Housing Market Spillovers in a System of Cities

Thom Malone*

Abstract

It is well-documented that housing leads the business cycle at the national level. Dating back as far as the great depression, nearly every recession or boom has been preceded by a respective drop or rise in residential investment (Green, 1997; Leamer, 2007). However, neither the business cycle nor the housing cycle is truly national in their nature. Historically, the severity and timing of both vary greatly across cities. This has been acknowledged in the literature (Ghent and Owyang, 2010), but the possibility that different cities housing markets and economies could be interacting, and spilling over amongst one another has not yet been considered.

I use a Global Vector Autoregression to model building permits and employment at the metro level for 78 cities in the U.S from 1990-2015, then link these models together in a system of cities. The model reveals several results of interest. First, the status of residential investment as a leading indicator for the business cycle is questionable at the metro level, and is found to be true in only 47 of 78 cities. Second, shocks to housing in cities do spill over into other cities. In fact, in many cases the responses to shocks are larger in other cities than in those where the shocks originated. Thus, the national relationship between housing and the economy appears to be created by a collection of cities housing markets spilling over into other cities. Third, the largest spillovers are often not found in adjacent cities to those where the shock occurred. This indicates that factors other than physical distance, such as trade and migration, are the channels shocks are transmitted through.

Overall, the results suggest a role for place-based policy, where stabilizing new housing development in particular cities could result in some stabilization of the national business cycle.

*Sol Price School of Public Policy, University of Southern California. Email: tdmalone@usc.edu. Mail: 117 13th Street NW, Charlottesville, VA 22903. Phone: 339-221-8633. Advisor: Christian Redfearn (Email:redfearn@price.usc.edu).

1 Introduction

The lead-lag relationship between housing and the business cycle at a national level has been well documented in the literature (e.g. Green (1997), Coulson and Kim (2000), Leamer (2007)). However, it is easy to see that there has been great variation by metropolitan region in house prices during recent upturns and downturns in the national housing market. While many works recognize this (Head et al. (2014), Malpezzi (1999)), few account for the possibility that the housing cycles of different metropolitan areas could potentially interact and lead to spillovers between one another. Such studies assume that a metropolitan area constitutes a housing market in and of itself, and choose to model different cities as being independent of one another. In this paper I choose to model this relationship from a different perspective; allowing shocks to cities housing and labor markets to transmit themselves to other cities. The paper is thus attempting to answer the questions: Why does the housing lead the business cycle in some cities but not others? Do shocks to local housing markets spill over into other local economies? And in turn, is the lead-lag relationship between housing and the business cycle driven by the housing markets of a particular set of cities? If spillovers from shocks to particular cities are prevalent, it could be that stabilizing residential investment in these areas would have some stabilizing effect on the business cycle as a whole. The work is, to my knowledge, the first to account for spillovers among different areas when looking at the relationship between the housing cycle and the business cycle.

The main empirical task of the paper is to model both the residential investment and the business cycle in a system of cities where one changes in one city are allowed to spill over into another city. DeFusco et al. (2013) and Greenaway-McGrevy and Phillips (2016) have used reduced form approaches to model spillovers between an MSAs housing market (exclusive of the rest of the macroeconomy) and its neighboring housing markets. This approach works well, but I instead opt to take an approach that allows all the cities in the sample to exist simultaneously in a system. The preferable way of doing this would be to use a large-scale VAR, that includes allows all the cities variables to enter into each other equations as endogenous regressors, but this cannot be done is this context due to the curse of dimensionality. To approximate such a large scale VAR, I use a Global Vector Auto Regression (GVAR). The GVAR works similarly to a VAR, but instead of allowing all the other cities variables to enter each other's equations as endogenous, only the variables of that own city are allowed to be endogenous. Other cities variables enter into each cities equation as weakly exogenous, and the individual models are then linked together by some predetermined factor meant to proxy for the unobserved factors that link cities housing and business cycles together.

The GVAR has mostly been used in a cross-country setting, as in Hiebert and Vansteenkiste (2009) or Cesa-Bianchi (2013), but can easily be brought used to model housing markets at lower levels. For instance, used the method Vansteenkiste (2007) found evidence of spillovers between the 31 largest US states. However, I place MSA level employment and building permits in the model so as to measure not just spillovers between cities housing markets but also the other elements of the business cycle. Spillovers can then be identified using generalized impulse response functions (GIRFs) to assess how shocks to building permits, in particular, manifest themselves in other cities employment over time.

Several interesting observations are found in the GIRFS. The first is consistent with the findings of Ghent and Owyang (2010); the lead-lag relationship between housing and GDP is not consistently observed across the metropolitan area. The GIRFs for the response of a city's employment to an innovation in that same city's building permits are only positive in 47 of 78 cities. For the other 31 cities, an innovation in permits actually elicits a negative response in employment.

Those cities where the response is positive are found to be different from those where the response is negative in a number of ways. First, they tend to be geographically concentrated in the Midwest, and South rather than on the coastlines. Second, cities were people have relatively little debt relative to their income are also more likely to have a positive response. This is likely because of the variations in marginal propensity to consume across regions. Shocks to housing lead to the business cycle in areas where people spend less deleveraging debts, and instead spend any shocks to income on purchasing durable goods.

The GIRFS also show that cities in the sample are not independent of one another. There are significant spillovers to other cities employment when building permits are shocked in a city. The spillovers do not seem to be particularly associated with the distance between places. Spillovers are more often larger between cities that have heavy migration flows between them and that has similar industrial compositions. Moreover, the biggest recipients of spillovers from most cities are New York, Los Angeles, and Chicago. This raises the possibility that these cities are 'hubs' of sorts, and contrasts with some existing results in the literature. For example, DeFusco et al. (2013) found that the greatest spillovers between housing markets to be between neighboring cities. Similar results have been found in other Saks (2008) and Zabel (2012) where cities are connected via labor markets and migration flows.

Third, in those cities where housing does lead the business cycle, spillovers mean that it also leads the business cycle in other cities. Thus, the mismatch between the national trend and the inconsistent city level trend is reconciled through this, as it appears the cities that are consistent with the national trend are in fact driving the national trend. Impulse responses from an innovation in permits in these cities show substantial effects in employment in all the cities in the sample. Impulse responses from a shock to permits in the other cities in the sample fail to elicit a statistically significant response in employment.

The paper proceeds as follows; sections 2 provides the background and reviews the current literature considering both spillovers between housing markets and the relationship between housing and the other components of the business cycles, section 3 provides some motiving facts about housing and business cycle, and differences between cities housing markets, section 4 gives information on the data to be used, section 5 outlines the estimation of the GVAR model, section 6 presents the results. Finally, in section 7 the paper is summarized, shortcomings are discussed, and potential future extensions to the work are provided.

2 Literature Review

The literature that this paper most directly follows is a purely empirical one concerned the temporal lead-lag relationship between residential investment and other components of the business cycle. The earliest contribution that I know of in this literature is Green (1997), that examines whether residential and non-residential investment granger cause GDP and vice versa using data from 1957 to 1992. The paper shows that residential investment granger cause GDP and GDP in turns granger causes non-residential investment. The paper is entirely empirical, but Green does provide some intuition by explaining that the differences in the individual tax treatments of residential investment are a path that it could 'exogenously' move through to affect GDP and create the observed relationship.

The most prominent paper that documents the temporal relationship between housing and other elements of the business cycle is Leamer (2007). This study presents strong evidence that residential investment leads other elements of the business cycle and recommends a modification to monetary policy that would replace the output gap with changes in housing starts. The biggest contribution of the paper is the presentation of a very interesting stylized fact; eight of the last ten recessions were preceded by a downturn in residential investment. Furthermore, in six of the last ten recessions, residential investment was the greatest contributor to the changes in GDP the year before the recession and only twice did it not contribute significantly.

Interestingly, both Green (1997) and Leamer (2007) find that only residential investment, and not house prices, lead the business cycle. Leamer suggests the reason behind residential investments importance it that housing has a persistent volume cycle, but not a persistent price cycle. He reasons that if house prices are sticky downwards, a decline in demand adjusts the volume of sales and not the prices, causing a decline in construction. If prices could quickly adjust when the housing cycle goes down, then normal sales volume would quickly reappear, but the sluggishness of prices makes the volume cycle very extreme and in turn, this makes housing so important in recessions. Given the rather extreme volatility in house prices shown since the paper was written, the idea that house prices are sticky downwards seems doubtful, however. Overall, rigorous testing of the mechanism that drives the lead-lag is still missing from the literature.

The most recent paper to document the relationship is Ghent and Owyang (2010). This paper is thematically the most similar to this one, it investigates the relationship between housing and the business cycle using cross-sectional variation across cities. Their key finding is that although housing appears to be an important driver of cyclical fluctuations at a national level, the local results cast doubt on the idea that there is a direct channel from housing markets to local employment. Despite this, the paper still finds that permits lead the cycle in over 80 percent of their sample. This paper builds upon these results; by accounting for potential spillovers between metros I allow for the possibility that the national trend could be driven by specific cities or regions.

This paper uses a Global Vector Autoregression (GVAR) to assess spillovers from shocks to housing permits in one city on employment in other cities. The method was originally developed in Pesaran et al. (2004), and di Mauro et al. (2007), and has mostly been used to assess financial linkages and dependencies between countries. However, there is a growing body of literature that applies this method to assess the linkages between housing markets at different geographical levels. At the cross-country level, Cesa-Bianchi (2013) used a GVAR to investigate whether or not shocks to housing demand in particular countries affect real economic activity in other countries. The author used orthogonalized impulse response functions (OIRFs) to show that US housing demand shocks are quickly transmitted to the domestic real economy, leading to short-term expansion of real GDP. Though this result is unsurprising given the literature showing that housing leads the business cycle, a more interesting result is that, these shocks have similar effects on the real economic activity of other advanced economies. The spirit of the paper is similar to this one, in that is primarily measuring the effect of changes in the housing market in one location on other locations economies.

The method was also used in Hiebert and Vansteenkiste (2009); where quarterly data on 10 Euro countries was used to construct a GVAR for real house prices, real per capita disposable income and the real interest rate. The generalized impulse response functions (GIRFs) contrast Cesa-Bianchi (2013), finding limited evidence long term spillovers between different countries in the Eurozone. A stronger shifter of house price activity is domestic long-term interest rates, which was found to have long-term effects on house prices.

The US housing market was modeled in a GVAR framework at the state level in Vansteenkiste (2007). This paper used data from the 31 largest US states over the period 1986-2005 to investigate the extent that house prices spillover between states. The GIRFs show some interesting results; cross state spillovers for house prices are found, and the effect of a shock is state dependent, with shocks to states that have a relatively lower land supply elasticities having a much stronger spillover effect than other states.

This paper is not the first to investigate interactions between different MSAs housing markets. DeFusco et al. (2013) investigates whether or not contagion was an important factor in the last housing cycle. They find evidence of house price contagion in the latest boom period. Particularly, price elasticities between one MSA and its closest to the neighbor (defined by euclidean distance) range from 0.1 to 0.27. Local fundamentals and expectations of future fundamentals have very limited ability to account for the estimated effect, indicating that the effect is a pure contagion effect driven by irrational forces.

There is also a growing literature that explores the connection between housing markets and labor markets at the metropolitan level that this paper adds to. In particular, Saks (2008) investigates the differences in housing supply between metro areas, and how it generates substantial variation in house prices across the United States. Using a VAR she finds that locations with fewer construction barriers experience more residential construction and smaller increases in house prices in response to an increase in housing demand. Specifically, an increase in labor demand is also associated with an increase in wages but no change in house prices. Interaction terms show when housing supply has constrained the increase in demand leads to higher house prices, along with a higher level of wages, and a smaller increase in employment (because migration flows into the area are constrained).

Zabel(2012) builds upon Saks (2008) to investigate how the housing market affects the flow of workers across cities through both the relative mobility of homeowners relative to renters and the relative cost of housing across metro areas. He estimates a VAR of migration, employment, wages, house prices and housing supply data from 277 US MSA. The impulse response functions show that the effects of labor supply and housing demand shocks exhibit substantial variation when evaluated at different values of the homeownership rate, the price elasticity of housing supply and relative house prices. This study did allow for spillover effects in the VAR, but such spillovers are only accounted for in the nearest neighbor.

Some attention should also be paid to housing and the business cycle in the general macroeconomic literature. While the topic has been written about by macroeconomists for decades, in recent years it has been most commonly featured in the literature using aggregate equilibrium

”real business cycle” (RBC) models in the spirit of Kydland and Prescott (1982).

Davis and Heathcote (2007) is the first paper to model shocks of production affecting housing in the RBC framework. In this model, there are 2 types of firms; those who produce intermediate goods and those that transform intermediate goods into a good for final consumption. Within these two types, there are three types of intermediate good producers; construction, manufacturing, and services. These ”industries” produce from a stock of capital and labor rented from households. There are also three types of final good producing firms within the model; a firm that produces a good for either household consumption or business investment, a firm that produces residential investment and a firm that produces housing. The first two types of firms use all three intermediate goods in production. The third firm uses residential investment and land to make housing. Only when one assumes that the amount of new land available in the economy in each new time period is fixed does the model have a closed form solution. This solution has great success at modeling the relative volatility of residential investment, accurately predicting it to be about twice as volatile as non-residential investment. However, the solution is less successful in matching the data along three dimensions that are very relevant to this study; the volatility of house prices is greatly underestimated, a negative correlation between residential investment and house prices is predicted. Most importantly residential investment is usually not found to lead GDP and non-residential investment is usually not found to lag GDP.

More recent works in the RBC literature have made some resolutions to the most unrealistic findings in Davis and Heathcote (2007). Fisher (2007) that shows when housing is instead modeled as a separate capital stock in and of itself in the production function, the lead-lag relationship of residential investment, GDP, and non-residential investment is restored. Van Nieuwerburgh et al. (2015) uses a heterogeneous agent equilibrium model with collateral constraints to generate considerably more volatility in house prices within the framework to Davis and Heathcote (2007). By doing so they improve the lead-lag relationship even more so than Fisher (2007). The reasons behind the lead-lag relationship are also explored in Kydland et al. (2012) that shows that fixed-rate mortgages may be the key. They find that the lead in residential investment is driven by those structures that rely on mortgage finance and it is specifically fixed-rate mortgages that account for growth in residential investment that leads the growth in real GDP. Furthermore, the dynamics of the 30-year mortgage interest rate suggest that mortgages are relatively cheap ahead of an economic upturn. The reasons behind this are not clear, however.

The results in Kydland et al. (2012), are compelling and have a believable and intuitive story behind them. However, it is doubtful that they explain the relationship between housing and the business cycle entirely. Particularly, the lead-lag relationship goes back as far as reliable data does. For instance, Leamer (2007) showed that the Great Depression was preceded by a crash in residential investment in the late 1920s. Fixed-rate mortgages, however, were not developed until after the Federal Housing Administration (FHA) was created as part of the National Housing Act in 1934. So this suggests that the relationship could be driven by other factors. Furthermore, fixed-rate mortgage law does not vary across the country, so it is unclear how they could cause the relationship to hold in some cities but not in others. Overall, while great strides have been made, both the empirical literature and the RBC literature still have a limited understanding of why it exists in the first place.

3 Data

The 'cities' in the analysis follow the 2009 Metropolitan Statistical Area (MSA) definitions created by the US Census Bureau. For parsimonious reasons, the analysis uses only a subset of cities in the United States. In an effort to keep the sample as representative as possible, the cities used are the top 50 cities in terms of population (not including cities in the non-contiguous United States), plus a group of 28 smaller cities picked to ensure adequate geographic coverage. These are listed in figure **??**, where we see the cities account for quite a large amount of the United States in terms of geographic area. They are similarly representative in terms of population, and GDP. Specifically, as of the 2010 Decennial Census, the total population of the sample is 182,012,755 thus accounting for around 3 quarters of the total 249,157,649 people living in urbanized areas in 2010. It is similarly representative in terms of GDP, accounting for \$9,503,012 million of total \$13,459,787 million in 2010 [1]. The time period covered runs from

[1]These numbers are according to the Bureau of Economic Analysis (BEA).

the first quarter of 1990 to final quarter of 2015 for a total of 100 time periods in each city. Subsets of the data can certainly be taken, but the period would have to be at least 9 years to ensure that enough degrees of freedom are available to estimate all the parameters in the model. Likewise, the data could also be extended back before 1990, but this would mean fewer cities could be included in the model.

I use the count of new building permits issued in each city to measure housing market activity. This is preferred to house prices since there are mixed results in the literature concerning the connection between house prices and GDP (Iacoviello (2005)). Leamer (2007) reckons this is because new housing construction exhibits a persistent cycle, rather than the price of housing because prices are sticky downwards. So building permits are a better measure than house prices for this study since they measure quantity.

Of course, it would be best to use residential investment itself, but the only available data on this is at the national level. However, housing permit data is made available at the metro level every month in the Census Bureaus' building permits survey. Each building permit represents a potential construction, so it carries a close relationship with residential investment. Indeed, Ghent and Owyang (2010) finds that the correlation between the band-passed nationwide time series for building permits and residential investment is over 98 percent. Another alternative would be to use the value of building permits issued instead of the count. I did experiment with this and it did not notably alter any results, so I choose to only present results from the count of building permits to save space, and because it has an easier interpretation since there is no need to adjust for inflation.

To represent the broader economy I use metro level employment data from the Bureau of Labor Statistics 'Smoothed Seasonally Adjusted Metropolitan Area Estimates' that counts all nonfarm employees in an MSA. Employment is preferred over metro level GDP, since the latter is only available annually, and not monthly. Nationwide employment carries a 90 percent correlation with GDP, so it reasonable to use as an indicator of the business cycle.

In order to control the greater macroeconomic environment, several nationwide variables that may affect both housing and employment are also included. For national monetary policy, the federal funds rate is included. Also, the national level core consumer price index (CPI), the

30-year conventional mortgage rate and, to measure credit turmoil, the spread between 3-month commercial paper and 90 day Treasury bills. I also include the West Texas Intermediate (WTI) price of Oil per barrel. It is important to take account of these factors, as evidence has been found to suggest that national level variables may be a better indicator of metro level economic changes than the metro level changes themselves Ghent and Owyang (2010).

The last metro level variable is the physical distance between two each city in the sample. This is for creating weighting matrices in the GVARS that are used to compute the weakly-exogenous outside city variables. I compute the physical distance between each city based on the latitude and longitude coordinates provided in the Census Bureau's 'Gazetter' files. I measure the distance between the two cities as the orthodromic or 'great circle' distance between two cities. I then take the inverse of this so that cities closer to one another have a stronger interdependence. These are then normalized so that sum of the inverse distance between any one city and all the others in the sample is equal to unity. This is essentially the same as when a distance matrix is used in spatial econometrics. A small caveat is that the coordinates used for the cities are based on the center of 'urbanized areas' rather than the strict CBSA definitions, so there will be a small amount of measurement error since these locations do not perfectly overlap with one another.

For robustness, I along performed the analysis with two other weighting variables; industrial distance and migration flows. This did not return any notably different results, so it is not presented here.

In order to explain the variation in results from city to city a number of other metro level variables are included in the analysis, but used outside of the main GVAR. These include housing supply elasticity from Saiz (2010), public loan to income ratios from Mian et al. (2013) that combines information from the Federal Reserve Bank of New York Consumer Credit Panel with IRS data on incomes, various demographic and socioeconomic variables found in the Census and American Community Survey, house prices from the Federal Housing Finance Agencies Metro Level House Price Indices, and information on the stringency of local housing regulation from the 'Wharton Residential Land Use Regulation Index'.

4 Motivating Facts

As mentioned, the literature shows a strong connection between residential investment and the other elements of the business cycle. This is shown in figure 3 where we see the lead-lag relationship from the last quarter of 1997 to last quarter of 2012. This figure is, to the best of my knowledge, the most up to date documentation of the relationship in the literature, as it shows that the relationship held throughout the 2007 crash, and this has not been documented formally yet. To isolate the cyclical component of the data, all 3 series are passed through a Baxter-King (BK) filter in the frequency domain. The low-pass BK filter is preferred over the more common Hodrick-Prescott (HP) filter because it is designed to remove the low-frequency variation in the data, rather than the high-frequency variation that characterizes booms and busts. Looking at the 3 series we clearly see that, though they all appear to have the same cycle, there a clear temporal displacement between residential investment, real GDP and non-residential investment. Residential investment clearly leads real GDP which in turn leads non-residential investment.

It is particularly interesting to see that the relationship appears to fit the data best for the most recent recession. This could be considered surprising given the great downturns in house prices during the last crash, as it had been suggested that the reason that housing leads the business cycle being that downward stickiness in house prices lead to a very extreme volume cycle. However, even with little downward friction in house prices, the latest recession was preceded by a downturn in residential investment, as was the recovery. This limits the appeal of this explanation of the phenomenon.

In addition to the casual observation in figure 3, more rigorous investigation for this can be performed by testing for granger causality. To do this I estimate a VAR(2) with all 3 series as endogenous variables, where the lag order 2 was chosen as according to the Alkaline Information Criterion (AIC). As suggested by Geweke et al. (1983), I use a simple F-Type test. The test statistic on the procedure is 7.2263, so the null that resident investment does not granger cause GDP is rejected[2].

[2] Green (1997) rigorously shows that the granger causal relationships between the 3 series in a longer data series, from 1957 to 1992.

It is important at this point to note that both the housing cycle and the business cycle have substantial regional variation. This is displayed in figures 4 and 5 that show time series for, respectively, house prices, building permits, unemployment and real GDP in the 11 largest MSAs in the US in terms of population[3] as well as the US as a whole. All series have been filtered with a BK band pass in the frequency domain so as to more easily detect the cycle rather than the trend. The great dispersion between the extent and timing of the cycles is clear. Though some appear to move in synchronous relationships (for example, Miami and Los Angeles in house prices) , others appear to have very little in common (for example, New York and San Francisco in permits).

This regional variation is important when looking at the lead-lag relationship While the evidence for the lead-lag relationship at the national level is clear, the evidence is weaker at the metro level. Figures 6 to 9 display the series for housing permits in place of residential investment[4] and metro level GDP for the Los Angeles, Dallas, New York and San Francisco Metropolitan Statistical Areas (MSA) [5]. In these figures, there appears to be great inconsistency both between cities and within the same cities over time in how well the relationship holds. In Dallas, permits appear to have been a strong indicator of the GDP throughout the entire period displayed, mirroring the national trend perfectly. The same could be said of San Francisco. However, in Los Angeles, the trend only seems to hold after the year 2000. In the 1990s, however, permits went through several upturns and downturns while employment moved smoothly upwards. New York shows no real evidence of the trend, except for in the 2007 crash.

A more thorough investigation into this was conducted in Ghent and Owyang (2010) which found that the lead-lag relationship held in 80 percent of their the MSAs in their sample, but also found that national level permits were, in fact, a better predictor of employment than a cities own permits. More rigorous investigation of this trend will be performed in this paper in the first section of the results, by assessing the response a cities employment to a shock in its building permits. However, the more information reasoning of these figures can motivate

[3]Population as of 2010 via the census bureau. See **??** for details.

[4]It is reasonable to swap out residential investment for building permits. The correlation between the two at a national level is over 98 percent.

[5]I use the 1999 OMB definitions of Metropolitan Statistical Areas. Details on these can be found at `http://www.census.gov/population/metro/data/pastmetro.html`

looking at this relationship at the metro level quite well. That it exists nationally, but not, but not in all metros, and even in some metros in some time periods but not others, all suggests that housing in particular set of cities could be driving this relationship.

Another important stylized fact to be noted is the connection between different housing markets. It is easy to establish that housing is the leading indicator of both upturns and downturns in the business cycle and that there is substantial variation in housing markets between different metropolitan areas. However, this variation does not exclude the possibility of co-movement and spillovers between the housing markets in different metros. While the most obvious places to look for these with any particular MSA would be those metros in a close physical proximity, this is not necessarily the where one should end their search. Figures 8-10 show some suggestive evidence of this. In figure 8 we see the time series of building permits in New York, Boston and Philadelphia (all series are passed through a Baxter-King so as to more easily detect the cycle rather than the trend) show all appear to follow an extremely similar pattern. Such a result would seem to suggest an important role for physical proximity in housing markets.

However, not all nearby metro areas move in sync with one another. Often, the movements in a city's housing market can have more in common with a city on the other side of the country than its neighbors. Figure 9 shows time series for house prices of California, San Diego, and San Francisco, despite their physical proximity being roughly similar to the cities in cities in figure 8 there do not appear to share common features in their cycles to the same extent. Figure 10 tells a different story, here we see that house prices in Los Angeles, Miami, Washington D.C. and Las Vegas have all shown remarkably similar patterns in spite of the vast physical distance between the cities. This motivates a model that allows for codependency all cities into a single system, rather than just linking one city to its one or two nearest geographic neighbors.

When we combine this idea that co-movement and spillovers could be quite common with established facts about the lead-lag relationship between the housing and GDP, it is reasonable to think that shocks to specific metros could have large spillovers not only into other cities housing markets but their broader economies as measured by metro level GDP or employment. Particularly, the reason why housing leads the business cycle at the national level, but not necessarily at the metro level could be because cities housing markets are not entirely independent of one another. If this were true, then we would expect to find responses in employment or GDP to shocks in residential investment , not only in that city but in others too.

That a shock to one isolated area of the economy could spread a larger (and not necessarily adjacent) area has important implications for the understanding of the business cycle and whether location specific attempts to stabilize specific housing markets could have stabilizing effects on the economy over a broader area. It is this that motivates next section of the paper that lays out the framework for a GVAR model that assesses these possibilities.

5 Estimating the GVAR

There are two common approaches when modeling housing and the rest of the economy: the VAR model and the structural model. While the structural model is typically taken from some applicable economic theory, the VAR is atheoretical using the lags of dependent variables as regressors. The choice between the two models has to do with the type of question being answered and the whether or not there are appropriate instruments available. While the VAR has the advantage of being more general, it has the disadvantage of its parameters often not having a clear practical interpretation, so they are not useful for testing hypothesis directly. However, since I am interested in the response of a cities housing permits and employment to exogenous shocks, and not the direct interpretation of coefficients a VAR would seem more appropriate in my case. Furthermore, it seems unlikely that one would be able to find valid instruments for a structural model.

Though my preferred strategy would be to model permits and employment as endogenous to one another in a large scale VAR, this would cause serious problems when attempting to answer this question. If one were to estimate an unrestricted VAR(p) model with, say, k endogenous variables covering N cities, the number of unknown parameters will be unfeasibly large, of order p(kN-1) [6]. So using a more traditional panel VAR approach is infeasible due to the often referred to 'curse of dimensionality' creating insurmountable computational limitations.

[6] In my case N=78, k=2, and, though it is at the disclosure of the researcher, p=2 is probably a good guess at what the AIC would choose, given previous results in the literature. So we parameters of order $2(78x2-1) = 144$. The data only covers 100 quarters, so there are not enough degrees of freedom to estimate this.

To deal with this I utilize a GVAR framework which avoids such problems by imposing an assumption of weak exogeneity on the variables not contained within the metro itself. That is, all other cities building permits, and employment, as well as national level variables, are assumed to affect the endogenous city-level variables of a particular metro contemporaneously, as well possibly by the lagged changes of inside and outside city variables. However, these variables are not affected by long run disequilibria in the city itself. Taking some license to say it more simply, this similar to assuming that any single city is 'small' with respect to the rest of the country.

The following subsections present the general details of the GVAR used in the analysis. The first deals with the estimation of each city's own individual VARX model one-by-one, while the second shows how the models are linked together as a system of cities. The remaining subsections consist of particulars of model specification, as well as testing the model's assumptions, and checking the model is robust and well-behaved to alternative specifications.

5.1 Individual VARX Models

I begin by modeling individual VARX models for each city in the sample. To motivate this, suppose we have $N+1$ cities indexed by $i = 0,1,...,N$. In order to approximate the large scale VAR that I would prefer to estimate, every cities employment and building permits are included in every other city's model in two vectors.

Speaking formally, for each city, i, the inside city and outside city variables are modeled in the VARX(2,2) structure:

$$x_{it} = a_{i0} + a_{i1}t + \Phi_{i1}x_{i,t-1} + \Phi_{i2}x_{i,t-2} + \Lambda_{i0}x^*_{it} + \Lambda_{i1}x^*_{i,t-1} + \Lambda_{i2}x^*_{i,t-2} + u_{it} \tag{1}$$

where x_{it} is a 2x1 vector containing city i's building permits and change in employment at time t. That is,$x_{it} = (\text{Permits}_{it}, \Delta\text{Employment}_{it})$. I refer to these as 'inside city' variables. x^*_{it} is similarly a 2x1 vector of employment and permits in every other city that I refer to as 'outside city variables', and u_{it} is a serially uncorrelated and cross-sectionally weakly dependent process. If preferred amount of inside and outside city variables are allowed to vary across cities to control for some city specific factors, though I choose not to do so here. The outside city variables are computed as weighted averages of the corresponding inside city variables of all cities, with the weights been specific for each different city. That is, $x^*_{it} = \sum_{j=0}^{N} w_{ij} x_{jt}$, where w_{ij}, $j = 0, 1, ..., N$ are a set of weights such that $w_{ii} = 0$ and $\sum_{j=0}^{N} w_{ij} = 1$. The weights are analogous to a spatial weighting matrix in a spatial auto regression or spatial error model in that they are determined by the modelers choice meant to convey the importance or connectedness of a given city i to some other city j. I use a physical distance for the main results, but also present robustness checks that weight by industrial distance and migration flows in subsection **??**. While physical distance is obviously time variant, the time varying nature of a cities industrial composition and its migration flows mean that a different weighting matrix which is additionally indexed by time, w_{iit}, is used in those VARXs.

Though the GVAR framework can be applied to either stationary or integrated variables or both, estimation, in this case, is performed taking into account the integration properties of each variables' time series in order to short-run and long-run relations. In this case, it is reasonable to interpret long run relations as cointegrating, perhaps reflecting the overall national trend. To do this, I rewrite equation (1) in VECMX form:

$$\Delta x_{it} = c_{i0} - \alpha_i \beta_i' [z_{i,t-1} - \gamma(t-1)] + \Lambda_{i0} \Delta x^*_{it} + \Gamma_i \Delta z_{i,t-1} + u_{it} \tag{2}$$

where $z_{it} = (x'_{it}, x^{*\prime}_{it})'$, α_i is a k_ixr_i matrix of rank r_i and β_i is a $(k_i + k^*_i)$xr_i matrix of rank r_i. If we partition β_i as $(\beta'_{ix}, \beta^{*\prime}_{ix})'$ conformable to z_{it}, the r_i error-correction terms defined by equation (2) can be written in the form:

$$\beta_i'(z_{it} - \gamma_i t) = \beta'_{ix} x_{it} + \beta^{*\prime}_{it} x^*_{it} - (\beta'_i \gamma_i)t \tag{3}$$

this allows for the possibility of co-integration both within x_{it} and between x_{it} and x^*_{it} and hence I am also allowing for the possibility of co-integration across x_{it} and x_{jt} for $i \neq j$.

Each VECMX model is estimated separately for each city conditional on x^*_{it} treating the vector x^*_{it} as I(1) weakly exogenous with respect to the parameters of the conditional model in equation (2). Specifically, the equation (2) is estimated using a reduced rank regression, taking into account the possibility of cointegration both within x_{it} and between x_{it} and x^*_{it}. Estimating the model this way allows one to obtain the number of co-integrating relations, r_i, the speed of adjustment coefficients, α_i, and the cointegrating vectors, β_i, for each individual city model.

Finally, the system is extended to include not only just inside city variables, x_{it}, and outside city variables, x^*_{it} but also a vector of country-level factors, d_t, that are the same in every city's economy. Here I have included the federal funds rate, oil prices, TED spread, and the 30 year fixed mortgage rate in each time period. Practically, this is done by including these variables as part of a city's outside city variables in every city, and not applying the weighting matrix to them. The weak exogeneity assumption is also applied to d_t.

5.2 Solving the Model across a System of Cities

Despite the estimation of each city's VARX model in equation (2) individually, the GVAR framework requires that the model is solved in terms of our entire system of cities as a whole.

To do this with the same set of VARX(2,2) models described in equation (1):

$$x_{it} = a_{i0} + a_{i1}t + \Phi_{i1}x_{i,t-1} + \Phi_{i2}x_{i,t-2} + \Lambda_{i0}x^*_{it} + \Lambda_{i1}x^*_{i,t-1} + \Lambda_{i2}x^*_{i,t-2} + u_{it}$$

As before, define z_{it} as the vector $(x'_{it}, x^{*\prime}_{it})'$ and rewrite (1) for each city as:

$$A_{i0}z_{it} = a_{i0} + a_{i1}t + A_{i1}z_{i,t-1} + A_{i2}z_{i,t-2} + u_{it} \tag{4}$$

where

$$A_{i0} = (I_{k_i}, -\Lambda_{i0})$$
$$A_{i1} = (\Phi_{i1}, \Lambda_{i1})$$
$$A_{i2} = (\Phi_{i2}, \Lambda_{i2})$$

The next is to use the matrices linking the cities, W_i, defined by the aforementioned weights w_{ij} to obtain the identity:

$$z_{it} = W_i x_t \tag{5}$$

where $x_t = (x'_{0t}, x'_{1t}, ..., x'_{Nt})'$ is a $kx1$ vector that collects all the endogenous variables in the system and hence W_i is a $(k_i + k_i^*)xk$ matrix.

From the identity in (5) it is straightforward to show:

$$A_{i0}W_i x_t = a_{i0} + a_{i1}t + A_{i1}W_i x_{t-1} + A_{i2}W_i x_{t-2} + u_{it} \tag{6}$$

stacking each of these individual models yields the model for x_t:

$$G_0 x_t = a_0 + a_1 t + G_1 x_{t-1} + G_2 x_{t-2} + u_t \tag{7}$$

where

$$G_0 = \begin{pmatrix} A_{00}W_0 \\ A_{10}W_1 \\ \vdots \\ A_{N0}W_N \end{pmatrix}, G_1 = \begin{pmatrix} A_{01}W_0 \\ A_{11}W_1 \\ \vdots \\ A_{N1}W_N \end{pmatrix}, G_2 = \begin{pmatrix} A_{02}W_0 \\ A_{12}W_1 \\ \vdots \\ A_{N2}W_N \end{pmatrix}, a_0 = \begin{pmatrix} a_{00} \\ a_{10} \\ \vdots \\ a_{N0} \end{pmatrix}, a_1 = \begin{pmatrix} a_{01} \\ a_{11} \\ \vdots \\ a_{N1} \end{pmatrix}, u_t = \begin{pmatrix} u_{0t} \\ u_{1t} \\ \vdots \\ u_{Nt} \end{pmatrix}$$

Since G_0 is known and is non-singular we can premultiply equation (7) by G_1^{-1} to yield the GVAR(2) model:

$$x_t = b_0 + b_1 t + F_1 x_{t-1} + F_2 x_{t-2} + \varepsilon_t \tag{8}$$

where

$$F_1 = G_0^{-1}G_1$$

$$F_2 = G_0^{-1}G_2$$

$$b_0 = G_0^{-1}a_0$$

$$b_1 = G_0^{-1}a_1$$

$$\varepsilon_t = G_0^{-1}u_t$$

finally equation (8) can be solved recursively for a number of purposely. I place no restrictions on the covariance matrix $\Sigma_\varepsilon = E(\varepsilon_t \varepsilon_t')$, but this could be done if one chose to do so.

When it is solved the GVAR model can model the interactions within the system through three different channels; the first is the contemporaneous dependence of inside city variables, x_it, on outside city variables, x_it^*, and its lagged values, the second is the dependence of inside city variables, x_{it}, on the national level variables d_t and third (and of the most interest to this study) the dynamic dependence of shocks in city i on the shocks in city j, as described by the between city covariances.

5.3 Specification of City Specific Models

As a first test of the dynamic stability of the system; I test for stationarity of variables in levels, first differences, second differences. This is done using an Augmented Dickey-Fuller (ADF) test, based on univariate autoregressions of each variable with the lag order set to 4. For log-levels, the tests use a regression that includes a linear trend. When performed, I found that permits and employment are I(1) in most cities. Those they do not appear to be I(1) appear to be I(2)[7]. That most of the variables do not appear to be stationary in levels supports the choice to include them in the model in levels.

The next step is the selection of lag orders and assessment of the cointegrating relations in the individual VARX models. These models are, as stated before, estimated in error correction form using a reduced rank regression. The rank of the co-integrating space is computed using the Johansen trace statistic following the procedure set out in Peseran, Shin and Smith (2000) for models that contain weakly exogenous I(1) regressors

The selection of p_i and q_i corresponding to the lag orders of the inside city and outside city variables in each individual cities respective models was made according to Akaike Information Criterion (AIC) with the maximum lag lengths, $p_{\max}$ and $q_{\max}$, set to 2. The most common are lag structure selected is, $(p_i, q_i) = (1,1)$, which is selected for 32 of the VARXs. The next most common structures $(p_i, q_i) = (2,1)$, and $(p_i, q_i) = (1,2)$, which are the number of lags chosen in 28 and 9 cities each. The lag structure $(p_i, q_i) = (2,2)$, was chosen in the remaining 8 cities. The system is quite different between different cities with a handful of the all the different possible

[7]Full documentation of the specifications are not presented in the paper to save space. If of interest, the complete results of all tests for stationarity, the lag orders of each model, and the number of cointegrating relations are available on request.

lag structures being selected. This reflects the strong heterogeneity among metro-level housing markets, that helps justify the use of a GVAR to assess spillovers. Other approaches such as a vector spatial autoregression, a panel VAR, or a reduced form approach are feasible alternatives to the use of a GVAR. However, such methods do not discriminate between different units, and rather report a single impulse response function, or a single coefficient reporting the average effect. By doing so they miss the substantial differences in each individual housing market, and in turn, the large variance in the extent of spillovers depending on both where shock that caused it originated from, and where it is spreading to.

Lastly, I looked at the number of cointegrating relations in each individual VARX. This is crucial since in order for there to be convergence in the impulse response functions that assess the level of spillovers it must be that the system as a whole is stable. This was done by computing the rank of the cointegrating matrix in each separate VARX using the Johansen trace statistic following the procedure set out in Peseran, Shin and Smith (2000) for models that contain weakly exogenous I(1) regressors. Using this it was found that, in total, the GVAR model contains 103 cointegrating relations.

For the GVAR to be well behaved, it must be that the rank of the cointegrating matrix does not exceed the number of cointegrating relations in all the individual VARX city models. The model contains 161 endogenous variables with a maximum lag order of 2 resulting in 322 eigenvalues. So at least 219 (that is, 322-103) eigenvalues fall within the unit circle on the complex plane. It happens that exactly 265 of the eigenvalues fall within the unit circle and the remainders have moduli less than unity. This indicates that the model is stable and that there do exist some shocks will have permanent effects on the endogenous variables.

5.4 Testing for Weak Exogeneity

The GVAR model assumes that the outside city variables, x_{it}^*, as well as the national level variables, d_t, are weakly exogenous with respect to the long-run parameters of the conditional model. Since this assumption is crucial for the model to be well behaved, a brief tangent is taken here to familiarize readers with the concept of weak exogeneity, and how it differs from other types of exogeneity.

The idea that the independent variables within a system must be regarded as 'exogenous' to the dependent variable for efficient estimation is at the heart of most econometric work. Practitioners have long had an understanding that variables being exogenous depend on whether or not it can be taken as 'given', or 'fixed' without losing any information relevant to the task at hand. However, more formal definitions of the concept did not exist until the seminal work of Engle et al. (1983), which differentiated between 3 types of exogeneity; weak exogeneity, strong exogeneity, and super exogeneity. Weak exogeneity is, predictably, the less strict of the 3, but also gives the econometrician less flexibility than the other 2. Speaking informally, a parameter can be regarded as weakly exogenous if knowledge of its value provides no additional information about the potential *range* of values the parameters of interest. So if weak exogeneity of a parameter holds, then the marginal distribution of that parameter can be ignored, and estimation will still be efficient. Strong exogeneity implies that a parameter is both weakly exogenous and does not granger cause the parameters of interest. This allows not only efficient estimation but also allows one to make valid n-step ahead forecasts. Super exogeneity is the strictest of the 3 definitions and probably the closest to the intuitive definition of exogeneity as something being 'outside' of the system. A parameter can be regarded as super exogenous to the parameters of interest if it is both weakly exogenous, and the parameters of interest are invariant with respect to the super exogenous parameter. This is much like weak exogeneity, only now addition information about the super exogenous parameter gives us no additional information about the parameter of interest itself, rather than just its range of values

In the context of this study, weak exogeneity is synonymous with 'long-run forcing', meaning that no long-run feedback from x_{it} to x^*_{it} is allowed. Short-run feedback, however, is allowed. Since my model is cointegrating, this means that the error correcting terms, $ECM_{ij,t-1}$, of the individual city VECMs do not enter into the marginal models of the weighted outside city variables, x^*_{it}. This can be formally tested. To do so, I carry out the procedure outlined in Johansen (1992) and perform a joint test of the significance of the estimated error correction terms in auxiliary equations for the outside city variables. Specifically, for the lth element of

x_{it}^* the following regression is run:

$$\Delta x_{it,l}^* = a_{i,l} + \sum_{j=1}^{r_i} \delta_{ij,l} ECM_{ij,t-1} + \sum_{k=1}^{s_i} \phi_{ik,l}' \Delta x_{i,t-k} + \sum_{m=1}^{n_i} \psi_{im,l}' \Delta \tilde{x}_{i,t-m}' + \eta_{it,l} \tag{9}$$

where $ECM_{ij,t-1},\ j = 1,2,...,r_i$ are the estimated error correction terms corresponding to the r_i co-integrating relations found for the ith country model and $\Delta\tilde{x}_{it} = (\Delta x_{it}'^*, \Delta(e_{it}^* - p_{it}^*), \Delta p_t^0)'$. The weak exogeneity test is an F-test with the (joint) null hypothesis; $\delta_{ij,l} = 0, j = 1,2,...,r_i$ in equation (9).

I conducted the test at the 5% level, in general, the variables of interest (i.e. permits, and employment), were to be well behaved, and only fail the test in 23 out of 156 cases [8]. In particular, only employment in Pensacola, permits in Louisville, and permits in Las Cruces cannot be said to be weakly exogenous to the weighted outside city variables, x_{it}^*, for each of their individual models. Since they are so few, and since each test was carried out independently of the others, I do not regard these rejections as particularly concerning.

The national variables display had mixed results, while the federal funds rate and the 30-year fixed mortgage rate appear to be well behaved, and only produce 1 rejection between the 2 variables (i.e. the 30-year fixed mortgage rate in Dallas). Results for the TED spread and the price of crude oil, however, are more concerning for the behavior of the model. The null is not rejected for the TED spread in 6 cities and is also not rejected for oil in 8 cities. These many failures indicate cast more doubt on whether or not the model will be well behaved, or if we can consider it reliable.

6 Results

The results show the generalized impulse response functions (GIRFs) for a one standard error shock to permits in each city. GIRFS are used instead of the traditional orthogonalized impulse response functions (OIRFs) in Sims (1980) because they do not require any ordering of the variables. The responses are cumulative and in levels for ease of interpretation.

[8] As with the model specification, I only summarize the weak exogeneity results in words to save space. Full documentation of the test and results is available on request.

6.1 Responses inside Cities.

Before assessing the evidence for spillovers, it is worth looking at the GIRFs within the city that is shocked itself. This will help tell us the extent to which the lead-lag relationship varies between cities, and use this regional variation to help explain why it exists in the first place.

The model allows me to pull out 4 different inside city responses; the response to permits from a 1 S.E shock to permits, the response of permits from a 1 S.E shock to employment, the response of employment from 1 S.E shock to employment, and, the response that is of the most interest, that of employment from a 1 S.E shock to permits. The dynamic profiles of this shock in all 78 cities are shown in four figures 13.

The first thing to note from figure 13 is that the responses all seem to display some sporadic behavior in the first few quarters, but generally appear to have converged to a stable value by 10 quarters. This means that the responses can be aptly summarized by looking a the point estimate of the response sometime after the 10th quarter. Thus, I present summary statistics for the responses after 20 quarters in 1. The point responses are in the top row, and in the second row, I show a dummy variable that takes the value 1 if the response is positive. The bottom row is equivalent to a dummy variable that takes the value 1 if housing leads the business cycle in that city.

When looking at table 1 in conjunction with figure 13 we can see that, consistent with previous results in the literature, the evidence that residential investment leads the business cycle at the metro level appears to be mixed. Shocks to permits only elicit a positive response in only 47 of 78 cities (i.e. 60 percent). In the other 31 cities, the response to a permits shock is actually negative, the opposite of what would be predicted by the national trend.

We can also look at the magnitude of responses. Table 1 shows that the mean response from a 1 S.E shock to permits is just under 600 jobs. This should be considered quite a high number, since the average standard error of city-wide permits issued is 238, this translates to an average of 2.5 jobs for every new permit issued.

However, concentrating on the mean alone masks the large variation between cities. The standard deviation in table 1 is 4,272, more than 7 times the size of the mean. This is because of a handful of cities (usually bigger ones), where the shock is quite substantially larger than average. For example, the largest response of 23,000 is found in New York where the standard error is 1277 permits, this normalizes to a response of roughly 18 new jobs for each permit issued. In Chicago, the corresponding numbers are 7,613 new jobs in response to a shock of 300 issued permits, or about 25 new jobs in response to each permit issued. By far the largest negative response is in Los Angeles, where innovation of 1005 new permits leads a response of about -13,231 job losses, or around 13 jobs lost for every new permit issued.

Figure 13, and table 1 effectively show that the lead-lag relationship between housing and the business cycle breaks down at lower geographic levels. However, neither is able to explain why this relationship holds in some cities but not others. Understanding this is important because in order to implement any kind of place based housing policies to stabilize the business cycle, the transmission mechanism between housing, and the business cycle must also be understood.

I first look at if the cities where housing leads the business cycle are concentrated in a particular region. Figure 14 shows the cities where permits lead employment in blue, and those where it does not in red (i.e. the variable in row 2 of table 1). From this we can see that Midwestern and Southern cities show are less likely to have housing lead the business cycle. Atlanta, St. Louis, Kansas City, and Cleveland among others all exhibit negative reactions in their employment. The cities that have large positive responses appear to mostly be located in the Northeast, Texas, and Northwest with New York, Boston, Dallas, Houston, and Seattle all exhibiting a positive response after 20 quarters. The geographic trend is not conclusive, however, as it is clear that even some adjacent cities (e.g. San Francisco and San Jose or Los Angeles and Riverside) respond in different directions.

Since there is no overwhelming geographic trend, the next step is to look if there are notable differences in the characteristics of the red and blue cities in figure 14. This can be done by regressing both the variables in table 1 on different city level variables.

The first set variables I use are the proportions of employees in a city who work in a certain industry. [9] It could be that cities were employees are more heavily concentrated in real estate

[9] My 'Industries' are 2012 2-digit NAICS codes downloaded from the Census Bureau's 'County Business Pattern' series. I use information from 2001 in my regressions, but the resuts are insensitive to year choice. More

related industries like construction, finance, or renting and leasing services are more likely to have housing lead their business cycles. Conversely, places that are heavily concentrated in other industries like oil, could be more likely to have any shocks to housing are 'washed out' by movements in those industries. This has some precedent at the national level, in 1951 and 1967 there were severe housing downturns that were, unusually, not followed by recessions. Leamer (2007) attributes this to huge fiscal spending on to the Korean and Vietnam War, respectively.

Leamer (2007) reckons that housing leads the business cycle because of the stickiness of house prices. If people refuse to sell their houses when prices fall, this means that shocks to housing will manifest themselves primarily through changes sales volumes. This means that shocks to housing will lead to less jobs in construction, brokerages, finance, etc. Strauss (2013) gives evidence for a similar story for housing leading the business cycle during booms, showing that expectations about the economy and the issuance of new building permits are highly correlated. He infers that when people have expect a boom, they buy houses/start renovations, which is shown in the many new permits being issued, however, it is not until sometime later that the permits is actually constructed and real economy activity takes place. While these reasons are plausible they cannot explain why the relationship is inconsistent from city to city.

Some of Leamer and Struass's stories may be captured in the aforementioned industrial compositions of cities. However, there may be another reason behind why permits lead employment in so areas but not others- variation in the chances that a permit will actually be built. A permit only represents potential construction, and it is common that it will never actually be followed through with, and thus not result in any response in employment. I use four variables that may create variation in this across; the average debt to income ratio in the city, the elasticity of its housing supply, how stringent its land use regulations are, and the price of housing.

The idea behind using debt to income ratios is that they can cause the marginal propensity to consume to differ from place to place. For instance, if a person expects a rise in their income, they may request a permit for some anticipated building that they intend to spend it on. However, they are a lot less likely to go through with this building if they have a large amount of debt they need to pay off. Alternatively, they may stretch out the construction over

information on the series can be found at .

a longer period of time (i.e. hire 1 workers for a 2 year project instead of 2 workers for a one year project) or even use their own labor. This is similar to the mechanisms displayed in papers such as Eggertsson and Krugman (2012), and Di Maggio et al. (2014), where the debt is often attributing as being the reason for a slow recovery, because people tend to deleverage before they spend more money on durable goods, dampening the effects of expansionary monetary policy [10].

My measures for elasticity of housing supply and stringency of regulation are the estimates found in Saiz (2010) and the Wharton Residential Land Use Regulation Index(Gyourko et al. (2008)), respectively. These are included because in places where the housing supply is less elastic, and regulation is stricter, it is less likely that a person will go through an approved permit because the costs of building and construction are higher. House Prices are also included for the same reason.

Results for the regression are shown in table 2. I run each regression for 2 different specifications, the first is a simple OLS regression where the point estimate of the response in the top row of table 1, and the second is a probit with the dummy variable in the bottom row of the same table. The regression are run for each type of variable separately in the first ten columns, then all at the same time in the final two columns.

The regressions produce ambiguous results for industries, house prices, regulation, and the elasticity of housing supply. The coefficients are almost always insignificant for these, and often change sign from specification to specification. Debt to income ratios (the top row of 2) appear to be the most consistent predictor of a whether or not housing leads the business cycle in a city. The OLS results indicate that an increase in the average ratio of a city by 1 will lower the response to a shock to permits by 4,643 jobs, or in other terms 5 less jobs per permit issued (see columns (1) and (2)). This is remains when I control for all the alternative variables of interest mentioned above (columns (7) and (8)) if a probit specification is used, but not OLS. The probit results indicate that raising the ratio by 1 will make a city more than 70 percent less likely to have housing lead the business cycle.

As mentioned, if a city has more debt relative to income, this does not do a great deal to

[10] Data on DTI ratios can be found at . I use the 2001 ratio in my regressions, but the results can be replicated using any of the years.

hinder the pursuit of building permit, since just having a permits issued does not cost much. However, it does hinder people from actually going through with having the permit built, since they are burdened with more debt, or have fewer means to pay off their debt with. If a permit is issued but not built, no actual economic activity will take place. Essentially, the argument is that a higher DTI does not dampen the propensity of a person to pursue a permit, but does inhibit their ability to build it, since they must also use their income to pay off debt.

There are reasons to doubt this story though, it is possible that DTI ratios could be picking up two-way causality. For instance, higher DTI ratios could reflect greater access to credit, which should give people greater means to be able to fulfill a permit–the opposite of the reasoning above. However, efforts to subject the results to greater scrutiny by introducing instruments have done little more than introduce a great deal of noise to estimation. Given that it is only significant in 3 of the 4 original specifications to begin with, the results remain inconclusive. They do suggest some interesting hypothesis though that will be explored more depth in future versions of the paper.

6.2 Spillovers

I next move on to assessing the responses in employment to an innovation in permits in all the cities besides the one where the shock originated. Tables 3 and 4 show the mean and standard deviation for each city's spillovers to the other 77 cities in the sample. As with the within city responses, the results in this table are calculated using the estimate of the point response in employment after 20. In most cities the average spillover is relatively small, but with a very large standard deviation. Indicating that the responses at any given time period follow negative binomial distribution with low means, where there a lot of relative small and inconsequential spillovers, and a few very large ones to some particular cities.

A few illustrative examples of the full distribution of spillovers are shown in figure 15. We see that, consistent with what might be suspected from the means and standard deviations, most spillovers are fairly small, indicated by the high peaks of distributions close to zero. The mass of all three cities distributions are very close to zero, but are skewed (right for New York and Houston, left for Los Angeles), indicating there are a few cities that the spillovers to are very large.

The full dynamic response of employment to an innovation in building to permits is shown for Chicago, Houston, Los Angeles, and New York in figure 16. From this we can see that the responses of cities spillovers generally have the same direction as the inside city spillovers. That is, if permits lead employment positively (negatively) within a city, they generally also lead positively (negatively) outside of the city. This could be useful for reconciling the mismatch between the seemingly iron-clad status of the lead-lag relationship at the national level, and its more murky status at the metro level. It could be that just a handful of cities are responsible for this relationship. For example, New York has a within city response of around 20,000 jobs to an innovation in employment. This is a large number, but looking at it alone ignores that the response is all other cities sums to around 28,000 jobs. So the response to permits in the whole system of cities is actually 48,000 jobs. In Houston the corresponding numbers are 10,000 and 70,000, and in Chicago they are 8,000, and 30,000. This will be looked at further in the next subsection.

I next look at what determines the size of a spillover from one city to another. To do this I first map the spillovers (after 20 quarters) for the same cities as in figures **??** and 16. This can be seen in figure 17 where the color represents the value the spillover (red for more positive, blue for more negative), and the size of each point is the absolute value of the spillover.

The most striking finding in figure 17 is that the spillovers do not appear to be to the closest neighbor of any of the cities. Take Houston for instance; the spillovers are for New York, Chicago, Detroit, and Los Angeles are all larger than Dallas, its closet large neighbor. Similar statements can be made about all the other cities. That the same set of cities receives the strongest spillovers from most cities indicates that they act as 'hubs' of sorts in economic activity, that are strongly connected with all other parts of the country.

To characterize the connections that determine spillovers between I regress the point estimates after 20 quarters (those summarized in tables 3 and 4) on three variables; the physical distance between the two cities, the migration flow between the two cities in both directions, and the similarity of the cities industrial compositions [11]. The results for this regression are

[11] Industrial similarities are measured by the taking the average squared differences between each cities employment share in 2-digit NAICS sectors.

shown in table 5. The first column uses the spillover as the dependent variable, while the second uses the absolute value of the spillover.

The results in table 5 show that the similarities between cities do not matter much for determining the direction of the spillover. This is seen in column 1, where all the coefficients are insignificant. However, different types of distance do matter when we look at the magnitude of a spillover, and ignore its direction. Physical distance has a counter-intuitive negative effect, the coefficient in column indicates that for every kilometer further apart two cities are, a shock to permits in the original city will result in 0.2 more jobs, on average. So for two cities 1000 kilometers apart (roughly the distance between L.A and Salt Lake City), every new 5 permits in each city will lead to 1 new job in the other. The migration coefficient indicates that if the city the spillover originated in sends 1,000 to the other city a year, then there we would expect to see 1 new job in the other city for every 2 new permits issued in the originator. The industrial distance coefficient is harder to interpret, but it roughly means that two cities with entirely different industrial compositions would have a spillover 800 jobs smaller than they would if their industries were identical.

Looking only at distance measures between two cities ignores the idea that some cities might be 'hubs' that always experience a spillover no matter where the originates. To look at this, I add the industrial composition, DTI ratio, average income, average house price, WRLURI score, housing supply elasticity, and population for both cities to the dependent variables in the regressions found in table 5. The results of these regressions are shown in table 6.

The first thing to note in the results is that though physical and industrial distance more or less maintain their effects on spillovers, migration connections are no longer significant. So the spillovers that occur are probably not due to people moving to another city in response to the shock to permits, taking jobs with them when they move.

Second, the characteristics of the recipient city seem to be the most important features for determining the magnitude of the spillover, but not the direction. Places where a higher proportion of people are employed in construction, manufacturing, and service based industries, tend to be larger recipients of larger spillovers. Bigger cities, also, will receive bigger spillovers. House prices, housing supply elasticity, and regulation are not significant. Higher DTI ratios predict a smaller spillover.

At first glance the importance of a recipient city's industries seems plausible. When a house is built, this will not only lead to increases in employment where the house is built, but also in the places that materials and services used when building are produced. A house built in Columbia, MO will create some jobs for local contractors, but it is also most likely financed with a loan from a bank headquartered in a city like New York that was then securitized in Washington D.C, perhaps used concrete produced in Houston, etc. Th results indicate that this, 'all roads lead to Rome' story, is more likely than the common stories of diffusion found in studies that look at house price spillovers such as DeFusco et al. (2013).

An important cavaet on this interpretation, however, is that table 6 shows negative coefficients for cities where more people ar employed in finance or real estate meaning that they actually receive smaller spillovers. Why such results exist is not totally clear, and indicates the need of a more nuanced approach to determining reasons behind spillovers in future versions of this paper.

Third, the characteristics of the origin city do not do much to determine the size of the spillover, but do have a detectable effect on its direction. These results are, for the most part, the same as those found in table 2. Suggest that the direction of spillover is, like the within city response itself, largely a function of the propensity of a permit to actually be built.

Overall, the most interesting result here is that some cities may act as 'hubs' of sorts in economic activity, that are strongly connected with all other parts of the country. Furthermore, that the strongest spillovers are in both adjacent cities and very distant ones indicates that things such as industrial similarities, migration, and common tastes and preferences could are also potential conduits that connect cities and transmit spillovers. This indicates that activities such as monetary policy that stimulate both the economy and the housing market, could have widely different effects across space. More testing is required to determine what specifically determines whether is a hub or more isolated. The results at the moment suggest that the most important characteristics are population and industrial composition.

6.3 Is the National relationship Driven by a Subset of Cities?

That housing leads the business cycle consistently at the national level, but inconsistently at the metro level is somewhat puzzling. If different metro housing markets were truly independent and the relationship was causal, this should not be possible. This leads us with two possibilities; the first is that in the markets where the relationship holds are large enough so that the trend shows up at a national level too. This would imply that metros are independent, and instead that there is some third initial condition present in those cities that interacts with residential investment causing a response in employment. The second is that metros are not independent, and instead that when residential investment rises in one place, the response in employment is found in other metros. This does not imply that the relationship is causal, but does not exclude it either.

Given previous results in the literature (Ghent and Owyang (2010)) have showed that national residential investment leads employment even in those cities where local residential investment does not, and the spillovers I have found, the second of those scenarios seems more likely. To look more closely at this, I define the cities where the relationship was found to hold in the above subsection as a particular region, and those cities where the relationship does not hold as another region and rerun the GVAR as detailed in section 5. I then look at the impulse response functions for a shock to each region on all employment in the cities in the sample. If it is true that the relationship is driven by a particular few cities, then the cities where the relationship did hold should elicit a positive response, while the cities that did not should elicit a negative response, or none at all. This is shown in figure 18, note that, as before, the response here are cumulative. As expected, the cities where the relationship held have a impulse response function that is positive, and large. It shows a marginally decreasing response, similar to the pattern seem in the individual cities spillovers, that settles into a permanent effect at around 20 quarters. The point estimate of this is between 2 and 2.5 million jobs, however, the 95 percent lower bound extends down to as low as 200,000. So the estimate is not particularly precise.

Compare this to the response of the cities where the relationship did not hold. This is shown by the red lines in figure 18. It is clear that at no point is the response significantly different from zero. This indicates that national lead-lag relationship is probably driven by the cities where the relationship does hold having spillovers, where shocks to their permits lead employment in other cities.

7 Conclusions

The housing market and the rest of the business cycle share a strong connection, with housing being a leading indicator for other components. Given that there is great dispersion between different MSAs housing markets, I have used a GVAR to model this relationship at the metropolitan level and investigated the possibility of spillovers from housing markets to labor markets.

The GVAR shows that, consistent with previous results in the literature, housing does not lead employment in every metro area, but only in a subset. For my sample the relationship is found to hold in 47 of 78 cities. I find that this mismatch can be somewhat reconciled by accounting for spillovers between metro areas. That is, I find that in those metros where the relationship holds there is also a substantial dynamic reaction in other areas employment. Thus, it appears that the lead-lag relationship at the national level is in fact driven by a handful of cities, where shocks to residential investment elicit large dynamic responses in employment across the nation. These responses are even found in cities where the relationship does not hold. For example, Los Angeles, a city where shocks to building permits actually have a negative response in employment, is nearly always the recipient of a positive spillover regardless of the origination city.

The reasons why some cities have housing lead their business cycles but others appear to be loosely related to geography, and to debt to income ratios. The latter indicates that the reason behind the inconsistency across cities is likely due to some places having a higher propensity of a building permits actually being built.

Another interesting result is that the spillovers between cities are not always biggest in adjacent cities. In particular, Los Angeles, New York, and Chicago are among the biggest responders for nearly all cities. This implies both that cities economies are connected by more than just distance, and raises the possibility that things such as industrial similarities, migration flows, demand drivers like amenities, or supply drivers like housing density could be more appropriate ways to connect cities.

There are important policy implications of these results. Since only some cities are responsible for driving the relationship between housing, and the rest of the business cycle this implies that there will be substantial cross regional effects in both national monetary policy, and more localized housing policies and regulations. Furthermore, if the relationship has a casual one, this implies that stabilization of particular housing markets could have result in some stabilization of the larger business cycle.

There work does have several shortcomings that must be acknowledged. Particularly, I am unable provide much insight as to whether or not the relationship is causal. There are reasons to believe that the relationship could be due to the income tax treatment of residential investment, or the regulatory environment of housing finance. For example, when tax law gives residential investment special treatment on things such as accelerated depreciation, or capital gains through property tax limits, it attracts capital to it. Indeed, Kydland et al. (2012) has suggested that the key connection lies in the 30 year fixed mortgage rate. The only insight this work does provide on that front is that whatever this condition is must vary across space. Meaning that the relationship cannot be explained entirely by something that rests in national policy. In this sense, it may be promising to look at cross country differences in the relationship between residential investment, and the business cycle, since this may reveal something that the city level data does not.

However, the reasons to believe that the relationship is only correlation are also very reasonable. It could be that housing is simply on the margin because it is durable, after all Leamer (2007) found that other types of durable investment, such as cars, are also a leading indicator. This again limits the feasibility of policy implications of work in this area, since without establishing a causal connection between the two it is not clear whether policy aimed at will actually be effective or not. Finding appropriate instruments for the regression in table 2 will be an important point to address in future drafts of the paper. Those that I have tried that are theoretically appropriate, such as the elasticity of housing supply, have a weak first stage.

Furthermore, the paper has yet to confirm what characteristics of cities predict that it will receive a large spillover. At this point, the best we can say is that characteristics of the origin city appear to determine the direction of the spillover, and the characteristics of the receiving city determine its magnitude. A more nuanced approach that the simple regressions run in this paper is likely needed to pin down exactly what these characteristics are. this will also be a focus in future versions of the paper.

References

Baker, S. R. (2015). Debt and the consumption response to household income shocks. *Available at SSRN 2541142*.

Cesa-Bianchi, A. (2013). Housing cycles and macroeconomic fluctuations: A global perspective. *Journal of International Money and Finance*, 37(0):215 – 238.

Coulson, N. E. and Kim, M.-S. (2000). Residential investment, non-residential investment and gdp. *Real Estate Economics*, 28(2):233–247.

Davis, M. A. and Heathcote, J. (2007). The price and quantity of residential land in the united states. *Journal of Monetary Economics*, 54(8):2595 – 2620.

DeFusco, A., Ding, W., Ferreira, F., and Gyourko, J. (2013). The role of contagion in the last american housing cycle. Mimeo.

Di Maggio, M., Kermani, A., and Ramcharan, R. (2014). Monetary policy pass-through: Household consumption and voluntary deleveraging. *Columbia Business School Research Paper*, (14-24).

di Mauro, F., Smith, L. V., Dees, S., and Pesaran, M. H. (2007). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics*, 22(1):1–38.

Eggertsson, G. B. and Krugman, P. (2012). Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics*, 127(3):1469–1513.

Engle, R. F., Hendry, D. F., and Richard, J.-F. (1983). Exogeneity. *Econometrica*, 51(2):277–304.

Fisher, J. D. M. (2007). Why does household investment lead business investment over the business cycle? *Journal of Political Economy*, 115(1):pp. 141–168.

Geweke, J., Meese, R., and Dent, W. (1983). Comparing alternative tests of causality in temporal systems : Analytic results and experimental evidence. *Journal of Econometrics*, 21(2):161–194.

Ghent, A. C. and Owyang, M. T. (2010). Is housing the business cycle? evidence from {US} cities. *Journal of Urban Economics*, 67(3):336 – 351.

Green, R. K. (1997). Follow the leader: How changes in residential and non-residential investment predict changes in gdp. *Real Estate Economics*, 25(2):253–270.

Greenaway-McGrevy, R. and Phillips, P. C. (2016). Hot property in new zealand: Empirical evidence of housing bubbles in the metropolitan centres. *New Zealand Economic Papers*, 50(1):88–113.

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, 45(3):693–729.

Head, A., Lloyd-Ellis, H., and Sun, H. (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review*, 104(4):1172–1210.

Hiebert, P. and Vansteenkiste, I. (2009). Do house price developments spill over across euro area countries? Evidence from a Global VAR. Working Paper Series 1026, European Central Bank.

Iacoviello, M. (2005). House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. *American Economic Review*, 95(3):739–764.

Johansen, S. (1992). Testing weak exogeneity and the order of cointegration in UK money demand data. *Journal of Policy Modeling*, 14(3):313–334.

Kydland, F., Rupert, P., and Sustek, R. (2012). Housing Dynamics over the Business Cycle. University of California at Santa Barbara, Economics Working Paper Series qt7bn5k73m, Department of Economics, UC Santa Barbara.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50(6):pp. 1345–1370.

Leamer, E. E. (2007). Housing is the business cycle. Working Paper 13428, National Bureau of Economic Research.

Malpezzi, S. (1999). A Simple Error Correction Model of House Prices. *Journal of Housing Economics*, 8(1):27–62.

Mian, A., Rao, K., and Sufi, A. (2013). Household balance sheets, consumption, and the economic slump*. *The Quarterly Journal of Economics*.

Pesaran, M., Schuermann, T., and Weiner, S. (2004). Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. *Journal of Business & Economic Statistics*, 22:129–162.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Saks, R. E. (2008). Job creation and housing construction: Constraints on metropolitan area employment growth. *Journal of Urban Economics*, 64(1):178 – 195.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1):1–48.

Strauss, J. (2013). Does housing drive state-level job growth? building permits and consumer expectations forecast a states economic activity. *Journal of Urban Economics*, 73(1):77–93.

Van Nieuwerburgh, S., Favilukis, J., Ludvigson, S. C., and Nieuwerburgh, S. V. (2015). The macroeconomic effects of housing wealth, housing finance, and limited risk-sharing in general equilibrium.

Vansteenkiste, I. (2007). Regional housing market spillovers in the US: lessons from regional divergences in a common monetary policy setting. Working Paper Series 0708, European Central Bank.

Zabel, J. E. (2012). Migration, housing market, and labor market responses to employment shocks. *Journal of Urban Economics*, 72(2):267–284.

Appendix A: Figures

Figure 1: Map of Cities Used in Analysis

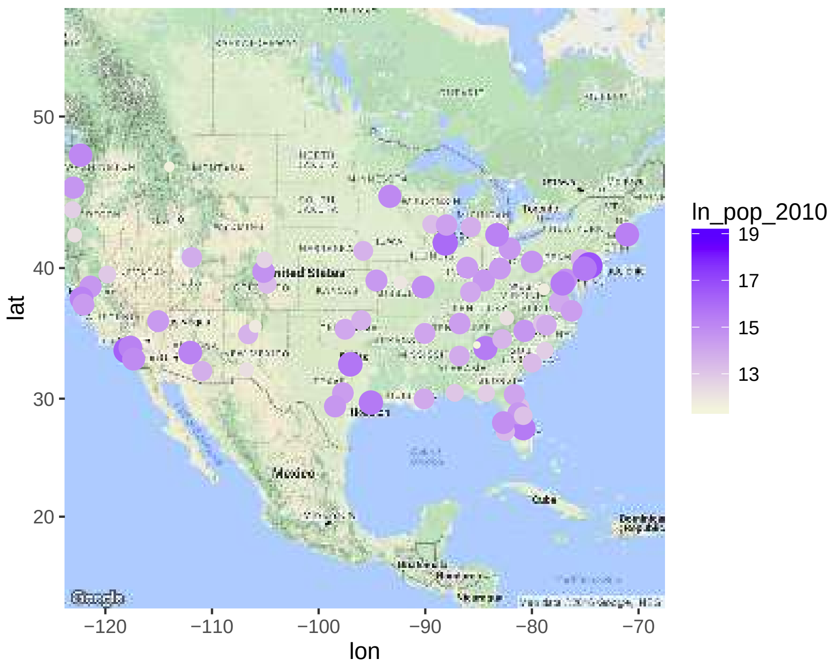


Figure 2: GDP, Residential Investment, and Non-Residential Investment at the National Level, 1997-2012

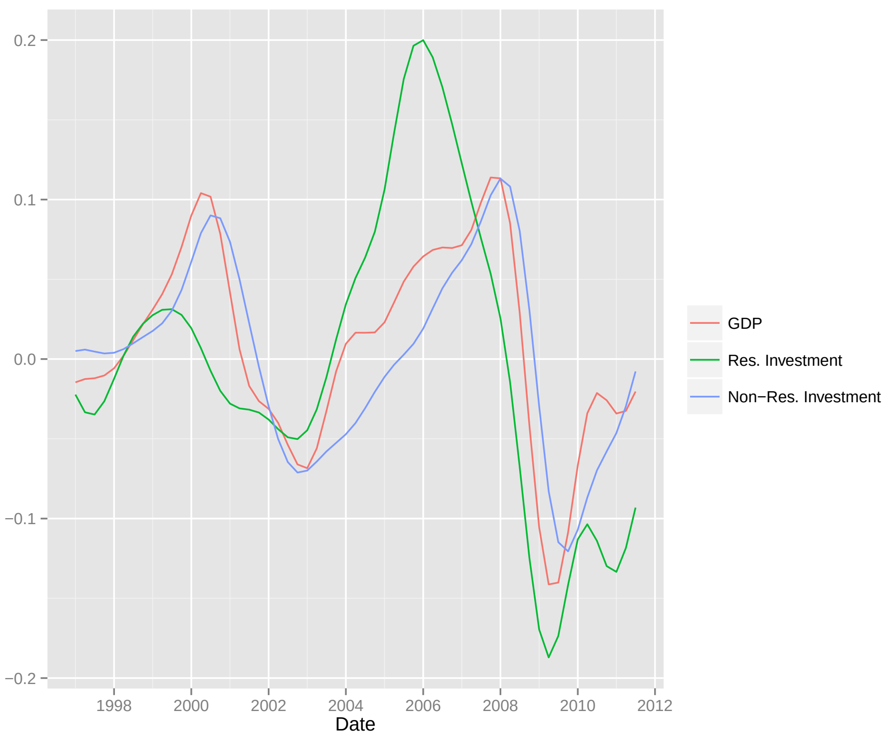


Figure 3: GDP, Residential Investment, and Non-Residential Investment at the National Level, 1997-2012

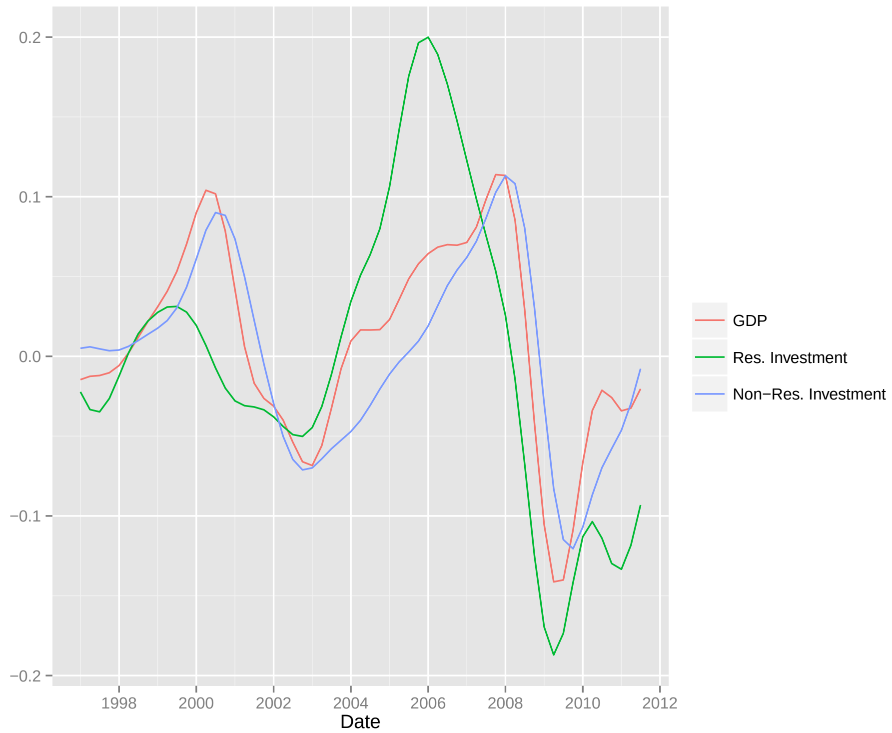


Figure 4: Permits in Eleven Biggest Metros,1988-2015

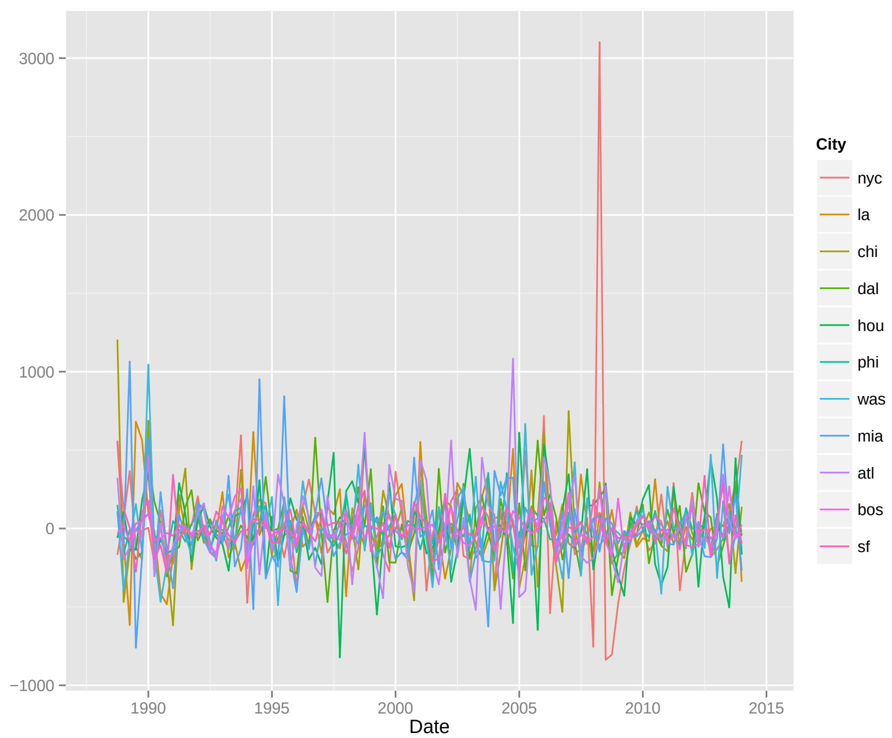


Figure 5: Employment in Eleven Biggest Metros, 1990-2015

Figure 6: Building Permits, and GDP in Los Angeles, 1990-2010

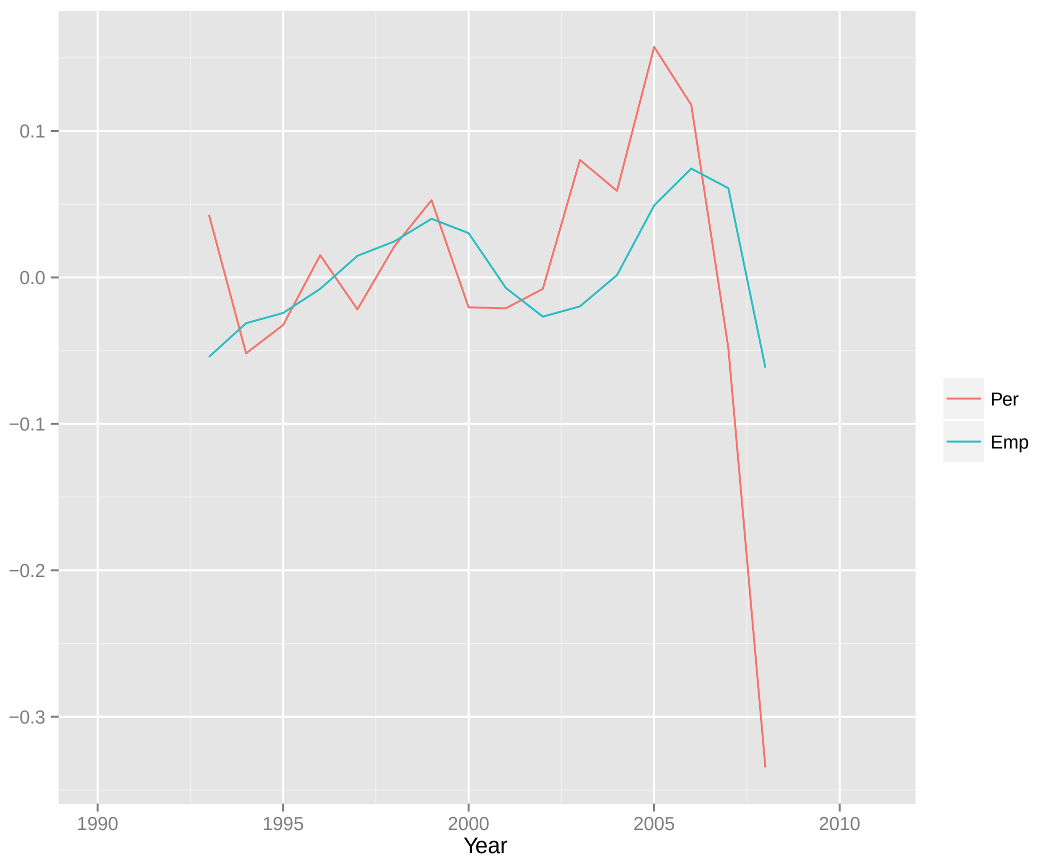


Figure 7: Building Permits, and GDP in New York, 1990-2010

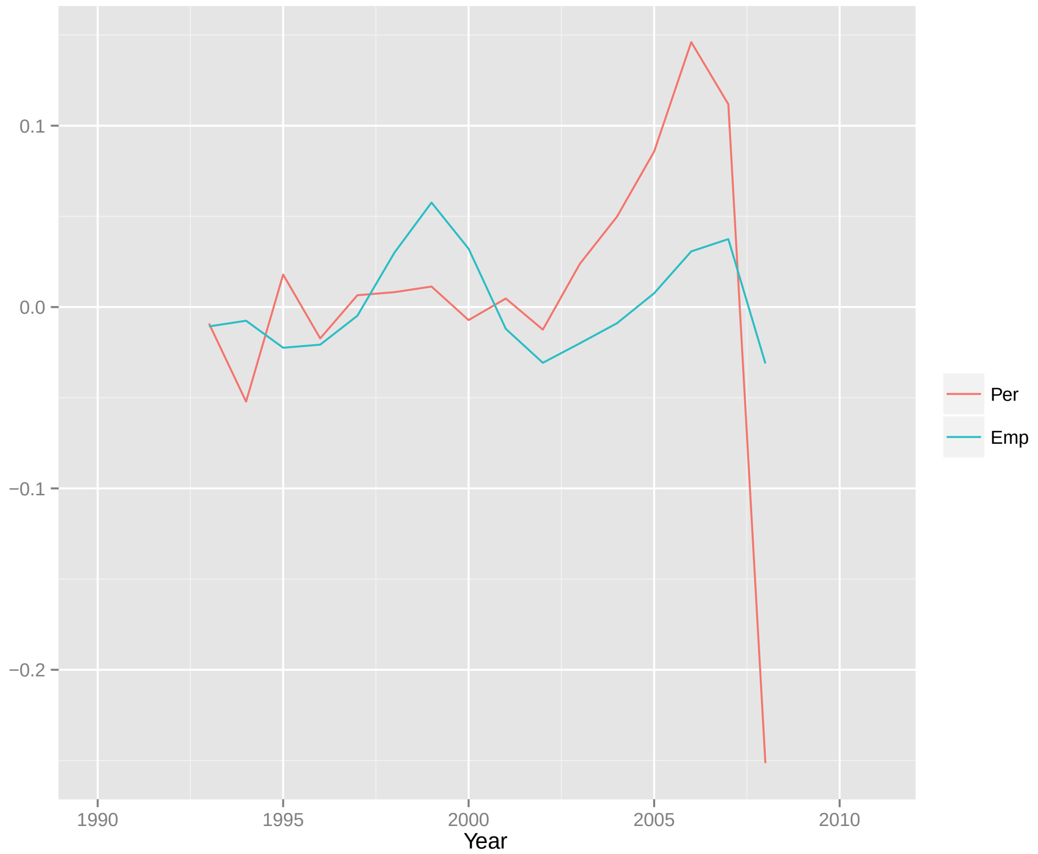


Figure 8: Building Permits, and GDP in San Francisco, 1990-2010

Figure 9: Building Permits, and GDP in Dallas, 1990-2010

Figure 10: Building Permits in Boston, New York City, and Philadelphia, 1994 - 2011

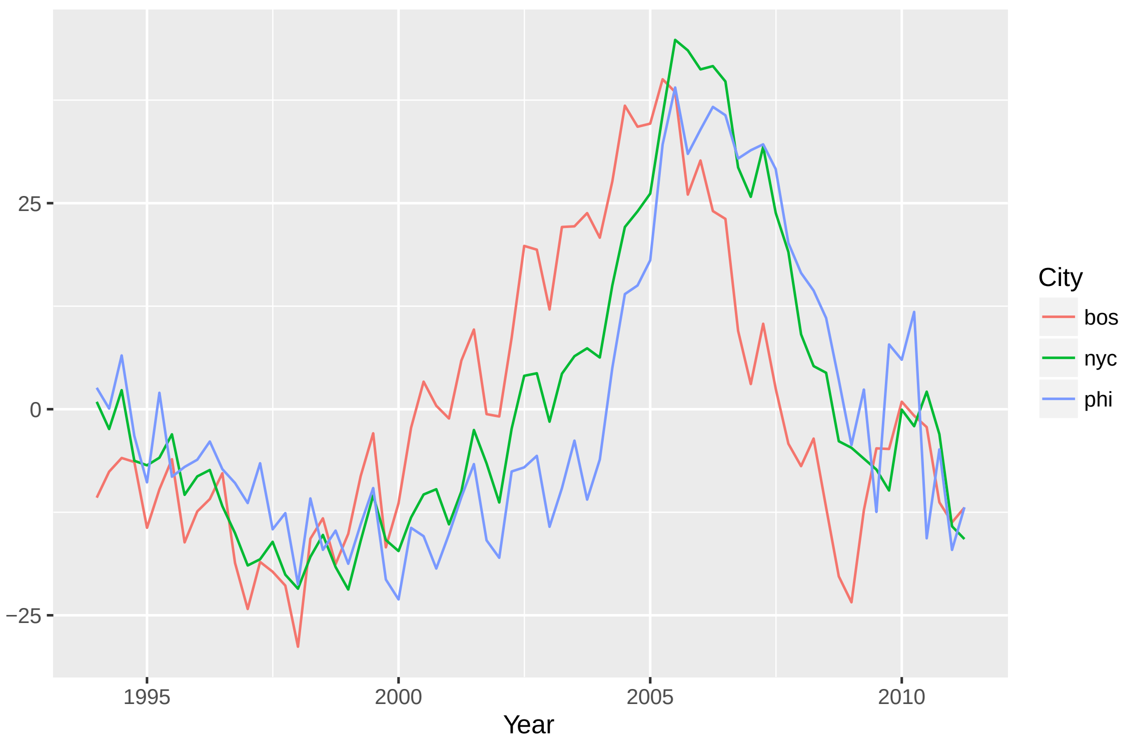


Figure 11: Building Permits in Los Angeles, San Francisco, and San Diego, 1994 - 2011

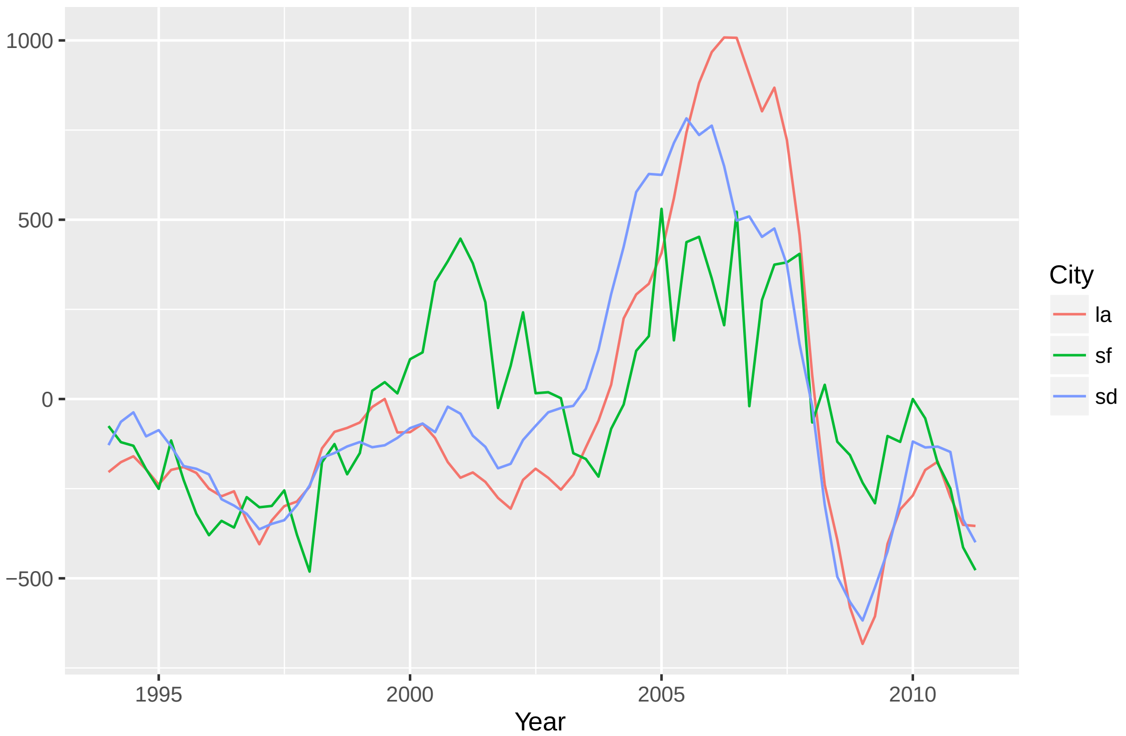


Figure 12: Building Permits in Los Angeles, Miami, Washington D.C, and Las Vegas, 1994 - 2011

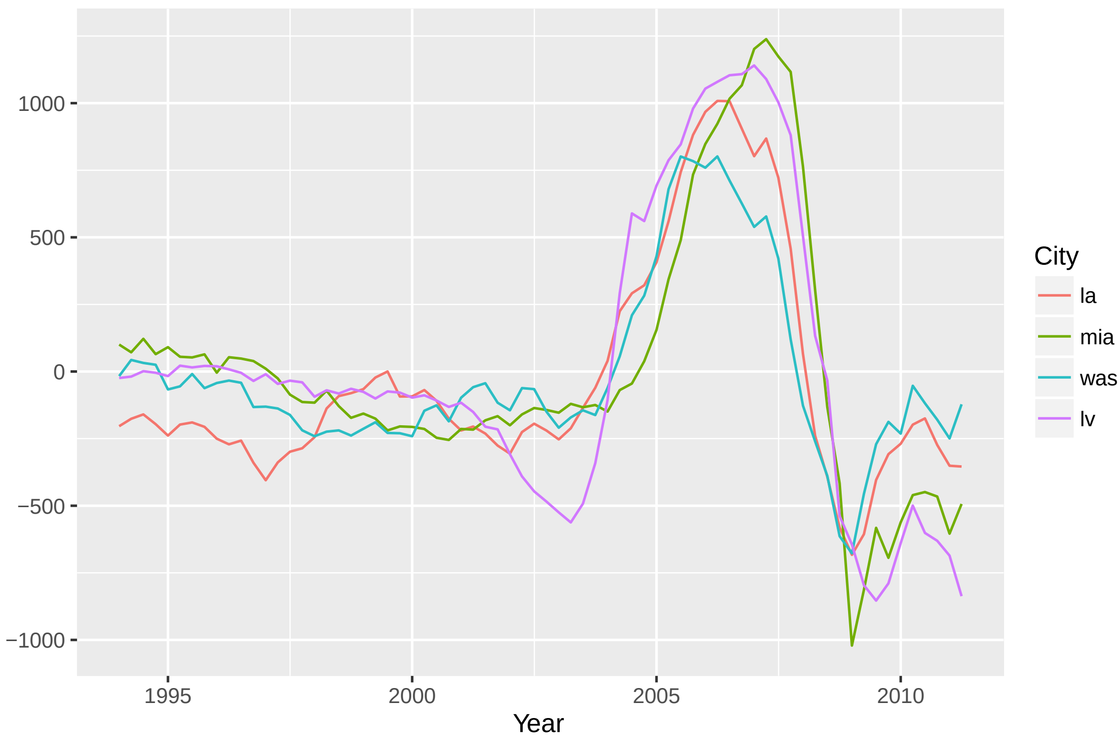


Figure 13: Response of Employment to a 1 S.E shock to Permits, All Cities.

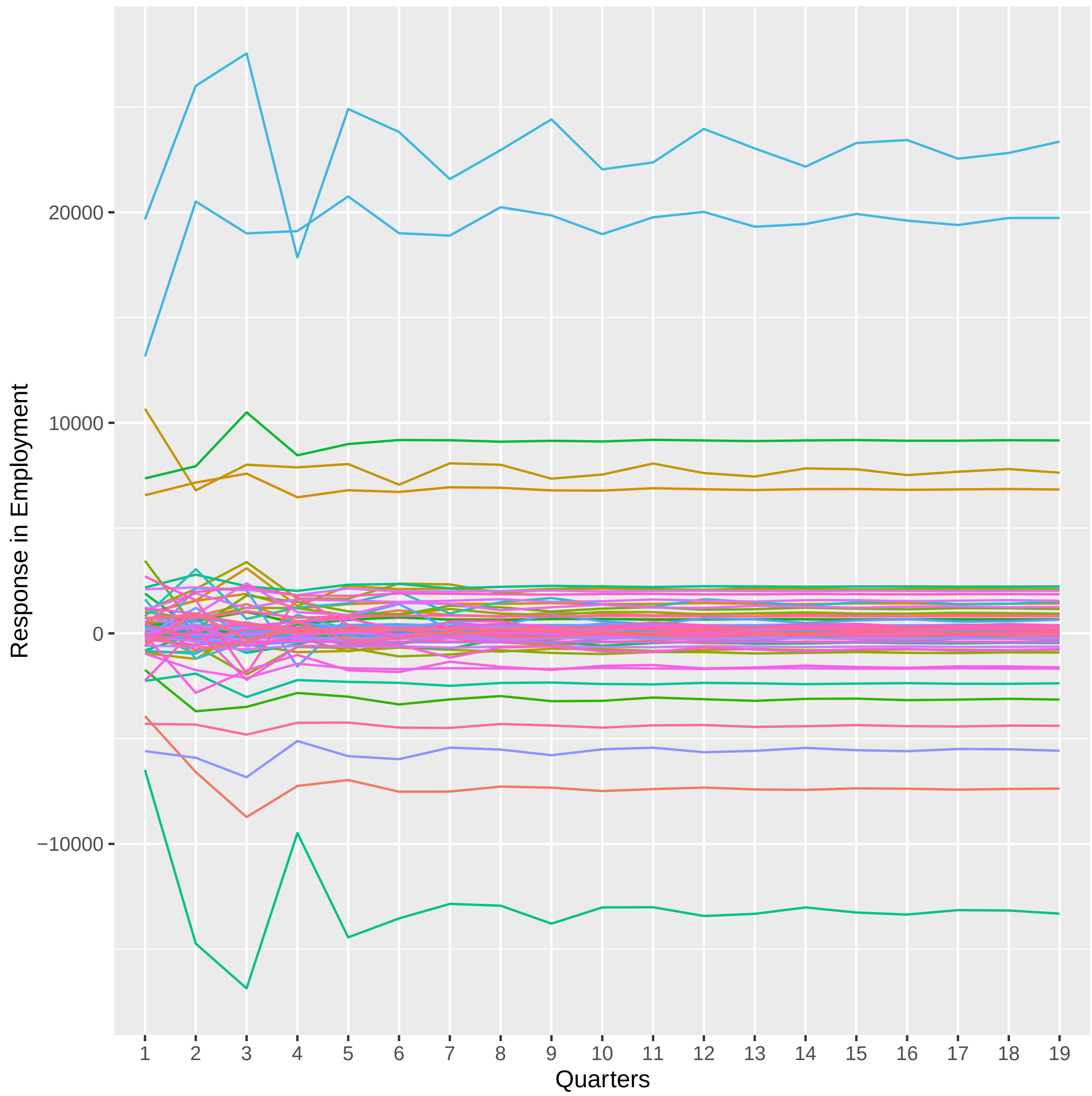


Figure 14: Cities where the Lead-Lag Relationship Holds (Blue) and Cities where it does not (Red).

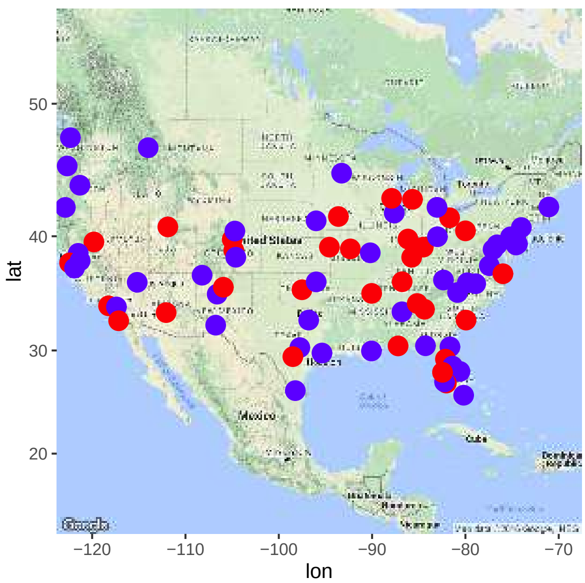


Figure 15: Distribution of Spillovers from a 1 S.E shock to Permits after 10 Quarters in a selection of Cities.

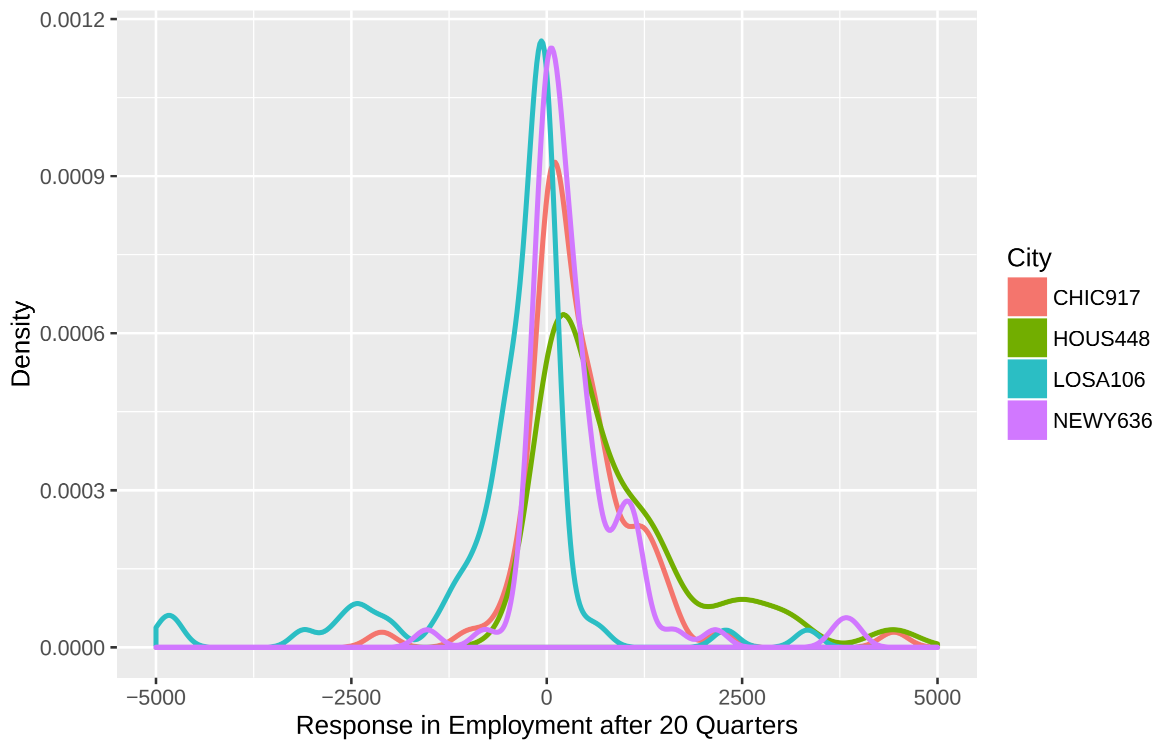


Figure 16: Spillovers from a 1 S.E Shock to Permits in Four Largest Cities.

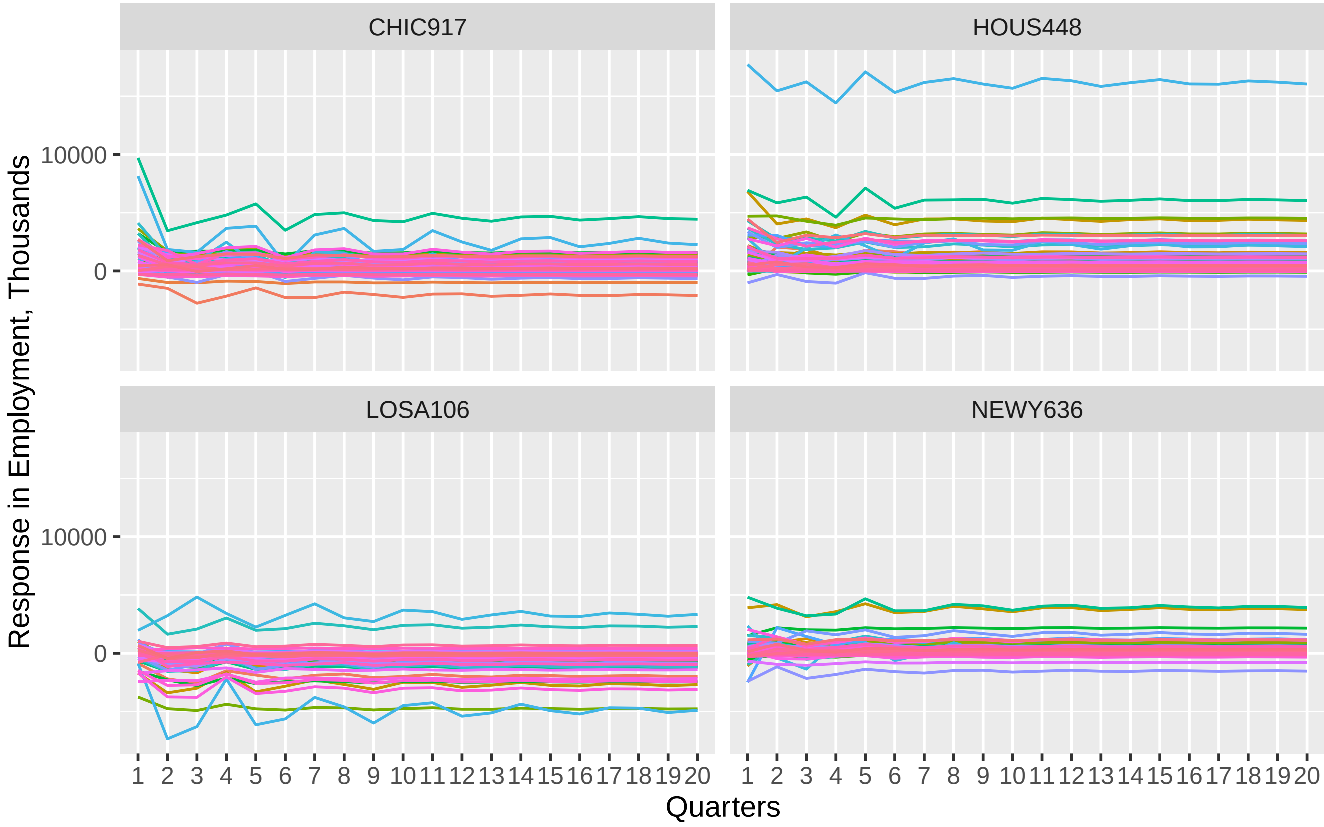


Figure 17: Spillovers from a 1 S.E Shock to Permits in Four Largest Cities.

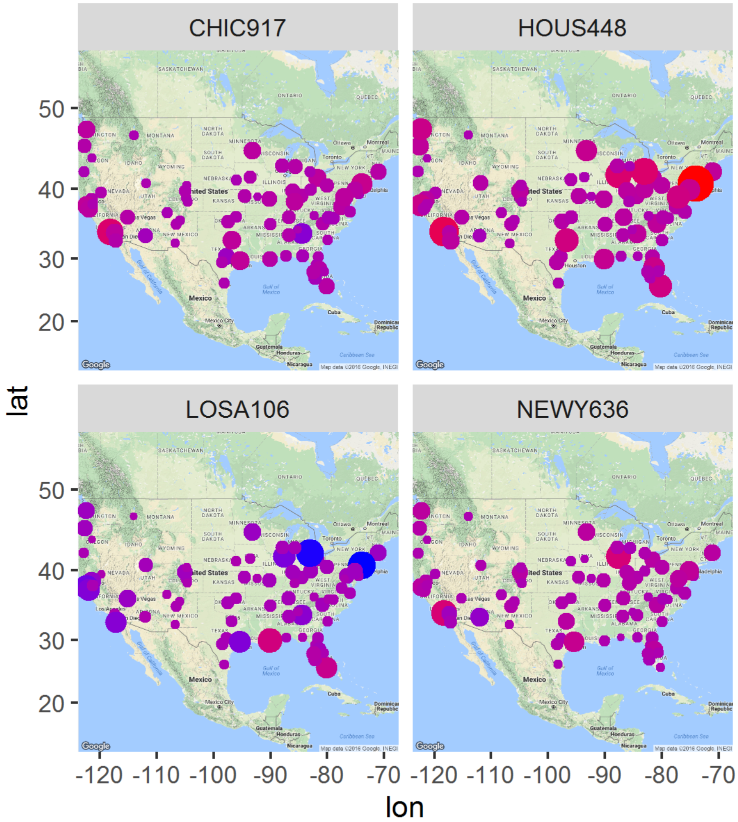
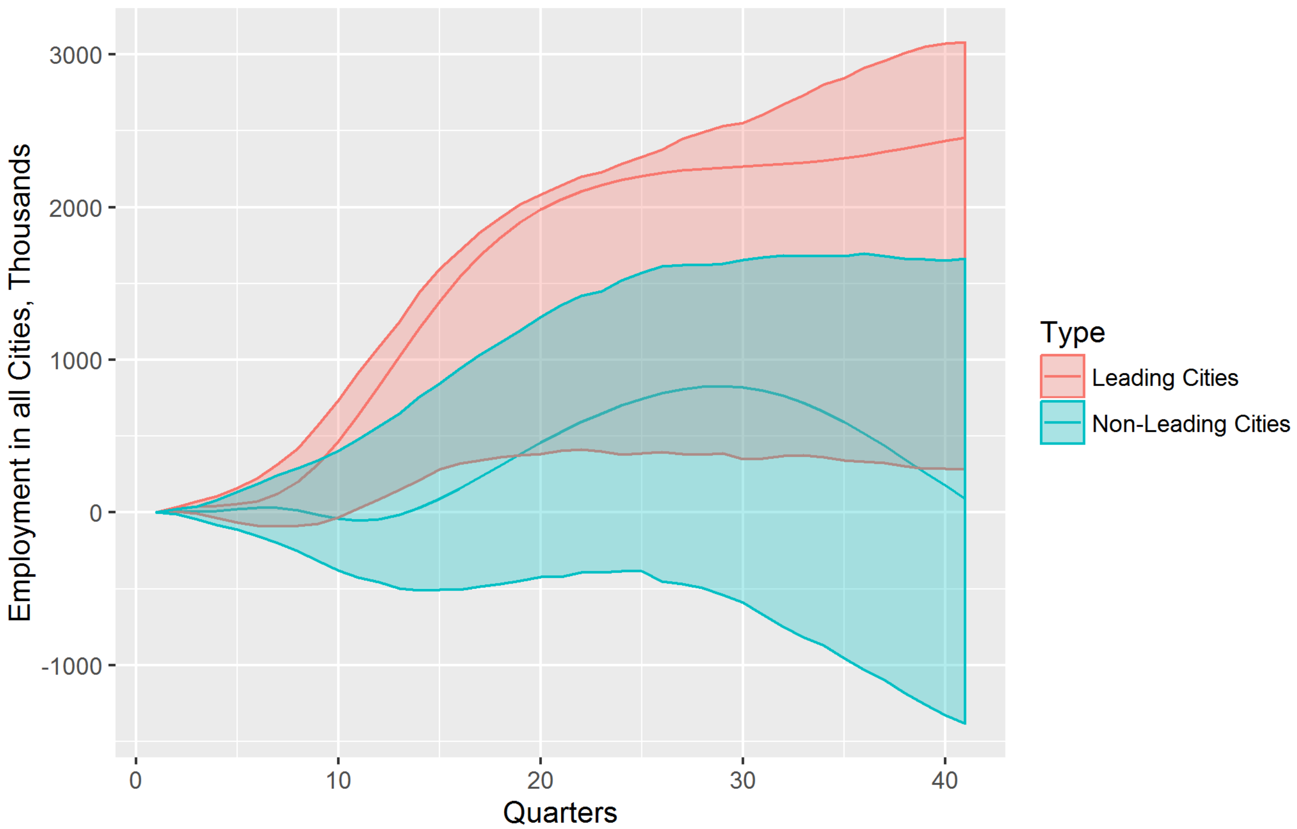


Figure 18: Response of Employment in All Cities to a 1 S.E Shock to Permits



8 Appendix B: Tables

Table 1: Summary Statistics of the Response Employment to a 1 S.E shock to Permits after 20 Quarters, All Cities.

	N	Mean	St. Dev.	Min	Max
Point Response	78	595.537	4,272.917	−13,231.590	22,918.930
Dummy Response	78	0.603	0.493	0	1

Table 2: Regressions of within City Response on City Characteristics

	Dependent variable:							
	OLS	Probit	OLS	Probit	OLS	Probit	OLS	Probit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DTI	−4,643.658**	−0.357***					−5,002.468	−0.717*
	(1,925.497)	(0.136)					(3,020.125)	(0.358)
NAICS 23			109,267.10	−0.01			103,741.20	1.04
			(274,698.50)	(9.26)			(294,470.70)	(10.00)
NAICS 21			245,119.60	−12.67			488,344.20	−7.40
			(463,048.50)	(15.62)			(544,117.70)	(18.47)
NAICS 3			7,615.75	−13.29			−138,121.30	−13.72
			(284,470.70)	(9.59)			(300,366.30)	(10.20)
NAICS 4			87,364.81	−7.00			84,761.03	−11.17
			(237,264.60)	(8.00)			(292,181.40)	(9.92)
NAICS 90			116,555.80	−38.17**			307,038.20	−36.03
			(561,041.90)	(18.92)			(652,810.60)	(22.16)
NAICS 53			−235,661.50	−39.58**			−536,152.50	−35.08
			(562,098.30)	(18.96)			(616,524.30)	(20.93)
Elasticity					−182.03	0.06	−803.40	−0.01
					(1,383.08)	(0.05)	(2,099.60)	(0.07)
W					−577.43	0.18**	−765.66	0.10
					(2,210.59)	(0.08)	(2,763.77)	(0.09)
HP					−2,307.56	0.52*	8,779.67	0.22
					(8,195.50)	(0.30)	(11,868.41)	(0.40)
Constant	91,473.06	−0.34	8,432.91	6.09	127,248.70	−6.12	−170,248.40	5.47
	(81,593.43)	(2.81)	(195,480.30)	(6.59)	(123,788.80)	(4.48)	(255,253.80)	(8.66)
Observations	79	79	79	79	79	79	79	79
R^2	0.17	0.16	0.37	0.39	0.30	0.22	0.43	0.44

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3

City	Mean	Std. Dev
Albuquerque, NM	-86.50458	847.7489
Atlanta-Sandy Springs-Roswell, GA	11.34522	809.2130
Atlantic City-Hammonton, NJ	124.73612	700.5778
Austin-Round Rock, TX	126.11491	756.4561
Baltimore-Columbia-Towson, MD	258.75992	943.5186
Bend-Redmond, OR	-375.21036	1188.4450
Birmingham-Hoover, AL	656.93631	1340.1245
Boston-Cambridge-Nashua, MA-NH	948.20969	1330.1801
Charlotte-Concord-Gastonia, NC-SC	183.65085	744.0768
Charleston-North Charleston, SC	80.54406	851.7843
Chicago-Naperville-Elgin, IL-IN-WI	395.86183	776.3000
Cincinnati, OH-KY-IN	-465.13223	1013.1053
Cleveland-Elyria, OH	-501.55995	1146.3006
Columbia, MO	-341.16792	1335.7104
Colorado Springs, CO	-350.65355	1107.6324
Columbus, OH	-93.94319	519.3398
Dallas-Fort Worth-Arlington, TX	-301.17712	1048.2297
Denver-Aurora-Lakewood, CO	269.65613	1067.8010
Des Moines-West Des Moines, IA	308.99835	756.0310
Detroit-Warren-Dearborn, MI	-69.82934	798.5872
Durham-Chapel Hill, NC	-112.21609	1570.0034
Fairbanks, AK	-57.92569	799.3063
Farmington, NM	280.30316	1165.7882
Grand Rapids-Wyoming, MI	-682.58890	1034.6986
Greeley, CO	-755.25658	1121.7647
Greensboro-High Point, NC	-359.60571	954.2729
Houston-The Woodlands-Sugar Land, TX	1126.33627	2086.8781
Indianapolis-Carmel-Anderson, IN	-125.64732	675.9441
Jacksonville, FL	140.60718	617.5841
Johnson City, TN	146.12105	734.4702
Kansas City, MO-KS	-618.49383	1241.0591
Las Cruces, NM	83.38441	1478.2859
Las Vegas-Henderson-Paradise, NV	-121.52502	778.5111
Los Angeles-Long Beach-Santa Ana, CA	-473.46856	1132.3642
Louisville/Jefferson County, KY-IN	-106.65875	1171.1868
McAllen-Edinburg-Mission, TX	26.13104	1023.0269
Medford, OR	69.06792	868.6539
Memphis, TN-MS-AR	-441.41877	1154.0195
Miami-Fort Lauderdale-West Palm Beach, FL	219.82284	500.7187
Milwaukee-Waukesha-West Allis, WI	127.45112	788.6264
Minneapolis-St. Paul-Bloomington, MN-WI	176.69546	615.7743
Missoula, MT	-515.73215	820.4533
Nashville-Davidson–Murfreesboro–Franklin, TN	-368.28021	725.1262
New Orleans-Metairie, LA	322.42364	798.7447
New York-Newark-Jersey City, NY-NJ-PA	367.62222	767.9606
Ocala, FL	-590.82967	1182.5434
Ocean City, NJ	286.49756	1064.6272
Oklahoma City, OK	-209.21387	1557.0437
Omaha-Council Bluffs, NE-IA	285.01835	1305.5851
Orlando-Kissimmee-Sanford, FL	-129.42567	1050.6353

Table 4

City	Mean	Std. Dev
Palm Bay-Melbourne-Titusville, FL	508.11034	809.5200
Pensacola-Ferry Pass-Brent, FL	615.81578	1020.6491
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	201.46448	1163.9119
Phoenix-Mesa-Scottsdale, AZ	-605.38813	916.1187
Pittsburgh, PA	36.12025	826.5619
Portland-Vancouver-Hillsboro, OR-WA	98.86769	629.5887
Pueblo, CO	75.08950	999.4793
Punta Gorda, FL	514.23968	1274.5071
Reno, NV	-139.61577	602.4797
Richmond, VA	399.00740	973.7698
Riverside-San Bernardino-Ontario, CA	26.26579	1209.1445
Rome, GA	98.47348	812.4426
Sacramento–Roseville–Arden-Arcade, CA	-183.05979	1197.3785
Salt Lake City, UT	-544.32012	779.5085
San Antonio-New Braunfels, TX	-206.32057	1357.3028
San Diego-Carlsbad, CA	122.60866	906.4371
San Francisco-Oakland-Hayward, CA	-467.74584	1036.6033
San Jose-Sunnyvale-Santa Clara, CA	302.71018	684.2586
Santa Fe, NM	-369.96662	2215.8776
North Port-Sarasota-Bradenton, FL	-329.77784	1381.1906
Seattle-Tacoma-Bellevue, WA	522.26830	1305.4519
St. Louis, MO-IL	115.22052	883.3373
Stockton-Lodi, CA	-802.84284	1326.3987
Tallahassee, FL	-445.24090	1165.9499
Tampa-St. Petersburg-Clearwater, FL	35.66828	704.9328
Tulsa, OK	622.51203	1495.8502
Virginia Beach-Norfolk-Newport News, VA-NC	-146.48945	662.1421
Washington-Arlington-Alexandria, DC-VA-MD-WV	-443.16427	1274.8921

Table 5: Regressions of Spillovers to Employment on 'Distance' Variables

	Dependent variable:	
	Y	abs(Y)
	(1)	(2)
Physical Distance	−0.014	0.028***
	(0.012)	(0.011)
Industrial Distance	552.775	−889.233***
	(352.448)	(310.746)
Migration Flow	0.116*	0.512***
	(0.066)	(0.058)
Constant	−57.987	468.093***
	(43.342)	(38.214)
Observations	6,634	6,634
R^2	0.001	0.015
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 6

	Dependent variable:			
	Y	abs(Y)	Y	abs(Y)
	(1)	(2)	(3)	(4)
Phy. Dist	−0.020	0.019**	0.003	0.017*
	(0.012)	(0.009)	(0.013)	(0.010)
Ind. Dist	760.697**	−688.712**	−318.266	−727.451**
	(375.670)	(281.661)	(401.877)	(305.154)
Mig. Flow	0.155**	−0.061	0.113	−0.076
	(0.068)	(0.051)	(0.070)	(0.053)
Rec. Prop Construction	−2,878.465	5,464.048***	−2,421.712	5,497.485***
	(1,834.013)	(1,375.064)	(1,810.631)	(1,374.855)
rec. Prop Manufacturing	−1,354.404	12,928.470***	−1,042.498	12,974.810***
	(1,596.431)	(1,196.934)	(1,575.886)	(1,196.607)
Rec. Prop Wholesale Trade	−3,719.300**	13,939.330***	−2,977.597*	13,965.320***
	(1,618.502)	(1,213.483)	(1,600.083)	(1,214.980)
Rec. Prop Finance	−841.914	−5,923.651**	−1,265.910	−6,017.678**
	(3,949.796)	(2,961.385)	(3,899.247)	(2,960.789)
Rec. Prop Real Estate	3,523.564	−9,578.830***	1,559.730	−9,567.607***
	(3,204.681)	(2,402.731)	(3,171.193)	(2,407.961)
Rec. Services	−3,011.379**	14,287.600***	−2,718.328**	14,317.410***
	(1,359.351)	(1,019.182)	(1,341.842)	(1,018.892)
Rec. DTI	147.705**	−144.481***	138.036**	−143.106***
	(69.382)	(52.020)	(68.448)	(51.974)
Rec. log(Pop)	−15.595	194.012***	−19.378	194.369***
	(12.973)	(9.727)	(12.867)	(9.770)
Rec. log(HP)	2.442	−34.992***	2.211	−35.069***
	(2.648)	(1.986)	(2.613)	(1.984)
Rec. WRLULI	−1.701	35.420***	−5.122	35.644***
	(16.353)	(12.260)	(16.136)	(12.252)
Rec. Elas	0.857	9.771	−2.634	9.983
	(13.567)	(10.172)	(13.396)	(10.172)
Orig. NAICS 23			1,913.002	885.244
			(1,810.269)	(1,374.580)
Orig. NAICS 22			5,685.173***	2,314.078
			(1,951.966)	(1,482.173)
Orig. NAICS 3			−3,230.205**	2,287.475*
			(1,575.785)	(1,196.530)
Orig. NAICS 4			715.216	810.524
			(1,600.028)	(1,214.938)
Orig. NAICS 52			−10,539.190***	−3,025.186
			(3,899.648)	(2,961.094)
Orig. NAICS 53			−18,016.650***	−303.413
			(3,170.362)	(2,407.330)
Orig. DTI			−428.547***	65.051
			(68.509)	(52.021)
Orig. log(pop)			29.638**	8.851
			(12.879)	(9.779)
Orig. log(hp)			1.131	−2.571
			(2.613)	(1.984)
Orig. WRLULI			−63.093***	14.533
			(16.141)	(12.256)
Orig. Elasticity			−77.884***	28.862***
			(13.397)	(10.173)
Constant	2,126.209*	−9,328.668***	3,450.735*	−10,050.030***
	(1,266.211)	(949.350)	(1,784.828)	(1,355.262)
Observations	6,634	6,634	6,634	6,634
R^2	0.005	0.291	0.034	0.294
Note:				*p<0.1; **p<0.05; ***p<0.01