University of Connecticut

Department of Economics Working Paper Series

Estimating U.S. Housing Price Network Connectedness: Evidence from Dynamic Elastic Net, Lasso, and Ridge Vector Autoregressive Models

by

David Gabauer
Software Competence Center Hagenberg

Rangan Gupta
University of Pretoria

Hardik A. Marfatia
Northeastern Illinois University

Stephen M. Miller
University of Nevada - Las Vegas

Working Paper 2020-08
August 2020

365 Fairfield Way, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
http://www.econ.uconn.edu/

Estimating U.S. Housing Price Network Connectedness: Evidence from Dynamic Elastic Net, Lasso, and Ridge Vector Autoregressive Models

David Gabauer[§], Rangan Gupta[¶], Hardik A. Marfatia[†], and Stephen M. Miller[‡]

[§]*Data Analysis Systems, Software Competence Center Hagenberg, Austria.*
[¶]*Department of Economics, University of Pretoria, South Africa.*
[†]*Department of Economics, Northeastern Illinois University, USA.*
[‡]*Department of Economics, University of Nevada, Las Vegas, USA.*

Abstract

This paper investigates the dynamic connectedness of random shocks to housing prices between the 50 U.S. states and the District of Columbia. The paper implements a standard vector autoregressive (VAR) model as well as three VAR models with shrinkage effects – Elastic Net, Lasso, and Ridge VAR models. The transmission of random shocks on a regional basis flows from Southern states to Western states to Midwestern states to Northeastern states. Since VAR models generally confront parameter values between zero and one, the Elastic Net and Lasso VAR models perform the best since the penalty involves the absolute value rather than he squared value as in the Ridge VAR model. Our results have important implications for investors and policymakers.

Keywords: Dynamic Connectedness, Elastic Net VAR, Lasso VAR, Ridge VAR, U.S. Housing.

JEL codes: C32, C52, R31.

1 Introduction

The housing markets play a critical role in the macroeconomy, affecting both the business cycle and the financial system. The important role of housing markets in the business cycle became dramatically clear during the sub-prime mortgage market collapse in late 2006 and the resulting Great Recession and financial crisis of 2007-2009, the worst recession since World War II until the COVID-19 recession of 2020. Shiller (2012), for example, argues that the housing bubble provided the major, if not the only, cause of the sub-prime mortgage crisis and the worldwide Great Recession and financial crisis of 2007–2009. Leamer et al. (2007) more provocatively asserts that "housing is the business cycle" in the United States or, more precisely, that house prices drive the U.S. business cycle.

One can argue that the recent Great Recession and financial crisis, more than any other macroeconomic event, makes a strong case for examining the dynamics of house prices, and, in particular, the role of persistence and the effect of shocks on house price dynamics. This paper considers the "connectedness" of housing markets across the 50 U.S. states and the District of Columbia. The basis of connectedness comes from the estimation of a vector autoregressive (VAR) model, where we examine how exogenous shocks to house prices in one state affect the exogenous shocks to house prices in another state, or vice versa. That is, connectedness measures cross-state relationships between exogenous shocks to house prices and not the cross-state relationships between the movements in the house prices themselves.

Given that the housing market leads the business cycles of the United States (Balcilar et al., 2014; Leamer, 2015; Nyakabawo et al., 2015; Emirmahmutoglu et al., 2016), analysis of regional housing market connectedness is an important question for policymakers. That is, determining which states and/or regions drive the overall housing market allows policy authorities to undertake policy decisions that target specific states or regions before implementing nationwide policies, which might not work, due to the heterogeneity of the housing market (Fairchild et al., 2015). Moreover, determination of which states or regions act as the main transmitter of shocks provides information to investors on acquiring housing assets in these markets rather than the ones that are more susceptible to such shocks as net receivers.

Dividing the United States into four Census regions (Northeast, South, Midwest, and West), we find that shocks in the South affect shocks in the Northeast, Midwest and the West. Shocks in the West affect shocks in the Midwest and the Northeast. Shocks in the Midwest affect shocks in the Northeast. Finally, shocks in the Northeast do not affect shocks in the other regions.

The remainder of this paper is organised as follows. Section 2 presents the relevant literature review and Section 3 describes the empirical methodologies applied in the study. Furthermore, Section 4 provides a short overview of the employed dataset whereas Section 5 illustrates the findings of the study and discusses the relevant arguments. Finally, Section 6 summarises the key elements, provides a framework for policy implications, and concludes the study.

2 Literature Review

A significant literature exists that considers the "ripple effect" in house prices in the United Kingdom and the United States. This literature begins with theorizing and empirical analysis in the United Kingdom. The ripple effect refers to the observation that house price increases in Southeastern United Kingdom generally led with some time lag to house price increases in Northwest United Kingdom (Meen, 1999). More recent work on the ripple effect in the United Kingdom includes Cook (2003, 2005), Holmes and Grimes (2008), and Tsai (2015). Cook (2003, 2005) tests for convergence and cointegration in house prices, introducing asymmetric responses to house price increases and decreases. Holmes and Grimes (2008) apply unit root tests to the first principal component of the set of regional to national house price differentials. Tsai (2015) examines regional and national housing market spillover effects in the United Kingdom.

In the United States, Gupta and Miller (2012*b*,*a*) consider the cointegration and Granger causality three metro areas (Los Angeles, Las Vegas, and Phoenix) in three different states (California, Nevada, and Arizona) and between house prices in Southern California counties. Chiang and Tsai (2016) examine regional and interregional ripple effects in the United States for eight metropolitan areas – Los Angeles, San Diego, San Francisco, Chicago, Boston, New York, Miami, and Tampa – and three regions – East, South, and West. Tsai (2018) considers the ripple effect for the four Census regions (South, West, Midwest, and Northeast) in the United States. Tsai (2019) considers the interrelationships between house prices in 10 U.S. metropolitan areas – Boston, Chicago, Denver, Las Vegas, Los Angeles, Miami, New York, San Diego, San Francisco, and Washington DC – at three price tiers – low, medium, and high price tiers.

The ripple effect literature relies on the concept of the Law of One Price (LOOP), which proposes that a homogeneous good that sells in two different markets should sell for the same price, after incorporation transaction and transportation costs. Fundamentally, the LOOP uses the arbitrage of goods prices between markets to generate the convergence of prices across regional markets. That is, for example, if one can transport the good between markets at relatively low cost, one can buy in the low-price market and sell it in the high-price market after transporting the good from the low-price market to the high-price market. Clearly, housing goods fail in, at least, two important areas (i) lack of homogeneity in housing goods and (ii) lack of transportability between markets. In addition, rather than comparing house prices, we compare house price indexes. Thus, comparing house price indexes, rather than individual home prices, across geographic regions, activates the idea of Purchasing Power Parity (PPP). PPP extends the LOOP to price indexes, implying that trade between geographic regions of goods leads to a convergence of the regions' price indexes for goods. Once again, the successful operation of PPP requires the arbitrage of goods between regions.

Housing economists address the issue of non-homogeneous housing goods by considering their characteristics. That is, hedonic housing good models allow the comparison of house prices based on the characteristics imbedded into the housing good, such as the number of bedrooms the number of baths, the square footage, and so on. In addition, researchers want to ensure that the house price index can accommodate the quality of the house. A "repeat-sales" index based on multiple sales of the same home attempts to address this last issue. To do so successfully requires that the repeat sales include information on renovations and depreciation. A "constant quality" index can compare housing good prices across time and space. Typically, the geographic reach of the housing market reflects the commuting shed for the metropolitan area. That is, houses compete with each other within the same metropolitan area. Original work on this hypothesis was done by Tirtiroğlu (1992) and Clapp and Tirtiroglu (1994) examining housing market efficiency in the Hartford, Connecticut spatial market.

Since moving houses from one metropolitan market to another proves too costly, does this necessarily imply that the housing markets in different MSAs do not exhibit linkages? Trade theory provides an answer. The factor price equalization theorem (Samuelson, 1948) shows that although capital and labor frequently do not flow freely between countries, goods and services do flow and can proxy for capital and labor flows. The flows of goods and services between countries cause the prices of labor and capital to equalize, albeit absent any flow of capital and labor between countries. Since housing goods cannot easily flow between markets, do other flows exist that can cause LOOP or PPP to hold? First, the migration of home buyers can link the housing markets different metropolitan areas. Second, large home builders operate in multiple metropolitan areas, allowing them to move operations in repose to shifts in metropolitan demands and differential returns on home building activity. In sum, the movement of home buyers and home builders' operations between regions in response to price differences can arbitrage the prices of homes, even though the homes themselves do not move between regions.

Home builders face two basic components to their cost of supplying new housing – construction (replacement) costs and land value. If the demand for housing rises in one region, that will draw resources, including construction labor, from other regions. As a result, construction costs in both regions will rise. It rises first in the market where the demand for housing rises to attract more resources and construction workers. And as a consequence, as the supply of resources and construction workers in the other region fall, their costs and wages will rise. The equalizing of construction costs tends to equilibrate house prices across regions.

Just as we cannot transport housing between regions, we cannot transport land as well. Thus, if a region faces a fixed, or extremely inelastic, supply of land, then that regions house prices and land values will rise. That is, since house prices include construction (replacement) costs and land prices, higher land prices will drive up house prices even though construction (replacement) costs may equilibrate between regions.

Antonakakis et al. (2018) analyzed the U.K. regional housing markets and Antonakakis et al. (2019) examined the four aggregated U.S. regional housing markets, using the connectedness approach. Connectedness examines the relationships between the random shocks in a vector autoregressive (VAR) system, identifying the pattern of transmission of shocks across housing markets in this case. Antonakakis et al. (2018) used quarterly data on 13 regions in the United Kingdom (North, North West, West Midlands, Outer South East, London, Wales, Northern Ireland, Yorkshire and Humberside, East Midlands, East Anglia, Outer Metropolitan, South West, and Scotland, examining the VAR model of annual nominal housing returns. Antonakakis et al. (2019) used monthly data on four U.S. regions (Midwest, Northeast, South, and West), examining the VAR model of annual housing returns and the growth rate of housing volumes.

This study provides a more comprehensive analysis of the transmission of real housing return shocks across all 50 U.S. states and the District of Columbia, applying the Lasso, Ridge, and Elastic Net methods of parameter estimation and selection. We also consider the four regions (Midwest, Northeast, South, and West) that Antonakakis et al. (2019) examined for comparison purposes, although we consider real housing returns and Antonakakis et al. (2019) examine nominal housing returns and the growth rate of housing volumes.

3 Empirical Methodology

3.1 Lasso, Ridge & Elastic Vector Autoregressions

Sims (1980) introduced one of the most popular and still widely used multivariate time series models in applied econometrics, namely the vector autoregressive (VAR) model. This framework can not only forecast the future of multiple time series in a dynamic way, but also comes with a sophisticated toolbox, including impulse response analysis and forecast error variance decomposition. The toolbox provides further information about the spillovers of shocks. Mathematically, a VAR(p) with a lag length of p can be formulated as follows:

$$\boldsymbol{y}_t = \boldsymbol{\beta}_0 + \sum_{i=1}^{p} \boldsymbol{\beta}_i \boldsymbol{y}_{t-i} + \boldsymbol{u}_t \qquad \boldsymbol{u}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}) \tag{1}$$

where $\boldsymbol{y}_t$, $t = 1, ..., T-p$, is an $k \times 1$ dimensional vector of the endogenous variables, $\boldsymbol{\beta}_0$ is an $k \times 1$ dimensional vector of all intercepts, and $\boldsymbol{\beta}_i$ represents the $k \times k$ dimensional VAR coefficient matrix of the ith lag. The $k \times 1$ dimensional error vector, $\boldsymbol{u}_t$, $t = 1, ..., T-p$, is multivariate normally distributed with means equal to zero and a variance-covariance matrix equals to $\boldsymbol{\Sigma}$. $\boldsymbol{\Sigma}$ is a diagonal matrix as $var(u_i) = \sigma_i^2$, $i = 1, ..., k$ and $cov(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0$, $i, j = 1, ..., k$ and $i \neq j$. Those characteristics allow to estimate the VAR model equation-by-equation often by the ordinary leasts squares method. This is good news as we are estimating the U.S. real estate return dynamics of all states, meaning that $k = 51$ and, hence, a joint estimation with a lag length of one[1] would mean that $k + k^2 p = 2,652$ parameters needs to be estimated, however, we only have $T = 418$ observations ($2,652 >> 418$), which would mean that this model cannot be estimated. Using the equation-by-equation estimation procedure makes this feasible as it just estimates, $k + 1 = 52$ parameters ($52 < 418$). Additionally, as we deal with this high-dimensional network, we employ different regularization methods to shrink and select parameters and compare the results with the standard equation-by-equation VAR model.

[1] The Bayesian information criterion suggests to use a lag length of one.

Hereby, the elastic net (Zou and Hastie, 2005) can be seen as the generalization of the OLS, Ridge, and Lasso (Tibshirani, 1996) estimation method which shrinks and selects parameters. The elastic net can be outlined for each variable j, $j = 1, ..., k$, as follows:

$$\underset{\boldsymbol{\beta}}{\text{argmin}} \left(\overbrace{\left[T^{-1} || y_{jt} - \beta_{j0} - \sum_{i=1}^{p} \sum_{j=1}^{k} \beta_{ji} y_{jt-i} ||_2 \right]}^{MSE} + \lambda \overbrace{\left[(1-\alpha) \underbrace{||\beta_{j0} + \sum_{i=1}^{p} \sum_{j=1}^{k} \beta_{ij} ||_2}_{\text{Ridge Penalty}} + \alpha \underbrace{||\beta_{j0} + \sum_{i=1}^{p} \sum_{j=1}^{k} \beta_{ij} ||_1}_{\text{Lasso Penalty}} \right]}^{P_\alpha} \right)$$

where the first part represents a loss function, which is in our case the mean squared error (MSE), and the second part is the elastic net penalty (P_α), which equals a weighted average of the Ridge penalty and the Lasso penalty. Whereas Lasso uses an ℓ^1-norm penalty to achieve a sparse solution, Ridge uses an ℓ^2-norm penalty. This means that the Lasso penalty term proves more effective than the Ridge penalty term if parameters lie between 0 and $|1|$whereas the Ridge penalty proves stronger than the Lasso penalty for parameters $> |1|$. As VAR coefficients usually lie in the range between 0 and $|1|$, the Lasso regression will shrink parameters more strongly than in the Ridge regression. In general, the ℓ^p-norm is defined by: $||\beta||_p = (\sum_{i=1}^{p} \sum_{j=1}^{k} |\beta_{ij}|^p)^{1/p}$.

The penalty parameters λ and α for each federal state j, $j = 1, ..., k$ are chosen based on 10-fold cross validation. The cross validation method is often used in machine learning and selects penalty parameters that improve a model's forecasting performance.

Besides estimating the full elastic net regression model, we also estimate restricted submodels, namely OLS ($\lambda = 0$), Ridge ($\alpha = 0$), and Lasso ($\alpha = 1$).

To the best of our knowledge and besides Demirer et al. (2018) who employed a Lasso-VAR model to estimate the connectedness measures of daily bank stock return volatilities, this is the first paper that uses regularization methods to shrink and select VAR parameters to compute connectedness measures.

3.2 Connectedness Measures

The starting point for the connectedness approach of Diebold and Yilmaz (2012) transforms the VAR(p) in (1) into its vector moving average representation using the Wold theorem: $\boldsymbol{y}_t = \sum_{j=0}^{\infty} \boldsymbol{\theta}_j \boldsymbol{u}_{t-i}$ where $\boldsymbol{\theta}_0 = \boldsymbol{I}_k$ and $\boldsymbol{\theta}_j = \boldsymbol{\beta}_1 \boldsymbol{\theta}_{j-1} + ... + \boldsymbol{\beta}_p \boldsymbol{\theta}_{j-p}$.

In next step, the Generalized Forecast Error Variance Decomposition (GFEVD)[2] of Koop et al. (1996) and Pesaran and Shin (1998) - which is invariant to the ordering of the variables in the VAR. The GFEVD ($\tilde{\phi}_{ij}^g(H)$) can be interpreted as variable j's contribution to variable i's H-step ahead forecast error variance. It is also called the pairwise directional connectedness and can be mathematically formulated by,

$$\phi_{ij}^g(H) = \frac{\sigma_{ii}^{-1} \sum_{t=1}^{H-1} (\boldsymbol{\iota}_i' \boldsymbol{\theta} \boldsymbol{\Sigma} \boldsymbol{\iota}_j)^2}{\sum_{j=1}^{k} \sum_{t=1}^{H-1} (\boldsymbol{\iota}_i \boldsymbol{\theta} \boldsymbol{\Sigma} \boldsymbol{\theta}' \boldsymbol{\iota}_i)} \qquad \tilde{\phi}_{ij}^g(H) = \frac{\phi_{ij}^g(H)}{\sum_{j=1}^{k} \phi_{ij}^g(H)},$$

where $\sum_{j=1}^{k} \tilde{\phi}_{ij}^g(H) = 1$, $\sum_{i,j=1}^{k} \tilde{\phi}_{ij,t}^g(H) = k$ and $\boldsymbol{\iota}_j$ is a zero vector with unity on the jth position.

Based upon the bilateral pairwise directional connectedness, aggregated connectedness measures

[2] As to the best of our knowledge, no economic theory exists that explains and justifies the variable ordering in case of real estate return dynamics, we have chosen the GFEVD instead of its orthorgonalized counterpart. Hereby, we follow the study of Wiesen et al. (2018) which emphasize that without a proper economic theory the GFEVD should be the model of choice.

are derived that provide an overview of the network spillover dynamics:

$$TO_j = \sum_{i=1, i\neq j}^{k} \tilde{\phi}_{ij}^{g}(H) \tag{2}$$

$$FROM_j = \sum_{j=1, i\neq j}^{k} \tilde{\phi}_{ij}^{g}(H) \tag{3}$$

$$NET_j = TO_j - FROM_j \tag{4}$$

$$TCI = k^{-1} \sum_{j=1}^{k} TO_j \equiv k^{-1} \sum_{j=1}^{k} FROM_j. \tag{5}$$

$$NPDC_{ij} = \tilde{\phi}_{ij}(H) - \tilde{\phi}_{ji}(H). \tag{6}$$

Equation (2) represents the aggregated impact a shock in variable j exerts on all other variables, which is defined as the total directional connectedness to others. Furthermore, Equation (3) formulates the aggregated influence all other variables exert on variable j, which is the so-called total directional connectedness from others. Subtracting the impact variable j exerts on others by the influence others exert on variable j generates the net total directional connectedness (4). This metric reveals whether a variable is a net transmitter or a net receiver of shocks. If $NET_j > 0$ ($NET_j < 0$), the effect of a shock in variable j on all others is larger (smaller) than vice versa. Thus, variable j is considered as a net transmitter (receiver) of shocks. Finally, the total connectedness index (TCI) (5), which measures the average effect state exerts on all other states or the average effect of all other states on a given state, is often considered as the systemwide connectedness and, hence, the market risk. A large (small) TCI means that the average propagation of a shock in one variable to all others is high (low) and, thus, the market risk is high (low). Finally, equation (6), which is the net pairwise directional connectedness ($NPDC_{ij}$), exhibits whether variable i drives or driven by variable j. A positive (negative) $NPDC_{ij}$ implies that variable j is dominates (is dominated by) variable i.

4 Dataset

The seasonally adjusted monthly nominal house price data for the 50 states and the District of Columbia come from the Freddie Mac, with the indices based on an ever expanding database of loans purchased by either Freddie Mac or Fannie Mae. The data are available for download from: http://www.freddiemac.com/research/indices/house-price-index.page. We generate corresponding real values by deflating the nominal house price indexes with the seasonally adjusted consumer price index (CPI) that come from the FRED database of the Federal Reserve Bank of St. Louis. The data cover the period from January, 1976 through November, 2019. We work with the annual growth rates of real house prices. That is, we examine an approximation to real housing returns.[3]

5 Empirical results

We employ a rolling-window estimation with a 120-month (10-year) window and a 10-month forecast horizon, with repeated 10-fold cross validation in each window. When reporting findings, we use the Net Effect method calculations, unless the exhibit clearly shows the reporting of all four methods (see Figure 1).

5.1 State-Specific Connectedness Results

We begin by examining total connectedness or the average effect a state exerts on all other states (equivalently, the average effect of all other states on a given state). The risk in the housing market, as measured by total connectedness, steadily increased since 1995, reaching its high in 2007 and remained at that level until it slowly began decreasing from 2015 onwards (see Figure 1). The three methods that penalized parameter estimates using the Lasso, Ridge, and Net Elastic methods of penalization trended generally together, not differing by too much. The standard OLS results,

[3]Complete details of the unit root tests are available upon request from the authors.

however, produced higher total connectedness than the Lasso, Ridge, and Net Elastic methods during the beginning of the sample period through 2007. Then, dynamic total connectedness did not differ much across the four models for the high volatility period followed by the Great Recession and Financial Crisis of 2007-2009.[4]

[Insert Figure 1 around here]

Figure 2 provides an overview of the total U.S. housing market spillover. That is, for each state, we calculate the total directional connectedness to other states minus total directional connectedness from other states. Thus, a positive (negative) number means that the shocks from the state in question exert a larger (smaller) effect on all other sates than the effect of all other states' shocks on this state. Further, Figure 2 uses lighter colors for more of a transmitter of shocks to other states and darker colors for more of a receiver of shocks from other states.

We can see that the main transmitters of shocks include Arkansas (AR), Colorado (CO), Ohio (OH), and Washington (WA) whereas the main receiver of shocks include Hawaii (HI), New York (NY), North Carolina (NC), West Virginia (WV), and Maine (ME). Even though this map provides a good overview of the spillovers, it will be interesting to see how the spillovers behave on a regional level, which we show below.

[Insert Figure 2 around here]

To see whether a state, on average, significantly transmits or receives shocks, we calculated t-tests for each state and determined the 99% confidence interval (see Figure 3). Figure 3 shows

[4]Heightened connection during the global financial crisis suggests that the housing market across the US states and regions despite its heterogeneities, comove strongly during periods of slowdown in the market. We investigated this issue further, by estimating estimating a Bayesian dynamic factor model (DFM) to deduce the importance of the common component in the house price movements relative to state-specific shocks as in Del Negro and Otrok (2007). Once we recovered this national component from our year-on-year growth rate of real house prices, and computed the correlation with the total connectedness measures derived from the four models, we found negative correlation in all the cases in a statistically significant manner. Complete details of these results are available upon request from the authors. It must be noted that, similar to our finding, Ngene et al. (2017) found support for the hypothesis that herding effects at the state-level US housing returns are more prominent during the recession periods.

that the VAR approach based on the OLS and Ridge regressions sometimes differ significantly from the VAR approach based on the Lasso and Elastic Net regressions. We can see this for SD, HI, and NE. This outcome occurs because the OLS does not penalize the parameter estimates at all and the Ridge regression squared parameters are penalized parameter estimates by a squared factor. We know that the parameters in a VAR model usually vary between zero and one. Hence, the penalization of the Ridge regression does not exert much of a penalty unless the parameter estimate is close to one. The Lasso regression, however, penalizes parameters between zero and one more severely as the absolute value and not its squared value is penalized. The fact that the Elastic Net based VAR model closely approximates the Lasso based VAR results indicates that the Elastic Net parameter gives the Lasso penalization term more weight than the Ridge regression parameter. Thus, the results of the Elastic Net approach should provide the most reliable findings as the estimation process allows in every step to decide whether the Elastic Net should weight the results of the Lasso or the Ridge regression more. This illustrates that the Lasso regression should be preferred over Ridge regression in applied time series econometrics that use the VAR model or other time-series modeling techniques, where the parameter estimates should lie between zero and one. This theoretical consideration is supported by the data as the Ridge regression results differ the most from all others and is a valuable information for researchers and practitioners.[5]

[Insert Figure 3 around here]

5.2 Regional-Specific Connectedness Results

We aggregate the state spillovers to the regional level to get an overview of the interdependencies across regions Figure 4 illustrates the dynamic total connectedness for each region without considering the interregional spillovers. The findings suggest that the Northeastern and Southern regions appear highly connected throughout the period of analysis whereas the Midwestern and

[5] Hence, from here onwards we will disregard the Ridge regression results and mainly focus on the findings of the Elastic Net model. Empirical results of the other models are available upon request.

Western regions appear less interconnected and, hence, exhibit a lower housing market risk. This has important implications for portfolios and risk management as the risk of MBSs and CDOs can be reduced by investing more into Western and Midwestern regional mortgages as the markets are not as highly synchronized as the markets in the Northeastern and Southern regions. Notably, the subprime market crisis of 2006 that led to the Great Recession and Financial Crisis 2008-2009. It appears that the subprime market crisis started in the Northeastern and Southern regions and fueled over time the risk in the Midwestern region. Moreover, for the Western region, the sharp increase could indicate that its risk was mainly caused by the contagion effect of the other three regions. Further, the Western region lived through another increase in market risk in 2012, which could mark the further deterioration of the U.S. housing market via defaulted loans. Finally, after 2015, all regions start decreasing together until the end of the sample period, whereby the level of risk concerning the Midwestern and Western regions still substantially exceeds the pre-crisis period.

[Insert Figure 4 around here]

In addition to intraregional connectedness, the spillovers across regions are also of major interest. Figure 5 illustrates interregional dynamic total connectedness, which shows a substantial increase in 2008 due to the U.S. subprime market crisis. This sudden increase in the U.S. housing interrelatedness shows the contagious effect the crisis had nationwide.

[Insert Figure 5 around here]

Table 1 reports the average connectedness measures on a regional level – Northeast, Midwest, South, and West. We find that the Northeastern region is the main receiver of shocks whereas the South is the main transmitter of shocks followed by the Western and Midwestern regions.

[Insert Table 1 around here]

The OLS, Lasso, and Elastic Net results confirm that the South is the net pairwise transmitter to all others followed by the West and Midwestern region living the Northeastern region as the sole receiver of all net pairwise spillover shocks. This behavior is visualized in Figure 6.

[Insert Figure 6 around here]

5.3 Northeastern, Southern, Midwestern, and Western Regional Spillovers

Finally, to complete the picture, we isolate each region and identify the spillover effects within each region by itself. Figures 7, 8, 9 and 10 show the spillovers from the Northeastern, Southern, Midwestern, and Western regions, respectively, on a state-by-state basis. In the Northeast region, New Hampshire (NH) is the largest net transmitter of shocks whereas New York (NY) is the largest net receiver of housing price shocks. Comparing our findings to all states (Figure 1) and regional (Figure 6) spillover effects, we note that New York State is the one constant as a net receiver of spillover effects from all states and all regions.

In the Southern region, Arkansas (AR) is the largest transmitter of shocks whereas North Carolina (NC) and West Virginia (WV) are the largest receivers of shocks. Comparing our findings to all states (Figure 1) and regional (Figure 6) spillover effects, we note that Arkansas is the one constant as a net transmitter of spillover effects from all states and all regions, although when we consider all states Arkansas trails behind Ohio as a net transmitter.

In the Midwestern region, Ohio (OH) is the largest transmitter of shocks whereas Iowa (IA) and Minnesota (MN) are the largest receivers of shocks. Comparing our findings to all states (Figure 1) and regional (Figure 6) spillover effects, we note that Ohio is a net transmitter of spillover effects from all states but a net receiver from the South and the West and a net transmitter to the Northeast when we consider the regions.

In the Western region, Colorado (CO) and Washington (WA) are the largest transmitters of shocks whereas Oregon (OR) is the largest receiver of shocks. Comparing our findings to all states

(Figure 1) and regional (Figure 6) spillover effects, we note that Colorado and Washington are net transmitters of spillover effects and Oregon is a net receiver of spillover effects from all states and all regions. In this case, however, the net transmitter status of Colorado and Washington and the net receiver status of Oregon strengthen when we examine the isolated region of the West.

[Insert Figure 7, 8, 9 and 10 around here]

In sum, returning to Figure 2 and the description of the largest senders and receives of shocks for all 50 states and the District of Columbia simultaneously, we see a large overlap with the regional findings.

6 Conclusions

We examine the connectedness between real housing return shocks across the 50 U.S. states and the District of Columbia as well as four regions (Midwest, Northeast, South, and West) using a VAR modeling approach that selects parameter estimates using Lasso, Ridge, and Elastic Net methods.

Total connectedness steadily increased since 1995, peaking in 2007 and remained at that level until 2005 when it slowly began decreasing. The Lasso, Ridge, and Net Elastic methods of penalization trended generally together, not differing by too much. The standard OLS results, however, produced higher total connectedness than the other three methods during the beginning of the sample period through 2007. Then, dynamic total connectedness did not differ much across the four models for the high volatility period followed by the Great Recession and Financial Crisis of 2007-2009.

On a state basis, the main transmitters of shocks include Arkansas, Colorado, Ohio, and Washington whereas the main receiver of shocks include Hawaii, New York, North Carolina, West Virginia, and Maine.

At the regional level, we find that the South is a net transmitter of real housing return shocks to the other three regions, whereas the Northeast is a net receiver of those shocks. The West is a net receiver of shocks from the South and a net transmitter of shocks to the Midwest and the Northeast. Finally, the Midwest is a net transmitter of shocks to the Northeast and a net receiver of shocks from the South and the West.

From the perspective of policy decisions, policymakers need to monitor the behavior of housing market movements in states like Arkansas, Colorado, Ohio, and Washington, and the South region closely, as these markets serve as transmitters of shocks. At the same time, these markets also provide investment opportunities for housing market investors. In this regard, note that with overall market connectedness being high during periods of crises, diversification opportunities for investors across regional markets diminish.

Future research could examine the national and regional factors that drive connectedness. Clearly, this information will help policymakers to design sector-specific policies.[6]

[6] Preliminary analysis using national-level variables indicate that measures of economic (Jurado et al., 2015) and real estate uncertainties Nguyen Thanh et al. (2018) are more important in driving connectedness than macroeconomic variables such as output growth, inflation, and interest rates. These results are available upon request from the authors.

References

Antonakakis, N., Chatziantoniou, I., Floros, C. and Gabauer, D. (2018), 'The dynamic connectedness of UK regional property returns', *Urban Studies* **55**(14), 3110–3134.

Antonakakis, N., Chatziantoniou, I. and Gabauer, D. (2019), A regional decomposition of US housing prices and volume: Market dynamics and economic diversification opportunities, Technical report, University of Portsmouth, Portsmouth Business School, Economics and Finance.

Balcilar, M., Gupta, R. and Miller, S. M. (2014), 'Housing and the great depression', *Applied Economics* **46**(24), 2966–2981.

Chiang, M.-C. and Tsai, I.-C. (2016), 'Ripple effect and contagious effect in the US regional housing markets', *Annals of Regional Science* **56**(1), 55–82.

Clapp, J. M. and Tirtiroglu, D. (1994), 'Positive feedback trading and diffusion of asset price changes: Evidence from housing transactions', *Journal of Economic Behavior & Organization* **24**(3), 337–355.

Cook, S. (2003), 'The convergence of regional house prices in the UK', *Urban Studies* **40**(11), 2285–2294.

Cook, S. (2005), 'Regional house price behaviour in the UK: Application of a joint testing procedure', *Physica A: Statistical Mechanics and its Applications* **345**(3-4), 611–621.

Del Negro, M. and Otrok, C. (2007), '99 Luftballons: Monetary policy and the house price boom across US states', *Journal of Monetary Economics* **54**(7), 1962–1985.

Demirer, M., Diebold, F. X., Liu, L. and Yilmaz, K. (2018), 'Estimating global bank network connectedness', *Journal of Applied Econometrics* **33**(1), 1–15.

Diebold, F. X. and Yilmaz, K. (2012), 'Better to give than to receive: Predictive directional measurement of volatility spillovers', *International Journal of Forecasting* **28**(1), 57–66.

Emirmahmutoglu, F., Balcilar, M., Apergis, N., Simo-Kengne, B. D., Chang, T. and Gupta, R. (2016), 'Causal relationship between asset prices and output in the United States: Evidence from the state-level panel Granger causality test', *Regional Studies* **50**(10), 1728–1741.

Fairchild, J., Ma, J. and Wu, S. (2015), 'Understanding housing market volatility', *Journal of Money, Credit and Banking* **47**(7), 1309–1337.

Gupta, R. and Miller, S. M. (2012*a*), 'The time-series properties of house prices: A case study of the southern California market', *Journal of Real Estate Finance and Economics* **44**(3), 339–361.

Gupta, R. and Miller, S. M. (2012*b*), '"Ripple effects" and forecasting home prices in Los Angeles, Las Vegas, and Phoenix', *Annals of Regional Science* **48**(3), 763–782.

Holmes, M. J. and Grimes, A. (2008), 'Is there long-run convergence among regional house prices in the UK?', *Urban Studies* **45**(8), 1531–1544.

Jurado, K., Ludvigson, S. C. and Ng, S. (2015), 'Measuring uncertainty', *American Economic Review* **105**(3), 1177–1216.

Koop, G., Pesaran, M. H. and Potter, S. M. (1996), 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics* **74**(1), 119–147.

Leamer, E. E. (2015), 'Housing really is the business cycle: What survives the lessons of 2008–09?', *Journal of Money, Credit and Banking* **47**(S1), 43–50.

Leamer, E. E. et al. (2007), Housing is the business cycle, *in* 'Proceedings-Economic Policy Symposium-Jackson Hole', Federal Reserve Bank of Kansas City, pp. 149–233.

Meen, G. (1999), 'Regional house prices and the ripple effect: A new interpretation', *Housing Studies* **14**(6), 733–753.

Ngene, G. M., Sohn, D. P. and Hassan, M. K. (2017), 'Time-varying and spatial herding behavior in the US housing market: Evidence from direct housing prices', *Journal of Real Estate Finance and Economics* **54**(4), 482–514.

Nguyen Thanh, B., Strobel, J. and Lee, G. (2018), 'A new measure of real estate uncertainty shocks', *Real Estate Economics* .

Nyakabawo, W., Miller, S. M., Balcilar, M., Das, S. and Gupta, R. (2015), 'Temporal causality between house prices and output in the US: A bootstrap rolling-window approach', *North American Journal of Economics and Finance* **33**, 55–73.

Pesaran, H. H. and Shin, Y. (1998), 'Generalized impulse response analysis in linear multivariate models', *Economics Letters* **58**(1), 17–29.

Samuelson, P. A. (1948), 'International trade and the equalisation of factor prices', *Economic Journal* **58**(230), 163–184.

Shiller, R. J. (2012), *The subprime solution: How today's global financial crisis happened, and what to do about it*, Princeton University Press.

Sims, C. A. (1980), 'Macroeconomics and reality', *Econometrica* pp. 1–48.

Tibshirani, R. (1996), 'Regression shrinkage and selection via the Lasso', *Journal of the Royal Statistical Society: Series B* **58**(1), 267–288.

Tirtiroğlu, D. (1992), 'Efficiency in housing markets: Temporal and spatial dimensions', *Journal of Housing Economics* **2**(3), 276–292.

Tsai, I.-C. (2015), 'Spillover effect between the regional and the national housing markets in the UK', *Regional Studies* **49**(12), 1957–1976.

Tsai, I.-C. (2018), 'The cause and outcomes of the ripple effect: Housing prices and transaction volume', *Annals of Regional Science* **61**(2), 351–373.

Tsai, I.-C. (2019), 'Interregional correlations in the US housing market at three price tiers', *Annals of Regional Science* **63**(1), 1–24.

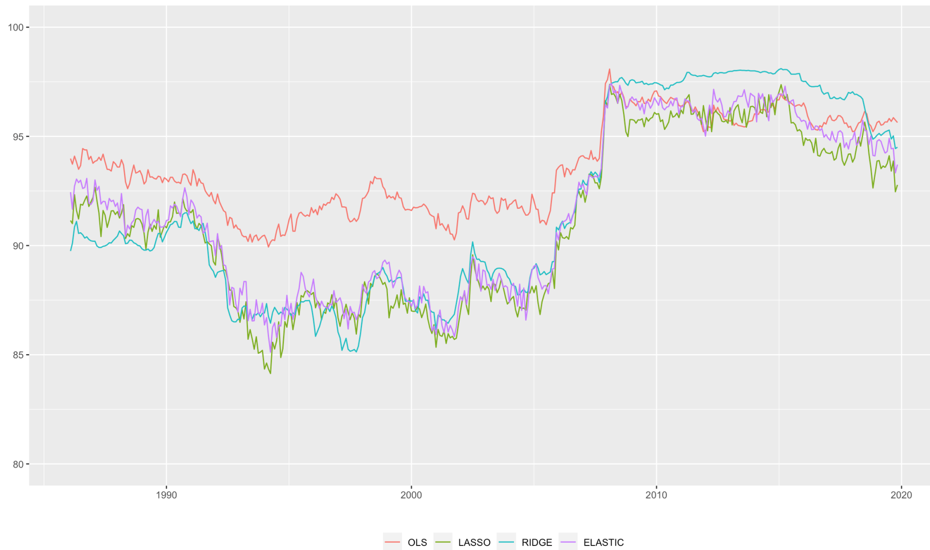
Wiesen, T. F., Beaumont, P. M., Norrbin, S. C. and Srivastava, A. (2018), 'Are generalized spillover indices overstating connectedness?', *Economics Letters* **173**, 131–134.

Zou, H. and Hastie, T. (2005), 'Regularization and variable selection via the Elastic Net', *Journal of the Royal Statistical Society: Series B* **67**(2), 301–320.

Table 1: Aggregated Regional Connectedness Table

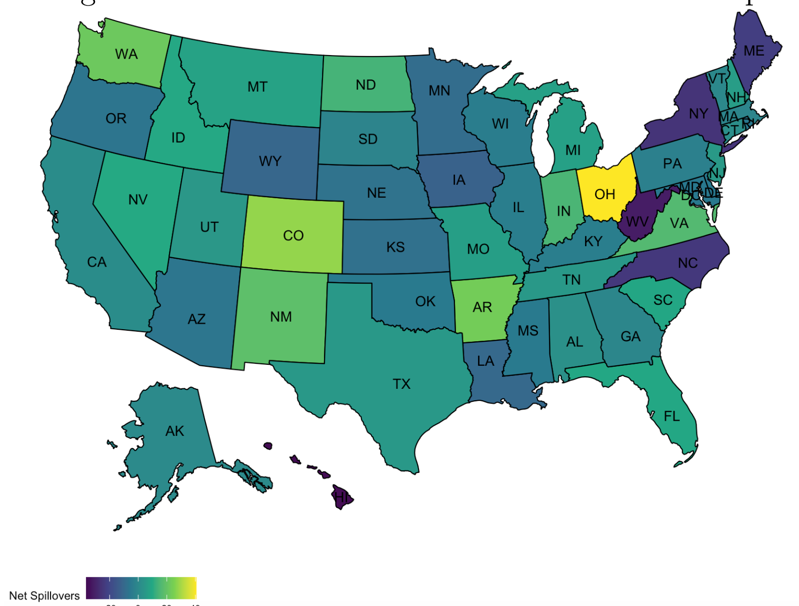
	Northeast	South	Midwest	West	FROM
Northeast	62.4	12.2	12.3	13.1	37.6
South	7.3	72.5	9.9	10.3	27.5
Midwest	6.9	11.6	70.5	11.0	29.5
West	6.7	11.1	10.6	71.6	28.4
TO	20.9	34.9	32.7	34.4	122.9
Net Spillovers	-16.7	7.5	3.2	6.0	TCI
Net Pairwise Transmission	0	3	1	2	41.0

Figure 1: Dynamic Total Connectedness



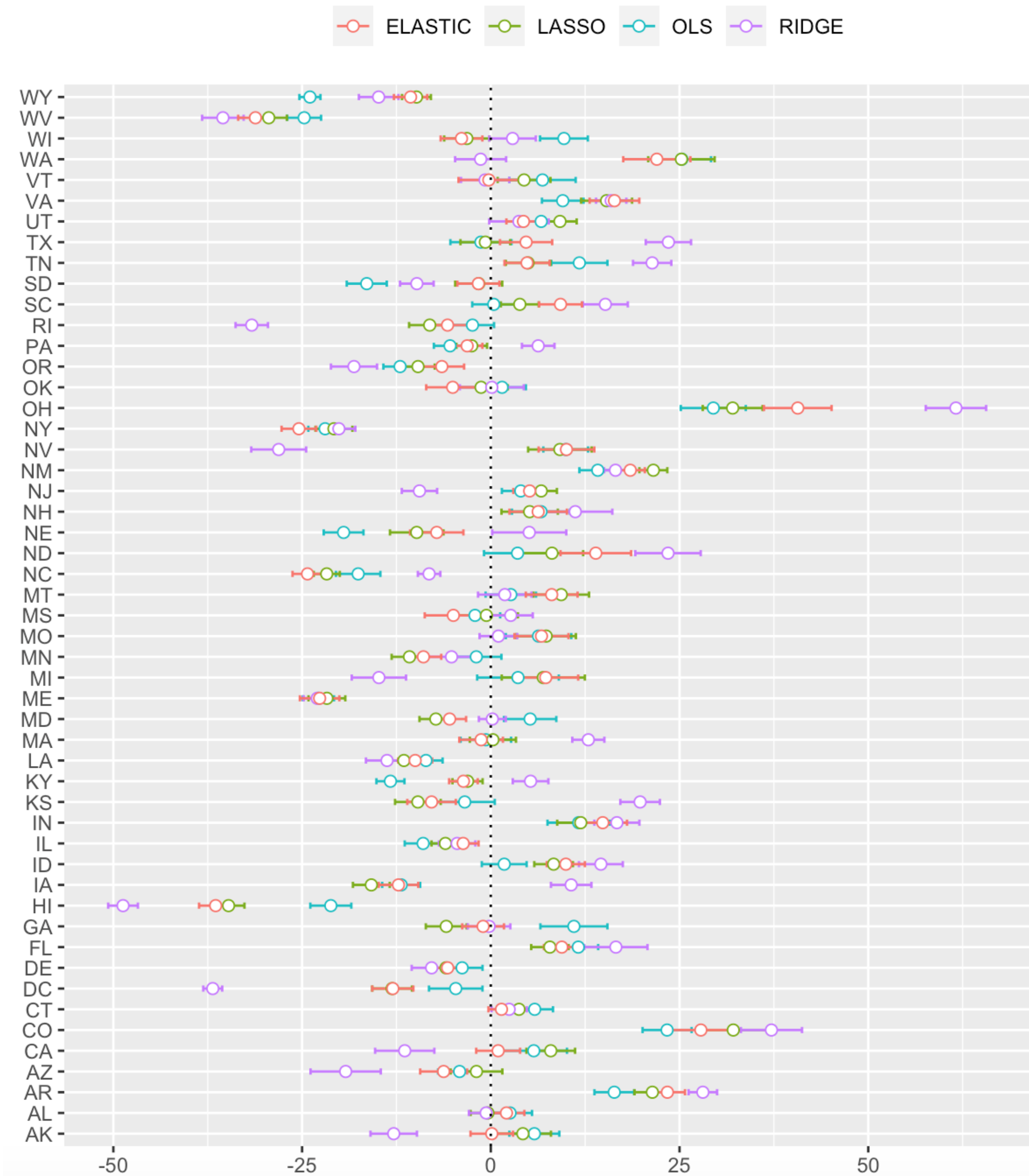
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 2: Net Total Directional Connectedness Map



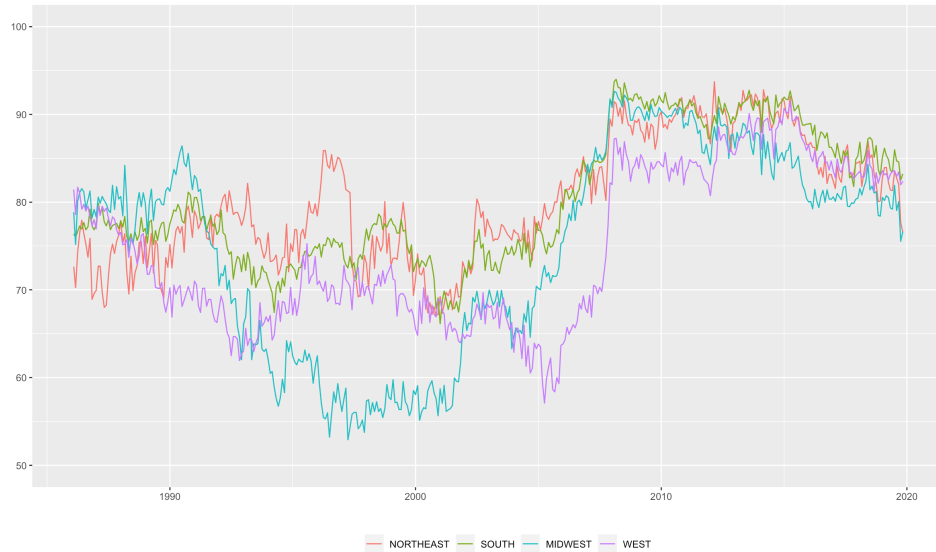
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 3: Average Net Total Directional Connectedness



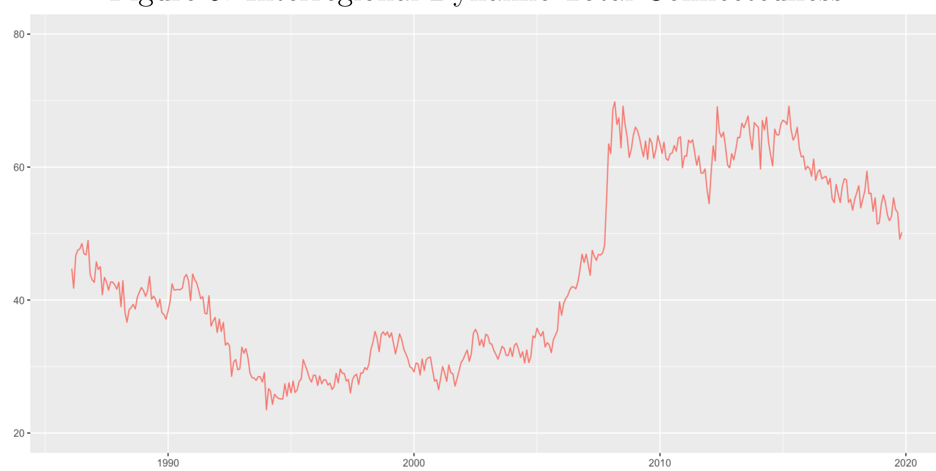
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon. The means and their corresponding 95% confidence intervals are shown.

Figure 4: Intraregional Dynamic Total Connectedness



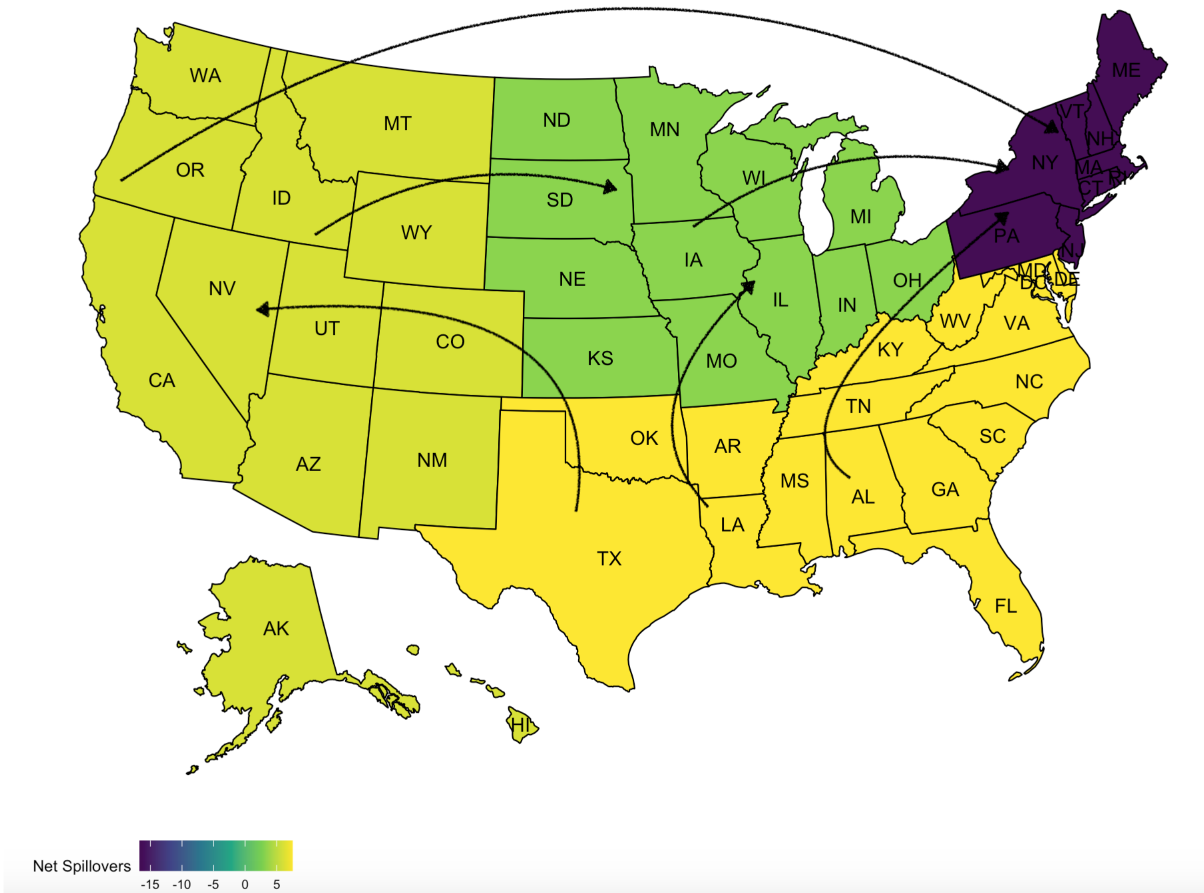
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 5: Interregional Dynamic Total Connectedness



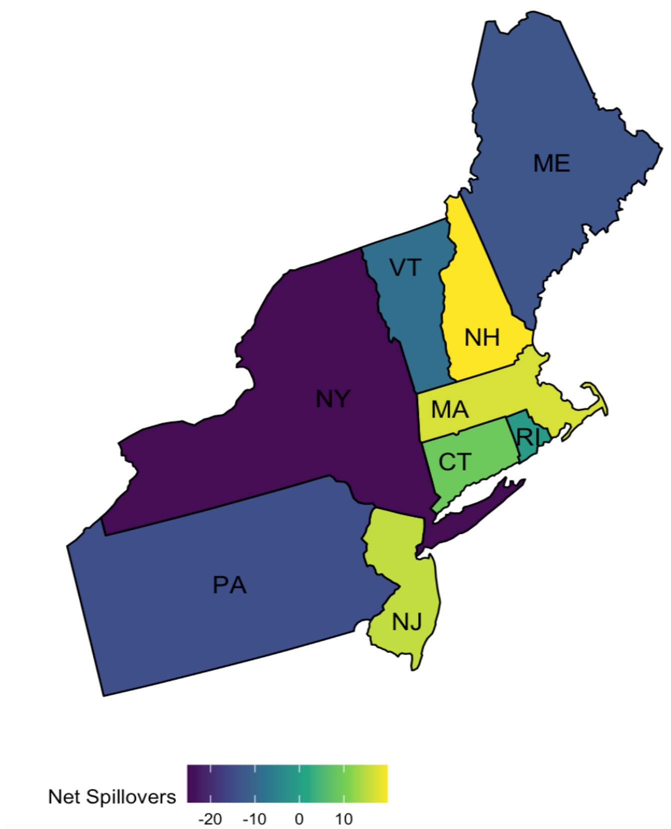
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 6: Regional Dynamic Total Connectedness



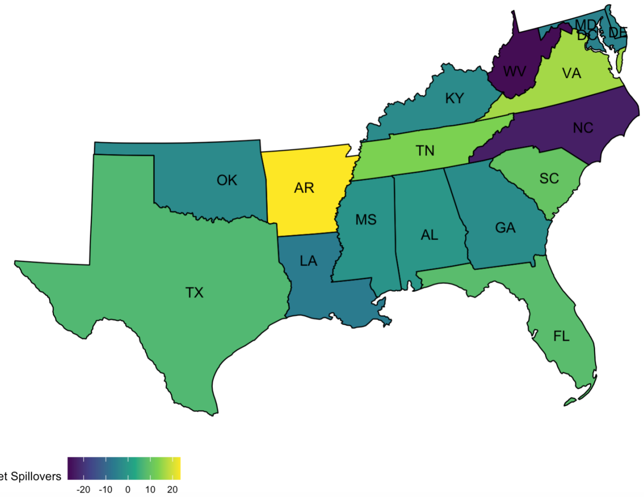
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 7: Northeastern Net Total Dynamic Connectedness (I)



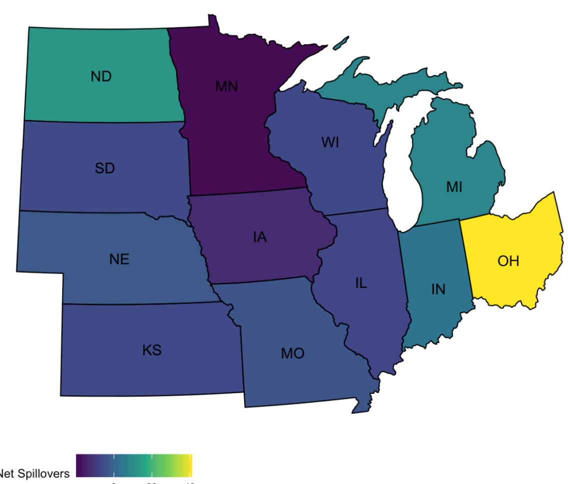
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 8: Southern Net Total Dynamic Connectedness (II)



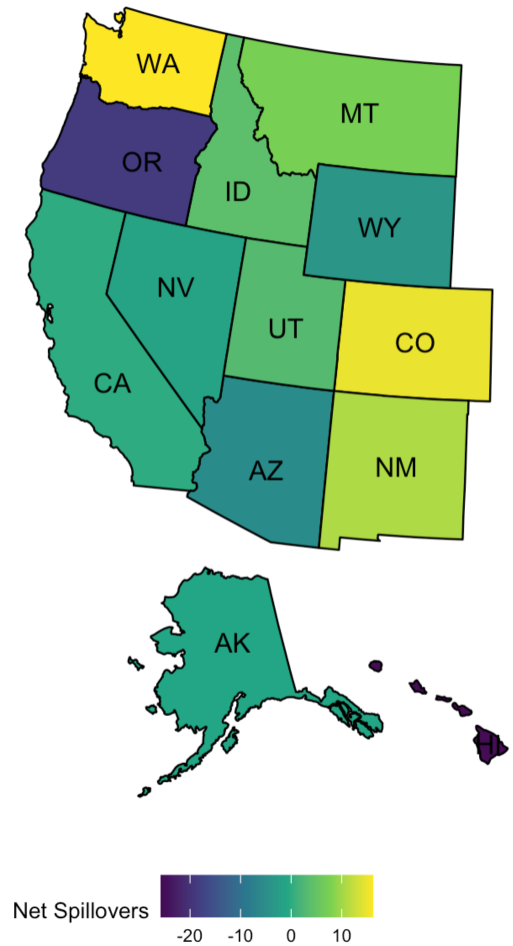
Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 9: Midwestern Net Total Dynamic Connectedness (III)



Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.

Figure 10: Western Net Total Dynamic Connectedness (IV)



Notes: Results are based on all 120 month rolling-window models with a 10 step-ahead forecast horizon.