# An Introduction to Housing Price Aggregates

William H. Rogers
University of Missouri at Saint Louis

January 27, 2011

# Why are Housing Price Aggregates Important?

- Shelter is important

# Why are Housing Price Aggregates Important?

- Shelter is important
  - Housing affordability is a major policy concern in most areas
  - Homeownership (also a major policy concern) may be affected by housing prices

# Why are Housing Price Aggregates Important?

- Shelter is important
  - Housing affordability is a major policy concern in most areas
  - Homeownership (also a major policy concern) may be affected by housing prices
- Real estate (mainly in the form of housing) is a major source of wealth
- Real estate markets may have a major influence on business cycles

# Why are Housing Price Aggregates Important?

- Shelter is important
  - Housing affordability is a major policy concern in most areas
  - Homeownership (also a major policy concern) may be affected by housing prices
- Real estate (mainly in the form of housing) is a major source of wealth
- Real estate markets may have a major influence on business cycles
  - The wealth effect: although most estimates suggest a weak connection

# Why are Housing Price Aggregates Important?

- Shelter is important
  - Housing affordability is a major policy concern in most areas
  - Homeownership (also a major policy concern) may be affected by housing prices
- Real estate (mainly in the form of housing) is a major source of wealth
- Real estate markets may have a major influence on business cycles
  - The wealth effect: although most estimates suggest a weak connection
  - Housing construction has been a good predictor of business cycles (post hoc ergo propter hoc?)

# Why are Housing Price Aggregates Difficult to Measure?

- Housing data is widely collected, but most of it is proprietary

# Why are Housing Price Aggregates Difficult to Measure?

- Housing data is widely collected, but most of it is proprietary
- Real estate data sources

# Why are Housing Price Aggregates Difficult to Measure?

- Housing data is widely collected, but most of it is proprietary
- Real estate data sources
  - National Association of Realtors (Multiple Listing Services)
    - Large sample; timely (quarterly); includes a seasonal adjustment; but much of the information is proprietary

# Why are Housing Price Aggregates Difficult to Measure?

- Housing data is widely collected, but most of it is proprietary
- Real estate data sources
  - National Association of Realtors (Multiple Listing Services)
    - Large sample; timely (quarterly); includes a seasonal adjustment; but much of the information is proprietary
  - The Census Bureau (new home construction)
    - Same advantages plus free historical data; new construction only (in what ways might this be a problem?)

# Price Index Type I: Median Price Index

- The simplest measure is the median

# Price Index Type I: Median Price Index

- The simplest measure is the median
  - Advantage: mean, median, and mode are easy to calculate
  - Median is most popular because it tempers extreme values

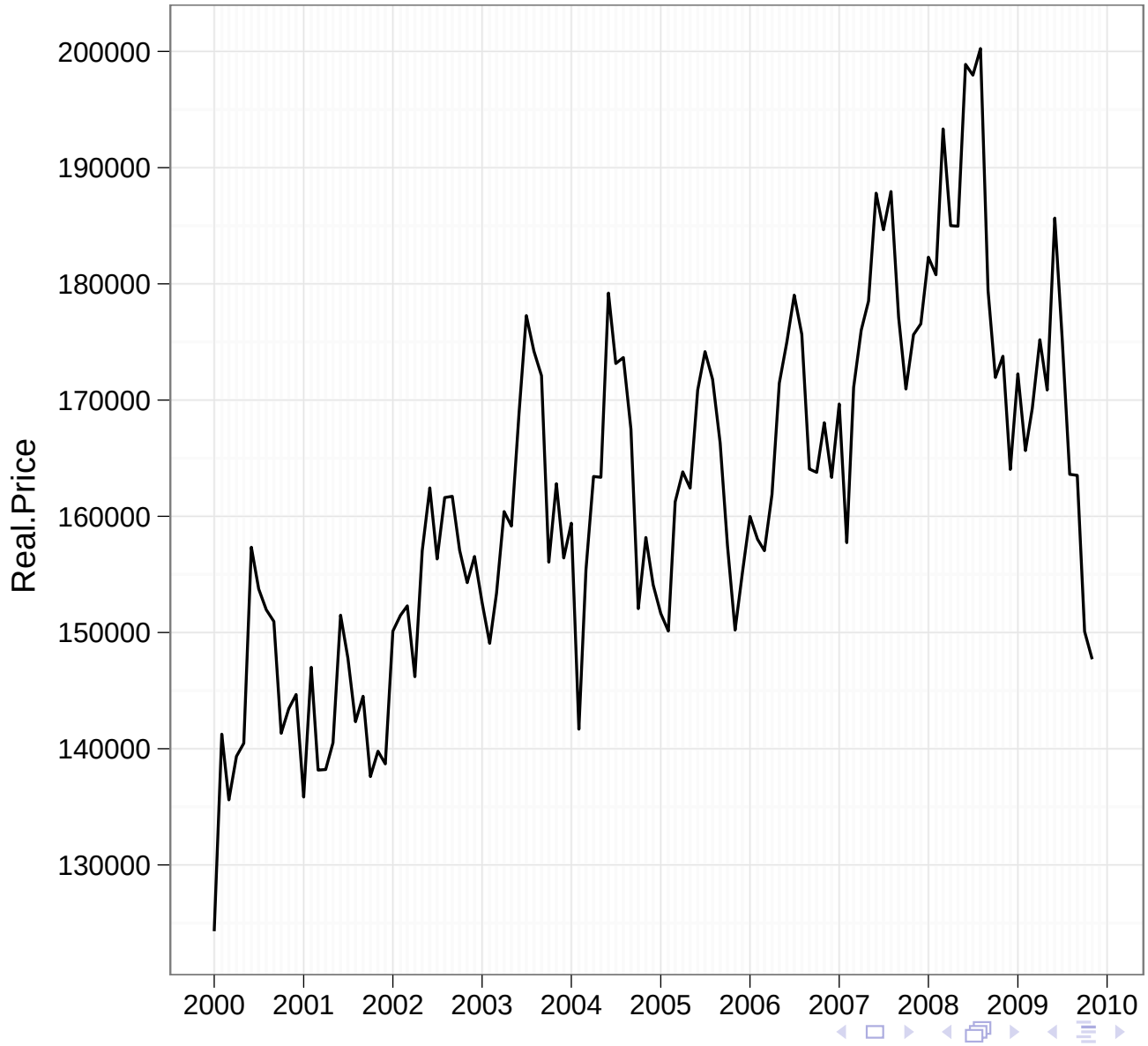
# Price Index Type I: Median Price Index

- The simplest measure is the median
  - Advantage: mean, median, and mode are easy to calculate
  - Median is most popular because it tempers extreme values
- So let's look at some data

# Price Index Type I: Median Price Index

- The simplest measure is the median
  - Advantage: mean, median, and mode are easy to calculate
  - Median is most popular because it tempers extreme values
- So let's look at some data
  - Saint Louis County single-family housing from 2000 through 2009 (interesting years!)
  - Saint Louis County Assessor's Office: actual transactions, not assessments

# Saint Louis Co. Median Housing Price: 2000 - 2009



# Price Index Type I: Median Price Index

- Notice the following:
  - General upward trend (Did you see a "bubble"?)
  - Noisy data; strong seasonal component
  - We could apply a seasonal adjustment to the series

# Price Index Type I: Median Price Index

- Notice the following:
  - General upward trend (Did you see a "bubble"?)
  - Noisy data; strong seasonal component
  - We could apply a seasonal adjustment to the series
- How should we present the data?

# Price Index Type I: Median Price Index

- Notice the following:
  - General upward trend (Did you see a "bubble"?)
  - Noisy data; strong seasonal component
  - We could apply a seasonal adjustment to the series
- How should we present the data?
- Indexes are a nice way to present results because they're unit free
  - $\frac{P_t}{P_0} \cdot 100$; units cancel

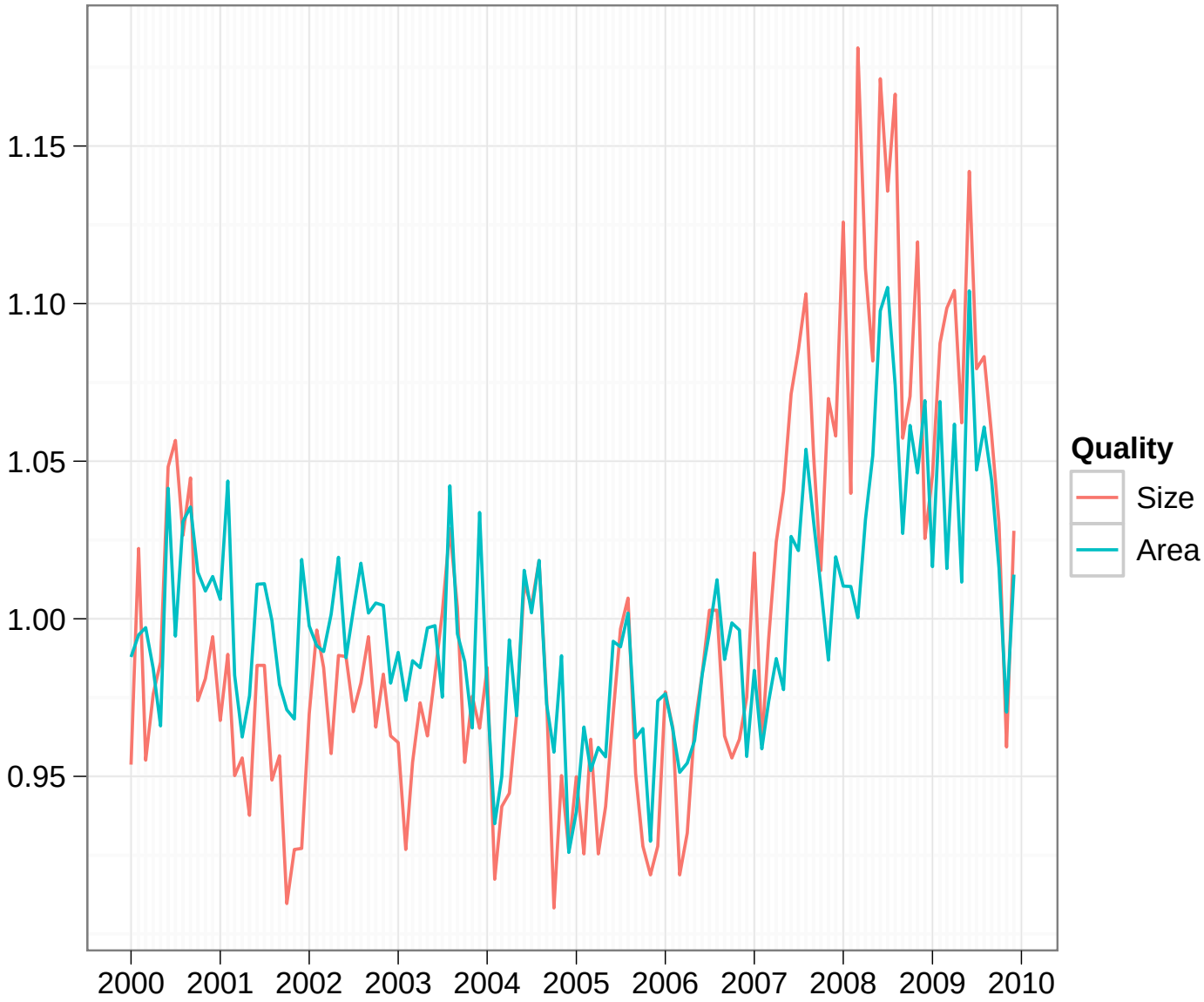
# Price Index Type I: Median Price Index

- Notice the following:
  - General upward trend (Did you see a "bubble"?)
  - Noisy data; strong seasonal component
  - We could apply a seasonal adjustment to the series
- How should we present the data?
- Indexes are a nice way to present results because they're unit free
  - $\frac{P_t}{P_0} \cdot 100$; units cancel
- Fine, but we're assuming constant housing quality across months

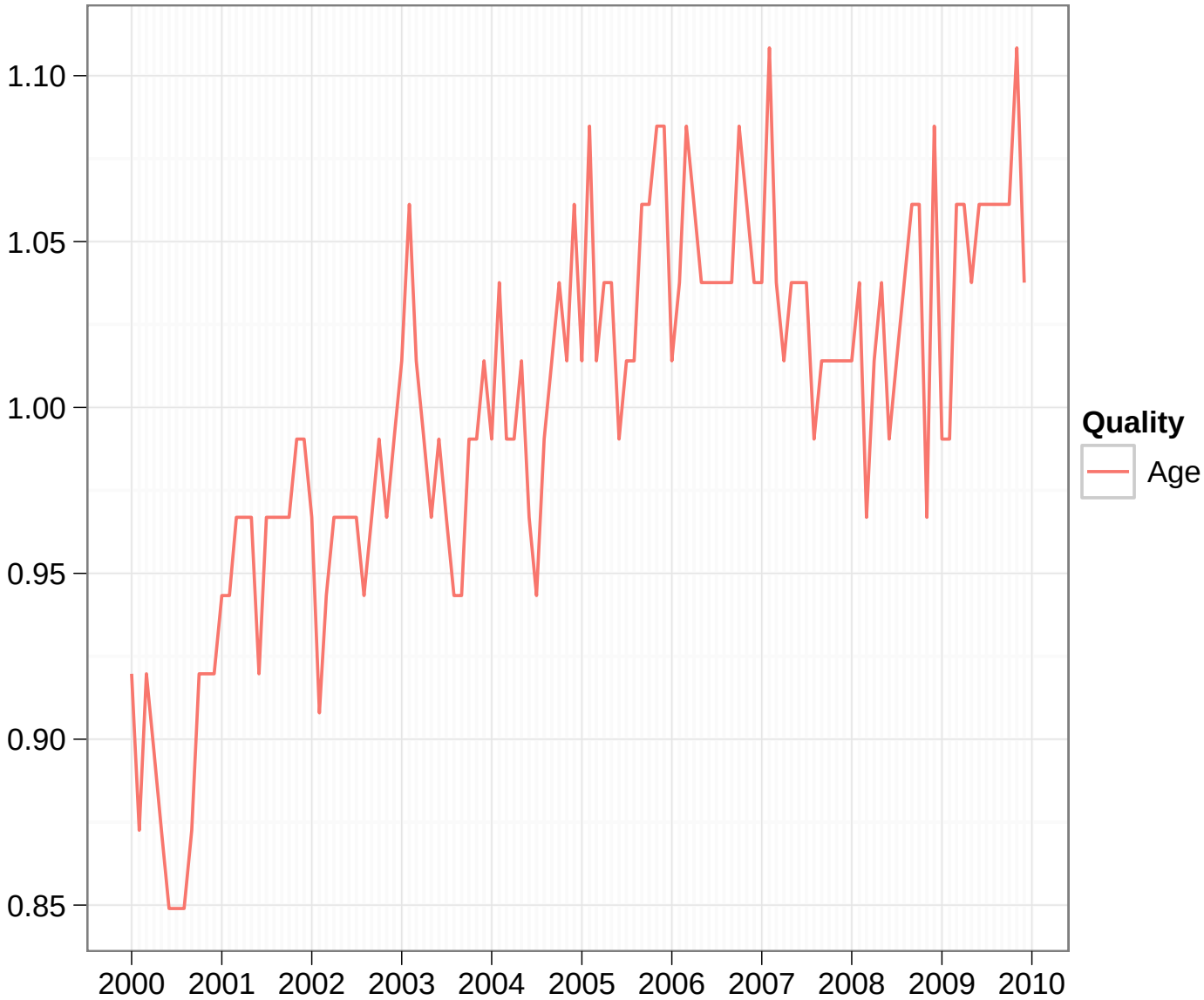
# Price Index Type I: Median Price Index

- Notice the following:
  - General upward trend (Did you see a "bubble"?)
  - Noisy data; strong seasonal component
  - We could apply a seasonal adjustment to the series
- How should we present the data?
- Indexes are a nice way to present results because they're unit free
  - $\frac{P_t}{P_0} \cdot 100$; units cancel
- Fine, but we're assuming constant housing quality across months
- Let's look at housing quality data

# Saint Louis Co. Median Housing Quality: 2000 - 2009



# Saint Louis Co. Median Housing Quality: 2000 - 2009



# Price Index Type I: Median Price Index

- Notice the following

# Price Index Type I: Median Price Index

- Notice the following
  - General upward and downward trends; thus quality changes
  - Noisy data; strong seasonal component

# Price Index Type I: Median Price Index

- Notice the following
  - General upward and downward trends; thus quality changes
  - Noisy data; strong seasonal component
- A median price index does not hold quality constant

# Price Index Type I: Median Price Index

- Notice the following
  - General upward and downward trends; thus quality changes
  - Noisy data; strong seasonal component
- A median price index does not hold quality constant
  - Tells us about the typical housing expenditure
  - Does not tell us about house price appreciation

# What is a Price Index?

- An ideal price index tells us about constant-quality price changes

# What is a Price Index?

- An ideal price index tells us about constant-quality price changes
- Laspeyres index: $I_t = \frac{P_t Q_0}{P_0 Q_0} \cdot 100$

# What is a Price Index?

- An ideal price index tells us about constant-quality price changes
- Laspeyres index: $I_t = \frac{P_t Q_0}{P_0 Q_0} \cdot 100$
- Paasche index: $I_t = \frac{P_t Q_t}{P_0 Q_t} \cdot 100$

# What is a Price Index?

- An ideal price index tells us about constant-quality price changes
- Laspeyres index: $I_t = \frac{P_t Q_0}{P_0 Q_0} \cdot 100$
- Paasche index: $I_t = \frac{P_t Q_t}{P_0 Q_t} \cdot 100$
- The above indexes generally overstate price changes due to the substitution bias

# What is a Price Index?

- An ideal price index tells us about constant-quality price changes
- Laspeyres index: $I_t = \frac{P_t Q_0}{P_0 Q_0} \cdot 100$
- Paasche index: $I_t = \frac{P_t Q_t}{P_0 Q_t} \cdot 100$
- The above indexes generally overstate price changes due to the substitution bias
  - Fisher Index: $I_t = \sqrt{\frac{P_t Q_0}{P_0 Q_0} \cdot \frac{P_t Q_t}{P_0 Q_t}} \cdot 100$
  - Chain-weighting has become popular... more on that later

# Price Index Type II: Hedonic Index

- Rationale: housing is like a bundle of homogeneous services

# Price Index Type II: Hedonic Index

- Rationale: housing is like a bundle of homogeneous services
- The hedonic model estimates the "shadow price" of each service in order to measure the overall change in housing series prices

# Price Index Type II: Hedonic Index

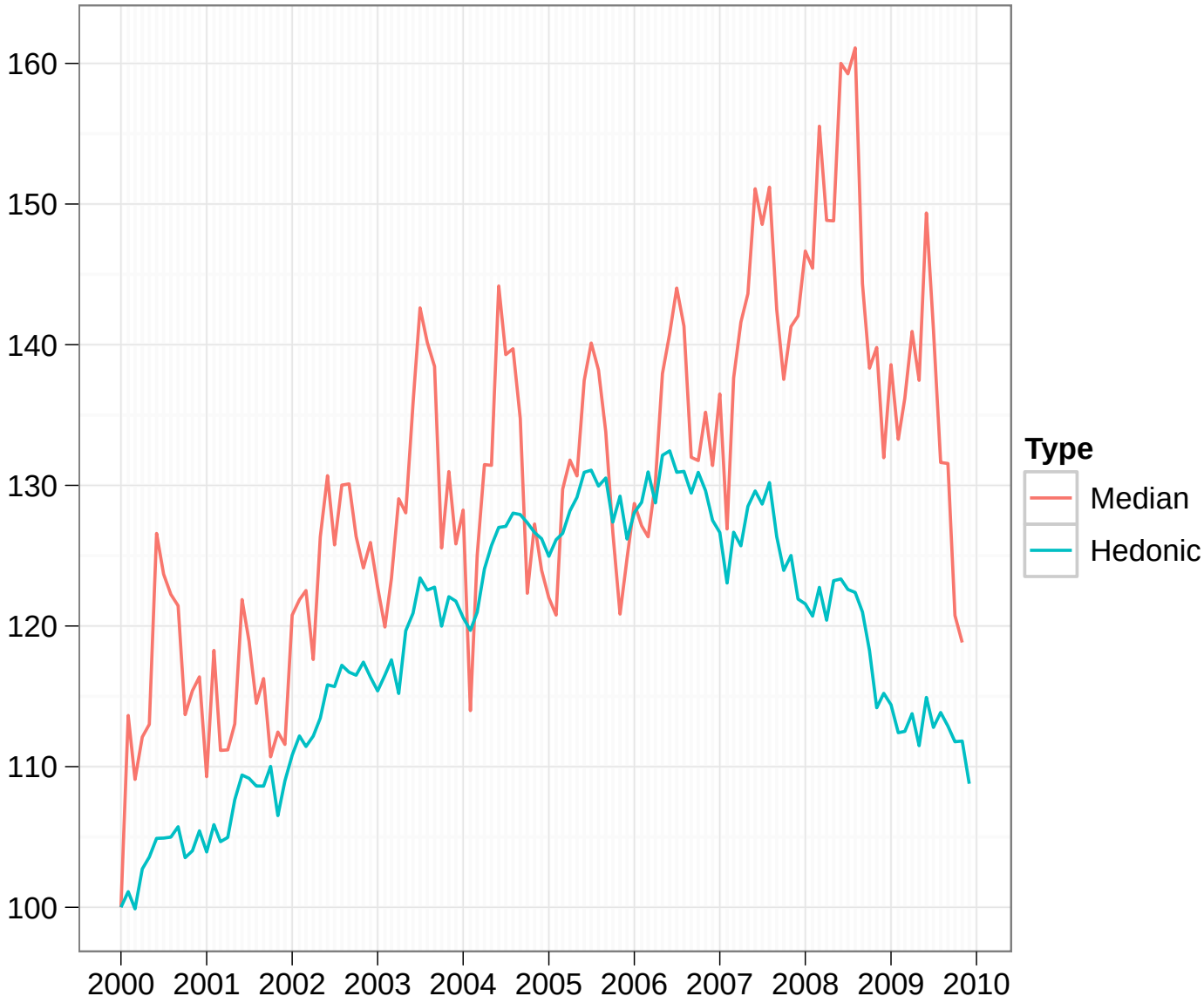
- Rationale: housing is like a bundle of homogeneous services
- The hedonic model estimates the "shadow price" of each service in order to measure the overall change in housing series prices
- Hedonic traditions: (1) explicit time and (2) imputation
  - Let's focus on the explicit time approach first

# Price Index Type II: Hedonic Index

- Rationale: housing is like a bundle of homogeneous services
- The hedonic model estimates the "shadow price" of each service in order to measure the overall change in housing series prices
- Hedonic traditions: (1) explicit time and (2) imputation
  - Let's focus on the explicit time approach first
- For simplicity, consider a semi-log hedonic model

$$lnP_{th} = \sum_{c=1}^{C} \beta_c z_{cth} + \sum_{t=1}^{T} \delta_t d_{th} + \varepsilon_{th} \qquad (1)$$

# Saint Louis Co. Hedonic Index: 2000 - 2009



# Hedonic Index: Explicit Time Model

- Notice the following

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker
- We still need to apply a seasonal adjustment to the series

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker
- We still need to apply a seasonal adjustment to the series
- What are we assuming about the housing market? (I.e. what are we holding constant?

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker
- We still need to apply a seasonal adjustment to the series
- What are we assuming about the housing market? (I.e. what are we holding constant?
  - Assuming accurate data and OLS assumptions – what's new?

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker
- We still need to apply a seasonal adjustment to the series
- What are we assuming about the housing market? (I.e. what are we holding constant?
  - Assuming accurate data and OLS assumptions – what's new?
  - Assuming coefficients constant over time and space

# Hedonic Index: Explicit Time Model

- Notice the following
  - You can see the downturn happen sooner
  - Seasonal component still exists, but is weaker
- We still need to apply a seasonal adjustment to the series
- What are we assuming about the housing market? (I.e. what are we holding constant?
  - Assuming accurate data and OLS assumptions – what's new?
  - Assuming coefficients constant over time and space
- BTW: What is the base year?

# Hedonic Index: Explicit Time Model

- There are three variants on the explicit time approach
  - (1) single equation, (2) overlapping equation, and (3) repeat sales

# Hedonic Index: Explicit Time Model

- There are three variants on the explicit time approach
  - (1) single equation, (2) overlapping equation, and (3) repeat sales
- The single equation model includes a time variable (or set of dummy variables)

# Hedonic Index: Explicit Time Model

- There are three variants on the explicit time approach
  - (1) single equation, (2) overlapping equation, and (3) repeat sales
- The single equation model includes a time variable (or set of dummy variables)
- The overlapping equation model uses a chain of overlapping time periods

# Hedonic Index: Explicit Time Model

- There are three variants on the explicit time approach
  - (1) single equation, (2) overlapping equation, and (3) repeat sales
- The single equation model includes a time variable (or set of dummy variables)
- The overlapping equation model uses a chain of overlapping time periods
- The repeat sales model is a ratio of two explicit-time models

# Hedonic Index: Repeat Sales Model

- The repeat sales model is not normally thought of as a hedonic model

# Hedonic Index: Repeat Sales Model

- The repeat sales model is not normally thought of as a hedonic model
- Consider a house that sells in 2000 and again in 2008

$$lnP_{2000} = \sum_{c=1}^{C} \beta_c z_{c,2000} + \delta_{2000} d_{2000} \tag{2}$$

$$lnP_{2008} = \sum_{c=1}^{C} \beta_c z_{c,2008} + \delta_{2008} d_{2008} \tag{3}$$

# Hedonic Index: Repeat Sales Model

- The repeat sales model is not normally thought of as a hedonic model
- Consider a house that sells in 2000 and again in 2008

$$lnP_{2000} = \sum_{c=1}^{C} \beta_c z_{c,2000} + \delta_{2000} d_{2000} \tag{2}$$

$$lnP_{2008} = \sum_{c=1}^{C} \beta_c z_{c,2008} + \delta_{2008} d_{2008} \tag{3}$$

- Assuming the house's bundle has not changed, consider the following equation

$$lnP_{2008} - lnP_{2000} = \delta_{2008} d_{2008} - \delta_{2000} d_{2000} \tag{4}$$

# Hedonic Index: Repeat Sales Model

- Federal Housing Finance Agency (FHFA)
  - Fannie & Freddie loans: large but biased sample
  - Estimates by the 9 Census divisions and weighted by housing units

# Hedonic Index: Repeat Sales Model

- Federal Housing Finance Agency (FHFA)
  - Fannie & Freddie loans: large but biased sample
  - Estimates by the 9 Census divisions and weighted by housing units
- Case-Shiller (published by Standard & Poor's)
  - Uses deed records, where suspected non-arms-length transactions are excluded
  - Estimates by the 9 Census divisions and weighted by estimated housing values: PE Ratio

# Hedonic Index: Repeat Sales Model

- Advantages of a repeat sales model
  - Some control of quality: depend on reinvestment
  - Data requirements are only slightly larger than the median measure

# Hedonic Index: Repeat Sales Model

- Advantages of a repeat sales model
  - Some control of quality: depend on reinvestment
  - Data requirements are only slightly larger than the median measure
- Disadvantages of the repeat sales model
  - Housing reinvestment and embedded depreciation
  - Repeat sales sample bias
  - Past measures change when new sales are added

# Hedonic Index: Imputation Method

- Single equation and repeat sales models assume a constant set of shadow prices (substitution bias), and change when new sales are added

# Hedonic Index: Imputation Method

- Single equation and repeat sales models assume a constant set of shadow prices (substitution bias), and change when new sales are added
- The imputation method allows shadow prices to change over time and space, and does not change when new sales are added

# Hedonic Index: Imputation Method

- Single equation and repeat sales models assume a constant set of shadow prices (substitution bias), and change when new sales are added
- The imputation method allows shadow prices to change over time and space, and does not change when new sales are added
- Shadow prices are then used to estimate the price ratio

$$\frac{p_{kt}}{p_{js}} = \frac{p_{kt}(z_{kt})}{p_{js}(z_{kt})} \times \frac{p_{js}(z_{kt})}{p_{js}(z_{js})} \tag{5}$$

# Hedonic Index: Imputation Method

- Single equation and repeat sales models assume a constant set of shadow prices (substitution bias), and change when new sales are added
- The imputation method allows shadow prices to change over time and space, and does not change when new sales are added
- Shadow prices are then used to estimate the price ratio

$$\frac{p_{kt}}{p_{js}} = \frac{p_{kt}(z_{kt})}{p_{js}(z_{kt})} \times \frac{p_{js}(z_{kt})}{p_{js}(z_{js})} \tag{5}$$

- The first term represents a constant quality price index, and the second is a quanity index

# Hedonic Index: Imputation Method

- How can one estimate separate shadow prices for each time period and region?

# Hedonic Index: Imputation Method

- How can one estimate separate shadow prices for each time period and region?
- A dummy variable for each period-region with interaction

# Hedonic Index: Imputation Method

- How can one estimate separate shadow prices for each time period and region?
- A dummy variable for each period-region with interaction
- Separate regression for each period-region

# Hedonic Index: Imputation Method

- How can one estimate separate shadow prices for each time period and region?
- A dummy variable for each period-region with interaction
- Separate regression for each period-region
- What are the advantages?