0.0428	-0.1297	-2.372	-0.0907	-1.4671
lag1_dlpinc	0.0275***	-0.3072***	0.2179	-0.1690***	1.0253**
	-0.0078	-0.0487	-0.2796	-0.0152	-0.3306
lag1_dlperm	0.0045***	-0.0044*	-0.4770***	0.0038**	-0.0685**
	-0.0007	-0.0021	-0.0674	-0.0013	-0.0264
lag1_dlrhpi	-0.0573***	0.1361***	0.1953	0.4055***	-0.7065*
	-0.0068	-0.0247	-0.3999	-0.0431	-0.3336
lag1_lfcl	0	0.0025	-0.0414*	-0.0132***	0.6029***
	-0.0006	-0.0032	-0.0193	-0.0023	-0.0319
lag2_dlemp	0.1493***	0.2063*	0.941	-0.3257**	-1.5619
	-0.0228	-0.0897	-1.0948	-0.1011	-1.1555
lag2_dlpinc	0.0206**	0.0805**	0.6547**	0.1766***	-0.5055
	-0.0078	-0.0277	-0.2315	-0.0267	-0.3264
lag2_dlperm	0.0054***	-0.0059*	-0.2683***	0.0120***	-0.1242***
	-0.0006	-0.0025	-0.0718	-0.0023	-0.0254
lag2_dlrhpi	0.0438***	0.0463	3.0295***	-0.2524***	-0.9125*
	-0.0085	-0.031	-0.6421	-0.0283	-0.3579
lag2_lfcl	0.0016***	0.0044	-0.0242	-0.0141***	0.1523***

	-0.0004	-0.0029	-0.0287	-0.002	-0.0336
lag3_dlemp	0.0074	0.0646	-4.8265**	0.4806***	-0.385
	-0.0267	-0.094	-1.7279	-0.1101	-1.1723
lag3_dlpinc	-0.0039	0.0001	-0.291	0.0880***	0.1601
	-0.0079	-0.0241	-0.3858	-0.0259	-0.3231
lag3_dlperm	0.0071***	0.0007	-0.1013	0.0074***	-0.1401***
	-0.0006	-0.0037	-0.0726	-0.0017	-0.0281
lag3_dlrhpi	0.0567***	-0.0358	1.6650***	0.0283	-0.9879*
	-0.0061	-0.0267	-0.4689	-0.03	-0.4023
lag3_lfcl	-0.0005	-0.0027	0.0495	0.0009	0.0465
	-0.0006	-0.0017	-0.0307	-0.0017	-0.0323
lag4_dlemp	-0.1659***	-0.6371***	-8.8992***	0.127	1.5232
	-0.0307	-0.0897	-1.6458	-0.0928	-0.8366
lag4_dlpinc	0.0241*	-0.0861**	0.0912	-0.0369	0.8381*
	-0.0094	-0.0281	-0.6306	-0.0308	-0.4101
lag4_dlperm	0.0080***	-0.0028	0.2958***	0.0002	-0.2056***
	-0.0008	-0.003	-0.039	-0.0017	-0.0403
lag4_dlrhpi	0.0364***	0.0595	0.6623	-0.0741*	-1.1314**
	-0.011	-0.0323	-0.5466	-0.0291	-0.4088
lag4_lfcl	-0.0002	-0.0036	-0.0136	0.0065**	0.1143***
	-0.0005	-0.002	-0.0257	-0.0021	-0.0274
lag5_dlemp	-0.0425	0.0962	0.5879	-0.2367***	-1.1427
	-0.0266	-0.1338	-1.1512	-0.0634	-1.2265
lag5_dlpinc	0.0053	-0.0572*	0.4997	0.0029	0.6269*
	-0.0078	-0.0274	-0.3056	-0.0207	-0.3024
lag5_dlperm	0.0059***	-0.0002	0.1127	-0.0071**	-0.1334**
	-0.0008	-0.0027	-0.0672	-0.0026	-0.0417
lag5_dlrhpi	0.0287***	0.0063	-0.9547***	0.0544*	-0.5745
	-0.0073	-0.0345	-0.2618	-0.0251	-0.3537
lag5_lfcl	0.0007	0.0056**	0.0292	0.0081***	0.0129
	-0.0006	-0.0019	-0.0295	-0.0019	-0.0233
lag6_dlemp	-0.0188	-0.1076	2.3433	0.0153	-0.7122
	-0.0306	-0.0928	-1.232	-0.0898	-1.0905
lag6_dlpinc	0.0215**	-0.0841**	0.5249	0.0470*	-0.9804***
	-0.0083	-0.0274	-0.3642	-0.0209	-0.2613
lag6_dlperm	0.0056***	0.003	0.0653	-0.0113***	-0.1385**
	-0.0007	-0.0026	-0.0429	-0.0034	-0.043

lag6_dlrhpi	0.0375***	0.1509***	-2.5350***	0.0585*	-0.0126
	-0.0103	-0.0291	-0.2382	-0.0259	-0.4082
lag6_lfcl	0.0002	0.0011	0.0038	-0.0090***	0.0254
	-0.0006	-0.0021	-0.0216	-0.0019	-0.025
lag7_dlemp	-0.0727**	-0.3248**	0.7768	0.3918***	-0.0973
	-0.0261	-0.1087	-1.0915	-0.1023	-1.094
lag7_dlpinc	-0.0133	-0.0981***	0.45	-0.0197	-0.3514
	-0.0095	-0.0256	-0.3698	-0.0231	-0.5115
lag7_dlperm	0.0027**	-0.003	0.028	-0.0039	-0.1142**
	-0.0008	-0.0027	-0.0457	-0.003	-0.0412
lag7_dlrhpi	-0.0525***	-0.0471*	-1.5671**	-0.0003	1.4592***
	-0.0074	-0.0232	-0.5933	-0.0234	-0.32
lag7_lfcl	0.0008	0.002	0.0503*	-0.0002	-0.0627
	-0.0006	-0.0021	-0.0245	-0.0014	-0.0361
lag8_dlemp	-0.1180***	0.1818*	-2.9672***	0.4247***	2.4385
	-0.0329	-0.0763	-0.874	-0.0634	-1.432
lag8_dlpinc	0.0141	-0.0248	0.0254	-0.1410***	-0.6935*
	-0.0085	-0.0298	-0.3497	-0.0221	-0.2887
lag8_dlperm	0.0046***	0.0142***	0.1649***	-0.0015	-0.1035**
	-0.0011	-0.0026	-0.0322	-0.0025	-0.0371
lag8_dlrhpi	0.0115	0.048	-0.1634	0.0565**	1.0628***
	-0.0071	-0.0279	-0.3353	-0.0207	-0.2928
lag8_lfcl	0.0005	0.0006	-0.0314	0.0057***	0.0337
	-0.0004	-0.0021	-0.0269	-0.0016	-0.0258
lag9_dlemp	-0.0161	-0.1409	-0.7629	0.0627	1.7948
	-0.0247	-0.0798	-0.8083	-0.0601	-1.2654
lag9_dlpinc	0.0059	0.0827**	0.1948	0.0285	0.5078
	-0.0093	-0.0254	-0.3749	-0.0201	-0.2724
lag9_dlperm	0.0051***	0.0187***	0.1836***	-0.0056*	-0.0873**
	-0.0011	-0.0026	-0.0281	-0.0027	-0.0291
lag9_dlrhpi	-0.0272***	0.1408***	-0.8998*	-0.0575**	-0.6910*
	-0.008	-0.0271	-0.3763	-0.0201	-0.3089
lag9_lfcl	-0.0004	0.0059**	-0.0564	0.0050*	0.0419
	-0.0007	-0.002	-0.0584	-0.0021	-0.0265
lag10_dlemp	0.0495	0.1143	1.2114	-0.1697**	-3.2371***
	-0.027	-0.151	-1.3497	-0.0638	-0.9408
lag10_dlpinc	0.0223***	-0.025	1.1664***	-0.1032***	-0.2892

	-0.0064	-0.0205	-0.2927	-0.0208	-0.2092
lag10_dlperm	0.0051***	0.0174***	0.0666*	-0.0088*	-0.0481
	-0.0008	-0.0027	-0.0311	-0.0045	-0.0294
lag10_dlrhpi	0.0132	-0.0354	1.7553***	-0.0167	-0.6371
	-0.0076	-0.0213	-0.4866	-0.0259	-0.3593
lag10_lfcl	0.0003	-0.0102***	0.0826**	-0.0029	0.0102
	-0.0007	-0.0018	-0.0311	-0.0021	-0.0257
lag11_dlemp	-0.0448*	-0.3182**	-3.7948***	0.3718***	3.7532***
	-0.0222	-0.1032	-0.807	-0.054	-0.9046
lag11_dlpinc	0.0012	-0.0724**	0.2399	0.0562**	-0.5252*
	-0.0066	-0.0259	-0.1889	-0.0191	-0.2459
lag11_dlperm	0.0060***	0.0152***	0.0826**	-0.0103**	-0.038
	-0.0007	-0.0031	-0.0304	-0.0036	-0.0263
lag11_dlrhpi	0.0192*	-0.1035***	0.7591	0.0182	0.9341**
	-0.0076	-0.0275	-0.4547	-0.0277	-0.3198
lag11_lfcl	0.0001	0.0054*	0.0102	0.0044*	-0.0826***
	-0.0005	-0.0022	-0.0269	-0.0021	-0.0195
lag12_dlemp	-0.1138***	-0.037	-3.3337*	-0.1767*	0.8647
	-0.0179	-0.0753	-1.3842	-0.088	-0.8378
lag12_dlpinc	0.0051	-0.0594**	-0.0262	-0.0057	-0.8895**
	-0.008	-0.0196	-0.4236	-0.0179	-0.2929
lag12_dlperm	0.0012*	0.0006	0.1154*	-0.0026	-0.0323
	-0.0005	-0.0018	-0.0544	-0.0019	-0.0263
lag12_dlrhpi	0.0222***	0.0122	1.6581***	-0.1344***	0.7040*
	-0.0054	-0.0209	-0.4859	-0.0282	-0.2822
lag12_lfcl	-0.0010*	-0.0091***	-0.0017	-0.0014	0.0823***
	-0.0005	-0.0017	-0.0213	-0.0014	-0.0232

Transformation: Forward orthogonal deviations

Group variable: fipn

Time variable: yq

Number of observations = 1920

Number of groups = 48

Obs per group: min = 40

Obs per group: avg = 40

Obs per group: max = 40

Number of instruments = 575

$p < 0.001$, $p < 0.01$, $p < 0.05$

Instruments for equation

Standard

GMM-type

Dependent vars: L(2,24))

Collapse = TRUE

Hansen test of overid. restrictions: chi2(275) = 8.22 Prob > chi2 = 1

(Robust, but weakened by many instruments.)

7 Bibliography

Agarwal, S., Ambrose, B. W., Chomsisengphet, S., Sanders, A. B., 2012. Thy neighbor's mortgage: Does living in a subprime neighborhood affect one's probability of default? Real Estate Economics 40 (1), 1–22.

Alvarez, J., Arellano, M., 2003. The time series and cross-section asymptotics of dynamic panel data estimators. Econometrica 71 (4), 1121–1159.

Anselin, L., 2008. Vspatial hedonics, v mills tc, patterson k (eds) palgrave handbook of econometrics: Volumn 2, applied econometrics.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. The review of economic studies 58 (2), 277–297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. Journal of econometrics 68 (1), 29–51.

Bajari, P., Chu, C. S., Park, M., 2008. An empirical model of subprime mortgage default from 2000 to 2007. Tech. rep., National Bureau of Economic Research.

Basu, S., Thibodeau, T. G., 1998. Analysis of spatial autocorrelation in house prices. The Journal of Real Estate Finance and Economics 17 (1), 61–85.

Calomiris, C. W., Longhofer, S. D., Miles, W. R., 2013. The foreclosure–house price nexus: a panel var model for us states, 1981–2009. Real Estate Economics 41 (4), 709–746.

Campbell, J. Y., Giglio, S., Pathak, P., 2011. Forced sales and house prices. American Economic Review 101 (5), 2108–31.

Canova, F., Ciccarelli, M., 2013. Panel vector autoregressive models: A survey. In: VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims. Emerald Group Publishing Limited, pp. 205–246.

Carroll, T., Clauretie, T., Neill, H., 1997. Effect of foreclosure status on residential selling price: comment. Journal of Real Estate Research 13 (1), 95–102.

Chomsisengphet, S., Kiefer, H., Liu, X., 2018. Spillover effects in home mortgage defaults: Identifying the power neighbor. Regional Science and Urban Economics 73, 68–82.

Clauretie, T. M., Daneshvary, N., 2009. Estimating the house foreclosure discount corrected for spatial price interdependence and endogeneity of marketing time. Real Estate Economics 37 (1), 43–67.

Cliff, A. D., Ord, J. K., 1973. Spatial autocorrelation, monographs in spatial environmental systems analysis. London: Pion Limited.

Elhorst, J. P., 2010. Dynamic panels with endogenous interaction effects when t is small. Regional Science and Urban Economics 40 (5), 272–282.

Fingleton, B., Le Gallo, J., 2008. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties. Papers in Regional Science 87 (3), 319–339.

Foote, C. L., Gerardi, K., Willen, P., 2008. Negative equity and foreclosure: Theory and evidence.

Foster, C., Van Order, R., 1984. An option-based model of mortgage default. Housing Fin. Rev. 3, 351.

Gerardi, K., Rosenblatt, E., Willen, P. S., Yao, V., 2015. Foreclosure externalities: New evidence. Journal of Urban Economics 87, 42–56.

Goodstein, R., Hanouna, P., Ramirez, C., Stahel, C., 2011. Are foreclosures contagious?

Guiso, L., Sapienza, P., Zingales, L., 2013. The determinants of attitudes toward strategic default on mortgages. The Journal of Finance 68 (4), 1473–1515.

Gupta, A., 2018. Foreclosure contagion and the neighborhood spillover effects of mortgage defaults.

Harding, J. P., Rosenblatt, E., Yao, V. W., 2009. The contagion effect of foreclosed properties. Journal of Urban Economics 66 (3), 164–178.

Hartley, D., 2014. The effect of foreclosures on nearby housing prices: Supply or dis-amenity? Regional Science and Urban Economics 49, 108–117.

Kelejian, H. H., Prucha, I. R., 2004. Estimation of simultaneous systems of spatially interrelated cross sectional equations. Journal of econometrics 118 (1-2), 27–50.

Kiefer, H., 2011. The house price determination process: Rational expectations with a spatial context. Journal of Housing Economics 20 (4), 249–266.

Lee, L.-f., Yu, J., 2014. Efficient gmm estimation of spatial dynamic panel data models with fixed effects. Journal of Econometrics 180 (2), 174–197.

LeSage, J. P., Pace, R. K., 2004. Models for spatially dependent missing data. The Journal of Real Estate Finance and Economics 29 (2), 233–254.

Makridis, C., Ohlrogge, M., 2018. The local effects of foreclosures.

Mian, A., Sufi, A., Trebbi, F., 2015. Foreclosures, house prices, and the real economy. The Journal of Finance 70 (6), 2587–2634.

Pace, R. K., Barry, R., Sirmans, C. F., 1998. Spatial statistics and real estate. The Journal of Real Estate Finance and Economics 17 (1), 5–13.

Revelli, F., 2001. Spatial patterns in local taxation: tax mimicking or error mimicking? Applied Economics 33 (9), 1101–1107.

Se Can, A., Megbolugbe, I., 1997. Spatial dependence and house price index construction. The Journal of Real Estate Finance and Economics 14 (1-2), 203–222.

Sigmund, M., Ferstl, R., 2017. Panel vector autoregression in r with the package panelvar. SSRN. URL https://ssrn.com/abstract=2896087

Towe, C., Lawley, C., 2013. The contagion effect of neighboring foreclosures. American Economic Journal: Economic Policy 5 (2), 313–35.

Yang, K., Lee, L.-f., 2018. Identification and estimation of spatial dynamic panel simultaneous equations models. Regional Science and Urban Economics.

Zhu, S., Pace, R. K., 2014. Modeling spatially interdependent mortgage decisions. The Journal of Real Estate Finance and Economics 49 (4), 598–620.