

Metropolitan Home Price Dynamics Untied from Observable Fundamentals and Their Linkages^{*}

Roland Füss[†], Bing Zhu[§], and Joachim Zietz[‡]

Working Paper

First version: April 2010

This version: December 2012

^{*} A previous version of this paper was circulated under the title: “U.S. Regional Housing Bubbles, Their Co-movements and Spillovers”.

[†] Corresponding author: Swiss Institute of Banking and Finance, University of St.Gallen, St. Gallen, Switzerland and Research Associate at the Centre for European Economic Research (ZEW), Mannheim, Germany; Phone: +41 (0)71 224 7055; Fax: +41 (0)71 224 7088; Email: roland.fuess@unisg.ch.

[§] Bing Zhu, Department of Real Estate Finance, University of Regensburg, 93053 Regensburg, Germany; email: bing.zhu@wiwi.uni-regensburg.de.

[‡] Department of Finance, Accounting and Real Estate, EBS Business School, EBS Universität für Wirtschaft und Recht, Gustav-Stresemann-Ring 3, D-65189 Wiesbaden, Germany; Middle Tennessee State University, Murfreesboro, TN, USA; email: joachim.zietz@gmail.com.

Acknowledgments: The authors thank John Glascock, Christian Hott, Jan Mutl, Seow Eng Ong, Andrew Paciorek, Marc W. Simpson, Bernd Süßmuth and participants of the doctoral seminar at the University of Leipzig, the American Real Estate and Urban Economics Association 2011 Annual Meeting, the 2011 Annual Meeting of the Midwest Finance Association, the 2011 Annual Meeting of German Finance Association, the ReCapNet Conference 2011, and the 2012 AREUEA Mid-Year conference for useful suggestions on previous versions of the paper. Part of this research was completed while the first author was visiting the ICMA centre at University of Reading. We alone are responsible for any errors.

Metropolitan Home Price Dynamics Untied from Observable Fundamentals and Their Linkages

Abstract

This paper examines the co-movements in unexplained run-ups in home price inflation among the 20 U.S. metropolitan statistical areas (MSAs) of the Case-Shiller housing price index over the period 1995 to 2008. We employ two stages. First, we use unobserved component modeling, separately by MSA, to decompose each observed home price index into a component explained by market fundamentals and an unexplained price run-up. Second, we analyze the interdependence among the estimated price run-ups using spatial panel data methods. We find that both demand- and supply-side influences play a role in linking run-ups in home price inflation across MSAs.

Keywords: Unobserved component modeling; excess home price inflation; spatial weight matrices; spatial panel data

JEL Classification: *C21; C23; R15*

“We’ve never had a decline in house prices on a nationwide basis.”

Mr. Bernanke on CNBC in 2005

1 Introduction

The run-up in housing prices prior to 2006 was far from evenly distributed across the U.S. Some states and MSAs experienced unprecedented boom times, others were hardly affected at all. For instance, in markets such as Los Angeles and Miami housing prices doubled on average from 1998 to 2006. Yet, very little price movement was observed over the same time horizon in some interior markets, such as Dallas and Denver.¹ These conditions made it natural to adopt the working hypothesis that local housing markets were adjusting to local demand and supply shocks.² The fact that national aggregate housing price indices showed clear signs of an unusual run-up in prices prior to 2006 could have been taken as a warning sign. But generally the idea prevailed that housing prices are a local or at best a regional phenomenon and not of national concern.³ As a consequence, no need was perceived by policy makers to intervene at the national level.

With hindsight, this interpretation was unfortunate and had strong consequences, as we now know. The starting point of this study is the idea that the run-ups in home prices in numerous regions and MSAs prior to 2006 were more than just coincidental and that in two

¹ Wheaton and Nechayev (2008) report that real home prices increased by 74% in Boston, 10% in Los Angeles, 11% in Chicago and decreased by 21% in Dallas and by 38% in Houston from 1980Q1 to 1998Q4. In contrast, from 1999Q1 to 2005Q4 the increases were 83% in Boston, 123% in Los Angeles, 42% in Chicago, but only 12% and 19% in Dallas and Houston, respectively. See also Glaeser, Gyourko, and Saks (2005) who argue that the dispersion in housing prices has increased substantially since 1970, and mainly in the upper tail of the housing price distribution.

² In this context, Greenspan argued in 2005 that the U.S. was not experiencing a nationwide housing bubble per se, but a number of local bubbles. However, in 2007 Greenspan admitted that “all the froth bubbles add up to an aggregate bubble.” *Financial Times*, September 17, 2007.

³ The fact that strong spatial interactions among house prices have been demonstrated empirically only for geographically close markets backed up this view (Tirtiroglu, 1992; Clapp and Tirtiroglu, 1994; Miao, Ramchander, and Simpson, 2011).

respects. First, MSAs or regions were not affected randomly; those experiencing unusual price run-ups had some common characteristics. Second, there exists a plausible link of the regional price run-ups to national policy choices, in particular the low interest rate policy of the Federal Reserve starting in 2000 and the de-regulation of mortgage origination and securitization.⁴

The national policy link can be briefly summarized as follows. Historically low interest rates induced investors to look for higher yielding investment. The housing market looked very attractive in this respect, in particular after the dot-com market crash at the beginning of the new millennium. Investments in housing supposedly have intrinsic value as land is finite, in particular in areas with natural barriers to further development and an attractive quality of life. Housing also had a very long track record of rising prices, in addition to offering favorable tax treatment for direct investors.

The entry of direct investors into the housing market, such as those looking for second homes for tax reasons or baby-boomers trying to secure retirement homes ahead of further price increases, would probably not have induced the types of price run-ups experienced in parts of the country. More of an influence on home prices originated likely with the large numbers of indirect investors who were lured into the housing market by the aggressive securitization of mortgages and the availability of convenient risk classes, with very competitive returns even on those securities branded as effectively risk free. As local and regional housing markets that looked attractive from an investor's perspective were flooded with funds, home prices were rising where supply could not keep pace. This in turn

⁴ The mid-2000s U.S. housing bubble was closely related to aggressive mortgage lending practices and relaxed mortgage requirements (Pavlov and Wachter, 2010; Dell'Araccia, Igan, and Laeven, 2009). The traditional banking model became less profitable and the banking system transformed from "originate and hold" to "originate and distribute." At the same time, the supply of ABS and the demand for alternatives to insured deposits led to strong growth of the shadow banking system. In addition, capital requirements were effectively removed for investment banks in 2004 for their securitization business (Calomiris, 2010).

convinced many more investors of the gains to be made in the housing market and laid the foundation for unusual home price run-ups.

Of key interest for this study is the fact that the funds from direct and indirect investors flowing into the housing market did not arrive in the form of a tsunami affecting all regions or MSAs indiscriminately and raising prices everywhere. Instead the funds flowed into regions or MSAs with significant potential for appreciating home values. If national investors are in fact involved, as we suggest, those MSAs with price run-ups should not be a random sample of all MSAs. Rather, they should have common characteristics that are observable to investors. As a consequence, we expect strong linkages of price run-ups among markets that have similar characteristics in the eyes of investors and far less pronounced linkages with markets without these characteristics.⁵

The purpose of this paper is to test for the existence and the nature of linkages among price run-ups across MSAs that go beyond those implied by local or regional economics and geography. If such linkages do indeed exist, one could argue that there is a national policy dimension to local or regional price run-ups in the sense that investors from outside the area are involved, who respond to financial incentives that are created by national policy choices. Learning more about the characteristics of the regional linkages of unusual price run-ups appears therefore a pre-requisite for differentiating in practice truly local boom markets, which are outside the scope of national economic policy, from those that may need the attention of national policy makers.

Only a few studies analyze the regional linkages among price movements that cannot be explained by market fundamentals. Fry (2009) attempts to empirically demonstrate

⁵ Spatial linkages may arise for a variety of reasons other than the potential for capital gains in the eyes of investors. Among those other reasons are geographical closeness (Topol, 1991), learning (Akerlof and Shiller, 2009), or the migration of home buyers and the responses of home suppliers (Gupta and Miller, 2010; Miao, Ramchander, and Simpson, 2011).

contagion effects among regional housing bubbles. Based on a bivariate stochastic model, he seeks to identify the contagion effects that developed during the U.K. housing market bubble over the period 2002 to 2007. He finds conclusive evidence for a nationwide bubble, but no evidence for regional contagion. Costello, Fraser, and Groenowold (2011) utilize a dynamic present value model within a VAR framework to estimate the fundamental housing price in Australian capital cities from 1984Q3 to 2008Q2. They demonstrate spillover effects among the non-fundamental prices, where the non-fundamental component is estimated as the difference between the actual price and fundamental price.

In a similar manner, Hott and Monnin (2008) compare estimated and actual prices to detect over- and undervalued international housing prices. They find that housing prices deviate substantially and persistently from their estimated fundamental values and return to their fundamental values only sluggishly. The authors further demonstrate that forecasting models that include fundamental prices outperform systematically those based on dynamic price models for medium- and long-term time horizons.

Kallberg, Liu, and Pasquariello (2011) examine 14 metropolitan areas for the period from 1992 to 2008. They distinguish between fundamental and excessive co-movements among housing prices, where the latter cannot be explained by common fundamental pricing factors. The authors find that the increasing covariation among the MSAs over their sample period is mainly related to systematic real and financial risk factors rather than to excess co-movements. They relate this phenomenon to an increasing integration of housing markets, similar to the integration of international financial markets.

Holly, Pesaran, and Yamagata (2011) analyze the spatio-temporal diffusion of housing prices between London and 11 regions from 1974 to 2008. By modeling the temporal and spatial dependence in a non-stationary dynamic system, the authors show that London plays a dominant role in propagating shocks contemporaneously and spatially to other U.K. regions.

However, using a spatial-temporal impulse response approach, the authors demonstrate that the decay of innovations is slower along the geographical dimension, i.e. for farther regions, than along the time dimension. In contrast, London is also affected by shocks stemming from New York's development in housing prices due to the role of both cities as global financial centers.

This study helps to fill an important gap in the literature by examining not only to what extent unexplainable price run-ups exist in several major housing markets, but also how these are connected across space. In particular, as unusual home price inflation occurred concurrently in MSAs far away from each other, such as San Diego, Miami, and Washington DC, we argue that simple geographical distance cannot be the key explanatory variable. Instead we argue that the linkages must be related to the potential of an MSA for price appreciation, because that is what ultimately is of interest to investors in the housing market. If the potential for price appreciation is the key driver of linkages across MSAs, then it is apparent that both demand- and supply-side variables must play a role in linking MSAs with unusual price run-ups. On the demand side, we would expect to find variables that make an MSA particularly attractive to live in; on the supply side, it would be variables that make for a low supply elasticity.

We identify linkages among unexplained price run-ups across MSAs through a two-step process. We first employ a standard unobserved component model (UCM)⁶ to identify what we call unusual price run-ups in the previous discussion. Subsequently, we take these local stochastic trend variables and estimate interactions among them with the help of a spatial panel model utilizing for that purpose a variety of weight matrices that try to capture our idea that price run-ups are driven in principle by a confluence of strong demand and a sluggish

⁶ UCMs are also known as structural time series models. They are a special case of general state space models. They were introduced to economics by Harvey (1989) and further developed, among others, by Durbin and Koopman (2001).

supply response. Our results identify population growth and supply inelasticity as the main linkages of price run-ups across MSAs. However, there are also linkages among the MSAs of a region that cannot be easily explained by observable variables with economic content.

2 Strategy to Identify Linkages among MSAs

To investigate the spatial linkages among the observed run-ups in housing prices of MSAs, we use a two-stage estimation strategy that is similar in principle although different in methodology to Costello, Fraser, and Groenowold (2011) and Kallberg, Liu, and Pasquariello (2011). At the first stage, we decompose each of our 20 observed MSA housing price series, one at a time, into (a) what can be predicted by local and national market fundamentals, i.e. an observed component related to market fundamentals, (b) what can be predicted but not with market fundamentals, i.e. a stochastic trend component, and (c) what cannot be predicted and, hence, must be classified as a white-noise error component. The stochastic trend component is our proxy for unusual price run-ups that are not explained by market fundamentals. It is based on an estimated model rather than just being a residual.⁷ At the second stage, we estimate, in a spatial panel data setting, the linkages among the price run-ups. By testing a large number of alternative spatial weight matrices at the second stage, we can identify the economic drivers that are responsible for the fact that some MSAs share similar price run-ups while others do not.

2.1 First Stage: Decomposing the Observed Housing Price

We use a decomposition method, alternatively known as structural time series or unobserved component modeling, to split the observed price series into (a) a regression

⁷ For example Kallberg, Liu, and Pasquariello (2011) utilize the SUR model to control for fundamental co-movements across 14 MSA return series. Correspondingly, they interpret return co-movements between two cities as excessive if the corresponding residuals from the linear factor structure are correlated. In the present paper, we are interested in explaining rather than merely detecting covariation among MSAs.

component associated with local and national market fundamentals, (b) a parameterized stochastic trend component to absorb non-stationary ups and downs of the price unrelated to local or national market fundamentals, and (c) a white noise error component.⁸ For computational tractability, this estimation is done for each MSA separately. The predicted parameterized stochastic trends are then used as dependent variables at the second estimation stage, where the nature of the spatial linkages among them is examined.

The regression component of our first-stage model intends to capture the impact of local and national market fundamentals. We derive the regressors from a standard equilibrium model for the single-family housing market (DiPasquale and Wheaton, 1996; Goodman and Thibodeau, 2008; Adams and Füss, 2010).⁹ Within this framework, we postulate the following long-run supply and demand model,

$$\ln Q^S = \alpha_1 \ln P^f + \alpha_2 \ln CS + \alpha_3 \ln HS, \quad (1)$$

$$\ln Q^D = \beta_1 \ln E - \beta_2 \ln rP^f - \beta_3 \ln UE + \beta_4 \ln IP. \quad (2)$$

Equation (1) represents the housing supply in an MSA market, where Q^S indicates the supply of housing units. Quantity supplied is a function of the fundamental housing price (P^f), as measured at the MSA level and, at the national level, the residential construction activity (CS). With the latter variable we try to capture national trends that drive the local housing market, such as building costs, changes in laws pertaining to housing construction and use,

⁸ We note that all model parameters, those of the regression component and those of the stochastic trend, are estimated simultaneously through a combination of maximum likelihood techniques and Kalman filtering, similar to Wu (1997), Elwood, Ahmed, and Rosser (1999), and Al-Anaswah and Wilfling (2009). This approach is different from estimating an OLS regression on hypothesized market fundamentals and treating the residual as the series to work with at the second stage.

⁹ For a more detailed discussion see DiPasquale and Wheaton (1996). For the use of different macroeconomic indicators compare the studies of DiPasquale and Wheaton (1994), Abraham and Hendershott (1993, 1996), Capozza, Hendershott and Mack (2004), Case and Shiller (2003), Malpezzi and Wachter (2005), Goodman and Thibodeau (2008) and Costello, Fraser, and Groenowold (2011), as well as Kallberg, Liu, and Pasquariello (2011).

and the general political and economic climate within which home builders make their decisions. As a final driver of housing supply we include single family housing starts at the MSA level (HS); α_1 , α_2 , and α_3 are the corresponding coefficients of the supply variables, with $\alpha_k > 0$.

Equation (2) specifies the determinants of MSA housing demand. Demand (Q^D) is taken to be a function of real income and population as proxied by employment at the MSA level (E), and the user cost of housing, as measured by the house price (P^f) at the MSA level multiplied by the effective mortgage rate at the national level (r). The size of the MSA housing market is also approximated by employment at the MSA level, while the MSA unemployment rate (UE) and industrial production at the national level (IP) reflect changes in general business conditions; β_1 to β_5 are the coefficients to be estimated, with $\beta_k > 0$.

In equilibrium, we have $\ln Q^S = \ln Q^D$, which implies that the equilibrium market price $\ln P^f$ can be written as a function of its local and national determinants,

$$\ln P^f = f(\ln CS, \ln r, \ln IP, \ln HS, \ln E, \ln UE) . \quad (3)$$

Although mortgage market conditions, such as subprime lending intensity, are also likely to affect prices, the data are not available at the required frequency. By not modeling these conditions explicitly through observable variables as part of Equation (3), their influence on price is absorbed by the stochastic trend component that we employ to capture variations in the unobserved drivers of housing prices. At the second stage of the estimation process, which is described in sub-section 2.2, we capture the influence of mortgage market similarities across MSAs via spatial weight matrices, which are far less demanding in terms of data availability.

One could argue that there is likely some spatial correlation among the first-stage local fundamentals across MSAs. We account for that by including in the regression component

three national series, industrial production, construction activity of one-family houses, and the effective interest rate. In analogy to the Frisch-Waugh-Lovell theorem we argue that the inclusion of these variables removes the impact of national trends and most of the spatial correlation related to these trends from the variables. Since we include national trend variables not only for the economy in general (IP and UE) but also one for the one-family housing sector, we capture interdependence at both the economy-wide and the sectoral level.

We decompose the observed housing price, $\ln P_{i,t}$, for each MSA i according to Equation (4.1) into (a) the fundamental house price equation ($\ln P_{i,t}^f$), as determined by Equation (3), (b) an unobservable stochastic trend component, as given by $B_{i,t}$ in Equation (4.1) and (c) an irregular or white noise component, ζ , which is assumed to be *i.i.d.* normally distributed,¹⁰

$$\ln P_{i,t} = \ln P_{i,t}^f + B_{i,t} + \zeta_{i,t}, \quad (4.1)$$

$$\begin{aligned} B_{i,t} &= B_{i,t-1} + \mu_{i,t-1} \\ \mu_{i,t} &= \mu_{i,t-1} + \xi_{i,t} \end{aligned} \quad (4.2)$$

with $\ln P_{i,t}^f = \gamma_i^{Con} \ln CS_{i,t-1} + \gamma_i^r \ln r_{i,t-1} + \gamma_i^{IP} \ln IP_{i,t-1} + \gamma_i^{Starts} \ln HS_{i,t-1} + \gamma_i^{Empl} \ln E_{i,t-1} + \gamma_i^{Ump} \ln UE_{i,t-1}$.

Equation (4.2) specifies the model's unobserved trend component B as a standard smooth stochastic trend. In particular, B_t follows a non-stochastic random walk with drift (μ), where the drift is specified as a standard random walk.¹¹ The shock to the drift (ξ_t) is an *i.i.d.* distributed error term that is assumed independent of ζ_t . We note that the stochastic trend is

¹⁰ Note that it is not feasible in our model with only 14 years of data to separately identify a cyclical component.

¹¹ We note that a smooth stochastic trend also underlies the well-known Hodrick-Prescott (HP) filter. Our model differs from the HP filter in that we do not restrict the variance of ξ . See Commandeur and Koopman (2007) for a discussion of this class of unobserved component model and its advantages over other modeling strategies. A related discussion can also be found in chapters 2 and 3 of Kim and Nelson (1999).

far more flexible than a typical deterministic trend; it can capture significant growth momentum, both in the upward and downward direction and adapts quickly to change. Yet it is highly parsimonious from a statistical point of view as it requires estimation of only one parameter, the variance of ξ .

Importantly, our unobserved stochastic component model can account for the fact that housing markets typically display search frictions and do not clear in the short term (see Wheaton, 1990). In particular, subject to the restriction of a common variance for ξ_t , all short-term disequilibria are captured through our unobserved stochastic component. The cumulative sum of these short-term deviations represents the stochastic trend B , which can evolve into a distinctive price run-up in case of several positive deviations in close succession.¹²

The model consisting of Equations (4.1) and (4.2) requires simultaneous estimation of the parameter vector that relates the market fundamentals of Equation (3) to the observed price and the model's two variance parameters, the one relating to ς and the other relating to ξ .¹³ The above model is estimated separately for each MSA. The predicted stochastic trend series B is retrieved for use at the second stage of the model estimation.

2.2 Second Stage: Explaining Spatial Linkages of Stochastic Price Trends

The purpose of the second estimation stage is to identify variables that can capture the co-

¹² In contrast, if the short-term fluctuations in the unobserved component compensate each other, there will be no stochastic trend and the market remains close to equilibrium.

¹³ Note that the first stage estimates are not based on cointegration principles, although there is some philosophical similarity in that the parameters relating to the market fundamentals can be interpreted as “long-run” coefficients similar to the coefficients in a cointegrating equation. The logic of this interpretation results from the inclusion of a stochastic trend component; it ensures that the coefficients of the regression component can be interpreted—in analogy to the Frisch-Waugh-Lovell theorem for a linear regression—as the parameters of a “de-trended” relationship. In contrast to cointegration analysis, however, we are not interested per se in the “long-run” coefficients of the regression component at the first stage. That means, we also do not worry about collinearity. What matters is that the regression variables *jointly* capture the long-run influence of the market fundamentals.

movement or contemporaneous interdependence across MSAs in the stochastic trend series identified for each MSA at the first stage. This interdependence is represented in our panel data model of Equation (5) by the spatial lag variable ($\mathbf{W}_N \Delta B_{N,t}$),

$$\Delta B_{N,t} = \rho \mathbf{W}_N \Delta B_{N,t} + \beta \mathbf{Z}_t + u_N + e_{N,t}, \quad (5)$$

where $\Delta B_{N,t}$ is a $N \times 1$ vector of the differenced stochastic price trends ($B_{N,t}$) from the first stage,¹⁴ with $t = 1, \dots, T$ and subscripts N and T representing the number of MSAs and periods, respectively. \mathbf{Z}_t is a vector of control variables, including the log values of the *Federal Funds Rate*, the *30yr Fixed Mortgage Rate*, the *University of Michigan Consumer Sentiment Index*, and the continuously compounded return series of the *S&P 500 stock index*. We use fixed effects (u_N) to control for MSA-specific unobserved heterogeneity, such as size or location.

\mathbf{W}_N in Equation (5) is a $N \times N$ non-stochastic spatial dependence matrix with zeros on the main diagonal and non-negative off-diagonal elements. Consistent with commonly used geographic weight matrices, the weight matrix for our variables is defined in terms of the inverse of the distance. The distance is calculated as the absolute difference between the values of a variable X in MSAs j and k ,

$$w_{j,k}^X = \begin{cases} \frac{1}{|X_j - X_k|} & \text{if } j \neq k \\ 0 & \text{if } j = k \end{cases}. \quad (6)$$

We consider a number of alternative variables for X that include proxies for housing supply and demand. The guiding principle in selecting variables for X is to capture the likely incentives of direct and indirect investors in the housing market in choosing among MSAs. The key idea is that investors are interested in markets with a high probability of fast

¹⁴ The stochastic price trends are differenced to ensure that the dependent variable is stationary when the model parameters are estimated by our spatial panel data model.

appreciating home values. This may be the result of strong demand or of apparent supply constraints. Proxies for either one are candidates for X and will be discussed in section 3.

Two aspects are noteworthy of Equation (5). First, we estimate the spatial panel data model with annual data rather than with the monthly data we employ for the first stage. We convert from monthly to annual data by simple averaging. The reason for the change to annual data is that our focus is on the economics of the linkages among the MSAs, not on issues of timing. Yet, timing issues tend to become dominant at the monthly frequency. Based on preliminary estimates with monthly data we feel that aggregation over time makes the identification of the linkages significantly more robust, that is, effectively independent of arbitrary specification choices related to the timing of linkages.¹⁵ Aggregating up to annual data is feasible at the second stage because there is no shortage of degrees of freedom in the panel data set. Second, the dependent variable in Equation (5) is generated at the first estimation stage; it is the first difference of the stochastic trend B . If a generated variable appears as an independent or right-hand side variable in an OLS regression, we would encounter a measurement error problem, which would require either an instrumental variables approach or bootstrapping. Fortunately, our generated variable appears on the left-hand side of the equation and the measurement error is, therefore, captured by the equation's error term.¹⁶

¹⁵ Timing issues, such as which MSAs react first or what differences exist in the timing of MSAs and the reasons behind them, are sufficiently complex in their own right that they may be best approached in a separate study. In this context it should also be noted that spatial panel data models are computationally significantly more demanding than ordinary spatial models with cross-section data; this applies in particular if there are many observations per cross-section unit. One way to reduce the dimensionality of the estimation problem is to move away from the fully populated weight matrices that we employ in this study to sparse matrices. This can be done by setting matrix entries below a particular value equal to zero.

¹⁶ The spatial lag makes it look as if the dependent variable is in fact also on the right-hand side. However, that is not true given that we are estimating Equation (5) not by OLS but by maximum likelihood, with the spatial lag factored and appearing on the left side of the equation.

3 Data and Decomposition of Observed Price Series

3.1 Decomposing the Observed Housing Price

Our analysis comprises the 20 MSAs that make up the well-known 20-city composite Case-Shiller housing price index. The Case-Shiller series we employ at the first stage are seasonally adjusted and at a monthly frequency.¹⁷ We realize that the MSA price indices are averages over a potentially rather diverse set of local communities inside an MSA and their evolution over time. But as our focus is on inter-MSA price linkages and to keep the model computationally tractable, we abstract from the issue of intra-MSA price differences. The sample period for our first-stage analysis covers the months from 1995:01 to 2008:12. The choice of our sample period is based on the consideration to cover a sufficient number of periods on either end of what is often identified as the housing bubble. As we rely on a stochastic trend model at the first stage, the length of the time series is sufficiently long given that a stochastic trend adapts effectively immediately to changes in the underlying variables.

In our first-stage model, we employ both MSA-specific and national variables to estimate the fundamental housing price trend at the MSA level. Inclusion of national variables ensures that we capture common trends in economic activity among MSAs, although we analyze each MSA individually. At the MSA level, we measure the *employment level* (E), the *unemployment rate* (UE), and *housing starts* (HS). Employment is taken to be a

¹⁷ The Case-Shiller (CS) index is a monthly index for the home prices in 20 U.S. MSAs. Similar to the OFHEO home price index, published by Federal Housing Finance Agency (FHFA), the CS index is derived using the repeat sales valuation approach. Compared to the FHFA index, the CS index has broader market coverage and applies a value-weighted method based on (repeat) transactions. In contrast, to increase the sample size OFHEO additionally includes appraisal data. The CS index uses a robust interval-value-weighted repeat sales procedure and thus reduces biases stemming from pricing anomalies, physical changes, local neighborhood effects, high turnover frequency, and time between transactions (see Miao, Ramchander, and Simpson, 2011; Kallberg, Liu, and Pasquariello, 2011). We note that the S&P/Case-Shiller Home Price Indices are calculated using three-month moving averages. This has the potential to induce some spurious autoregression in the data generating process (DGP).

proxy for both the level of population and economic activity as no monthly data on GDP or population are available. The variables *industrial production (IP)*, private residential *construction spending (CS)* and the *effective interest rate (r)* are measured at the national level. Most of our first-stage variables are commonly used to identify long-run housing values (see Abraham and Hendershott, 1993; Capozza, Hendershott and Mack, 2004; Case and Shiller, 2003; Goodman and Thibodeau, 2008) and, therefore, require little explanation. To eliminate potential problems of endogeneity, we include only lagged variables on the right side of Equation (4). Collinearity among the variables is of no concern at the first stage of our estimation process since we are not interested in the precision or efficiency of the individual variable coefficients, but only in their joint ability to predict housing prices. Price movements that are not picked up by our proxies for market fundamentals either enter the smooth stochastic trend, which we take to represent persistent price deviations from market fundamentals, or the disturbance term of the model, which captures idiosyncratic shocks that cannot be modeled.

Figure 1 illustrates for one MSA with a significant price appreciation (San Diego) and for one with only a moderate price increase (Denver) how the observed price index, the price index predicted by market fundamentals, and the unobserved price component (B in 4.2) evolve over time. Both the observed price index and the unobserved price component use the same vertical scale for both MSAs to reveal the differences in their behavior over time. The predictions on the basis of market fundamentals are scaled differently between San Diego and Denver to highlight that the predictions from the market fundamentals are similar over time between the two MSAs, although the amplitude of the predictions is less for Denver than for San Diego.¹⁸

¹⁸ We note that the product of the graphs in the bottom panel of Figure 1 and the center panel forms the prediction of the observed price index. This prediction deviates from the observed price index in the top panel of Figure 1 only by an idiosyncratic error term, which is not shown in the figure.

The price indices for San Diego and Denver share a similar behavior up until 2001. Thereafter, the price index for San Diego starts to deviate strongly from the predictions of the fundamental variables; a strong price run-up results.

<< Figure 1 about here >>

In general, we find that significant price run-ups of the type apparent for San Diego appear more likely in coastal MSAs, which typically have stronger income increases, lower unemployment rates, and more inelastic housing supplies than MSAs in the interior of the country. Figure 2 illustrates this point. The size of the circles in Figure 2 identify the deviations from market fundamentals based on our first-stage estimates. We calculate these deviations by subtracting the smallest from the largest value of the stochastic trend (B) over the estimation period 1995-2008. Since B is measured in logs, we derive an approximate percentage deviation from market fundamentals.¹⁹ It is apparent from Figure 2 that these deviations are disproportionately concentrated among MSAs along the east and west coast, while some MSAs in the interior do not have measurable deviations from market fundamentals; Dallas is a case in point. These findings are consistent with the majority of studies on bubbles in housing markets (see, e.g., Abraham and Hendershott, 1993, 1996; DiPasquale and Wheaton, 1994; Case and Shiller, 2003; Capozza, Hendershott and Mack, 2004; Goodman and Thibodeau, 2008; Wheaton and Nechayev, 2008).

<< Figure 2 about here >>

3.2 Data for the Second Estimation Stage

At the second stage of our estimation process the smooth stochastic trend from the first stage of our estimation process becomes the dependent variable, although in first difference form to assure stationarity. The key aspect of the second stage is the specification of the

¹⁹ In practice, we take the exponent of the log difference of B and subtract one.

spatial lag.

The weight matrix of the spatial lag connects the stochastic trend of a given MSA with the stochastic trends of all other MSAs. In standard applications, the weight matrix consists of distances in miles. We reject the idea that distance is a sensible weighting matrix for our data because price run-ups can be observed in locations far from each other, such as San Diego and Miami, but not necessarily in locations closer to each other, such as San Diego and Dallas. Although distance may not be sensible, we do not reject the notion that regional association may be relevant. It is well known that investors do have regional preferences and investment and management companies or mortgage originators, which translate the preferences of investors into purchase and selling activity, also have regional emphases for their operations. Some form of regional clustering is also apparent from Figure 2. However, before we consider regional clustering via contiguity matrices, we focus on observable economic variables in line with e.g. Case, Rosen, and Hines (1993) and Fingleton (2001, 2008) and define the elements of our spatial weight matrices in terms of economic distance or similarity. In particular, we try to capture the tendency of an area to experience unusual house price appreciations by using both demand-side variables, such as population growth as an indicator of revealed preference for an area, and supply side variables, such as physical or administrative constraints on urban sprawl.

3.2.1 Demand Factors

Population. The attraction of new residents is an important driver of housing demand. Mulder (2006) describes the complex relationship between *population* and housing. Undeniably, *population growth* via migration leads to an increase in housing demand, and because of a greater demand for housing, home and land prices increase.²⁰ In addition, if land

²⁰ See, e.g., Saiz (2003, 2007) who demonstrates that prices and rents in housing markets which are characterized by immigrant population shocks undergo price appreciations, which then have an impact on labor mobility of current residents (Ottoviano and Peri, 2007).

is scarce, households demand smaller housing units, building heights increase, and *population density* is higher. In contrast, a higher supply of housing attracts more people to immigrate or form new households and therefore, induces population to grow. It is also obvious that moving for household or retirement reasons is more closely related to higher quality housing than moving for education or work (Mulder, 2006). More generally, a growing population coupled with growing income leads to higher housing demand. However, rising income levels also increase the willingness to pay for high-amenity or low-density neighborhoods (see Glaeser, Gyourko, and Saks, 2005).

Socio-economic Indicators. Apart from neighborhood variables, Freeman (1979) highlights the importance of *socio-economic variables* as determinants of property values. The levels of income in addition to house prices and interest rates are the key indicators for housing affordability. Normally, the *median income family* qualifies for the median value home (Gyourko and Linneman, 1993). However, since the beginning of the 21st century monetary policies of lower long-term interest rates and mortgage subsidy programs have improved housing affordability substantially. Therefore, relaxed financial regulation and credit standards in combination with aggressive lending practices make it possible for larger numbers of people to enter the market for real estate (Dell’Ariccia, Igan, and Laeven, 2009; Pavlov and Wachter, 2010). In particular, in the subprime mortgage market with a dominant proportion of low income households and a high number of unemployed household members the amount of owner-occupied homes increased. Hence, in markets where these mortgage market conditions prevail and *subprime lending intensity* is high, we can expect price appreciations to develop more easily.

Gyourko and Linneman (1993) also point to the affordability problem for less well educated and lower income households, i.e. there is little supply of lower quality housing that is affordable to low-skilled households. Paired with strict building codes, approval delays,

stringent low density zoning, and impact fees make lower quality housing economically less profitable to developers. It follows that owning single-family homes is inevitably associated with *educational attainment*, i.e. having more than a high school education.²¹ Moreover, socioeconomic characteristics are also relevant in the context of neighborhood effects. Owner-occupied residents value neighborhood education and income levels because of positive externalities of neighborhoods such as superior school quality, a lower property crime rate, and positive environmental characteristics. High income households are willing to pay a premium for these effects in terms of a higher property price, which in turn prevents further unabated in-migration into these areas (Gibbons, 2001). The educational level and job status also reflect the technology-based productivity change which has favored skilled residents' cities (Moretti, 2003; Glaeser and Saiz, 2004).

Amenities. In a similar vein, increasing income makes individuals value local amenities more highly. Hence, households are willing to pay for amenities, such as mild seasons, sunshine, hills, coastal proximity, safety, clean air, arts and culture. Albouy (2011) finds these amenities are key determinants of *quality of life (QOL)* and are even more sought after in growing cities than previously thought. Many of the states and MSAs that saw dramatic increases in home prices in the run-up to the housing crash of 2006 rank highly on just these quality of life criteria. This suggests that a more formal test of the quality of life link between prices in different areas is warranted. Interestingly, neither population size nor population density appears to affect a city's *QOL* negatively, but most differences in quality of life can be explained by natural amenities, such as coastal proximity, which in turn is often correlated

²¹ Gyourko and Tracy (1999) find that 55% of homeowners in the U.S. have at least some college education with 30% holding a college degree. In a similar vein, high school graduates with more than five years of work experience own homes at approximately 90% of the rate in 1974. According to the 2011 American Housing Survey, the percentage of homeowners holding a high school degree or higher (bachelor degree or higher) equals 86.8% (30.9%). However, it can be assumed that this numbers are highly diluted by relaxed mortgage lending practices during subprime market boom.

with a limited housing supply due to scarce land and zoning restrictions.

The *QOL* weights are derived from the adjusted quality of life indices calculated by Albouy (2011).²² The intuition behind calculating a proximity measure between two MSAs on the basis of their closeness in terms of a quality of life index is as follows. As previously mentioned, housing prices in an MSA are likely to be affected by housing prices in other MSAs because investors looking for a high return are pushing up demand first in areas with an almost certain potential for value appreciation. Those areas are very likely to be those that are attractive to final home buyers. The Albouy's quality of life index consists of natural and artificial amenities. The former includes climate and geographic characteristics such as heating and cooling degree days per year, sunshine, coastal proximity, and the average slope of the land. The latter are determined by local inhabitants such as restaurants and bars per capita, the arts and culture index from Places Rated Almanac, air quality index, and safety (violent crime and property crime per capita).²³ Note that Albouy's quality of life index cannot be classified as pure demand-side proxy since it also includes geographic constraints such as coastal proximity and the steepness of land, i.e. non-buildable land.²⁴

3.2.2 Supply Factors

Regulatory Environment and Undevelopable Area. Glaeser, Gyourko, and Saks (2005) emphasize that housing supply constraints arise because of changes in regulatory regimes rather than the lack of developable land. Naturally, new construction increases the supply of

²² Unlike many previous studies of quality of life across the United States, the calculations of Albouy (2011) also consider cost-of-living expenses that are not related to housing costs, better relate city wage levels to a household's buying power and also net out federal taxes. As a consequence, generally favored MSAs, such as Honolulu or San Francisco do not end up at the bottom of the quality of life index, but are ranked according to typical expectations.

²³ See Appendix B.4 of Albouy (2011) for a detailed description of the amenity data.

²⁴ However, in the study of Albouy (2011) a higher inverse distance to the coast and a higher average slope of land are interpreted as positive valuations of amenities and thus reflect demand-side variables. In his amenity-value estimates Albouy shows that households are willing to pay 1.7 and 2.7% of income to live in close distance to the coast and in areas where the average slope is 10% higher, respectively.

housing and both home values and rents will decrease. However, because of the negative externality of higher population density and the expectation of declining housing values associated with new developments, current residents try to restrict zoning via organized community groups (Glaeser, Gyourko, and Saks, 2005). Moreover, as mentioned above, high income households are willing to pay for high-amenity and, in particular, low-density neighborhoods.

Inelastic supply due to scarcity of developable land can be explained by geographic factors, such as the proximity to the ocean, a lake and a river, steep topography, as well as wetlands. Regulatory barriers to development are related to housing zoning which explicitly limits the availability of land and other land-related regulatory procedures and building restrictions, such as the political process of approval, which causes significant costs through delays in the development of new projects.

Supply Elasticity. According to Saiz (2010), it is thus obvious that housing supply rather than demand shocks may account for most differences in the pricing of home values across cities. He also emphasizes that the value of the housing supply elasticity is well captured by both physical and regulatory constraints. Glaeser, Gyourko, and Saiz (2008) demonstrate the pivotal role of housing supply in shaping the course of housing bubbles. In particular, under the assumption of irrational exuberance, supply inelastic MSAs have larger price increases along with a smaller impact on the housing stock and longer lasting bubbles. In contrast, U.S. cities with more elastic housing supplies have fewer and shorter bubbles, but tend to overbuild in response to bubbles. Hence, in case of inelastic supply an endogenous bubble acts as a short-term demand shock, where rising demand translates into rising prices. Glaeser, Gyourko, and Saiz (2008) find that average estimated real prices appreciate by 81% in relatively inelastic MSAs compared to 34% in relatively elastic ones during the boom period 1997 to 2006.

As proxies for housing supply we use the measures of supply side conditions developed by Saiz (2010) and Gyourko, Saiz, and Summers (2008): land scarcity, land regulation, and supply elasticity. Saiz (2010) estimated first *undevelopable land (UDA)* due to adjacency to the ocean and great lakes in a radius of 50 kilometers around the geographic centroid of each metropolitan area by utilizing GIS techniques. He then used satellite-based geographic data on land provided by the United States Geographic Service (USGS) to account for area lost to minor water bodies, wetlands, permanent ice caps, and bare-rock desert areas. Finally, he derived slope maps for rings around the centroid of each city from USGS Digital Elevation Model at 90 square meter parcel of land to determine irreclaimable land with slopes above 15 degrees. Note that this supply measure is exogenous from market conditions since it is solely based on natural land constraints.

In addition, we use the Wharton Residential Land Use Regulatory Index (WRLURI) developed by Gyourko, Saiz, and Summers (2008) as a further supply-side proxy. The measure is based on a 2005 survey of over 2000 localities across the U.S. and includes a number of dimensions of residential land regulations, such as zoning, permit approval, state and judicial activism, and other aspects of local regulatory environment. A factor analysis is applied to the different dimensions to derive a composite index with higher standardized values indicating a more restrictive regulatory environment. In contrast to the physical constraint measure, this regulation-based index is affected by the fact that zoning and land use policies are endogenous (see McMillan and McDonald, 1991; Pogodzinski and Sass, 1994). There is a reverse causality from prices to the regulatory process, i.e. residents and politicians choose the level of regulatory development constraints based on actual and future housing prices.

Our final supply proxy is again derived by Saiz (2010) based on a function of both physical and regulatory constraints. Because of the severe endogeneity of regulation, Saiz

estimated a simultaneous equation system which provides local supply elasticities by jointly determining housing supply, demand, and regulations. Hence, market clearing prices and quantities in final equilibrium reflect the final regulation level (Saiz, 2010).

The vacancy rate reflects the balance between supply and demand of housing, i.e. excess supply and excess demand defined as the deviation of actual from natural home or rental vacancy rate determines the change in median house or rental value. Accordingly, the level of vacancies is given by the difference between supply and demand of owner-occupied or rental housing units. Furthermore, the *price-to-rent ratio* equals the average user cost of owner occupied housing divided by the rent being paid for the same dwelling size. Because rents are strongly tied to supply and demand fundamentals, an increase in the price-to-rent ratio indicates deviations of house prices from fundamental values. In contrast, if the user cost of owner-occupied housing falls below the rental price of housing services, households choose to purchase rather than renting their home.

Table 1 provides a detailed description of our variables.

<< Table 1 about here >>

4 Empirical Results

4.1 One Weight Matrix Based on Individual Variables

Table 2 shows the estimation results from the spatial lag panel data model (SLPDM) of Equation (5), with the stochastic trend component in first differences as the dependent variable. The co-movements of the dependent variables, i.e. the linkages in price dynamics unexplained by local fundamentals, are captured by a spatial lag based on alternative demand and supply measures as described in Table 1. To control for remaining influences from national fundamentals, we include log values of the Fed Funds Rate, the 30 year. Fixed Mortgage Rate, the University of Michigan Consumer Sentiment Index, and the continuously

compounded S&P 500 return. We further account for unobserved heterogeneity by using MSA fixed effects.

Most of our demand and supply variables are significant with the exception of family income per capita and indicators for unemployment and educational attainment. The highest log-likelihood values are obtained for the demand variables *population growth* and *median rental value*, as well as for *supply elasticity*, with corresponding ρ coefficients of 0.55, 0.54, and 0.43, respectively. The regression results indicate that population growth prior to the beginning of the substantial price run-ups after 2000 and inelastic supply due to natural and regulatory restrictions tie together MSA price run-ups.

<< Table 2 about here >>

4.2 One Weight Matrix Based on Aggregated Variables

To test the joint impact of demand and supply variables in the SLPDM, we want to work with spatial weight matrices only of variables that are not highly correlated. To test which of our weight matrix variables are correlated, we calculate Spearman's rank correlation coefficients for all 2-variable combinations. Based on these results we aggregate those variables that are highly interdependent. By aggregating variables rather than simply eliminating variables we retain potentially useful information yet avoid highly correlated weight matrices. Some variables are aggregated by way of a Mahalanobis distance measure; others are combined by simply forming a ratio.²⁵

From Table 3 it is apparent that the unemployment rates and the number of families below poverty rate are significantly correlated. We create a new aggregated variable *Socially Deprived* by using the Mahalanobis distance measure. In contrast, population growth is only

²⁵ The Mahalanobis distance measure identifies the distance between any two points in k -dimensional space. It is scaled by the covariance matrix of the k variables. If the covariance matrix is the identity matrix, the Mahalanobis distance measure reduces to the Euclidean distance measure.

related to supply elasticity. People prefer to live in regions where the amenity value is high. However, these areas are mainly characterized by scarce land or undevelopable areas. Interestingly, population growth and population density are negatively correlated, which confirms the previous remarks that high-amenity regions attract immigrants, however, high-income households are willing to pay for low-density neighborhoods. We thus aggregate population level and population density to the demand variable *Population*, while we keep population growth as an individual demand-side indicator.

Household income (family income per capita and median family income) is inevitably highly correlated with educational attainment (percent with bachelor or graduate degree, percent with graduate degree) as well as the share of employment in skill intensive service sector industries and managerial occupations. We call the new variable which combines these socio-economic determinants *Social Status*. We further keep the uncorrelated subprime lending intensity as a separate variable (*Subprime Lending*).

Other than median house and rental values, we find the quality of life index to be correlated with the share of skill intensive service sector employment. This is likely driven by the salaries or incomes of those employed in skill intensive jobs. Not surprisingly, our amenity indicator is also highly correlated with our supply factors *undevelopable area* and *regulatory environment*, because the former variable is also considered in the construction of the QOL index. However, because of the multi-dimensionality of QOL and its interpretation as a demand variable, we decide to leave the *Quality of Life* index as an individual variable in our model specifications.

<< Table 3 about here >>

Even though our three supply variables are uncorrelated, we decide to aggregate undevelopable land and the regulatory index WRLURI in a similar vein as Saiz (2010) did for the variable *Supply Elasticity*. However, in comparison to the estimated supply elasticity, the

combined indicator *Scarce Land* is unadjusted for endogeneity, i.e. the reverse channel that arises via demand.

Most of the socio-economic variables are correlated with median house and rental values, with the latter two being almost perfectly correlated. To combine both price variables we construct the *Price-to-Rent Ratio* which is commonly accepted as a measure to detect bubbles in the housing market. Finally, we aggregate home and rental vacancy rates (*Vacancy Rates*) to reflect the balance of demand and supply.

Table 4 provides a summary of the estimation results with just one of the aggregate weight matrices. In addition to the regression results for our variable aggregates the table shows for the purpose of comparison the regression results for three variables that proved to be particularly promising as weight matrix variables in Table 2. These are *population growth*, *subprime lending intensity*, and *supply elasticity*.

<< Table 4 about here >>

We notice that the demand-side variable *population growth* has more explanatory power than any of the regressions with spatial weight matrices based on combined variables. Next in explanatory power are the two supply-side variables, *scarce land* and *supply elasticity*. Also of interest are the weight matrices based on the *price-to-rent ratio* and the variable combining population level and density.

Table 4 also reports the coefficient estimates of the control variables. The coefficients have typically the expected sign and do not vary much across specifications, although their statistical significance is somewhat sensitive to the chosen spatial weight matrix.

4.3 Two Weight Matrices

The joint estimation of more than one weight matrix allows us to incorporate more than one dimension in linking MSAs. In particular, we can consider demand-side and supply-side influences simultaneously. The estimation proceeds with the standard spatial lag model of the

last section, again estimated by maximum likelihood, but with the difference that we now have two spatial lags with two separate coefficients (ρ_1 and ρ_2),

$$\Delta B_{N,t} = \rho_1 \mathbf{W}_{1,N} \Delta B_{1,N,t} + \rho_2 \mathbf{W}_{2,N} \Delta B_{2,N,t} + \beta \mathbf{Z}_t + u_N + e_{N,t} . \quad (7)$$

<< Table 5 about here >>

The results are summarized in Table 5. The models are ordered by their log likelihood value.²⁶ The two models with the best fit both contain the demand-side variable *population growth* and one of the two supply-side variables, *scarce land* or *supply elasticity*. In fact, all models of Table 5, except for one, contain either the variable *population growth* or one of the two supply variables. This suggests that these variables are important determinants of the linkages among the price run-ups of the 20 MSAs we consider.

Table 6 translates the estimates of Model Ib of Table 5 into individual impact coefficients for the 11 MSAs with the largest price run-ups. The second row, for example, shows that the stochastic trend of house price inflation calculated for San Diego is connected with the equivalent house price inflation in Los Angeles with a coefficient of 0.393, and with house price inflations in San Francisco, Boston and New York with coefficients of 0.203, 0.098, and 0.075, respectively. This suggests that the influence from the neighboring MSA in Los Angeles is about twice as large as that from San Francisco, the MSA that is in the same region but somewhat further away. The influence from across the continent, from Boston or New York, is again about half the size that is exercised by San Francisco.

<< Table 6 about here >>

The connections that are reported numerically in Table 6 are illustrated graphically in the first panel of Figure 3. Thicker connections identify larger coefficients. It is apparent from the graph that the strongest connections are not necessarily among MSAs in the same

²⁶ A complete set of results, with all combinations of weight matrices reported individually in Table 4, is provided in Table A1.

regional cluster. For example, Washington DC has the strongest connection with Phoenix, Miami with Los Angeles, and Seattle with Miami. Strong connections exist between MSAs in the Northwest and in Florida and between MSAs in the Northeast and in the Southwest. The pattern visible from the first panel of Figure 3 is repeated in the subsequent panels, which use different variables for the weight matrices, although there is some variation in the strengths of the various connections across panels.

4.4 Three Weight Matrices

From the results with two weight matrices it is apparent that there are regional connections among MSAs that are not fully captured by the variables entering the weight matrices. To quantify the regional impact we combine the best fitting weight combination of section 4.3 with alternative contiguity matrices.²⁷ A contiguity matrix is a weight matrix consisting of only zeros and ones, where all MSAs in the same regional cluster are connected with a one. Different contiguity matrices identify different ways to allocate MSAs to regional clusters. A summary of our estimates based on various contiguity matrices, representing alternative regional clusters, is given in Table 7.

<< Table 7 about here >>

We find that the regional cluster identified as *contiguity 1* adds by far the most information to the two matrices based on *population growth* and *supply elasticity*, with a log likelihood value of 1117.84. *Contiguity 1* identifies the regional clustering that is behind the color pattern of Figure 2. *Contiguity 2* and *3* are minor variations of this clustering. It is noticeable from Table 7 that the rho values of the spatial lags for *population growth* and *supply elasticity* decline perceptively when the spatial lag based on the regional cluster

²⁷ We find that all efforts to add a third weight matrix on the basis of economic variables tested separately or jointly to be unsuccessful. For example, the thought that subprime lending or quality of life would add a new dimension to be rejected by the data. The coefficient values of both are zero. This applies to the other variables to varying degrees as well.

identified by *contiguity 1, 2, or 3* is added. The decline is particularly pronounced for the weight matrix based on *population growth*, less so for the weight matrix based on *supply elasticity*. The coefficients of the weight matrices for *population growth* and *supply elasticity* are much less reduced in size and remain statistically significant for insignificant contiguity matrices, for example the two matrices that identify regional clusters on the basis whether an MSA is located on the ocean or not (*coast 1* and *coast 2*).

We conclude from these results that our two preferred weight matrices from the last section, *population growth* and *supply elasticity*, do not capture all of the relevant linkages among the price run-ups of MSAs. There is more that connects price run-ups in different locations. Regional preferences of investors do seem to play a role. How to translate these regional preferences into economically measurable variables is a question for further research. We are facing here an issue that is common in much of regional, urban, and real estate research: there is spatial autocorrelation, but it is difficult to find variables that make the underlying linkages explicit.

Figure 4 illustrates the linkages identified by Model I of Table 7. The methodology is the same as that for Figure 3, only that we now utilize three weight matrices and their estimated coefficients rather than only two. The strongest links occur now within the regional clusters. However, the general pattern and links visible in Figure 3 are also noticeable in Figure 4. In other words, even if we account for regional linkages there remains room for demand and supply linkages that are active beyond the region.

5 Conclusion

In this paper, we aim to answer the question whether local house price inflations beyond the level that can be explained by market fundamentals are similar across U.S. metropolitan real estate markets and what the nature of possible linkages are. For our analysis

we make use of the Case-Shiller home price index for 20 U.S. metropolitan statistical areas (MSAs) for the period 1995:01 to 2008:12.

We employ a two-stage approach for our empirical analysis. At the first stage, we decompose for each MSA separately the observable housing price index into its fundamental value, which is determined by standard housing demand and supply factors, and a stochastic trend that represents persistent deviations of the housing price index from its fundamental value. At the second stage, we combine the stochastic trends of all 20 MSAs in a spatial panel and estimate numerous spatial lag panel data models with one, two, and three spatial weight matrices. We define the weight matrices in terms of economic variables representing demand and supply influences on house price inflation. In the models with three weight matrices we add a spatial lag on the basis of a contiguity matrix that captures regional influences that are difficult to identify with economic variables.

In line with the existing literature, we find that price run-ups untied from market fundamentals vary significantly across the 20 MSAs that are monitored by the Case-Shiller housing price index. Many of these price run-ups are similar between neighboring MSAs, such as between San Francisco, Los Angeles, San Diego, Phoenix, and Las Vegas or between Boston, New York, and Washington DC. However, we also find that there are significant linkages across regions, such as between the MSAs in the Northeast and the MSAs in the Southwest or between the MSAs in the Northwest and those in Florida.

We find that the price dynamics among U.S. MSAs are tied together by a combination of demand-side indicators, such as *population growth* and to a lesser extent the aggregated *population* measure, *price-to-rent ratio*, and *social status*, and supply-side indicators, such as *scarce land* and *inelastic supply*. These links in price dynamics are complemented by regional linkages that are difficult to explain by economic variables.

Summing up, we demonstrate that the run-ups in home prices in numerous regions and

MSAs prior to 2006 were more than just coincidental and far from being only a local or regional phenomenon. Because simultaneous price inflations do have some common characteristics, either economic or geographic, policy makers should potentially consider regional run-ups in housing prices as being of national concern.

Appendix A:

Table A.1: Alternative Combinations of Spatial Lag Panel Data Models including Two Weight Matrices

Weight Matrix 1	Weight Matrix 2	LLF	Weight Matrix 1			Weight Matrix 2		
			ρ_1	S.E.	P-value	ρ_2	S.E.	P-value
Population Growth	Supply Elasticity	1085.757	0.408	0.102	0.000	0.277	0.100	0.006
Population Growth	Scarce Land	1084.230	0.409	0.145	0.005	0.313	0.159	0.050
Social Status	Population Growth	1081.548	0.229	0.133	0.086	0.460	0.104	0.000
Population	Population Growth	1080.891	0.259	0.165	0.117	0.439	0.143	0.002
Scarce Land	Supply Elasticity	1080.080	0.354	0.120	0.003	0.305	0.107	0.004
Price-to-Rent Ratio	Population Growth	1078.884	0.222	0.153	0.146	0.433	0.141	0.002
Population Growth	Subprime Lending	1077.727	0.503	0.101	0.000	0.138	0.106	0.194
Population	Supply Elasticity	1076.963	0.307	0.141	0.030	0.328	0.125	0.009
Socially Deprived	Population Growth	1075.790	0.042	0.129	0.743	0.534	0.111	0.000
Social Status	Supply Elasticity	1075.705	0.244	0.115	0.033	0.347	0.083	0.000
Quality of Life	Population Growth	1075.702	0.020	0.099	0.842	0.546	0.104	0.000
Vacancy Rate	Population Growth	1075.680	-0.010	0.090	0.911	0.558	0.089	0.000
Supply Elasticity	Price-to-Rent Ratio	1074.585	0.348	0.109	0.001	0.231	0.082	0.005
Price-to-Rent Ratio	Scarce Land	1073.939	0.282	0.117	0.016	0.370	0.132	0.005
Price-to-Rent Ratio	Supply Elasticity	1073.482	0.250	0.109	0.021	0.311	0.109	0.005
Social Status	Price-to-Rent Ratio	1072.018	0.271	0.132	0.039	0.371	0.106	0.000
Subprime Lending	Supply Elasticity	1071.809	0.148	0.097	0.129	0.385	0.098	0.000
Population	Scarce Land	1071.567	0.237	0.152	0.120	0.378	0.161	0.019
Social Status	Scarce Land	1071.336	0.185	0.182	0.308	0.411	0.175	0.019
Socially Deprived	Supply Elasticity	1070.541	0.119	0.116	0.305	0.390	0.094	0.000
Quality of Life	Supply Elasticity	1070.237	0.089	0.070	0.200	0.404	0.086	0.000
Population	Price-to-Rent Ratio	1070.073	0.307	0.159	0.053	0.317	0.152	0.037
Population	Social Status	1069.824	0.358	0.154	0.020	0.256	0.171	0.134
Vacancy Rate	Supply Elasticity	1069.554	-0.005	0.073	0.946	0.434	0.074	0.000
Subprime Lending	Scarce Land	1069.336	0.086	0.090	0.340	0.485	0.120	0.000
Vacancy Rate	Scarce Land	1069.148	-0.079	0.072	0.273	0.580	0.106	0.000
Quality of Life	Scarce Land	1068.787	-0.041	0.069	0.558	0.561	0.115	0.000
Socially Deprived	Scarce Land	1068.681	0.005	0.150	0.972	0.532	0.152	0.000
Social Status	Subprime Lending	1066.091	0.345	0.131	0.008	0.219	0.098	0.026
Price-to-Rent Ratio	Subprime Lending	1065.281	0.428	0.114	0.000	0.107	0.124	0.392
Price-to-Rent Ratio	Vacancy Rate	1064.877	0.540	0.090	0.000	-0.088	0.081	0.277
Socially Deprived	Price-to-Rent Ratio	1064.717	0.083	0.139	0.549	0.442	0.110	0.000
Population	Subprime Lending	1064.601	0.440	0.107	0.000	0.094	0.110	0.391
Socially Deprived	Population	1064.439	0.099	0.167	0.554	0.442	0.155	0.004
Quality of Life	Price-to-Rent Ratio	1064.339	0.004	0.078	0.963	0.488	0.089	0.000
Quality of Life	Population	1064.136	-0.066	0.144	0.646	0.550	0.166	0.001
Population	Vacancy Rate	1064.004	0.521	0.112	0.000	-0.034	0.093	0.714
Socially Deprived	Social Status	1063.117	0.185	0.144	0.200	0.327	0.154	0.034
Quality of Life	Social Status	1061.936	0.117	0.075	0.121	0.355	0.131	0.007
Social Status	Vacancy Rate	1060.989	0.387	0.121	0.001	0.036	0.078	0.650
Socially Deprived	Subprime Lending	1058.161	0.263	0.136	0.052	0.213	0.115	0.064
Quality of Life	Subprime Lending	1055.511	0.162	0.089	0.070	0.234	0.123	0.058
Quality of Life	Socially Deprived	1055.244	0.133	0.089	0.132	0.287	0.141	0.042
Vacancy Rate	Subprime Lending	1054.386	0.089	0.089	0.318	0.272	0.120	0.024
Socially Deprived	Vacancy Rate	1054.251	0.351	0.137	0.010	0.030	0.093	0.748
Quality of Life	Vacancy Rate	1051.366	0.247	0.077	0.001	0.065	0.085	0.442

Notes: The dependent variable is the stochastic trend component in first differences according to Equation (4) for the 20 MSAs. The estimates are derived from Equation (7) based on annual data for the period 1995 to 2008. LLF is the value of the log-likelihood. ρ_1 and ρ_2 are the estimated coefficients of the spatial weight matrices.

References

- Abraham, J.M., and P. Hendershott (1993). "Patterns and Determinants of Metropolitan House Prices, 1977 to 1991, Real Estate and the Credit Crunch." In: Browne and Rosengren (Eds.), *Proceedings of the 25th Annual Boston Fed Conference*, 18–42.
- Abraham, J.M., and P. Hendershott (1996). "Bubbles in Metropolitan Housing Markets." *Journal of Housing Research* 7(2), 191–207.
- Adams, Z., and R. Füss (2010). "Macroeconomic Determinants of International Housing Markets." *Journal of Housing Economics* 19(1): 38–50.
- Akerlof, G.A., and R.J. Shiller (2009). "Animal Spirits: How Human Psychology Drives the Economy and why it Matters for Global Capitalism." Princeton, New Jersey, Princeton University Press 2009.
- Al-Anaswah, N., and B. Wilfling (2009). "Identification of Speculative Bubbles using State-Space Models with Markov-Switching." *Working Paper*, Center for Quantitative Economics, University of Münster.
- Albouy, D. (2011). "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." *NBER Working Paper* No. 14472, University of Michigan, Ann Arbor.
- Calomiris, C.W. (2010). "Financial Innovation, Regulation, and Reform." In *Globalization and Growth: Implications for a Post-Crisis World*, Commission on Growth and Development, The World Bank, 47–68.
- Capozza, D.R., P.H. Hendershott, C. Mack (2004). "An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets." *Real Estate Economics* 32(1): 1–32.
- Case, A.C., H.S. Rosen, and J.R. Hines (1993). "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics* 52(3): 285–307.
- Case, K.E., and R.J. Shiller (2003). "Is there a Bubble in the Housing Market? An Analysis." *Brookings Papers on Economic Activity*, 2: 299-362.
- Clapp, J.M., and D. Tirtiroglu (1994). "Positive Feedback Trading and Diffusion of Asset Price Changes: Evidence from Housing Transactions." *Journal of Economic Behavior and Organization* 24(3): 337–55.
- Commandeur, J.J.F., and S.J. Koopman (2007). *An Introduction to State Space Time Series Analysis*. Oxford and New York, Oxford University Press.
- Costello, G., P. Fraser, and N. Groenowold. (2011). "Housing Prices, Non-fundamental Components and Interstate Spillovers: The Australian Experience." *Journal of Banking and Finance*. 35(3):653-669.
- Dell’Ariccia, G., D. Igan, and L. Laeven (2009). "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market, European Banking Center Discussion Paper No. 2009-14S, April 2009.
- DiPasquale, D., and W.C. Wheaton, (1994). "Housing Market Dynamics and the Future of Housing Prices." *Journal of Urban Economics* 35(1): 1–27.
- DiPasquale, D., and W.C. Wheaton (1996). *Urban Economics and Real Estate Market*. New York: Prentice Hall, Englewood Cliffs.
- Durbin, J. and S.J. Koopman, (2001). *Time Series Analysis by State Space Methods*. Oxford and New York, Oxford University Press.
- Elwood, S.K., E. Ahmed, and J.B. Rosser, Jr. (1999). "State-Space Estimation of Rational Bubbles in the Yen/Deutsche Mark Exchange Rate." *Weltwirtschaftliches Archiv* 135(2): 317–31.

- Fingleton, B. (2001). "Equilibrium and Economic Growth Spatial Econometric Models and Simulations." *Journal of Regional Science* 41(1): 117–41.
- Fingleton, B. (2008). "A Generalized Method of Moments Estimator for a Spatial Model with Moving Average Errors, with Application to Real Estate Prices." *Empirical Economics* 34(1): 35–57.
- Freeman, M.A. (1979). "The Hedonic Price Approach to Measuring Demand for Neighbourhood Characteristics." In D. Segal (ed.), *The Economics of Neighborhood*, Academic Press: New York.
- Fry, J. M. (2009). "Bubbles and Contagion in English House Prices." MPRA Working Paper.
- Gibbons, S. (2001). "Paying For Good Neighbours? Neighbourhood Deprivation and the Community Benefits of Education." Centre for the Economics of Education, *Working Paper*, October.
- Glaeser, E.L., J. Gyourko, and R.E. Saks (2005). "Why Have Housing Prices Gone Up?" *American Economic Review* 95(2): 329–33.
- Glaeser, E.L., J. Gyourko, and A. Saiz (2008). "Housing Supply and Housing Bubbles." *Journal of Urban Economics* 64(2): 198–217.
- Glaeser, E.L., and A. Saiz (2004). "The Rise of the Skilled City." *Brookings-Wharton Papers on Urban Affairs* 5: 47–105.
- Goodman, A.C., and T.G. Thibodeau (2008). "Where are the Speculative Bubbles in US Housing Market." *Journal of Housing Economics* 17(2): 117–37.
- Gupta, R., and S. M. Miller. (2010). "The Time-Series Properties on Housing Prices: A Case Study of the Southern California Market." *The Journal of Real Estate Finance and Economics*. DOI: 10.1007/s1146-010-9234-7.
- Gyourko, J., and P. Linneman (1993). "The Affordability of the American Dream: An Examination of the Last 30 Years." *Journal of Housing Research* 4(1): 39–72.
- Gyourko, J., A. Saiz, and A. Summers (2008). "A New Measure of the Local Regulatory Environment for Housing Markets." *Urban Studies* 45(3): 693–721.
- Gyourko, J., and J. Tracy (1999). "A Look at Real Housing Prices and Incomes: Some Implications for Housing Affordability and Quality." *FRBNY Economic Policy Review*, September.
- Harvey, A.C. (1989). *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- Holly, S., M.H. Pesaran, and T. Yamagata (2011). "The Spatial and Temporal Diffusion of House Price in the UK." *Journal of Urban Economics* 69(1), 2–23.
- Hott, C., and P. Monnin (2008). "Fundamental Real Estate Prices: An Empirical Estimation with International Data." *Journal of Real Estate Finance and Economics* 36(4): 427–450.
- Kallberg, J.G., C.H. Liu, and P. Pasquariello (2011). "On the Price Comovement of U.S. Residential Real Estate Markets." *Working Paper*, August 12, 2011.
- Kim, C.-J., and C.R. Nelson (1999). *State-Space Models with Regime Switching*. Cambridge, MA, and London, MIT Press.
- Malpezzi, S., and S. Wachter, (2005). "The Role of Speculation in Real Estate Cycles." *Journal of Real Estate Literature* 13(2): 143–64.
- McMillan, D., and J. McDonald (1991). "Urban Land Value Functions with Endogenous Zoning." *Journal of Urban Economics* 28(1): 14–27.
- Miao, H., S. Ramchander, and M.S. Simpson (2011). "Return and Volatility Transmission in U.S. Housing Markets." *Real Estate Economics*. DOI: 10.1111/j.1540-6229.2010.00303.x.
- Moretti, E. (2003). "Human Capital Externalities in Cities." In *Handbook of Regional and Urban Economics*, Vol. 4, J.V. Henderson and J.-F. Thisse (ed.), Amsterdam: North-Holland: 2243–2291.

- Mulder, C.H. (2006). "Population and Housing: A Two-sided Relationship." *Demographic Research* 13(13): 401–412.
- Ottoviano, G.I.P., and G. Peri (2007). "The Effects of Immigration on US Wages and Rents: A General Equilibrium Approach." *CEPR Discussion Paper No. DP6551*.
- Pavlov, A.D., and S.M. Wachter (2010). "Subprime Lending and Real Estate Prices." *Real Estate Economics* 39(1), 1–17.
- Pogodzinski, J.M., and T.R. Sass (1994). "The Theory and Estimation of Endogenous Zoning." *Regional Science and Urban Economics* 24(5): 601–630.
- Saiz, A. (2003). "Room in the Kitchen for the Melting Pot: Immigration and Rental Prices." *The Review of Economics and Statistics* 85(3): 502–521.
- Saiz, A. (2007). "Immigration and Housing Rents in American cities." *Journal of Urban Economics* 61(2): 345–371.
- Saiz (2010). "The Geographic Determinants of Housing Supply." *The Quarterly Journal of Economics* 125(3), 1253–1296.
- Topol, R. (1991). "Bubbles and Volatility of Stock Prices: Effect of Mimetic Contagion." *The Economic Journal* 407(101): 786–800.
- Tirtiroglu, D. (1992). "Efficiency in Housing Markets: Temporal and Spatial Dimensions". *Journal of Housing Economics* 2(3): 276–92.
- Wheaton, W.C. (1990). "Vacancy, Search, and Prices in a Housing market Matching Model." *Journal of Political Economy* 98(6): 1270–1292.
- Wheaton, W.C., and G. Nechayev (2008). "The 1998–2005 Housing "Bubble" and the Current "Correction": What's Different This Time?" *Journal of Real Estate Research* 30(1): 1–26.
- Wu, Y. (1997). "Rational Bubbles in the Stock Market: Accounting for the U.S. Stock-Price Volatility." *Economic Inquiry* 35(2): 309–19.

Table 1: Variable Definitions

Variable	Abbrev.	Description
Panel A. MSA-Level and National Fundamentals – Stage 1		
House Price	P	MSA level, S&P/Case-Shiller Home Price Index seasonally adjusted, monthly frequency
Employment	E	MSA level, in thousand persons, seasonally adjusted, monthly frequency
Housing Starts	HS	MSA level, privately owned housing starts, authorized by building permits, seasonally adjusted, monthly frequency
Unemployment	UE	MSA level, unemployment rate (%)seasonally adjusted, monthly frequency
Construction Spending	CS	National level, monthly, seasonally adjusted
Industrial Production	IP	National level, monthly, seasonally adjusted
Interest Rate	r	National level, effective interest rate, monthly, seasonally adjusted
Panel B: Weight Matrix Variables at MSA Level – Stage 2		
<i>Demand Factors</i>		
2005 Unemployment Rate	UE2005	Rate for civilian labor force, ACS 2005
Avg. Unemployment Rate	ØUE	Average unemployment rate, 1998-2008
Population Growth	ΔPOP	Population growth, 1990-2000
Population Density	POPD	Density estimates, mid 2000s, various sources
Population Level	POP	population size
Managerial Positions	MP	Share of civilian employment above age 16 in management, professional and related occupations, ACS 2005
Service Sector Employment	SSE	Share of civilian employment above age 16 in the following industries: information; professional, scientific, management; finance, insurance, real estate; educational services, health care; public administration; ACS 2005
Home Vacancy Rate	HVACR	Total housing units - Homeowner vacancy rate, ACS 2005
Rental Vacancy Rate	RVACR	Total housing units - Rental vacancy rate, ACS 2005
Median House Value	MHVAL	Median house value for owner-occupied units, ACS 2005
Median Rental Rate	MRVAL	Median gross rent for renter-occupied units, ACS 2005
Subprime Lending Intensity	SLI	Annual rate of increase from 2004 to 2006
Median Family Income	MFINC	In 2005 dollars, ACS 2005

Family Income per Capita	FINC	In 2005 dollars, ACS 2005
Families below Poverty Rate	F<PR	Percentage of all families with incomes in past 12 months below the poverty rate, ACS 2005
Quality of Life Index	QOL	Adjusted index, Albouy (2011), Table A1.
% Bachelor/Graduate Degree	%BC/GD	Percentage of population 25 years and over with bachelor's, master's, professional or doctorate degree
% Graduate Degree	%GD	Percentage of population 25 years and over with master's, professional or doctorate degree

Supply Factors

Undevelopable Area		Saiz (2007), Table 1
WRLURI		Wharton Residential Land Use Regulatory Index ; Saiz (2007), Table 1
Supply Elasticity		Saiz (2007), Table 8

Panel C: Control Variables – Stage 2

Ln(Fed Funds Rate)		Federal funds rate
Ln(LT FMR)		30yrs. Fixed Mortgage Rate
Δ Ln(S&P 500)		Log difference of closing price of S&P 500 stock price index
Ln(Sentiment Index)		University Michigan Consumer Sentiment Index

Notes: ACS stands for American Community Survey. For the weight matrix variables we record one observation per MSA. All other variables are recorded as time series per MSA.

Table 2: Spatial Lag Panel Data Models of Stochastic Trends with Alternative Individual Weight Matrices

Variables for Weight Matrices	Log-Likelihood	Rho	Std. Error	P-value
Demand Factors				
Population Indicators				
Population Level	1055.991	0.386	0.090	0.000
Population Growth	1075.670	0.555	0.077	0.000
Population Density	1058.588	0.368	0.092	0.000
Socio-economic Indicators				
2005 Unemployment Rate	1046.594	0.139	0.092	0.132
Avg. Unemployment Rate	1045.539	0.087	0.079	0.271
Median Family Income	1050.016	0.256	0.121	0.035
Family Income per Capita	1047.781	0.178	0.157	0.255
Families below Poverty Rate	1049.049	0.236	0.070	0.001
% Bachelor/Graduate Degree	1045.487	0.097	0.110	0.379
% Graduate Degree	1046.890	0.162	0.114	0.153
Managerial Positions	1055.115	0.368	0.124	0.003
Service Sector Employment	1052.501	0.322	0.082	0.000
Subprime Lending Intensity	1053.779	0.303	0.110	0.006
Amenity Indicator				
Quality of Life Index	1051.114	0.284	0.076	0.000
Housing Market Related Indicators				
Home Vacancy Rate	1047.016	0.125	0.066	0.059
Rental Vacancy Rate	1054.645	0.301	0.063	0.000
Median House Value	1056.116	0.359	0.089	0.000
Median Rental Value	1075.586	0.538	0.068	0.000
Supply Factors				
Undevelopable Area	1055.671	0.346	0.086	0.000
WRLURI	1057.558	0.382	0.089	0.000
Supply Elasticity	1069.552	0.432	0.080	0.000

Notes: The estimation results are derived from Equation (5) based on annual data from 1995 to 2008 with the stochastic trend component in first differences as dependent variable, $\Delta B_{n,t}$. The co-movements of the dependent variables across the 20 MSAs are captured by a spatial lag based on alternative demand- and supply-side variables, which are described in Table 1. . . ***, **, and * denote significance at the 1, 5, and 10% level. Standard errors are reported in parentheses.

Table 3: Spearman Rank Correlations over 20 MSAs of Variables Intended for Spatial Weight Matrices

	UE2005	ØUE	ΔPOP	POPD	POP	MP	SSE	HVACR	RVACR	MHVAL	MRVAL
2005 Unemployment Rate	1										
Avg. Unemployment Rate	0.863*	1									
Population Growth	-0.256	-0.305	1								
Population Density	0.500*	0.460*	-0.574*	1							
Population Level	0.058	0.111	-0.311	0.570*	1						
Managerial Positions	-0.329	-0.353	-0.325	0.080	0.212	1					
Service Sector Employment	-0.447*	-0.439	-0.436	0.111	0.208	0.605*	1				
Home Vacancy Rate	0.275	0.186	0.423	-0.312	-0.349	-0.549*	-0.443	1			
Rental Vacancy Rate	0.427	0.153	0.338	-0.141	-0.274	-0.435	-0.457*	0.645*	1		
Median House Value	-0.429	-0.237	-0.239	0.086	0.280	0.549*	0.550*	-0.529*	-0.874*	1	
Median Rental Rate	-0.484*	-0.407	-0.159	0.102	0.394	0.526*	0.587*	-0.385	-0.744*	0.910*	1
Subprime Lending Intensity	0.063	0.179	0.037	0.131	0.433	-0.621*	-0.335	0.248	-0.017	-0.120	0.006
Median Family Income	-0.288	-0.303	-0.349	0.047	0.208	0.959*	0.547*	-0.474*	-0.381	0.555*	0.517*
Family Income per Capita	-0.286	-0.307	-0.356	0.209	0.283	0.949*	0.637*	-0.505*	-0.463*	0.590*	0.554*
Families below Poverty Rate	0.543*	0.659*	-0.081	0.372	0.269	-0.709*	-0.381	0.227	0.192	-0.371	-0.396
Quality of Life Index	-0.190	-0.032	-0.129	0.065	0.026	0.318	0.472*	-0.487*	-0.670*	0.712*	0.548*
Undevelopable Area	0.077	0.211	-0.448*	0.256	0.101	-0.027	0.317	-0.449*	-0.595*	0.522*	0.396
Supply Elasticity	-0.093	-0.255	0.630*	-0.473*	-0.399	-0.205	-0.473*	0.603*	0.693*	-0.628*	-0.498*
WRLURI	-0.275	-0.165	-0.190	0.081	0.293	0.451*	0.559*	-0.524*	-0.542*	0.624*	0.484*
% Bachelor/Graduate Degree	-0.321	-0.383	-0.175	0.061	0.203	0.946*	0.643*	-0.477*	-0.350	0.532*	0.521*
% Graduate Degree	-0.246	-0.219	-0.465*	0.132	0.334	0.916*	0.723*	-0.488*	-0.492*	0.648*	0.605*

Table 3 continues on the next page.

Table 3 (continued)

	SLI	MFINC	FINC	F<PR	QOL	UDA	SUPEL	WRLURI	%BC/GD	%GD
Subprime Lending Intensity	1									
Median Family Income	-0.558*	1								
Family Income per Capita	-0.519*	0.936*	1							
Families below Poverty Rate	0.560*	-0.744*	-0.687*	1						
Quality of Life Index	-0.302	0.233	0.326	-0.121	1					
Undevelopable Area	0.095	-0.056	-0.014	0.188	0.663*	1				
Supply Elasticity	-0.077	-0.175	-0.245	-0.188	-0.696*	-0.873*	1			
WRLURI	-0.224	0.423	0.499*	-0.191	0.691*	0.406	-0.640*	1		
% Bachelor/Graduate Degree	-0.665*	0.903*	0.922*	-0.702*	0.350	-0.096	-0.165	0.518*	1	
% Graduate Degree	-0.461*	0.922*	0.926*	-0.574*	0.427	0.110	-0.402	0.598*	0.890*	1

Notes: The variable descriptions are provided in Table 1. * denotes significance at the 5% significance level. Coefficients in bold denote significantly positive correlations.

Table 4: Spatial Lag Panel Data Models of Stochastic Trends of 20 MSAs with Alternative Weight Matrices

	Model Ia	Model IIa	Model IIIa	Model IVa	Model Va	Model VIa	Model VIIa	Model VIIIa	Model IXa	Model Xa
ρ (Socially Deprived)	0.369*** (0.121)									
ρ (Population)		0.502*** (0.091)								
ρ (Social Status)			0.401*** (0.116)							
ρ (Quality of Life)				0.284*** (0.075)						
ρ (Price-to-Rent Ratio)					0.490*** (0.086)					
ρ (Vacancy Rates)						0.202** (0.082)				
ρ (Population Growth)							0.555*** (0.077)			
ρ (Subprime Lending)								0.303*** (0.110)		
ρ (Scarce Land)									0.536*** (0.097)	
ρ (Supply Elasticity)										0.432*** (0.080)
Ln(Fed Funds Rate)	-0.002** (0.001)	-0.002* (0.001)	-0.002** (0.001)	-0.002** (0.001)	-0.002* (0.001)	-0.003*** (0.001)	-0.002* (0.001)	-0.003*** (0.001)	-0.002 (0.001)	-0.002* (0.01)
Ln(LT FMR)	-0.010** (0.005)	-0.009 (0.005)	-0.010* (0.005)	-0.011** (0.005)	-0.009* (0.005)	-0.012** (0.005)	-0.008 (0.006)	-0.011** (0.006)	-0.007 (0.005)	-0.009* (0.005)
Δ Ln(S&P 500)	-0.014 (0.018)	-0.017 (0.020)	-0.015 (0.018)	-0.018 (0.018)	-0.013 (0.017)	-0.019 (0.017)	-0.007 (0.017)	-0.019 (0.018)	-0.012 (0.020)	-0.012 (0.018)

Ln(Sentiment Index)	0.036 ^{***} (0.007)	0.031 ^{***} (0.008)	0.036 ^{***} (0.009)	0.041 ^{***} (0.007)	0.031 ^{***} (0.007)	0.045 ^{***} (0.008)	0.028 ^{***} (0.009)	0.042 ^{***} (0.007)	0.026 ^{***} (0.007)	0.032 ^{***} (0.009)
MSA-FE	Yes									
Log Likelihood	1054.194	1063.917	1060.888	1051.114	1064.338	1047.972	1075.670	1053.779	1068.68	1069.552

Notes: This table shows the estimation results of the panel data model of Equation (5) based on annual data for the period 1995 to 2008. The spatial lag variables consist of ratios or of Mahalanobis distance measures created from the variables reported in Table 3. Aggregation is based on the highly significant correlation coefficients of Table 3. From the 21 supply and demand variables we form the following 7 aggregated variables, Socially Deprived (Unemployment rate 2005, unemployment rate (avg.), percentage of families below poverty rate), Population (population and population density), Social Status (family income per capita, median family income, share of employment in managerial positions, percent with bachelor or graduate degree, percent with graduate degree), Amenity (Quality of Life), Price-to-Rent Ratio (median house value divided by median rental value), Vacancy Rates (home vacancy rate and rental vacancy rate), and Scarce Land (undevelopable area and WRLURI); and three variables taken from Table 2, Population Growth, Subprime Lending Intensity, and Supply Elasticity. ^{***}, ^{**}, and ^{*} denote significance at the 1, 5, and 10% level. Standard errors are reported in parentheses.

Table 5: Spatial Lag Panel Data Models of Stochastic Trends of 20 MSAs with Two Weight Matrices

	Model Ib	Model IIb	Model IIIb	Model IVb	Model Vb	Model VIb	Model VIIb	Model VIIIb	Model IXb	Model Xb
ρ (Socially Deprived)									0.042 (0.129)	
ρ (Population)				0.259 (0.165)				0.307** (0.141)		
ρ (Social Status)			0.229* (0.133)							0.244** (0.115)
ρ (Quality of Life)										
ρ (Price-to-Rent Ratio)						0.222 (0.153)				
ρ (Vacancy Rates)										
ρ (Population Growth)	0.408*** (0.102)	0.409*** (0.145)	0.460*** (0.104)	0.439*** (0.143)		0.433*** (0.141)	0.503*** (0.101)		0.534*** (0.111)	
ρ (Subprime Lending)							0.138 (0.106)			
ρ (Scarce Land)		0.313** (0.159)			0.354*** (0.120)					
ρ (Supply Elasticity)	0.277*** (0.100)				0.305*** (0.107)			0.328*** (0.125)		0.347*** (0.083)
Log Likelihood	1085.757	1084.230	1081.548	1080.891	1080.080	1078.884	1077.727	1076.963	1075.790	1075.705

Table 6 continues on the next page.

Table 5 (continued)

	Model XIb	Model XIIb	Model XIIIb	Model XIVb	Model XVb	Model XVIb	Model XVIIb	Model XVIIIb	Model XIXb	Model XXb
ρ (Socially Deprived)										0.119 (0.116)
ρ (Population)								0.237 (0.152)		
ρ (Social Status)						0.271** (0.132)			0.185 (0.182)	
ρ (Quality of Life)	0.020 (0.099)									
ρ (Price-to-Rent Ratio)			0.231*** (0.082)	0.282** (0.117)		0.371*** (0.106)				
ρ (Vacancy Rates)		-0.010 (0.090)			0.250** (0.109)					
ρ (Population Growth)	0.546*** (0.104)	0.558*** (0.089)								
ρ (Subprime Lending)							0.148 (0.097)			
ρ (Scarce Land)				0.370*** (0.132)				0.378** (0.161)	0.411** (0.175)	0.390*** (0.094)
ρ (Supply Elasticity)			0.348*** (0.109)		0.311*** (0.109)		0.385*** (0.098)			
Log Likelihood	1075.702	1075.680	1074.585	1073.939	1073.482	1072.018	1071.809	1071.567	1071.336	1070.541

Notes: This table shows results of the panel data model of Equation (5) based on annual data for the period 1995 to 2008. From all possible pairwise combinations of the seven aggregated variables and the variables population growth, subprime lending, and supply elasticity, those with the highest Log-Likelihood value are selected. ***, **, and * denote significance at the 1, 5, and 10% level. Standard errors are reported in parentheses.

Table 6: Estimated Spatial Lag Impact of All Other MSAs on the 11 MSAs with the Largest Deviation from Market Fundamentals

Impacted MSAs	MSAs impacting the MSA in the first column with the given spatial lag coefficient																			
	SD	MI	SF	DC	CHI	NY	BO	TPA	MPLS	SEA	DET	CLV	PD	DE	DA	CL	PH	AT	LV	
LA	0.405	0.364	0.146	0.033	0.011	0.008	0.007	0.005	0.004	0.004	0.003	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.000
	LA	SF	BO	NY	CHI	DC	SEA	MI	CLV	TPA	DET	PD	MPLS	DE	PH	DA	AT	LV	CL	
SD	0.392	0.203	0.098	0.075	0.066	0.032	0.030	0.028	0.014	0.012	0.010	0.010	0.009	0.006	0.005	0.003	0.003	0.003	0.003	0.002
	LA	SEA	MPLS	PD	SF	TPA	CL	SD	DA	DC	DV	CHI	NY	BO	DET	CLV	AT	PH	LV	
MI	0.389	0.139	0.054	0.050	0.046	0.044	0.032	0.031	0.031	0.029	0.027	0.026	0.024	0.021	0.016	0.014	0.014	0.011	0.004	
	DC	AT	DE	MPLS	DA	PD	CL	TPA	MI	SEA	DET	SD	CHI	SF	LA	LV	CLV	NY	BO	
PH	0.290	0.117	0.072	0.050	0.049	0.047	0.046	0.035	0.034	0.033	0.029	0.026	0.026	0.026	0.026	0.024	0.024	0.024	0.023	
	SD	LA	MI	NY	DC	BO	CHI	SEA	TPA	CLV	MPLS	DET	PD	DE	PH	DA	AT	CL	LV	
SF	0.323	0.314	0.119	0.050	0.044	0.041	0.031	0.015	0.012	0.009	0.009	0.008	0.007	0.005	0.004	0.003	0.002	0.002	0.002	0.002
	BO	SD	CHI	SF	LA	DET	DC	CLV	MI	TPA	SEA	MPLS	PD	DE	DA	CL	PH	AT	LV	
NY	0.300	0.094	0.094	0.085	0.073	0.065	0.045	0.043	0.043	0.031	0.031	0.027	0.016	0.012	0.011	0.010	0.008	0.008	0.004	
	PH	LA	SD	SF	MPLS	CHI	TPA	DE	DET	NY	SEA	PD	BO	CLV	MI	DA	AT	CL	LV	
DC	0.266	0.164	0.153	0.132	0.044	0.039	0.035	0.030	0.020	0.019	0.016	0.015	0.015	0.014	0.012	0.009	0.007	0.006	0.006	0.006
	DET	MPLS	PD	DC	SD	LA	SF	SEA	CHI	MI	CLV	NY	DE	BO	DA	CL	PH	AT	LV	
TPA	0.212	0.163	0.114	0.066	0.056	0.055	0.053	0.050	0.041	0.033	0.028	0.027	0.024	0.023	0.014	0.014	0.013	0.009	0.005	
	MI	CHI	MPLS	SD	TPA	SF	LA	PD	NY	BO	CLV	DC	DET	DE	DA	CL	PH	AT	LV	
SEA	0.145	0.113	0.090	0.075	0.074	0.053	0.052	0.052	0.051	0.050	0.049	0.042	0.032	0.031	0.026	0.026	0.017	0.015	0.008	
	DA	AT	PH	DE	CL	PD	DC	MPLS	TPA	SEA	MI	DET	SD	CHI	SF	LA	CLV	NY	BO	
LV	0.186	0.119	0.068	0.051	0.048	0.045	0.044	0.043	0.040	0.039	0.039	0.036	0.035	0.035	0.035	0.035	0.034	0.034	0.033	
	NY	DET	SD	CHI	SF	CLV	LA	MI	DC	SEA	TPA	MPLS	PD	DE	DA	CL	PH	AT	LV	
BO	0.298	0.139	0.095	0.068	0.065	0.062	0.056	0.037	0.034	0.030	0.026	0.023	0.015	0.012	0.010	0.010	0.008	0.008	0.004	

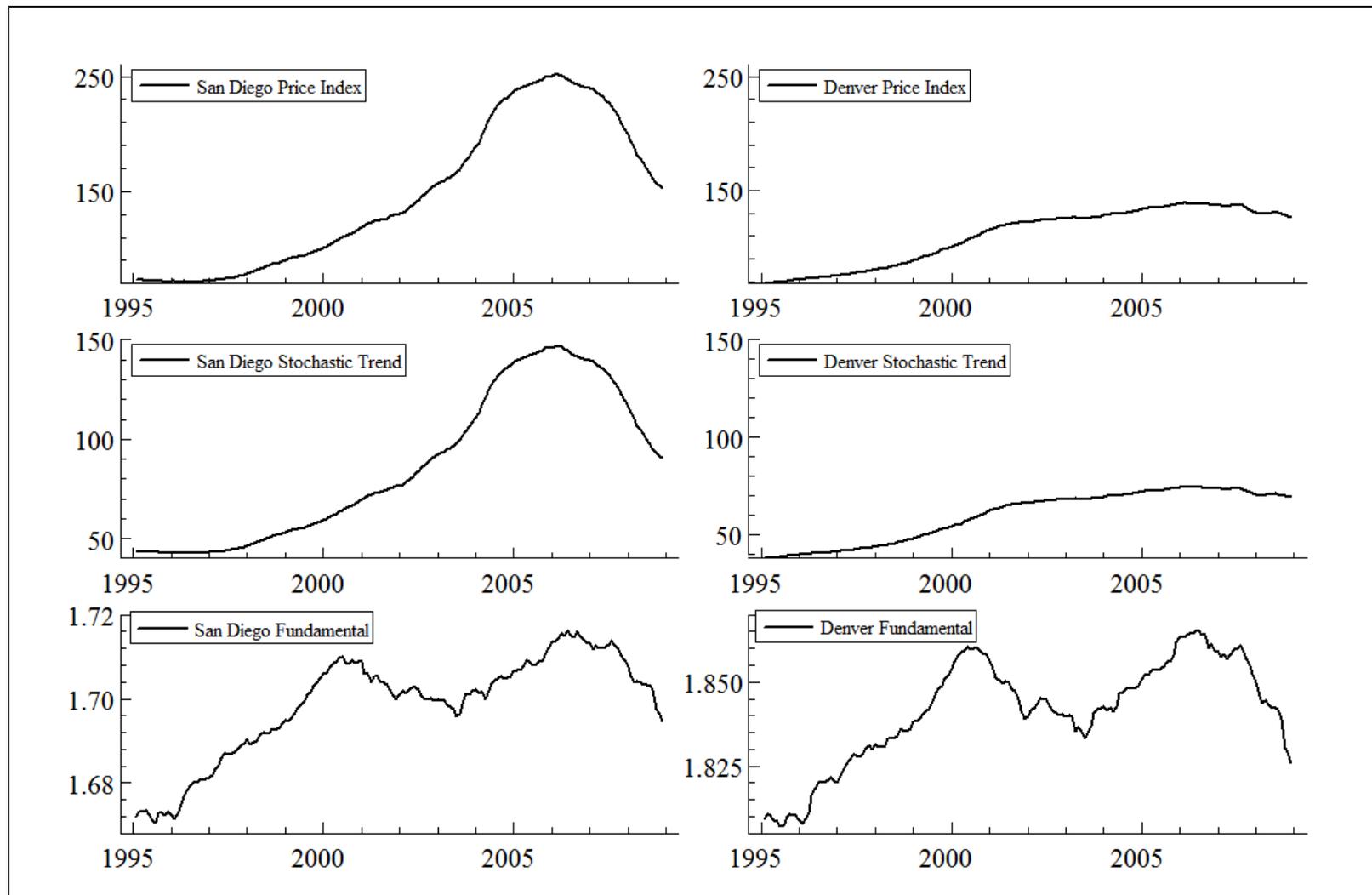
Notes: Estimates are derived from the best fitting model of Table A.1 based on annual data from the period 1995 to 2008. The impact coefficients are based on two weight matrices, one for population growth, with estimated parameter 0.408, and one for the supply elasticity, with estimated parameter 0.277, which is the first model shown in Table A.1. The twenty MSA are: Atlanta (AT), Boston (BO), Charlotte (CL), Chicago (CHI), Cleveland (CLV), Dallas (DA), Denver (DV), Detroit (DET), Las Vegas (LV), Los Angeles (LA), Miami (MI), Minneapolis (MPLS), New York (NY), Phoenix (PH), Portland (PD), San Diego (SD), San Francisco (SF), Seattle (SEA), Tampa (TPA), and Washington (DC).

Table 7: Spatial Lag Panel Data Models of Stochastic Trends of 20 MSAs with Three Weight Matrices

	Model I	Model II	Model III	Model IV	Model V	Model VI	Model VII	Model VIII
ρ (Population Growth)	0.135 (0.107)	0.128 (0.112)	0.159 (0.111)	0.330** (0.158)	0.240* (0.141)	0.379*** (0.138)	0.384*** (0.138)	0.460*** (0.141)
ρ (Supply Elasticity)	0.208** (0.083)	0.190** (0.083)	0.192** (0.085)	0.237** (0.109)	0.184** (0.093)	0.260** (0.105)	0.261** (0.110)	0.315*** (0.085)
Contiguity 1	0.406*** (0.056)							
Contiguity 2		0.420*** (0.068)						
Contiguity 3			0.400*** (0.067)					
Contiguity 4				0.167 (0.154)				
Contiguity 5					0.324*** (0.099)			
Contiguity 6						0.062 (0.118)		
Coast 1							0.049 (0.067)	
Coast 2								-0.118 (0.147)
Log Likelihood	1117.840	1115.383	1113.828	1088.170	1099.747	1086.263	1085.868	1086.402

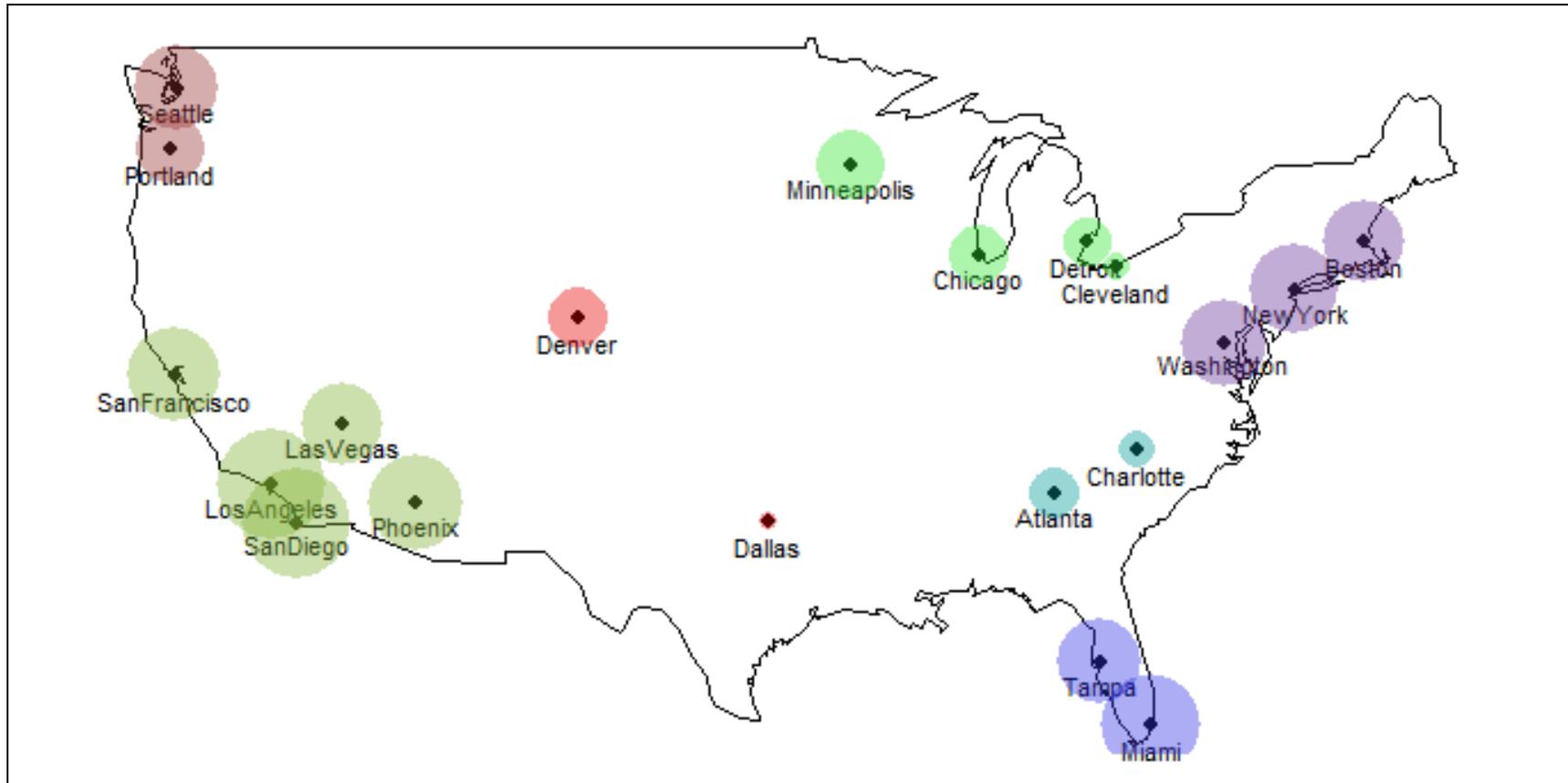
Notes: The estimation results are derived from Equation (5) based on annual data from 1995 to 2008. ***, **, and * denote significance at the 1, 5, and 10% level. Standard errors are reported in parentheses. Alternative spatial contiguity matrices are added as the third spatial weight matrix in addition to one matrix for population growth and one for supply elasticity. Each contiguity matrix identifies a particular set of regional clusters. For example, Contiguity 5 below lumps all western MSAs including Denver into one regional cluster, all MSAs on or close to the East Coast in another, and all interior MSAs between Charlotte and Dallas into a third regional cluster. **Contiguity 1** = (SEA-PD; SF-LA-SD-LV-PH; DE-DA; MPLS-CHI-DET-CLV; BO-NY-DC; AT-CL; TPA-MI); **Contiguity 2** = (SEA-PD-DE; SF-LA-SD-LV-PH; MPLS-CHI-DET-CLV; BO-NY-DC; AT-CL-DA; TPA-MI); **Contiguity 3** = (SEA-PD-DE; SF-LA-SD-LV-PH; DA-MPLS-CHI-DET-CLV; BO-NY-DC; AT-CL; TPA-MI); **Contiguity 4** = (SEA-PD-DE-SF-LA-SD-LV-PH; DA-MPLS-CHI-DET-CLV; BO-NY-DC-AT-CL-TPA-MI); **Contiguity 5** = (SEA-PD-DE-SF-LA-SD-LV-PH; DA-MPLS-CHI-DET-CLV-AT-CL; BO-NY-DC-TPA-MI); **Contiguity 6** = (SEA-PD- SF-LA-SD; DE-LV-PH; MPLS-CHI-DET-CLV-AT-CL-DA; BO-NY-DC-TPA-MI); **Coast 1** = (SEA-CHI-CLV-SF-LA-SD-BO-NY-TPA-MI; PD-DE-LV-PH-MPLS-DET-CLV-AT-CL-DA-DC); **Coast 2** = (SEA-SF-LA-SD-BO-NY-TPA-MI; PD-DE-LV-PH-MPLS-CHI-DET-CLV-AT-CL-DA-DC).

Figure 1: Observed Price Indices Decomposed into Stochastic Trend Components and Predictions of Market Fundamentals for San Diego and Denver



Notes: This figure shows the observed S&P/Case-Shiller Home Price Indices for the MSAs San Diego and Denver in the upper panels, the stochastic trend components not explained by market fundamentals in the center (series B of equation 4.2) and the price predictions based on market fundamentals in the bottom panels. Estimates are based on monthly data for the period January 1995 to December 2008. For each MSA, the product of the series in the bottom and center panels gives the predicted price index, which deviates from the observed price index by the exponent of the residual series ζ in equation 4.1.

Figure 2: Maximum Deviations from Market Fundamentals



Notes: The size of the circles indicates the percentage difference between the maximum and minimum value of the stochastic trend B (Equation 4.2), which we interpret as the deviation from market fundamentals. Estimates are based on monthly data for the period January 1995 to December 2008. Cleveland and Dallas both show negligible deviations from market fundamentals. The color pattern identifies a particular regional clustering that works well empirically at the second estimation stage. For instance, Denver and Dallas share one cluster as do Atlanta and Charlotte.

Figure 3: Estimated Linkages of Stochastic Trends among the 20 MSAs Conditional on Two Spatial Weight Matrices

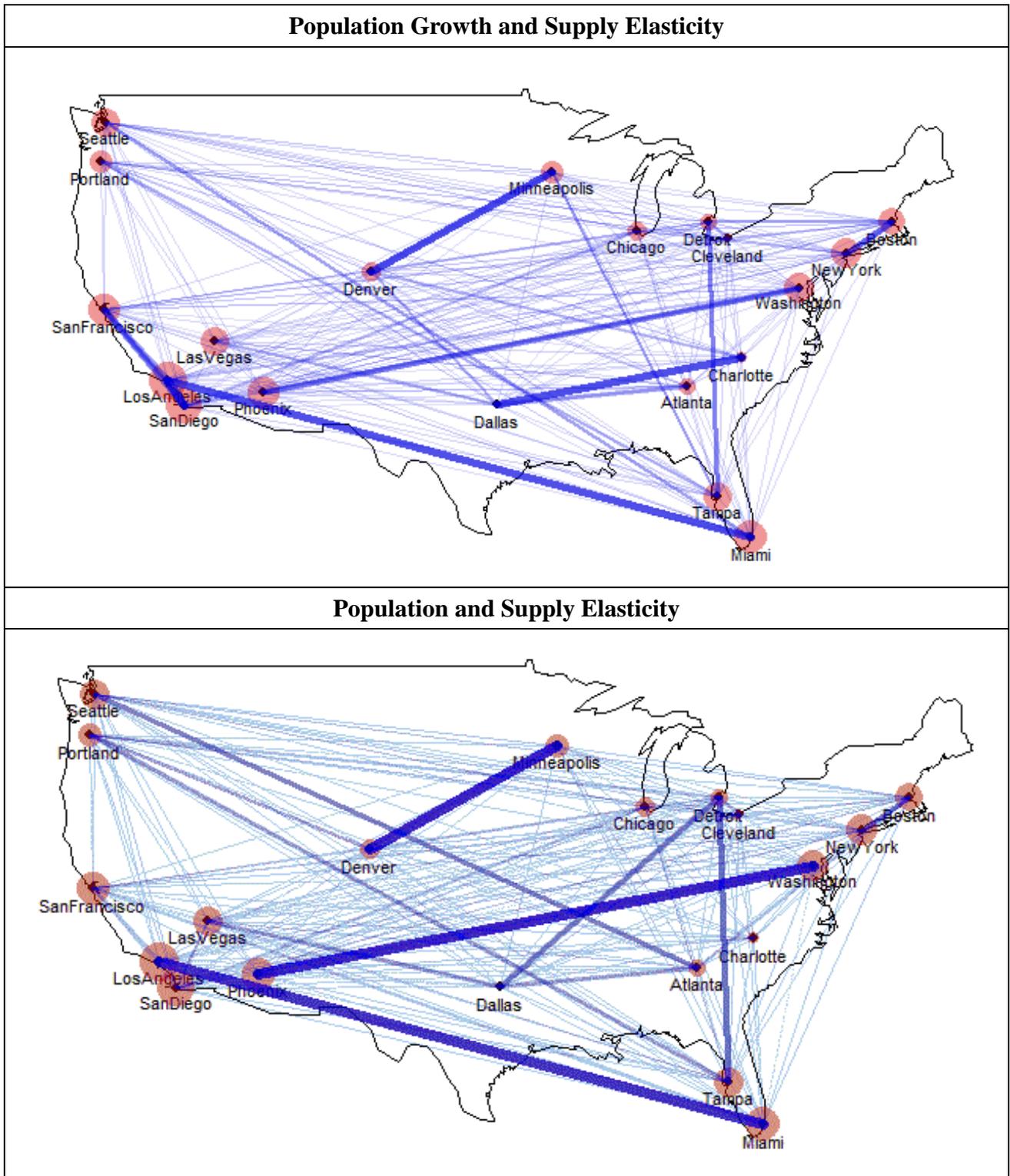


Figure 3 continues on the next page.

Figure 3 (continued)

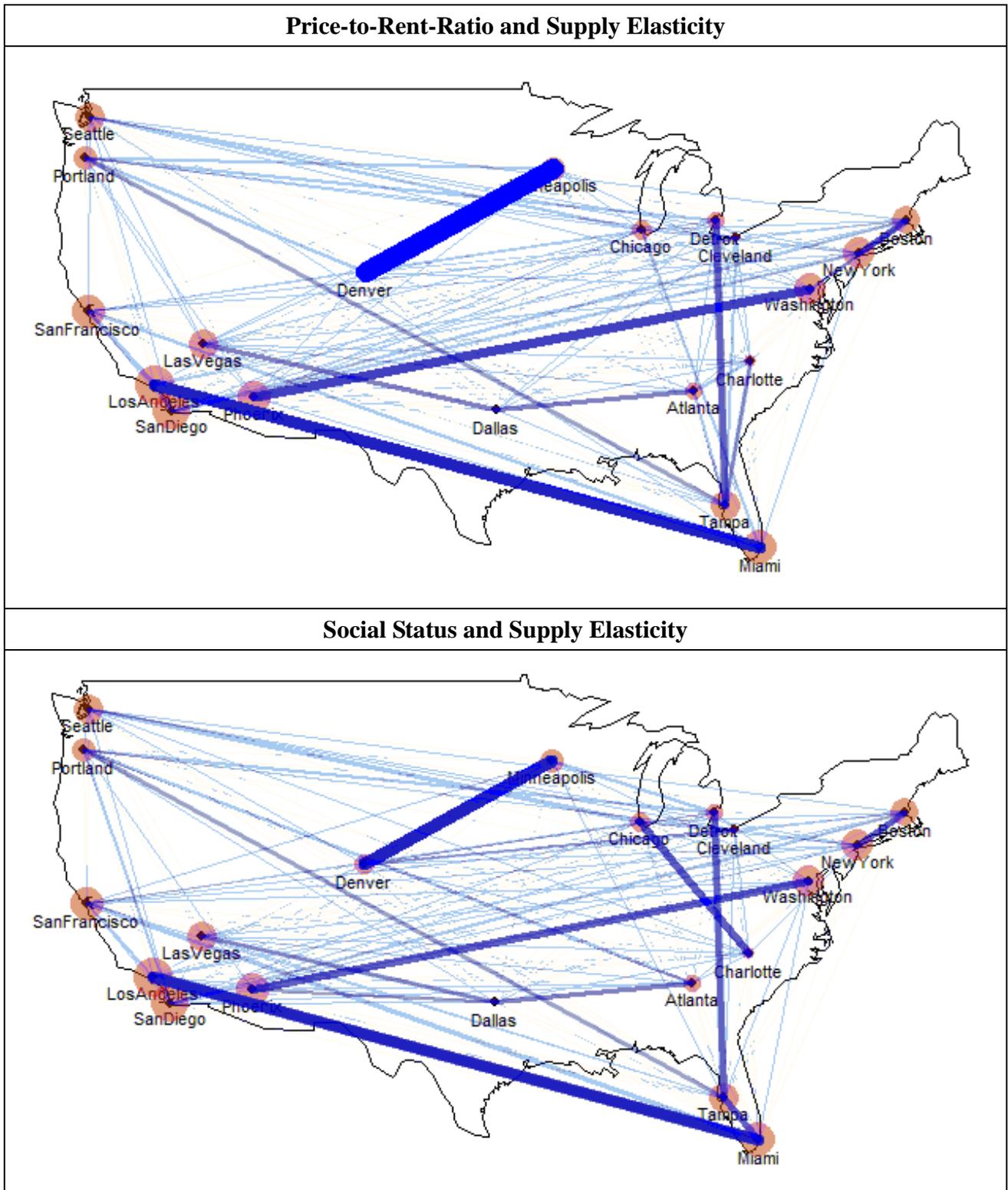
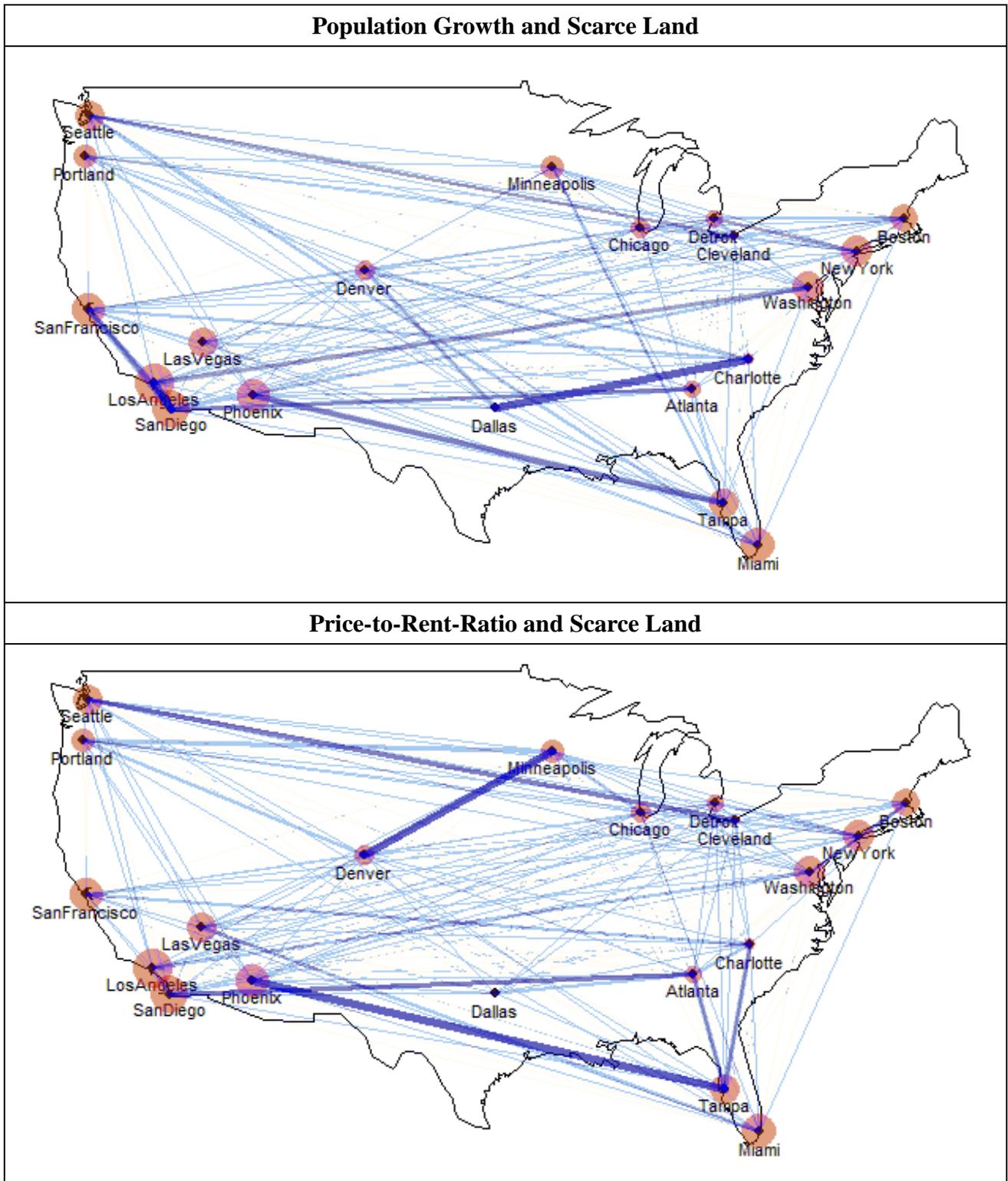


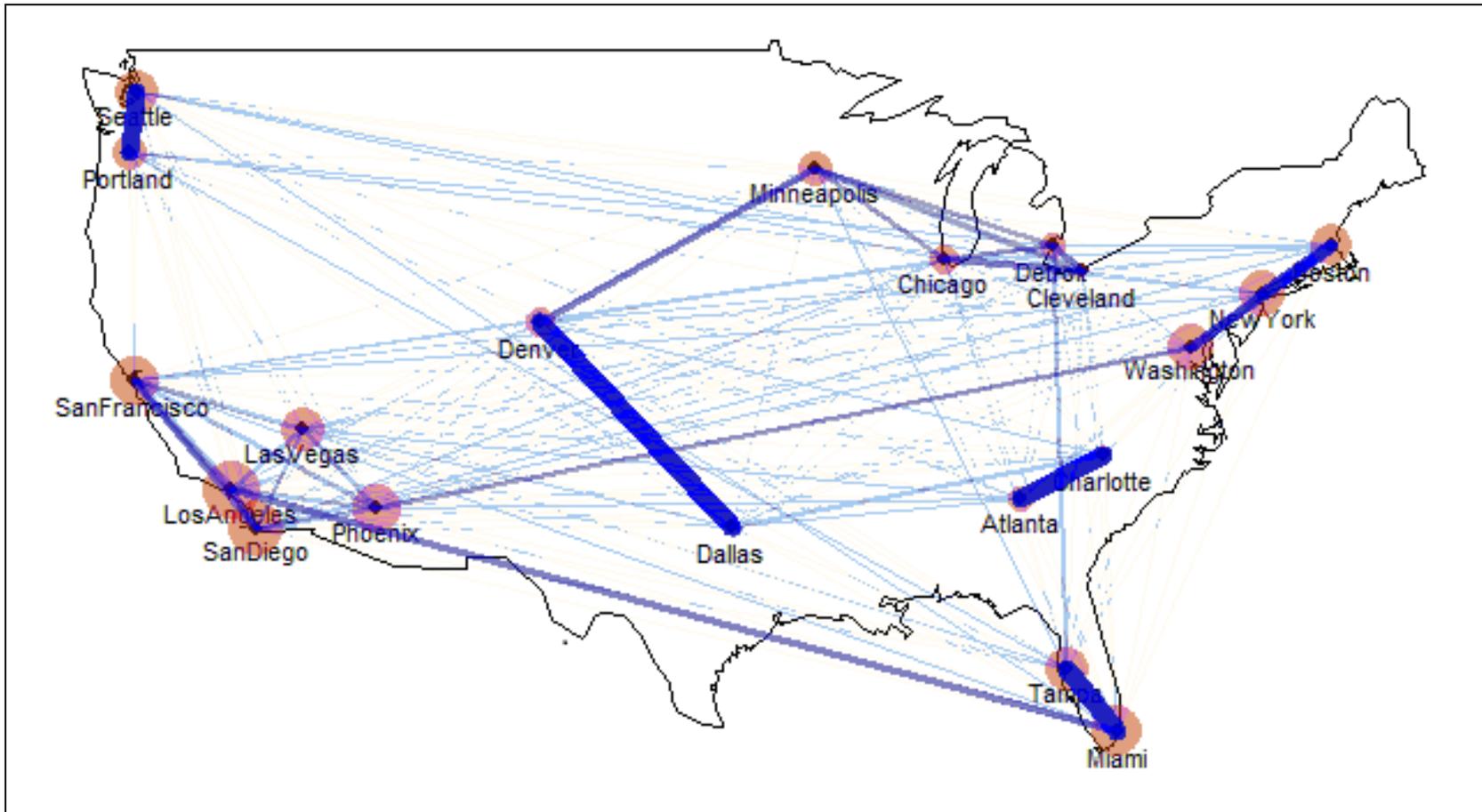
Figure 3 continues on the next page.

Figure 3 (continued)



Notes: Each graph is based on two weight matrices as identified in the heading. One represents the demand side and the other the supply side. The linkages shown in Table 6 are graphed for all 20 MSAs. Larger coefficients are identified by thicker lines.

Figure 4: Estimated Linkages of Stochastic Trends among the 20 MSAs Conditional on Three Spatial Weight Matrices



Notes: Population growth represents the demand side, the supply elasticity the supply side, and a contiguity matrix implements the regional clustering shown by different coloring in Figure 1.