# Local House Price Diffusion

Jeff Cohen
University of Connecticut
and
Jeff Zabel
Tufts University

**WEIMER SCHOOL OF ADVANCED STUDIES**
**IN REAL ESTATE AND LAND ECONOMICS**
***Big Data and Spatial Issues in Real Estate***
May 16, 2015

# General Principles

- We are interested in modeling the process by which house prices evolve over time and space.
- In particular, if there is a shock to house prices, how does that shock play out across time and space?
- Is there a
  - "ripple effect"?
  - "persistence effect"?

# Questions

- Does the house price diffusion process depend on:
  - The level of analysis?
  - The frequency of the data?
  - The housing cycle?
  - Supply-side factors (ease of building)?
- Is it different across borders of jurisdictions such as towns?

# Motivation

- This should be useful for
  - Understanding the full impact on prices of housing cycle fluctuations.
  - Understanding the impact of local shocks on prices in nearby jurisdictions.
  - Evaluating the full impact of policies that target specific areas such as enterprise or redevelopment zones or policies that create affordable housing in certain locations.

# Data

Data on house prices at three levels of aggregation:

- City:
  - Quarterly data from 1991:q1 – 2014q4 for 100 CBSAs (FHFA)

# Data

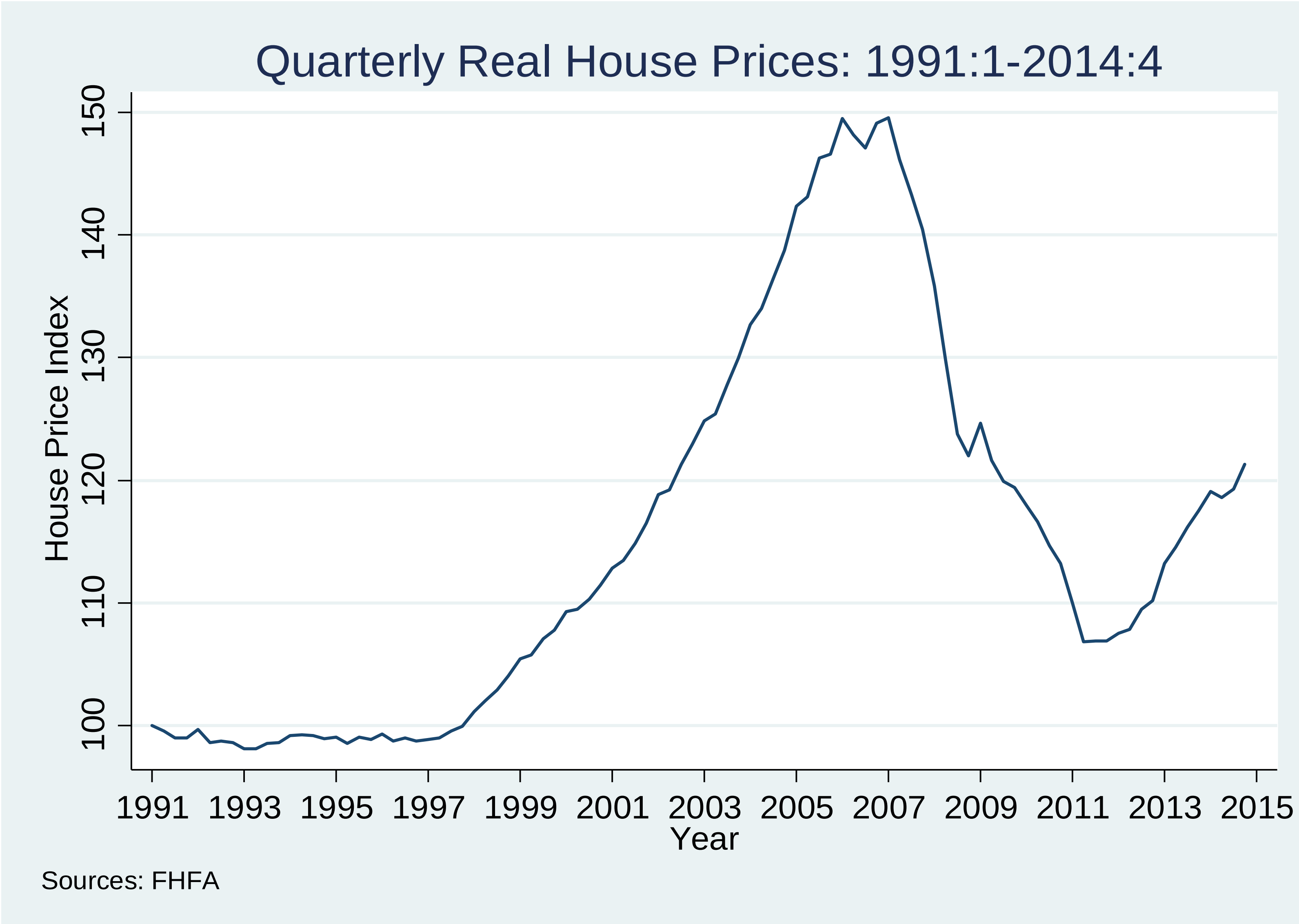
Data on house prices at three levels of aggregation:

- City:
  - Quarterly data from 1991:q1 – 2014q4 for 100 CBSAs (FHFA)
- Town:
  - Annual data on transactions of single family houses from 1987-2012 for 145 towns in the Greater Boston Area

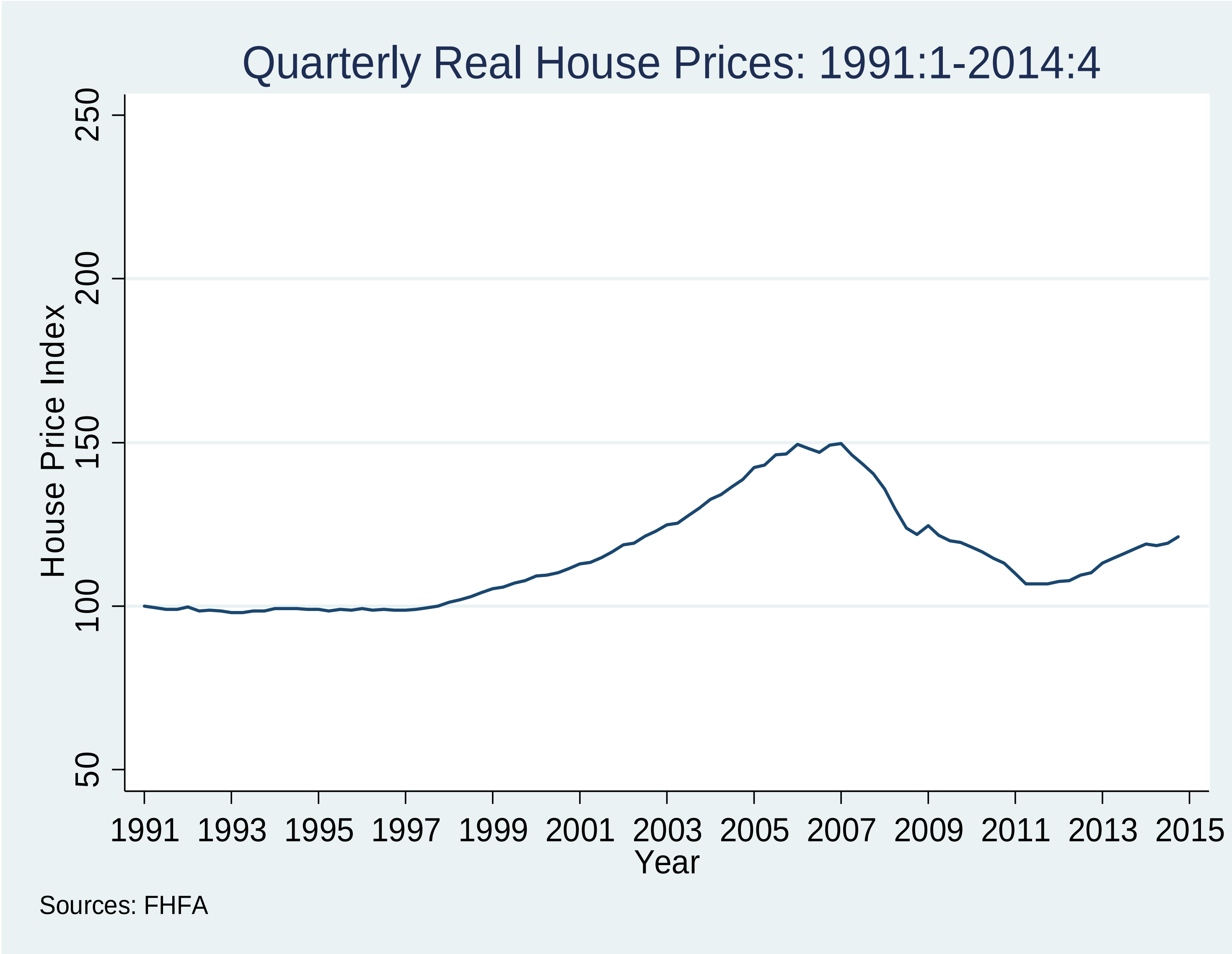
# Data

Data on house prices at three levels of aggregation:

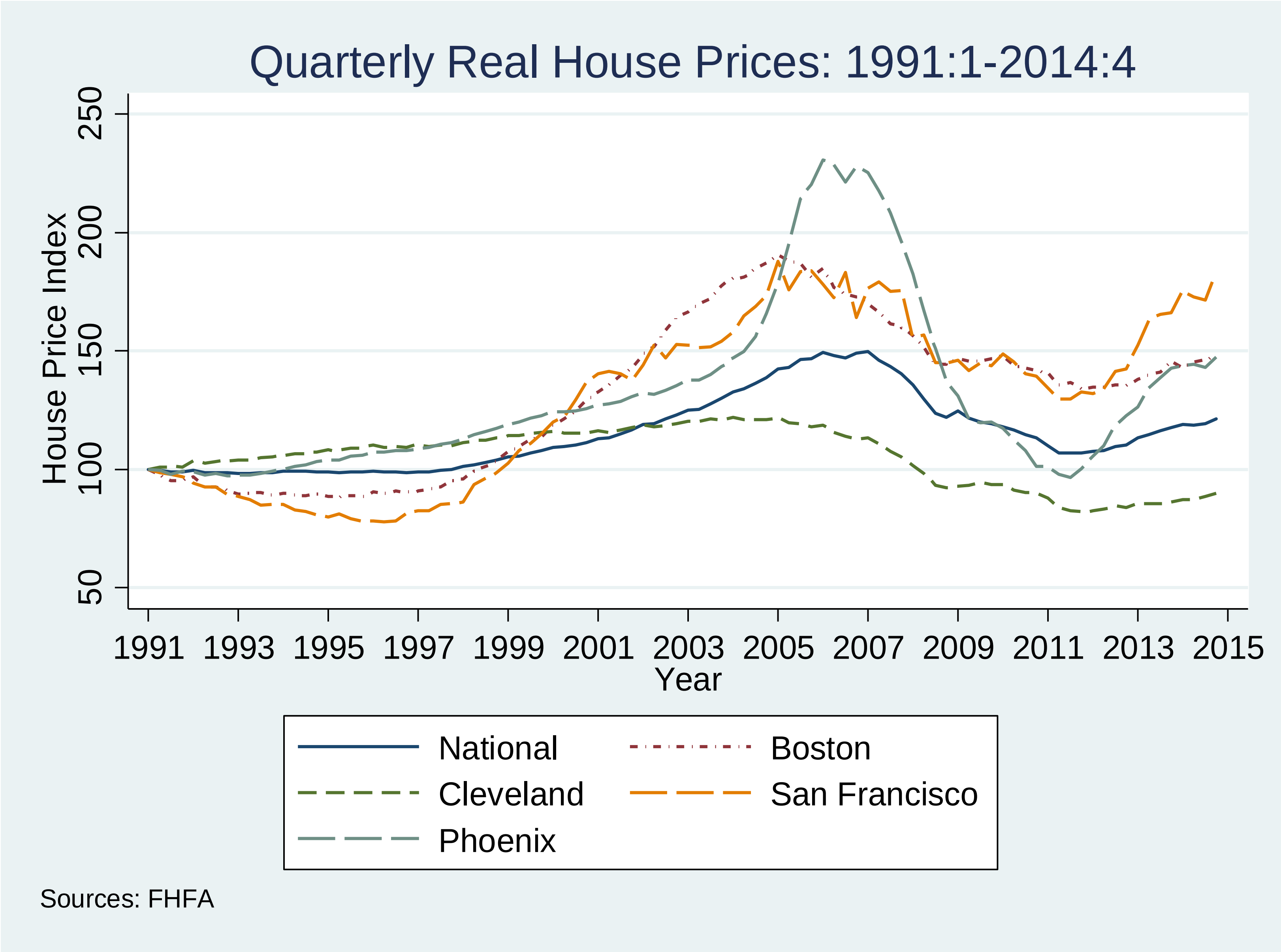
- City:
  - Quarterly data from 1991:q1 – 2014q4 for 100 CBSAs (FHFA)
- Town:
  - Annual data on transactions of single family houses from 1987-2012 for 145 towns in the Greater Boston Area
- Census Tract:
  - The same data at the census tract level



Quarterly Real House Prices: 1991:1-2014:4
House Price Index
100
110
120
130
140
150
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
Year
Sources: FHFA



Quarterly Real House Prices: 1991:1-2014:4
House Price Index
250
200
150
100
50
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
Year
Sources: FHFA



Quarterly Real House Prices: 1991:1-2014:4
House Price Index
50
100
150
200
250
1991
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
Year
National
Boston
Cleveland
San Francisco
Phoenix
Sources: FHFA

# House Price Model

$$\ln\left(\mathrm{P}_{\mathrm{injt}}\right) = \beta_0 + \mathrm{X}_{\mathrm{injt}}\beta_1 + \mathrm{u}_{\mathrm{njt}} + \mathrm{u}_{\mathrm{jt}} + \mathrm{u}_{\mathrm{t}} + \mathrm{e}_{\mathrm{injt}}$$

| | |
|---|---|
| $\mathrm{P}_{\mathrm{injt}}$ | Real price for house *i*, in neighborhood *n*, jurisdiction *j*, in year *t,* |
| $X_{injt}$ | vector of house characteristics, |
| $u_{njt}$ | neighborhood effect, |
| $u_{jt}$ | jurisdiction effect, and |
| $u_t$ | time effect. |

# House Price Index

$$\ln\left(P_{injt}\right) = \beta_0 + X_{injt}\beta_1 + u_{njt} + u_{jt} + u_t + e_{injt}$$

- We want to develop a house price index to model the diffusion process.
- This essentially controls for the structure of the house to focus on the value of the land.

# House Price Index

- At the CBSA level, the FHFA estimates a repeat-sales index for each of the 100 CBSAs – 1991q1 is set to 100

# House Price Index

- At the town level, we have three indices
  - A repeat-sales index from the Boston Fed
  - Using our transactions data to estimate annual town fixed effects
  - Using a nonparametric procedure known as geographically weighted regression (GWR)

# House Price Index

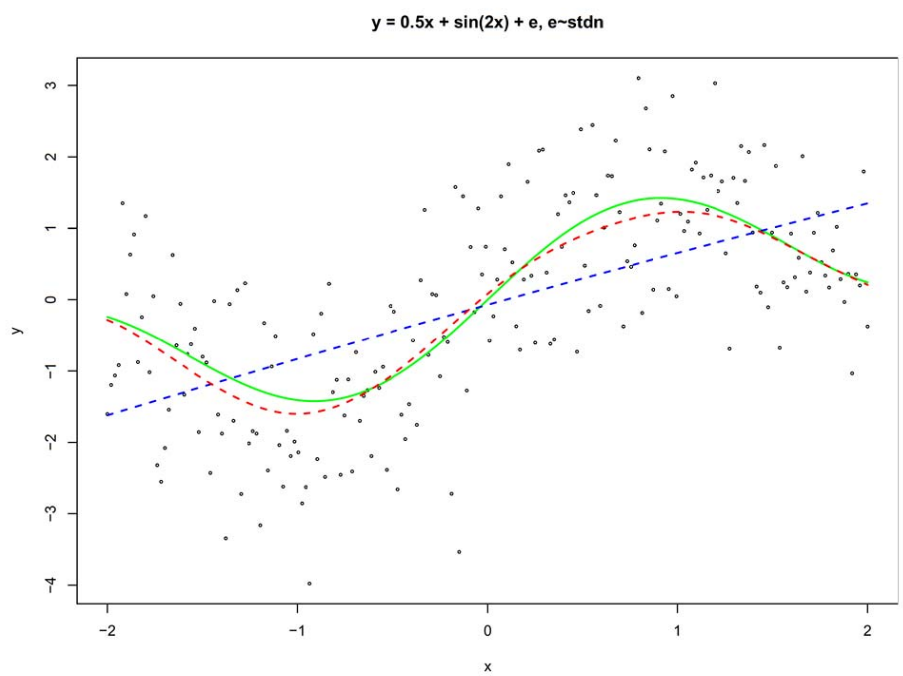
- At the census tract level, we have one approach
  - Using GWR
  - This procedure is particularly important at the tract level given that there can be few sales in a given census tract in a given year

# Our Nonparametric Approach:

- Geographically Weighted Regressions (McMillen and Redfearn, 2010)
- Separate regressions to be estimated around each observation in the sample
- Allows for nonlinear spatial structure
- Leads to a “smooth” approximation to the true function (McMillen and Redfearn, 2010)

# Parametric vs. Nonparametric: A Simple Example

- Generate data points through the data generating process on the right for the true function (Yang, 2011)
- Blue dotted line: OLS
- Green line: true function
- Red dotted line: results of nonparametric estimation
- Nonparametric more closely approximates the true function than OLS



# Model and Approach:

- $Y = f(X) + u$ ,
- $Y \equiv \ln(\text{average sale price})$,
- $X$=[residuals of OLS hedonic regression]

- Geographically Weighted Regressions:

(equivalently: Locally Weighted Regressions)

- $\beta_i = (\sum_j w_{ij} X_j X'_j)^{-1} (\sum_j w_{ij} X_j Y_j)$ , where $w_{ii} = 0$

# Kernel Weights

- We use Gaussian kernel:

$$w_{ij} = \boldsymbol{K}\left(\frac{d_{ij}}{b}\right) = \begin{Bmatrix} \boldsymbol{e}^{-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^{2}} & if\ year_i \geq year_j \\ 0 & otherwise \end{Bmatrix},$$

- K(•) is kernel function: tract j's weight on tract i
- $d_{ij}$ is distance between tracts j and i in geographic space,
- b is the bandwidth
- For nonparametric tract-level estimates, there are 13,910 separate coefficient estimates (535 tracts, 26 years)

# Diffusion Process

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(GR_{n_j,t-k}\beta_k\right) + \sum_{k=1}^{N} f\left(GR_{t-k}\gamma_k\right) + e_{jt}$$

| | |
|---|---|
| $GR_{jt}$ | Growth rate in jurisdiction j at time t |
| $GR_{n_j,t}$ | Growth rate in jurisdiction j's neighbors at time t |
| $GR_t$ | Growth rate in aggregate area at time t |

# Stylized Facts

- Glaeser et al (2011) – Housing Dyamics
  - SR: positive persistence
  - LR: mean reversion
  - Most variation in house prices is local not national

# CBSA Level Analysis

- Data are the FHFA house price index from 1991q1 to 2014q4 for 100 CBSAs
- We then generate quarterly and annual growth rates

# Diffusion Process

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(GR_{n_j,t-k}\beta_k\right) + \sum_{k=0}^{N} GR_{t-k}^{NAT}\gamma_k + e_{jt}$$

| | |
|---|---|
| $GR_{jt}$ | Growth rate in CBSA j at time t |
| $GR_{n_j,t}$ | Growth rate in CBSA j's neighbors at time t |
| $GR_t^{NAT}$ | National Growth rate at time t |

# Diffusion Process

$$\mathrm{GR}_{jt} = \beta_0 + \sum_{k=1}^{M} \mathrm{GR}_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(\mathrm{GR}_{n_j,t-k}\beta_k\right) + \sum_{k=0}^{N} \mathrm{GR}_{t-k}^{NAT}\gamma_k + e_{jt}$$

- Previous literature has often used excess returns
- The coefficient estimate for $\gamma_0$ (for the current national growth rate) is insignificantly different from 1 (either with quarterly or annual data)
- So the model can be rewritten as

# Diffusion Process

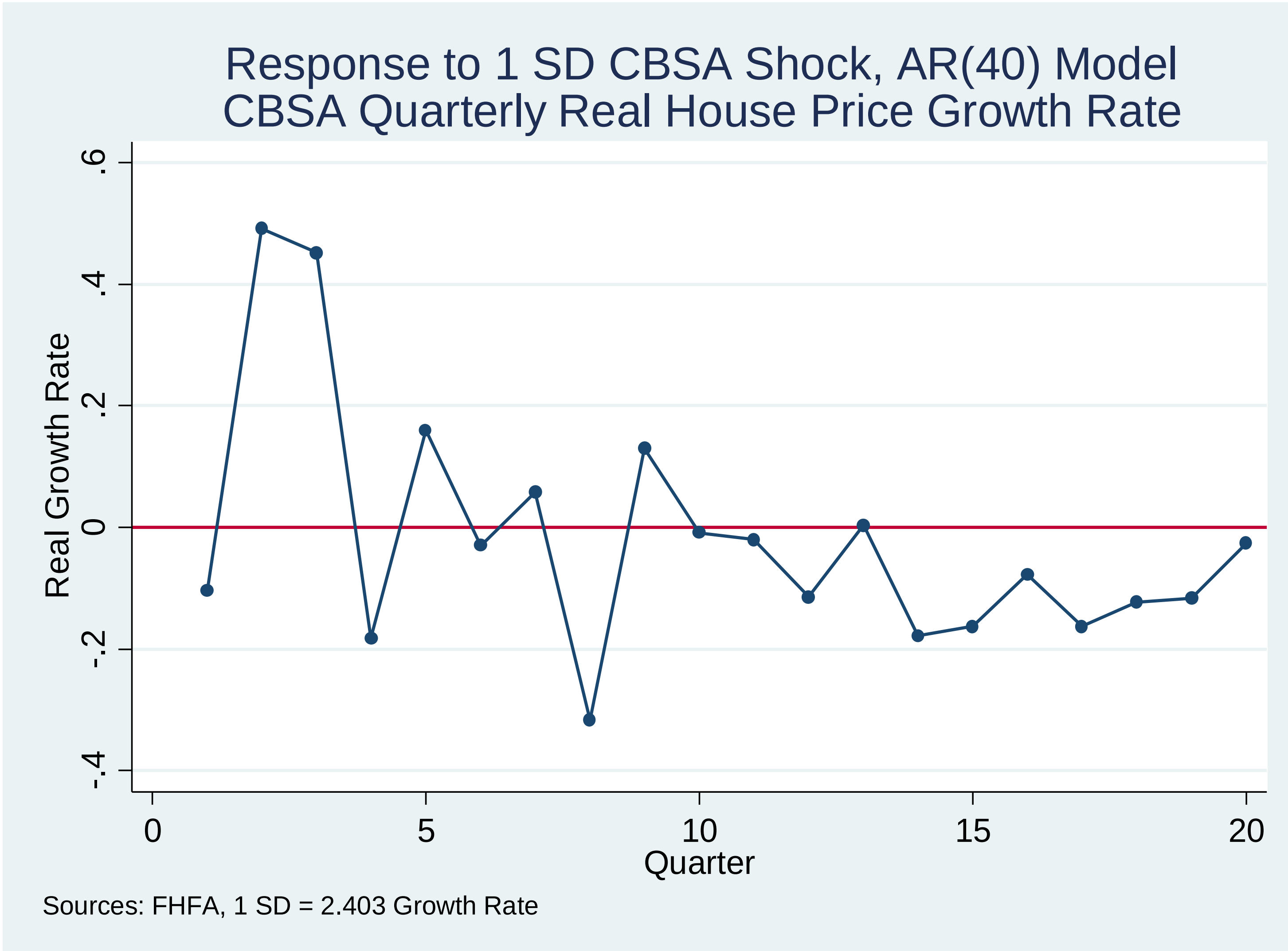
$$\mathrm{GR}_{jt}^{E} = \beta_0 + \sum_{k=1}^{M} \mathrm{GR}_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(\mathrm{GR}_{n_j,t-k}\beta_k\right) + \sum_{k=1}^{N} \mathrm{GR}_{t-k}^{NAT}\gamma_k + e_{jt}$$

where $\mathrm{GR}_{jt}^{E} = \mathrm{GR}_{jt} - \mathrm{GR}_{t}^{NAT}$

# First: Persistence Process

$$GR_{jt}^{E} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k$$

- Quarterly growth rates
- M = 40
- To show the persistence in growth rates consider a 1 standard deviation shock to CBSA growth rates = 8.13 Real Growth Rate

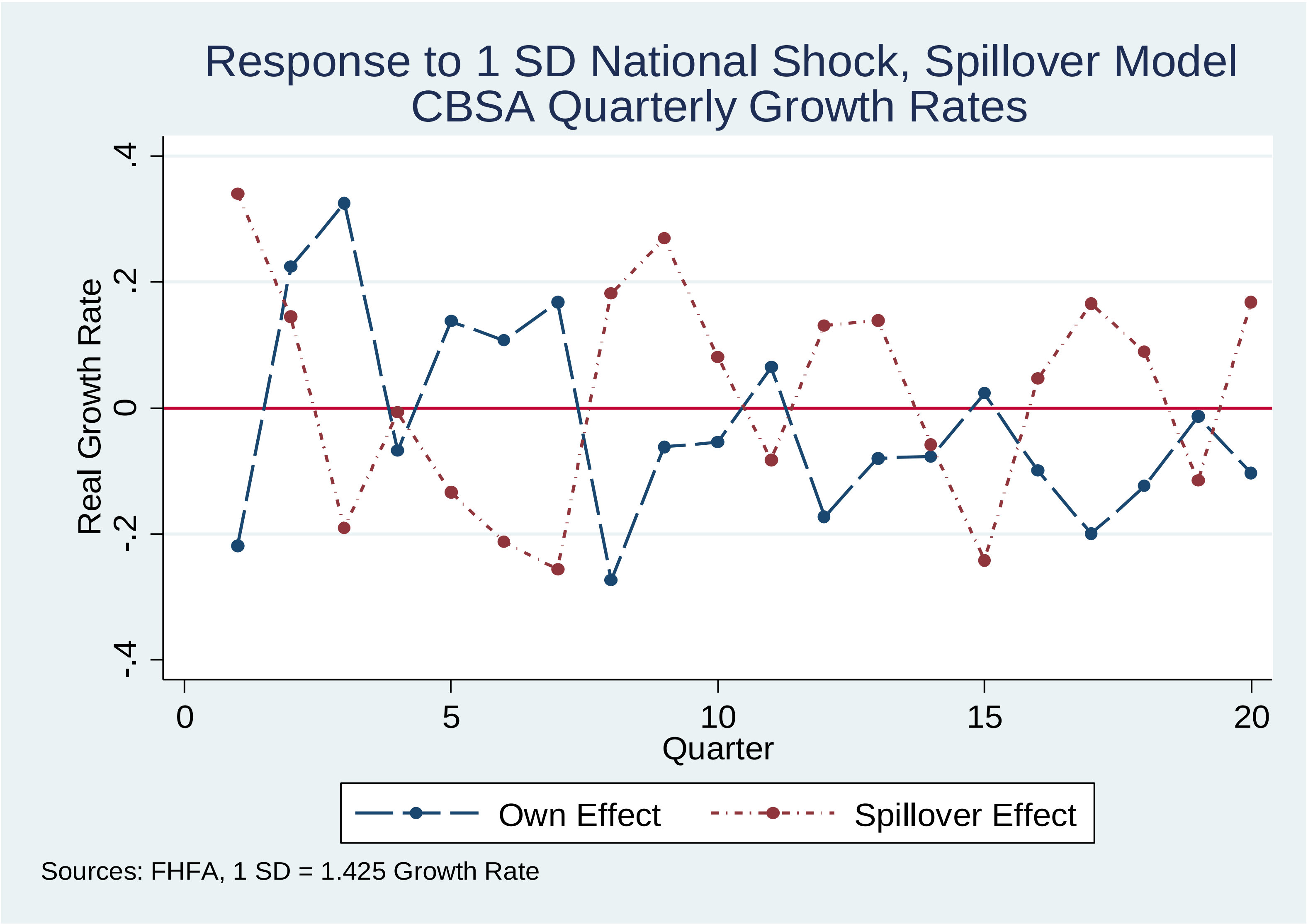

Response to 1 SD CBSA Shock, AR(40) Model
CBSA Quarterly Real House Price Growth Rate
Real Growth Rate
.6
.4
.2
0
-.2
-.4
0
5
10
15
20
Quarter
Sources: FHFA, 1 SD = 2.403 Growth Rate

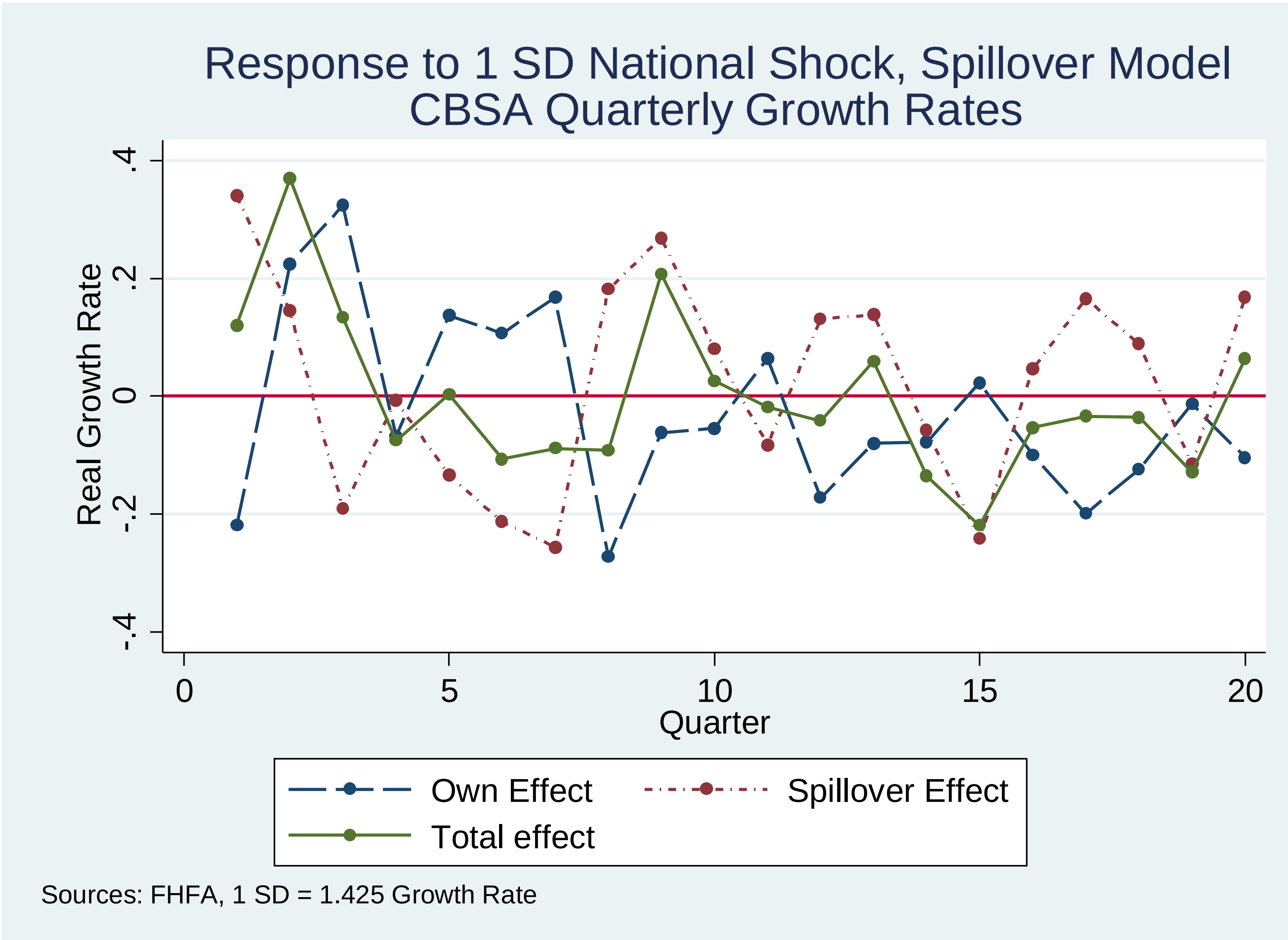
# Second: Spillover Effect

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(GR_{n_j,t-k}\beta_k\right) + e_{jt}$$

$f(\cdot)$ distance weighted average of 5 closest CBSAs

- N= 20
- consider a 1 standard deviation shock to National growth rate = 1.425 Real GR
- So own and nearby CBSAs get the same shock


Response to 1 SD National Shock, Spillover Model
CBSA Quarterly Growth Rates
Real Growth Rate
-.4
-.2
0
.2
.4
0
5
10
15
20
Quarter
Own Effect
Spillover Effect
Sources: FHFA, 1 SD = 1.425 Growth Rate

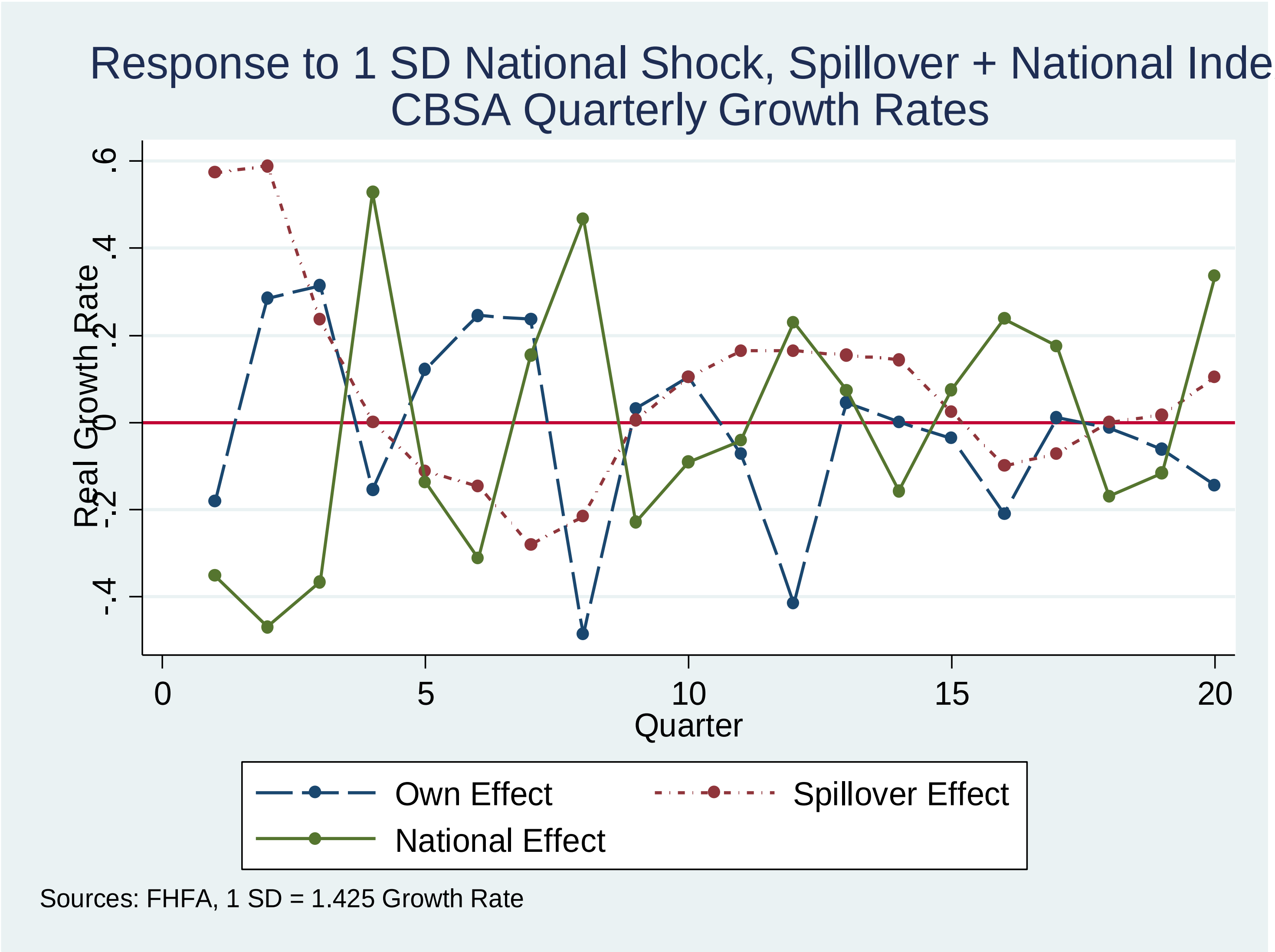


Response to 1 SD National Shock, Spillover Model
CBSA Quarterly Growth Rates
Real Growth Rate
-.4
-.2
0
.2
.4
0
5
10
15
20
Quarter
Own Effect
Spillover Effect
Total effect
Sources: FHFA, 1 SD = 1.425 Growth Rate

# Third: National Effect

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{N} f\left(GR_{n_j,t-k}\beta_k\right) + \sum_{k=0}^{P} GR_{t-k}^{NAT}\gamma_k + e_{jt}$$

- P=40 lags
- consider a 1 standard deviation shock to National growth rate = 1.425 Real GR



Response to 1 SD National Shock, Spillover + National Index
CBSA Quarterly Growth Rates
Real Growth Rate
-.4
-.2
0
.2
.4
.6
0
5
10
15
20
Quarter
Own Effect
Spillover Effect
National Effect
Sources: FHFA, 1 SD = 1.425 Growth Rate

# Options

- Maybe using national house price index is not appropriate since there is no national housing market
- Maybe estimate at regional level
- OR include interest rate, unemployment rate, new housing supply instead of national house price index

# Suggestions???

- Other ways to capture the ripple effect?

# Town Level Data

- Data on single family transactions for 145 towns (but not Boston) in the Greater Boston Area for 1987-2012

# House Price Model

$$\ln\left(\mathrm{P}_{\mathrm{injt}}\right) = \beta_0 + \mathrm{X}_{\mathrm{injt}}\beta_1 + \mathrm{u}_{\mathrm{njt}} + \mathrm{u}_{\mathrm{jt}} + \mathrm{u}_{\mathrm{t}} + \mathrm{e}_{\mathrm{injt}}$$

- Estimate house price hedonic by town fixed effects and use residual to control for structural characteristic
- Town price is average over transactions in each year
- Use GWR weighted by number of transactions where kernel is based on spatial and temporal “closeness”
- We then generate annual growth rates

# Town-level Growth Rates

- Two other sources of town-level growth rates:

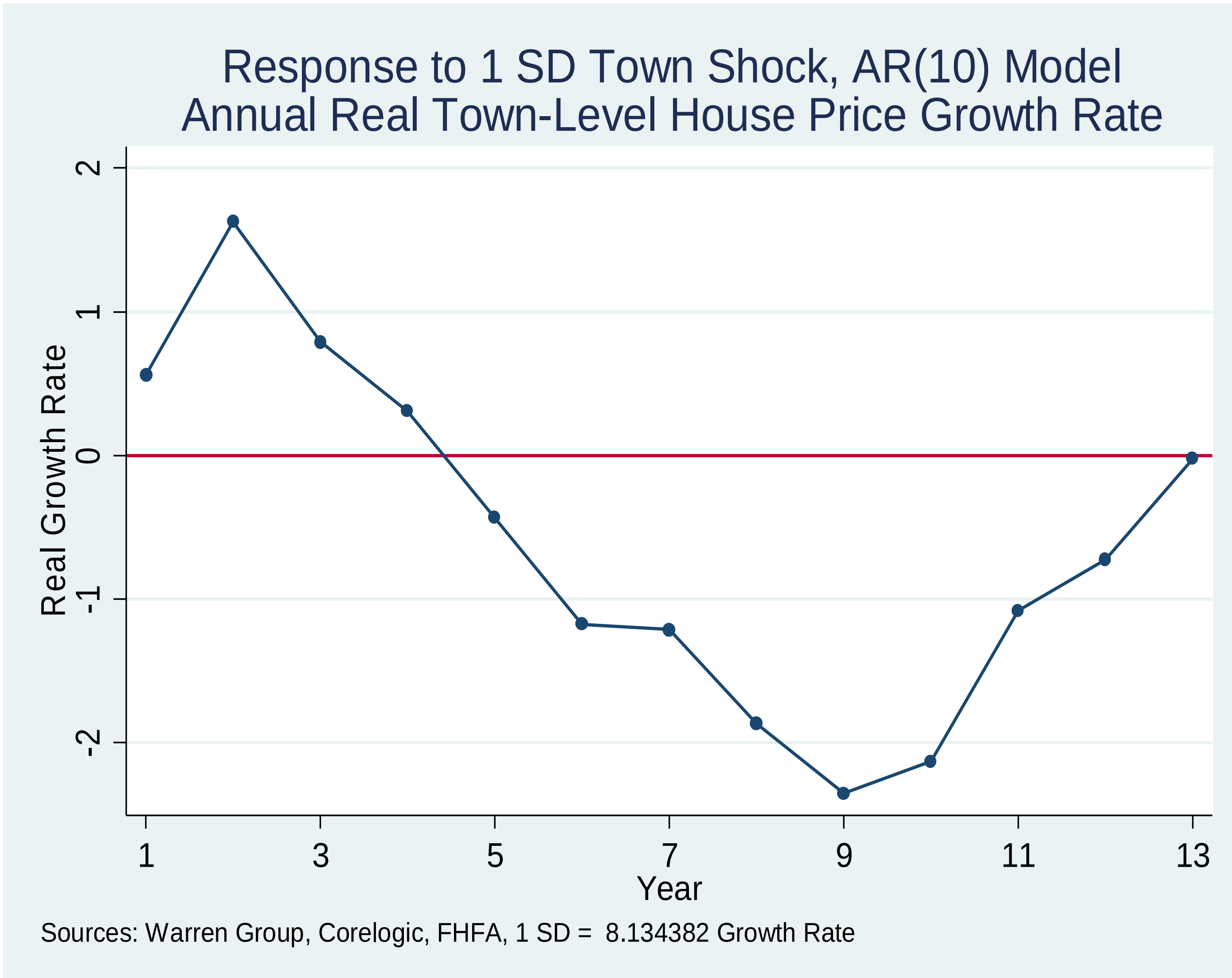
1. Boston Fed Repeat Sales Index
2. Town Fixed Effects

| Correlations Between Town-Level Growth Rates | | | |
|---|---|---|---|
| | GWR | Boston Fed | Town FE |
| **GWR** | 1.00 | | |
| **Boston Fed** | 0.60 | 1.00 | |
| **Town FE** | -0.15 | 0.18 | 1.00 |

# First: Persistence Process

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k$$

- First: just a persistence process.
- Annual growth rates
- M = 10
- To show the persistence in growth rates consider a 1 standard deviation shock to GBA growth rates = 2.403 Real Growth Rate


Response to 1 SD Town Shock, AR(10) Model
Annual Real Town-Level House Price Growth Rate
Real Growth Rate
2
1
0
-1
-2
1
3
5
7
9
11
13
Year
Sources: Warren Group, Corelogic, FHFA, 1 SD = 8.134382 Growth Rate

# Second: Full Model

$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{A} f\left(GR^{A}_{j\;,t-k}\beta_k\right) + \sum_{k=1}^{B} f\left(GR^{A^2}_{j\;,t-k}\delta_k\right)$$
$$+ \sum_{k=0}^{P} GR^{GBA}_{t-k}\gamma_k + e_{jt}$$

$GR^{A}_{j,t}$ average growth rate for towns adjacent to j

$GR^{A^2}_{j,t}$ average growth rate for towns adjacent to towns adjacent to j (excluding j)

$GR^{GBA}_{t}$ growth rate for GBA (from FHFA)

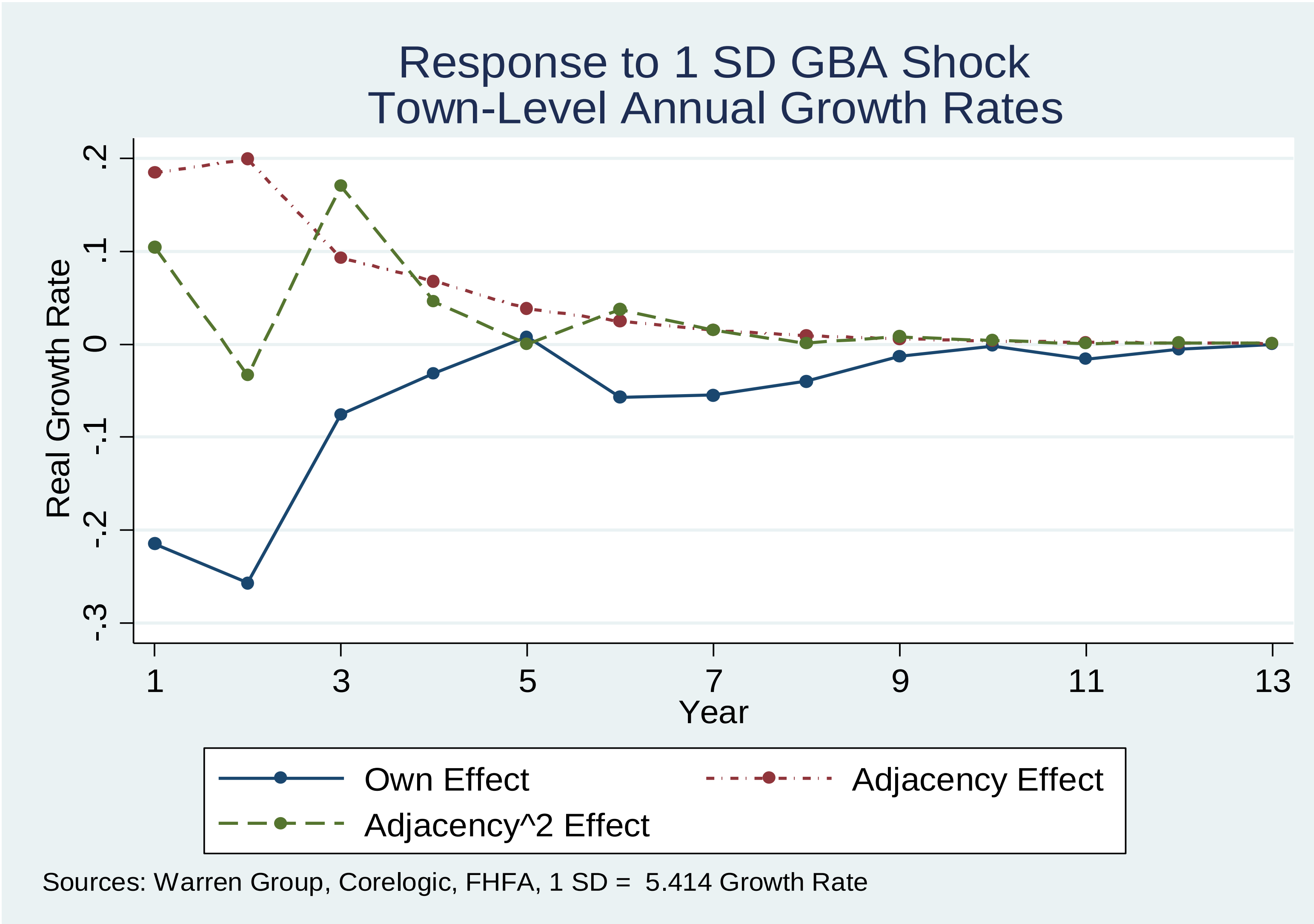
# Second: Full Model

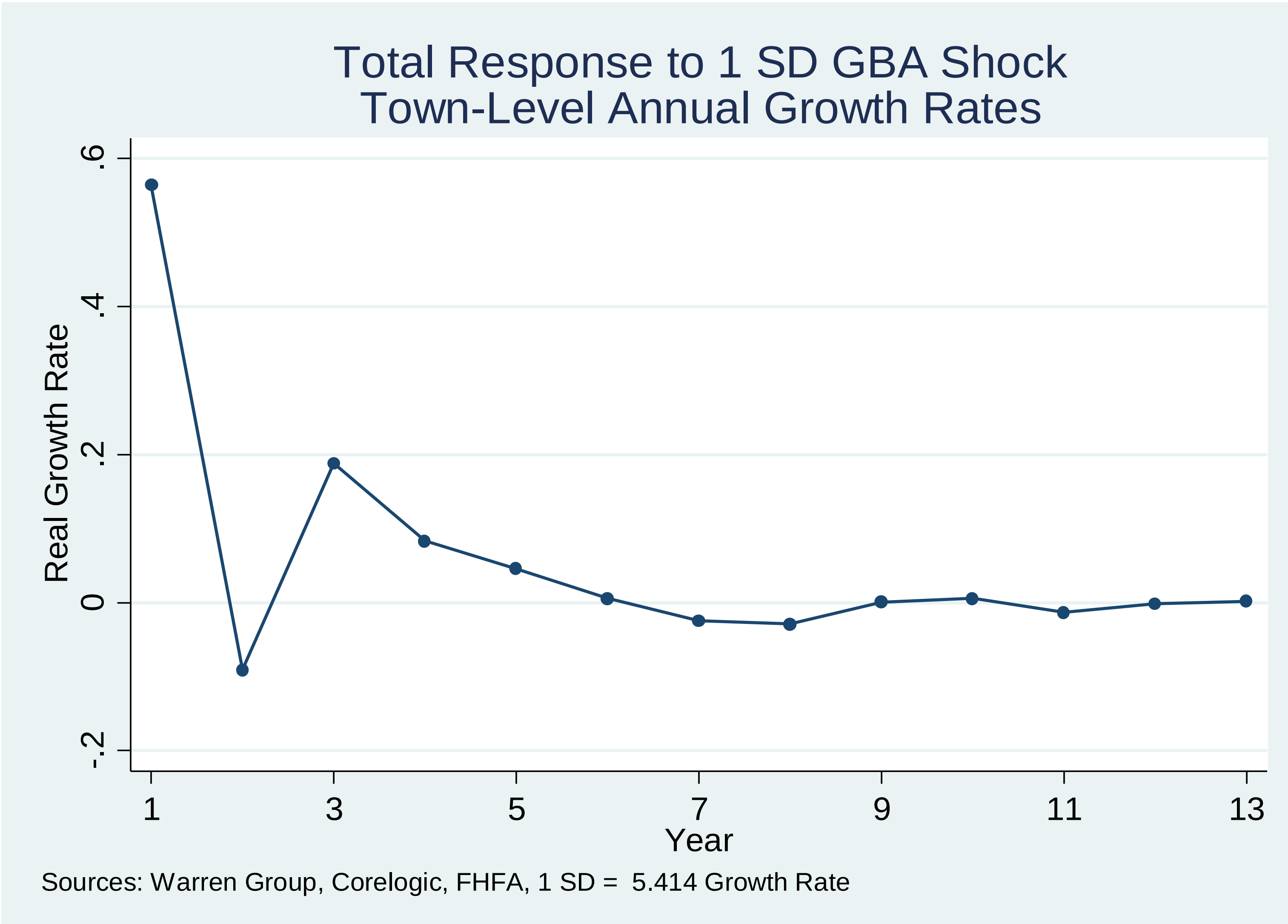
$$GR_{jt} = \beta_0 + \sum_{k=1}^{M} GR_{j,t-k}\alpha_k + \sum_{k=1}^{A} f\left(GR_{j}^{A}{}_{,t-k}\beta_k\right) + \sum_{k=1}^{B} f\left(GR_{j}^{A^2}{}_{,t-k}\delta_k\right)$$

$$+ \sum_{k=0}^{P} GR_{t-k}^{GBA}\gamma_k + e_{jt}$$

In this case, the coefficient estimate for $\gamma_0$ is different from one (around 0.15) so current value for $GR_t^{GBA}$ is included as explanatory variable.

M=8, A=2, B=3, P=1


Response to 1 SD GBA Shock
Town-Level Annual Growth Rates
Real Growth Rate
.2
.1
0
-.1
-.2
-.3
1
3
5
7
9
11
13
Year
Own Effect
Adjacency Effect
Adjacency^2 Effect
Sources: Warren Group, Corelogic, FHFA, 1 SD = 5.414 Growth Rate


Total Response to 1 SD GBA Shock
Town-Level Annual Growth Rates
Real Growth Rate
-.2
0
.2
.4
.6
1
3
5
7
9
11
13
Year
Sources: Warren Group, Corelogic, FHFA, 1 SD = 5.414 Growth Rate

# Census Tract-Level

- We next look at price diffusion at the census tract level
- This allows for diffusion within and across towns
- Data on single family transactions for 535 census tracts in the Greater Boston Area with at least one sale in each year for 1987-2012.

# Consistent Census Tracts

- We needed to deal with the problem that census tracts change over time; they can split or merge.
- Using GIS, we have determined the largest origin tract as the consistent tract.
- For example, if tract A splits into B and C in 2000 we use A as the consistent tract and aggregate sales in B and C starting in 2000

# Estimating Census Tract Prices

- We use the GWR technique to get estimates of tract prices for each year.
- Here, GWR is particularly useful since there can be only a few sales in a given tract in a given year
- GWR uses sales in "nearby" tracts in space and time to better estimate annual tract prices.

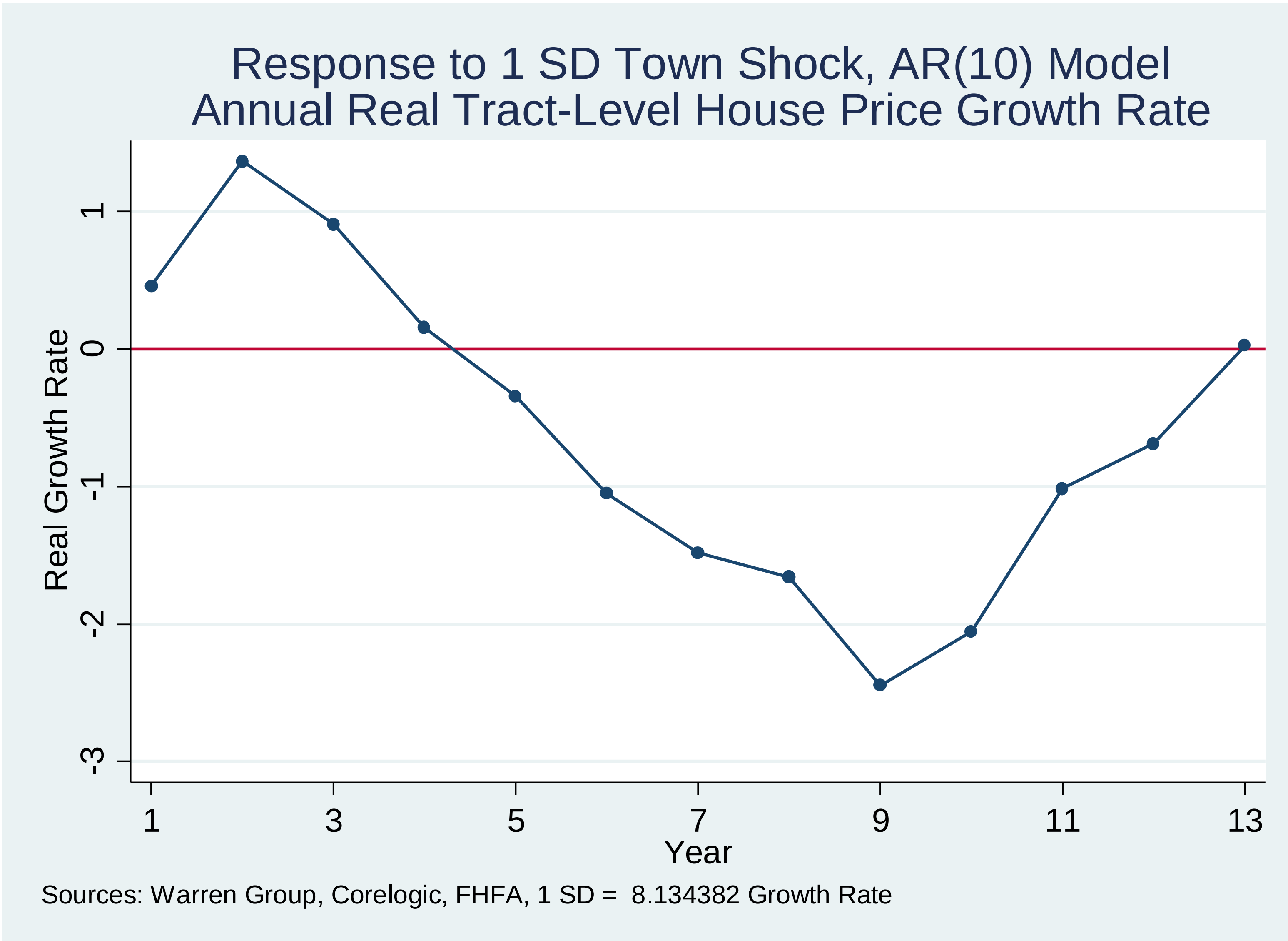
# Estimating Census Tract Prices

- GWR uses sales in “nearby” tracts in space and time to better estimate annual tract prices.
- Q1: If prices are rising over time, should we “deflate” by price increase before using kernel that uses temporally close sales?
- Q2: Should we weight tracts in same town differently than those in adjacent towns?

# First: Persistence Process

$$GR_{cjt} = \beta_0 + \sum_{k=1}^{M} GR_{cj,t-k}\alpha_k$$

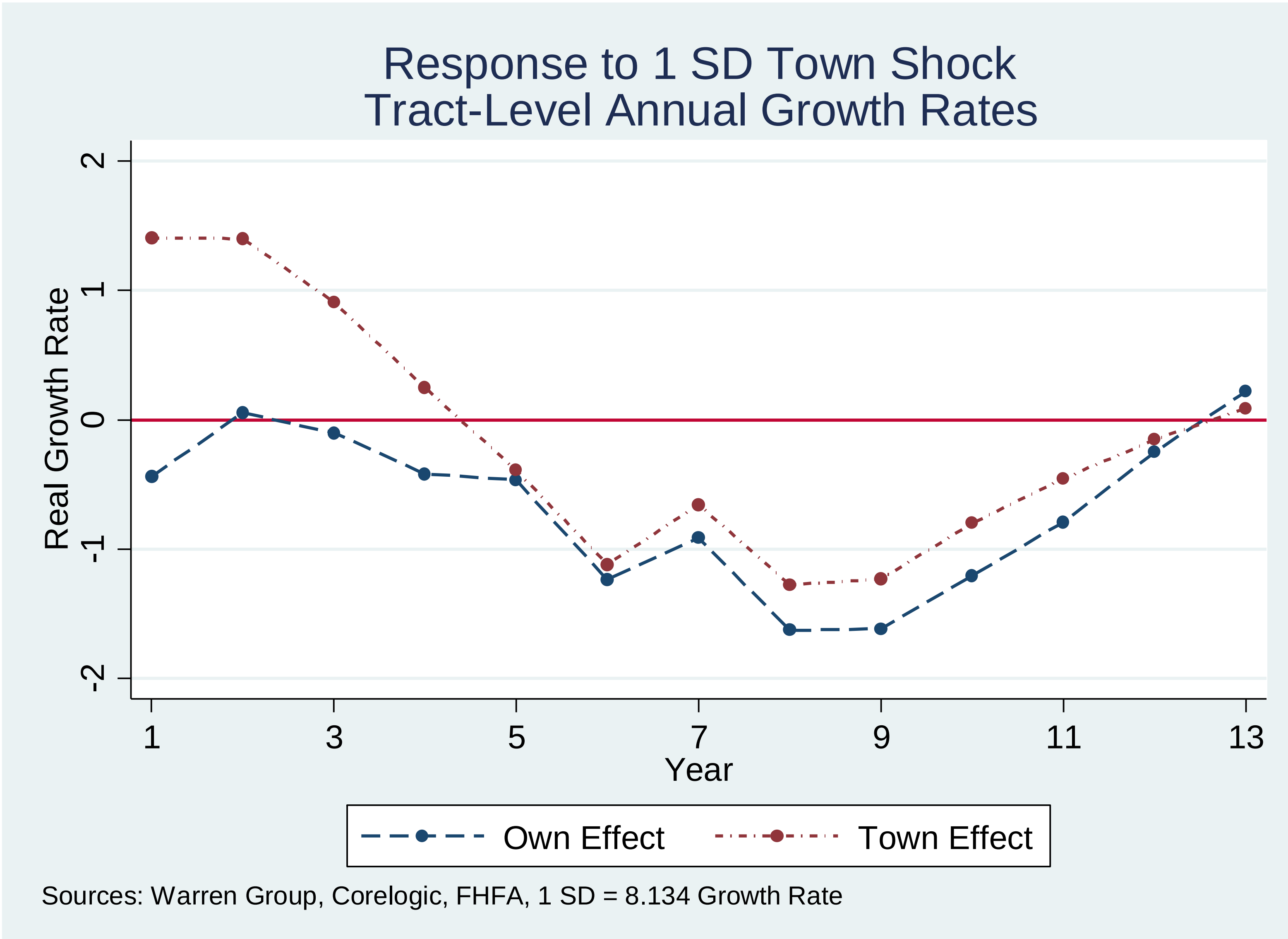
- $GR_{cjt}$ - Annual Real growth rate in tract c, town j, time t
- M = 10
- To show the persistence in growth rates consider a 1 standard deviation shock to Town growth rates = 8.134 Real Growth Rate


Response to 1 SD Town Shock, AR(10) Model
Annual Real Tract-Level House Price Growth Rate
Real Growth Rate
1
0
-1
-2
-3
1
3
5
7
9
11
13
Year
Sources: Warren Group, Corelogic, FHFA, 1 SD = 8.134382 Growth Rate

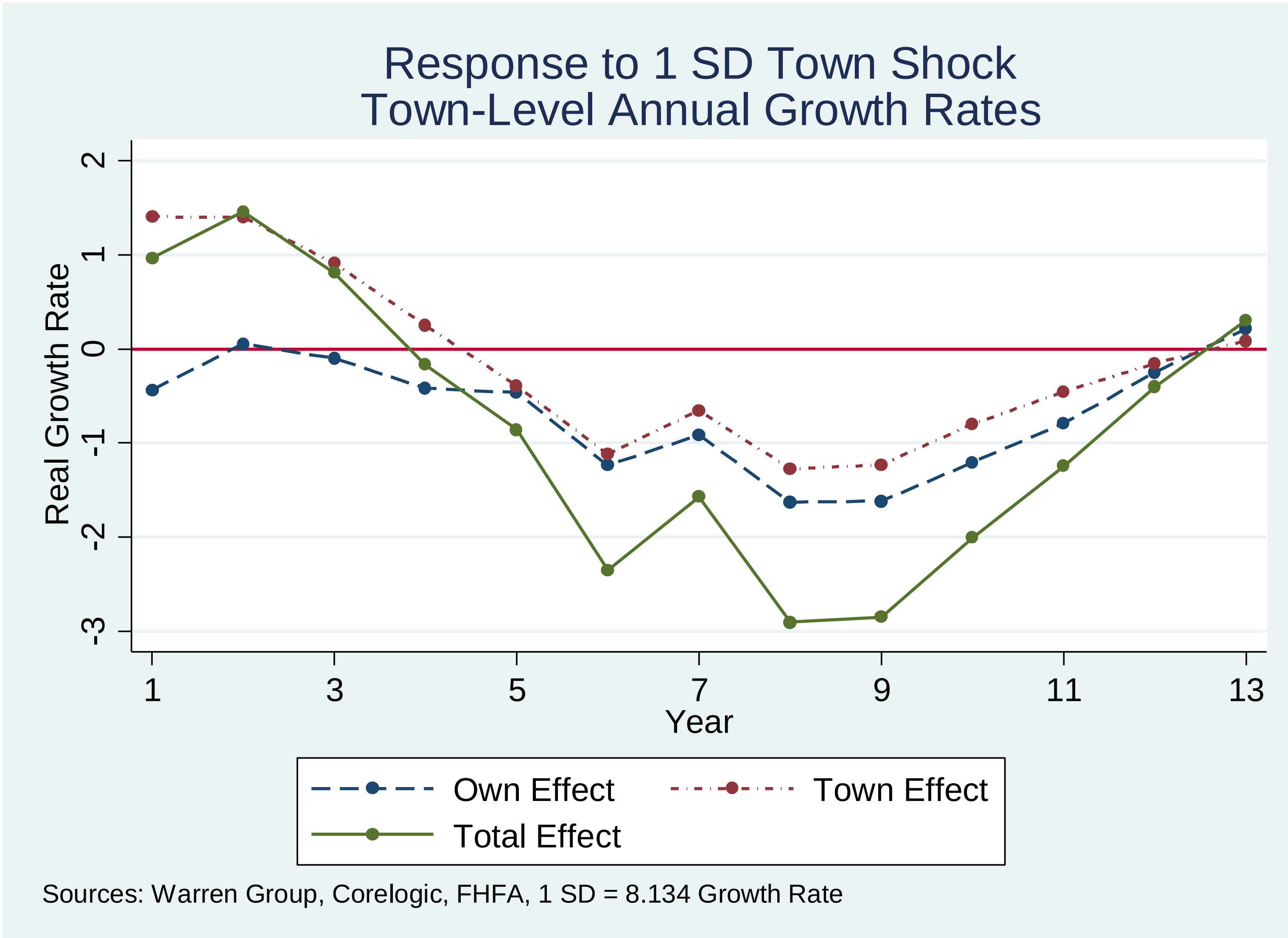
# Second: Tract + Town Model

$$GR_{cjt} = \beta_0 + \sum_{k=1}^{M} GR_{cj,t-k}\alpha_k + \sum_{k=1}^{P} GR_{j,t-k}\gamma_k + e_{jt}$$

- $GR_{cjt}$ - Annual real growth rate in tract c, town j, time t
- $GR_{jt}$ - Annual real growth rate in town j, $\gamma_0$ time t
- M=10, P=10
- consider a 1 standard deviation shock to Town growth rates = 8.134 Real Growth Rate



Response to 1 SD Town Shock
Tract-Level Annual Growth Rates
Real Growth Rate
2
1
0
-1
-2
1
3
5
7
9
11
13
Year
Own Effect
Town Effect
Sources: Warren Group, Corelogic, FHFA, 1 SD = 8.134 Growth Rate


Response to 1 SD Town Shock
Town-Level Annual Growth Rates
Real Growth Rate
2
1
0
-1
-2
-3
1
3
5
7
9
11
13
Year
Own Effect
Town Effect
Total Effect
Sources: Warren Group, Corelogic, FHFA, 1 SD = 8.134 Growth Rate